Binomial distribution in a video game automation model

Distribución binomial en un modelo de automatización en videojuegos.

Barrios-de la Cruz, Francisco[a], Aparicio-Urbano, José [b] and González-Dominguez, Marcos Crescencio*[c]

[a] TecNM: Tecnológico de Estudios Superiores de Jocotitlán • LTD-8902-2024• 0009-0001-4361-1966.
[b] TecNM: Tecnológico de Estudios Superiores de Jocotitlán • 0000-0002-6323-2605
[c] TecNM: Tecnológico de Estudios Superiores de Jocotitlán• JWP-9663-2024 • 0000-0002-0933-6374 • 666986.

CONAHCYT classification:

DOI: https://doi.org/10.35429/H.2024.12.56.66

Area: Physics-Mathematics and Earth Sciences
Field: Mathematics
Discipline: Analysis and functional analysis
Subdiscipline: Applied mathematics

Key Handbooks

To provide a clear vision on binomial distributions in different contexts, from the approach of video games, in which it is required to produce resources in the most efficient way and in daily life with the production of resources seeking the highest efficiency against a shorter time of obtaining maximizing the results against a production time. Knowing that with simulation tools and video games can be obtained and modeled results applicable to real life, in this case, the efficient production of resources. It was achieved to satisfy the reduction of time in 83% corresponding to the improvements implemented in order to reduce production time and based on the above shown in the research it was concluded that you can get to lose a block of ice every 8 hours. Therefore, it is concluded that the improvement and optimization of conditions can maximize and make production more efficient.

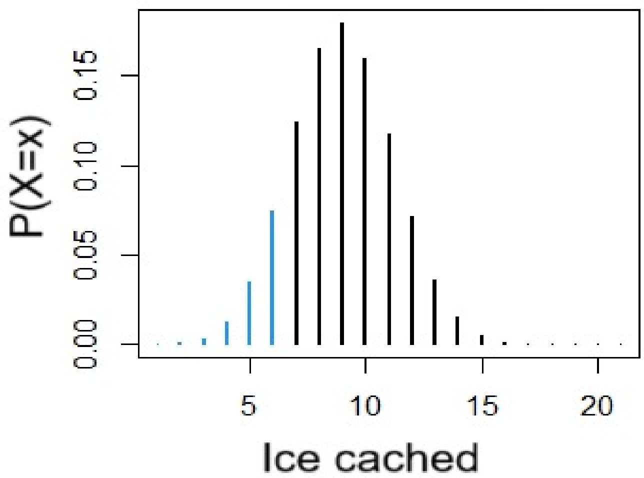
Citation: Barrios-de la Cruz, Francisco, Aparicio-Urbano, José and González-Dominguez, Marcos Crescencio. 2024. Binomial distribution in a video game automation model. 56-66. ECORFAN.

* ✉ [marcos.gonzalez@tesjo.edu.mx]

Handbook shelf URL: https://www.ecorfan.org/handbooks.php

Abstract

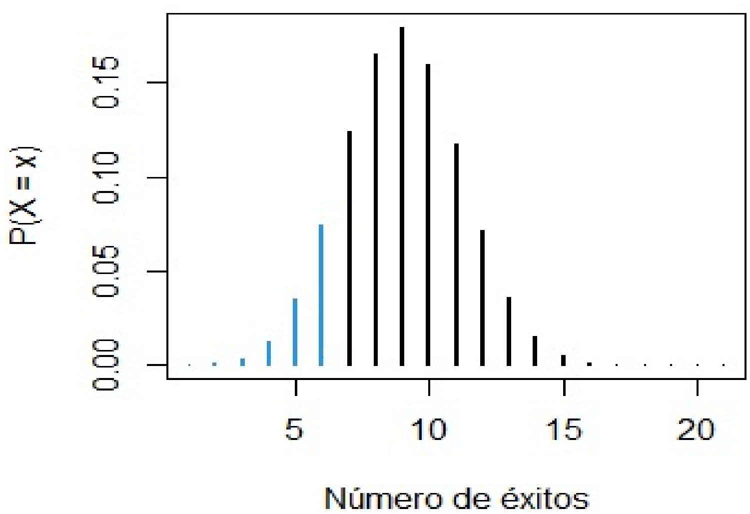
This article presents the application of the Hybrid approach that contemplates an improvement in the collection times of materials by the user as well as a calculation method to avoid loss of raw material and waste of time, this was done in the videogame called "Minecraft" the calculations were developed using Matlab® software through a statistical analysis of a binomial distribution; evaluating the conditions to obtain these materials, for the section of improvements in the times a program was used to calculate a before and after, which resulted in the probability that a block is lost in a cycle is approximately $2.1x10^{-6}$ and on average one block escapes every 8 hours, which coincides with the experimental result.



Statistical, Analysis, Distribution, Probability, Application.

Resumen

En este artículo se presenta la aplicación de un enfoque híbrido que contempla una mejora en los tiempos de recolección de materiales por parte de los usuario asi como un método de cálculo para evitar pérdidas de materia prima y desperdicio de tiempo, esto se realizó en el videojuego llamado “Minecraft” los cálculos se desarrolló mediante el software Matlab® a través de un análisis estadístico de una distribución binomial; evaluando las condiciones para obtener estos materiales, para la sección de mejoras en los tiempos se utilizó un programa para calcular un antes y después, con lo cual se obtuvo como resultado que la probabilidad de que se pierda un bloque en un ciclo es de aproximadamente 2.1x10-6 y por término medio se escapa un bloque cada 8 horas con lo que coincide el resultado experimental.



Estadística, Análisis, Distribución, Probabilidad, Aplicación

Introduction

This article was developed using binomial distribution methods and analysis of the time cycles that an average player spends in the game to collect raw materials for their projects or constructions within the Minecraft video game, in order to improve, optimize and automate the time spent gathering materials.

The development of probabilistic reasoning is one of the important objectives, especially in the process that is and needs to be optimized (Jones, Langrall and Mooney, 2007); while the binomial is the most important discrete distribution studied in the university probability course, this combined with the fact that we are living a revolution where the industry evolves day by day, mainly in the field of automation.

However, few explorations have been made on problems that attempt to develop probabilistic reasoning taking as a starting point the concept of distribution, in particular, the binomial distribution (Budde et al., 2024).

This study presents an exploration of the development of probabilistic reasoning that can be applied in a model to reduce the time to obtain materials in a video game , and which of course have not been introduced to the formal study of probability and statistics (Radzvilas, 2024), taking as a starting point the notion of binomial distribution in its simplest form, when doing physical (coins) and computational (Fathom) simulations.

This interest arises from one of the eleven questions formulated by Pfannkuch and Reading (2006) for the development of the notion of distribution: "How is reasoning about distributions developed from their simplest forms or aspects to more complex ones?" Based on what was proposed by it is raised how we can use the binomial distribution to see the way of improving times in obtaining resources and in this way knowing that the time used in this task can be reduced and used in other types of activities reducing the lost material and above all the wasting time again in obtaining these (Khatri, 1961).

In (Ramiro Vásquez et al., 2018) they study how to perform a data analysis with a binomial distribution in a data simulation, along with this emphasis is placed on the use of the Monte Carlo method for the use of random variables with a binomial distribution for its later application in the article as a way of knowing the proportion of the sample and this can be used to simulate random data from a given distribution, which allows estimating the distribution and proportion of the data in a sample .

Mainly when we now want to use 7 different users and model the before and after of the model in terms of the time used. Now based on this (Gómez Sarahi et al., 2019) propose an application of analysis alternatives in experiments with repeated measures where one of the most used research methods is to measure the response variable in the same experimental unit at different points in time.

These are more efficient than using a different experimental unit for each measurement, since they require fewer units, which reduces the sample size and the estimates over time are more accurate.
In (Olkin et al., 1981) it is shown that depending on the variation of the number of N trials in a given sample and replicating the problem with N trials is less studied, so it is considered more difficult in the binomial distribution, this given if p is unknown and small, while if N is large this becomes very unstable having a large variation. Some properties of the binomial distribution, such as its symmetry, displacement, flattening and dispersion were exposed in the work of Santamaría M. and Malla M. (2007) as well as a brief introduction on the definition and notation of a binomial random variable, B (n, k). The following procedure shows a simple way to fit a binomial distribution by a Monte Carlo method function and use areas to represent the probabilities given by the ordinates of the discrete distribution (Weisstein, E. W., 2002; Wang et al., 2024).

2. Development

Within the game there is a mechanic where water blocks that meet the following conditions have a chance of freezing:

- In a snow-capable area of a cold biome, such as a snow biome and a mountain.
- There are no blocks above it.

- The brightness of blocks in water is less than 10.
- At least one block that does not contain water is adjacent to it.

Blocks in water that meet the conditions in this way have a chance to freeze every game tick (gametick, *gt*). The specific mechanism is: in a block, each *gt* has a 1/16 chance to select the topmost block in the 16x16 horizontal position and attempt to freeze, if the necessary conditions are met we can make it freeze, then the probability of a water grid freezing per *gt* is:

$$\rho = (\frac{1}{16})^3 \approx 0.000244 \quad [1]$$

2.1 Ice probability analysis

When starting with the probability analysis there are different types of probabilities, each with a specific objective, in this case the binomial probability distribution was selected (Johnson et al., 2004), because in it there are only two results that need, that there is a success or not, that is why it was considered to use it within this work. Suppose that in a grid of water blocks they freeze and turn into ice, each grid of water meets the freezing conditions. Assuming that the freezing time of a water tank is *T gt* and a gt is equivalent to 1/20 that is, every 0.05 sec, knowing that every 20 minutes in the real world a whole day has passed in the game, therefore in 1 hour there are 72000 *gt* then the probability that a water grid freezes within *T gt* is:

$$P = (T) = 1 - (1 - \rho)^T \quad [2]$$

Assuming there are n water grids that can be frozen in a water reservoir, since ρ in equation (1) is very small, the freezing events of the n water grids can be approximately considered to be independent and identically distributed, then, after *T gt* , the amount of new ice follows a binomial distribution as shown in equation (3):

$$P(X = k) = C_n^k P^k(T)(1 - P(T))^{n-k} \quad [3]$$

where C_n^k is the probability of success.

2.2 Analysis of the amount of ice stored in the tank

In “Ice Farm”, the length of water tank of one unit of ice farm is L, and the ice stored in the water tank has a space b, then the amount of water that can be frozen is:

$$n = L - b \quad [4]$$

The working process of the ice farm cycle of one unit is as follows: firstly, after the stationary ice formation through *T gt*, the flying machine is used to push at most M pieces of ice, where M=12 is the upper limit of the piston thrust. So after *T gt* freezes, the total amount of ice in the storage is b+k. After the flying machine collects the ice, the amount of ice remaining is Max(b+kM, 0). Therefore, at the beginning of the next cycle, the amount of ice stored is:

$$b_{next} = \max(b + k - M, 0) \quad [5]$$

It can be seen that the amount of ice stored in one cycle is only related to the previous cycle, so the idea of state transition can be used for analysis. We divide the water tank into different states according to the different amounts of ice stored b, and set the probability that the water tank is in a state with a stored amount of ice b as F(b), and it satisfies Equation 6:

$$\sum_{b=0}^{L} F(b) = 1 \quad [6]$$

If the initial amount of ice stored in a cycle is b, from equations (3), (4) and (5), the probability that the amount of ice stored becomes b-after a cycle is:

$$P(b, b_{next}) = P(X = k) = C_{L-b}^{b_{next}-b+M} P^{b_{next-b+M}}(T)(1 - P(T))^{L-b_{next}-M} \quad b_{next} > 0$$

$$P(b,0) = \sum_{k=0}^{M-b} P(X=k) \sum_{k=0}^{M-b} C_{L-b}^{k} P^{k}(T)(1-P(T))^{L-b-k} \quad b_{next} = 0$$ [7]

Then, after one cycle, the probability distribution F(b) of ice storage becomes:

$$F_{next}(b) = \sum_{b'}^{L} F(b')\,(Pb', b)$$ [8]

From where the expression (8) can be represented in matrix form as follows:

$$\begin{bmatrix} P(0,0) & P(1,0) & \dots & P(L,0) \\ P(0,1) & P(1,1) & \dots & P(L,1) \\ \vdots & \vdots & & \vdots \\ P(0,L) & P(1,L) & \cdots & P(L,L) \end{bmatrix} \begin{bmatrix} F(0) \\ F(1) \\ \vdots \\ F(L) \end{bmatrix} = \begin{bmatrix} F_{next}(0) \\ F_{next}(1) \\ \vdots \\ F_{next}(L) \end{bmatrix}$$ [9]

2.3 Ice Leak Probability Analysis

Ice leakage occurs when the amount of ice coming out of the tank reservoir is less than 12, considering that the maximum is 12 in one cycle, which results in a final ice mining throughput of less than 72000/hr, i.e. one mined block every *gt* and it is known that 1 hour in real life is 72000 game ticks. In the "ice farm" if the sum of stored ice and new ice is less than 12 pieces, ice leakage will occur, i.e. b+k<12.

Assuming that the ice production in a cycle is *M'*, then according to expression (5), there is:

$$\begin{matrix} M' = M & b+k \geq M \\ M' = b+k & b+k < M \end{matrix}$$ [10]

Then the probability distribution of ice production is:

$$P(M') = P(X = M' - b) = C_{L-b}^{M'-b} P^{M'-b}(T)(1-P(T))^{L-M'} M' < M$$

$$P(M) = \sum_{k=M-b}^{n} P(X=k) \sum_{k=M-b}^{L-b} C_{L-b}^{k} P^{k}(T)(1-P(T))^{L-b-k} \quad M' = M$$ [11]

In equations (10) and (11), the first row represents the normal output case M'=12 pieces of ice, and the second row represents the ice leakage case M'<12. Then, the probability of ice leakage is:

$$P_{miss} = \sum_{b=0}^{L} F(b)(1-P(M))$$ [12]

3. Results

For the ice probability analysis, we consider T=1320 and n=40 (The ice farm norms), the probability distribution of the number of frosts is shown in the following Figure:

Box 1

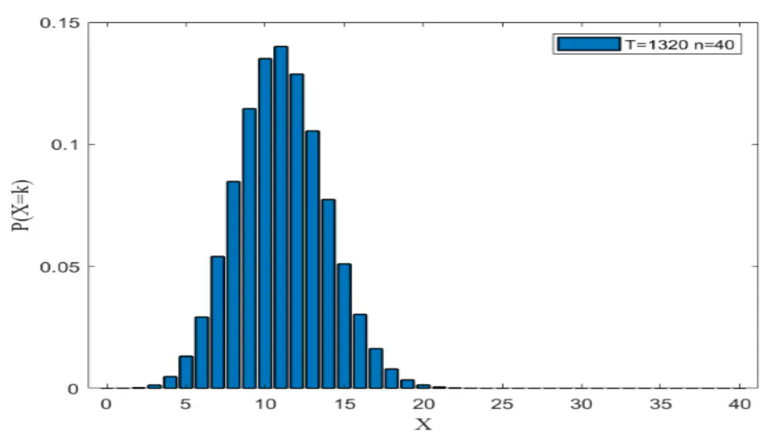


Figure 1

Probability of distribution with T=1320 n=40

Source [The author]

In Figure 1 we can see that using the data with n=40 and T=1320 we can obtain a binomial distribution, in which we can see its maximum point at 12 units on the ordered axis. The data given with n=40 and T=1320 are existing data given by the player, where n equals the number of modules that can form ice and T the time it takes for each module to produce ice again. The selected distribution is binomial because we only deal with two important results, the probability of success or failure, in this case that there is a material leak and the player has to do rework.

The binomial distribution gives the discrete probability distribution $P_p(n|k)$ of obtaining exactly *n* successes out of *k* Bernoulli trials (where the result of each Bernoulli trial is true with probability *p* and false with probability $T = 1 - p$).

For the case of the analysis of the amount of ice stored in the tank, we will take the "ice farm" as an example, considering that the initial state is non-freezing of the water, that is, F(0)=1, after a cycle, the probability distribution F(b) of ice storage is shown in the following figure:

Box 2

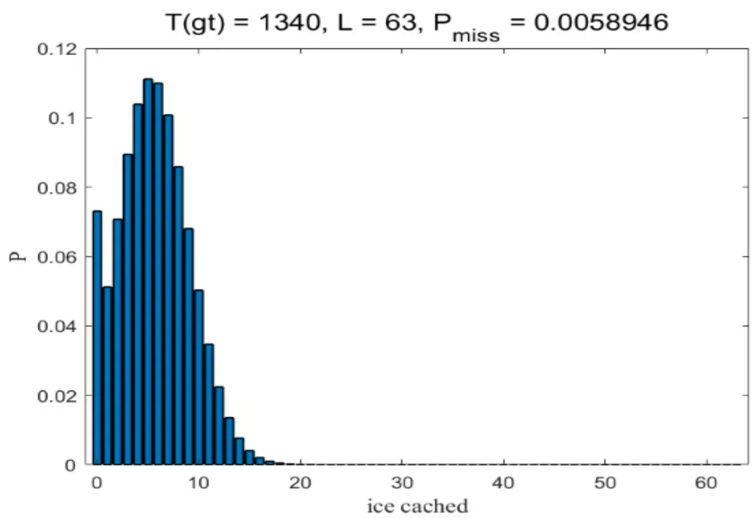


Figure 2

Probability of distribution with T=1340 and L=63

Source [The author]

Figure 2 we can observe a variation between units 6 and 7 where the maximum is 11 units on the ordered axis. The Figure behaves in such a way that when in the initial state 0 there are no blocks that feed the demand of the machine and as it increases the units on the ordered axis decrease. After a sufficient time (100 cycles), F(b) stabilizes, as shown in the following Figure:

Box 3

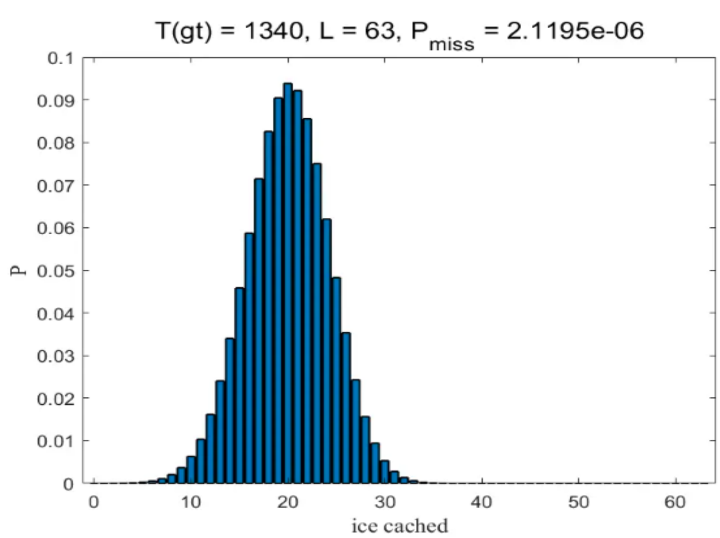


Figure 3

Probability of distribution with T=1340 and L=63

Source [The author]

In Figure 3 it can be observed that at 20 units we find stability in terms of demand for produced materials, as a consequence of the surplus of blocks.

For the ice leakage probability analysis, we will select the water tank length as L=40, 50, 60, 70 to do simulation calculation respectively, and the results are shown in Graphic 4, it can be seen that when the water tank length is less than 60, there is a high probability of ice leakage. When the tank length is more than 70, it is almost impossible for ice leakage to occur. Considering the relevant factors such as the time taken by the flying machine to go back and forth, the water tank length is finally selected as L=63.

Box

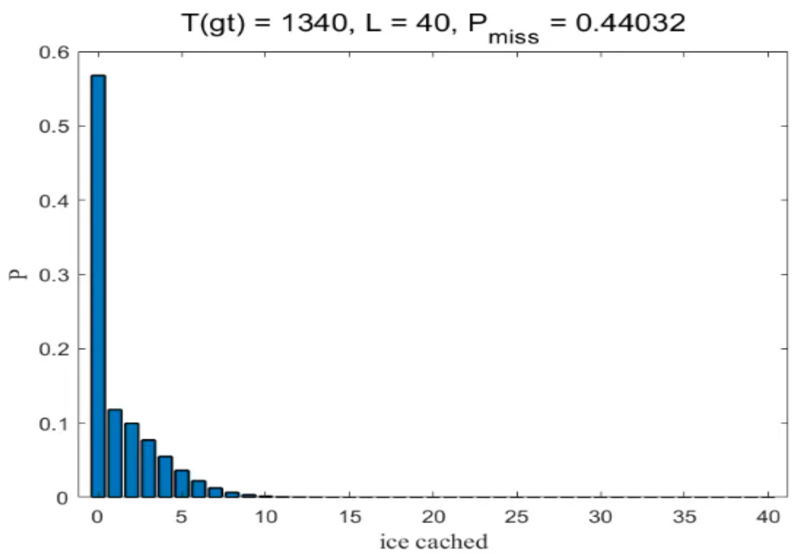


Figure 4
Probability of distribution with T=1340 and L=40

Source [The author]

Figure 4 shows in detail that when the tank length L = 40 there is a very high probability of losses, but being in the initial state 0 there is no material to satisfy the demand and in turn as time progresses the probability decreases, because there is already material in formation.

Box

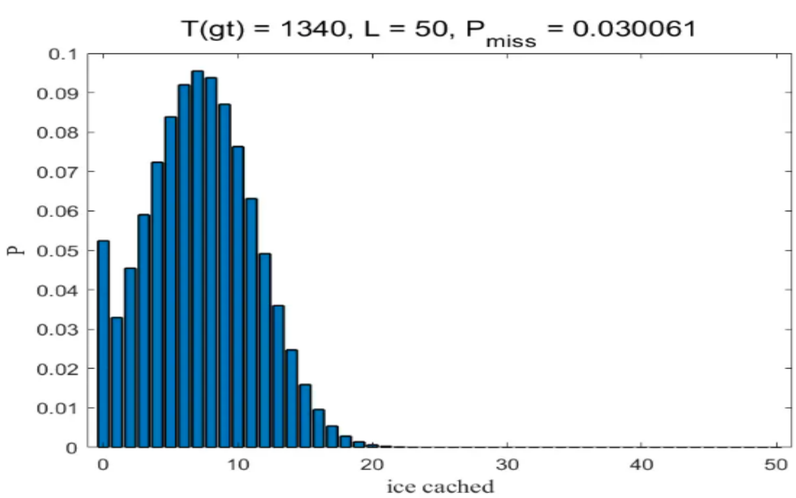


Figure 5
Probability of distribution with T=1340 and L=50

Source [The author]

In Figure 5 we can see that if the tank has a length L = 50, there is a variation between units 7, 8 and 9 within the ordered axis, which indicates that from these existing units there is a high probability of a leak, as a consequence in the initial state 0 there is a probability of a leak, but because there is not enough time for the generation of material there is no material that is lost.

Box 6

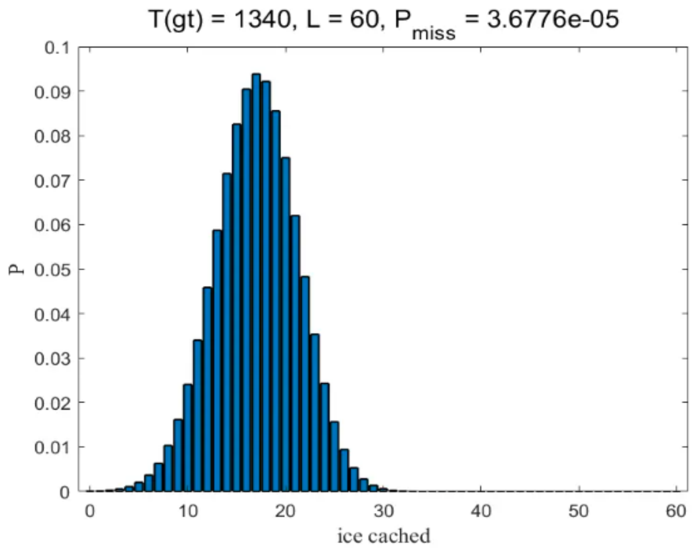


Figure 6

Probability of distribution with T=1340 and L=60

Source [The author]

Within Figure 6 we can see that if the length of the tank L is equal to 60 the graph stabilizes, this happens because there are more necessary generation points that meet the ideal conditions for the formation of ice, taking this into account it can be observed that between units 17, 18 and 19 of the ordered axis there is a probability of leakage, because enough time has already passed for the material to form, which indicates that it is more likely that a material leak will occur between these units.

Box

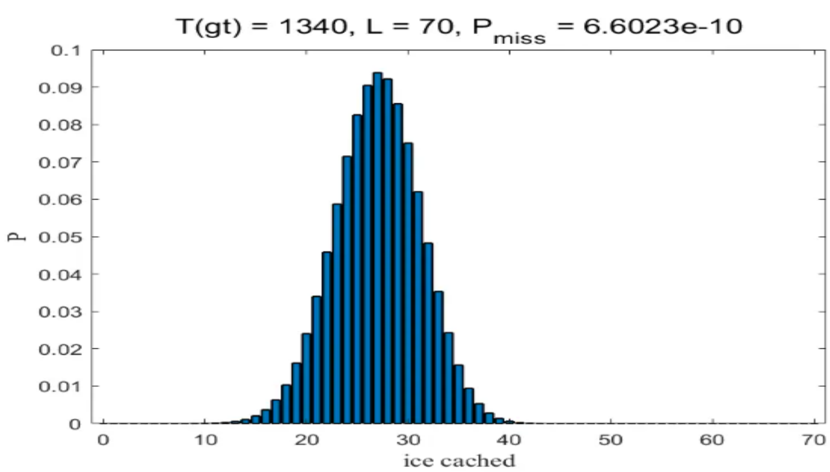


Figure 7

Probability of distribution with T=1340 and L=70

Source [The author]

Taking into account that the length of the tank L is equal to 70, we can see that Figure 7 is stable. This is because the length of the tank increased and the generation points also increased, so now the probability of leakage is between units 28, 29 and 30 of the ordered axes. And in turn, it can be seen that the maximum points are close to the central axis, so that at the midpoint there is already enough material for ice formation, which means that there is a very high probability of leakage in these units.

Before testing, an analysis was performed on 7 users under different software and component conditions to determine the usage time and to begin estimating the data with an average given by the Monte Carlo method (Matlab, 2022) where the average usage time can be expressed and how the data is related.

Box 8

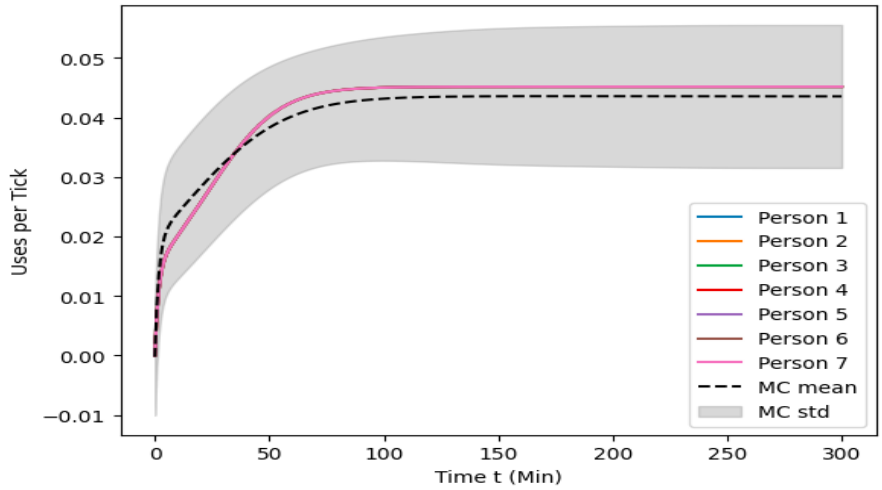


Figure 8

Distribution of results with 7 users before testing

Source [The author]

In the previous tests, an average close to the study done with Monte Carlo can be observed, where it can be found that it stabilizes after 95 minutes after exponential growth.

Box 9

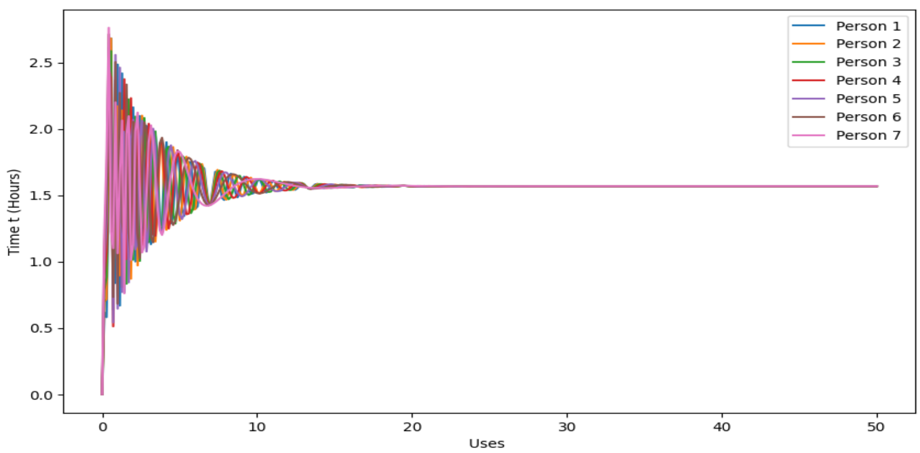


Figure 9

Distribution of results with 7 users with results after testing.

Source [The author]

In Figure 8, it can be observed how at the beginning of the use, the data begins to have a random behavior, but after the number of uses it stabilizes, which shows us that there is an improvement in terms of the time invested in the collection and how this can be used in other activities, while in the previous model it did not stabilize until after 90 minutes, which shows us a significant improvement in time savings and a better use of raw material without wasting it.

4. Discussion

The approach of this analysis has provided the best alternative based on the time/cost ratio, which entails producing these materials, as well as better optimization and automation when obtaining materials for later use. The proposed methodology resulted in the player spending less time collecting raw materials, on the other hand, the number of losses is reduced. In addition, it was found that, of the 6 tests carried out, they do not exceed 60% efficiency, compared to 95% of the final test. In turn, for practical purposes, the improvements can be used and implemented in order to benefit all types of players, whether casual or advanced. In the following tests, different conFiguretions of computer components from 7 users were used for a more effective demonstration of the efficiency and effectiveness of the system in reducing the time used to obtain resources. With the data obtained, it can be observed and the conclusion was reached that with the data collected we are above average, which means an average reduction in time.

5. Conclusions

The time reduction was achieved with 83% that correspond to the improvements implemented in order to reduce the time of production and obtaining materials, in addition the probability of a material leak was established and, in this case, repeating the task, which entails a loss of time, evaluating the established criteria so that obtaining the resources is more efficient, giving better results for the players. The above through the graph that corresponds to the experimental data on which is the most optimal system for the production of blocks, also in the graphs the analysis was made in the variation of parameters to find the cost / time relationship, taking as a reference the demand for raw material based on a small group of around 10 advanced players. This article presents an analysis of the time taken to produce materials and whether there is a probability of leakage, where the player has to repeat the task by reworking and wasting time, so when the mechanism length L=63, the probability of the water mechanism losing an ice block in one cycle is approximately $2.1x10^{-6}$, and on average, an ice block leaks every 8 hours, which basically coincides with the experimental result of an ice block leaking every 13 hours. The contribution of this article contemplates an improvement in time and optimization of processes in a video game. We would like to thank the Tecnológico Nacional de México (TECNM): Tecnológico de Estudios Superiores de Jocotitlán (TESJo) for the facilities provided for the preparation of this article.

Declarations

Conflict of interest

The authors declare that they have no conflicts of interest. They have no financial interests or personal relationships that could have influenced this book.

Authors' contributions

Barrios-de La Cruz, Fransisco: Structure and development of the data to be analyzed, in addition to this, he contributed in the writing of the paper and part of the mathematical analysis.

Aparicio-Urbano, José: Cleanliness of the writing and the structure of the work to be developed, in addition to this, he contributed in the elaboration of the graphs presented in the work.

Gonzalez-Dominguez, Marcos Crescencio: Analysis of the results obtained, translation of the paper and analysis of the graphs presented.

References

Background

Budde, C. E., Hartmanns, A., Meggendorfer, T., Weininger, M., & Wienhöft, P. [2024]. Sound Statistical Model Checking for Probabilities and Expected Rewards. arXiv preprint arXiv:2411.00559.

Gómez Sarai, Torres Verena, García Yoleisy, Herrera Magaly, Medina Yolaine, & Rodríguez, R.. [2019]. Procedimiento estadístico para el análisis de experimentos con medidas repetidas en el tiempo en la esfera agropecuaria. Cuban Journal of Agricultural Science, 53(4), 353-360. Epub 05 de diciembre de 2019.

Johnson, R. y P. Kuby , [2004]: Estadística Elemental. Thomson Learning, México.

Jones, G. A., Langrall, C. W. y Mooney, E. S. [2007]. Research in probability. En F. K. Lester, Jr. (Ed.), Second handbook of research on mathematics teaching and learning, pp. 909-955. Charlotte, NC, USA: Information Age-NCTM

Basics

Khatri, C. G. [1961]. On the distributions obtained by varying the number of trials in a binomial distribution. *Annals of the Institute of Statistical Mathematics*, *13*(1), 47–51.

Olkin, I., Petkau, A. J., & Zidek, J. V. [1981]. A Comparison of n Estimators for the Binomial Distribution. *Journal of the American Statistical Association*, *76*(375), 637–642.

Pfannkuch y Reading [2006]. Reasoning about distribution: a complex process. Statistical Education Research Journal, 5(2), pp. 4-9..

Support

Radzvilas, M., Peden, W., Tortoli, D., & De Pretis, F. [2024]. A comparison of imprecise Bayesianism and Dempster–Shafer theory for automated decisions under ambiguity. Journal of Logic and Computation, exae069.

Ramiro Vásquez, E., Caballero Núñez, A., & Herrera Villafranca, M. [2018]. Alternativas de análisis de datos con distribución binomial en diseño de bloques al azar. Cultivos Tropicales, 39(4), 66-70.

Santamaria M. y Malla M. [2007], Comprensión de las Propiedades de la Distribución Binomial mediante Gráficos con R". XXXV Coloquio Argentina de Estadística, Mar del Plata, Argentina

Differences

The MathWorks Inc. [2022]. MATLAB version: 9.13.0 (R2022b), Natick, Massachusetts: The MathWorks

Wang, J., & Hu, G. [2024]. Game-Based Optimal Aggregation of Energy Prosumer Community With Mixed-Pricing Scheme in Two-Settlement Electricity Market. IEEE Transactions on Automation Science and Engineering.

Discussions

Weisstein, E. W. [2002]. Binomial distribution.