

# Determining the Number of Seeds To Sow per Cell: An Application of the Geometric Distribution

John T. Harrington and Patrick A. Glass

*Assistant professor and senior research assistant, New Mexico State University Mora Research Center, Mora, New Mexico*

*Using germination data, specifically total germination percentage, to determine sowing rates is necessary for effective resource allocation in container nurseries. Traditionally, binomial probability tables and, more recently, microcomputer programs are employed to determine sowing rates in container nurseries. Existing computer programs may require several iterations to deduce sowing rates, depending on test germination percentage. When used with the binomial probability function, the geometric probability function alleviates the need to run multiple iterations of a program. We provide programming language for a spreadsheet program that combines these two probability functions to determine sowing rates. Tree Planters' Notes 48(1/2): 28-34; 1997.*

Effective resource allocation in container nurseries can be improved by using germination percentage to determine sowing rates. Seed sowing and germinate thinning, both intrinsically linked to sowing strategy, are primary candidate areas for cost reduction in container nurseries (Wenny 1993). Germination percentage is a mathematical probability (Lipschultz 1968; Schwartz 1993). Initially, seed sowing guides were based on interpretations of binomial probability tables (Tinus and McDonald 1979). By using the factorial expansion of the binomial distribution, growers can refine their calculations using scientific calculators (Schwartz 1993). The advent of personal computers and user-friendly spreadsheet programs further simplified the use of the binomial distribution for seed sowing calculations (Wenny 1993). All these methods, when used appropriately, can improve several aspects of container production, including reducing material, labor, and greenhouse costs.

However, using approaches based on the binomial distribution function requires nursery managers to ask the question, "If X number of seeds are sown per cell, how many cells will have X number of germinates; X-1 germinates; X-2 germinates; and so on?" Depending on germination percentage, nursery managers may have to ask the question several times to achieve a satisfactory distribution. Often, the nursery managers are con-

cerned with the question "How many seeds need to be sown per cell to achieve X number of filled cells?" This question can be answered by running several iterations of the binomial distribution. The question can also be answered with one iteration, using the geometric probability distribution. Geometric distribution is easily adaptable to user-friendly commercial spreadsheet programs that are commonly available for personal computers.

## Microcomputer Application

Instructions for using the program language for the geometric probability distribution and the binomial probability distribution in concert are listed in the appendix (see page 33). Nursery managers can use this program with any commercially available spreadsheet that has natural logarithm, binomial distribution, and exponential functions. This particular formulation was developed for programs with @FUNCTION capability. It will be necessary to format your spreadsheet program as follows:

- Column A—Text format and column width of 50 characters
- Column B—Fixed numeric format with 4 decimal places
- Column C—Requires no special formatting
- Columns D through P—Fixed numeric with 2 decimal places and column width of 7 characters

Only 2 numeric entries are necessary to run this program. First, in cell B5, seedlot germination percentage is entered as an integer. A nursery manager then enters the allowable number of empty cells per 100 cells sown into cell B9 as either an integer or a real number.

## Examples of Model Applications

For our first example, let's start with a seedlot having a tested germination percentage of 95% and an accept-

able empty cell rate of 1.5%. The spreadsheet program at the end of this paper provides a value for the number of seeds necessary to sow to achieve this objective (1.4 seeds/cell; cell B13). Because sowing 1.4 seeds is impractical, the program provides a floor value, the nearest whole number below the value (in this case 1 seed/cell; cell B17) and a ceiling value, the next highest whole number to the calculated value (in this case 2 seeds/cell; cell B15). On the right-hand side of the output the program generates binomial probability distributions for these 2 recommended values (figure 1).

In a second example, with a germination percentage of 75% and an acceptable empty cell rate of 1.5%, the spreadsheet recommends sowing 3.03 seeds/cell with floor and ceiling sowing rates of 3 seeds/cell and 4 seeds/cell, respectively (figure 2).

In the last example, let's use a seedlot with a germination percentage of 60% and an acceptable empty cell rate of 1.5%. The spreadsheet recommends sowing 4.58 seeds/cell with floor and ceiling sowing rates of 4 seeds/cell and 5 seeds/cell, respectively (figure 3).

These examples demonstrate microcomputer applications of the geometric probability distribution. Understanding why the geometric distribution works requires a review of some probability distributions and their inherent differences.

### Technical Background on Probability Distributions

Three discrete probability distribution functions are useful in determining seed sowing rates. These are the Bernoulli distribution function, the binomial distribution function, and the geometric distribution function. The probability ( $P$ ) of a seed from a tested seedlot germinating (success) equals the germination percentage for that seedlot. Subtracting germination percentage, expressed in decimal form, from 1 is the probability ( $q$ ) of non-germinating (failure). Planting a single seed with these probabilities ( $p$  and  $q$ ) assigned to the outcome is an expression of the simplest probability distribution function, the Bernoulli distribution, mathematically expressed as (Dudewicz and Mishra 1988):

$$\text{EQUATION 1} \quad P_x = p^x q^{1-x} \quad x = 0,1$$

where:  $P_x$  = the probability of germinating  
 $p$  = the germination percent in decimal form  
 $q = 1 - p$

For nursery managers, the utility of this discrete probability distribution is limited because germination percentage is usually already known or assumed. However,

Determining Seed/Cell Requirements		Binomial Distributions of Ceiling and Floor Number of Seeds to Sow						
From the Geometric Probability Density Function		Number of Seeds Sown		Percentages of Number of Germinants				
		2	1	0				
Seedlot Germination %	95	2	90.25	9.50	0.25			
		1		95.00	5.00			
Acceptable Number of Empty Cells per 100	1.5							
Exact Number of Seeds to Sow per Cell	1.4019							
Ceiling Number of Seeds to Sow	2							
Floor Number of Seeds to Sow	1							
Ceiling Percentage of Empty Cells	0.2500							

Figure 1—Spreadsheet outputs, example A: seedlot with 95% germination.

Determining Seed/Cell Requirements		Binomial Distributions of Ceiling and Floor Number of Seeds to Sow						
From the Geometric Probability Density Function		Number of		Percentages of Number of Germinants				
		Seeds Sown	4	3	2	1	0	
Seedlot Germination %	75	4	31.64	42.19	21.09	4.69	0.39	
		3		42.19	42.19	14.06	1.56	
Acceptable Number of Empty Cells per 100	1.5							
Exact Number of Seeds to Sow per Cell	3.0294							
Ceiling Number of Seeds to Sow	4							
Floor Number of Seeds to Sow	3							
Ceiling Percentage of Empty Cells	0.3906							

Figure 2—Spreadsheet outputs, example B: seedlot with 75% germination.

Determining Seed/Cell Requirements		Binomial Distributions of Ceiling and Floor Number of Seeds to Sow						
From the Geometric Probability Density Function		Number of		Percentages of Number of Germinants				
		Seeds Sown	5	4	3	2	1	0
Seedlot Germination %	60	5	7.78	25.92	34.56	23.04	7.68	1.02
		4		12.96	34.56	34.56	15.36	2.56
Acceptable Number of Empty Cells per 100	1.5							
Exact Number of Seeds to Sow per Cell	4.5834							
Ceiling Number of Seeds to Sow	5							
Floor Number of Seeds to Sow	4							
Ceiling Percentage of Empty Cells	1.0240							

Figure 3—Spreadsheet outputs, example C: seedlot with 60% germination.

this probability distribution provides the foundation for 2 discrete probability distributions—binomial and geometric—applicable to developing sowing strategies.

Planting 2 or more seeds from a seedlot (that is, possessing the same germination percentage) is a practical demonstration of the binomial probability distribution. The binomial probability distribution calculates the probability of a specific number of germinations (successes) without regard to the order of those successes (Dudewicz and Mishra 1988; Lipschultz 1968). The general form of the binomial probability distribution (Dudewicz and Mishra 1988; Lipschultz 1968) is

EQUATION 2 
$$P_x = \binom{n}{x} p^x q^{n-x} \quad x = 0, 1, 2, \dots, n$$

where:  $P_x$  = the probability of occurrence (percentage of cells)  
 $n$  = the number of independent trials (number of seed sown per cell)  
 $x$  = the number of the  $i^{\text{th}}$  trial (number of germinates)  
 $p$  = the probability of germinating (success)  
 $q$  = the probability of not germinating (failure;  $1-p$ )

Each probability includes 2 components: the binomial coefficient and the Bernoulli probability raised to the power of the trial ( $x$ ). The binomial coefficient can be derived using the following factorial expansion (Schwartz 1993):

EQUATION 3 
$$(X + Y)^n = X^n + aX^{n-1}Y + \dots + aXY^{n-1} + Y^n$$

where:  $n$  = the number of independent trials (seeds)  
 $a$  = binomial coefficient  
 $X$  = germination probability  
 $Y$  =  $1 -$  germination probability

Another way to derive the binomial coefficient is to use factorial notation. The general form for determining the binomial coefficient (Dudewicz and Mishra 1988; Lipschultz 1968) is

EQUATION 4 
$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

where:  $n$  = the number of independent trials (seeds)  
 $x$  = the number of the  $i^{\text{th}}$  trial (number of germinates)

Most statistics texts denote this formula using the notation  $\binom{n}{x}$ ,  $nCx$ , or  $C(n,x)$ . This formula is also used to determine the number of combinations from  $n$  objects (seeds) choosing  $x$  at a time (Lipschultz 1968). The exponent of the Bernoulli probability signifies the number of

the trial being conducted. Therefore, when  $n = 5$ , the term  $p^3q^2$  indicates that the third of 5 trials is being computed. Note that the exponents of  $p$  and  $q$  sum to  $n$  (5 in this case). Also, the probability of all successes is given by  $p^n$  and conversely the probability of all failures is  $q^n$ , because in both cases the binomial coefficient is one (Dudewicz and Mishra 1988; Lipschultz 1968).

The geometric probability distribution has the general form (Dudewicz and Mishra 1988):

EQUATION 5 
$$P_x = (1-p)^x \quad x = 1, 2, 3, \dots$$

where:  $P_x$  = the probability of occurrence (percentage of empty cells)  
 $x$  = the number of the  $i^{\text{th}}$  trial (number of failed germinates)

This probability distribution is useful because it computes the number of trials to achieve the first success. It differs from the Bernoulli probability distribution by not restricting the number of trials to one. However, if only 1 trial is performed, the geometric probability distribution is equivalent to the Bernoulli probability distribution. The geometric probability distribution differs from the binomial probability distribution in two aspects: trials place failures before the first success and an infinite number of trials can be run.

The geometric probability distribution can be algebraically rearranged to the following form (Dudewicz and Mishra 1988; Mood and others 1974):

EQUATION 6 
$$X = \frac{\ln(P_x)}{\ln(1-p)}$$

where:  $X$  = the number of the  $i^{\text{th}}$  trial (number of seed to sow)  
 $\ln$  = the natural logarithm  
 $P_x$  = the probability of occurrence (percentage of empty cells)  
 $p$  = the germination probability

This form is useful when  $P_x$ , the desired probability of occurrence (percentage of empty cells), is known, and  $X$ , the number of trials (number of seed to sow) necessary to achieve that probability, is unknown.

The geometric probability distribution can determine the number of seeds to sow per cell to achieve a specific cell occupancy. By itself, the geometric probability distribution provides no information on distribution frequency of multiple germinates per cell. Using the two probability distributions in concert can reduce the number of iterations required when using the binomial probability distribution alone, and provide the distribution

of germination frequencies lacking in the geometric probability distribution.

### Recommendations

The most limiting feature of production varies between container nurseries, and within nurseries based on crops and cropping schedules. Sometimes, growing space is more valuable than the labor associated with thinning. Sometimes, the converse is true. For example, at New Mexico State University's Mora Nursery, growing space is at a premium during spring and early summer months. At these times, greater emphasis is placed on sowing multiple seeds per cell (that is, using ceiling values) and lowering the acceptable empty cell rate. Later in summer, when a second crop is being sown, sufficient space is usually available to over-sow the total number of cells and sow using the floor sowing rate. Using geometric and binomial probability distributions in concert provides a tool for nursery managers to evaluate different options for producing their crops.

**Address correspondence to:** Dr. John Harrington, New Mexico State University, Mora Research Center, Box 359, Mora, NM 87732; e-mail: joharrin@nmsu.edu

### Acknowledgments

This research was funded, in part, though McIntire-Stennis Grant No. 01527980 and the New Mexico Agricultural Experiment Station.

### Literature Cited

- Dudewicz EJ, Mishra SN. 1988. Modern mathematical statistics. New York: John Wiley & Sons: chapters 2 & 3.
- Lipschultz S. 1968. Schuam's outline of theory and problems of probability. New York: McGraw-Hill: 153.
- Mood AM, Graybill FA, Boes DC. 1974. Introduction to the theory of statistics. New York: McGraw-Hill: chapter 2.
- Schwartz M. 1993. Germination math: calculating the number of seeds necessary per cavity for a given number of live seedlings. *Tree Planters' Notes* 44(1): 19-20.
- Tinus RW, McDonald SE. 1979. How to grow seedlings in containers in greenhouses. Gen. Tech. Rep. RM-60. Fort Collins, CO: USDA Forest Service Rocky Mountain Forest and Range Experiment Station. 256 p.
- Wenny DL. 1993. Calculating filled and empty cells based on number of seeds sown per cell: a microcomputer application. *Tree Planters' Notes* 44(2): 49-52.

## Appendix

## Spreadsheet Programming Language

Establish the column width of column A as 50 characters wide.

Cell	Entry
A1	Determining Seed/Cell Requirements
A2	From the Geometric Probability Density Function
A5	Seedlot Germination %
A9	Acceptable Number of Empty Cells per 100
A13	Exact Number of Seeds to Sow per Cell
A15	Ceiling Number of Seeds to Sow
A17	Floor Number of Seeds to Sow
A19	Ceiling Percentage of Empty Cells
A21	Floor Percentage of Empty Cells
B13	=@LN(B9/100)/@LN(1-(B5/100))
B15	=@INT(B13)+1
B17	=B15-1
B19	=(1-(B5/100))^B15*100
B21	=(1-(B5/100))^B17*100
D1	Binomial Distributions of Ceiling and Floor Number of Seeds to Sow
D3	Number of
D4	Seeds Sown
D5	=B15
D7	=B17
F4	=D5
F5	=@IF(F4<>"",@BINOMDIST(F4,\$D\$5,\$B\$5/100,0)*100,"")
G3	Percentages of Number of Germinations
G4	=@IF(F4-1>=0,F4-1,"")
G7	=@IF(G4<>"",@BINOMDIST(G4,\$D\$7,\$B\$5/100,0)*100,"")

This spreadsheet formulation was developed on Corel Quattro Pro® for Windows 95® and has been tested on Microsoft Excel® running under Windows 95®. **The character " " represents a blank space enclosed within double quotes.** The @BINOMDIST function may be named differently depending upon the spreadsheet you are using, possible synonyms include BINOMIAL and BINOM. Refer to the help files or user's guide for the exact syntax. The formula entries in cells F5, G4, and G7 should be copied to the right within the same row into each subsequent cell through column P. For spreadsheets operating on Windows® platforms the copying procedure should be done with the clipboard copy and paste functions.

Use the following procedure to perform the copying:

1. Select the cell you wish to copy by clicking it one time with the left mouse button.
2. Then click the copy button on the tool bar or select copy on the pull-down edit menu.
3. Then select the cell where you want to place the copy by clicking it one time with the left mouse button.
4. Then click the past button on the tool bar or select paste on the pull-down edit menu.

For applications running on other operating platforms, refer to your help files or the user's guide for copying cells.

**Do not manually enter the formula into subsequent cells that you would copy the formula into, as this destroys the relative cell referencing and makes the formula invalid.**

Two entries are made into the spreadsheet:

- ▶ The seedlot germination percentage is entered into cell B5.
- ▶ The acceptable number of empty cells per 100 cells sown is entered into cell B9.

Some spreadsheets may display error messages in certain cells because of the interpretation of the formula. This is

caused by not having a value available upon which to perform the operation that the formula specifies. The error messages will occur in rows 4, 5, and 7 in the columns to the right of where the binomial distribution reaches a value of zero. If error messages occur in any of these cells—B13, B15, B17, B19, B21, D5, D7, F4, F5, G4, or G7—verify that the formula is entered correctly, specifically making sure that it is not entered as text.

---