Journal of Statistics, Optimization and Data Science
Vol. 1 No: 1 (June 2023); pp 50-60

Combinatorial Probability Distribution on Sequence of Overlapping Squares within a Step Formed Square

Surnam Narendra[1*], Tirupati Rao Padi[2]

1. Research Scholar, Dept. of Statistics, Pondicherry University, surnamnarendra@gmail.com
2. Professor. Department of statistics, Pondicherry University, India. drtrpadi@gmail.com

ABSTRACT

A Square is a quadrilateral with four sides and equal angles. The diagonals of a square are equal and bisect each other. The step-formed square is also a square with different unit lengths. In this paper, we model the probability distribution of a sequence of overlapping squares within a step-formed square in the case of odd-numbered courts on the baseline. We explored the direct functional relationships for the formulated distribution's probability mass function (PMF). We have verified the regulations of the probability mass function (PMF). This study has focused on deriving different mathematical relations of the statistical measures for location, scaling, shaping, peakedness etc. Further, we discussed the inter-probability distribution properties, generating functions, characteristic functions, etc. Sensitivity analysis uses suitable numerical illustrations to understand the model behaviour. This study has numerous real-time applications in the context of Combinatorics.

Keywords: Discrete probability distribution, squares within a step formed Squares, Pearson's coefficients, sensitivity analysis.

****Corresponding Author**: Surnam Narendra,* surnamnarendra@gmail.com

1. Introduction

This work aims to formulate the probability mass function of the distributions for a sequence of overlapping squares within a step-formed court and explicit statistical properties. The step-formed squares will be started with one, two, three, and so on …, n numbered squares on the baseline. There are two possible formulations that the step-formed courts will have, either odd-numbered or even-numbered yards on the baseline. This study deals with the initial understanding of the problem as the maximum possible baseline length of a step-formed square with an odd number. The study has identified the possible sample space and favourable sub-spaces by observing the diversified patterns. Probability functions are defined with the help of subspace and sample space domains. The notions of probability mass function, cumulative distribution function, etc., will be used to understand further detailed concepts of the formulated distribution. Model behaviour is studied with Pearson's coefficients and other characteristics such as mean, variance, Skewness, Kurtosis, MGF, PGF, characteristics function, etc. using tabular/graphical representations that are formulated with randomly generated data through R-studio.

1.1. Squares with Step-Formed Squares

A Square is a shape with four equal sides. There are several ways of square formats, out of which, very few have been considered here for our study. Chess Boards, Stamps, Floor and wall tiles, Photo frames, Clocks, etc., are examples of square shapes. Similarly, step-formed squares will resemble the conditions of Ladders, Temple Steps, Gopuram, etc.

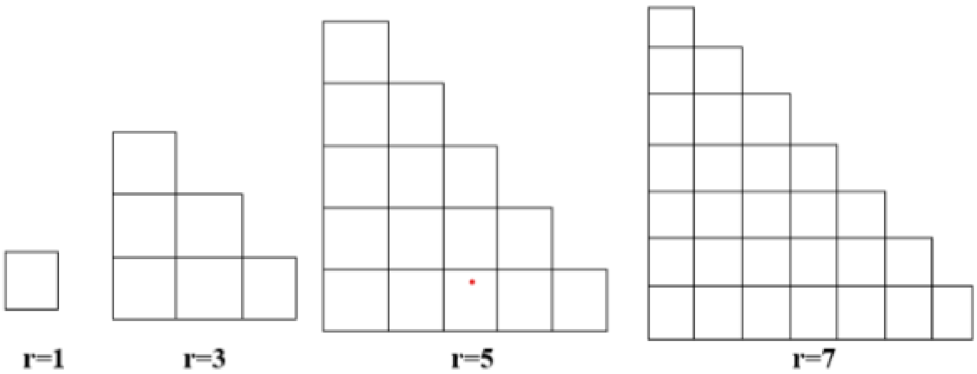


Figure 1 different forms of step-formed squares

Step-formed squares look like stairs starting with a single square, then 2, 3, 4, etc. From these different unit-length step-formed squares, we aim to determine the patterns of the step formed with the maximum possible length. In the preceding section, we discuss how patterns will change for the different unit length step formed squares (baseline is odd) with diagrams and explanations.

1.2. About the Context

This section discusses how patterns will change for the different unit length step formed squares (baseline odd) with diagrams and explanations. Now let us consider r = other unit length step formed squares. By observing Figure 1 of the step forming square, when r=1, the number of squares on the baseline is 1, and the maximum possible length is 1. When r =3, the Number of squares with 1-unit length = 1+2+3=6, the number of squares with 2-unit length = 1, number of squares with 3-unit length = 0. Similarly, for higher order, let r=5, i.e. the baseline length is five and the maximum possible length is 3. For this pattern, the number of squares with 1-unit length = 1+2+3+4+5 = 15, the number of squares with 2-unit length = 1+2+3 = 6, the number of squares with 3-unit length = 1 and the Number of squares with 4-units and 5 units length is = 0.

Let r =7, i.e. the baseline length is seven, and the maximum possible length is four as the number of squares with 1-unit length = 1+2+3+4+5+6+7= 28, number of squares with 2-unit length = 1+2+3+4+5= 15, number of squares with 3-unit length = 1+2+3= 6, number of squares with 4-unit length = 1 and the number of squares with 5, 6 and 7 –units of length = 0. For r=9, i.e. the baseline is nine, and the maximum possible length is 5. Number of squares with 1-unit length = 1+2+3+4+5+6+7+8+9= 45, Number of squares with 2-unit length = 1+2+3+4+5+6+7=28, Number of squares with 3-unit length = 1+2+3+4+5 =15, Number of squares with 4-unit length = 1+2+3=6, Number of squares with 5-unit length = 1 and Number of squares with 6-, 7-, 8- and 9-units length = 0. Hence, the baseline will generally be extended to (r+1)/2 with the maximum number of possible squares equal to 1.

1.3. Relevant Reported Research Studies

Klamkin. M.S. (1987), From this book, learned how to implement mathematical modelling in real-life applications like traffic flow, electronic networks, medical, sports, and so on [3]. A.A. Samarskiiet al. (2001), explained different mathematical models in nature with differential equation techniques [6]. F. Blanchet-Sadri et al. (2009), introduced the counting of distinct squares in partial words, i.e., in DNA sequence string [2]. Koh Khee Meng et al. (2013), From this book, got ideas on developing mathematical modelling in different techniques like the addition principle, multiplication principle, bijection principle, etc., in combinatorics. And also learned how to apply counting techniques in well-defined shapes like squares, triangles, circles, etc., with examples [7]. Elise Lockwood et al. (2015) described the problem involving combinations. They have expressed an essential aspect of their activity that refers to the combinatorial encoding of outcomes and use this language to analyze the work. Bruce E. Sagan (2020) [4], From this book, learned how to make combinations and permutations from different cases with different from the methodology inclusive and exclusive principle, matrix tree method, exponential functions, and so on [5]. Vito Barbarani (2021), This text is about a research paper divided into two parts. In the first part, the article looks at a new class of combinatorial objects and examines how they relate to the distribution of prime numbers. It also looks at the probability distribution of the n-th prime number and provides an estimate of the prime-counting function. The second part of the paper looks at generalizing the model to investigate the conditions that enable both the Prime Number Theorem and the Riemann Hypothesis. Finally, it discusses a heuristic version of the model related to the sequence of primes [1].

1.4. Research Gap and Motivation of Study

After a thorough search of the literature on Combinatorics probability theory, it is observed that more work needs to be reported on the formulation of probability distributions and modelling of Combinatorics theory with discrete stochastic processes. Most of the time, the researchers have attempted to spell out computing different probabilities using the Combinatorics theory. Until now, probability theory has been the review stuff using combinatorial mathematical modelling, permutations, and counting techniques. However, there must be evidence of building probability distribution models using the Combinatorics theory, specifically, to the number of squares within step-formed squares. Hence the working domain is of a pure virgin. No attempt has been reported on the formulation of combinatorial probability distributions for the sequences of odd-numbered baseline step-forming squares. These studies are innovative and exciting. All these factors provoked us to make the thought processing in this direction. The study has discussed applying Combinatorics techniques to explore new probability distributions. Understanding the model behaviour through its statistical properties is a comprehensive study. The robustness of the probability distribution theory has been carried out with several statistical properties. The appropriate analysis is carried out with tabular and graphical representations.

2. Mathematical Model

2.1. Sample Space and Subspace

Observing the schematic diagrams, the generalized sample space of different unit-length squares from the step-formed square (with baseline odd number) is discussed below with mathematical equations and results.

The number of squares with a 1-unit length is $1 + 2 + 3 + 4 + \ldots\ldots + r = \frac{r(r+1)}{2}$

The number of squares with a 2-unit length is $1 + 2 + 3 + 4 + \ldots\ldots + (r-2) = \frac{(r-2)(r-1)}{2}$

The number of squares with a 3-unit length is $1 + 2 + 3 + 4 + \ldots\ldots + (r-4) = \frac{(r-4)(r-3)}{2}$

The number of squares with a 4- unit length is $1 + 2 + 3 + 4 + \ldots\ldots + (r-6) = \frac{(r-6)(r-5)}{2}$

The number of squares with a 5-unit length is $1 + 2 + 3 + 4 + \ldots\ldots + (r-8) = \frac{(r-8)(r-7)}{2}$

The number of squares with a 6-unit length is $1 + 2 + 3 + 4 + \ldots\ldots + (r-8) = \frac{(r-8)(r-9)}{2}$

The number of squares with a 7-unit length is $1+2+3 + 4 + \ldots\ldots + (r-10) = \frac{(r-10)(r-9)}{2}$

The number of squares with k-unit length is 1+2+3+4+………. + (r-2(k-1)) $=\frac{(r-2k+2)(r-2k+3)}{2}$

The number of squares with $\frac{r+1}{2}-1$ length is 6

The number of squares with $\frac{r+1}{2}$ length is 1

The sequence of possible cases is given by

$$T = 1 + 6 + 15 + \ldots + \frac{(r-2k+2)(r-2k+3)}{2} + \ldots\ldots + \frac{(r-2)(r-1)}{2} + \frac{r(r+1)}{2} = \sum_{l=1}^{\frac{r+1}{2}} \frac{(r-2l+2)(r-2l+3)}{2}$$

Now, the total possible number of unit lengths is T and is given by $T=\frac{3+10r+9r^2+2r^3}{24}$

Favourable Subspace is obtained by observing the patterns from the diagram and possible cases the favourable space for different unit length step formed squares N is given by

$$N = \frac{(r-2l+2)(r-2l+3)}{2}$$

2.2. Probability Mass Function

Let us assume X is a non-negative discrete random number. The probability mass function of the discrete combinatorial distribution of step-formed squares with odd baseline numbers. By using probability definition PMF is given by

$$E(X^u) = \frac{12}{3+10r+9r^2+2r^3} \sum_{x=1}^{\frac{r+1}{2}} (4x^{u+2} - x^{u+1}(4r+10) + r^2 + 2r + 6), \forall u = 1,2,3,4$$

Here x= the number of squares with different unit lengths, and r is the parameter.

Verification of PMF

From the above probability mass function x and r takes any positive integers so p(x)>0 $\forall x$

$$\sum_{x=1}^{(r+1)/2} P(X=x) = \sum_{x=1}^{(r+1)/2} \frac{12(r-2x+2)(r-2x+3)}{2r^3+9r^2+10r+3} = \sum_{x=1}^{(r+1)/2} \frac{72+60r+12r^2-120x-48rx+48x^2}{2r^3+9r^2+10r+3} = 1$$

2.3. Cumulative Distribution Function

The cumulative distribution function of the given probability mass function is given by

$$F(X) = P(X \le t) = \sum_{x=1}^{t} \frac{12(r-2x+2)(r-2x+3)}{2r^3+9r^2+10r+3}$$

$$F(X) = \begin{cases} 0, & t < 1 \\ \frac{4(5t+9rt+3r^2t-9t^2-6rt^2+4t^3}{2r^3+9r^2+10r+3}, & 1 \le t \le \frac{r+1}{2} \\ 1, & t > \frac{r+1}{2} \end{cases}$$

2.4. Statistical Characteristics

The statistical characteristics are follows

$$E(X^u)=\frac{12}{3+10r+9r^2+2r^3}\sum_{x=1}^{\frac{r+1}{2}}(4x^{u+2}-x^{u+1}(4r+10)+r^2+2r+6), \forall u=1,2,3,4$$

$$\mu=\frac{5+6r+r^2}{4(1+2r)}$$

$$V(X)=3.\frac{r^4+4r^3-4r^2+4r-5}{80(1+2r)^2}$$

$$\mu_3=\frac{r(-110+33r+56r^2+14r^3+6r^4+r^5}{160(1+2r)^3}$$

$$\mu_4=\frac{3(385+1832r-3024r^2-1960r^3+2378r^4+152r^5+120r^6+104r^7+13r^8}{8960(1+2r)^4}$$

$$\beta_1=\frac{20r^2(22+11r+2r^2+r^3)^2}{27(-1+r)(5+r)(1+r^2)^3}$$

$$\gamma_1=\frac{2}{3}\sqrt{\frac{5}{3}}\sqrt{\frac{r^2(22+11r+2r^2+r^3)^2}{(-1+r)(5+r)(1+r^2)^3}}$$

$$\beta_2=\frac{-77-428r+247r^2+504r^3-23r^4+52r^5+13r^6}{21(1+r^2)^2(-5+4r+r^2)}$$

$$\gamma_2=-\frac{2(-35-1196r+901r^2+1008r^3+37r^4+4r^5+r^6)}{21(1+r^2)^2(-5+4r+r^2)}$$

$$M_x(t)=\sum_{x=1}^{(r+1)/2}e^{tx}\frac{12(r-2x+2)(r-2x+3)}{2r^3+9r^2+10r+3}$$

$$P_x(s)=\sum_{x=1}^{(r+1)/2}s^{x}\frac{12(r-2x+2)(r-2x+3)}{2r^3+9r^2+10r+3}$$

$$\phi_x(t)=\sum_{x=1}^{(r+1)/2}e^{itx}\frac{12(r-2x+2)(r-2x+3)}{2r^3+9r^2+10r+3}$$

3. Numerical Illustration and Graphical Representation

3.1. Probability Mass Function

The behaviour of the probability distribution is analysed through the numerical data illustrations as a common point of view. Let us take one numerical example with a finite case. Let r = 5, then the total possible chances are 22 by using the formula, and now favourable cases for different unit square length follow, and the Maximum possible square is 3-unit length.

T=22 and $N = \frac{(5-2x+2)(5-2x+3)}{2}$

The probability distribution

x	1	2	3
P(X=x)	15/22	6/22	1/22

$E(X) = 1.3636, V(X) = 0.3223, \mu_3 = 0.2367, \mu_4 = 0.3826, \beta_1 = 1.6727, \gamma_1 = 1.2933, \beta_2 = 3.6825, \gamma_2 = 0.6825$

The probability mass function graph and table with r = 1 to 199 are considered for identifying the patterns of the distribution.

x	1	29	49	69	89	109	129	149	159	179	199
P(x)	0.02 931	0.01 972	0.01 124	0.00 515	0.00 141	0.00 001	0.00 094	0.00 423	0.00 986	0.01 784	0.02 812

When the number of unit-length step-formed squares(X) in the baseline increases, the probability value decreases to half of the part. Again, the probability value increases to the remaining half part of the lengths. Half of the curve is a positive relationship, and half is a negative one. Finally, the probability distribution is shaped like the letter 'U' with the highest probability at the two extremes and not necessarily symmetrically.

3.2. Cumulative Distribution Function

Let us assume we have a 5-unit length step-formed square in these results. Here the maximum 3-unit length square exists. So, we have a discrete integer probability of 1/22. The cumulative probability that 1-unit length step formed squares is 15/22. The cumulative probability that 2-unit length steps formed squares is 21/22. Similarly, for 3-unit square length is 1. Again, the table and graphical representation are shown when r = 99.

x	5	9	15	19	25	29	35	39	45	49
P(X≤x)	0.26 74	0.48 23	0.650 50	0.77 76	0.86 94	0.93 17	0.97 01	0.99 05	0.99 85	0.99 99

When the number of unit-length step-formed squares(X) values increases, cumulative probability also increases positively to the 39-unit length. Then the incremental value is fixed at one and constant till 49.

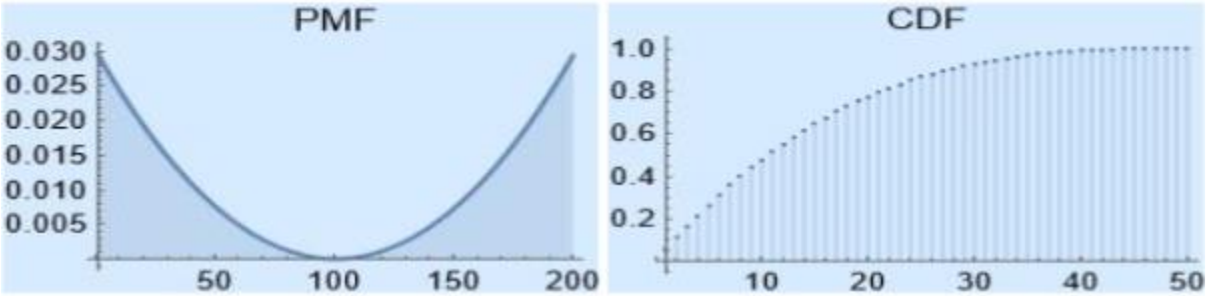
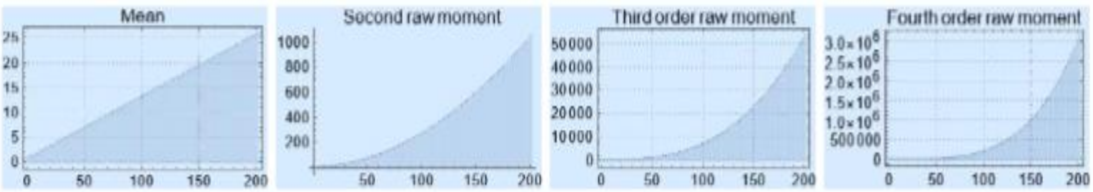


Figure 2 probability distribution

3.3. Statistical Characteristics Table

X	1	29	49	69	89	109	129	149	159	179	199
μ	1	3.076923	5.569620	8.067227	10.566038	13.065327	15.564854	18.064516	20.564263	23.064067	25.563910
σ^2	0	3.85562	15.290723	34.227413	60.664507	94.601763	136.039099	184.976483	241.413895	305.351327	376.788773
μ_3	0	6.75594	51.92090	173.02274	407.56193	793.03857	1366.95268	2166.80428	3230.09337	4594.31996	6296.98403
μ_4	0	4.2984	56.2170	263.5080	799.1929	1901.5815	3874.2733	7086.1579	11971.4151	19029.5147	28825.2168
$\beta 1$	4.1666	0.79632	0.7540505	0.7465934	0.7440176	0.7428326	0.7421911	0.7418051	0.7415550	0.7413837	0.741263
$\gamma 1$	2.0412	0.89236	0.8683608	0.8640563	0.8625645	0.8618774	0.8615051	0.8612811	0.8611359	0.8610364	0.86093
$\beta 2$	5.1666	3.11606	3.099005	3.096705	3.096005	3.095707	3.095554	3.095464	3.095408	3.095371	3.095344
$\gamma 2$	2.1666	0.11606	0.09900502	0.09670494	0.09600536	0.09570702	0.09555352	0.09546448	0.09540835	0.09537073	0.09534431

3.4.Statistical Characteristics Graphs



(a) (b) (c) (d)

Figure 3First four Raw moments of the distribution

By observing the above four raw moments, firstly, we have a mean plot when the number of different unit length step formed squares(X) values increases, the mean value also increases ideally, so there is a perfect positive relationship between X values and mean. Finally, it looks like an ideal positive increment line. Similarly, the second raw moment is slightly constant and again increasing positively. Third and fourth-order raw moments look like 'J' shaped curves.

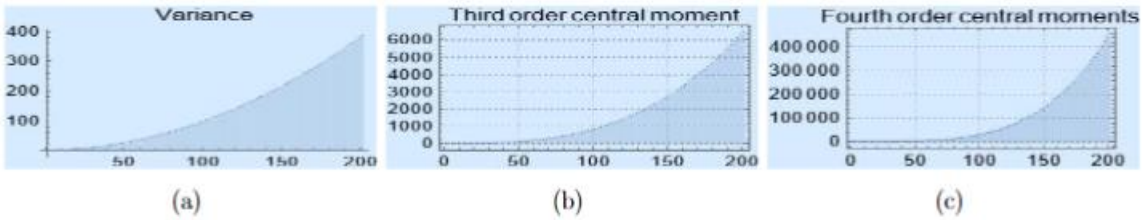


Figure 4First four central moments of the distribution

All the central moments start from zero. When the number of different unit-length step-formed squares increases, the variance curve also increases positively. Finally, it looks like the letter "J" shaped curve and exponential positive. Similarly,μ_3 and μ_4 also.

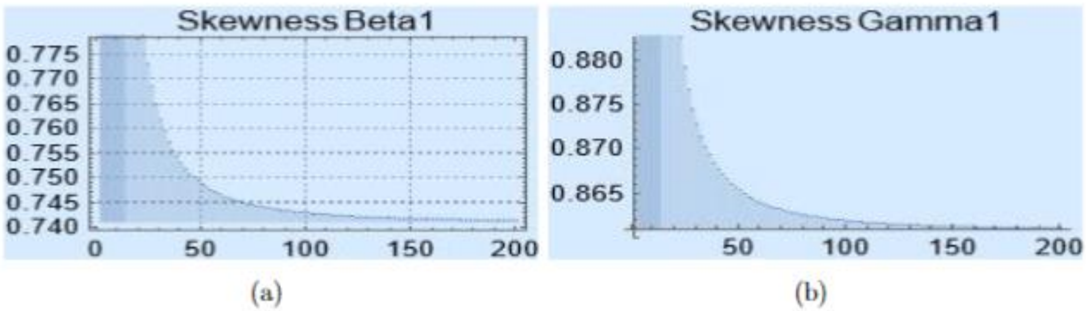


Figure 5coefficient of Skewness

When the number of different unit length step formed squares increases, then $\beta_1$value starts at 4.1667. Downwards to approximately 0.796321 at x=29, then is approximately constant from 69 onwards till 100, indicating positive skewness, i.e., mean>median>mode. when x values increase then $\beta_1$value decreases. Similarly, When the number of different unit length step formed squares increases, then γ_1 value starts at 2.04124 and then downwards to approximately 0.86836 at x=29. It is approximately constant till 100, so the right tail is more, and it indicates positive Skewness, i.e., mean>median>mode. When x values increase, then γ_1 value decreases.

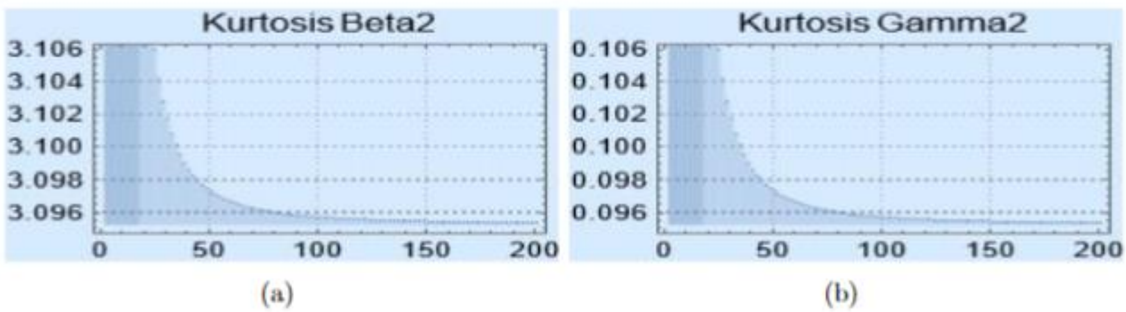


Figure 6 coefficient of Kurtosis

When the number of different unit length step formed squares increases, then the $\beta_2$value starts at 5.6667, then downwards to approximately 3.116061 at x=29, then from x=49 onwards constant till 100. when X values increase then β_2 value decreases. Similarly, the number of different unit length step formed squares increases then γ_2 the value starts at 2.16667 and then downwards to approximately 0.1160611 at x=29, then from x=69 on wards constant till 100, and it indicates positive Kurtosis.

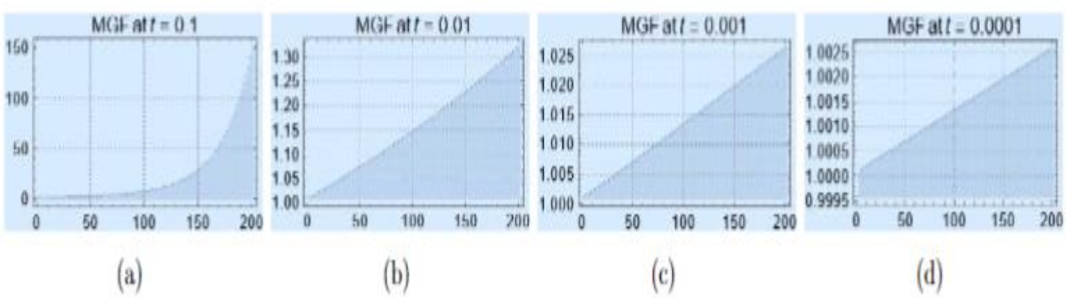


Figure 7Moment generating function with different values of t

Observing the MGF plot with different t values, moment values also increase when the r-value increases.

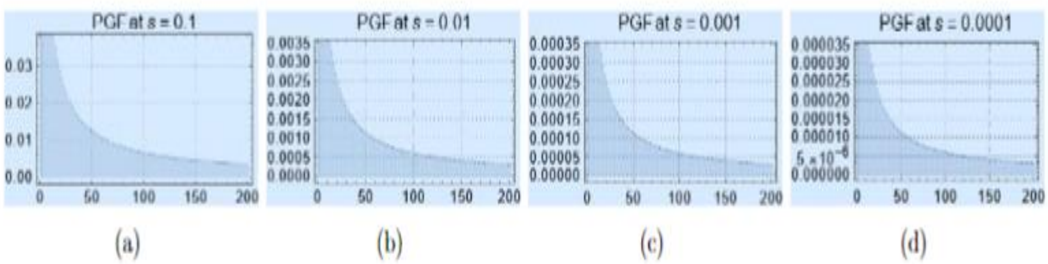


Figure 8Probability generating function with different values of S

By observing the PGF plot with different S values when the r-value increases probability value decreases.

4. Results and Conclusion

This paper finds the combinatorial probability distribution of a sequence of overlapping squares within a step-formed square. It is observed that if the baseline unit length is an odd-numbered square, then the maximum possible unit length square is $\frac{(number+1)}{2}$, and for finding different unit length squares is $\frac{(r-2x+2)(r-2x+3)}{2}$. Where r = different unit length, say 1,3, 5, , ,. and the sample is $T=\frac{3+10r+9r^2+2r^3}{24}$ and favorable space is

$N = \frac{(r-2x+2)(r-2x+3)}{2}$. By using the definition of probability, probability mass function is $P(X = x) = \frac{12(r-2x+2)(r-2x+3)}{3+10r+9r^2+2r^3}, x = 1,2,3,4,5,...,.r > 0$

The randomly generated data from the R-studio discussed the statistical properties. The probability mass function is a 'U' shaped curve, and the probability maximum attains at two points and is not necessarily symmetric. The Cumulative distribution function slightly increases still 40-unit square length, then stationary till 50. The mean value is growing positively, and the variance is a 'J' shaped curve, concluding that the variance is greater than the mean. Similarly, both μ_3 and μ_4 are positive and 'J' shaped curves. Skewness is positive, and Kurtosis is leptokurtic. From the MGF with different t values, when the r-value increases, then moment values also increase. From the PGF with different S values, when the r-value increases probability value decreases.

References

[1] R VitoBarbarani. "Combinatorial Models of the Distribution of Prime Numbers". inMathematics:9.11 (2021).

[2] Francine Blanchet-Sadri, Robert Merca and Geoffrey Scott. "Counting distinct squares in partialwords". InActa Cybernetica: 19.2 (2009), pages 465–477.

[3] Murray S Klamkin. Mathematical Modelling: Classroom Notes in Applied Mathematics. SIAM,1987.

[4] Elise Lockwood, Craig A Swinyard and John S Caughman. "Modeling outcomes in combinatorialproblem solving: The case of combinations". InProceedings for the eighteenth special interest group of the MAA on research on undergraduate mathematics education: (2015), pages 601–696.

[5] Bruce E Sagan. Combinatorics: The art of counting. volume 210. American Mathematical Soc.,2020.

[6] Alexander A Samarskii and Alexander P Mikhailov. Principles of mathematical modelling: Ideas,methods, examples. CRC Press, 2001.

[7] Tin Lam Toh andothers. Making mathematics more practical: Implementation in the schools.World Scientific Publishing Company, 2013.