Stochastic Modeling and Applications
Vol. 24 No. 2 (July-December, 2020)

Received: 15th September 2020 Revised: 19th October 2020 Accepted: 24th October 2020

MARKOV MODEL FOR SEED FERTILIZATION IN MULTI-SEED FRUITS USING TRUNCATED EXPONENTIAL DISTRIBUTION

TIRUPATHI RAO PADI, MOHAMMED HISHAM M*, AND ABARNA. S

Abstract. Understanding the dynamics of flower pollination in the process of seed formation through stochastic modelling is attracting much attention from researchers in mathematical biology. Pattern identification based on prior probability structures is a significant approach for assessing the success of fertilization processes among plants. In this study, the probability of successful fertilization in the Markov models derived from transition probabilities using conditional probabilities [?] is assumed to follow the truncated exponential distribution. The derived model explains the random sequence formed in ovule fertilization in multi-seed fruits using the Markov chain property. The models also describe the chances of successful fertilization as a function of time. Sensitivity analysis is carried out to understand the model behaviour.

1. Introduction

Flower pollination and ovule fertilization are fundamental processes in plant reproduction. Modeling of such a phenomenon to predict success in reproduction has always remained as an important field of research. Mathematical modeling and stochastic modeling has helped us in understanding the various mechanisms in plants[?, ?]. The human interest lies in finding the yield of crop production and the quality of seed formation.

Sequence modeling is one of the major research areas in which many researchers are trying to find new methods to answer various problems. Markov models have a strong ability to address these problems. In this paper, we try to explain the success rate of seed formation after the process of ovule fertilization. In fruits that give rise to more than one seed, fertilization occurs multiple times as a sequence. As the ovules are stacked, we can consider them arranged in an array. Only those ovules that successfully mate with male gametophyte convert themselves into full-fledged seeds[?, ?]. Full-fledged seeds are the quality ones that can germinate[?]. Formation of full-fledged seeds are considered as successes and, the remain remaining cases are a failures.

2000 *Mathematics Subject Classification.* Primary 91B70; Secondary 65C40, 60J25, 60J27.
Key words and phrases. Stochastic models, Multi-Ovule Fertilization, Markov Chains, Transition Probabilities, Truncated Exponential Distribution.
* Corresponding Author.

In plants, the number of flowers at the starting stage will be high, and not all the flowers turn into fruits. Those flowers that turn into multi-seed fruits will have multiple numbers of ovules[**?**]. The principal objective of this study is to find the probability associated with the successful ovule fertilization in multi-seed fruits. Consider that the ovules in multi-seed fruits are in an array, and the fertilization happens in succession. The model is derived in Section 2, and its behavior is studied in Section 3.

2. Stochastic Model

A simple two-state Markov model with probabilities p_1 and p_2 is shown in Figure 1. State 1 and 0 represent success and failure respectively. A two-state Markov model with assumed probability p_1 can be used to predict the n^{th} random event occurring in sequence. Two conditional distributions arrived from the transition probability matrix for finding the success of an n^{th} trial based on the first trial, assuming that the success and failure happen independently of any phenomena[**?**]. The two distributions are shown in Equation 2.1 and 2.2.

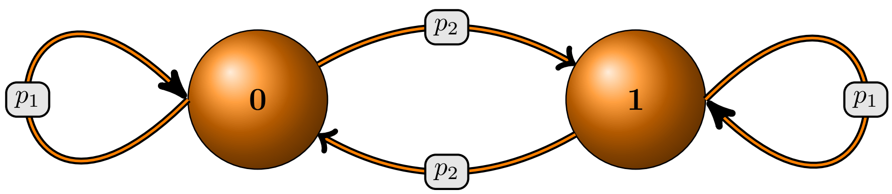


FIGURE 1. Schematic diagram for Two-State Markov Model

$$p_{0j}^n = \Pr\{X_n = j | X_0 = 0\} = \begin{cases} \frac{1}{2} + \frac{1}{2}\sum_{r=0}^{n}\binom{n}{r}(-p_2)^r p_1^{n-r} & ; j = 0 \\ \frac{1}{2} - \frac{1}{2}\sum_{r=0}^{n}\binom{n}{r}(-p_2)^r p_1^{n-r} & ; j = 1 \\ 0 & ; otherwise \end{cases} \tag{2.1}$$

where $0 \le p_1 \le 1$ and $p_1 + p_2 = 1$.

$$p_{1j}^n = \Pr\{X_n = j | X_0 = 1\} = \begin{cases} \frac{1}{2} - \frac{1}{2}\sum_{r=0}^{n}\binom{n}{r}(-p_2)^r p_1^{n-r} & ; j = 0 \\ \frac{1}{2} + \frac{1}{2}\sum_{r=0}^{n}\binom{n}{r}(-p_2)^r p_1^{n-r} & ; j = 1 \\ 0 & ; otherwise \end{cases} \tag{2.2}$$

where $0 \le p_1 \le 1$ and $p_1 + p_2 = 1$.

The random variable associated is a binary variable, which takes values 0 and 1 defined as,

$$X(\omega) = \begin{cases} 0 & ; \text{If the } n^{th} \text{ ovule fertilization is failure} \\ 1 & ; \text{If the } n^{th} \text{ ovule fertilization is success} \end{cases}$$

The arrival of male and female gametophytes plays a vital role in the success of ovule fertilization. Exponential distributions are adopted to model the inter-arrival time of a process[?, ?, ?]. In the case of ovule fertilization, time interval during which the successful fertilization can happen is 18 days[?]. Hence using truncated exponential distribution will be more appropriate than exponential distribution as a prior probability distribution.

The probability density function of truncated exponential distribution[?] is given by,

$$f(x) = \frac{\frac{1}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}};\ 0 < x \leq b,\ \theta > 0 \tag{2.3}$$

Assuming p_2 follows truncated exponential distribution given by Equation 2.3, the Equation 2.1 can be modified as,

$$p_{0j}^{n} = \Pr\left\{X_n = j | X_0 = 0\right\} = \begin{cases} \frac{1}{2}+\frac{1}{2}\left(1-\frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n & ; j = 0 \\ \frac{1}{2}-\frac{1}{2}\left(1-\frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n & ; j = 1 \\ 0 & ; otherwise \end{cases} \tag{2.4}$$

and Equation 2.2 as,

$$p_{1j}^{n} = \Pr\left\{X_n = j | X_0 = 1\right\} = \begin{cases} \frac{1}{2}-\frac{1}{2}\left(1-\frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n & ; j = 0 \\ \frac{1}{2}+\frac{1}{2}\left(1-\frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n & ; j = 1 \\ 0 & ; otherwise \end{cases} \tag{2.5}$$

The Equations 2.1 and 2.2 explain the success rate of ovule fertilization in a multi-seed fruit when the fertilization does not depend on how and when the process takes place. The derived new models (Equations 2.4 and 2.5) will explain how the success rate increases or decreases as a function of time(new variable x).

As Equations 2.4 and 2.5 satisfies all the necessary assumptions for being a probability distribution, we can derive all its statistical measures and hence comment on how these models behave. Let Equation 2.1 and Equation 2.2 be Model-1 and Model-2 respectively. Various statistical properties of Model-1 and Model-2 are derived in Section 2.1 and 2.2 respectively.

2.1. Statistical Measures/Properties of Model-1. In this section, various statistical measures like mean, variance, coefficient of variation, moments, etc. of Model-1 are derived.

The mean and variance of model-1 is given by

$$\text{Mean} = \frac{1}{2} - \frac{1}{2}\left(1-\frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n \tag{2.6}$$

$$\text{Variance} = \frac{1}{4} - \frac{1}{4}\left(1-\frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n} \tag{2.7}$$

Coefficient of Variation of Model-1 is

$$\mathrm{CV} = \left(\frac{1 + \left(1 - \frac{\frac{2}{\theta} e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n}{1 - \left(1 - \frac{\frac{2}{\theta} e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n} \right)^{\frac{1}{2}} \times 100 \tag{2.8}$$

The third and fourth central moments are

$$\mu_3 = \frac{1}{4}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n - \frac{1}{4}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{3n} \tag{2.9}$$

$$\mu_4 = \frac{1}{16}\left(1 - \left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}\right)\left(1 + 3\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}\right) \tag{2.10}$$

Shaping measure(β_1) and Peakedness measure(β_2) of Model-1 are

$$\beta_1 = \frac{4\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}}{1 - \left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}} \tag{2.11}$$

$$\beta_2 = \frac{1 + 3\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}}{1 - \left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}} \tag{2.12}$$

Characteristic Function($\Phi_x(t)$) is given by

$$\Phi_x(t) = \frac{1}{2} + \frac{1}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n + \frac{e^{it}}{2} - \frac{e^{it}}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n$$

Moment Generating Function($M_x(t)$) is given by

$$M_x(t) = \frac{1}{2} + \frac{1}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n + \frac{e^{t}}{2} - \frac{e^{t}}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n$$

Probability Generating Function($P(s)$) is given by

$$P(s) = \frac{1}{2} + \frac{1}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n + \frac{s}{2} - \frac{s}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n$$

2.2. Statistical Measures/Properties of Model-2. Different statistical measures of Model-2 like mean, variance, coefficient of variation, moments, etc. are derived in this section.

The Mean of Model-2 is

$$\mathrm{Mean} = \frac{1}{2} + \frac{1}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n \tag{2.13}$$

The variance of Model-2 is

$$\text{Variance} = \frac{1}{4} - \frac{1}{4}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^{2n} \tag{2.14}$$

Coefficient of Variation of Model-1 is

$$\text{CV} = \left(\frac{1 - \left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n}{1 - \left(1 + \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^n}\right)^{\frac{1}{2}} \times 100 \tag{2.15}$$

The third and fourth central moments are

$$\mu_3 = \frac{1}{4}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^{3n} - \frac{1}{4}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^{n} \tag{2.16}$$

$$\mu_4 = \frac{1}{16}\left(1 - \left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^{2n}\right)\left(1 + 3\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^{2n}\right) \tag{2.17}$$

Shaping measure(β_1) and Peakedness measure(β_2) of Model-1 are

$$\beta_1 = \frac{4\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}}{1 - \left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}} \tag{2.18}$$

$$\beta_2 = \frac{1 + 3\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}}{1 - \left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1-e^{-\frac{b}{\theta}}}\right)^{2n}} \tag{2.19}$$

Characteristic Function is given by

$$\Phi_x(t) = \frac{1}{2} - \frac{1}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^n + \frac{e^{it}}{2} + \frac{e^{it}}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^n$$

Moment Generating Function is given by

$$M_x(t) = \frac{1}{2} - \frac{1}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^n + \frac{e^{t}}{2} + \frac{e^{t}}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^n$$

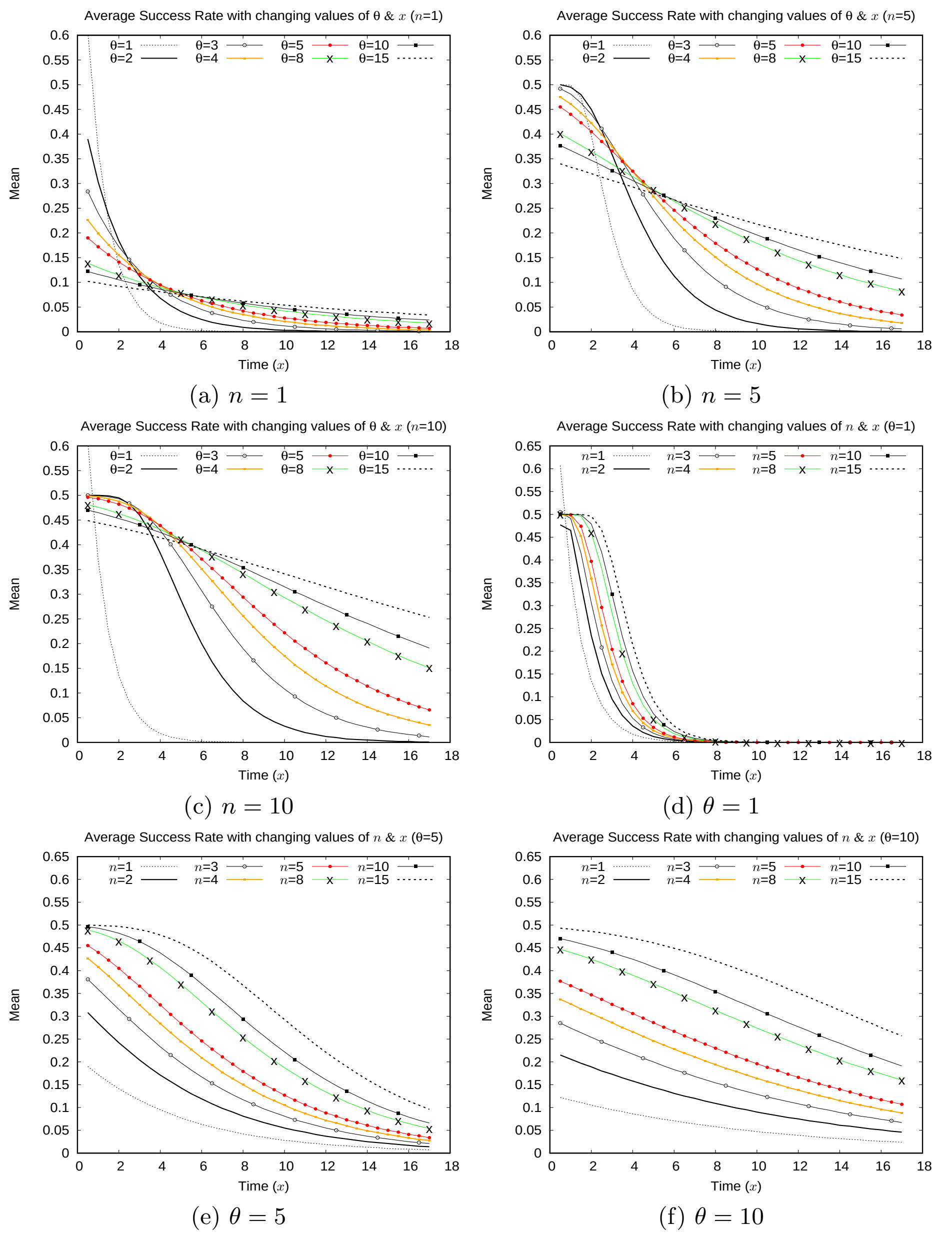
Probability Generating Function is given by

$$P(s) = \frac{1}{2} - \frac{1}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^n + \frac{s}{2} + \frac{s}{2}\left(1 - \frac{\frac{2}{\theta}e^{-\frac{x}{\theta}}}{1 - e^{-\frac{b}{\theta}}}\right)^n$$

3. Results and Discussion

In this section, the behavior of the models is explored graphically. The mean and variance of the two models are studied using different combinations of parameter values. Mean of Model-1 and Model-2 are explained in section 3.1 and 3.2 respectively. The variance of both the models are same (see equation 2.7 and 2.14). Hence the variance of both the models is explained together in section 3.3.

3.1. Mean of Model-1. The mean of Model-1 explains the average success rate of an arbitrary (say n^{th}) ovule fertilization when the initial ovule fertilization is a failure. Graphs are drawn with different values of mean of exponential distribution(θ), ovule number (n)and time (x) as shown in Figure 2.



(a) $n = 1$ (b) $n = 5$

(c) $n = 10$ (d) $\theta = 1$

(e) $\theta = 5$ (f) $\theta = 10$

FIGURE 2. Average Success Rate of Model-1

Figure 2a shows the average success rate of fertilization of the immediate next (second) ovule when the first ovule fertilization is a failure. Separate curves are drawn using different θ values. The average success rate of fertilization decreases as the time(x) increases. When $n = 1$, the mean decreases drastically and flattens as time becomes 4. As n, increases, the initial point of the average success rate in fertilization comes down, and the curves decrease gradually as the time progresses.The drastic decline in average success rate shifts its position (see Figure 2b and 2c), making a gradual decrease in the beginning and ending except for high values of $\theta(\theta = 10$ and $\theta = 15)$ where the decreases are uniform.

When $\theta = 1$, the mean decreases fast, making probability zero when time reaches 6 (see Figure 2d). Whereas when θ is 5 or 10 (Figure 2e and 2f), the mean decreases gradually over time. Also, the average success rate of fertilization decreases as the ovule number(n) increases.

3.2. Mean of Model-2. The mean of Model-2 explains the average success rate of an n^{th} ovule fertilization when the initial ovule is a success. Graphs are drawn with different values of mean of exponential distribution(θ), ovule number (n)and time (x) shown in Figure 3.

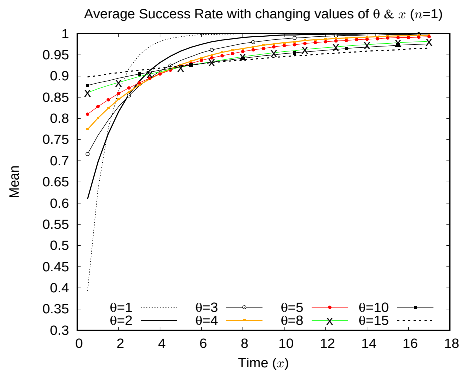
The average success rate in Model-2 increases as the time(x) progresses. The mean increases fast and flattens to 1 as time(x) reaches 4, when $n = 1$ and $\theta = 1$ (Figure 3a). The drastic increase in the average success rate gradually becomes uniform increase as θ increases. Figure 3a speaks about the immediate next(second) ovule fertilization. It is clear from the graphs as the theta value increases, the increase of the average success rate becomes smooth. Whereas for $n = 5$ (Figure 3b) and $n = 10$ (Figure 3c), the drastic increase is observed after some time.

When $\theta = 1$, the average success rate becomes 1 after the time(x) becomes 8(see Figure 3d). As θ increases, a uniform increase is observed in the mean and does not reach 1 while the time progresses. The success rate of the immediate next ovule is high and decreases as the ovule number increases. That is, the chance of fertilization of the second ovule is higher than the third ovule when the first ovule fertilization is a success. When θ is 5 and 10, we have uniformly increasing smooth curves. The model shows that the increase in the success rate of fertilization is high when the ovule number(n) is small.

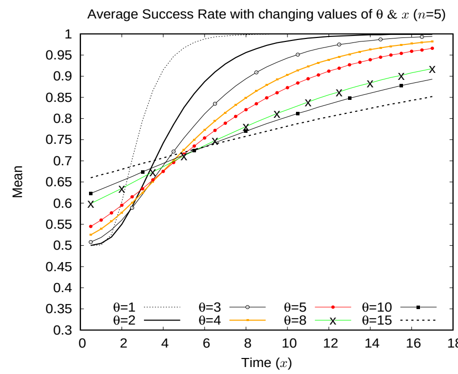
3.3. Variance of Model. Variances of both the model will behave similarly as the equations are the same. Hence the graphs are similar. Figure 4 shows the behavior of the variance for the success rate of fertilization with different values of ovule number(n), θ, and time(x). The means are either decreasing from 0.5 to 0 (in Model-1) or increasing from 0.5 to 1 (in Model-2) while the variance is decreasing from 0.25 to 0 for both the models.

The variance curves for different values of θ shows a sharp decline for small values of θ and flattens at 0 when $n = 1$ (Figure 4a). As n increases ($n = 5$ in Figure 4b and $n = 10$ in Figure 4c), a gradual shift of the sharp decline is towards right is observed for small values of θ. For high values of θ, the variance decreases gradually.

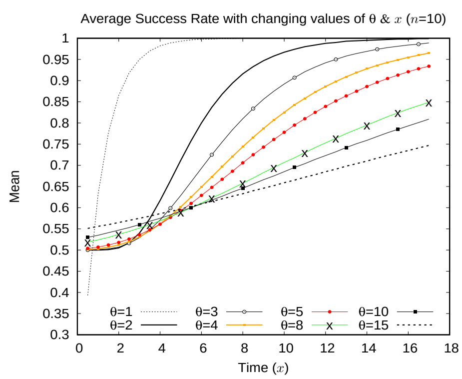
In the case of $\theta = 1$ and varying n the curves decline is very sharply and the variance goes to 0. Whereas when $\theta = 5$ (in Figure 4e) and $\theta = 10$ (in Figure 4f) gradual decreases are observed and no overlapping.



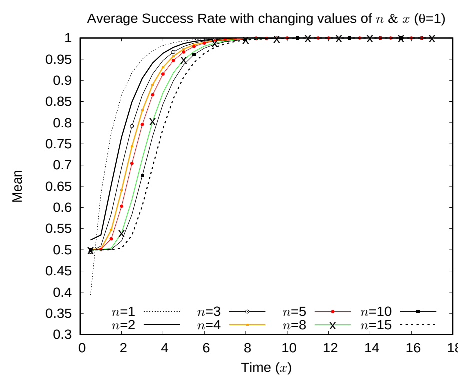
(a) $n = 1$



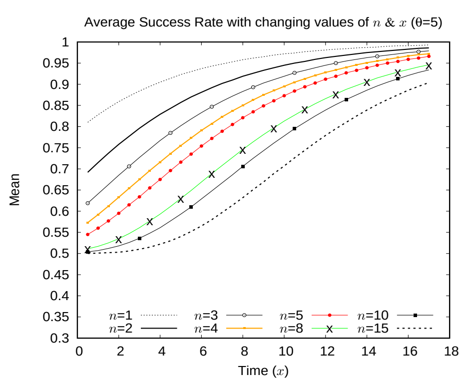
(b) $n = 5$



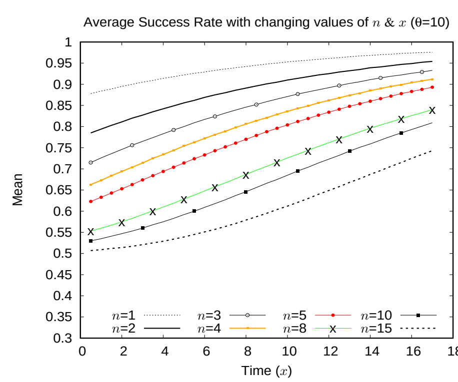
(c) $n = 10$



(d) $\theta = 1$



(e) $\theta = 5$



(f) $\theta = 10$

FIGURE 3. Average Success Rate of Model-2

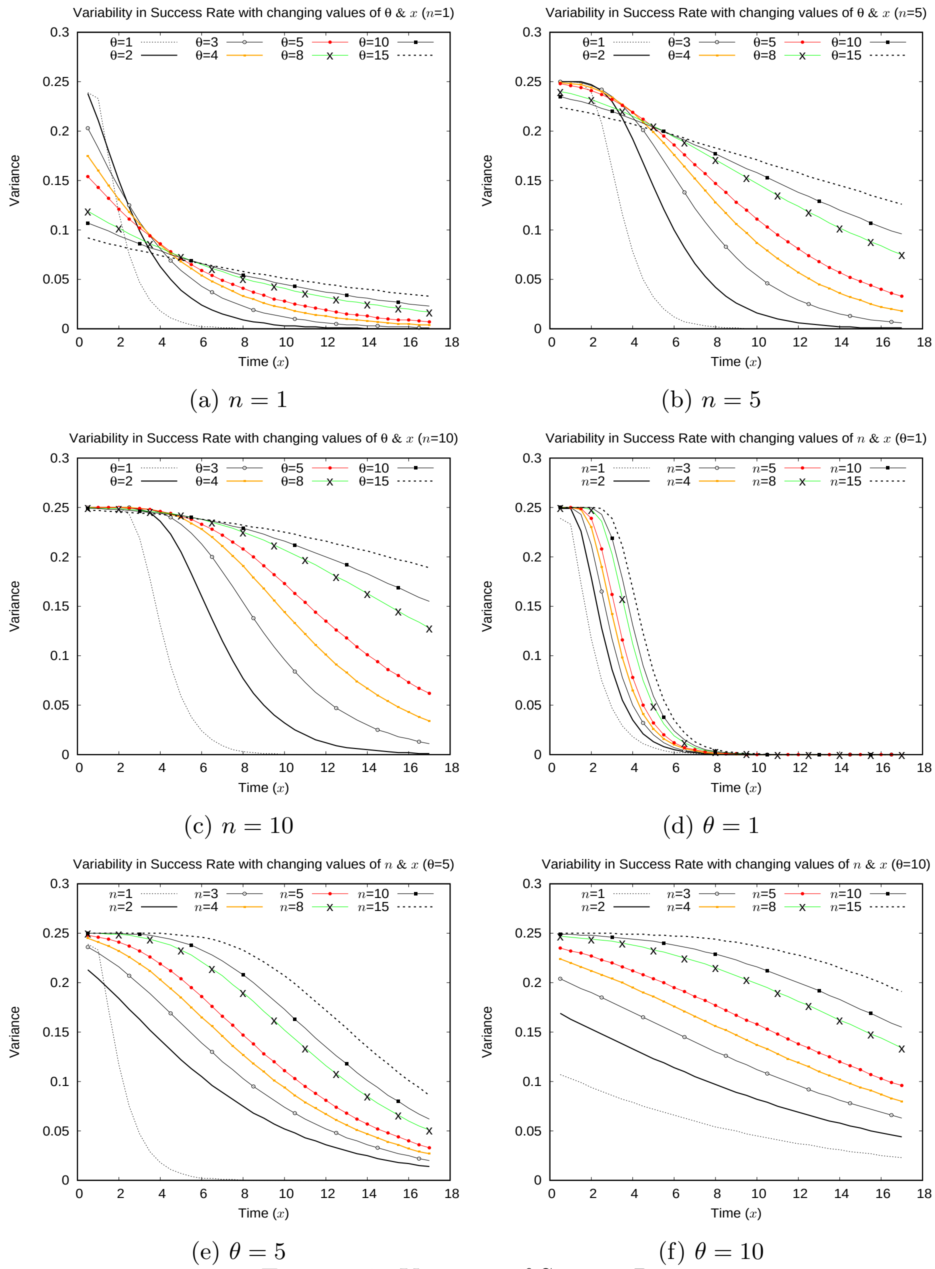


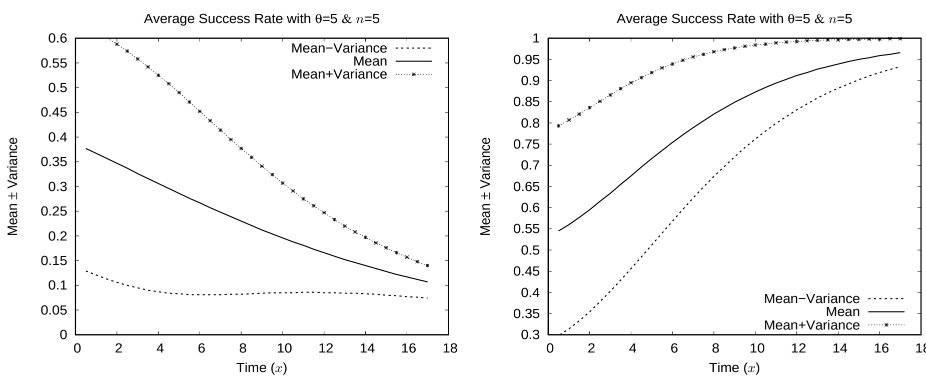
FIGURE 4. Variance of Success Rate

3.4. Comparison of Model-1 and Model-2. In this section, a comparison between Model-1 and Model-2 is carried out. Table 1 show the mean and one standard deviation range about the mean for Model-1 and Model-2. Figure 5 shows the graph of mean $\pm$ variance of Model-1 and Model-2. The Mean of

Model-1 shows a negative slope (Figure 5a) while the mean of Model-2 shows a positive slope (Figure 5b). Since the variance of both the model had a negative slope, the spread between mean+variance and mean-variance narrows down as the time progress.

Table 1. Mean and Range of Mean ± Variance for Model-1 and Model-2 with $\theta = 5$ & $n = 5$

	Model-1		Model-2	
Time(x)	Mean	Mean ± Variance	Mean	Mean± Variance
1	0.367	(0.121,0.614)	0.56	(0.314,0.807)
3	0.326	(0.095,0.558)	0.634	(0.403,0.866)
5	0.286	(0.083,0.49)	0.716	(0.512,0.919)
7	0.248	(0.081,0.414)	0.789	(0.623,0.956)
9	0.212	(0.084,0.341)	0.849	(0.72,0.977)
10	0.196	(0.085,0.307)	0.873	(0.762,0.984)
12	0.166	(0.085,0.247)	0.912	(0.831,0.992)
14	0.140	(0.083,0.197)	0.939	(0.882,0.996)
16	0.117	(0.077,0.157)	0.959	(0.919,0.998)
17	0.107	(0.074,0.14)	0.966	(0.933,0.999)



(a) Model-1 (b) Model-2

Figure 5. Mean ± Variance of Model-1 and Model-2

4. Conclusion

This model can be used to predict the probability of an ovule transforming into a seed after successful fertilization basing on the outcome of the first ovule fertilization process. Model-1 speaks about the success rate of fertilization when the first ovule is a failure, while Model-2 deals with the case that the first ovule fertilization is a success. The Model-1 indicates that if the first ovule fertilization is a failure, then the probability of successful fertilization of the second ovule becomes low. Also, if the fertilization did not happen on the first day, then the possibility of happening on the second day will be lower. So as time progresses, the chance of successful fertilization decreases. Whereas the Model-2 states that, if the first ovule fertilization is a success, then there is a high chance that the second ovule fertilization will be a success. Also, if the fertilization did not happen on the first day, then the possibility of happening on the second day is higher. So as time progresses, the probability of successful fertilization increases, creating more likelihood for success. The flower with the Model-1 has a very less chance to transform into fruit, whereas the flower with Model-2 has a high chance to become a fruit.

Department of Statistics, Pondicherry University, Puducherry-605 014, India

Email address: drtrpadi@gmail.com

Department of Statistics, Pondicherry University, Puducherry-605 014, India

Email address: mohammed.hisham19@gmail.com

Department of Statistics, Pondicherry University, , Puducherry-605 014, India

Email address: abarna.0791@gmail.com