Presented at the DLSU Research Congress 2015
De La Salle University, Manila, Philippines
March 2-4, 2015

Success Probability of an n-Step Process With n Independent Step Probabilities (Developing and Applying an Algorithm)

Andrew Philip A. Wee and Ederlina G. Nocon
De La Salle University
**Corresponding Author: ederlina.nocon@dlsu.edu.ph*

Abstract:
A process is defined as a series of actions or steps taken in order to achieve a particular end. Currently, there is a wide range of studies involving different types of processes ranging from engineering, business, biology, to information theory. The authors of this paper are interested in a new type of process study, labeled as Process Based Strategy Model and is the first study of its kind. The process specifically looks into the success probability of an n-step process with n independent step probabilities. In our model, there are exactly n steps that lead to the desired goal Xn a success in step i leads to step $i + 1$ but a failure in it only leads to goal $X_{i?1}$ and thereby, a failure in achieving the end goal X_n . We want to maximize the success probabilities of each step in order to assure the fulfillment of the end goal X_n . The paper accomplishes this by developing theorems that adjust the success probabilities of each process steps. Another method of achieving our objective is by replacing a certain step of the process with one or more steps that results to a higher overall success probability. We develop and apply an algorithm that will yield an optimal result in terms of the success of the n-step process.

Keywords: success probability; n-step process; alternate step(s) approach

1. INTRODUCTION

This study is the first of its kind in looking into the success probability of an n-step process with n independent step probabilities. This paper will lay a foundation on dealing with this type of model with an approach that uses probabilities. The mathematical theorems and proofs of this paper are proven to be solid and can be applied in the future for real-life applications in statistics, economics, industrial engineering, business processes and even game theory. Of course further research needs to be done in those areas but the mathematics of this paper will prove to be very helpful. This paper can also help managers, strategists and the like with guiding principles in creating strategy models based on processes.

The scope of this paper includes the following:

1. Provide a clear definition of a process based strategy model.
2. Develop theorems, with rigorous mathematical proofs, that determine ways that can optimize the final output of our model by way of adjusting the probability values of some steps or replacing certain steps with additional "smaller" steps, all done with the consideration of possible limited resources.
3. Show how the process based strategy model can be applied to a certain scenario involving elements in a population.

2. THE PROCESS BASED STRATEGY MODEL

In order to clearly describe the process based strategy model, let us consider an element E (which may be a person, a company, or an entity) that is aiming to reach the intended goal X_n by going through n successive steps. E achieves only X_i by going through steps $1, 2, \ldots, i$ successfully while failing to go through the rest of the steps $i + 1, \ldots, n$. Now, for every $i = 1, 2, \ldots, n$ we associate the probability of success (s_i) with step i (S_i). This tells us that the probability of success for our desired end output X_n is given by

$$x_n = \prod_{i=1}^{n} s_i \tag{1}$$

with this equation, we assume that the probability of success of the events pertaining to all n steps are pairwise independent. The probability that the goal will fail on the first step is represented by $x_o = 1 - s_1$.

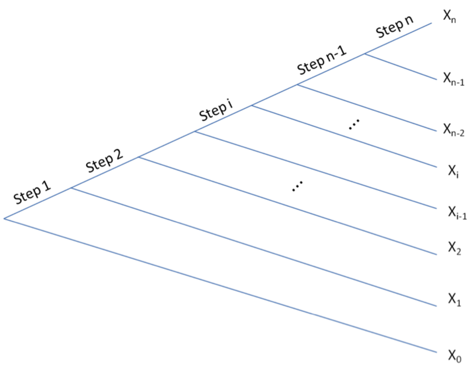

Figure 2.1: Process Based Strategy Model

The probability for a result X_j, where j is not equal to 0 or n, is the complement of the probability of step $i+1$ times the product of the probabilities of steps 1 to i. Thus, the probability of reaching goal X_i is represented by

$$x_i = (1 - s_{i+1}) \prod_{j=1}^{i} s_j \quad where \quad j = 1, 2, 3, ..., n-1.$$

Our interest is in Equation (1) with the aim of increasing its value by adjusting the different success probabilities (s_i) where $i = 1, \ldots, n$. We adjust the success probabilities of the steps in our model that will have a direct impact on (1) and develop theorems that will help us in achieving our objective.

Theorem 1. *If $(s_1, s_2, \ldots, s_n)$ is the sequence of independent step probabilities in an n-step process, then the success probability x_n of the desired output X_n satisfies*

$$a^{n-1}b \leq x_n \leq ab^{n-1}$$

where $a = min\{s_1, s_2, \ldots, s_n\}$ and $b = max\{s_1, s_2, \ldots, s_n\}$.

Proof. The value of x_n is given by $x_n = \prod_{i=1}^{n} s_i$ with $0 \leq s_i \leq 1$ for all $i = 1, 2, \ldots, n$. If we let $a = min\{s_1, s_2, \ldots, s_n\}$ and $b = max\{s_1, s_2, \ldots, s_n\}$ then we have, $x_n = s_1 s_2 \cdots a \cdots b \cdots s_n$ but $s_i \geq a$ for all $i = 1, 2, \ldots, n$ and $b = b$. Thus, we have $a^{n-1}b \leq x_n$ and also at the same time we have $s_i \leq b$ for all $i = 1, 2, \ldots, n$ and $a = a$. Thus, we have $x_n \leq b^{n-1}a$ and therefore, we have the inequality

$$a^{n-1}b \leq x_n \leq ab^{n-1}$$

where $a = min\{s_1, s_2, \ldots, s_n\}$ and $b = max\{s_1, s_2, \ldots, s_n\}$.

□

Corollary 1. *If $(s_1, s_2, \ldots, s_n)$ is the sequence of independent step probabilities in an n-step process, then the success probability x_n of the desired output X_n satisfies*

$$a^n \leq x_n \leq a$$

where $a = min\{s_1, s_2, \ldots, s_n\}$.

Proof. From Theorem 1 we have $a^{n-1}b \leq x_n$ but $a^n \leq a^{n-1}b$. Therefore, we have $a^n \leq x_n$. Now for any pair, s_j, s_k we have $s_j s_k \leq min\{s_j, s_k.\}$ Therefore,

$$x_n = \prod_{i=1}^{n} s_i \leq min\{s_1, s_2, \ldots, s_n\} = a$$

.

Thus, we have $a^n \leq x_n \leq a$.

□

Theorem 2. *An increase $s_i + F$ resulting to a decrease $s_k - F$ produces the highest increase in the desired output x_{new} and is achieved when $s_i = min\{s_1, s_2, \ldots, s_n\}$ and $s_k = max\{s_1, s_2, \ldots, s_n\}$.*

Proof. We note here that $s_i + F \leq 1$ and $s_k - F \geq 0$. To maintain our assumptions that these are probability values associated with the success in steps i and k respectively.

$$\begin{aligned} x_{new} &= s_1 s_2 \cdots (s_i + F) \cdots (s_k - F) \cdots s_n \\ &= s_1 s_2 \cdots (s_i s_k - F s_i + F s_k - F^2) \cdots s_n \\ &= \prod_{j=1}^{n} s_j - \frac{\prod_{j=1}^{n} s_j}{s_k} F + \frac{\prod_{j=1}^{n} s_j}{s_i} F - \frac{\prod_{j=1}^{n} s_j}{s_k s_i} F^2 \\ &= x_n - \frac{x_n}{s_k} F + \frac{x_n}{s_i} F - \frac{x_n}{s_k s_i} F^2 \\ &= \left[x_n + \frac{x_n}{s_i} F \right] - \frac{F}{s_k} \left[x_n + \frac{x_n}{s_i} F \right] \\ &= \left[x_n + \frac{x_n}{s_i} F \right] \left[1 - \frac{F}{s_k} \right] \end{aligned}$$

We see that x_{new} is maximised when we choose $s_i = min\{s_1, \ldots, s_n\}$ and $s_k = max\{s_1, \ldots, s_n\}$. □

In the following theorem, we factor in the possibility of having limited resources in undergoing through a process set to fulfil a certain goal.

Theorem 3. *Suppose we can increase and decrease the probability values of each s_j by certain values such that the following equations are satisfied:*

$$\sum_{j=1}^{n} s_j = W = \sum_{j=1}^{n} s_{jnew}$$

where s_{jnew} are now the new success probability values of each step depending whether they had an increase or decrease of probability values.

The new probability of success for our desired output X_n which will is denoted by x_{new}, will have the biggest increase if we choose to make each $s_j = \frac{W}{n}$.

Proof. Suppose we have s_j's where $0 \leq s_i \leq 1$ for all $i = 1, 2, \ldots, n$

$$\sum_{j=1}^{n} s_j = W$$

where W is a constant.

There are an infinite number of possibilities for the s_j's which will satisfy the equation above. By the AM-GM inequality we have:

$$\begin{aligned} \frac{\sum_{j=1}^{n} s_j}{n} &\geq \sqrt[n]{s_1 s_2 \cdots s_n} \\ \frac{W}{n} &\geq \sqrt[n]{\prod_{j=1}^{n} s_j} \\ \left(\frac{W}{n}\right)^n &\geq \prod_{j=1}^{n} s_j \end{aligned}$$

If we have $s_1 = s_2 = ... = s_n$, by the AM-GM inequality, we have

$$\left(\frac{W}{n}\right)^n = \prod_{j=1}^{n} s_j$$

Thus, all other combinations of the s_i's tell us that:

$$\prod_{j=1}^{n} s_j < \left(\frac{W}{n}\right)^n$$

However,

$$\prod_{j=1}^{n} s_j = (s_i)^n$$

Thus we have,

$$\begin{aligned} (s_i)^n &= \left(\frac{W}{n}\right)^n \\ s_i &= \frac{W}{n} \end{aligned}$$

□

Example 1. *Suppose we can reallocate the probabilities of our given model such that the sum of the new probabilities of success for all four steps will still be equal to*

$$\sum_{i=1}^{4} s_i = s_1 + s_2 + s_3 + s_4 = 0.05 + 0.45 + 0.30 + 0.22 = 1.02$$

What would be the new values of the probability of success now associated for each step so that it will give us the highest possible chance of success for X_n?
By Theorem 4, the probability of success that should now be associated for each step s_i should be 0.255 because we have $W = 1.02$ and $n = 4$ thus $\frac{W}{n} = \frac{1.02}{4} = 0.255$.
$x_n = (s_{1new})(s_{2new})(s_{3new})(s_{4new}) = (0.255)(0.255)(0.255)(0.255) = 0.004228$.

3. AN ALGORITHM FOR OPTIMIZATION

All of our above basic theorems show us fundamental principles when dealing with a process based strategy model. Based on the above theorems, we have created an algorithm which will be very helpful for maximizing and optimizing our process based strategy model. This can be seen in Algorithm 1 which is one of the main results of this paper.

The Algorithm.

Suppose we can increase $\sum_{j=1}^{n} s_j = W$ where $0 \leq W \leq n$ by a total of F by choosing to increase any combination of the s_j's such that

$$\sum_{j=1}^{n} s_{jnew} = W + F$$

where $0 \leq F \leq n - W$

Then the new probability of success for our desired output X_n which will be denoted by x_{new} will have the biggest increase if we choose to follow the following step-by-step procedure:

Step 1. List the probabilities $s_1, s_2, \ldots, s_n$ in non-decreasing order, say u_1,u_2,..., u_n. Thus,

$$u_1 \leq u_2 \leq \ldots u_n \quad and \quad \sum_{i=1}^{n} u_i = W.$$

Step 2. If $nu_n \leq W + F$, we let $u'_i = \frac{W+F}{n}$ be the new value of u_i, for $i = 1, 2, \ldots, n$. Otherwise,

Step 3. We have $nu_n > W + F$. We determine the largest integer l satisfying

$$lu_l + \sum_{i=l+1}^{n} u_i \leq W + F$$

Clearly, l exists and $1 \leq l \leq n$. We update the values of $u_1, u_2, \ldots, \mu_l$ to

$$u'_j = \frac{W + F - \sum_{i=l+1}^{n} u_i}{l}, \quad j = 1, 2, \ldots, l.$$

On the other hand, for each $i > l$, we retain the value of u_i.

Rationale for the Algorithm

Step 1 prepares us into finding the appropriate changes in the values of all probability values s_i, $i = 1, 2, \ldots, n$. We aim to attain the highest possible obtainable value for x_{new} by changing all or some of the u_i's. Our goal is to reach the sum $\sum_{j=1}^{n} u'_j = W + F$ by taking actions based on the value nu_n (increasing all of u_i to its current maximum values of n) as compared to $W + F$. This leads to two cases: $nu_n \leq W + F$ (do step 2) and $nu_n > W + F$ (do step 3).

Suppose we encounter the case $nu_n \leq W + F$. Then by Theorem 4, the highest attainable value for x_n in this situation is to set each u_i' to $u'_i = \frac{W+F}{n}$ (step 2). Now, if $nu_n > W + F$ we only have to choose some u_i to change in order to achieve the sum $\sum_{j=1}^{n} u'_j = W + F$. By applying Theorem 2, we should choose (the lower values) $u_1, u_2, \ldots, \mu_l$ with l as the largest integer satisfying

$$lu_l + \sum_{i=l+1}^{n} u_i \leq W + F.$$

We then remove the excess $nu_n - (W + F) = T > 0$ and we do this by retaining the last $(n - l)$ higher values $u_{l+1}, \ldots, u_n$ and changing the first l lower values $u_1, \ldots, u_l$ so that

$$\sum_{j'=1}^{l'} u_{j'} + \sum_{j=l+1}^{n} u_j = W + F.$$

Taking $M = W + F - \sum_{i=l+1}^{n} u_i$, we are left to determine the values of u'_j, $j' = 1, \ldots, l$ to obtain M and maximize x_n. By way of choosing the integer l and applying Theorem 4, we define $u'_j = \frac{M}{l}$, $j' = 1, 2, \ldots, l'$. (step 3).

Example 2. *In our given process based strategy model we have* $\sum_{i=1}^{4} s_i = 1.02$. *Suppose we can increase the sum by a constant value of* 0.50 *by increasing any combination of the success probabilities of our steps. What combination of increase should we choose in order to give us the highest value for* $x_n = \prod_{i=1}^{4} s_i$*?*

We apply Algorithm 1 to our problem above. According to Algorithm 1, our first step is to list the probabilities $s_1, s_2, \ldots, s_n$ *in non-decreasing order, thus* $\{s_1, s_2, s_3, s_4\} = \{0.05, 0.45, 0.30, 0.22\}$ *will now be listed as*

$$\{u_1, u_2, u_3, u_4\} = \{0.05, 0.22, 0.30, 0.45\}$$

Clearly, $\sum_{i=1}^{4} u_i = 1.02$. *Since we have* $4(u_4) = 4(0.45) = 1.8 > 1.02$ *then according to Algorithm 1 we determine the largest integer* l *satisfying* $lu_l + \sum_{i=l+1}^{n} u_i \leq W + F$ *as shown in the table below.*

l	u_l	lu_l	$\sum_{i=l+1}^{n} u_i$ $(= A)$	$lu_l + \sum_{i=l+1}^{n} u_i$ $(= B)$	$A + B \leq W + F$?
1	0.05	0.05	0.97	1.02	YES
2	0.22	0.44	0.75	1.19	YES
3	0.30	0.9	0.45	1.35	YES
4	0.45	1.8	-	-	-

The largest integer l satisfying $lu_l + \sum_{i=l+1}^{n} u_i \leq W + F$ *is* $l = 3$. *Thus, the new value for* u_1, u_2, u_3 *is now given by the equation*

$$u'_j = \frac{W + F - \sum_{i=l+1}^{n} u_i}{l} = \frac{1.02 + 0.50 - 0.45}{3} = \frac{107}{300},$$

$j = 1, 2, 3$.

For u_4 *we retain its original value of* 0.45. *The combination of increase that will give us the highest value for* x_n *is given by* $u_1 = \frac{107}{300}$, $u_2 = \frac{107}{300}$, $u_3 = \frac{107}{300}$, *and* $u_4 = 0.45$.
Thus $x_{new} = (\frac{107}{300})(0.45)(\frac{107}{300})(\frac{107}{300}) = 0.02042$

4. ALTERNATE STEP(S) APPROACH

In real life situations there will be cases wherein a step in our model will reach a ceiling in terms of its success probability. No matter how hard you try to increase the training and efficiency of a certain step, the act becomes too big of a step to facilitate the movement in the model. A proposed solution to that challenge is to create one or more alternative step(s) to facilitate the movement of elements in our model.

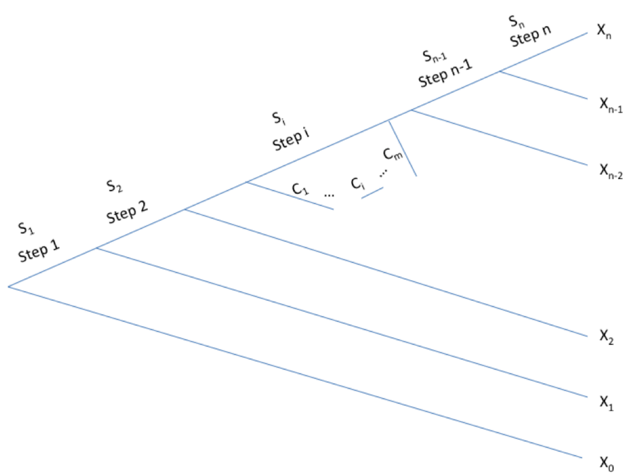

Figure 2.2: Alternate Step(s) Approach

The probabilty of success for our desired output X_n is x_n, which is equal to the product of the probability of success for all the steps $x_n = \prod_{j=1}^{n} s_i$.. However, with the given alternative step(s) in our model replacing s_i the probability of success for our desired output (x_n) is now given by the following equation:

$$x_n = \prod_{v=1}^{i-1} s_v \prod_{j=1}^{m} c_j \prod_{w=i+1}^{n} s_w$$

where the c_j's $j = 1, 2, ..., m$ are the probability of success for the alternate step(s) replacing step S_i. Our interest in this section is the product of the success probabilities of the alternate steps replacing s_i: $\prod_{j=1}^{m} c_j$.

We want to maximize its value and make sure the value is equal to or greater than s_i.

Case 1 Suppose we can find an alternate step to S_i and replace it with a different step to achieve S_{i+1}. This alternate step is denoted as C_1. The value c_1 should exceed s_i and a parameter is now set as to how larger c_1 ought to be compared to s_i. Let F be the desired incremental increase of probabilty from s_i. We want $c_1 \geq s_i + F$, $0\% \leq c_1 \leq 100\%$. Thus the possible target values for c_1 can be seen with the graph below (Figure 2.6).

In choosing c_1 to be anywhere in the interval: $[s_i + F, 100\%]$ we definitely increase the chances of moving forward to step S_{i+1} and x_n overall.

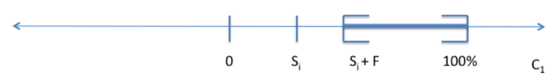

Figure 2.6: Case 1

In general we are trying to find an m consecutive alternate step(s) to replace S_i to move on to S_{i+1} that should satisfy the inequality below.

$$\prod_{j=1}^{m} c_j \geq s_i + F \quad where \quad 0 \leq c_j \leq 1 \quad \forall i \qquad (2)$$

There are an infinite number of possible combinations that can satisfy (2). If we can just choose any combination then we should go for the combination that will give us a product that is as close to 100% as possible. However, in real-life, reaching that target requires too much effort if not impossibility. The higher we go in our success rate the more time or effort is required to achieve that.

Satisfying the equation of the inequality (2), should be the minimum required effort that we want to achieve. The equation is given below.

$$\prod_{j=1}^{m} c_j = s_i + F$$

This inequality will still yield infinite possible combinations of c_i's that will satisfy the equation above. However, the sum of the combinations of the c_i's will not all be the same: $\sum_{j=1}^{m} c_j$.

Our goal is to find the combination that will satisfy the above equation and keep the summation at a minimum. The reason we want to keep the summation at a minimum is because the higher the summation is, this could possibly imply more time or effort on the part of the one who designs the model.

Theorem 4. *The combinations of c_j's that will give us the smallest possible sum for $\sum_{j=1}^{m} c_j$ that satisfies the equation $\prod_{j=1}^{m} c_j = s_i + F$ is given by*

$$c_j = \sqrt[n]{s_i + F}, \ \ \forall j = 1, \ldots, n.$$

Proof. Let $c_j = \sqrt[n]{s_i + F}$ $\forall i$ by the AM-GM inequality we have

$$\frac{\sum_{j=1}^{n} c_j}{n} \geq \sqrt[n]{\prod_{j=1}^{n} c_j}$$

$$\frac{c_1 + \ldots + c_n}{n} \geq \sqrt[n]{\prod_{j=1}^{n} \sqrt[n]{s_i + F}}$$

$$\frac{c_1 + \ldots + c_n}{n} \geq \sqrt[n]{(\sqrt[n]{s_i + F})^n}$$

$$\frac{c_1 + \ldots + c_n}{n} \geq \sqrt[n]{s_i + F}$$

but $c_1 = c_2 = \ldots = c_n$. According to the AM-GM inequality

$$\frac{c_1 + \ldots + c_n}{n} = \sqrt[n]{s_i + F}.$$

Thus, $c_j = \sqrt[n]{s_i + F}$ gives us the smallest possible combination because all other combinations gives us

$$\frac{\sum_{j=1}^{n} c_j}{n} > \sqrt[n]{s_i + F}.$$

□

5. CONCLUSION AND RECOMMENDATIONS

This paper's main interest is on X_n and increasing its overall success probability. Further study may consider computing the success probabilities of X_i where $i = 0, 1, 2, \ldots, n-1$. Also, the process based strategy model and its overall success probability x_n assumed that the probability of success of steps 1 to n are pairwise independent. The reader may wish to extend the process based strategy model where the success probabilities of the steps are dependent.

Many applications in mathematics and statistics may be applied in our model (i.e. bootstrapping, variance estimations, monte-carlo modelling etc.). One can possibly apply these techniques in mathematics and statistics with the process based strategy model. This will certainly help bridge real life applications to the discoveries of this paper.

6. REFERENCES

[1] A. K. Dixit, B. J. Nalebuff. *The Art of Strategy.* New York, London: W. W. Norton and Company.

[2] J. Herman, R. Kucera, J. Simsa, K. Dilcher (Translator) *Equations and Inequalities: Elementary Problems and Theorems in Algebra and Number Theory (CMS Books in Mathematics).* Page 151. New York: Springer-Verlag New York, Inc., 2000.