A SunCam online continuing education course

Engineering Methods in *Microsoft Excel*

Part 4: Simulation and Systems Modeling I

by

Kwabena Ofosu, Ph.D., P.E., PTOE

Abstract

This course is part of a series that presents *Microsoft Excel* tools that are useful for a wide range of engineering analyses and data management. This course covers an introduction to simulation and systems modeling. Simulation is a technique for conducting experimentation of a system or process, virtually, on a computer, using statistical models. This course presents a detailed review of statistical distributions that are widely used in simulation studies. Real-life examples are formulated and implemented in *Microsoft Excel* and worked using the various *Excel* tools, spreadsheet techniques, and built-in functions. Examples from various engineering fields are used to demonstrate the concepts and methods learned throughout this course. Upon completion of this course, practitioners will be able to apply the methods learned to a variety of engineering problems, and also to identify situations in their fields of specialty where the innovative application of these tools and methods will be advantageous to their output and to their work product.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

List of Figures

Figure 1. 1: Framework for simulation 5
Figure 4. 1: The Uniform distribution.....58
Figure 4. 2: The Exponential distribution 63
Figure 4. 3: The Normal distribution 64
Figure 4. 4: Cumulative density.....65
Figure 4. 5: Log-Normal probability density function.....78
Figure 4. 6: Log-Normal cumulative density function.....79
Figure 4. 7: Log-Normal percent point function.....80
Figure 4. 8: Gamma probability density function.....83
Figure 4. 9: Beta probability density function.....86
Figure 4. 10: Weibull probability density function.....89
Figure 4. 11: Logistic probability density function 91
Figure 4. 12: Log-Logistic probability density function.....93
Figure 4. 13: Triangular probability density function.....94

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

TABLE OF CONTENTS

Abstract ii
List of Figures iii
1. INTRODUCTION 1
1.1 Overview 1
2. SIMULATION 2
2.1 Introduction 2
2.2 Illustration of a Simulation Model 2
3. RANDOM VARIABLES 7
3.1 Definition 7
3.2 Discrete Random Variables 8
3.3 The Binomial Distribution 11
3.3 The Poisson Distribution 14
3.4 Other Discrete Distributions 17
3.5 Discrete Random Variables in *Excel* 19
4. CONTINUOUS RANDOM VARIABLES 55
4.1 Definition 55
4.2 Properties of Continuous Random Variables 55
4.3 The Uniform Distribution 57
4.4 The Exponential Distribution 62
4.5 The Normal Distribution 64
4.6 The Log-Normal Distribution 76
4.7 The Gamma Distribution 81
4.8 The Beta Distribution 84
4.9 The Weibull Distribution 87
4.10 The Logistic Distribution 90

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4.11 The Log-Logistic Distribution 92
4.12 The Triangular Distribution 94
4.13 Other Continuous Distributions 96
4.14 Continuous Random Variables in *Excel* 100
5. CONCLUSION 140
REFERENCES 141

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

1. INTRODUCTION

1.1 Overview

Engineers are constantly challenged with solving a wide range of complex analytical and computational problems in their fields of specialty. These problems involve analyses methodologies and the management of data. The application of computers enables repetitive, time-consuming and often tedious calculations to be conducted rapidly, efficiently, and less prone to errors. The application of computer tools also enables the results and outputs of such engineering analyses to be readily transferred and incorporated into reports and other engineering documents. An even greater advantage, in terms of productivity and efficiency, is realized when these calculations and outputs are replicated across numerous projects. Competence in computer skills predisposes engineers to pursue and develop more creative and innovative solutions to problems than otherwise.

Microsoft Excel is widely and increasingly being used as a tool to assist engineers in conducting and replicating intricate calculations and analyses, designing complex systems, and managing large data sets. *Microsoft Excel* is an electronic spreadsheet program developed by the *Microsoft* Company, and part of the *Microsoft Office* suite of products. A spreadsheet is a grid that organizes data and calculations into columns and rows. The intersection of a column and a row is called a cell. An electronic spreadsheet enables users to store, organize, manipulate, and analyze data in the cells of the spreadsheet. As of this publication, the software is into the 2019 release.

This course presents fundamental principles and engineering applications of simulation and systems modeling, and demonstrates the *Microsoft Excel* tools, methods, and strategies that can be used to simulate and model real-life engineering systems. Simulation techniques involve conducting virtual experimentation of a system or process, on a computer, using mathematical and statistical models. This course presents how to formulate, implement and solve simulation and systems modeling problems in *Microsoft Excel*.

Upon completion of this course, participants will have gained insight into applying *Excel* tools, methods, and strategies in formulating, implementing and analyzing simulation models. Participants will also be able to identify professional situations where the application of these innovative *Excel* techniques will be of great benefit and advantage, and will enable practitioners to significantly improve their productivity and the quality of their work product.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

2. SIMULATION

2.1 Introduction

Simulation is a statistical analysis tool used in many fields such as the sciences, engineering, business, management and many others. Simulation techniques have been used to study a wide variety of problems and complex physical phenomena such as traffic congestion, the spread of public health epidemics, weather forecasting, the performance of financial markets, military operations, emergency response scenarios, the quality of product or duration of an industrial or manufacturing process, etc., etc. Simulation has been used to solve mathematical problems for which a direct solution is impractical.

If the processes being **simulated** involve an element of probabilistic behavior (random chance) the simulation is referred to as **Monte Carlo simulation**. Many complex and large-scale engineering problems are amenable to Monte Carlo simulation due to the fact that the simulation technique can handle large numbers of **random variables**, numerous statistical distributions, and nonlinear mathematical models.

Monte Carlo simulation is useful in situations where direct experimentation of a system or process is impractical, infeasible, time or cost prohibitive, or simply impossible. For example, it is not possible to conduct an experiment on the spread or impacts of a highly contagious disease outbreak on a large or densely populated U.S. city. Likewise, it would be impractical or time prohibitive for a light bulb manufacturer to test a large sample of a product over the product's entire design life. In Monte Carlo simulation, the experimentation is conducted rapidly and many times over on a computer using a simulation model. The incorporation of random variables in the simulation model to describe the discreet elements of the system or process enables many multiple scenarios of the system or process to be modelled, thus synthesizing model output data. The model output data set is then analyzed by appropriate statistical methods, to draw conclusions and to make decisions and recommendations about the system or process.

Thus, Monte Carlo simulation enables one to see all possible outcomes of a process and make decisions that take into account the probability (uncertainty or risk) associated with the discrete elements of the process.

2.2 Illustration of a Simulation Model

In this section a very simple example shall be used to illustrate a Monte Carlo simulation model.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Consider a small local contractor hired to prepare a site for a national company to come in and build a luxury apartment complex. The local contractor's work items include clearing the site and setting up trailers for site offices and other facilities. The contractor breaks down the project into the following discrete tasks in sequential order, and their estimated completion times.

	Work Item	Description	Completion Time (days)
1.	Clearing and Grubbing	Clear the site of vegetation, pull tree stumps, and remove all debris.	10
2.	Furnish and install trailers	Supply trailers and all ancillary work to complete installation.	5
3.	Connect services	Hook up power, utilities, and telecommunication services	7
4.	Furnish and install interior amenities	Install bathrooms, kitchenette equipment, office furniture, security devices	4
5.	Final cleanup	Final cleanup of interior and exterior	2
	Project completion		28

The completion times are based on the local contractor's recent experience with similar projects, clients and suppliers. In this approach, the local contractor considers these times to be set and fixed to yield the overall project completion time. This is called a **deterministic** approach. Prior to any knowledge or experience in Monte Carlo methods, the deterministic approach would be the intuitive approach to handle this type of problem.

The fact that the deterministic approach considers the work items' completion times to be set and fixed is a major weakness of the approach. Practically, all the work items' completion times are inherently **random** to varying degrees. For example, the completion of the clearing and grubbing is subject to randomness (or **uncertainty**) due to factors such as weather, equipment reliability, availability and punctuality of earth moving equipment and the operators, etc. etc. The work item to furnish and install the trailers is subject to randomness due to traffic conditions, availability or schedule of escort vehicles required to accompany the truck(s) that bring the trailer(s) to the site, as well as the randomness associated with obtaining the permits from governmental agencies to

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

convey oversized cargo on state roads. Once the trailer(s) reaches the site, further uncertainty may be due to soil conditions that may slow down (or speed up for that matter) the process of anchoring and securing the trailer(s). There is also randomness associated with the quality of the workmanship, or lack thereof, that may necessitate rework or additional inspections prior to approval. The uncertainty associated with each work item contributes to the uncertainty associated with the overall project completion time. The degree of uncertainty is called **risk**, and exposes the local contractor to potential unfavorable outcomes, particularly financial loss.

It is therefore justified to take a **probabilistic** approach to the analysis in order to incorporate the uncertainty associated with each work item's completion time. This can be done by describing each work item's completion time with an appropriate **random variable** (also called a **probability distribution**). The work items' completion times can be given at certain probabilities based on the appropriate probability distribution for that work item. Overall project completion times can then be determined at given probabilities. The local contractor can now select an acceptable level of uncertainty (or risk) and reach a more informed conclusion regarding the project completion time.

Using assumed probability distributions, the work items' completion times for this project can be described probabilistically as follows

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

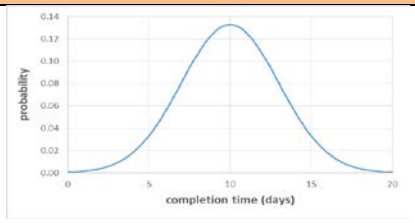
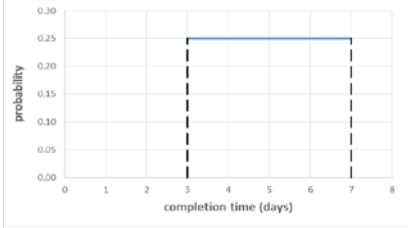
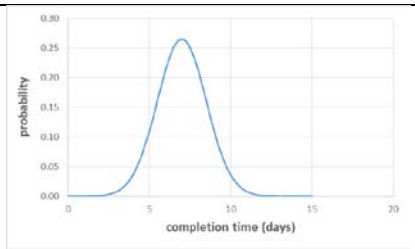
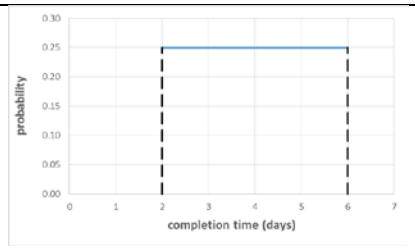
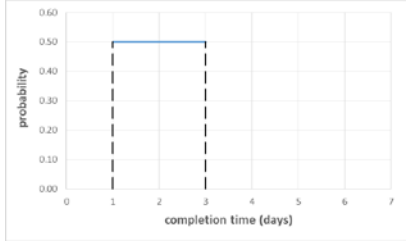
	Work Item	Probability Distribution	Completion Time (days)
1.	Clearing and Grubbing	Normal distribution Parameters: Mean (μ) = 10 Standard deviation (σ) = 3	
2.	Furnish and install trailers	Uniform distribution Parameters: α = 3 β = 7	
3.	Connect services	Normal distribution Parameters: Mean (μ) = 7 Standard deviation (σ) = 1.5	
4.	Furnish and install interior amenities	Uniform distribution Parameters: α = 2 β = 6	
5.	Final cleanup	Uniform distribution Parameters: α = 1 β = 3	

Figure 1. 1: Framework for simulation

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

To generate a possible scenario for a particular work item's completion time, the contractor can randomly pick a probability value and read off the completion time from the probability distribution graph (or calculate it from the graph function). The process is repeated for each work item, and the project completion time for the scenario is the aggregate of the individual completion times obtained. This process can now be replicated several times over to synthesize a data set of project completion times. The project completion time data can be analyzed to obtain the descriptive statistics, quartiles and percentiles, or other statistical measures which are used for appropriate decision making.

At this stage it can be seen that in order to gain expertize in Monte Carlo simulation, one must have a grasp (or refresher) on random variables (probability distributions). The following chapters will present the fundamentals of random variables, the types of random variables and how they are applied. A subsequent question is how does one select an appropriate random variable for a work item, or how does one know that a given random variable appropriately describes the random behavior of a specific activity?

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

3. RANDOM VARIABLES

3.1 Definition

A random variable can be defined as a variable that has a single numerical value, determined by chance, for each outcome of the procedure.

To elaborate on this definition, consider the number of vehicles that queue up behind a traffic light on an approach of a city intersection each time the light goes red. By observation, it will be seen that the number of vehicles that queue each time can take on values (numerical outcomes) such as 3, 1, 6, 0, 10, 2, etc. etc. For each occurrence of the red light, the numerical outcomes of the queue are random. If the observer studies this system long enough, the observer shall realize that each numerical outcome is associated with some probability of occurrence. Thus, the number of vehicles that queue up behind the red light each time it occurs is a random variable.

Let X denote the random variable, i.e. the number of vehicles in queue up at the red light.
Let x_i denote a specific numerical outcome of the random variable, i.e. 3, 1, 6, 0, 10, 2, etc. etc.
Let p_i denote the probability of occurrence of numerical outcome x_i.
Using the notation, the probability of the random variable X realizing a numerical outcome x_i can be written as

$$\boldsymbol{P(X = x_i) = p_i} \quad \text{..... Equation (3. 1)}$$

A tabular or graphical description of the possible numerical outcomes (x_i) and their corresponding probability values (p_i) is called the **probability distribution** of the random variable. If the probability values (p_i) can be computed from some known mathematical function ($\boldsymbol{f(x_i)}$), this function is known as the **probability distribution function**.

That is,

$$\boldsymbol{f(x_i) = P(X = x_i)} \quad \text{..... Equation (3. 2)}$$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

For example, the probability distribution could be,

Number of Vehicles in queue (x_i)	$f(x_i) = P(X = x_i)$
0	0.02
1	0.10
2	0.15
3	0.25
4	0.35
5	0.10
6	0.03

A random variable may be discrete or continuous. A **discrete random variable** can take on a finite number of numerical outcomes or a countable number of numerical outcomes. For example, the number of vehicles that queue up behind a red light will always be some limited range of values, typically positive integer values and zero.

On the other hand, a **continuous random variable** has infinitely many possible numerical outcomes and those values can be associated with measurements on a continuous scale without gaps. For example, if one was interested in the random variable,

Y = waiting time (in minutes) of a vehicle at the red light,

this would be a continuous random variable. Possible numerical outcomes could include 1 minute, 1.5 minutes, 1.55 minutes, 1.555 minutes and so on and so forth and there are infinitely many possible values, and the values can be associated with measurements on a continuum with no gaps.

3.2 Discrete Random Variables

The properties of a discrete random variable as follows,

1. If $f(x_i) = P(X = x_i)$, then for all x_i values,

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

$$\mathbf{0 \leq f(x_i) \leq 1} \quad \text{. . . . . Equation (3. 3)}$$

and

$$\sum_{\mathbf{all}\ x_i} \mathbf{f(x_i) = 1}$$

. . . . . **Equation (3. 4)**

The distribution function $f(x_i)$ for a discrete random variable is called the **probability mass function**.

2. The **cumulative probability** of an x value of interest (x_k) is the sum of the probabilities of the x_i values up to and including the x_k value. This is mathematically expressed as,

$$P(X \leq x_k) = \sum_{i=1}^{k} f(x_i) = F(x)$$

. . . . . **Equation (3. 5)**

$F(x)$ is called the **cumulative distribution function**.

3. The **expectation** (or **expected value** or **mean**) (μ) of the random variable X is defined as

$$\mu = \sum_{i=1}^{n} x_i . f(x_i)$$

. . . . . **Equation (3. 6)**

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4. The variance (σ^2) of a discrete random variable is defined as

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 . f(x_i)$$

. . . . . **Equation (3. 7)**

5. The standard deviation (σ) of the discrete random variable is defined as

$$\sigma = \sqrt{\sum_{i=1}^{n} (x_i - \mu)^2 . f(x_i)}$$

. . . . . **Equation (3. 8)**

The following tabulated calculations demonstrate the properties of a discrete random variable

x_i	$f(x_i)$	$F(x)$	$x_i . f(x_i)$	$(x_i - \mu)^2 . f(x_i)$
0	0.02	0.02	0.00	0.209
1	0.10	0.12	0.10	0.497
2	0.15	0.27	0.30	0.227
3	0.25	0.52	0.75	0.013
4	0.35	0.87	1.40	0.208
5	0.10	0.97	0.50	0.313
6	0.03	1.00	0.18	0.230
$\sum$	1.00		3.23	1.697

The expectation μ = 3.23
The variance σ^2 = 1.697
The standard deviation $\sigma = \sqrt{1.697}$ = 1.302

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

3.3 The Binomial Distribution

The binomial distribution is a discrete random variable that meets the following conditions:

1. The procedure has a fixed number of trials.
2. The outcomes of the trials are independent.
3. There are only two possible outcomes of a trial - pass versus fail, good versus bad, success versus failure, accept versus reject, etc., referred to as **Bernoulli trials**.
4. The probability of success (or failure) is the same for all trials.

Examples of the Binomial distribution include:

- The number of heads obtained by tossing a coin a certain number of times
- the number of manufactured parts from a sample that fail a quality control inspection
- the number of drivers from a sample, that obey a “reduce speed” sign on a rural highway
- the number of candidates from a sample that pass a standardized test

Consider a binomial distribution (X).
Let S and F denote the outcomes for success and failure respectively,
If the probability of success, $P(S) = p$, then
The probability of failure $P(F) = q = 1 - p$
Let n denote the fixed number of trials of the experiment
Let x denote a specific number of successes out of the n trials
Thus, $P(X = x)$ is the probability of getting x successes from the n trials

It can be shown that the distribution function ($f(x)$) of the Binomial distribution is given by

$$f(x) = P(X = x) = \frac{n!}{(n-x)!} p^x q^{n-x} \quad \text{. . . . Equation (3. 9)}$$

From Equation (3.6), it can be shown that the expected value (μ) of the Binomial distribution is given by,

$$\mu = np \quad \text{. . . . . Equation (3. 10)}$$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

From Equation 3.7, it can be shown that the variance (σ^2) of the Binomial distribution is given by,

$$\sigma^2 = npq \quad \text{..... Equation (3. 11)}$$

The standard deviation (σ) of the Binomial distribution is given by,

$$\sigma = \sqrt{npq} \quad \text{..... Equation (3. 12)}$$

Example:
A traffic engineer is studying the effectiveness of a yield sign placed on an exit ramp. A sample size of 10 vehicles was used. The probability that a driver will obey a yield sign is $2/3$.
a) What is the probability that exactly 3 vehicles will obey the sign?
b) What is the probability that less than 3 vehicles will obey the sign?
c) What is the probability that 2 or more vehicles will obey the sign?
d) What is the expected number of vehicles that will obey the sign?

Solution:
1. There is a fixed number of trials. $n = 10$
2. If each driver's decision to obey or not obey the sign is not influenced by other drivers, the outcomes of the trials are independent.
3. There are only two possible outcomes of a trial: Obey or Not Obey the sign.
4. The probability of Obey is the same for all trials, with a value of $p = 2/3$.

This experiment meets the conditions of a Binomial distribution.

a) From Equation (3.9),

$$f(x) = P(X = x) = \frac{n!}{(n-x)!}p^x q^{n-x}$$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

$$P(X=3)=\frac{10!}{(10-3)!}\left(\frac{2}{3}\right)^{3}\left(\frac{1}{3}\right)^{10-3}=0.016\ \ or\ 1.6\%$$

b) From Equation (3.5), the cumulative probability

$$P(X\le x_k)=F(x)=\sum_{i=1}^{k}f(x_i)$$

$$P(less\ than\ 3)=\ P(X<3)=P(X\le 2)=F(2)$$

or

$$P(X<3)=P(X=0)+P(X=1)+P(X=2)$$

$$P(X<3)=\frac{10!}{(10-0)!}\left(\frac{2}{3}\right)^{0}\left(\frac{1}{3}\right)^{10-0}+\frac{10!}{(10-1)!}\left(\frac{2}{3}\right)^{1}\left(\frac{1}{3}\right)^{10-1}+\frac{10!}{(10-2)!}\left(\frac{2}{3}\right)^{2}\left(\frac{1}{3}\right)^{10-2}$$

$$P(X<3)=0.000017+0.00034+0.003=0.00357\ or\ 0.36\%$$

c) $P(2\ or\ more)=\ P(X\ge 2)=P(X=2)+P(X=3)+\cdots+P(X=10)$

or we can take advantage of the **complementation** rule of probability

$$P(2\ or\ more)=\ P(X\ge 2)=1-P(X<2)=1-[P(X=0)+P(X=1)]$$

$$P(2\ or\ more)=\ 1-[P(X=0)+P(X=1)]$$

$$P(2\ or\ more)=1-[\ 0.000017+0.00034]=0.999643\ or\ 99.96\%$$

d) From Equation (3.10),

$$\mu=np=10*2/3=6.667$$

This is the expected number of drivers from the sample that will obey the sign.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

3.3 The Poisson Distribution

The Poisson distribution is a special limiting case of the Binomial distribution where $\boldsymbol{n}$ is large and p is small. In other words, the Poisson distribution is more appropriate for computing probabilities associated with rare events. In general if $\boldsymbol{n} > 50$, and $p < 0.1$, the Poisson is appropriate.

It can be shown that the distribution function ($f(x)$) of the Poisson distribution is given by

$$f(x) = P(X = x) = \frac{\lambda^x}{\mathrm{x!}} e^{-\lambda} \quad \text{. . . . . Equation (3. 13)}$$

where λ is the expected value expressed per unit of time, and $\lambda > 0$.

From Equation (3.6), it can be shown that the expected value (μ) of the Poisson distribution is given by,

$$\mu = \lambda = np \quad \text{. . . . . Equation (3. 14)}$$

From Equation (3.7), it can be shown that the variance (σ^2) of the Poisson distribution is given by,

$$\sigma^2 = \lambda \quad \text{. . . . . Equation (3. 15)}$$

In other words, the expected value and the variance are equal. This is a unique property of the Poisson distribution, and can be used to suggest that a given data set exhibits the behavior of a Poisson distribution. The Poisson distribution is widely used in industrial engineering and also in traffic engineering (a branch of civil engineering). In traffic engineering, the Poisson distribution is used to model vehicle arrivals, as in traffic flows, or arrivals at an intersection or a toll booth. Other applications in traffic engineering include parking analysis and modeling the frequency of crashes.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Example:
The average number of vehicles randomly arriving at a toll both follows a Poisson distribution with an average of 6 vehicles arriving per minute.
a) What is the probability of exactly 3 vehicles arriving in a minute?
b) What is the probability of 3 or more vehicles arriving in a minute?
c) What is the probability of 3 or more vehicles arriving in 30 seconds?
d) What is the 95th percentile number of vehicles arriving in 30 seconds?

Solution:

a) From Equation (3.13),

$$f(x) = P(X = x) = \frac{\lambda^x}{\text{x!}} e^{-\lambda}$$

$$\lambda = 6\ veh/min$$

$$P(X = 3) = \frac{6^3}{3!} e^{-6}) = 0.0924\ or\ 9.24\%$$

It is pertinent to note that when using the Poisson model, one does not have to know the sample size in order to proceed with the calculations.

b) $P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) + \cdots$

This cumulative probability calculation will essentially go on forever. So we must use the complementation rule of probability.

$$P(X \geq 3) = 1 - P(X < 3)$$

$$P(X \geq 3) = 1 - P(X \leq 2)$$

$$P(X \geq 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$P(X \geq 3) = 1 - \left[\frac{6^0}{0!} e^{-6} + \frac{6^1}{1!} e^{-6} + \frac{6^2}{2!} e^{-6}\right]$$

$$P(X \geq 3) = 1 - [0.00248 + 0.0149 + 0.0446] = 0.938\ or\ 93.8\%$$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

c) The original average vehicle arrival rate (λ) is given per minute. This problem seeks to calculate a probability based on a different time interval of 30 seconds. Therefore the average value (λ) must be converted to a new average value (λ_{new}) that reflects the 30 second time interval. The conversion is done by simple proportion.

$$\lambda_{new} = \frac{30}{60}\lambda = \frac{30}{60} * 6 = 3\ vehicles\ per\ 30\ seconds$$

This conversion is referred to as the **Poisson process**.

Thus,

$$P(X \geq 3_{in\ 30s}) = 1 - P(X < 3_{in\ 30s})$$

$$P(X \geq 3_{in\ 30s}) = 1 - [P(X = 0_{in\ 30s}) + P(X = 1_{in\ 30s}) + P(X = 2_{in\ 30s})]$$

$$P(X \geq 3_{in\ 30s}) = 1 - \left[\frac{3^0}{0!}e^{-3} + \frac{3^1}{1!}e^{-3} + \frac{3^2}{2!}e^{-3}\right] = 0.762$$

$$P(X \geq 3_{in\ 30s}) = 1 - [0.0498 + 0.149 + 0.224] = 0.423\ or\ 42.3\%$$

d) The 95th percentile is a value (x) that has 95% of the distribution less than or equal to it. In terms of (cumulative) probabilities, it is the value (x) such that,

$$P(X \leq x) = 0.95$$

To "solve" for x, a trial and error approach can be employed.

$$P(X \leq 4) = \left[\frac{3^0}{0!}e^{-3} + \frac{3^1}{1!}e^{-3} + \frac{3^2}{2!}e^{-3} + \frac{4^3}{3!}e^{-4}\right] = 0.647$$

$$P(X \leq 5) = 0.916$$

$$P(X \leq 6) = 0.966$$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The 95th Percentile is approximately $x = 6$. (Interpolation of the values may be conducted if warranted by the level of accuracy desired). One is therefore 95% confident that the number of vehicle arrivals in a 30 second period will not exceed 6. Computing a percentile is effectively calling the inverse of a cumulative probability function to compute a value x for a given cumulative probability.

3.4 Other Discrete Distributions

For the sake of brevity, this discussion on discrete random variables is limited to the Binomial distribution and the Poisson distribution. There are several other discrete random variables presented in the engineering statistics literature. The Binomial distribution assumes a specimen is pulled for a Bernoulli trial and "put back in the box". This is referred to as **sampling with replacement**. The **Hypergeometric** distribution differs from the Binomial distribution in that it is based on **sampling without replacement**. The **Geometric distribution** differs from the Binomial distribution in that the number of independent trials (n) is not fixed. The **Negative Binomial** distribution describes the number of Bernoulli trials to obtain a specified number of successes. The **Multinomial** distribution is a generalization of the Binomial distribution whereby a trial can have more than two possible outcomes. Students are encouraged to review the details of other discrete random variables on their own.

The following is a non-exhaustive list of other discrete distributions that are in use in various fields of specialty.

- The Beta-Binomial distribution
- The Beta Geometric distribution
- The Beta Negative Binomial distribution
- The Borel-Tanner distribution
- The Consul (Generalized Geometric) distribution
- The Discrete Uniform distribution
- The Discrete Weibull distribution
- The Geeta distribution
- The Generalized Logarithmic Series distribution
- The Generalized Lost Games distribution
- The Generalized Negative Binomials distribution
- The Hermite distribution
- The Katz distribution

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

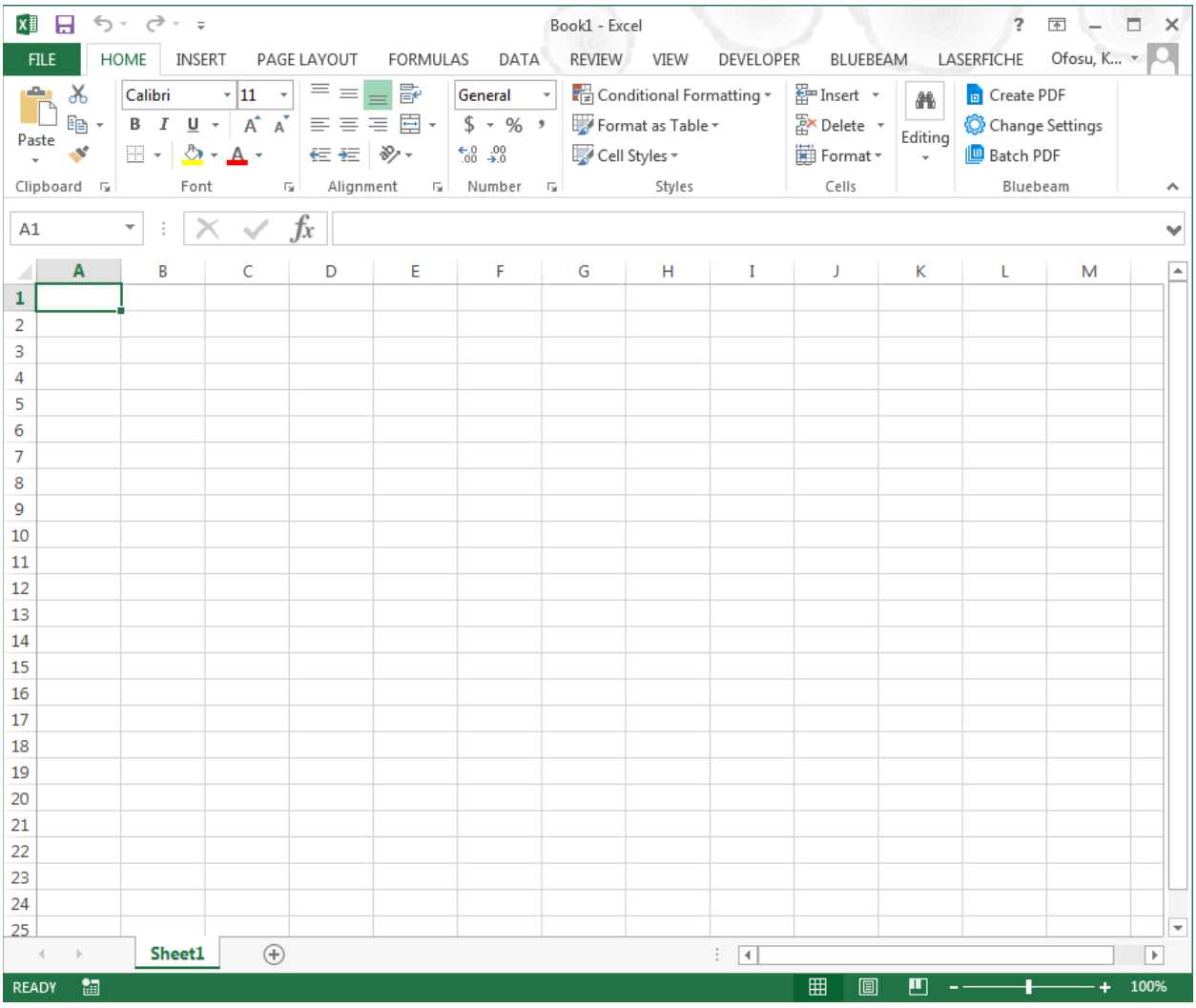
- The Lagrange-Poisson distribution
- The Leads in coin tossing distribution
- The Logarithmic Series distribution
- The Lost Games distribution
- The Matching distribution
- The Polya-Aeppli distribution
- The Quasi Binomial Type I distribution
- The Truncated Gene Negative Binomial distribution
- The Waring distribution
- The Yule distribution
- The Zeta distribution
- The Zipf distribution

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

3.5 Discrete Random Variables in *Excel*

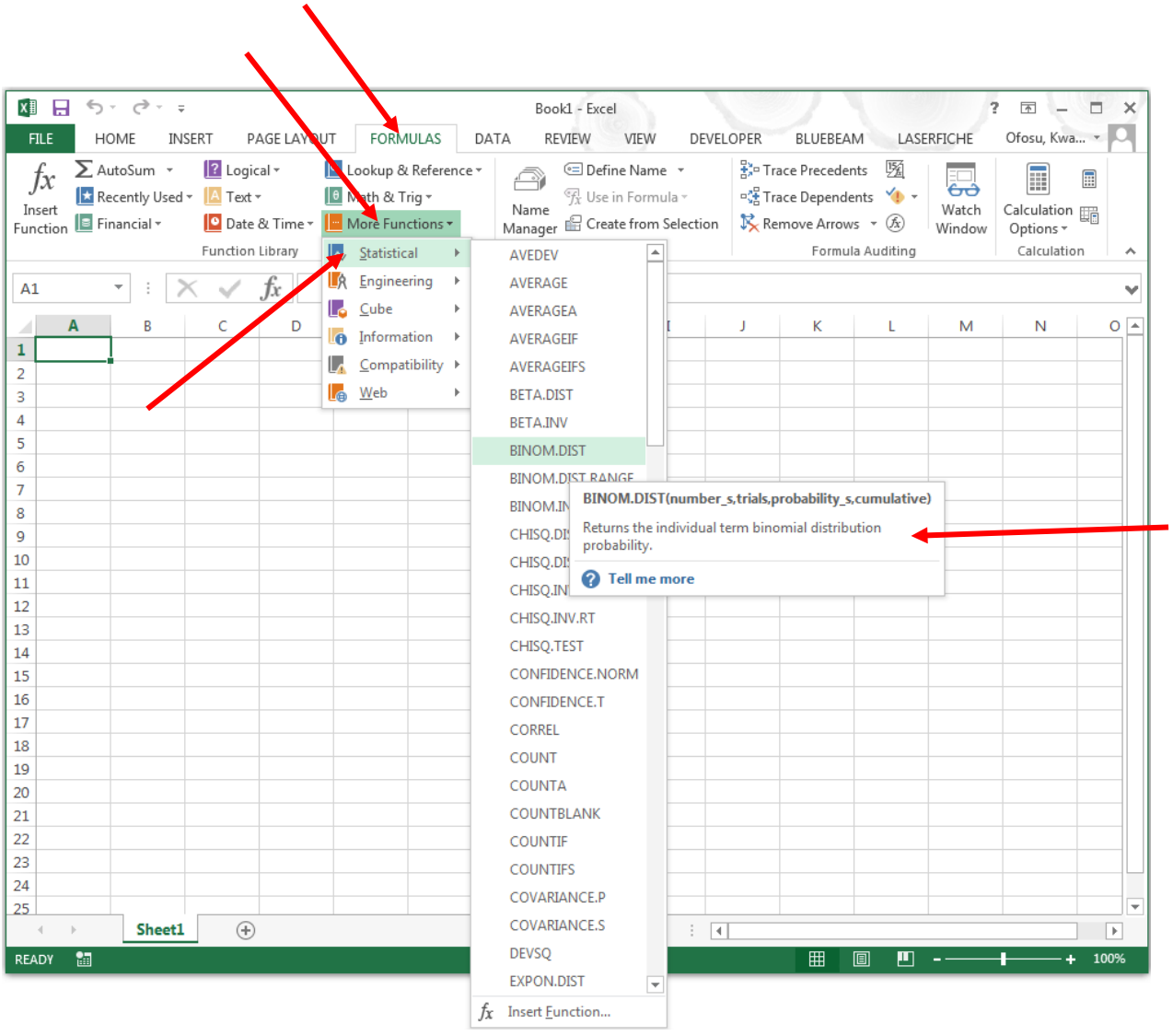
In this section the discrete random variables worked in Section 3.3 shall be reworked using *Microsoft Excel*.

Open a new session of *Excel*.



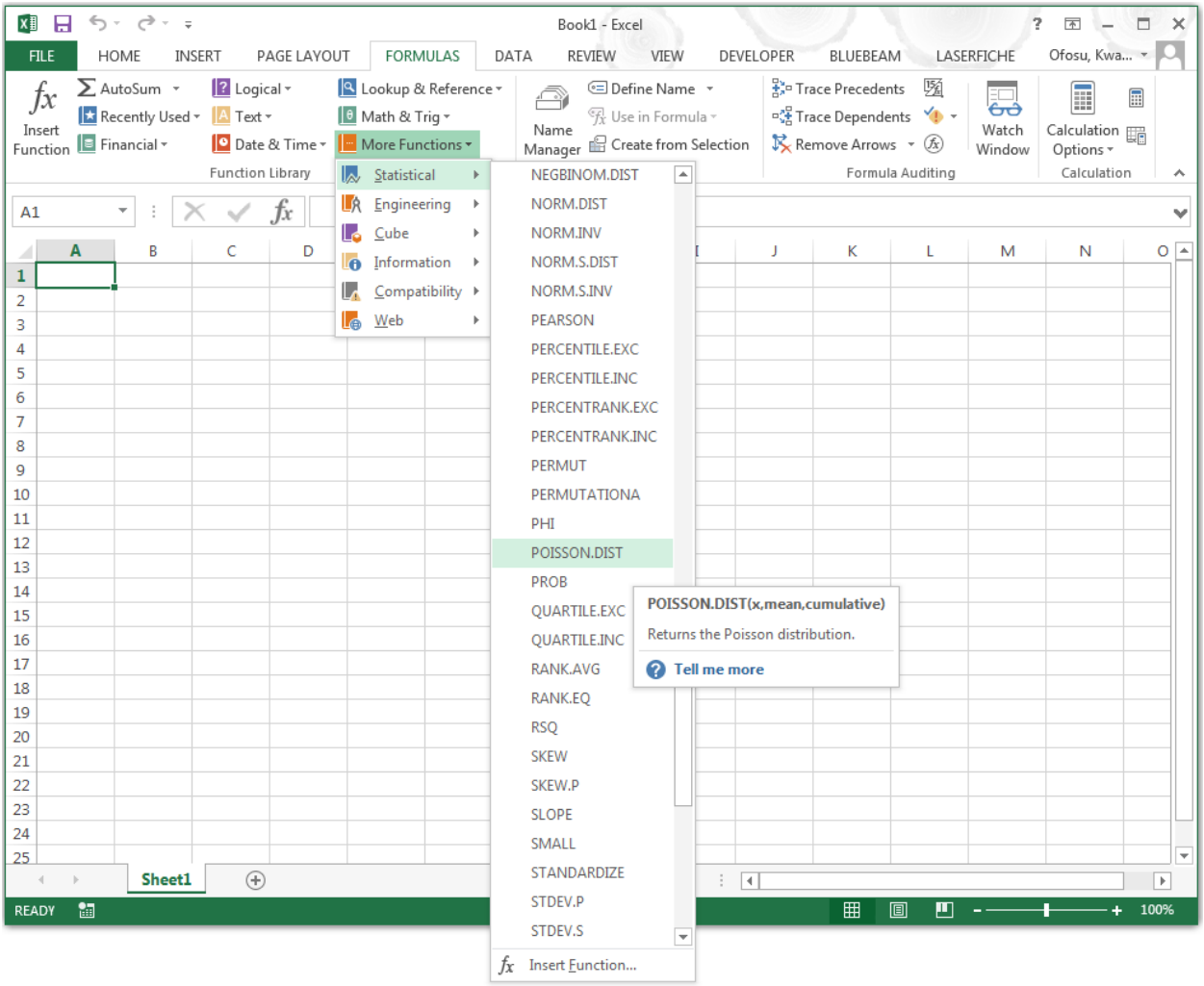
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Click on FORMULAS.
Click on **More Functions**.
Hover over **Statistical**.
Hover over a function, a tool tip appears.
Review the information on the tool tip which describes the function and how to implement it.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

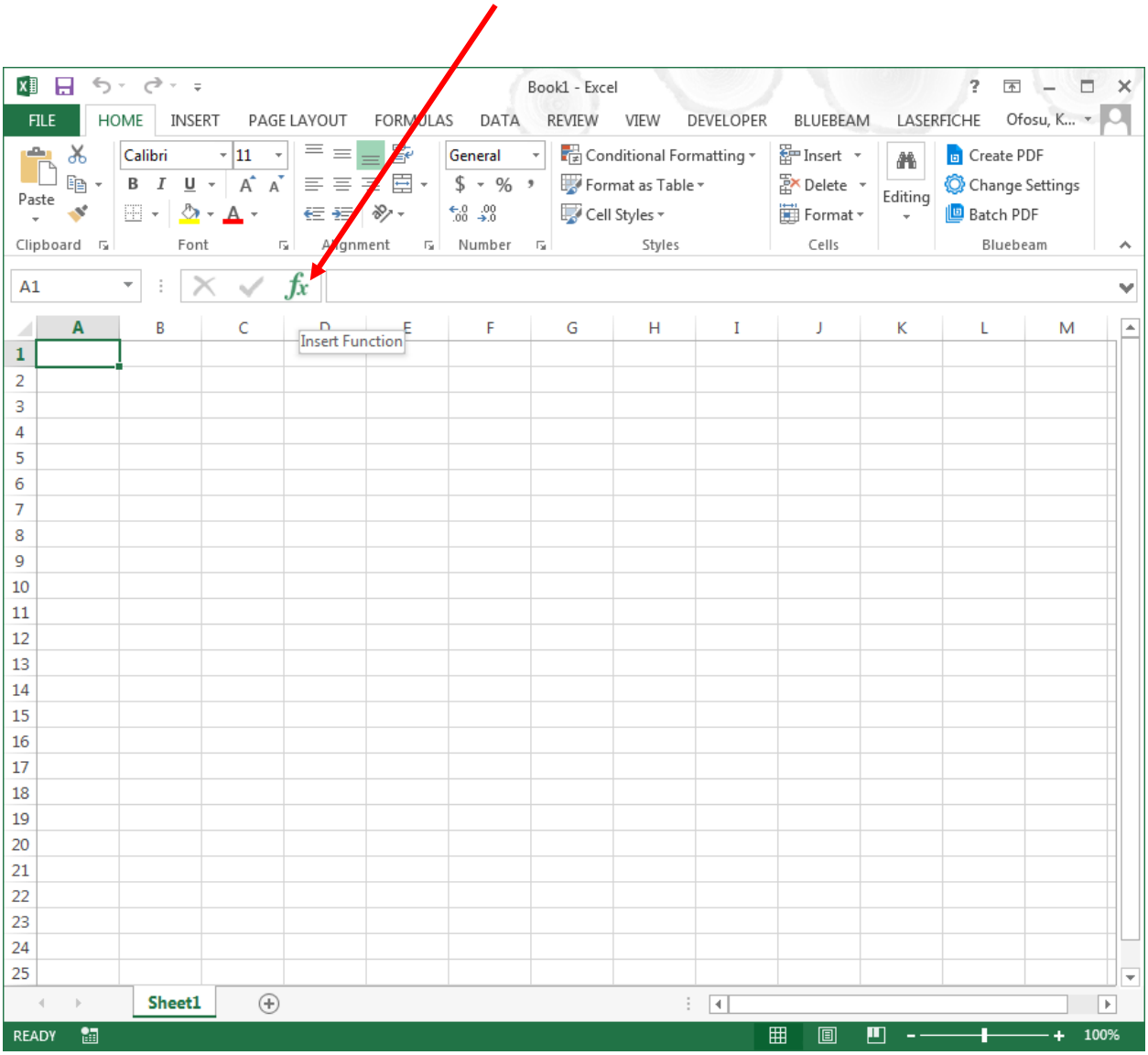
Drag the scrollbar and scroll down the list to review the statistical functions available in *Excel*.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

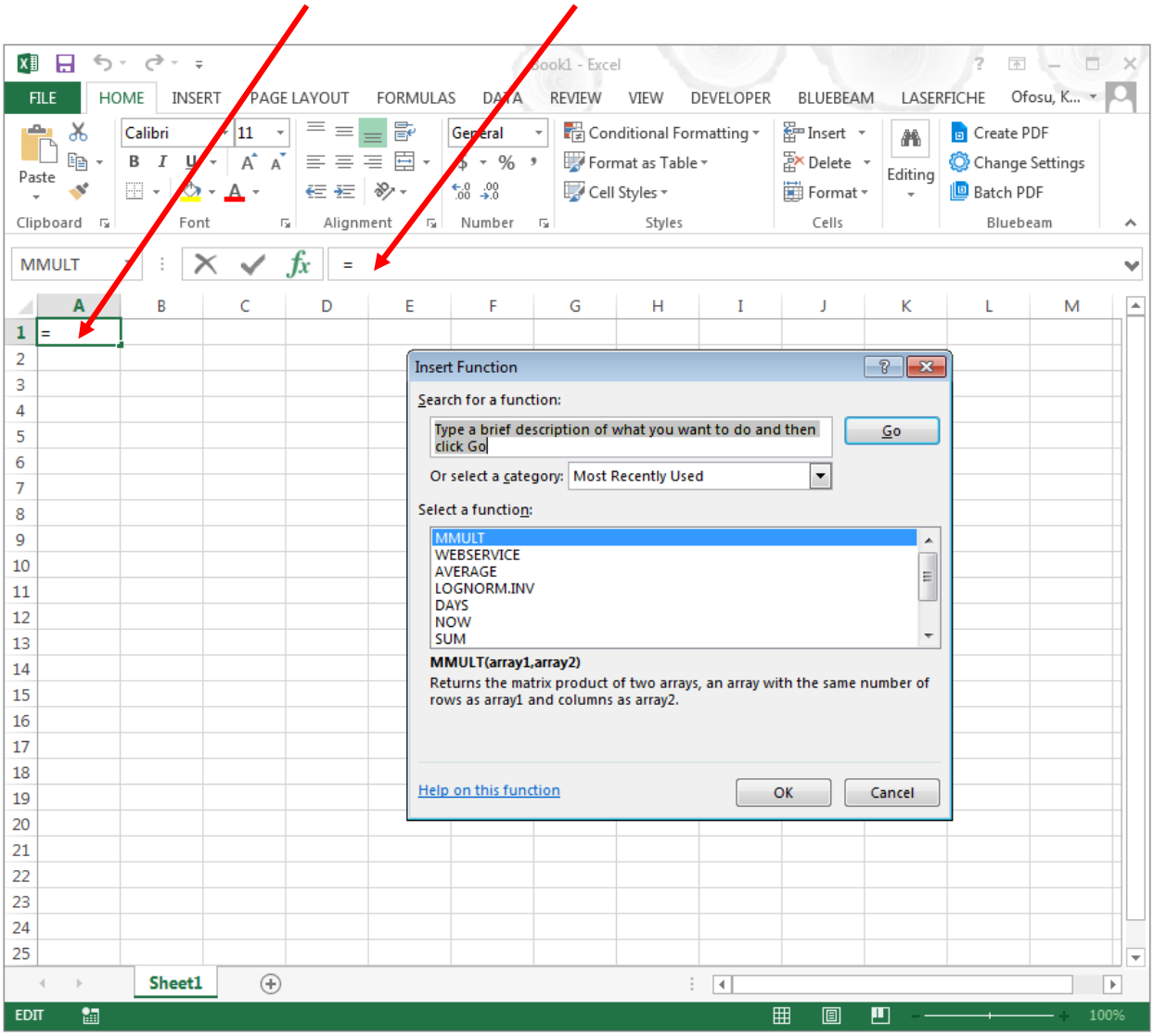
A second way of accessing the statistical functions in *Excel* is via the **Insert Function** icon.

Click on **Insert Function**.



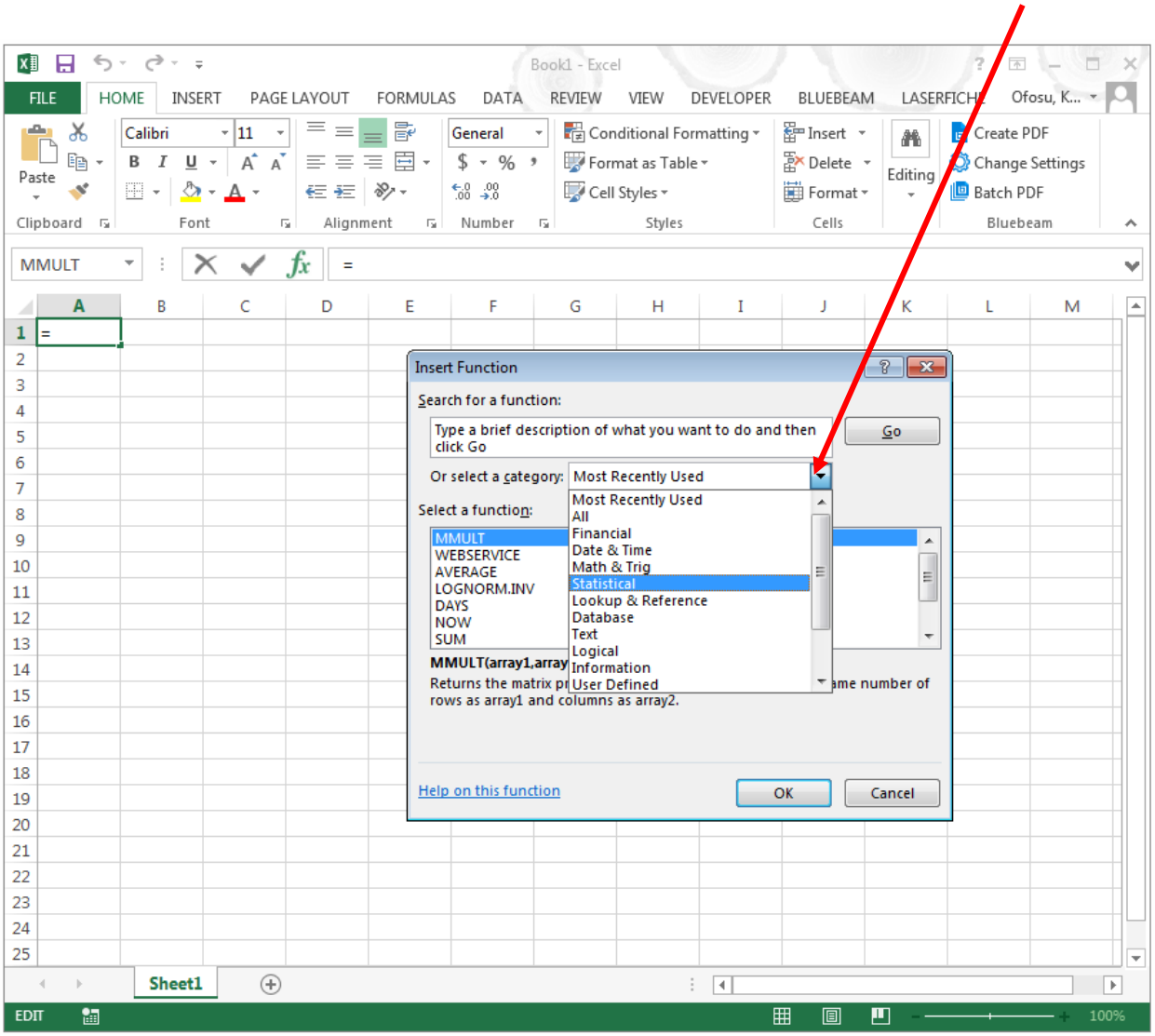
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The **Insert Function** window opens.
Note the equal (“ = ”) sign that appears in the **Formula Bar** and also in the current spreadsheet cell when the **Insert Function** icon is clicked on.



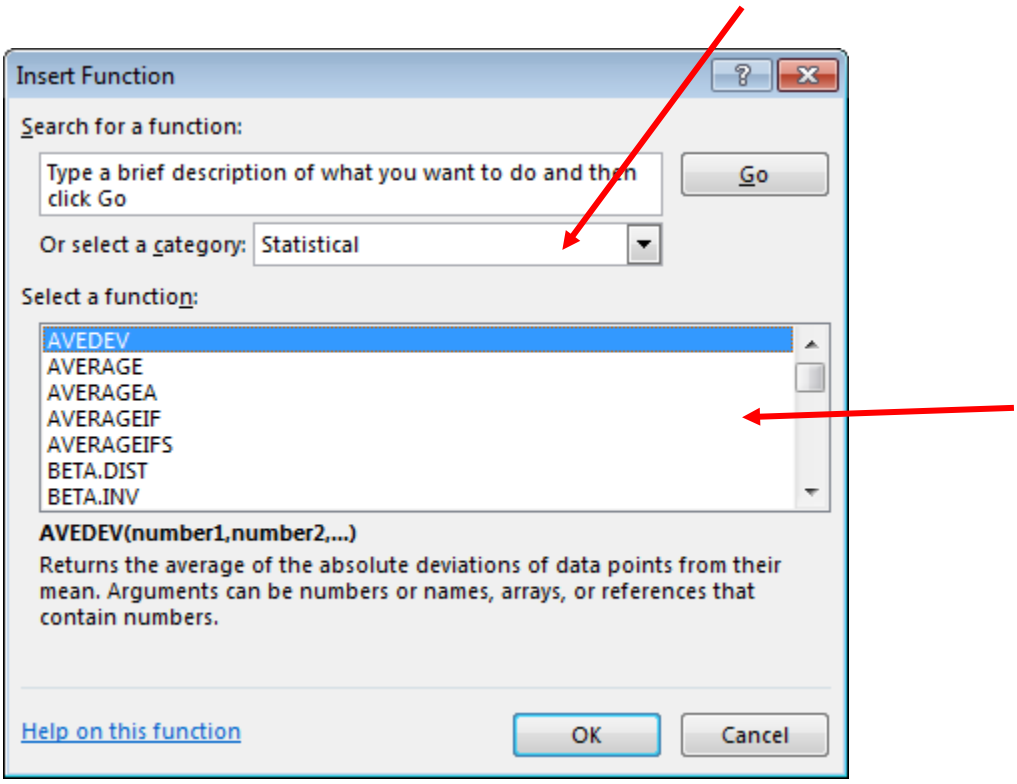
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Under **Or select a category**, click on the drop-down to select a category.



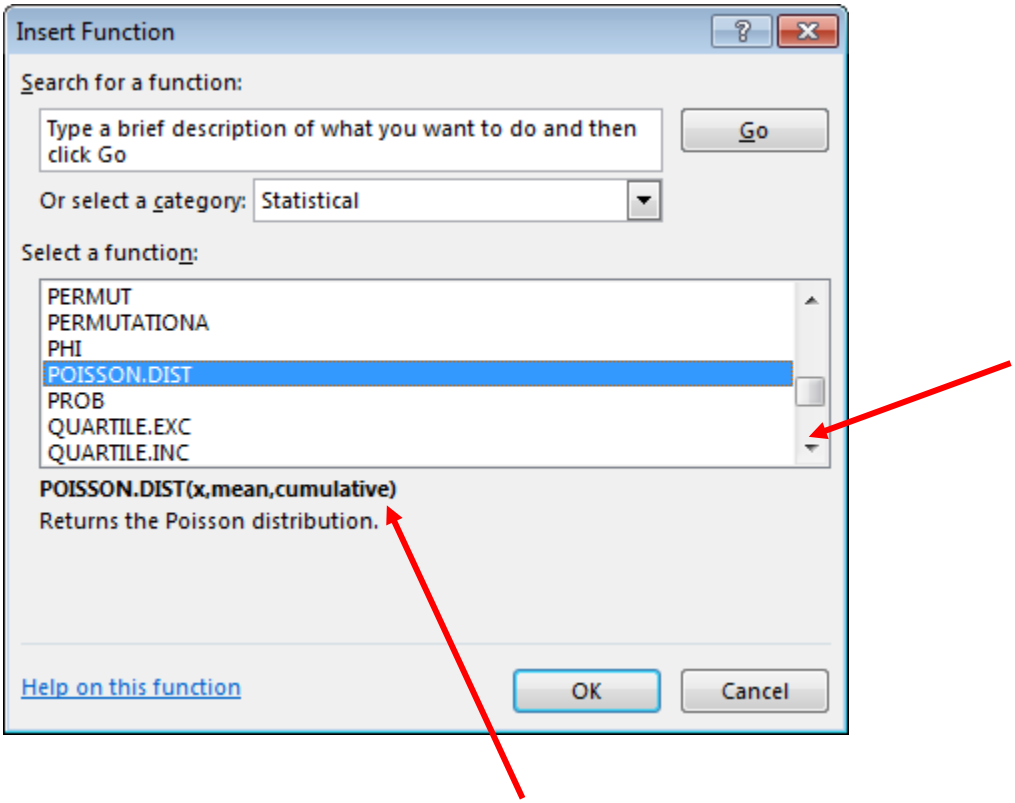
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

In the category list, select **Statistical**.
The statistical functions only, are now displayed in alphabetic order.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

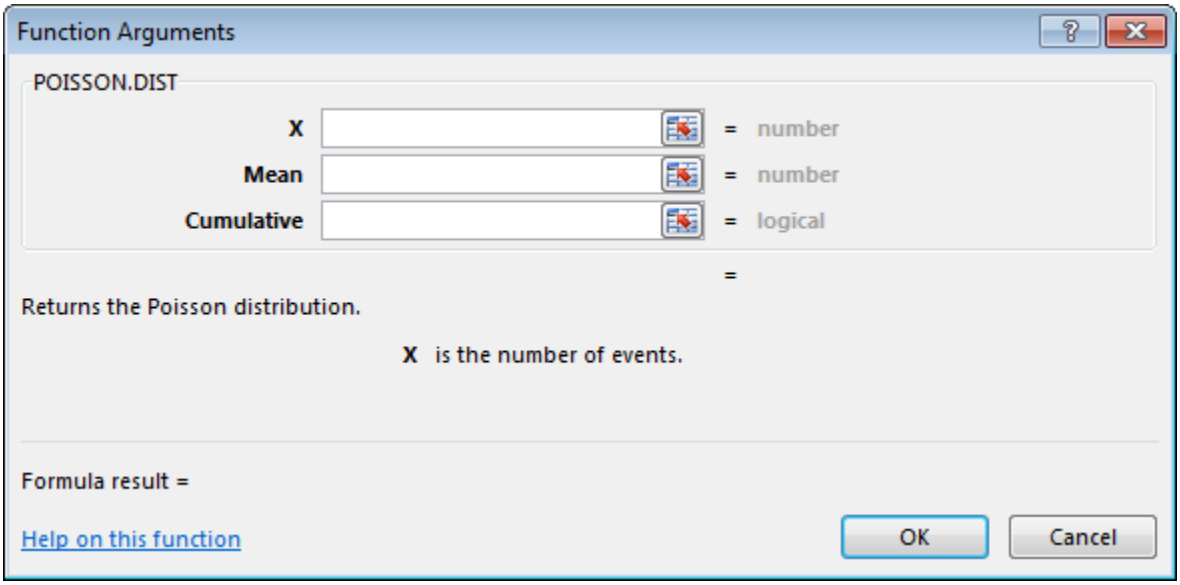
Under **Select a function**, scroll through the list of statistical functions.
Select POISSON.DIST
Note the description of the selected function that guides the user on what input is needed, and what result is **returned**.



Click on **OK**.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The **Function Arguments** window opens.
This is where the user supplies data to complete a calculation and return a result to the current spreadsheet cell.

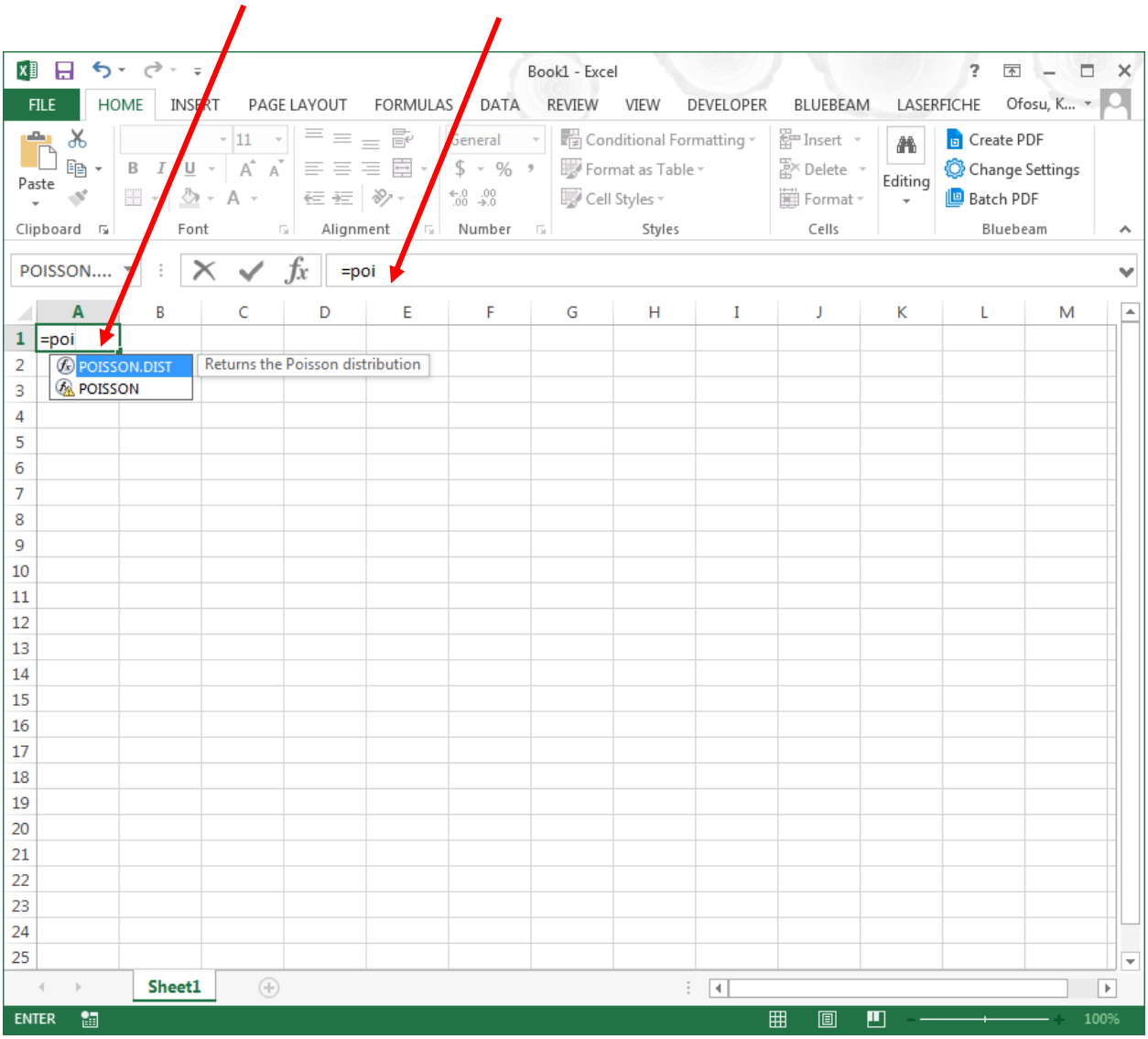


For now, hit **Cancel** to close out of the **Function Arguments** window.
Hit **Cancel** to close out of the **Insert Function** window, and return to the spreadsheet.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

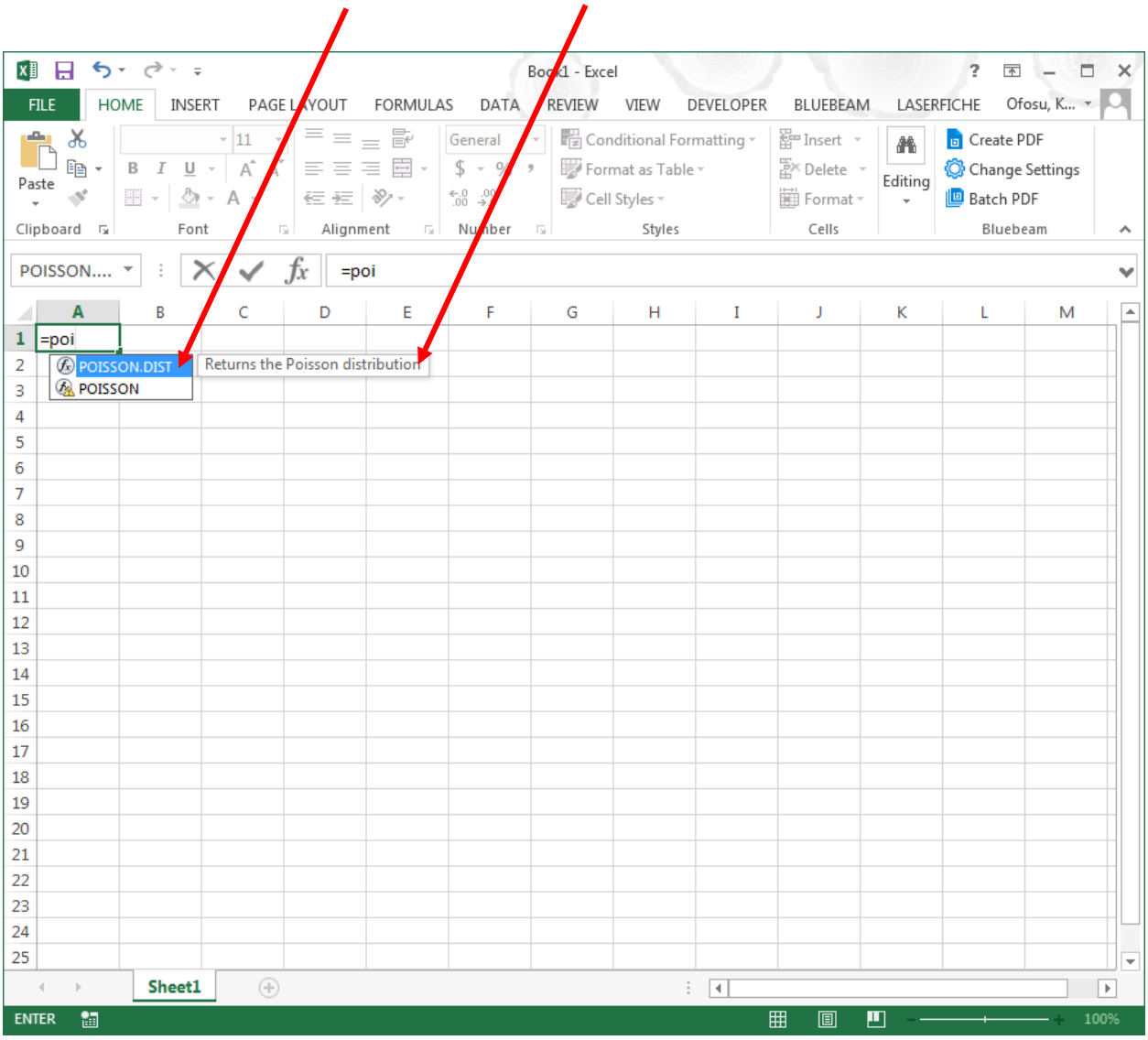
A third way of accessing any function is by typing directly into a cell the equal (“ = ”) sign followed by the function name as listed in the **Insert Function** window.

Pick a cell on the spreadsheet.
In the cell, type “ = ”, followed by “POISSON” (upper case or lower case does not matter), to look up the Poisson function.
As one begins typing in the cell, the entry also appears in the **Formula Bar**.



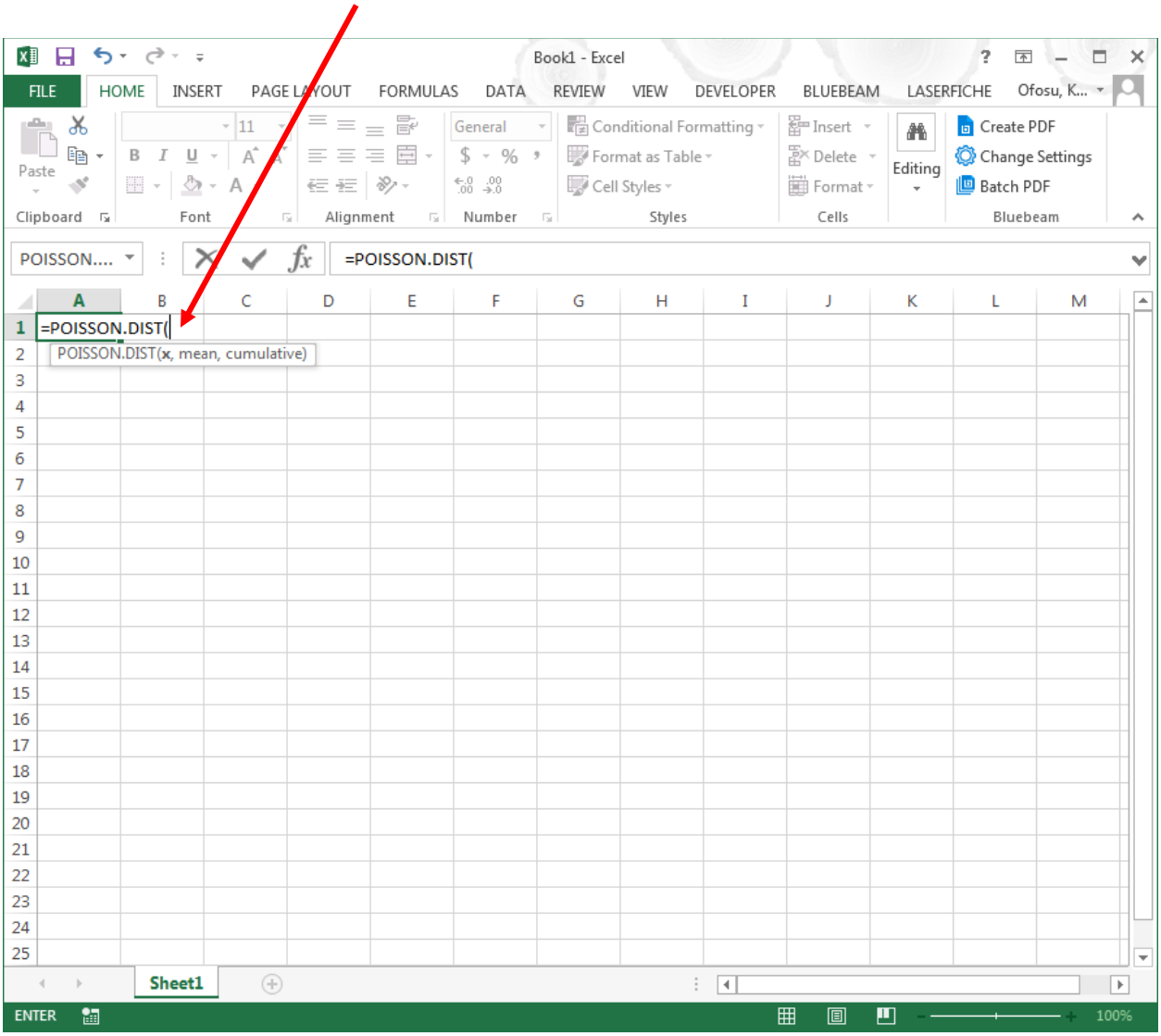
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Notice that as one starts typing, a shortlist of candidate functions appears from the cell.
Select a candidate function.
Review the tooltip that appears and gives a brief description of that function.



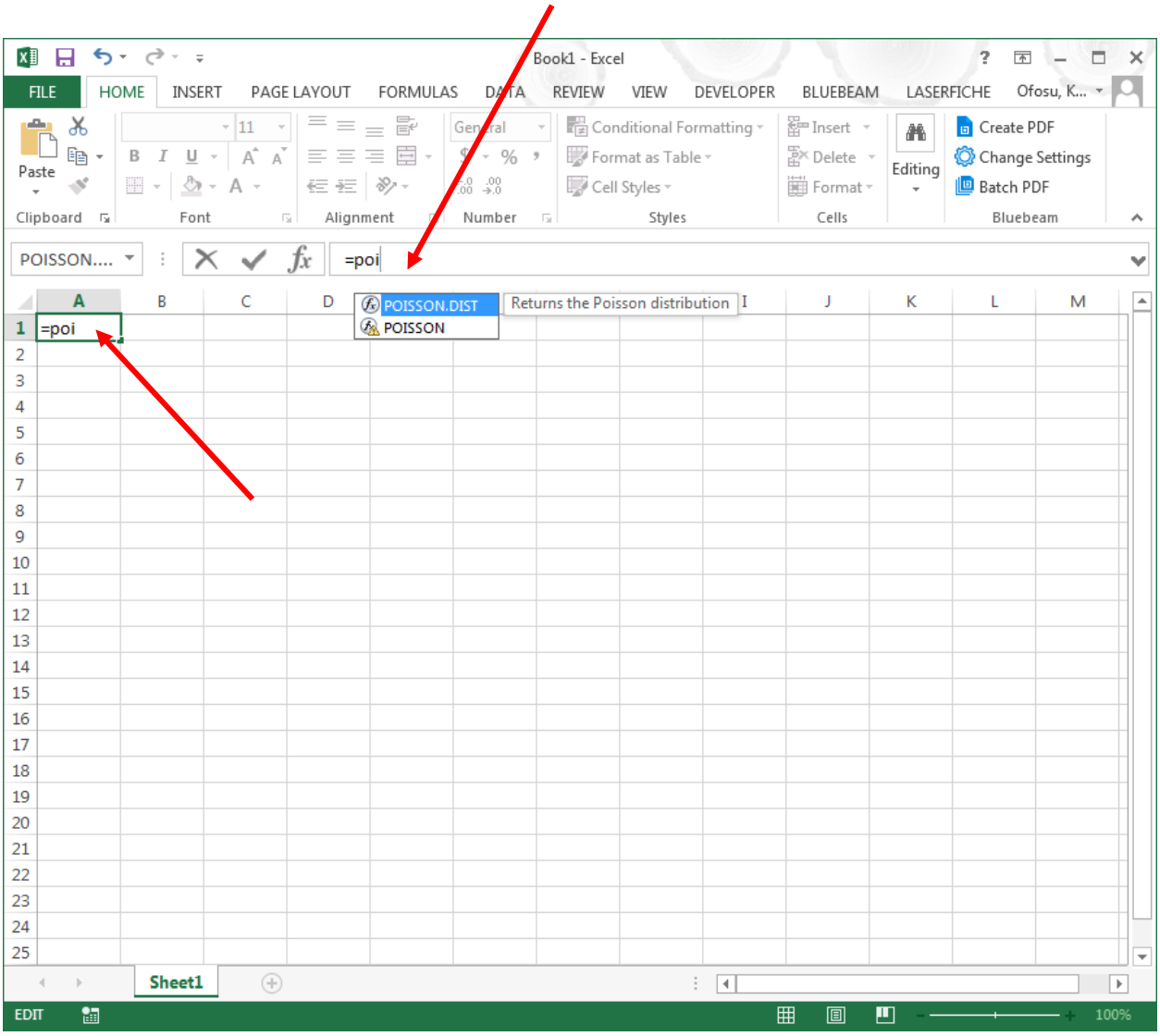
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

To call a function to conduct a calculation in the cell, double click on the function name in the candidate list.
In this demonstration, double click on POISSON.DIST.
The function name followed by open parenthesis ("(") prompt the user to supply the data needed conduct the calculation.
The tooltip guides the user to enter the required data in the required order, in order to return a result to the current spreadsheet cell.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

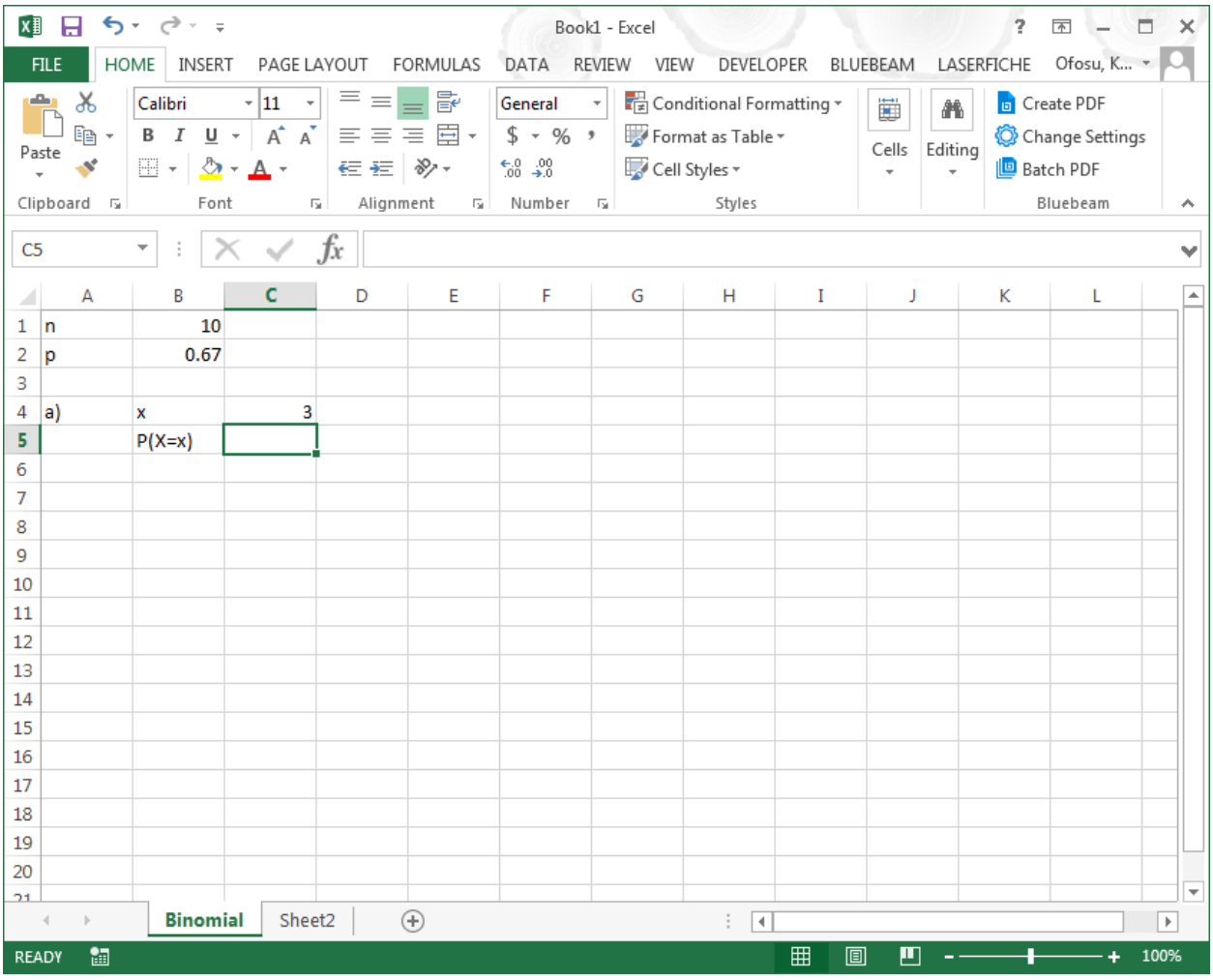
For now, hit the **Esc** button on the keyboard to abort the process.
A full blown calculation process will conducted later in this tutorial.
Notice that in the above method, the typing was conducted in the cell, resulting in the candidate list of functions, the tooltips and data input prompt appearing from the cell, and the function call replicated in the **Formula Bar**.
Alternately, one may type into the **Formula Bar**, in which case, the candidate list of functions, the tooltips and data input prompt appear from the **Formula Bar**, and the function call is replicated in the current cell.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

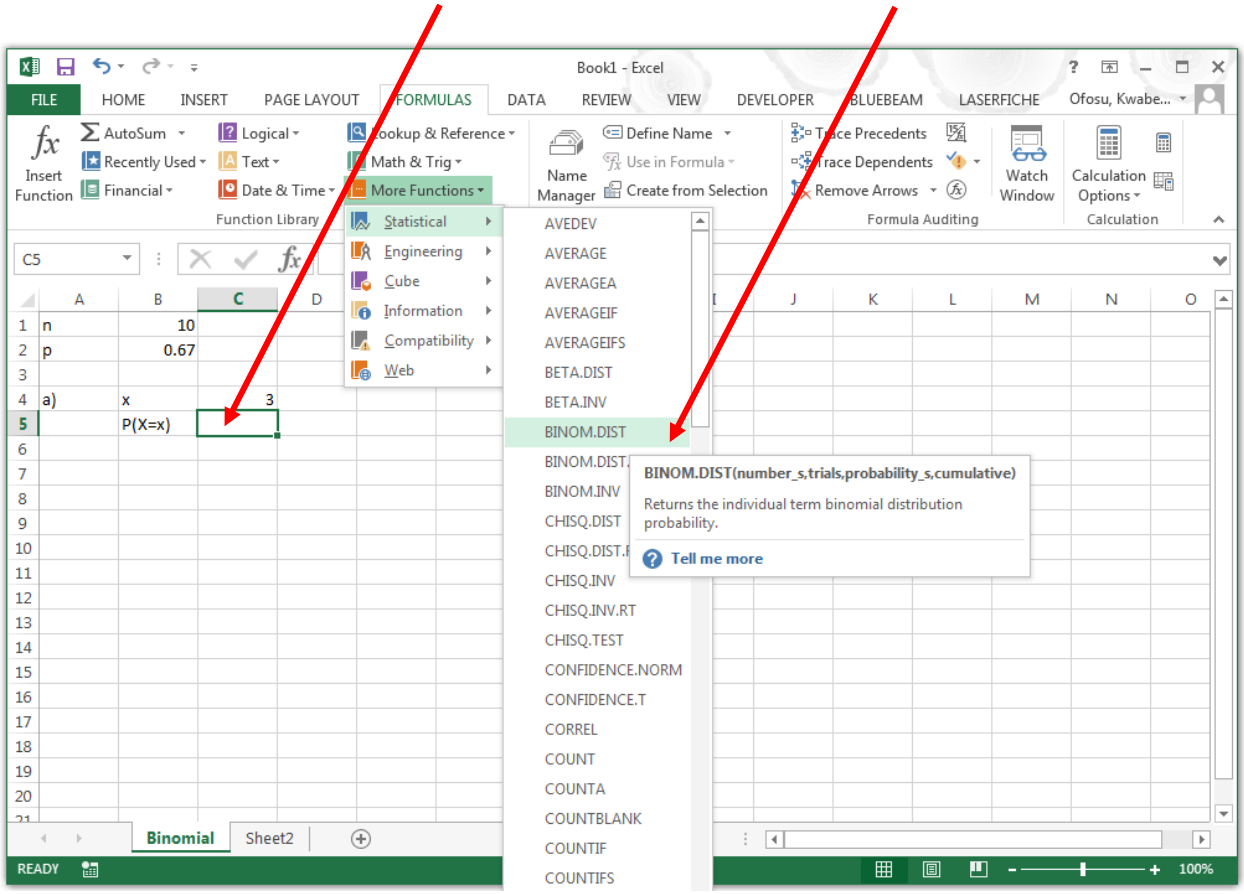
Revisiting the Binomial Distribution Example (Probability Calculations only):
A traffic engineer is studying the effectiveness of a yield sign placed on a ramp. A sample size of 10 vehicles was used. The probability that a driver will obey a yield sign is $2/3$.
a) What is the probability that exactly 3 vehicles will obey the sign?
b) What is the probability that less than 3 vehicles will obey the sign?
c) What is the probability that 2 or more vehicles will obey the sign?

Solution:
a) Based on the data given, set up the spreadsheet as follows:



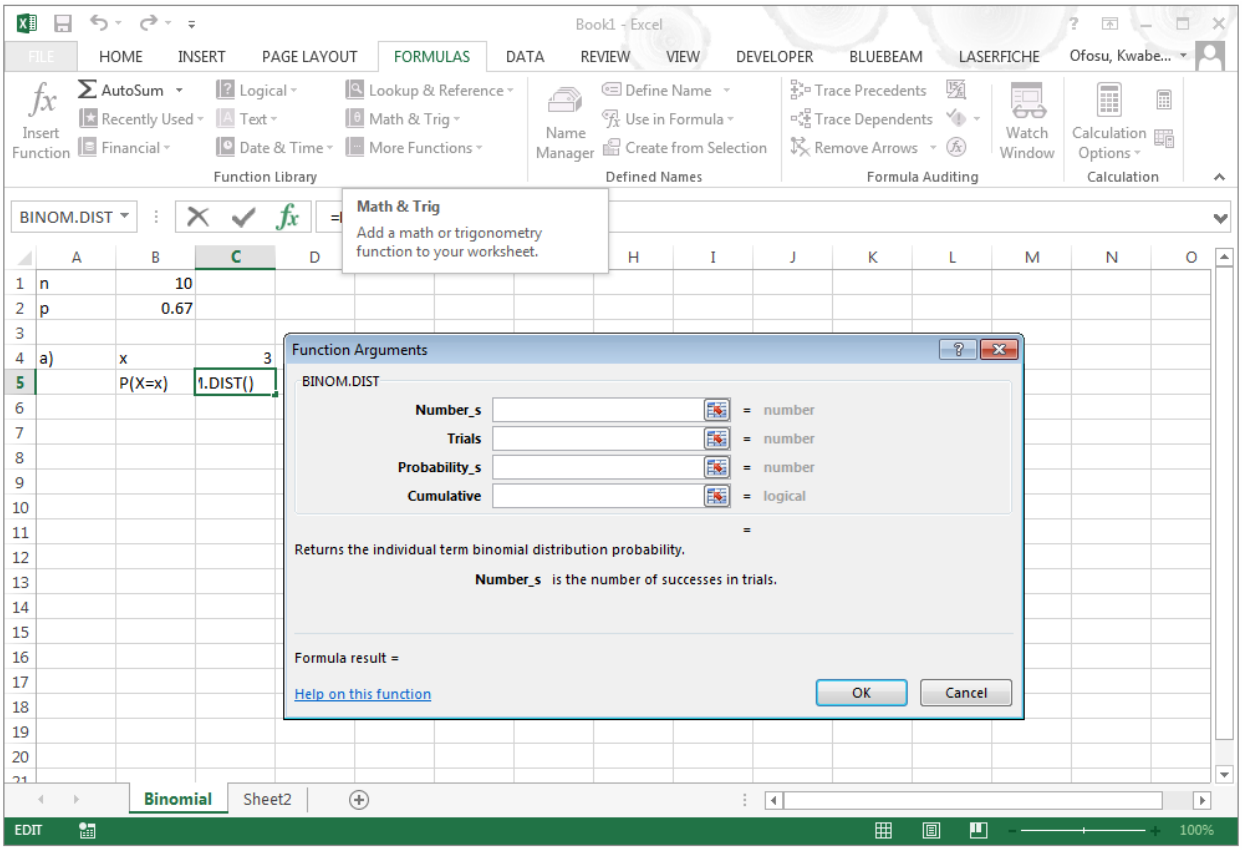
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

For the calculation to be conducted in the current cell (C5) as shown, click on FORMULAS.
Click on **More Functions**.
Click on **Statistical**.
Scroll down to look up the Binomial distribution function.
Hover over BINOM.DIST.
Review the tooltip and confirm this is indeed the relevant function.



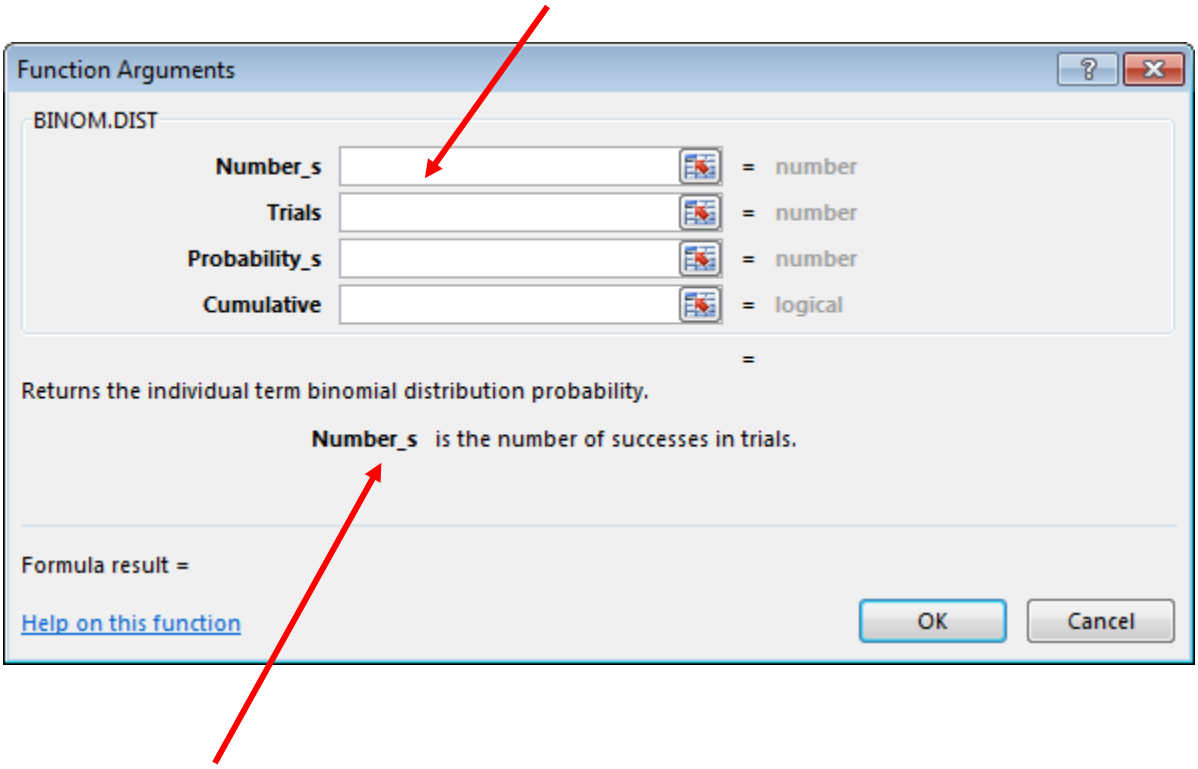
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Click on BINOM.DIST
The **Function Arguments** window opens.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

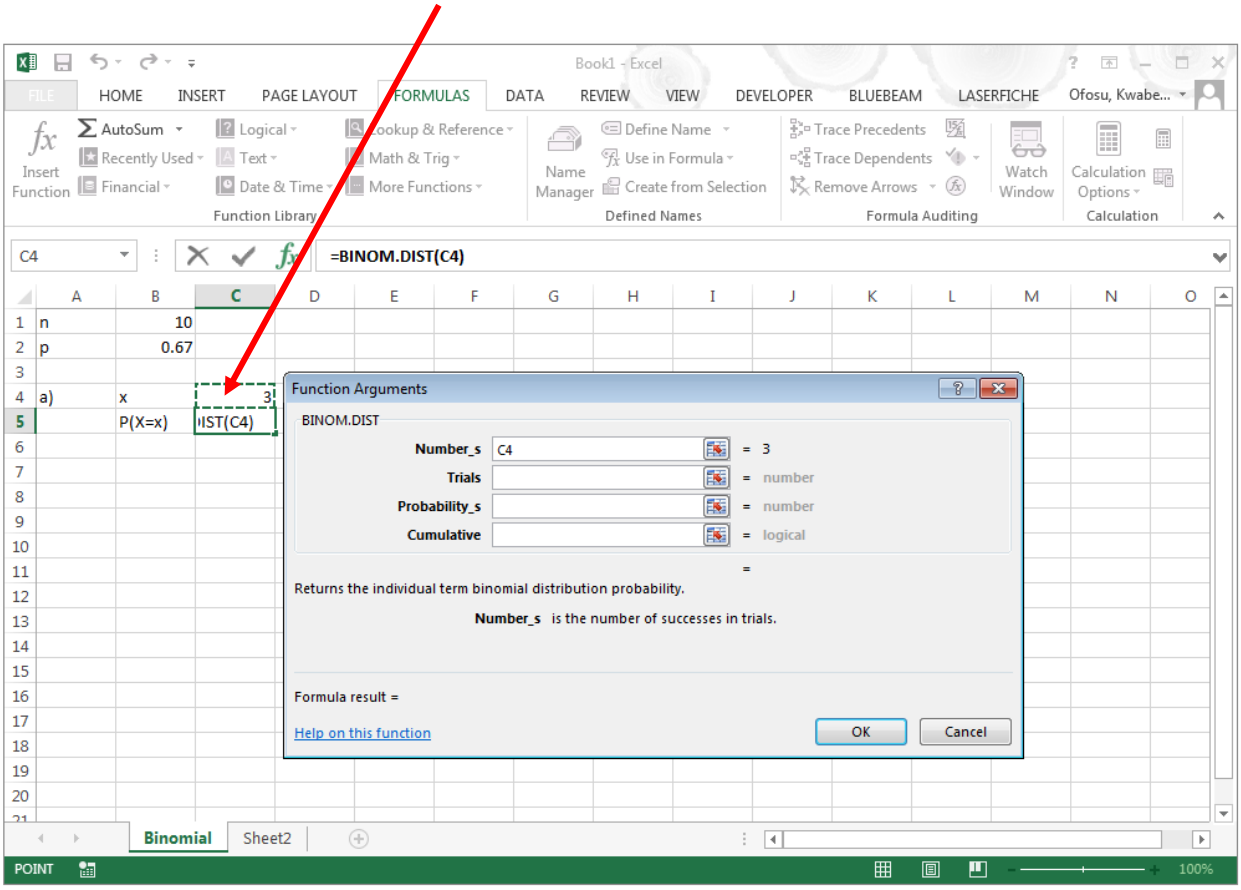
Click in the **Number_s** textbox.
Review the description.



For us, **Number_s** is our x value.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

With the cursor in the **Number_s** textbox, click on the cell holding the x value on the spreadsheet.



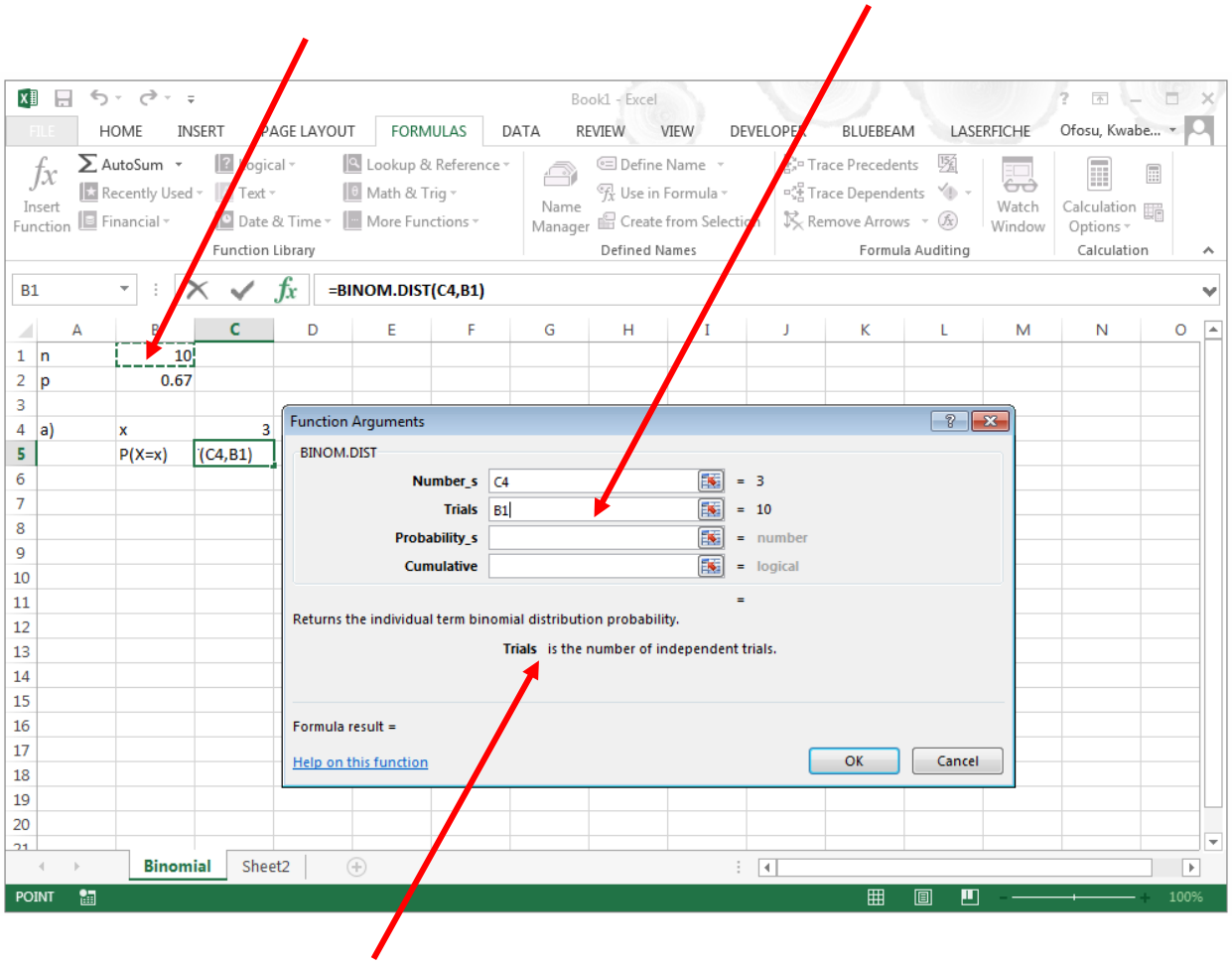
The cell reference is captured in the textbox.
Note that alternately one may type the cell reference into the textbox, however the **point-and-click** method is more efficient and less error prone, and therefore preferred.

Also, one may alternately enter the numerical value of the input (or **argument**) rather than capture the cell reference. However, as will become apparent throughout this course or for any *Excel* effort for that matter, using cell references enables data updates on spreadsheets to automatically recalculate results without re-entering the "hard" values. Using cell references also enables spreadsheets and incorporated formulas to be re-used for calculations on other projects.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

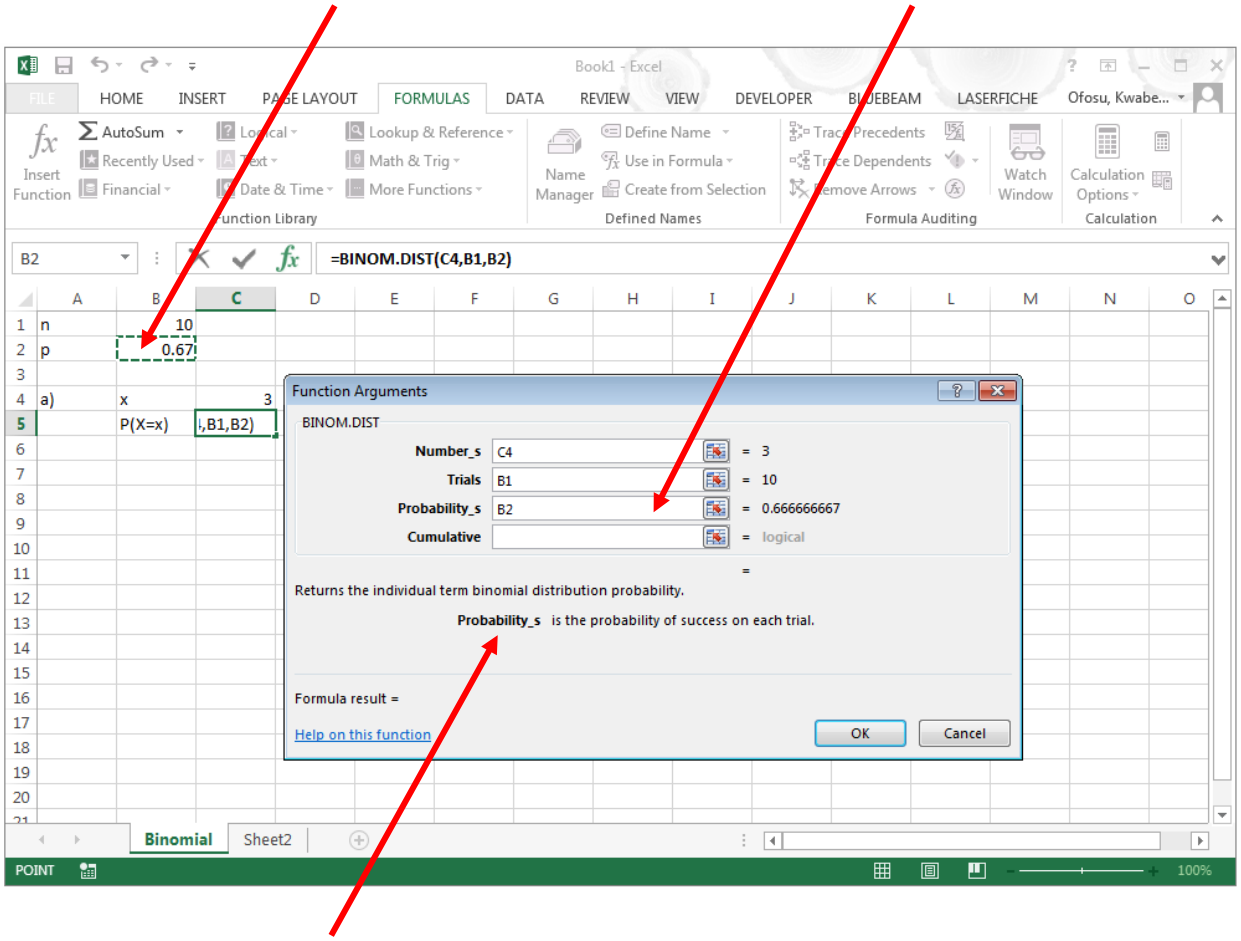
The automation, reuse, and replication of calculations across multiple repetitive-type projects is indeed a fundamental goal of engineers applying computer programs to begin with.

Click in the **Trials** textbox.
Review the description.
For us, **Trials** is the number of trials n.
With the cursor in the **Trials** textbox, click on the cell holding the n value on the spreadsheet.
(Note the updates in the **Formula Bar** as the function arguments are sequentially added.)



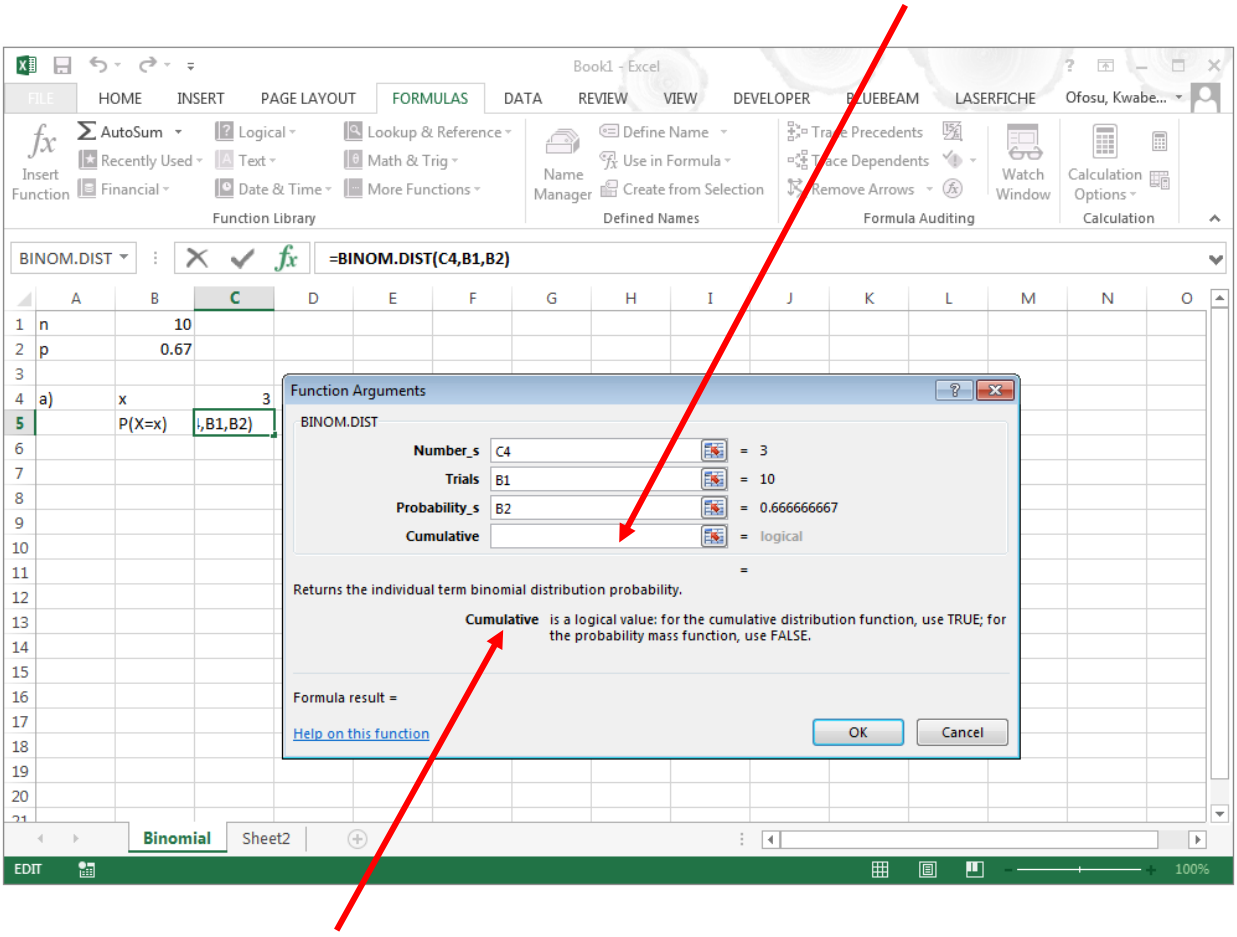
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Click in the **Probability_s** textbox.
Review the description.
For us, **Probability_s** is the probability p.
With the cursor in the **Probability_s** textbox, click on the cell holding the p value on the spreadsheet.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

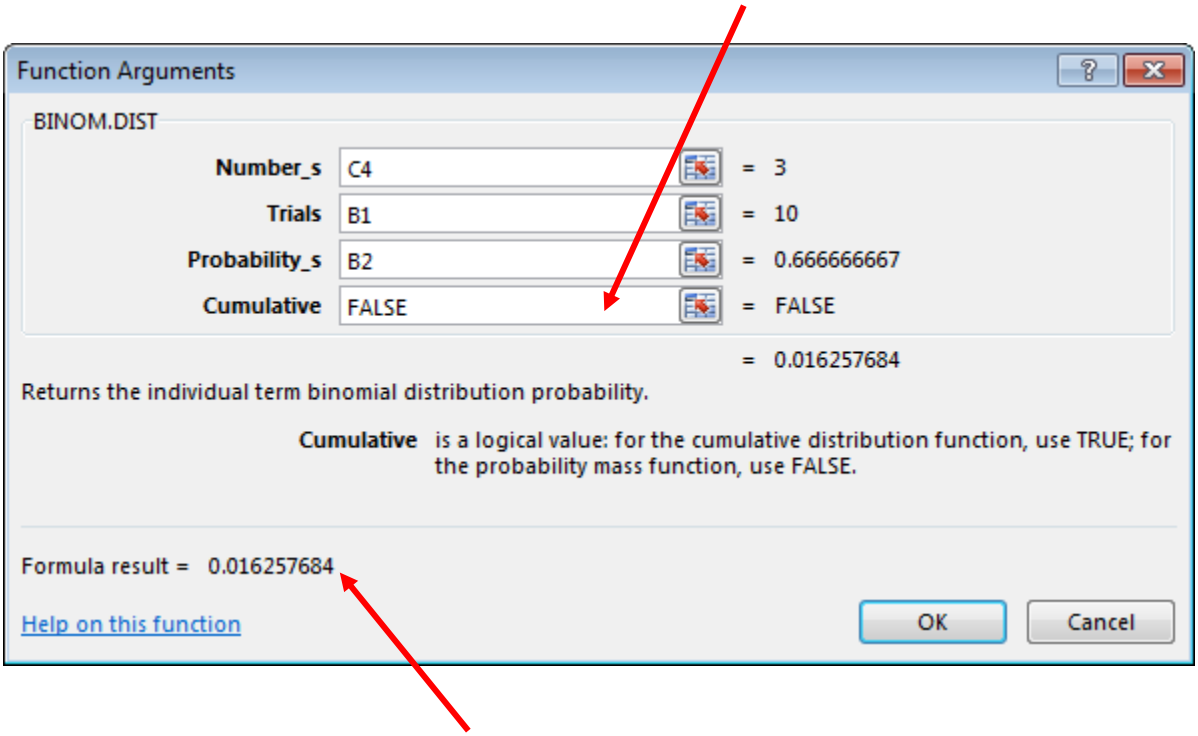
Click in the **Cumulative** textbox.
Review the description.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

In this part a) of the problem, we are not calculating a cumulative probability.
The probability mass function is the relevant one.
Thus, type in “FALSE”.

Review the **Formula result**. This serves as a preview.

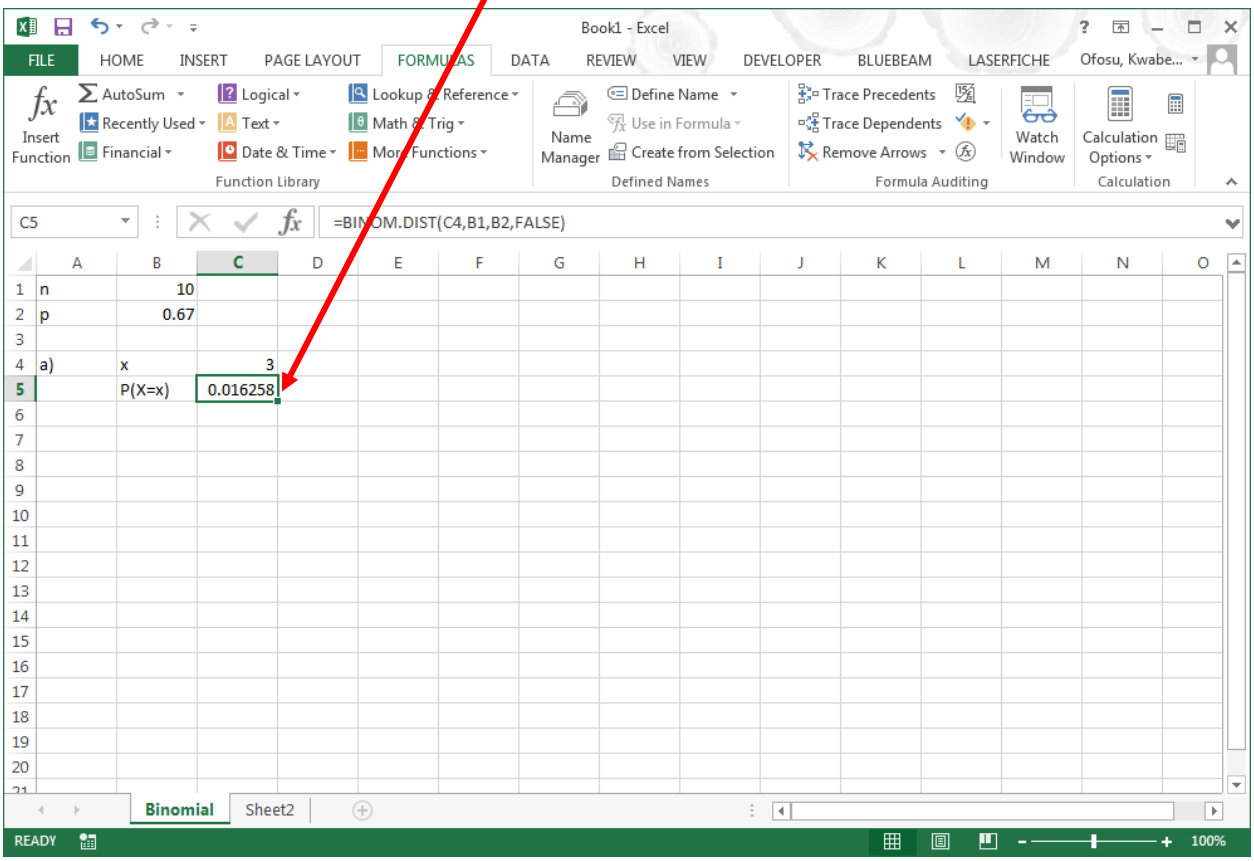


As an engineer or scientist, one is inclined to ask if this result makes sense. From fundamental principles, a probability value will be between 0 and 1. Thus, for example, a negative value, would indicate some input or other error of some form. In this case, the value previewed appears to be acceptable and correct.

Hit **OK**.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

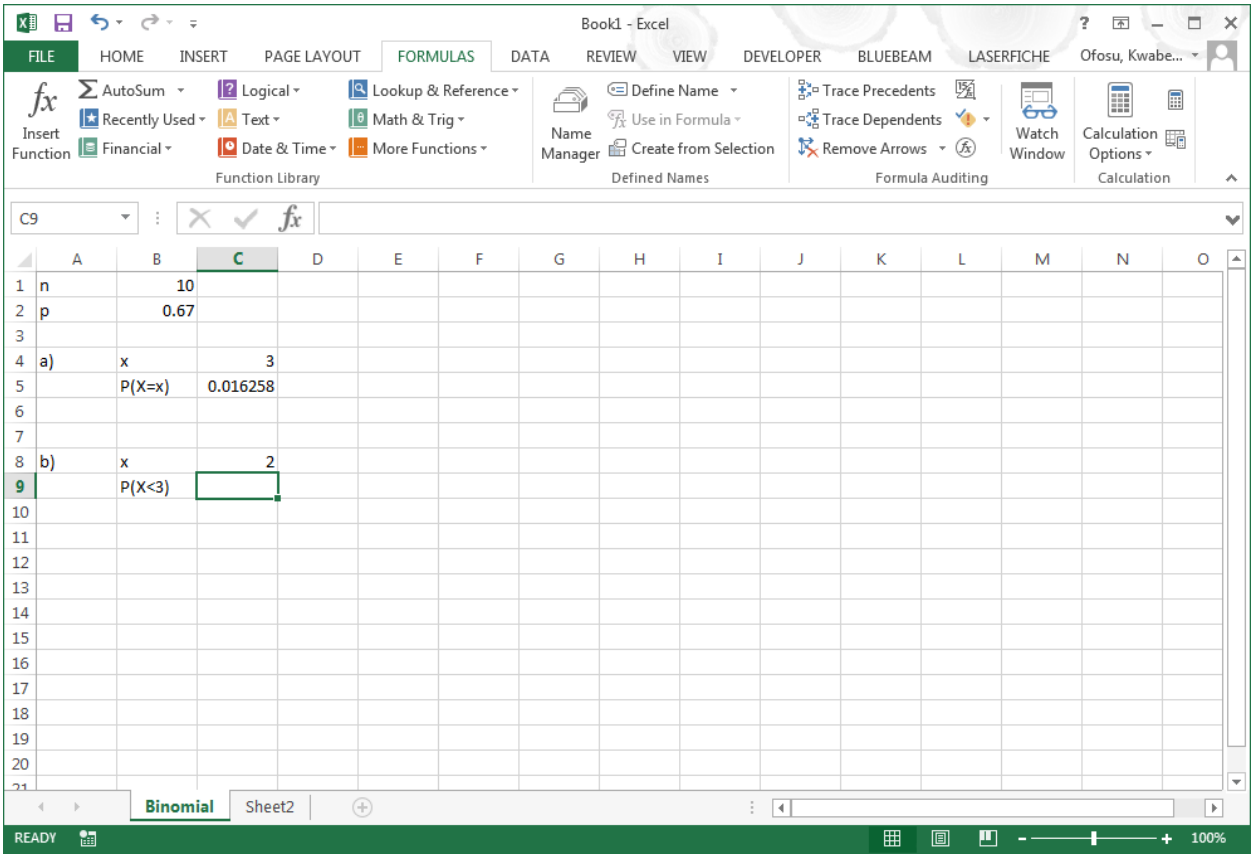
The **Function Arguments** window is dismissed.
The result is displayed in the spreadsheet cell.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

b) What is the probability that less than 3 vehicles will obey the sign?
It was previously demonstrated that $P(X < 3) = P(X \leq 2)$

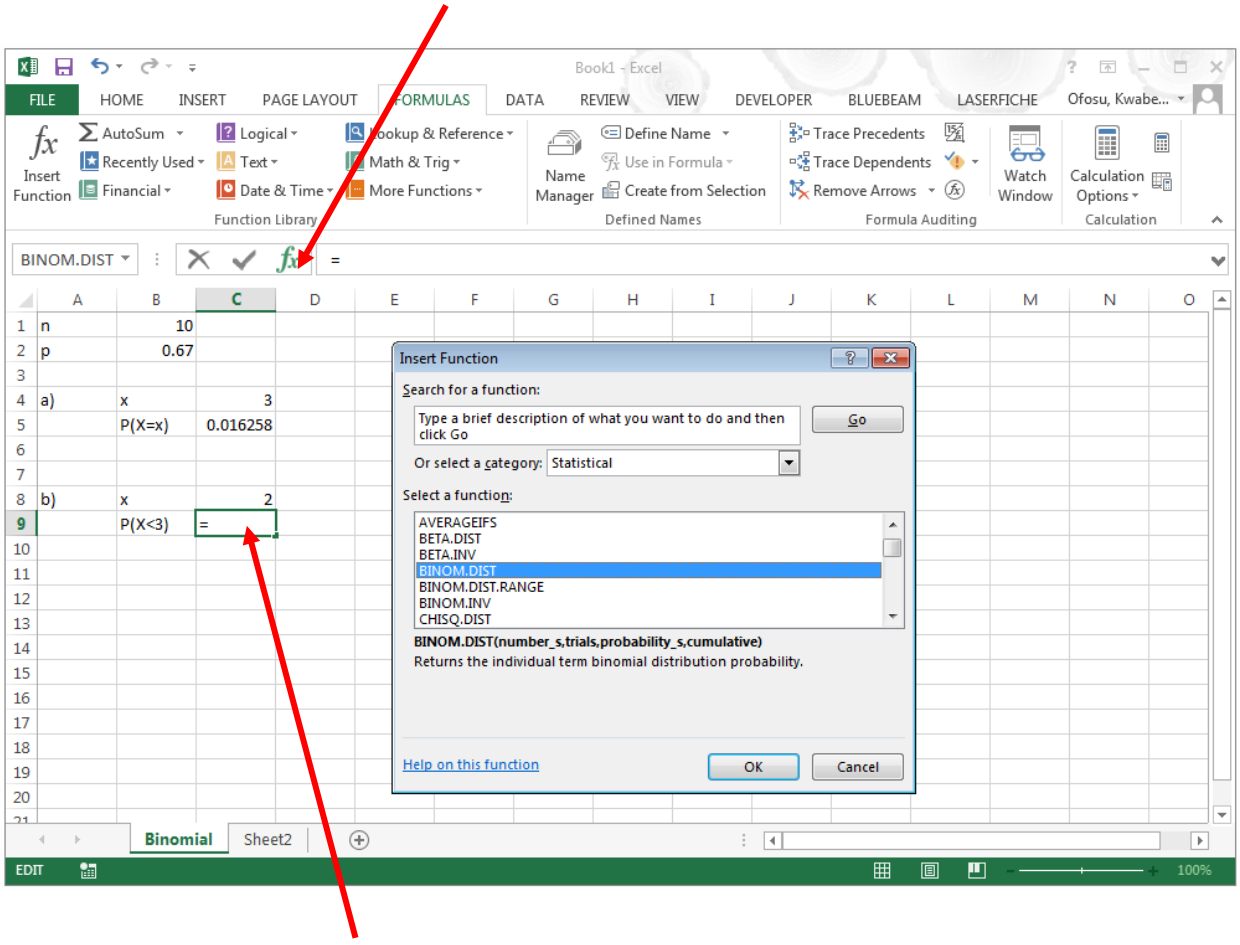
Update the spreadsheet as follows:



For this calculation use the **Insert Function** icon to access the statistical functions and thus the Binomial distribution function.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

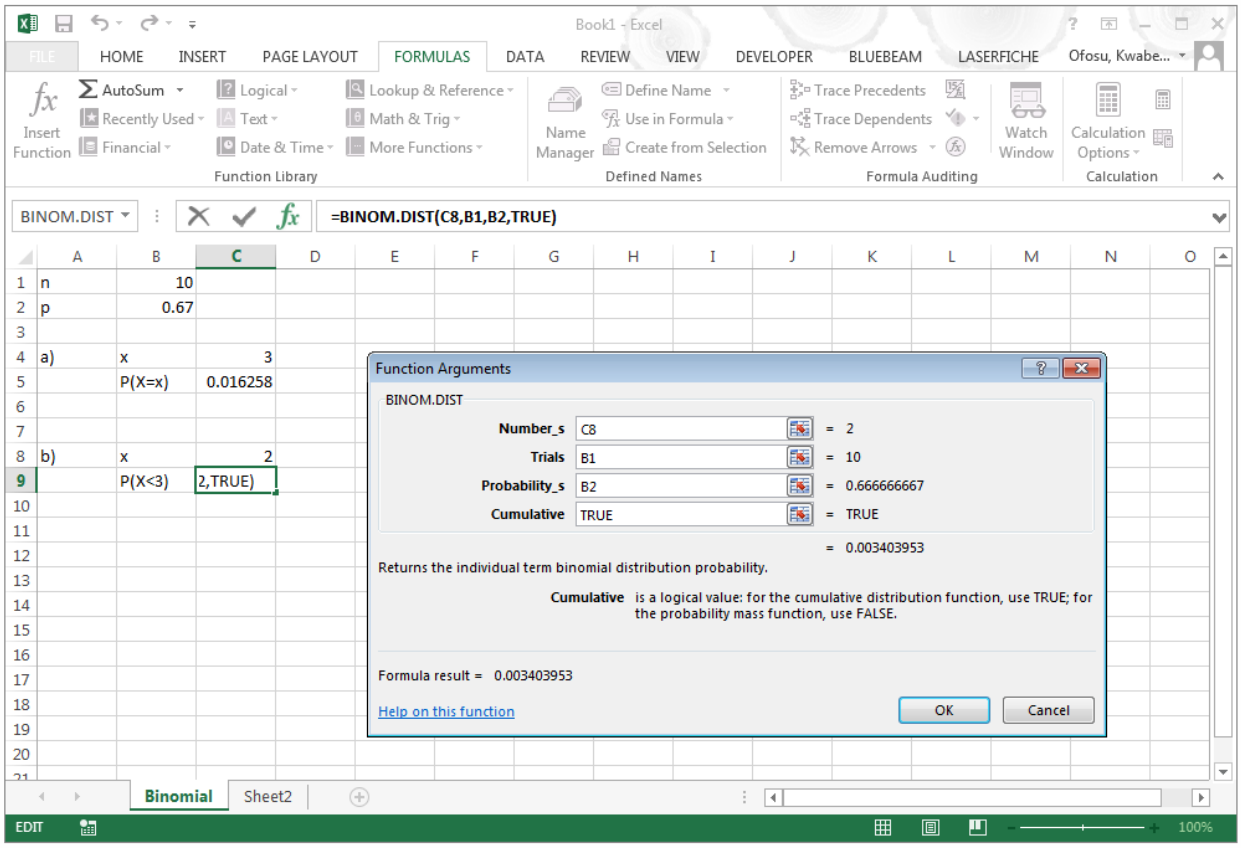
Click on **Insert Function**.
Under **Or select a category**, select **Statistical**.
Under **Select a function**, scroll to BINOM.DIST.
Review the description and confirm this is the relevant function.



Hit **OK** to open the **Function Arguments** window.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

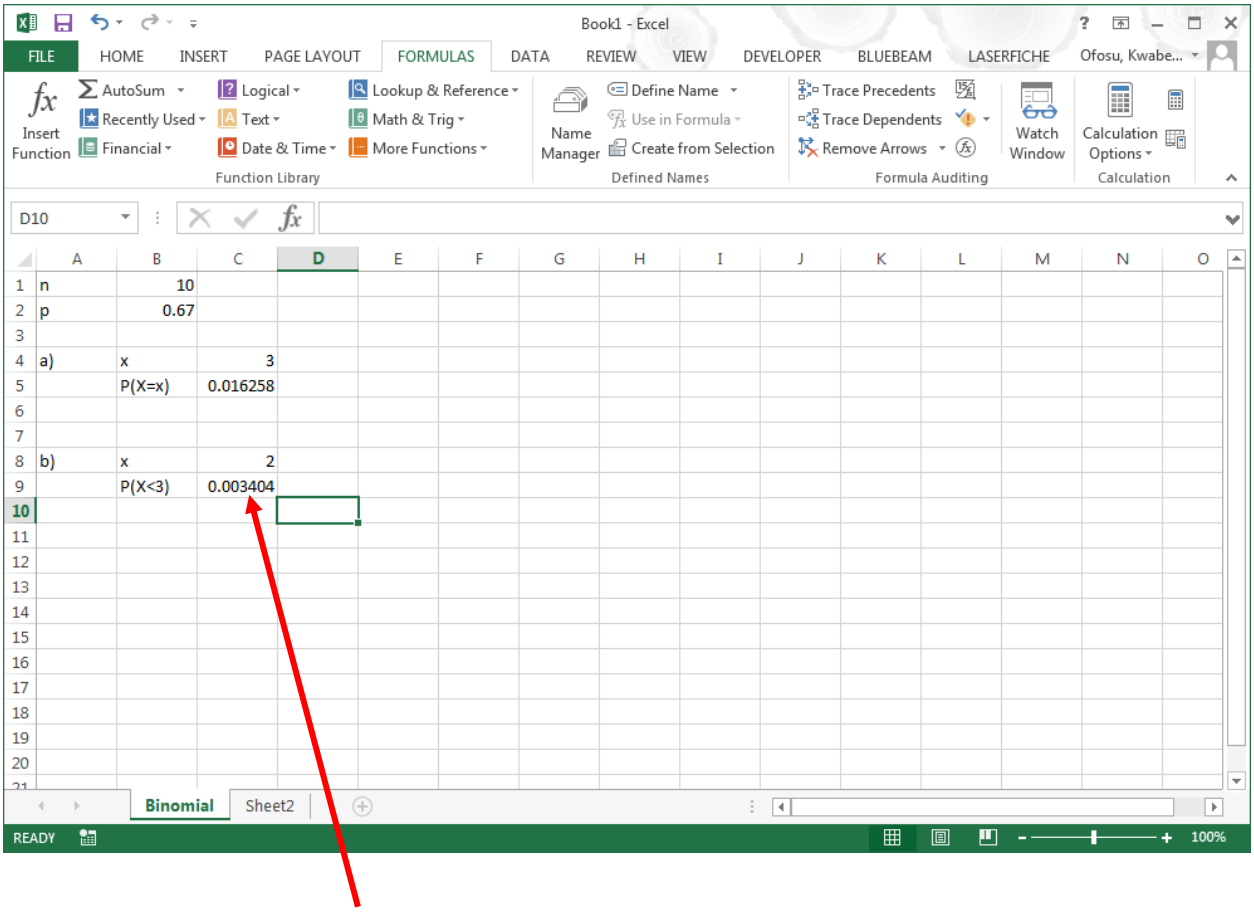
In the **Function Arguments** window, point-and click to select the arguments.
For this computation, the goal is to compute a cumulative probability.
For the **Cumulative** argument, type in "TRUE".



Review and concur with **Formula result**.
Hit **OK**.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

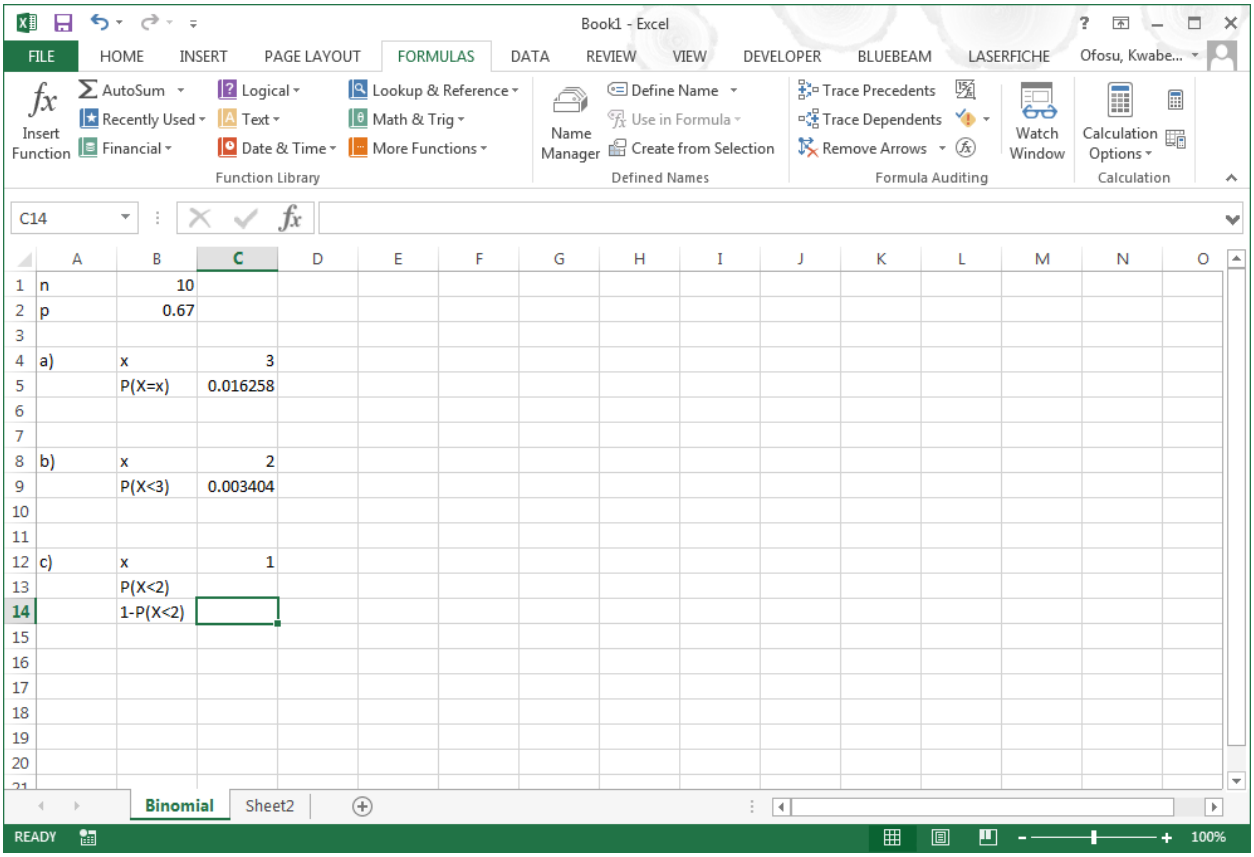
Review the result on the spreadsheet.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

c) What is the probability that 2 or more vehicles will obey the sign?
It was previously demonstrated that $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1)$
In this demonstration we shall use the third method of accessing functions which is by typing the function name directly into the cell or into the **Formula Bar** once a cell to display the result is selected.

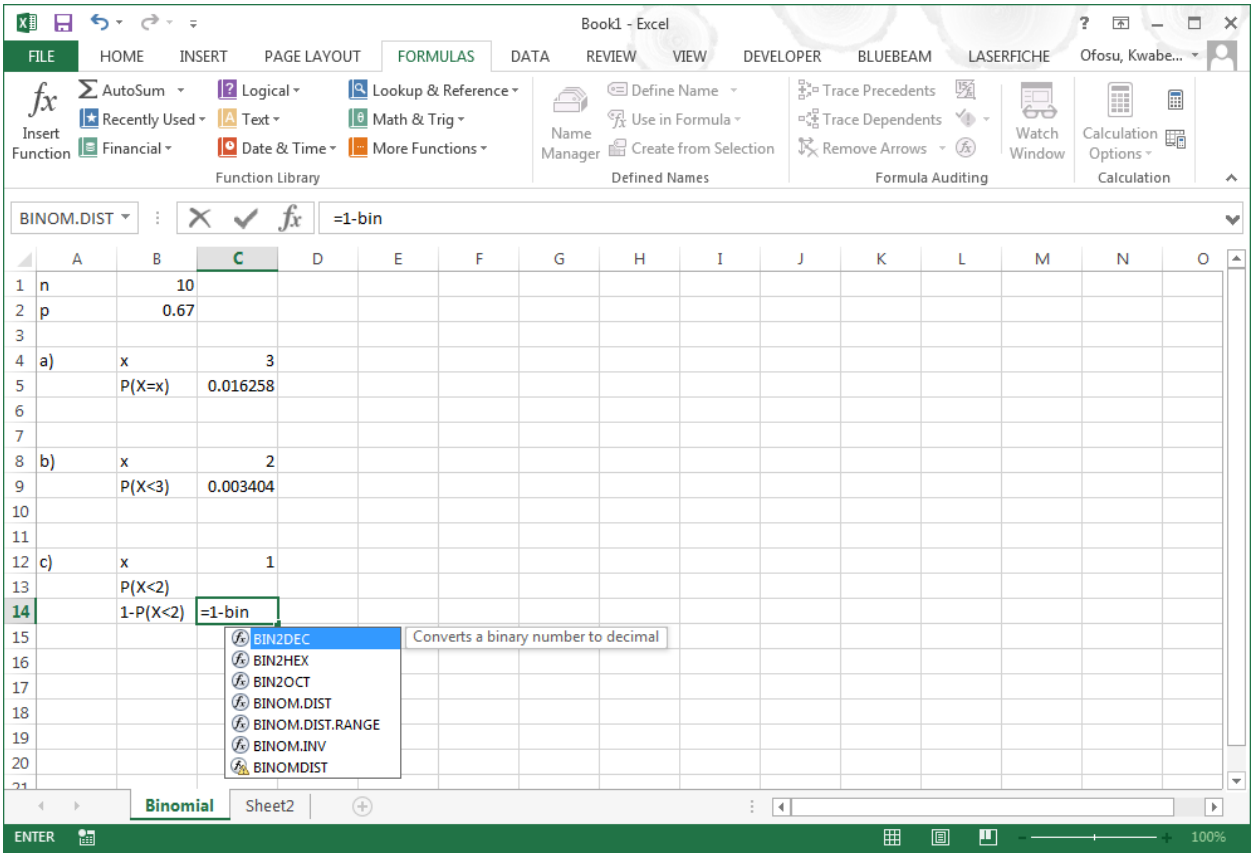
Update the spreadsheet as follows:



	A	B	C
1	n	10	
2	p	0.67	
3			
4	a)	x	3
5		P(X=x)	0.016258
6			
7			
8	b)	x	2
9		P(X<3)	0.003404
10			
11			
12	c)	x	1
13		P(X<2)	
14		1-P(X<2)	
15			
16			
17			
18			
19			
20			

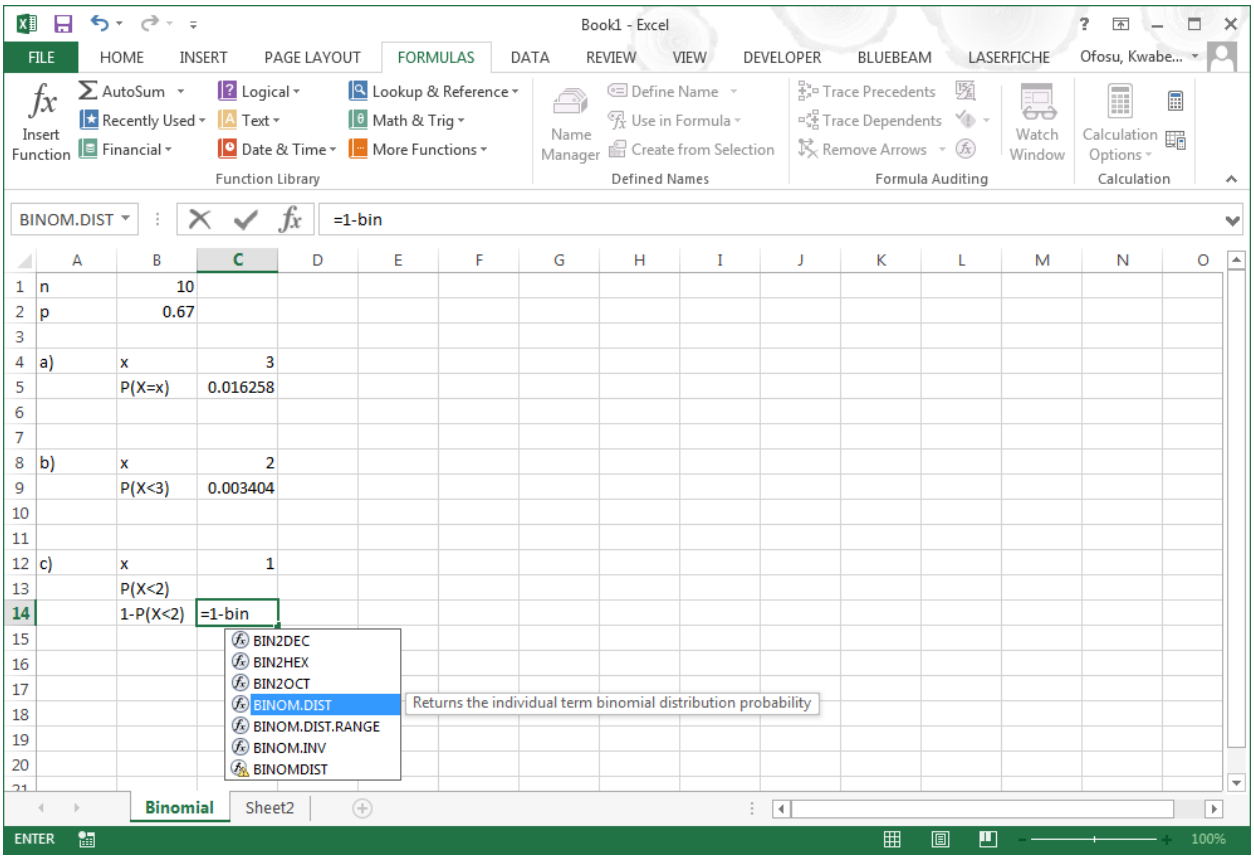
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

In the selected cell type the following
" = 1 - BINOM.DIST "
Before completing the typing, a candidate list of functions appears.



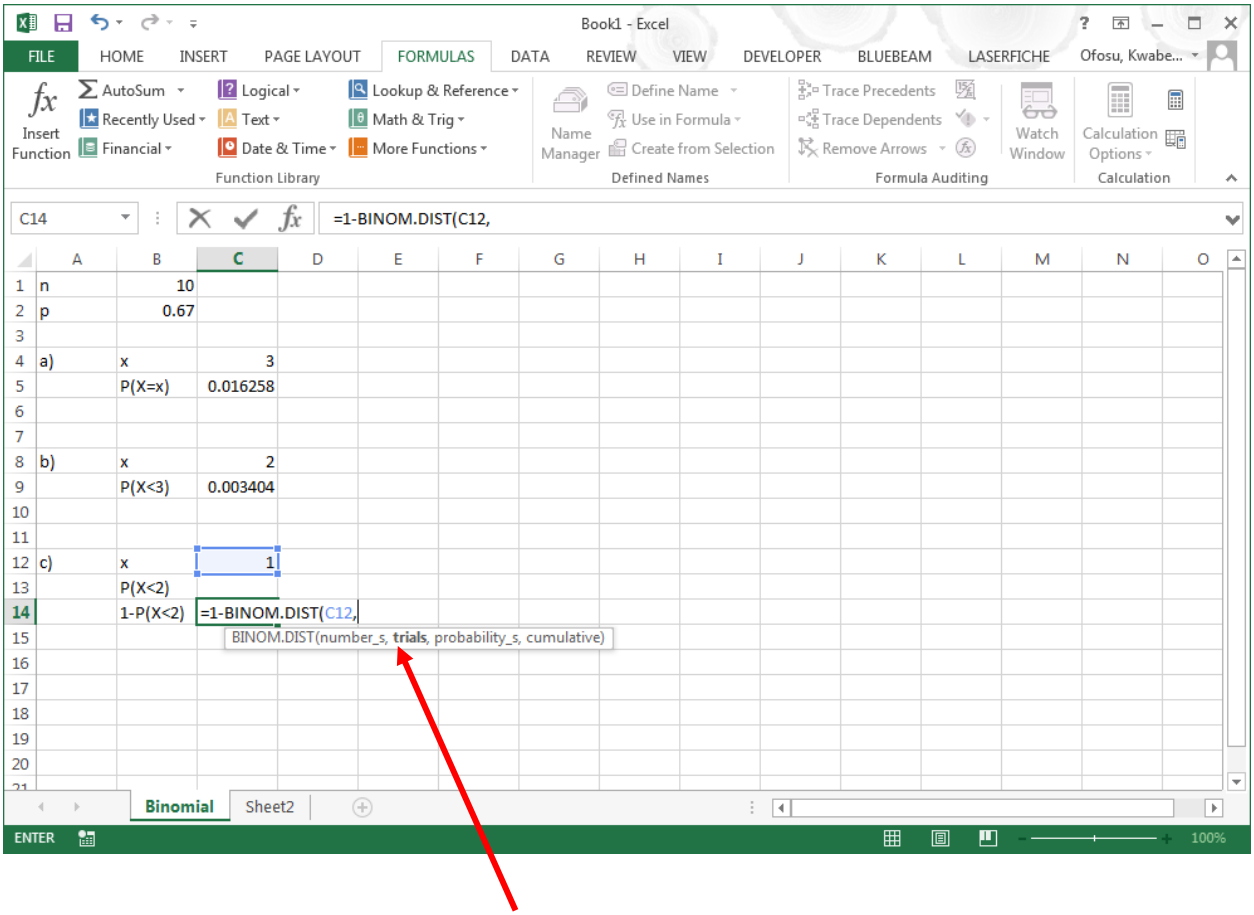
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Select BINOM.DIST.
Review the tooltip and confirm this is the relevant function.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

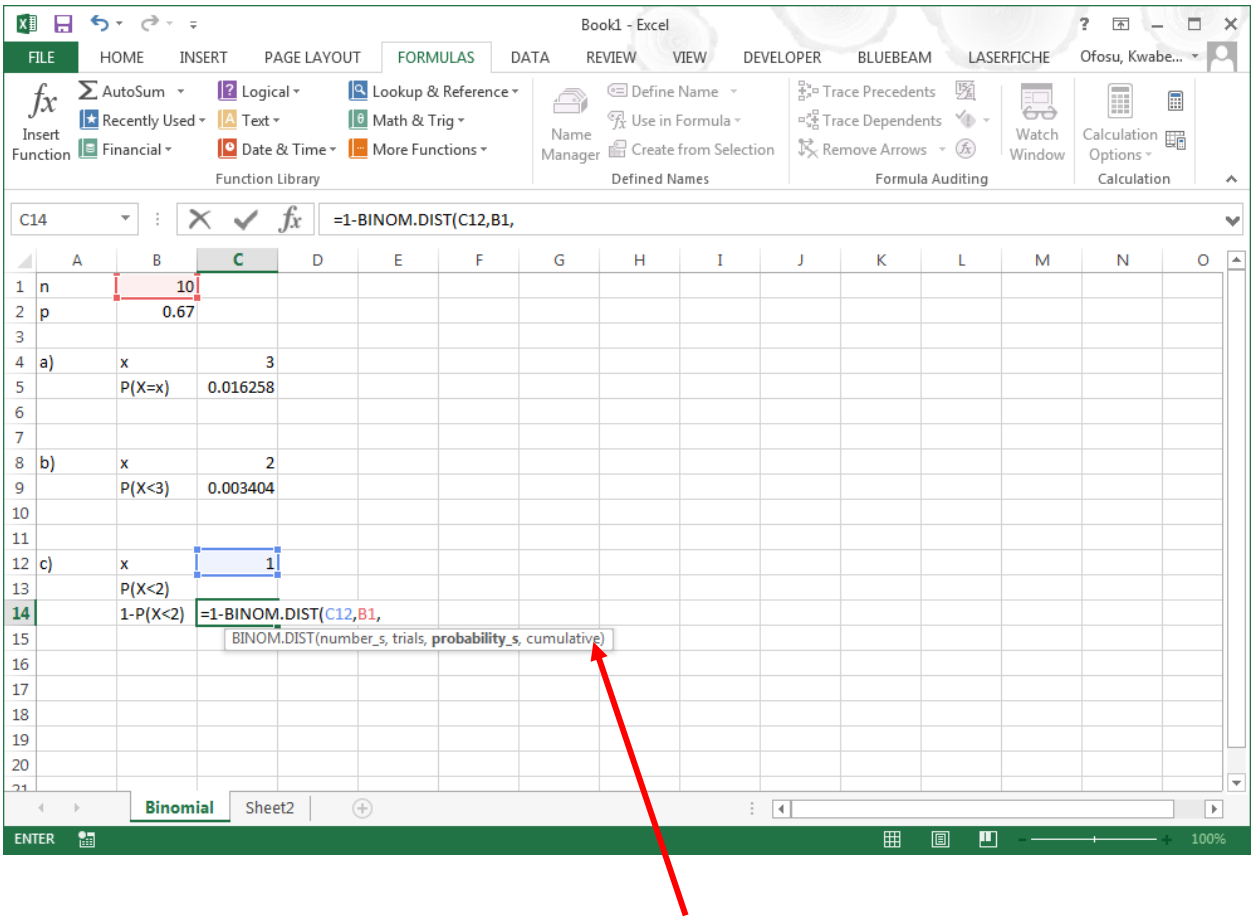
Double click on BINOM.DIST.
The prompt opens.
Following the tooltip, point-and-click to select the cell holding the **number_s** argument.
(Or type in the cell reference for the **number_s** argument).
Type a comma (“,”)



The comma pushes the prompt to the next argument.

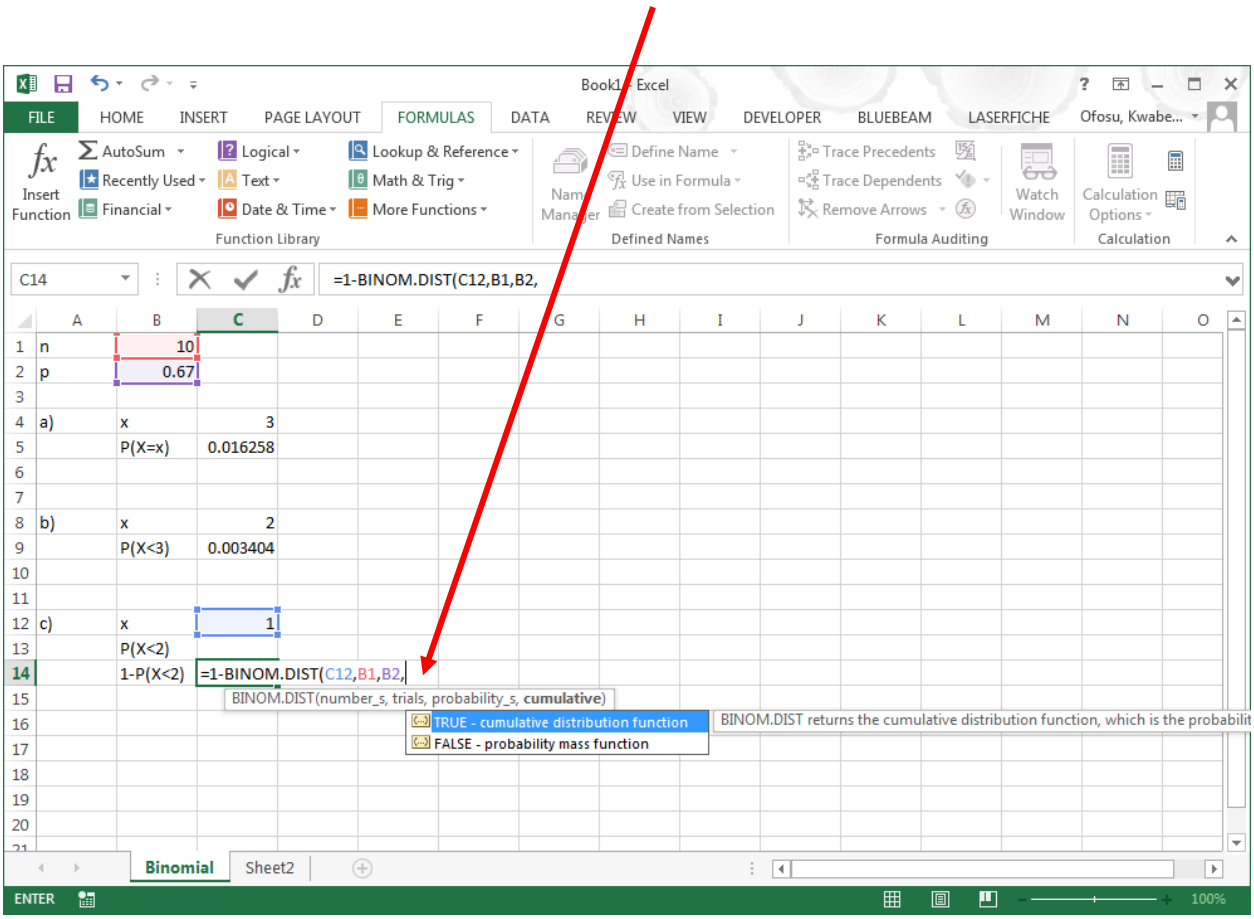
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Point-and-click (or type in the cell reference) to select the cell holding the **trials** argument. Type a comma (“,”)



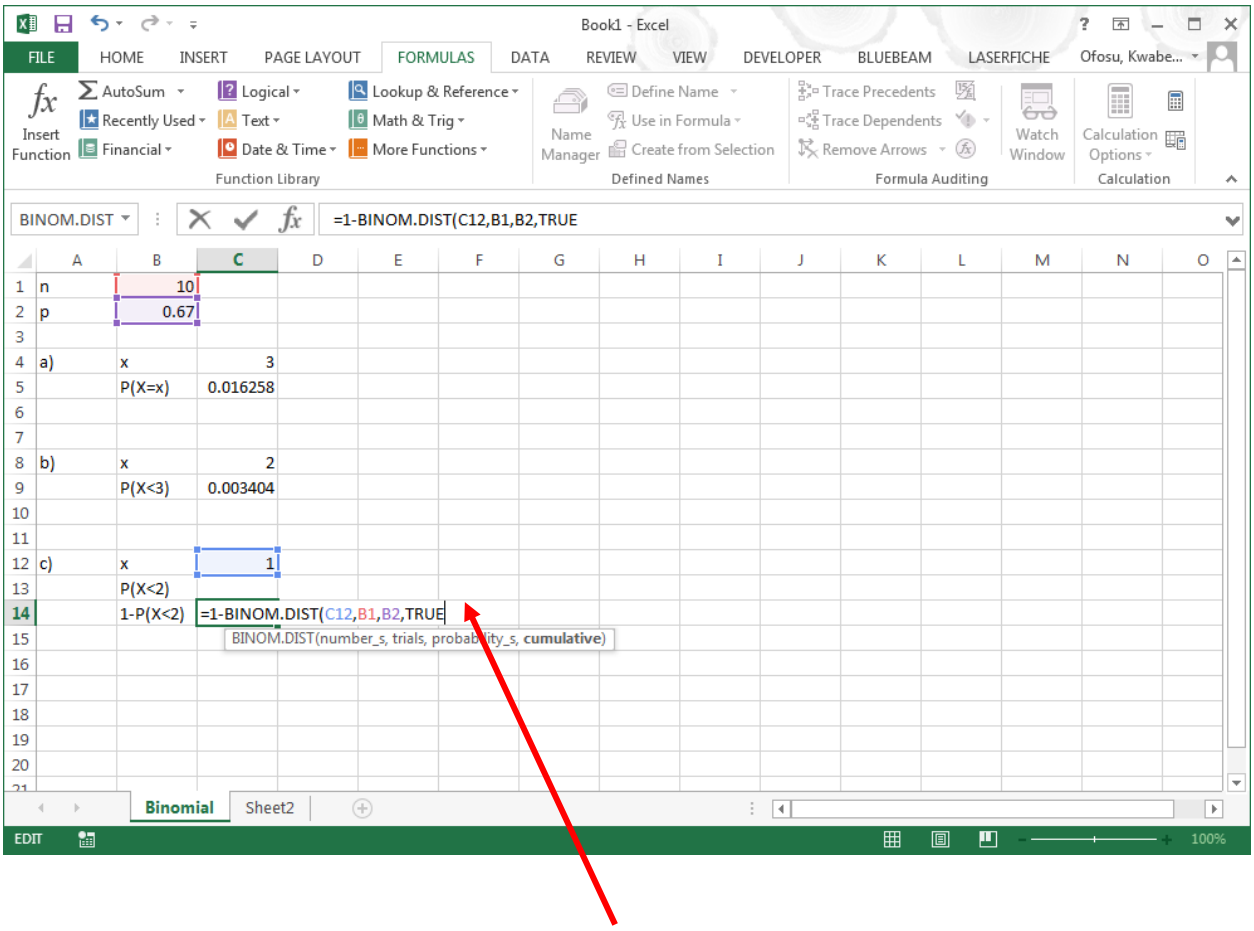
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Point-and-click (or type in the cell reference) to select the cell holding the **probability_s** argument.
Type a comma (“,”)



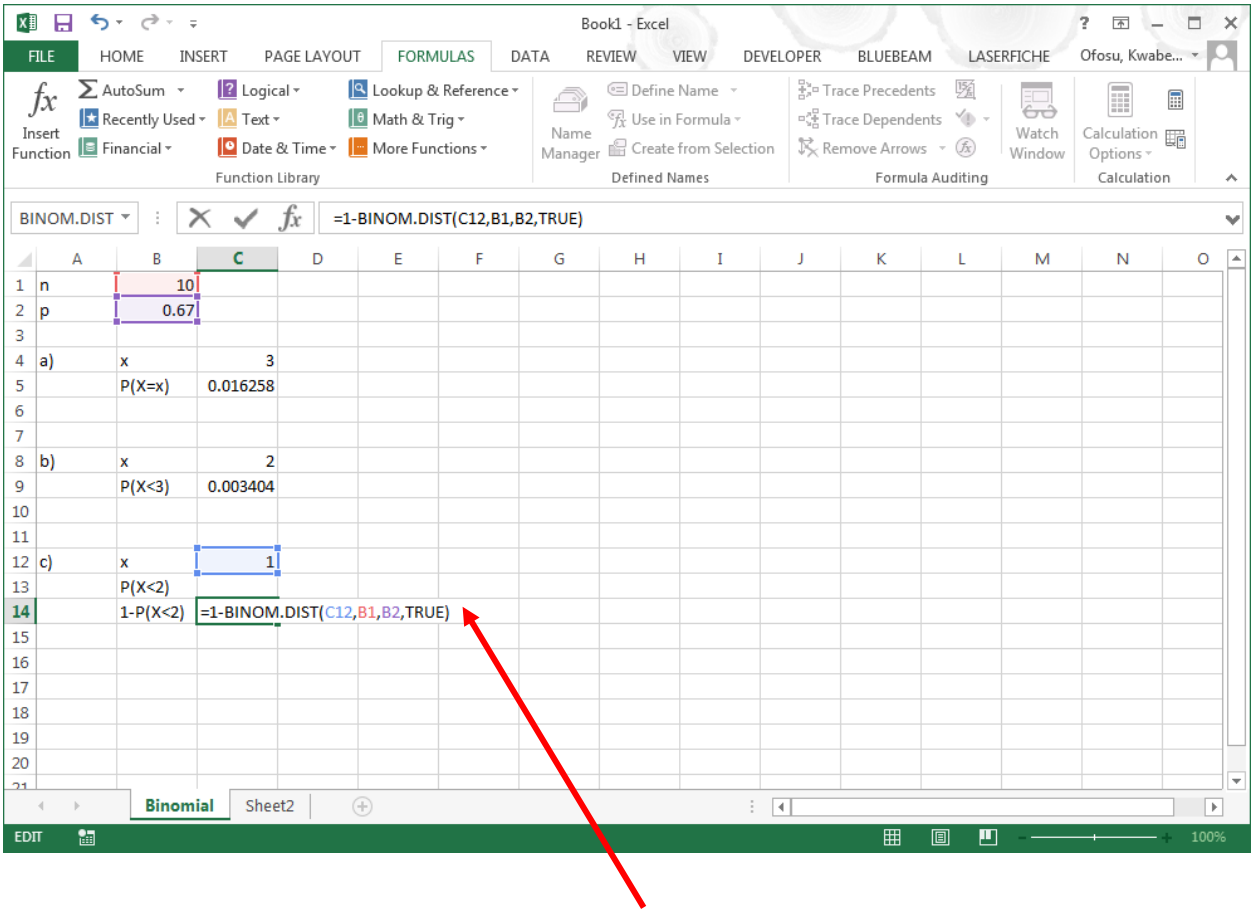
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

As this is a cumulative probability calculation, double click on “TRUE” from the list (or simply type in “TRUE” directly).



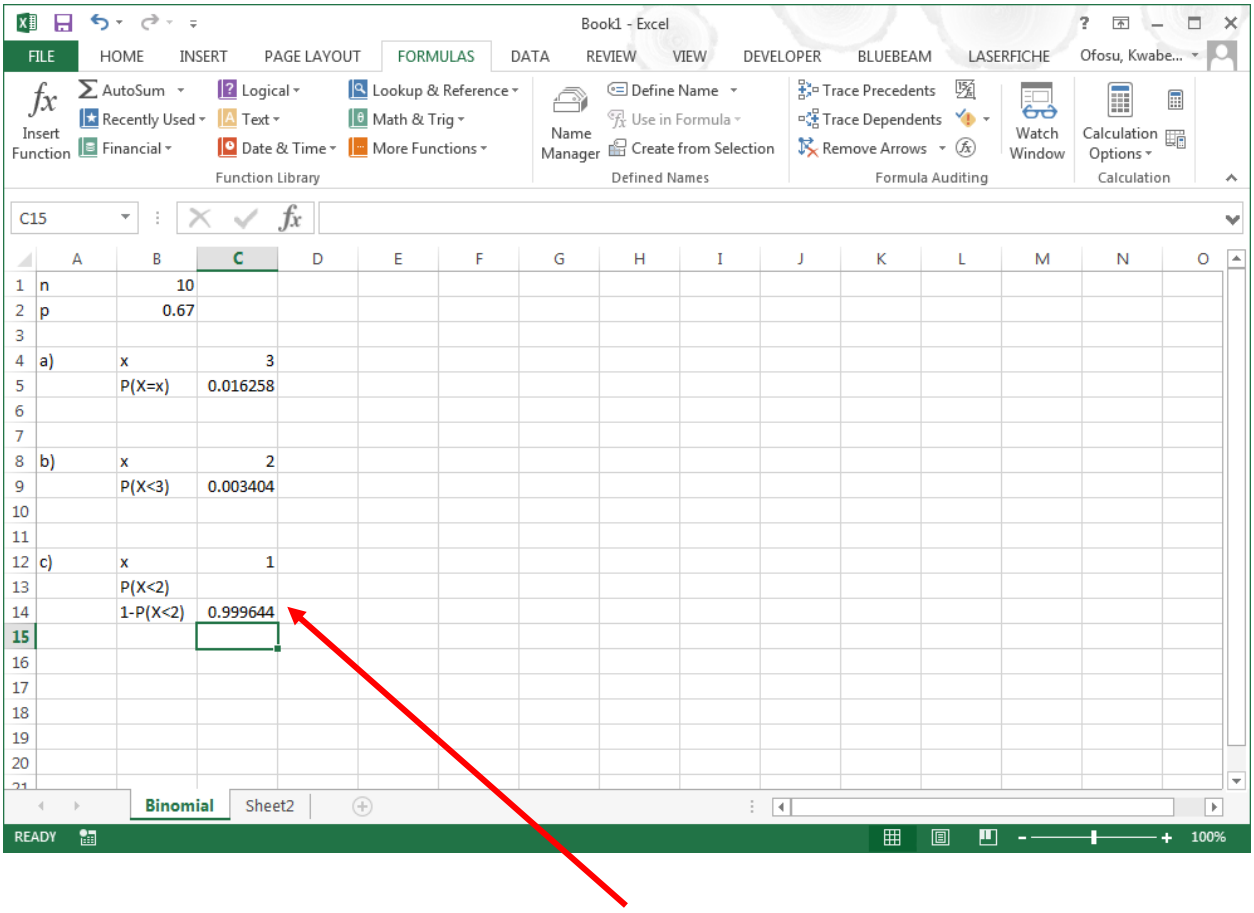
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Close parentheses.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Hit **Enter** on the keyboard.



	A	B	C
1	n	10	
2	p	0.67	
3			
4	a)	x	3
5		P(X=x)	0.016258
6			
7			
8	b)	x	2
9		P(X<3)	0.003404
10			
11			
12	c)	x	1
13		P(X<2)	
14		1-P(X<2)	0.999644
15			
16			
17			
18			
19			
20			

The calculation is complete.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4. CONTINUOUS RANDOM VARIABLES

4.1 Definition

In contrast to discrete random variables that have values (x) that are limited to zero and positive integer values, continuous random variables can take on integer and decimal values over their domains. Therefore, over a given range of values, there will be infinitely many possible values (x) the continuous random variable can actualize.

4.2 Properties of Continuous Random Variables

1. From Equation (3.5), it follows that the cumulative probability that a discrete random variable (X) will take on a value (x_i) over a range of values [a ,] is given by,

$$P(a \le X \le b) = \sum_{i=a}^{b} f(x_i)$$

. . . . . **Equation (4. 1)**

But for a continuous random variable there are infinitely many x_i values over the range [a , b], and if the distribution function $f(x_i)$ is smooth and can be integrated, then

$$P(a \le X \le b) = \int_{a}^{b} f(x)dx$$

. . . . . **Equation (4. 2)**

The distribution function of a continuous random variable is called the **probability density function** (or **density function**).

2. The density function satisfies the following requirements.

$f(x) > 0$ for all x,

and

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

. . . . . **Equation (4. 3)**

In other words, the total area under the density curve equals unity. Therefore $P(a \leq X \leq b)$ is the area under the density curve between $x = a$ and $x = b$.

3. As with discrete random variables, the summation of the probabilities over a range of x values is given by the **cumulative density function** ($F(x)$).

$$P(X \leq x) = \int_{-\infty}^{x} f(x)dx = F(x)$$

. . . . . **Equation (4. 4)**

Thus,

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$$

. . . . . **Equation (4. 5)**

Equation (4.5) implies that if the derivative of the cumulative density function exists, then

$$\frac{d[F(x)]}{dx} = f(x)$$

. . . . . **Equation (4. 6)**

In contrast to discrete random variables, calculations with continuous random variables generally involve computing cumulative probabilities.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4. Since the probabilities for continuous random variables are obtained through integration, the inequality conditions are not relevant for continuous random variables. Therefore, for all intents and purposes,

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a < X < b) = P(a \leq X < b)$$

5. The expected value (μ) of a continuous random variable is defined as,

$$\mu = \int_{-\infty}^{\infty} x.f(x)dx$$

. . . . . **Equation (4. 7)**

6. The variance (σ^2) of a continuous random variable is defined as,

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2.f(x)dx$$

. . . . . **Equation (4. 8)**

where σ is the standard deviation of the random variable.

4.3 The Uniform Distribution

The Uniform distribution is a continuous random variable that has equal probability of occurrence for values in a range $[\alpha, \beta\ ]$ and a probability of zero for all other values.

Graphically, the uniform distribution looks as follows,

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

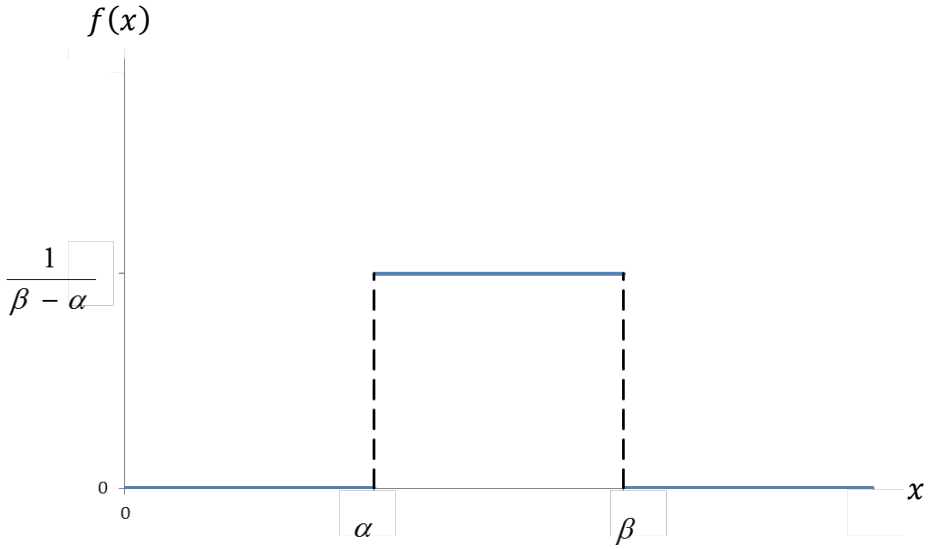


Figure 4. 1: The Uniform distribution

Mathematically, the probability density function of the uniform distribution is,

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & for\ \alpha \le x \le \beta \\ \\ 0 & for\ all\ other\ x \end{cases}$$

. . . . . **Equation (4. 9)**

From Equation (4.5), the cumulative probability over a range $[a, b\ ]$ is the area under the density function between a and b.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

From Equation (4.7), it can be shown that the expected value (μ) of the Uniform distribution is given by,

$$\mu = \frac{\alpha + \beta}{2} \quad \text{. . . . . Equation (4. 10)}$$

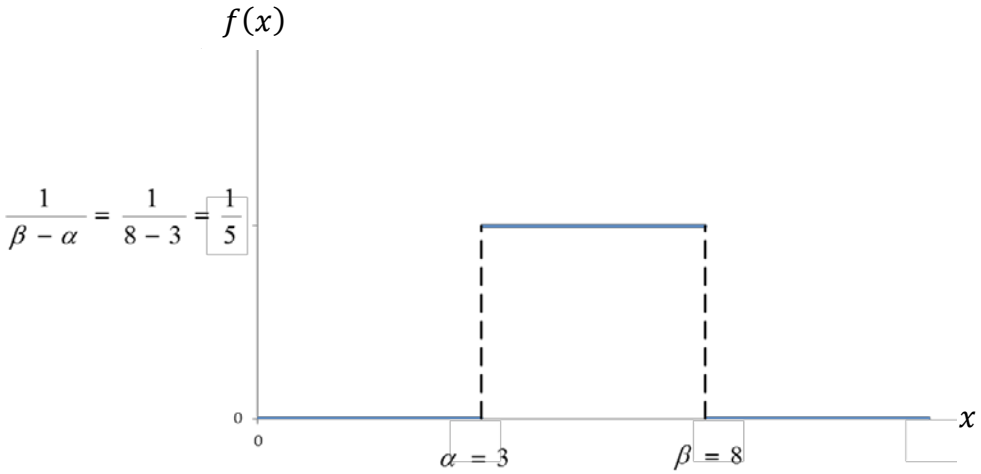
From Equation (4.8), it can be shown that the variance (σ^2) of the Uniform distribution is given by,

$$\sigma^2 = \frac{(\beta - \alpha)^2}{12} \quad \text{. . . . . Equation (4. 11)}$$

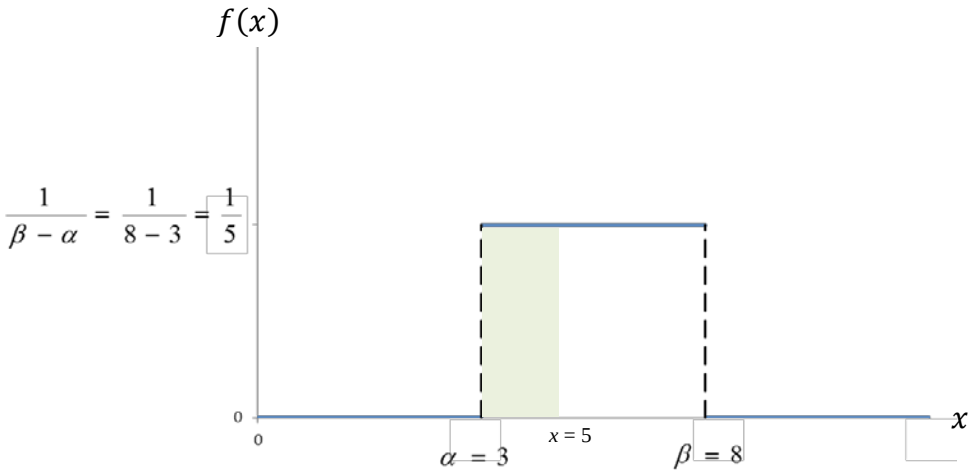
Example:
An emergency response team for a sector of a city estimates that the response time to any incident on the sector is between 3 to 8 minutes with any specific response time in the range equally likely.
a) Calculate the probability that a response time will be less than 5 minutes.
b) Calculate the probability that the response time will be between 4 and 6 minutes.
c) What is the 95^{th} percentile response time?

Solution:
Depicting the distribution graphically,

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE



a) This is the cumulative area under the density function up to $x = 5$.

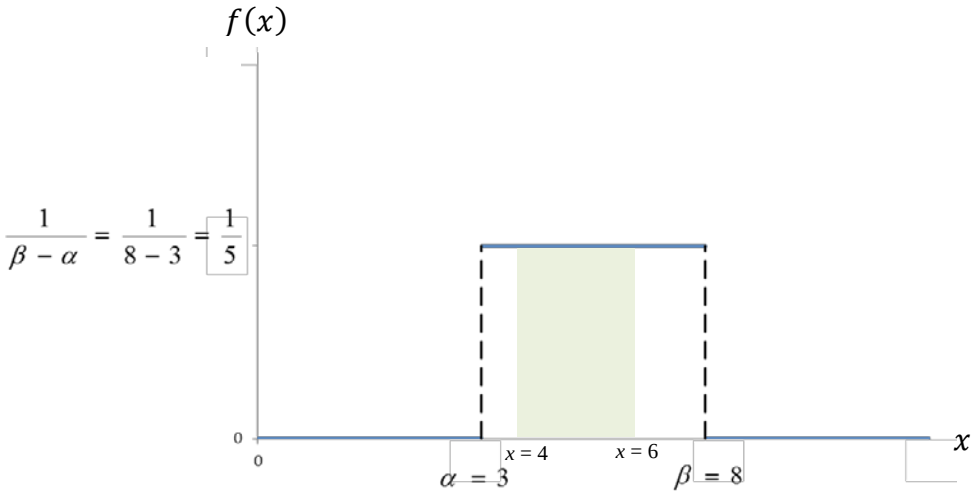


Due to the simple rectangular shape of the area enclosed by the density function, there is no need to conduct integration to calculate the area.

$$P(X \leq 5) = \int_{-\infty}^{5} f(x)dx = \frac{1}{5} * (5 - 3) = 0.4$$

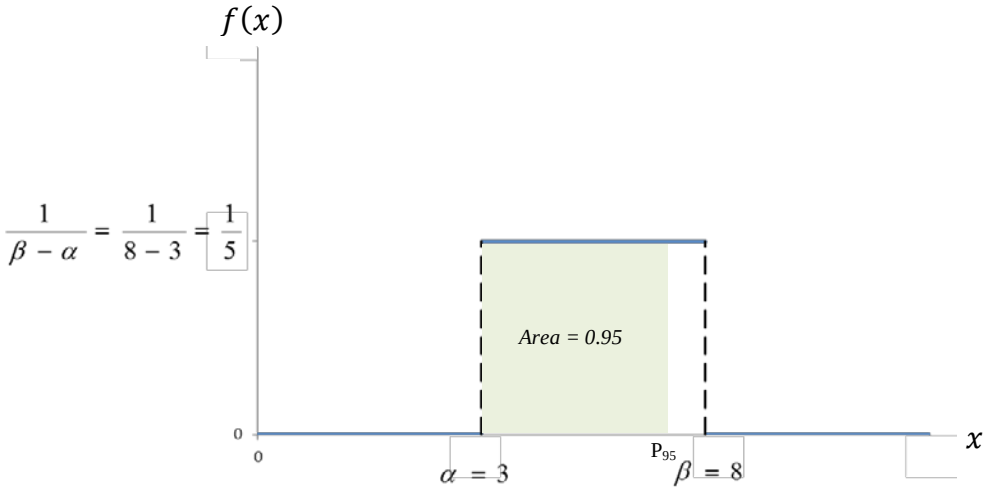
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

b) This is the cumulative area under the density function from $x = 4$ to $x = 6$.



$$P(4 \le X \le 6) = \int_4^6 f(x)dx = \frac{1}{5} * (6 - 4) = 0.4$$

c) The 95th percentile is the response time that has 95% of the distribution less than it. Thus, it is the value that has 95% of the area under the density function to the left of it.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

$$\frac{1}{5} * (P_{95} - 3) = 0.95$$

Solving for P_{95},

$$P_{95} = 7.75\ mins$$

4.4 The Exponential Distribution

The Exponential distribution (also called the Negative Exponential distribution) has the density function of the form,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & for\ x \geq 0 \\ 0 & for\ all\ other\ x \end{cases} \quad \text{..... Equation (4. 12)}$$

where λ is the **distribution parameter** (also called the **rate parameter**)

The cumulative density function is given by,

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x} & for\ x \geq 0 \\ 0 & for\ all\ other\ x \end{cases} \quad \text{..... Equation (4. 13)}$$

The expected value (μ) is given by,

$$\mu = \frac{1}{\lambda} \quad \text{..... Equation (4. 14)}$$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The variance (σ^2) is given by,

$$\sigma^2 = \frac{1}{\lambda^2}$$

. . . . . **Equation (4. 15)**

Equation (4.15) implies that the standard deviation (σ) is equal to the mean. The Exponential distribution is commonly used to describe data related to natural growth or decay type processes. It can be shown mathematically that if a process exhibits Poisson behavior, then the time between the events follows an Exponential distribution. For example, for a traffic flow (number of vehicle arrivals per unit of time) that is Poisson, the vehicular inter-arrival times will follow an Exponential distribution. A graphical depiction of an exponential distribution is as follows.

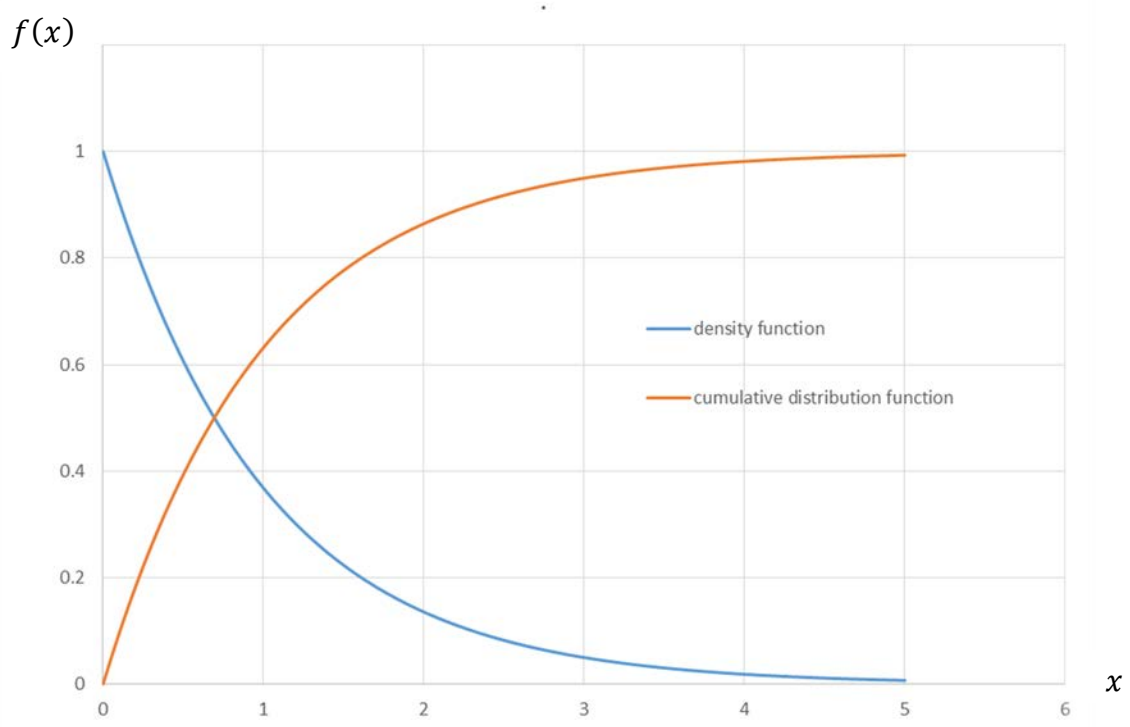


Figure 4. 2: The Exponential distribution

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4.5 The Normal Distribution

The Normal distribution, also called the Gaussian distribution, is a continuous distribution with a symmetric bell-shaped density function.

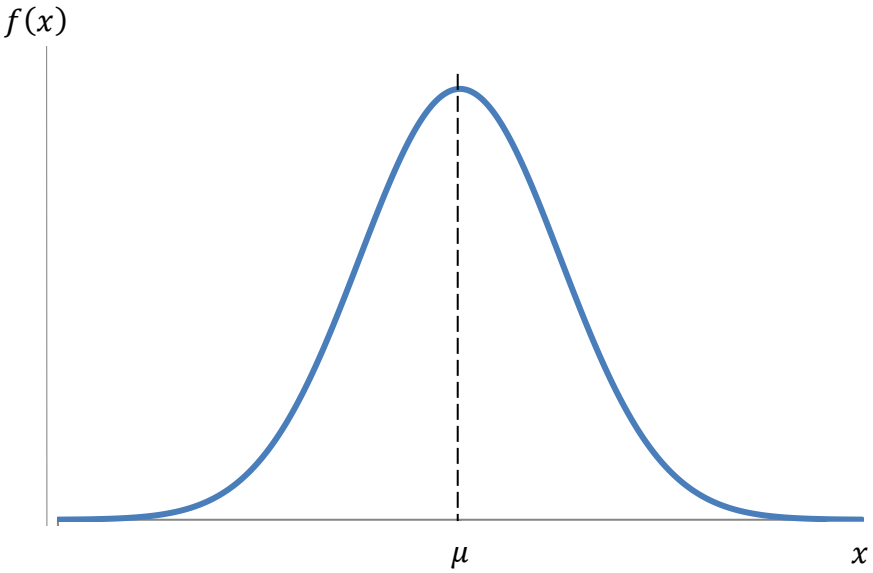


Figure 4. 3: The Normal distribution

The expected value of the distribution corresponds to the apex of the bell curve.

The density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

. . . . . **Equation (4. 16)**

where $-\infty \leq x \leq +\infty$

The expected value (mean) and the standard deviation are always needed to specify a Normal distribution. In practice they are computed from the data. The mean and standard deviation are referred to as the **location parameter** and the **shape parameter** respectively. The location

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

parameter provides the location of the apex of the bell curve, whereas the shape parameter determines the shape or the spread of the distribution. Lower values of the shape parameter will result in a tighter and steeper bell curve, whereas larger values of the shape parameter will result in a flatter and more spread out bell curve.

The cumulative probability (cumulative density) of a value x_i is the area under the bell-curve up to the value x_i.

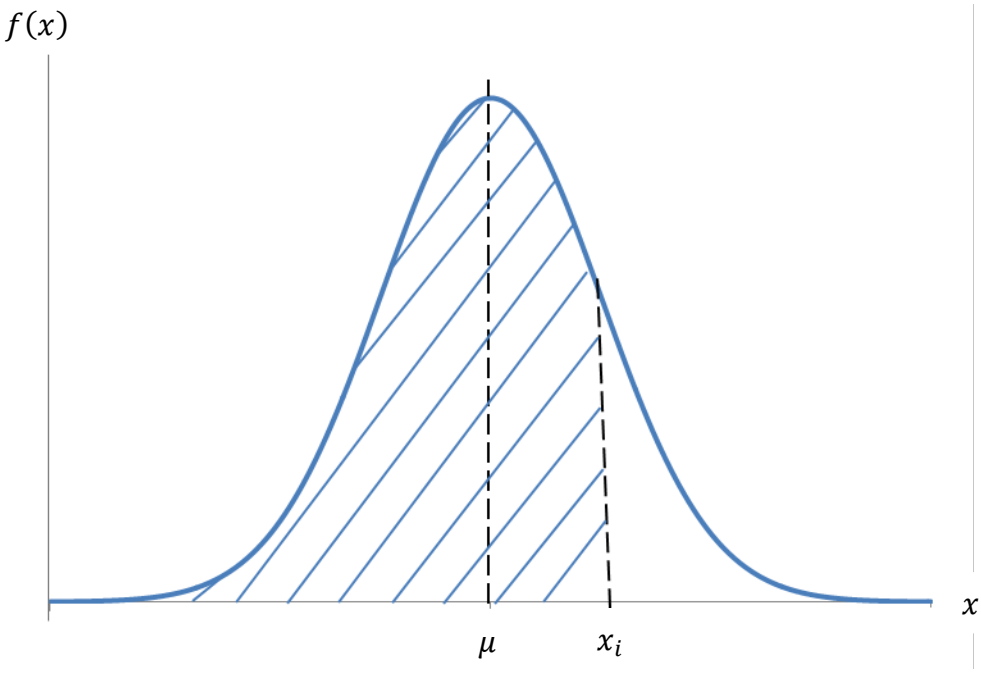


Figure 4. 4: Cumulative density

The cumulative density is given by,

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

. . . . . **Equation (4. 17)**

Equation (4.17) is a difficult integration to perform. It can be simplified by setting μ = 0 and

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

μ = 1.

From Equation (4.16), this results in a density function,

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

. . . . . **Equation (4. 18)**

where $-\infty \leq z \leq +\infty$

Equation (4.18) is the density function of what is called the **Standard Normal distribution**. The integration is easier, but still not that easy. To assist practitioners, the integration has been done by others and tabulated into **standard normal distribution tables**, which are available in all statistics texts, and engineering reference manuals. Standard normal tables are also abundant and available for use, as needed, on the internet. Most modern scientific calculators have statistical capabilities and can produce Normal distribution or Standard Normal distribution values.

In practice, ANY normal random variable (X) is converted to a **standard normal variate** (Z), (commonly called a **z-score**) and the corresponding cumulative density (cumulative probability) value ($F(z)$) is read from a standard normal table (commonly referred to as a **z-table**).

The conversion from normal variate (X) to the standard normal variate (Z) is given by,

$$z = \frac{x - \mu}{\sigma}$$

. . . . . **Equation (4. 19)**

Example:
The spot speed of vehicles monitored at a traffic count station on an expressway indicates the speeds are normally distributed with a mean of 65 mph and a standard deviation of 7 mph.

a) What is the probability that the speed of a vehicle is less than 55 mph?
b) What is the probability that the speed of a vehicle is greater than 80 mph?
c) What is the probability that the speed of a vehicle is between 60 mph and 75 mph?
d) What is the 85th percentile speed?

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

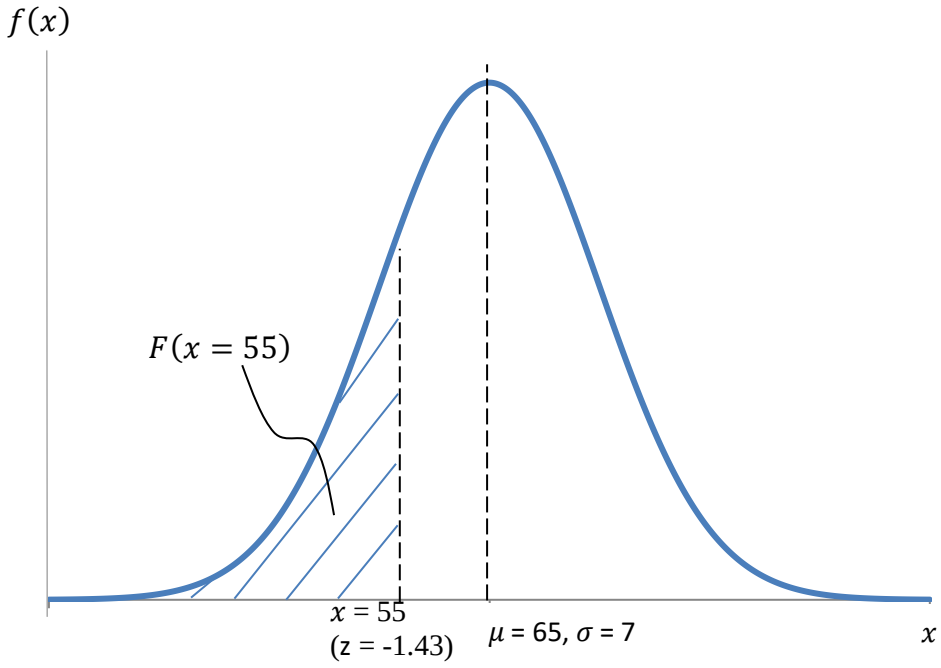
Solution:
a) This is a normal distribution with μ = 65 and σ = 7.
The objective is to compute $P(X \leq 55)$
The $x = 55$ value must be converted to a standard normal variate (z-score). The cumulative probability for the z-score shall then be read from a standard normal distribution table, in lieu of performing integration of Equation (4.17) or Equation (4.18).

From Equation (4.19),

$$z = \frac{x - \mu}{\sigma} = \frac{55 - 65}{7} = -1.43$$

$$P(X \leq 55) = P(Z \leq -1.43)$$

Or, graphically,



Using a Standard Normal distribution table from a random online source,

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Standard Normal Probabilities

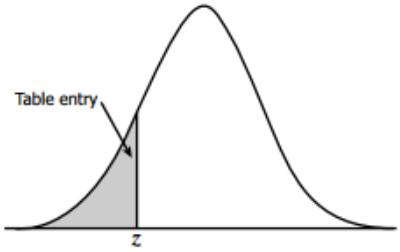


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Source: http://www.stat.ufl.edu/~athienit/Tables/Ztable.pdf

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Thus,
$P(X \leq 55) = P(Z \leq -1.43) = 0.0764 = 7.64\%$

The following shall be considered when using any published standard normal distribution tables:

1. Always look at the diagram associated with the table, and confirm that the area shaded under the curve is consistent with the shaded area of interest for the specific problem being worked on. In this example the area shaded under the curve on the table diagram and the shaded area of interest are consistent and to the left of the z-score.
2. If the area shaded under the curve on the table diagram and the shaded area of interest are not consistent, that is, on opposite sides of the z-score, then the complementation rule of probability shall be applied to obtain the (cumulative) probability value of interest.
3. Most modern texts have separate tables for positive and negative z-scores.
4. When reading a probability value for a z-score, the second decimal place of the z-score is taken by moving across the table.
5. If the z-score involves more than two decimal places, then interpolation may be used to obtain the probability value. However, for most engineering intents and purposes, rounding the z-score to two decimal places is adequate.

b) The objective is to compute $P(X \geq 80)$
The value $x = 80$ shall be converted to a z-score. The probability (cumulative probability) of the z-score may then be looked up on a Standard Normal distribution table, in lieu of performing integration of the Normal distribution density function.

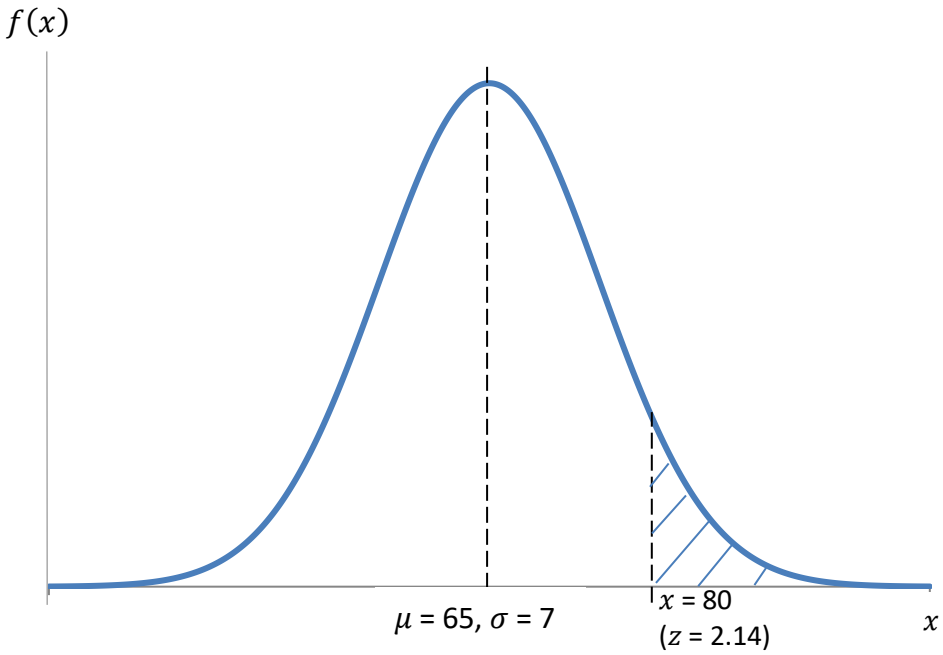
From Equation (4.19),

$$z = \frac{x - \mu}{\sigma} = \frac{80 - 65}{7} = 2.14$$

$P(X \geq 80) = P(Z \geq 2.14)$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Or, graphically,



From the Standard Normal distribution table,

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Standard Normal Probabilities

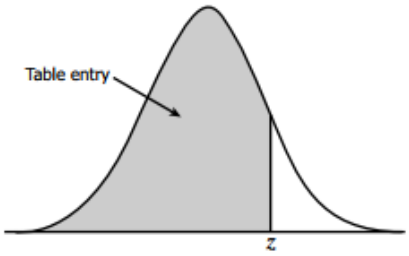


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Source: http://www.stat.ufl.edu/~athienit/Tables/Ztable.pdf

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The area shaded under the curve on the Standard Normal distribution table diagram and the shaded area of interest are on opposite sides of the z-score. The area shaded on the Standard Normal distribution table diagram is to the left of the z-score, whereas the shaded area of interest is to the right of the z-score. Thus, the complementation rule of probability shall be applied to probability value read from the Standard Normal distribution table for this z-score.

$P(X \geq 80) = P(Z \geq 2.14) = 1 - 0.9838 = 0.0162 = 1.62\%$

c) $P(60 \leq X \leq 75)$

From Equation (4.19),

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{60 - 65}{7} = -0.71$$

and

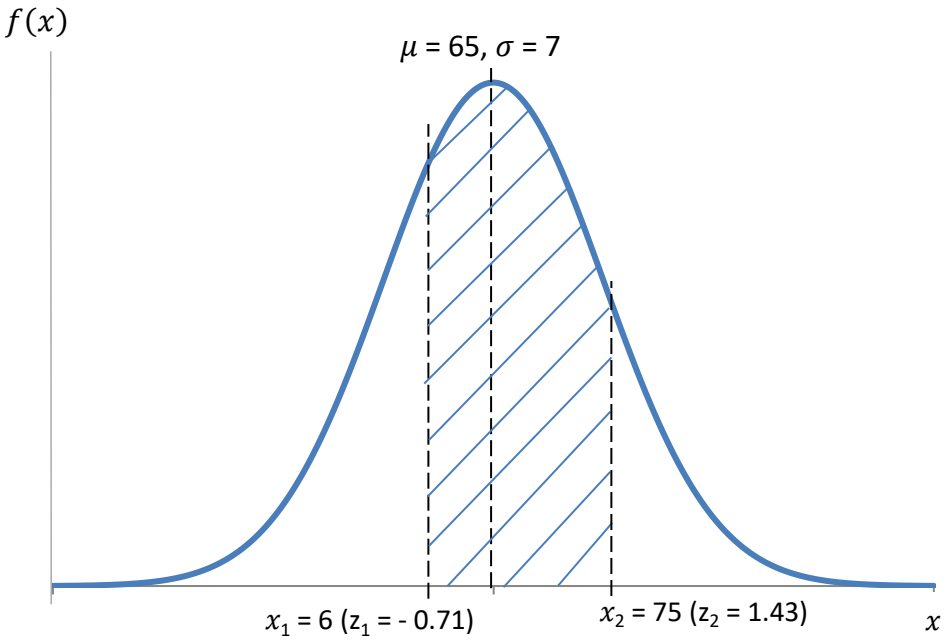
$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{75 - 65}{7} = 1.43$$

So,

$P(60 \leq X \leq 75) = P(-0.71 \leq Z \leq 1.43)$

Graphically, this is the area under the curve between $x_1 = 60$ and $x_2 = 75$ (or $z_1 = -0.71$ and $z_2 = 1.43$).

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE



By looking at the areas under the curve associated with each z-score, it can be deduced that,

$$P(60 \leq X \leq 75) = P(X \leq 75) - P(X \leq 60)$$

Or using the z-scores,

$$P(-0.71 \leq X \leq 1.43) = P(X \leq 1.43) - P(X \leq -0.71)$$

which can be generalized as,

$$P(z_1 \leq Z \leq z_2) = P(Z \leq z_2) - P(Z \leq z_1) \quad \text{. . . . . } \textbf{Equation (4. 20)}$$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

or,

$$P(z_1 \leq Z \leq z_2) = F(z_2) - F(z_1)$$

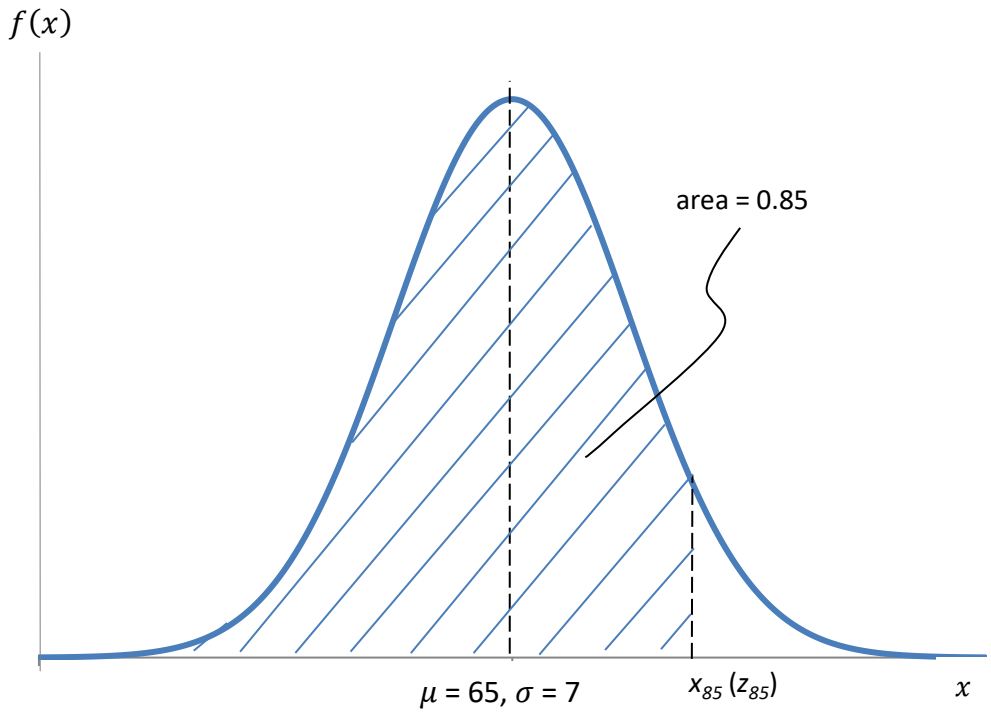
. . . . . **Equation (4. 21)**

From the Standard Normal distribution tables,

$P(-0.71 \leq X \leq 1.43) = P(X \leq 1.43) - P(X \leq -0.71) = 0.9236 - 0.2389$

$P(-0.71 \leq X \leq 1.43) = 0.6847$ or 68.47%

d) By definition, the P^{th} percentile is the x value that has P% of the data less than or equal to it. In other words, the P^{th} percentile will have P% of the area under the density curve to its left.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Scanning through the probability values in the z-tables, the closest value to 0.85 is 0.8508 which corresponds to a z-score of 1.04. Note that interpolation may be applied if an increased level of accuracy is warranted.

Standard Normal Probabilities

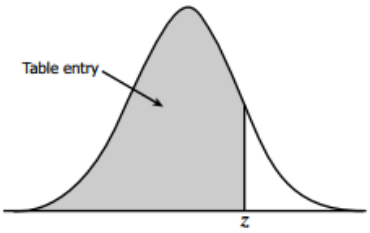


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Source: http://www.stat.ufl.edu/~athienit/Tables/Ztable.pdf

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The z-score must be converted to an x value.

From Equation (4.19),

$$z_{85} = \frac{x_{85} - \mu}{\sigma}$$

$$1.04 = \frac{x_{85} - 65}{7}$$

Solving for x_{85}, the 85th percentile,

$x_{85} = 72.28$ mph

4.6 The Log-Normal Distribution

A random variable X has a Log-Normal distribution if $Y = \ln(X)$ is Normally distributed. A random variable that is Log-Normally distributed takes on only positive real values. There are many formulations of the probability density function of the Log-Normal distribution. One commonly used formulation is,

$$f(x) = \frac{1}{\sqrt{2\pi}\beta}\frac{1}{x}e^{-\frac{[\ln(x)-\alpha]^2}{2\beta^2}}$$

. . . . . **Equation (4. 22)**

for $x > 0, \beta > 0$

where
α is the mean of the $\ln(x)$ values, and
β is the standard deviation of the $\ln(x)$ values.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

It can be shown that the cumulative density function of the Log-Normal distribution is given by,

$$F(x) = P(X \leq x) = \Phi\left(\frac{\ln(x) - \alpha}{\beta}\right)$$

. . . . . **Equation (4. 23)**

for $x > 0, \beta > 0$

where
Φ is the cumulative density function of the Standard Normal distribution.

The cumulative density that the random variable will take on an x value in the range $[a, b]$ is given by,

$$P(a \leq X \leq b) = F(b) - F(a) = \Phi\left(\frac{\ln(b) - \alpha}{\beta}\right) - \Phi\left(\frac{\ln(a) - \alpha}{\beta}\right)$$

. . . . . **Equation (4. 24)**

The inverse function ($G(p)$) of the cumulative density function enables a value of the variate (x) to be computed for a given cumulative probability value (p). Rearranging Equation (4.23), the inverse cumulative density function is given by,

$$x = G(p) = e^{\beta.\Phi^{-1}(p)}$$

. . . . . **Equation (4. 25)**

for $0 \leq p < 1, \beta > 0$

where
Φ^{-1} is the inverse function of the cumulative density function of the Standard Normal distribution.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The inverse function of a cumulative density function is the called the **Percent Point function** (also called the **Quantile function**) of the distribution. The Percent Point function is used extensively in conducting simulations and modeling systems as shall be presented later in this course.

Figure 4.5 shows an example of a probability density function (**pdf**) of a Log-Normal distribution.

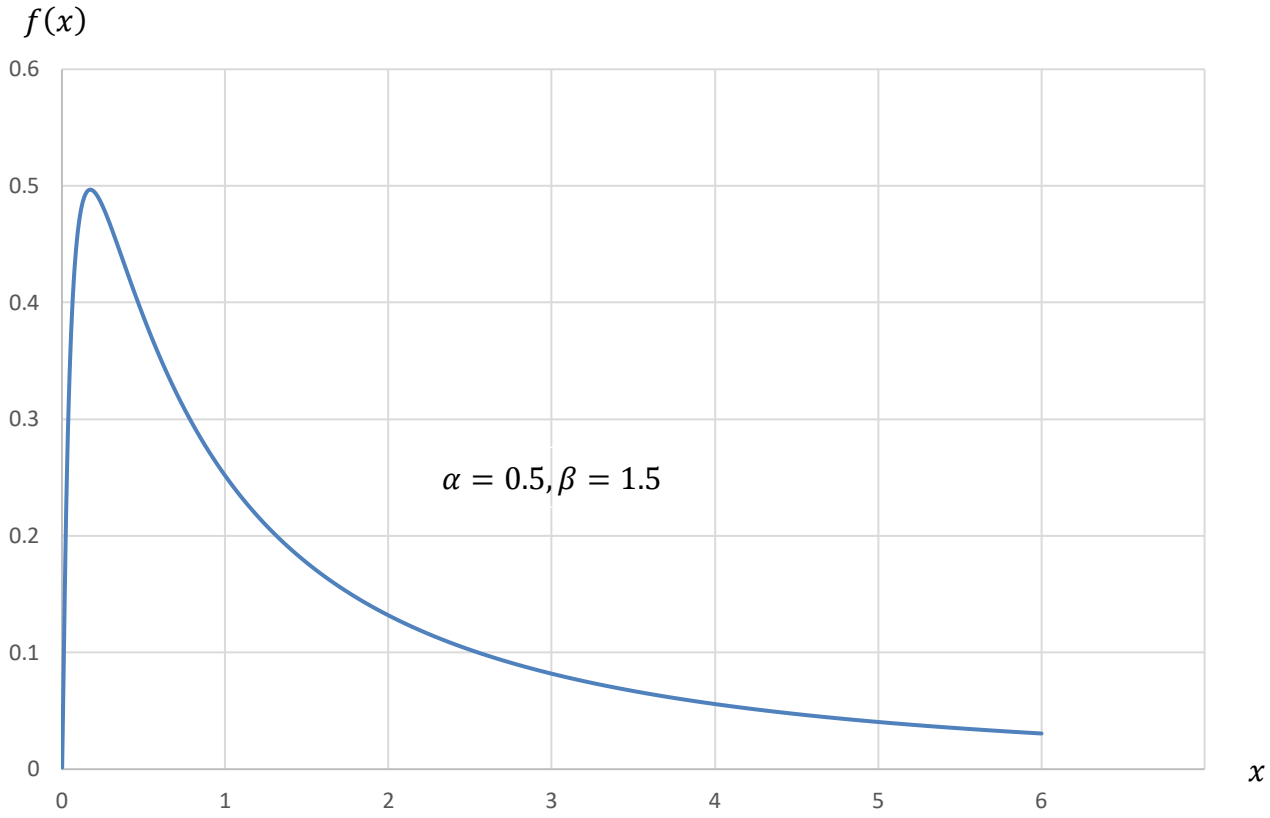


Figure 4. 5: Log-Normal probability density function

The Log-Normal distribution has a long "tail" to the right. The Log-Normal distribution said to be a **skewed distribution** and **skewed to the right** (or **positively skewed**).

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Figure 4.6 shows a cumulative density function (**cdf**) of a Log-Normal distribution.

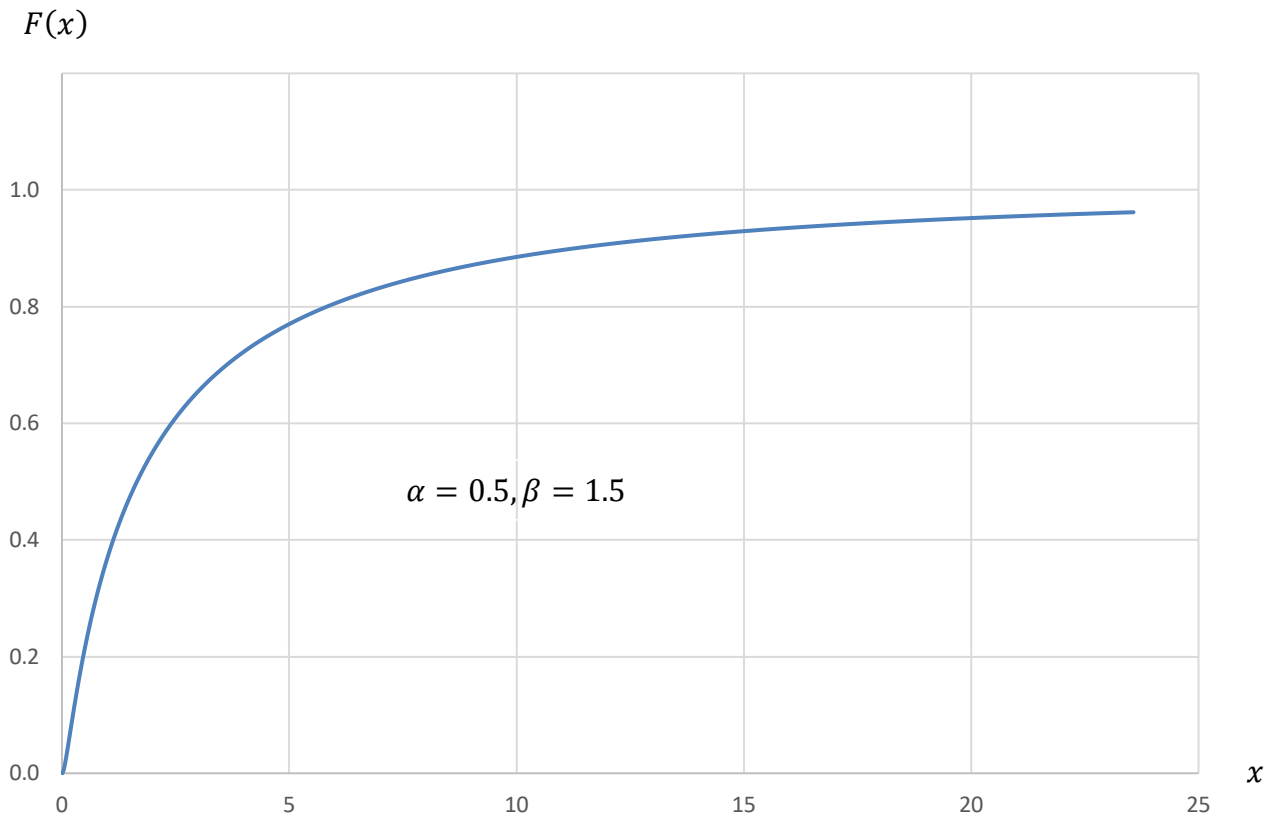


Figure 4. 6: Log-Normal cumulative density function

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Figure 4.7 shows an example of a percent point function.

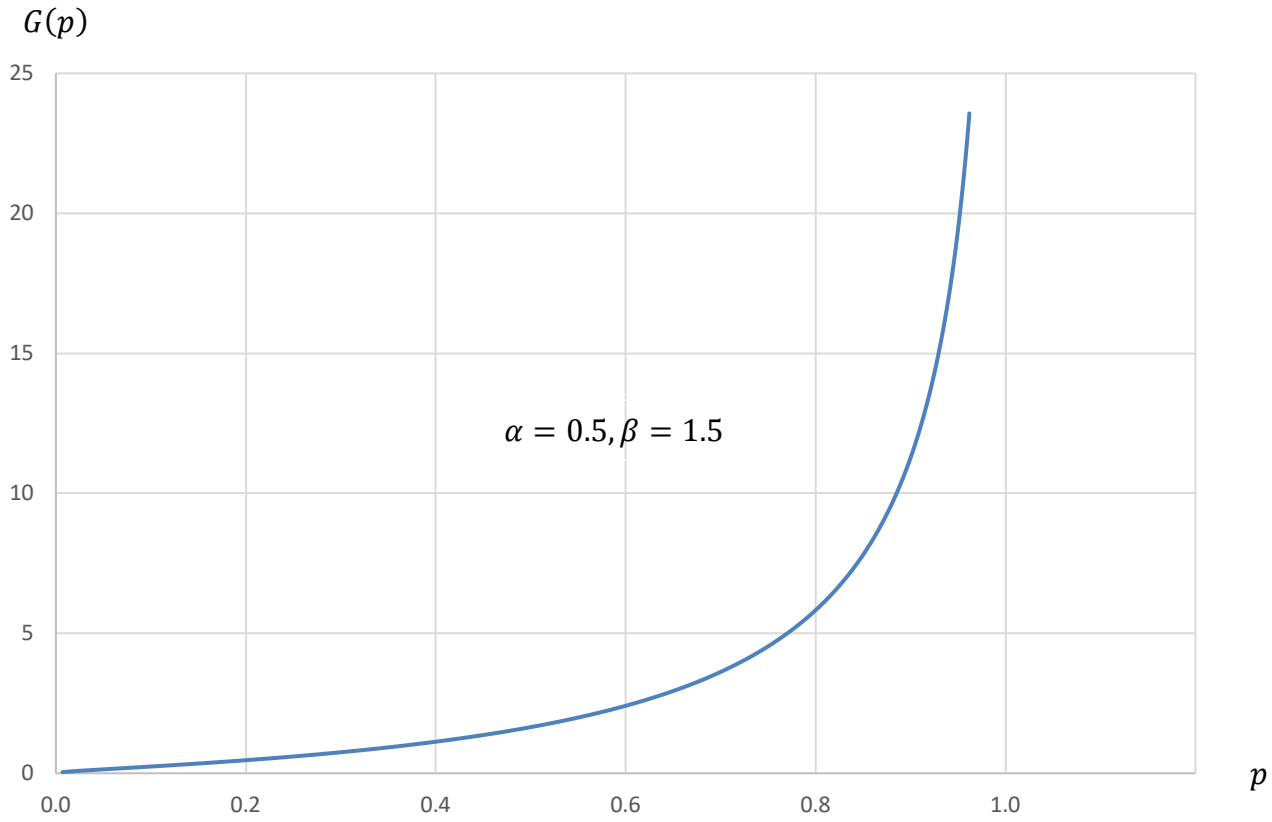


Figure 4. 7: Log-Normal percent point function

The relationship between the mean (μ) of the x values and the mean (α) of the $\ln(x)$ values for a Log-Normal distribution is derived as,

$$\mu = e^{\alpha + \frac{\beta^2}{2}}$$

. . . . . **Equation (4. 26)**

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The relationship between the standard deviation (σ) of the x values and the standard deviation (β) of the $\ln(x)$ values is,

$$\sigma^2 = e^{2\alpha + \beta^2}(e^{\beta^2} - 1)$$

. . . . . **Equation (4. 27)**

The Log-Normal distribution has many applications. It has been used to model inter-arrival times of industrial processes. In reliability engineering, the Log-Normal distribution is used to model the time to failure of machine parts that are subject to modes of failure that are of a fatigue-stress nature. The Log-Normal distribution is applied in establishing warranties of products. It is also used in the fields of finance and risk management.

4.7 The Gamma Distribution

The Gamma distribution has a probability density function given by,

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} . \frac{1}{x^{\alpha-1}} . e^{-\frac{x}{\beta}}$$

. . . . **Equation (4. 28)**

for $x > 0, \alpha > 0,\ \beta > 0$

where,

the expected value (μ) of the distribution is,

$$\mu = \alpha . \beta$$

. . . . . **Equation (4. 29)**

and the variance (σ^2) of the distribution is,

$$\sigma^2 = \alpha . \beta^2$$

. . . . . **Equation (4. 30)**

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The term $\Gamma(\alpha)$ is called the **Gamma function**, and is defined as,

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

. . . . . **Equation (4. 31)**

The cumulative density function of the Gamma distribution works out to,

$$F(x) = \frac{\Gamma_x(\alpha)}{\Gamma(\alpha)}$$

. . . . . **Equation (4. 32)**

where

$$\Gamma_x(\alpha) = \int_0^{x} x^{\alpha-1} e^{-x} dx$$

. . . . . **Equation (4. 33)**

The percent point function of the Gamma distribution does not exist in a closed form. Percent point values for the Gamma distribution can only be approximated by numerical methods.

Figure 4.8 shows examples of probability density functions of Gamma distributions.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

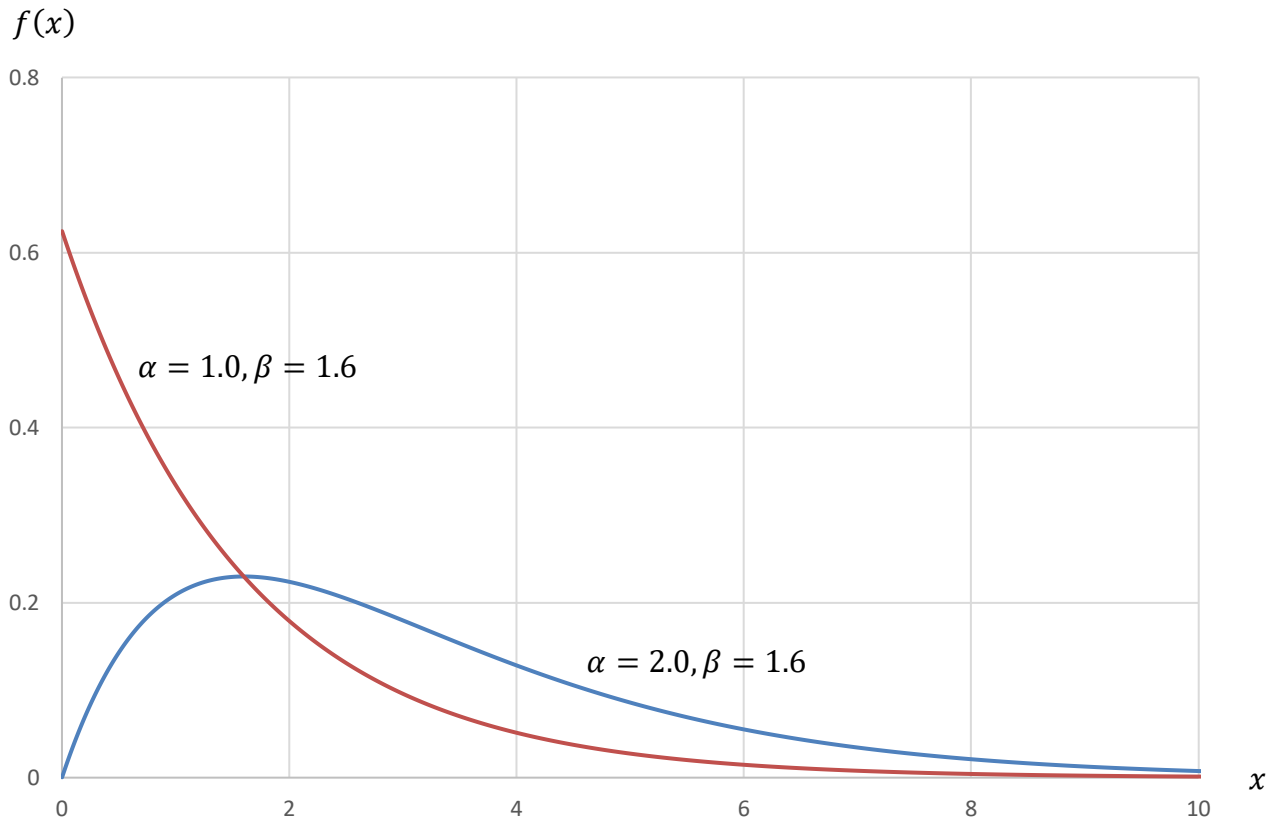


Figure 4. 8: Gamma probability density function

The Gamma distribution is used to model "waiting time" type variables in a variety of fields such as industrial and manufacturing processes, financial services and climate science. The Gamma distribution is in fact not a single distribution but actually a "family" of distributions each with unique and special properties. Each member of the Gamma distribution family is generated by a special combination of the distribution parameters ($\alpha,\ \beta$). For example, α = 1 generates the Exponential distribution covered earlier in this course.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4.8 The Beta Distribution

The Beta distribution is a member of the Gamma distribution "family". The Beta distribution has a probability density function given by,

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)}.x^{\alpha-1}(1-x)^{\beta-1}$$

. . . . . **Equation (4. 34)**

for $0 < x < 1, \alpha > 0,\ \beta > 0$

where,

the expected value (μ) of the distribution is,

$$\mu = \frac{\alpha}{\alpha+\beta}$$

. . . . . **Equation (4. 35)**

and the variance (σ^2) of the distribution is,

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

. . . . . **Equation (4. 36)**

Equation (4.34), incorporates the Beta function,

$$\mathrm{B}(x) = \frac{\Gamma(\alpha).\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

. . . . . **Equation (4. 37)**

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The cumulative density function is given by,

$$F(x) = I(x; a, b)$$

. . . . . **Equation (4. 38)**

where I is the regularized beta function.

$$I(x) = \frac{\mathrm{B}(x; a, b)}{\mathrm{B}(a, b)}$$

. . . . . **Equation (4. 39)**

where $\mathrm{B}(x; a, b)$ is the Incomplete Beta function.

$$\mathrm{B}(x; a, b) = \int_0^x t^{\alpha-1}(1-t)^{b-1}dt$$

. . . . . **Equation (4. 40)**

A closed form of the percent point function does not exist. Thus, percent point values for Beta distributions must be evaluated by numerical methods.

Figure 4.9 shows an example of a probability density function of a Beta distribution.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

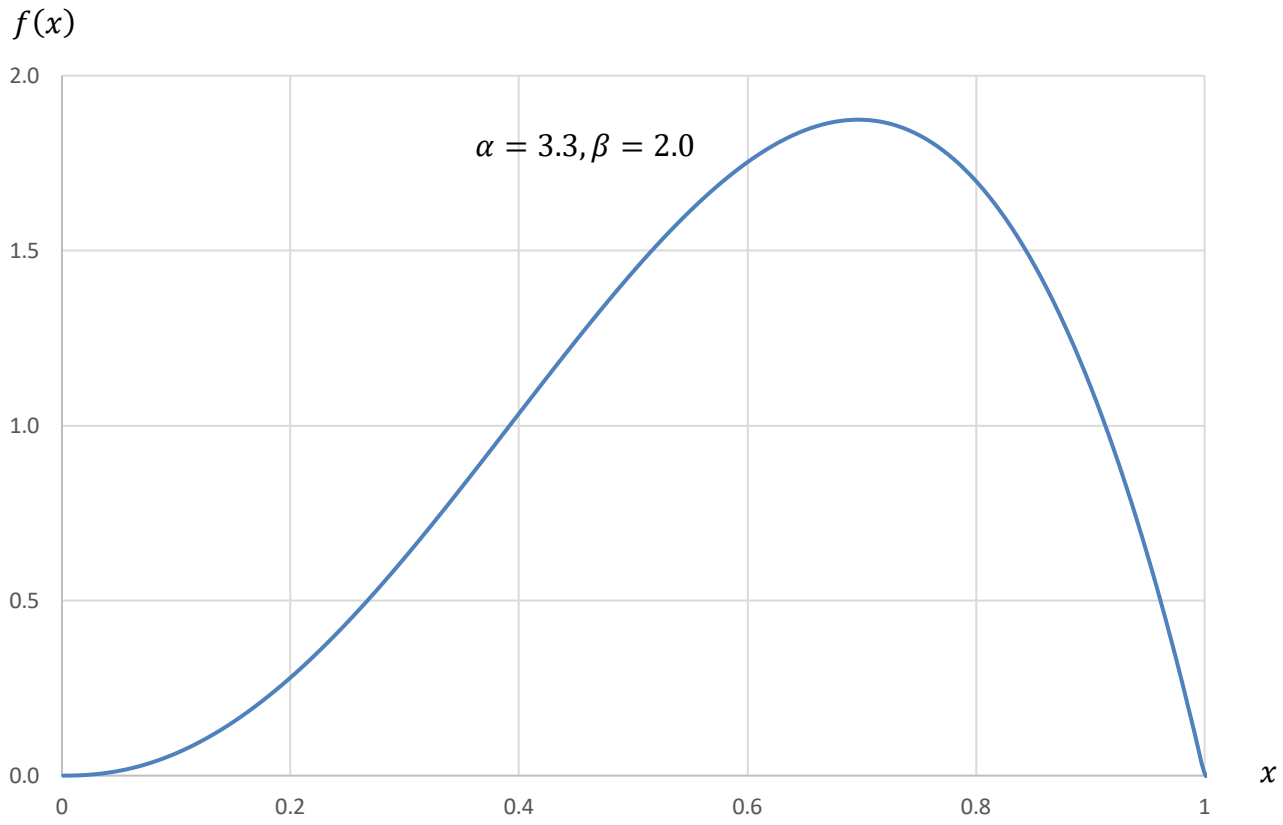


Figure 4. 9: Beta probability density function

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

From Equation (4.34), the x values the Beta distribution can actualize are limited to the range $(0, 1)$, in other words proportions, percentages, ratios, etc. Examples of the application of the Beta distribution include computing probabilities associated with the proportions of rear-end collisions at an intersection over a given period, the percentage of the annual maintenance budget spent on equipment XYZ, the proportion of product that fails QA/QC checks per production cycle etc., etc.

A special case of the Beta distribution occurs if $\alpha = 1$ and $\beta = 1$. The distribution reduces to a Uniform distribution.

4.9 The Weibull Distribution

The Weibull distribution is a member of the Gamma distribution "family". The Weibull distribution has a probability density function given by,

$$\Gamma_x(\alpha) = \int_0^x \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}} dx$$

. . . . . **Equation (4. 41)**

for $x > 0, \alpha > 0,\ \beta > 0$

where,

the expected value (μ) of the distribution is

$$\mu = \alpha^{-\frac{1}{\beta}} . \Gamma\left(1 + \frac{1}{\beta}\right)$$

. . . . . **Equation (4. 42)**

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

and the variance (σ^2) of the distribution is,

$$\sigma^2 = \alpha^{-\frac{2}{\beta}}\left\{\Gamma\left(1+\frac{2}{\beta}\right) - \left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2\right\}$$

. . . . . **Equation (4. 43)**

The cumulative density function works out to,

$$F(x) = 1 - e^{-\alpha x^{\beta}}$$

. . . . **Equation (4. 44)**

for $x \geq 0, \alpha > 0,\ \beta > 0$

Thus, the percent point function is given by,

$$x = G(p) = \left\{-\ln\,(1-p)^{\frac{1}{\alpha}}\right\}^{\frac{1}{\beta}}$$

. . . . **Equation (4. 45)**

for $0 \leq p < 1, \beta > 0$

Figure 4.10 shows examples of probability density functions of Weibull distributions.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

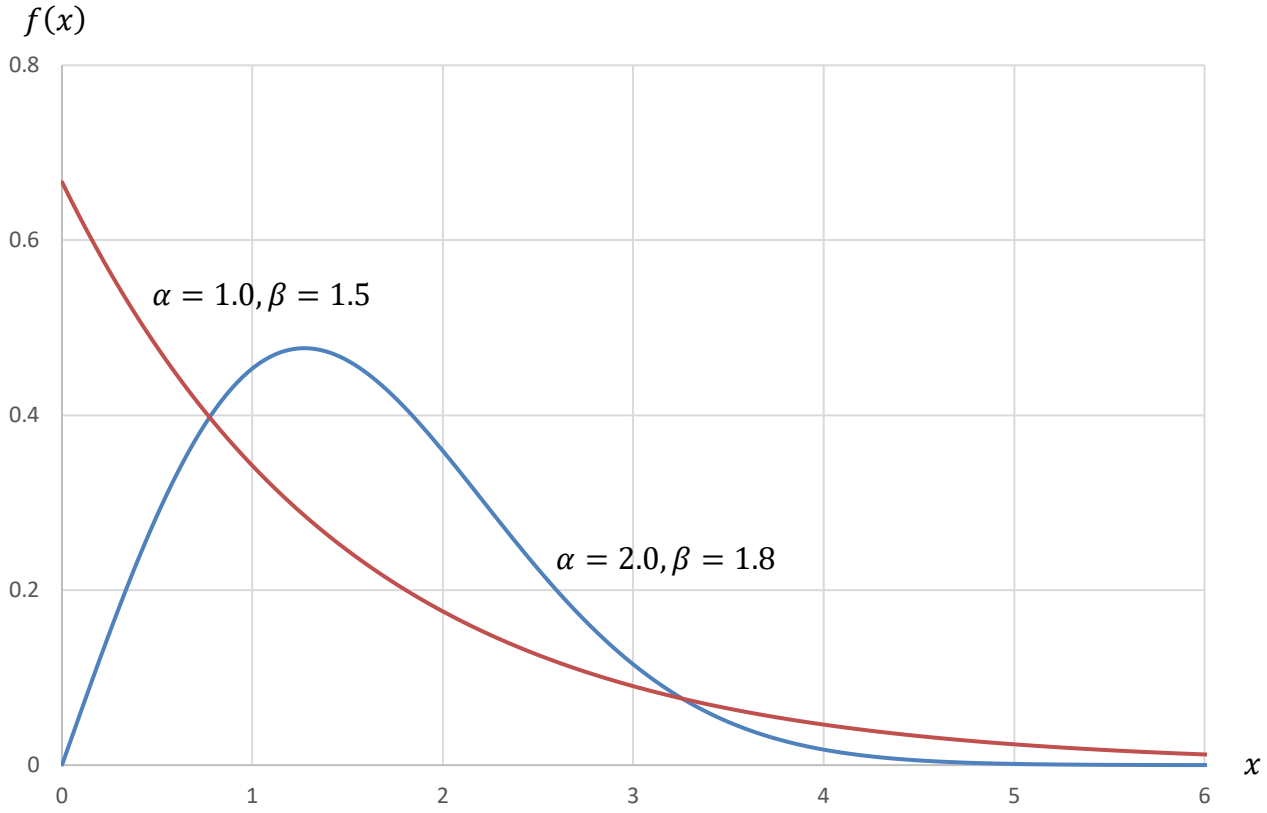


Figure 4. 10: Weibull probability density function

The Weibull distribution is used extensively in reliability engineering to model lifetimes or failure times of manufactured components.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4.10 The Logistic Distribution

The Logistic distribution is a continuous random variable with a probability density function given by,

$$f(x) = \frac{e^{\frac{x-\mu}{\sigma}}}{\sigma\left(1 + e^{\frac{x-\mu}{\sigma}}\right)^2} \quad \text{. . . . . Equation (4. 46)}$$

where $-\infty \leq x \leq +\infty$ and $\sigma > 0$

The cumulative density function of the Logistic distribution works out to,

$$F(x) = \frac{e^{\frac{x-\mu}{\sigma}}}{1 + e^{\frac{x-\mu}{\sigma}}} \quad \text{. . . . . Equation (4. 47)}$$

where $-\infty \leq x \leq +\infty$ and $\sigma > 0$

The percent point function of the Logistic distribution works out to

$$x = G(p) = \mu + \left[ln\left(\frac{p}{1-p}\right)\right]\sigma \quad \text{. . . .. Equation (4. 48)}$$

for $0 \leq p < 1$

Figure 4.11 shows an example of a probability density function of a Logistic distribution.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

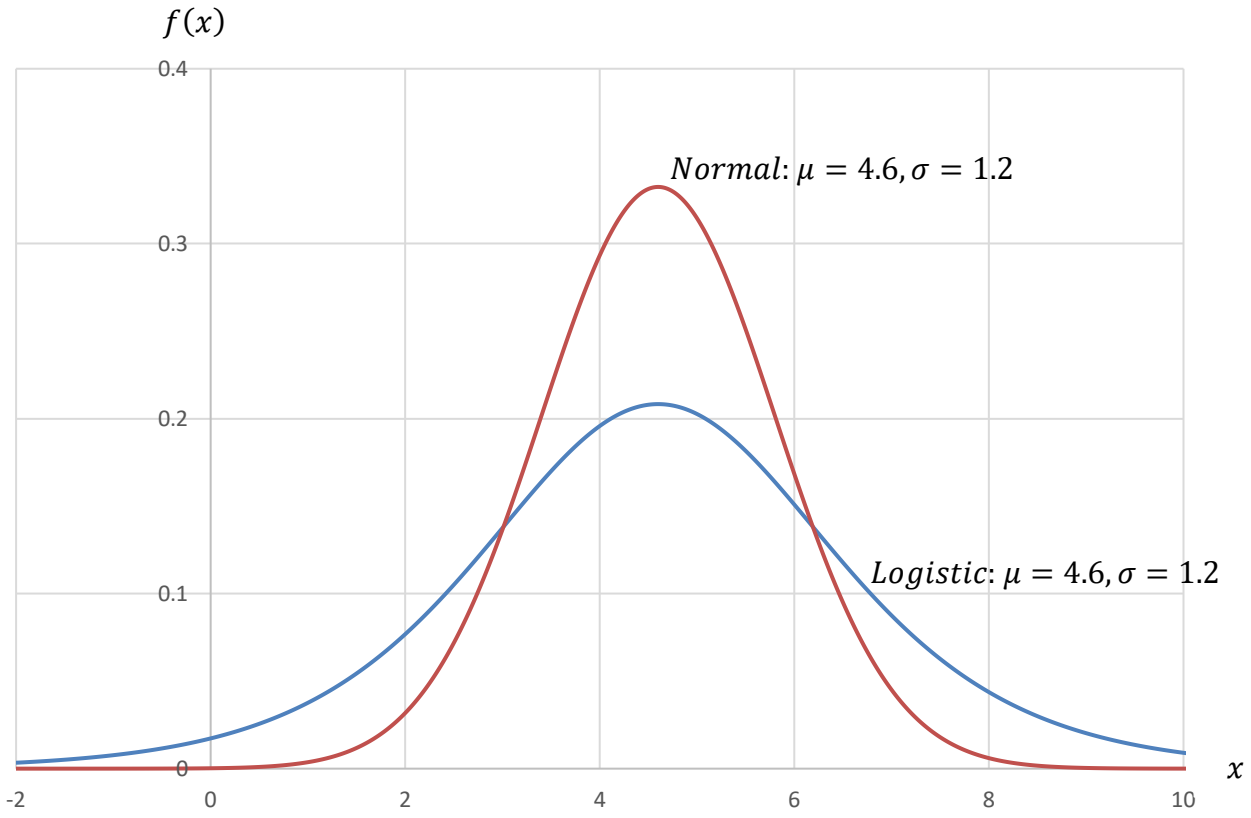


Figure 4. 11: Logistic probability density function

It can be seen that the shape of the Logistic distribution resembles that of the Normal distribution, but has heavier tails. Also, the cumulative density function of the Logistic distribution, Equation (4.47), is simpler and easier to work with, compared to the Normal distribution. The Logistic distribution functions can be solved easily, manually. As a result, in the "good old days" when most calculations were done by hand, the Logistic distribution was often used in place of the Normal distribution. In the current era of computers and software, the Logistic distribution is still preferred for some purposes.

The Logistic distribution is still useful in modeling growth related processes and phenomena and is applied in many fields of economics, finance, public health, and agriculture science.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4.11 The Log-Logistic Distribution

The relationship between the Logistic distribution and the Log-Logistic distribution is analogous to that of the Normal distribution and the Log-Normal distribution. The probability density function of the Log-Logistic distribution is given by,

$$f(x) = \frac{e^{\frac{\ln(x)-\alpha}{\beta}}}{x\beta\left(1+e^{\frac{\ln(x)-\alpha}{\beta}}\right)^2} \quad \text{. . . . . Equation (4. 49)}$$

where $x > 0$ and $\beta > 0$, and
α is the mean of the $\ln(x)$ values, and
β is the standard deviation of the $\ln(x)$ values.

The cumulative density function of the Log-Logistic distribution works out to,

$$F(x) = \frac{e^{\frac{\ln(x)-\alpha}{\beta}}}{1+e^{\frac{\ln(x)-\alpha}{\beta}}} \quad \text{. . . . . Equation (4. 50)}$$

The percent point function of the Log-Logistic distribution works out to

$$x = G(p) = e^{\alpha+\left[ln\left(\frac{p}{1-p}\right)\right]\beta} \quad \text{. . . .. Equation (4. 51)}$$

for $0 \leq p < 1$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Figure 4.12 shows an example of a probability density function of a Log-Logistic distribution.

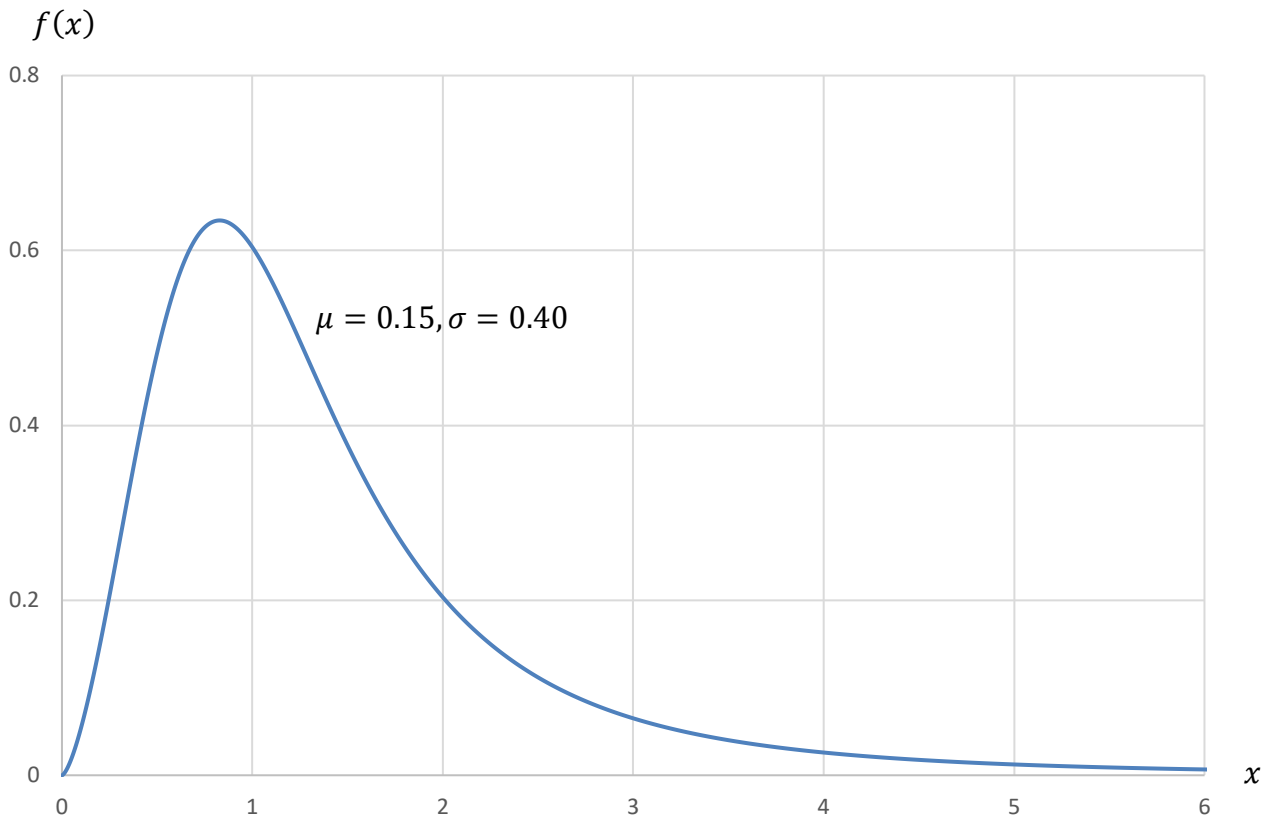


Figure 4. 12: Log-Logistic probability density function

The Log-Logistic distribution has been used to model the lifetime of organisms as well as service times in industrial processes.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

4.12 The Triangular Distribution

The Triangular distribution is a triangular shaped continuous distribution where the actualized values (x) are over a range $[a, b]$, with a mode of c. The parameters of the Triangular distribution are such that,

$-\infty \leq x \leq +\infty$
$a < b$, and
$a \leq c \leq b$

A graphical example is as follows,

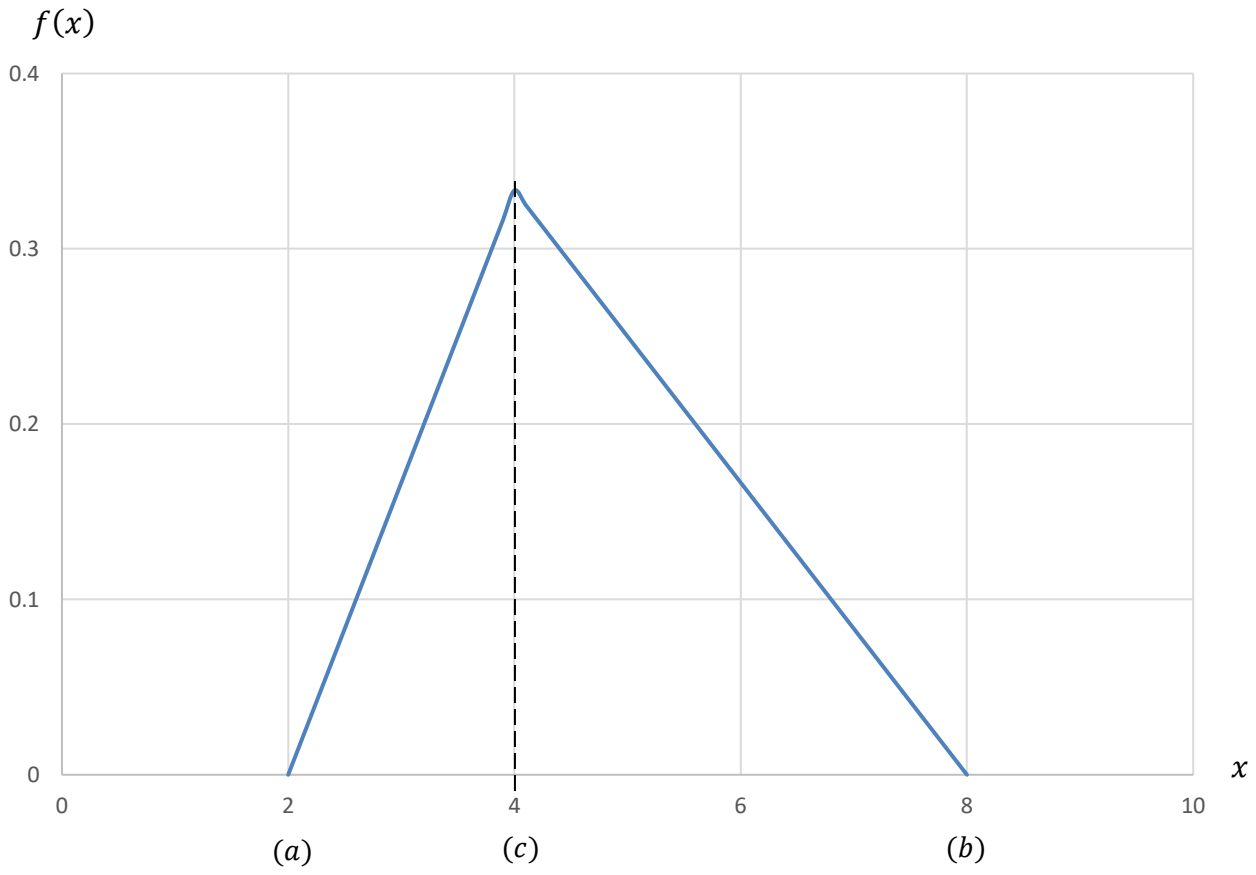


Figure 4. 13: Triangular probability density function

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Mathematically, the probability density function is given by,

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & for\ a \le x < c \\ \frac{2}{(b-a)} & for\ x = c \\ \frac{2(b-x)}{(b-a)(b-c)} & for\ c < x \le b \\ 0 & for\ all\ other\ x \end{cases}$$

. . . . **Equation (4. 52)**

The cumulative density function works out to,

$$F(x) = \begin{cases} \frac{(x-a)^2}{(b-a)(c-a)} & for\ x \le c \\ 1 - \frac{(b-x)^2}{(b-a)(b-c)} & for\ x > c \end{cases}$$

. . . . **Equation (4. 53)**

The percent point function works out to,

$$x = G(p) = \begin{cases} a + \sqrt{p(b-a)(c-a)} & for\ p \le \frac{(c-a)}{(b-a)} \\ b - \sqrt{(1-p)(b-a)(b-c)} & for\ p > \frac{(c-a)}{(b-a)} \end{cases}$$

. . . . . **Equation (4. 54)**

where $0 < p < 1$

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The expected value (μ) of the distribution is

$$\mu = \frac{a + b + c}{3}$$

. . . . . **Equation (4. 55)**

The variance (σ^2) of the distribution is,

$$\sigma^2 = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18}$$

. . . . **Equation (4. 56)**

The Triangular distribution is generally used when there is insufficient data to justify using some other distribution. The values of the parameters used are typically subjective. In other words, a Triangular distribution is an educated guess of the correct description of the random variable. Nevertheless, the triangular distribution is used in simulations related to decision-making methodologies in business, finance and project management.

4.13 Other Continuous Distributions

The following is a non-exhaustive list of other continuous distributions that are in use in various fields of specialty. Students are strongly encouraged to review the statistics and probability literature and study the properties of these distributions and their utility for various types of analyses, such as simulation and systems modeling.

- The Alpha Distribution
- The Anglit Distribution
- The Arcsin Distribution
- The Asymmetric Double Exponential Distribution
- The Asymmetric Log Double Exponential Distribution
- The Bessel I-Function Distribution
- The Bessel K-Function Distribution
- The Beta Normal Distribution
- The Bi-Weibull Distribution

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

- The Bradford Distribution
- The Brittle Fracture Distribution
- The Burr Type 2 Distribution
- The Burr Type 3 Distribution
- The Burr Type 4 Distribution
- The Burr Type 5 Distribution
- The Burr Type 6 Distribution
- The Burr Type 7 Distribution
- The Burr Type 8 Distribution
- The Burr Type 9 Distribution
- The Burr Type 10 Distribution
- The Burr Type 11 Distribution
- The Burr Type 12 Distribution
- The Cauchy Distribution
- The Chi Distribution
- The Chi-Squared Distribution
- The Cosine Distribution
- The Double Exponential Distribution
- The Double Gamma Distribution
- The Double Weibull Distribution
- The Doubly Non-Central Beta Distribution
- The Doubly Non-Central F Distribution
- The Doubly Non-Central T Distribution
- The Doubly Pareto Uniform Distribution
- The Error (Subbotin) Distribution
- The Exponential Power Distribution
- The Exponentiated Weibull Distribution
- The Extreme Value Type 1 (Gumbel) Distribution
- The Extreme Value Type 2 (Frechet) Distribution
- The F Distribution
- The Fatigue Life Distribution
- The Folded Cauchy Distribution
- The Folded Normal Distribution
- The Folded T Distribution
- The G-And-H Distribution

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

- The Generalized Asymmetric Laplace Distribution
- The Generalized Exponential Distribution
- The Generalized Extreme Value Distribution
- The Generalized Gamma Distribution
- The Generalized Half Logistic Distribution
- The Generalized Logistic Type 2 Distribution
- The Generalized Logistic Type 3 Distribution
- The Generalized Logistic Type 4 Distribution
- The Generalized Logistic Type 5 Distribution
- The Generalized Logistic Distribution
- The Generalized McLeish Distribution
- The Generalized Pareto Distribution
- The Generalized Topp And Leone Distribution
- The Generalized Tukey Lambda Distribution
- The Generalized Inverse Gaussian Distribution
- The Generalized Trapezoid Distribution
- The Geometric Extreme Exponential Distribution
- The Gompertz Distribution
- The Gompertz-Makeham Distribution
- The Gompertz-Makeham Distribution
- The Half-Cauchy Distribution
- The Half-Normal Distribution
- The Hyperbolic Secant Distribution
- The Inverse Gaussian Distribution
- The Inverted Beta Distribution
- The Inverted Gamma Distribution
- The Inverted Weibull Distribution
- The Johnson Sb Distribution
- The Johnson Su Distribution
- The Kappa Distribution
- The Kumaraswamy Distribution
- The Landau Distribution
- The Log Beta Distribution
- The Log Double Exponential Distribution
- The Log Gamma Distribution

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

- The Logistic-Exponential Distribution
- The Log-Skew-Normal Distribution
- The Log-Skew-T Distribution
- The Maxwell Distribution
- The McLeish Distribution
- The Mielke's Beta-Kappa Distribution
- The Muth Distribution
- The Non-Central Beta Distribution
- The Non-Central Chi-Square Distribution
- The Non-Central F Distribution
- The Non-Central T Distribution
- The Normal Mixture Distribution
- The Ogive Distribution
- The Pareto Second Kind Distribution
- The Pareto Distribution
- The Pearson Type 3 Distribution
- The Power Function Distribution
- The Power Lognormal Distribution
- The Power Normal Distribution
- The Rayleigh Distribution
- The Reciprocal Inverse Gaussian Distribution
- The Reciprocal Distribution
- The Reflected Generalized Topp Leone Distribution
- The Reflected Power Distribution
- The Semi-Circular Distribution
- The Skew Double Exponential Distribution
- The Skew Normal Distribution
- The Skew T Distribution
- The Slash Distribution
- The Slope Distribution
- The T Distribution
- The Topp and Leone Distribution
- The Trapezoid Distribution
- The Truncated Exponential Distribution
- The Truncated Normal Distribution

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

- The Truncated Pareto Distribution
- The Tukey Lambda Distribution
- The Two-Sided Ogive Distribution
- The Two-Sided Power Distribution
- The Two-Sided Slope Distribution
- The Uneven Two-Sided Power Distribution
- The Von Mises Distribution
- The Wakeby Distribution
- The Wald Distribution
- The Wrapped Cauchy Distribution

4.14 Continuous Random Variables in *Excel*

In this section the continuous random variables worked in Section 4.5 shall be reworked using *Microsoft Excel*.

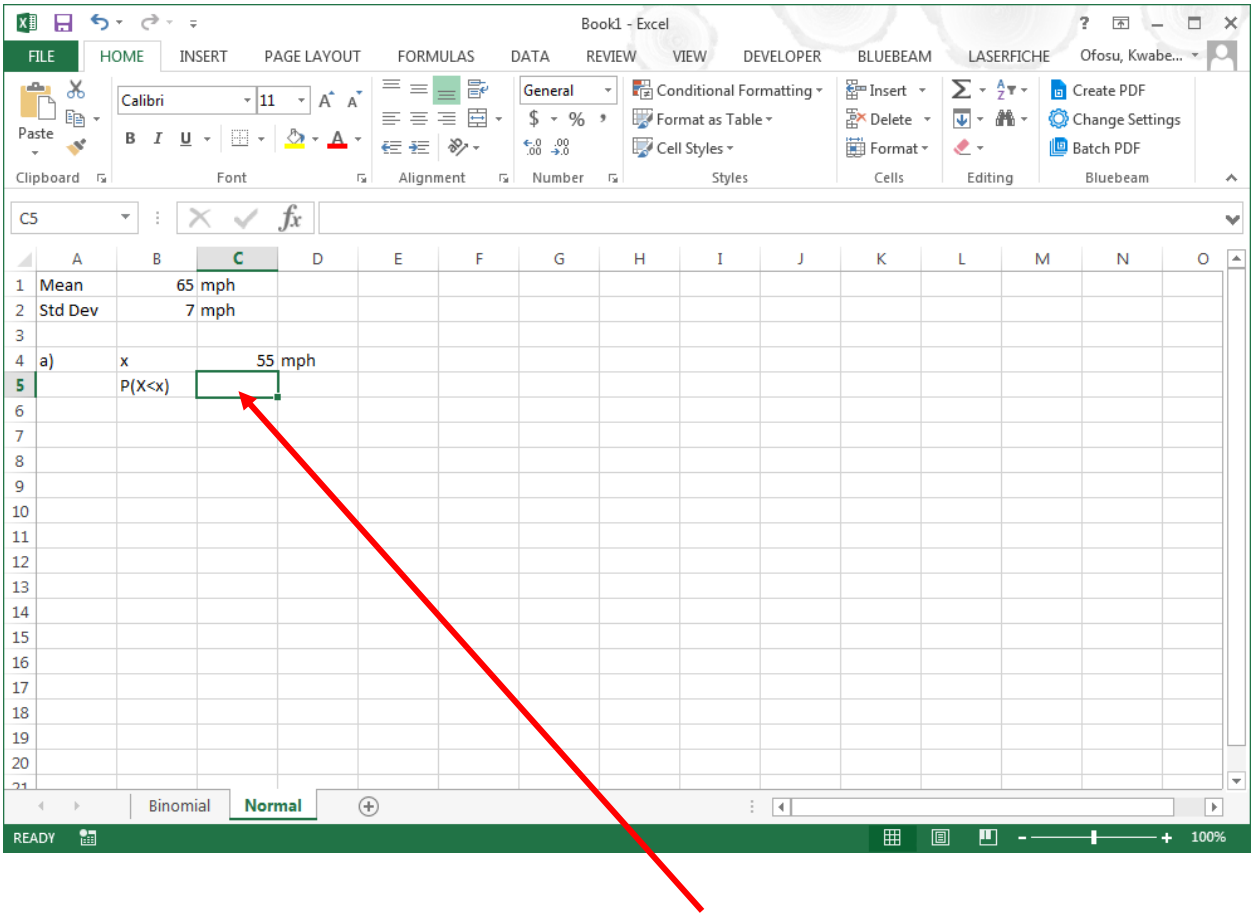
Revisiting the Normal Distribution Example:
The spot speed of vehicles monitored at a traffic count station on an expressway indicates the speeds are normally distributed with a mean of 65 mph and a standard deviation of 7 mph.
a) What is the probability that the speed of a vehicle is less than 55 mph?
b) What is the probability that the speed of a vehicle is greater than 80 mph?
c) What is the probability that the speed of a vehicles is between 60 mph and 75 mph?
d) What is the 85^{th} percentile speed?

Solution:
a) We have with μ = 65 and σ = 7.
The objective is to compute $P(X \leq 55)$

Open a new session of *Excel*.

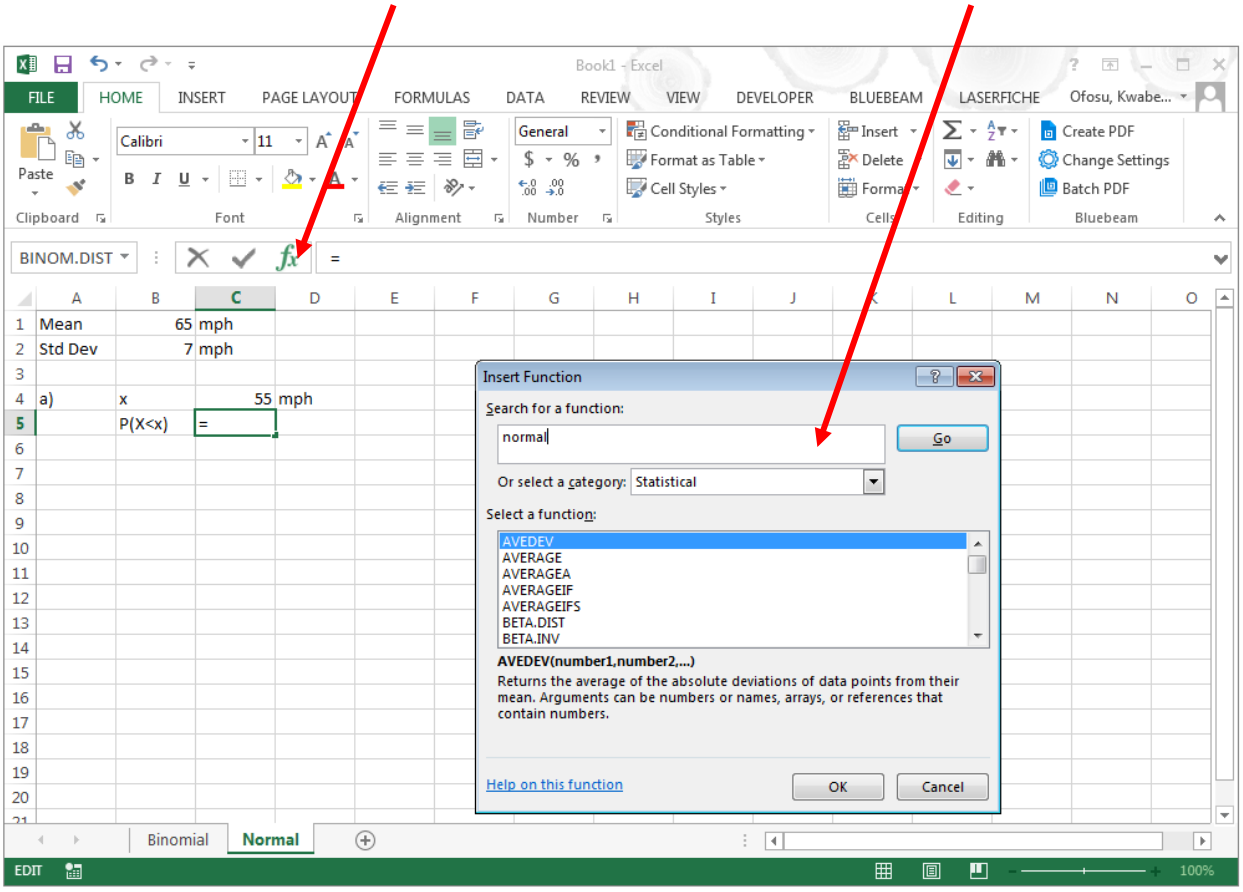
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Enter the data as follows:



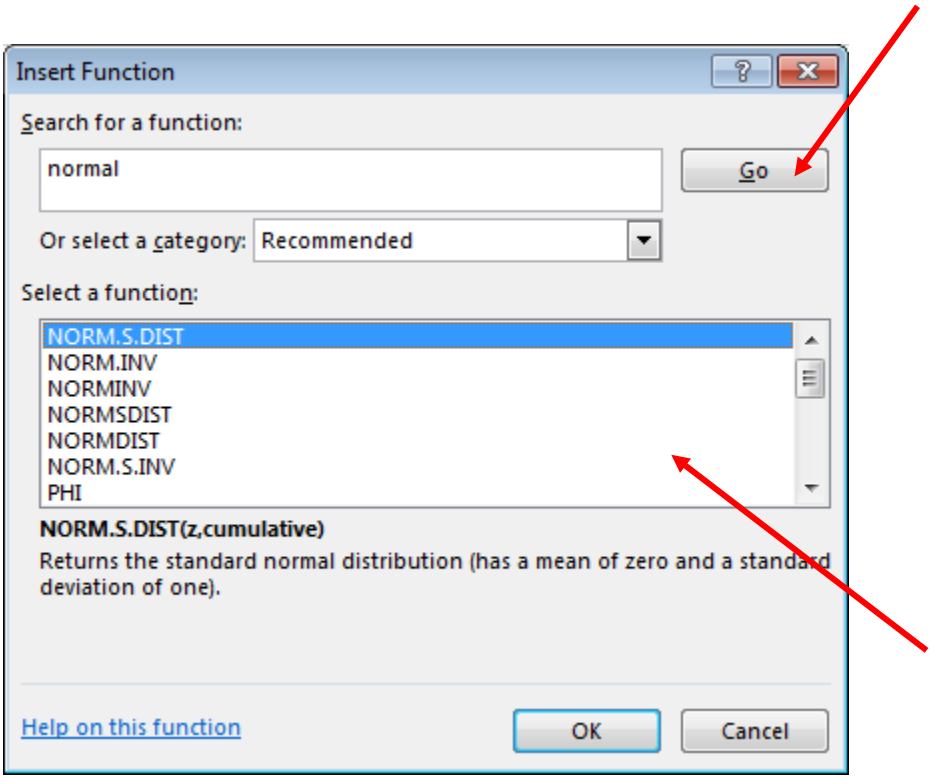
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Click on **Insert Function**.
Under **Or select a category**, select **Statistical**.
Under **Search for a function**, type “Normal”.



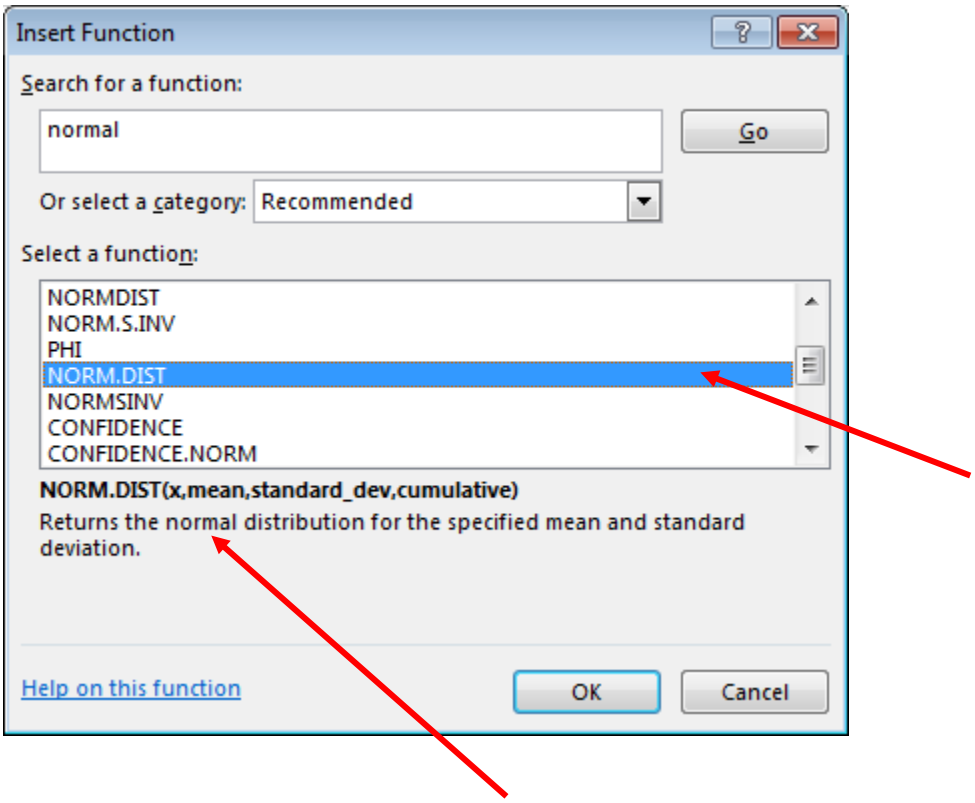
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Hit **Go**.
The list narrows down.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

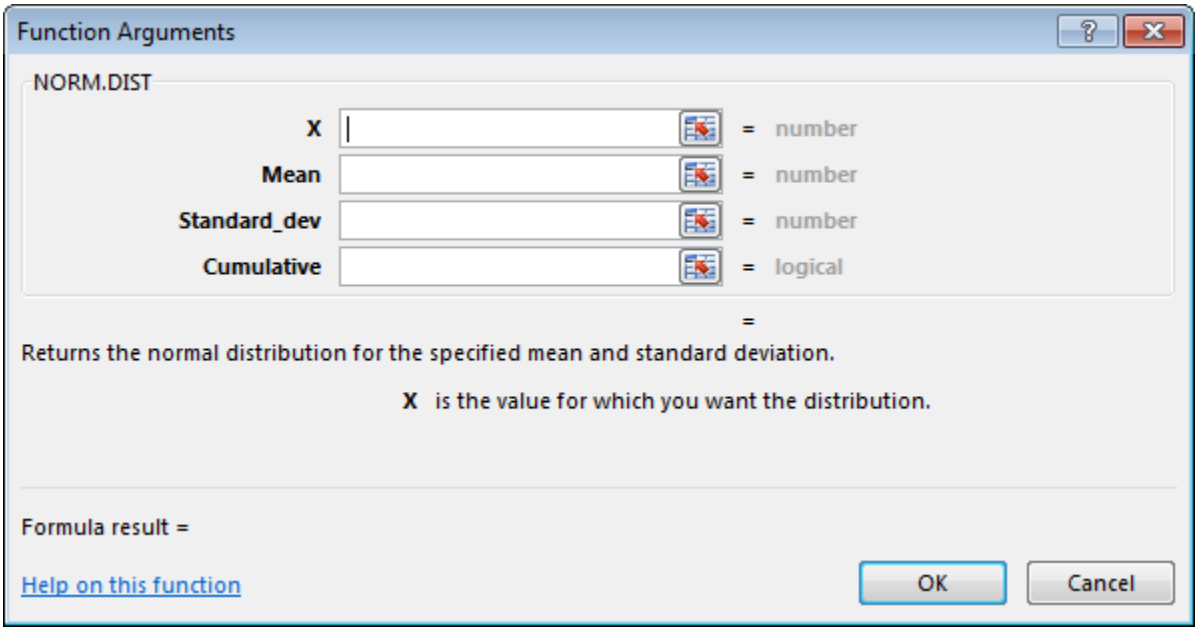
Scroll to NORM.DIST
Select NORM.DIST.
Review the description.
For this function, there is no need to convert an x value to a z-score, the function handles the conversion.



Hit **OK**.

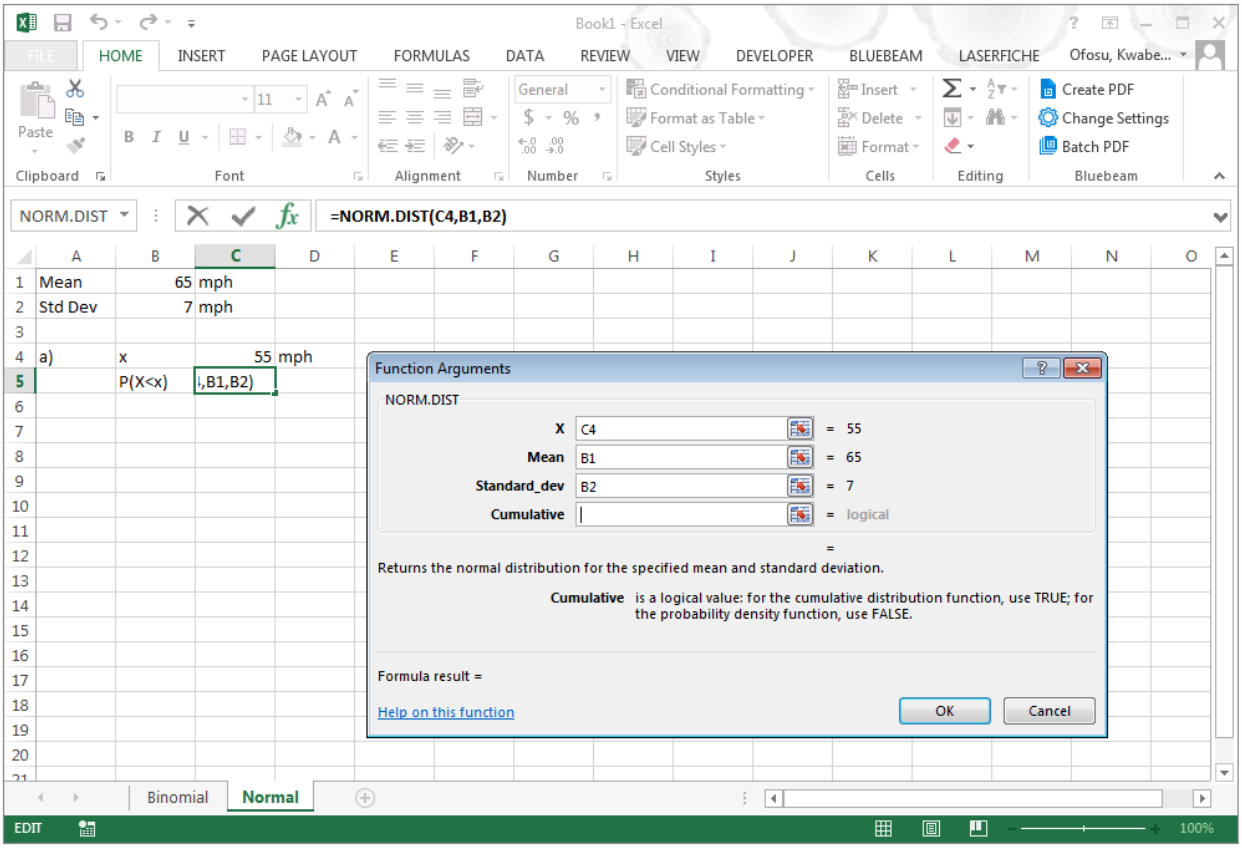
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The **Function Arguments** window opens.



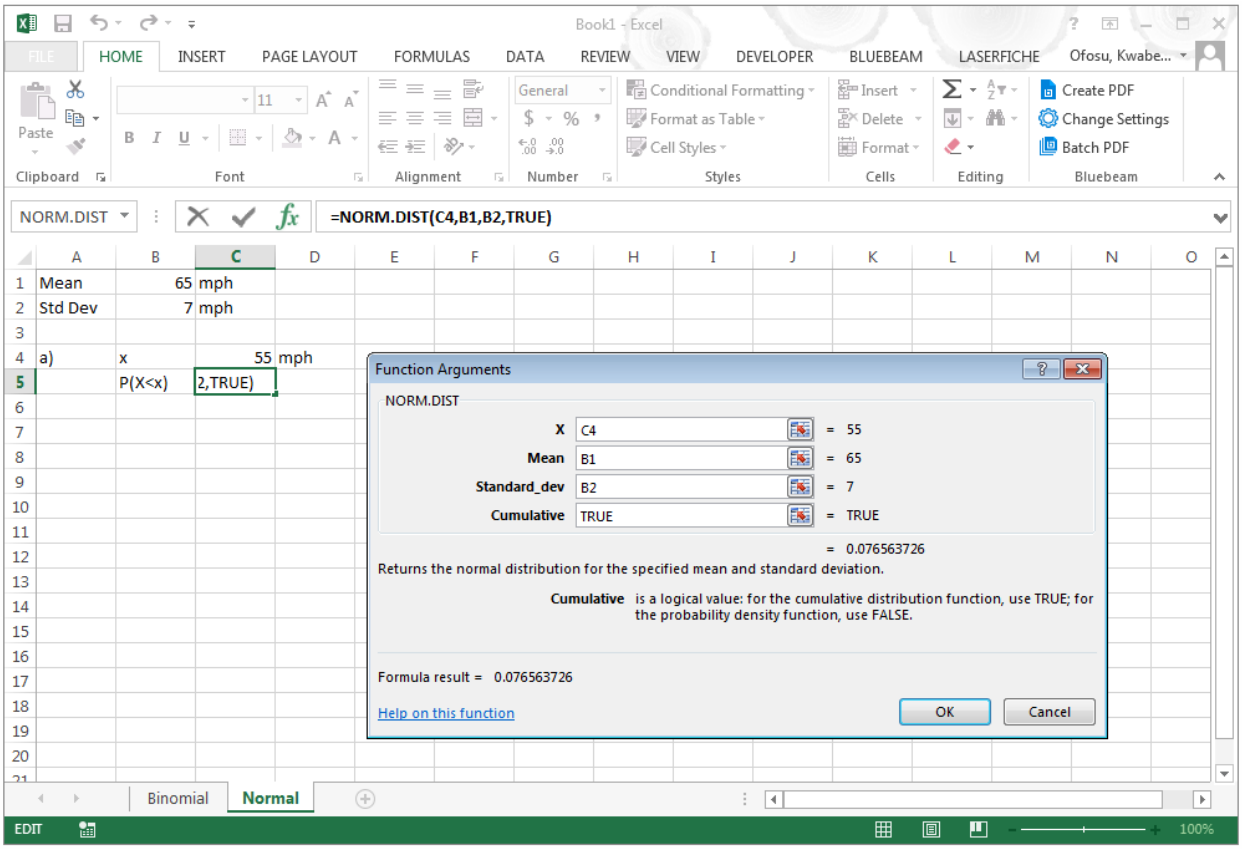
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Complete the arguments entries as follows.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

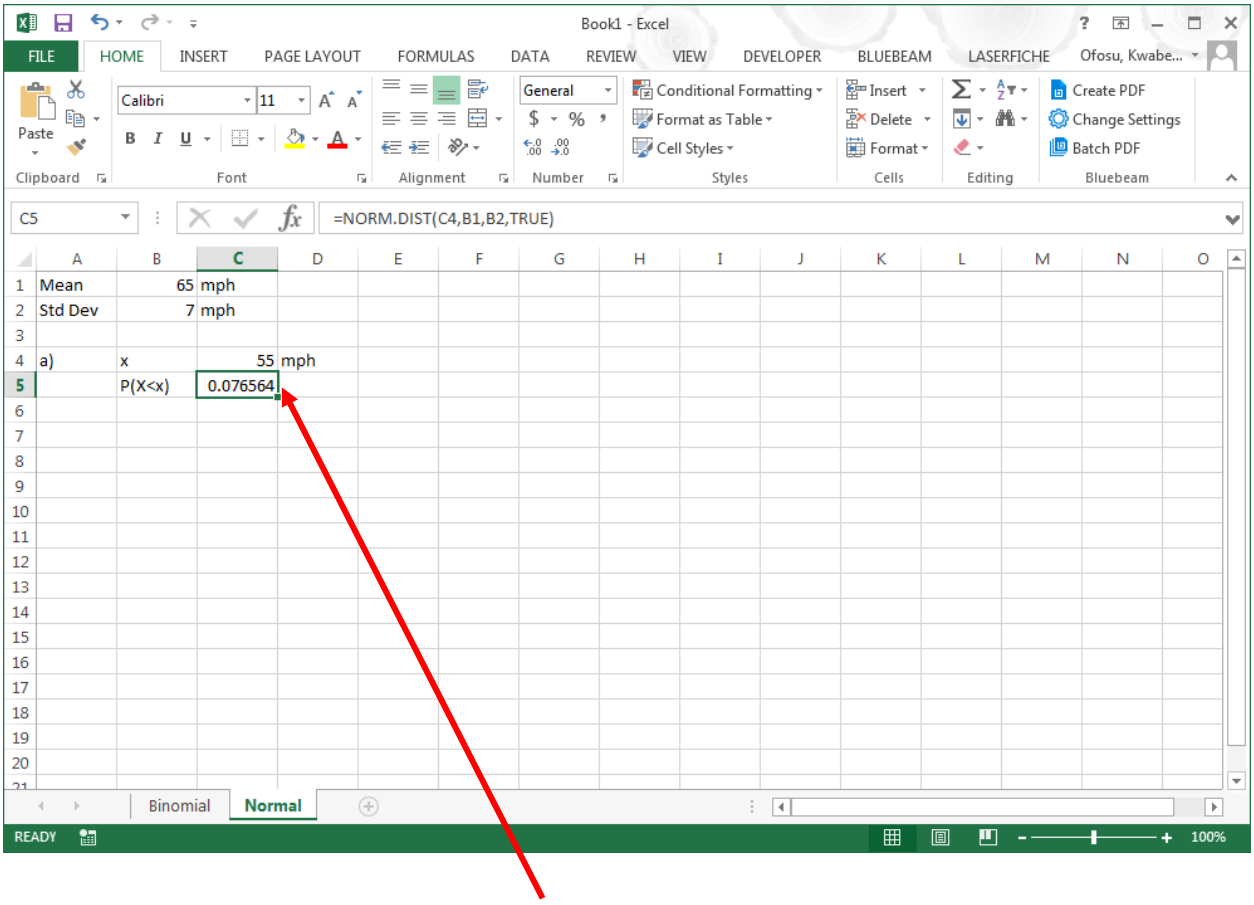
This is a cumulative probability calculation.
Under **Cumulative**, type in “TRUE”.



Review the **Formula result**.
Hit **OK**.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

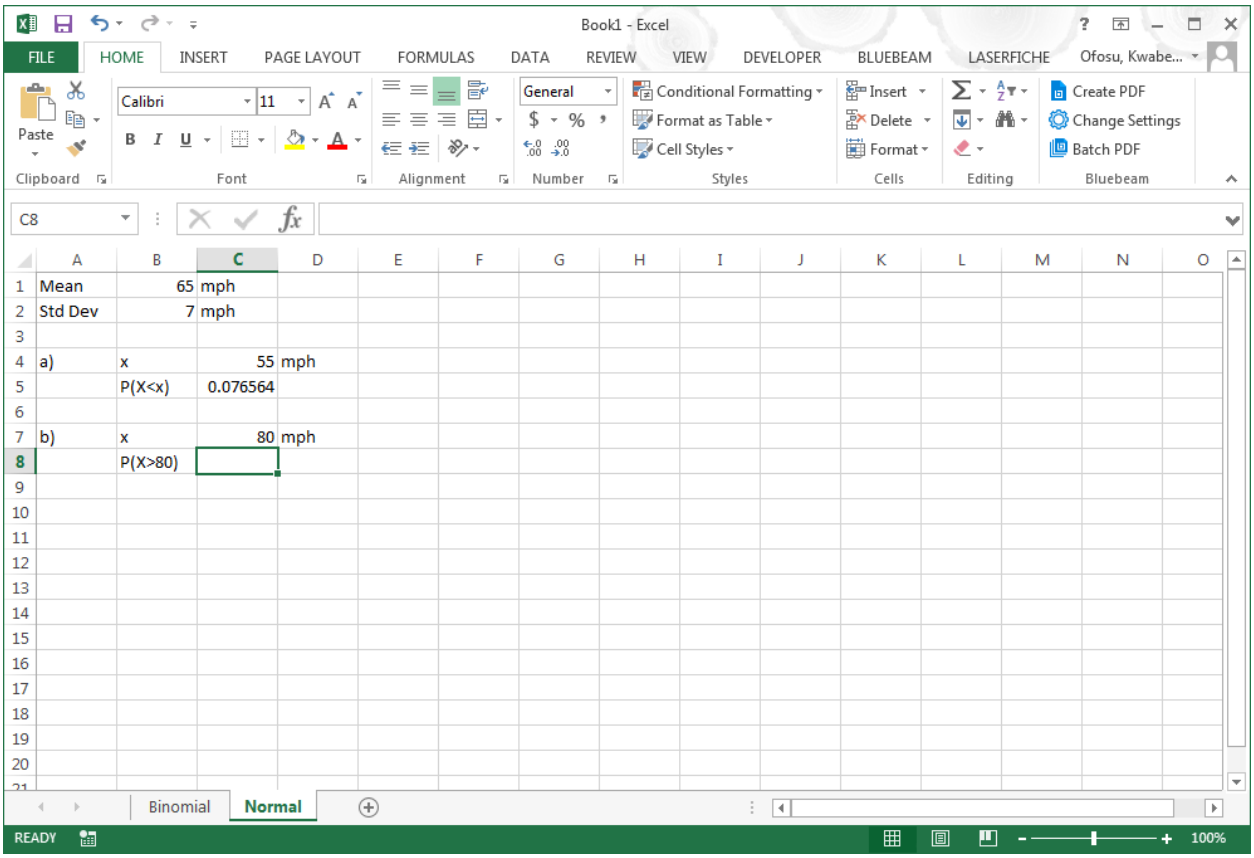
The result is posted to the current cell.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

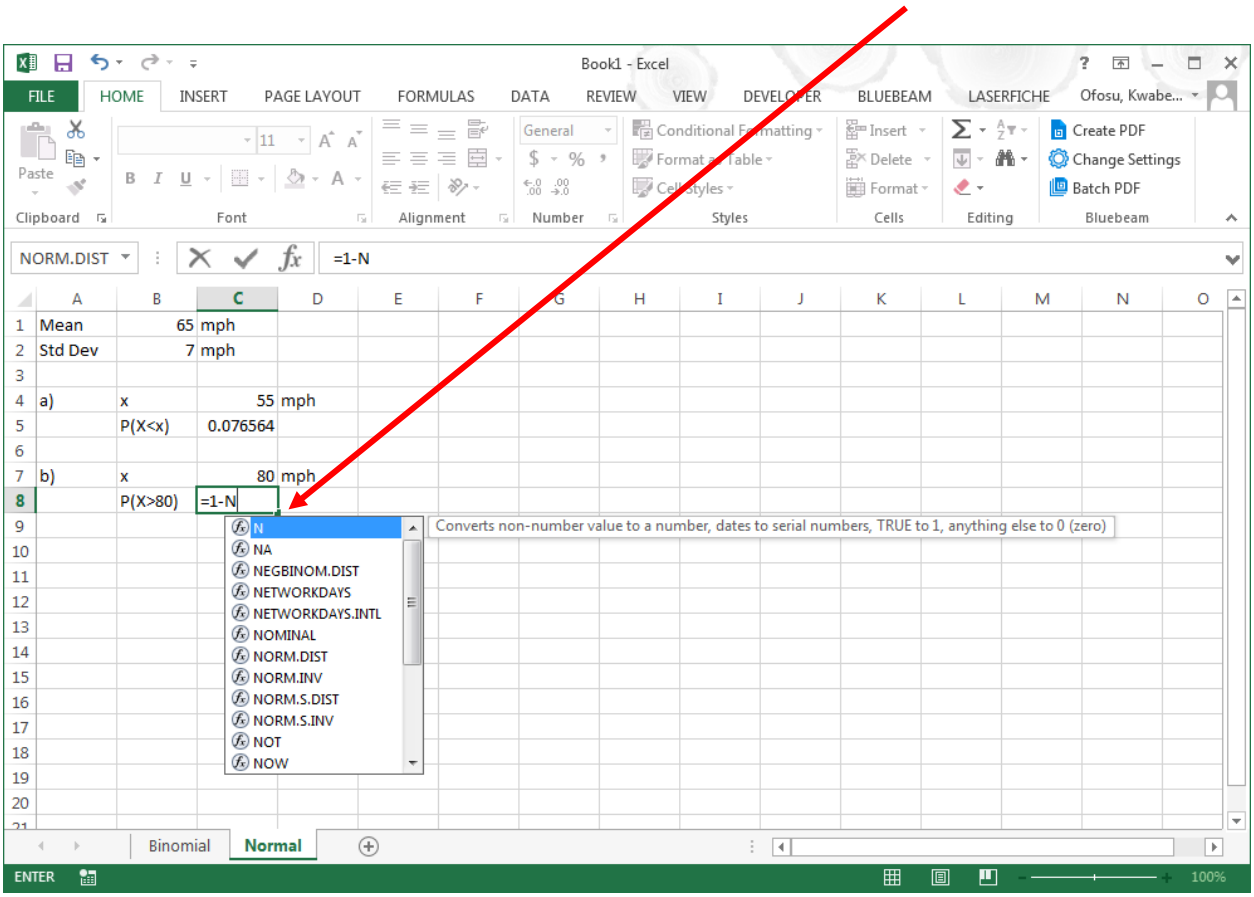
b) $P(X \geq 80) = 1 - P(X < 80)$

Update the spreadsheet as follows:



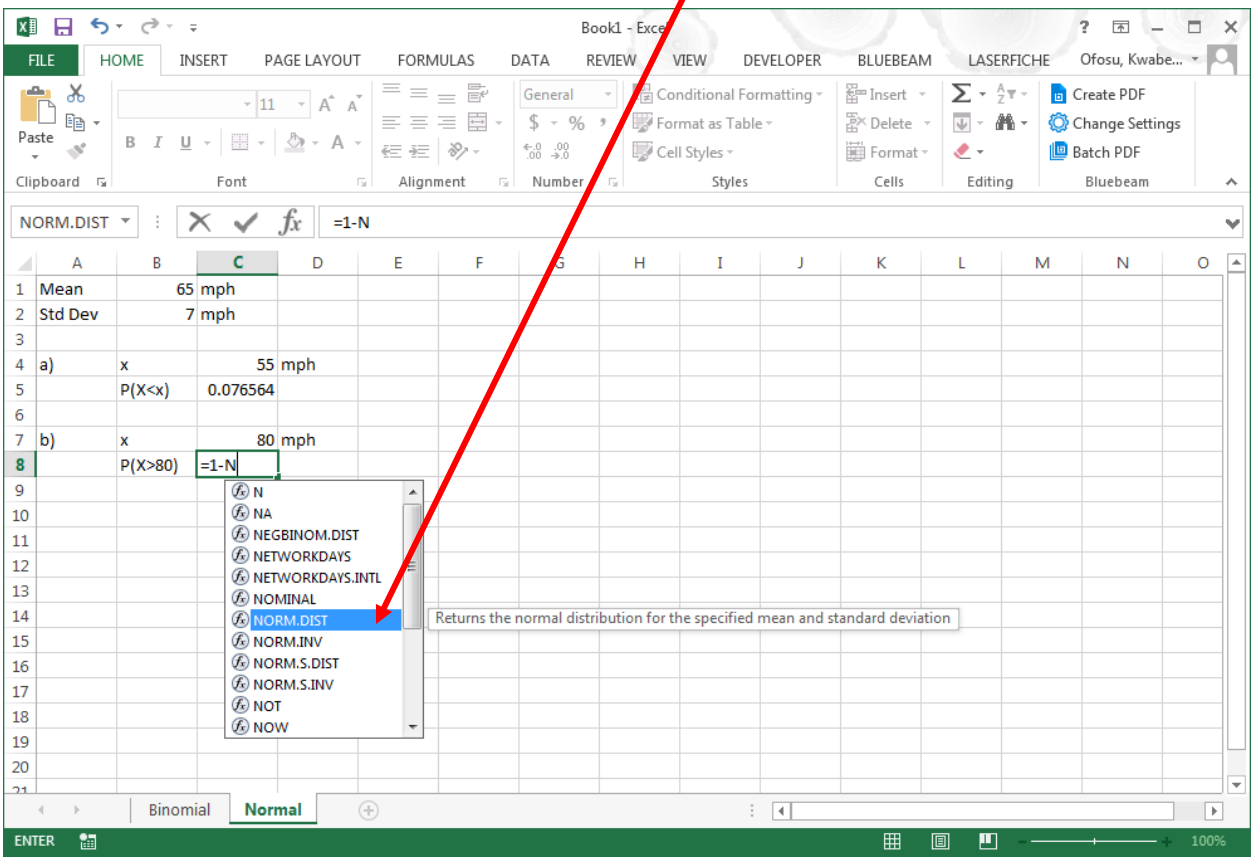
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

In the current cell, type the following:
" = 1 – Normal"
However, before you complete the entry the candidate list appears.



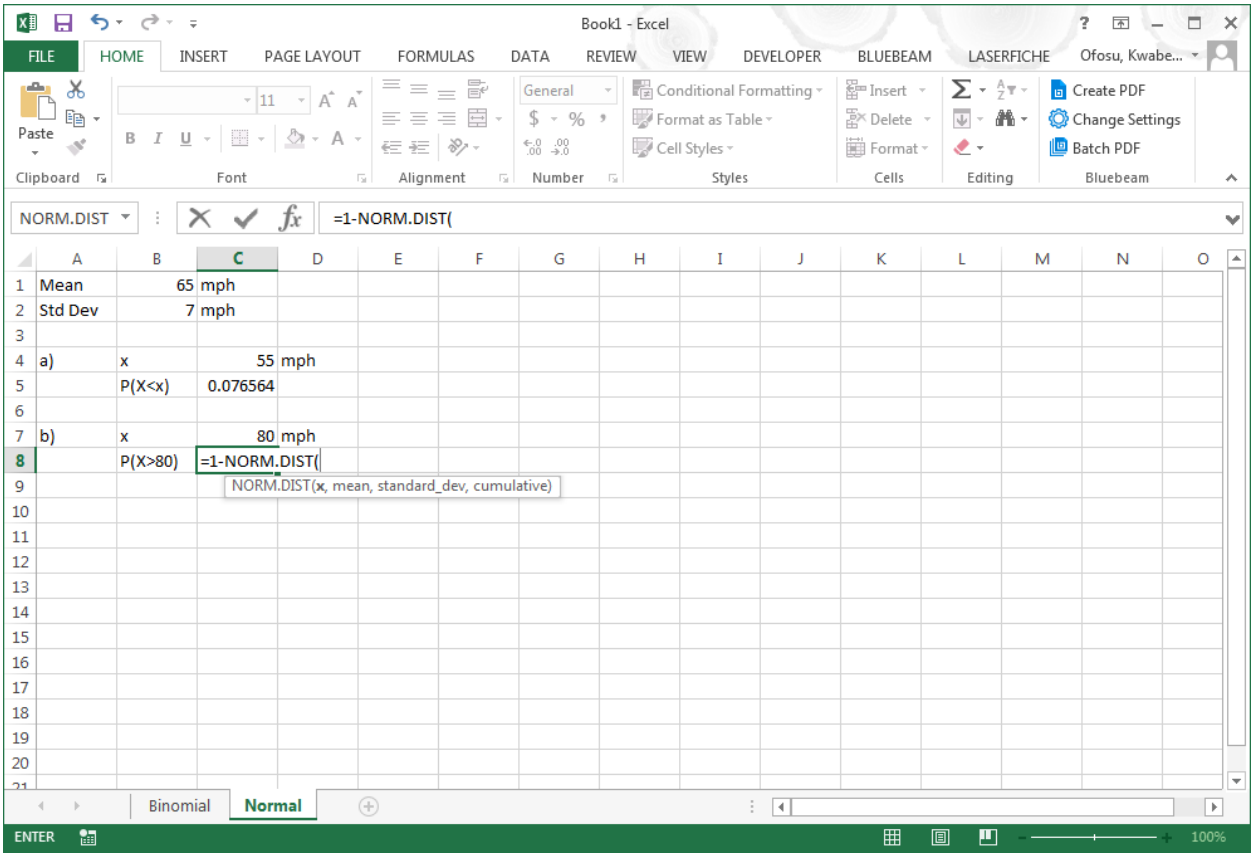
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Select NORM.DIST.
Double click on NORM.DIST.



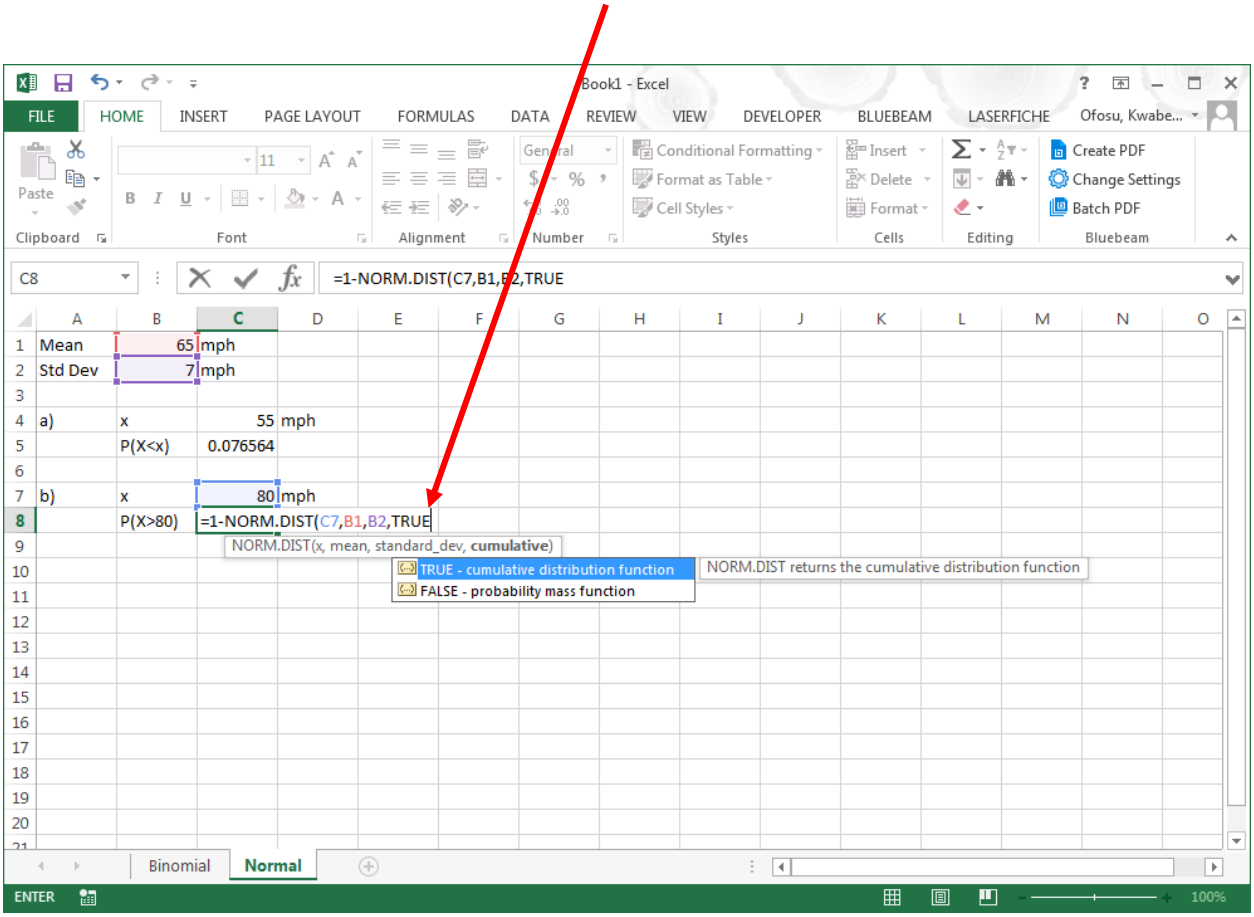
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The function prompt appears.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

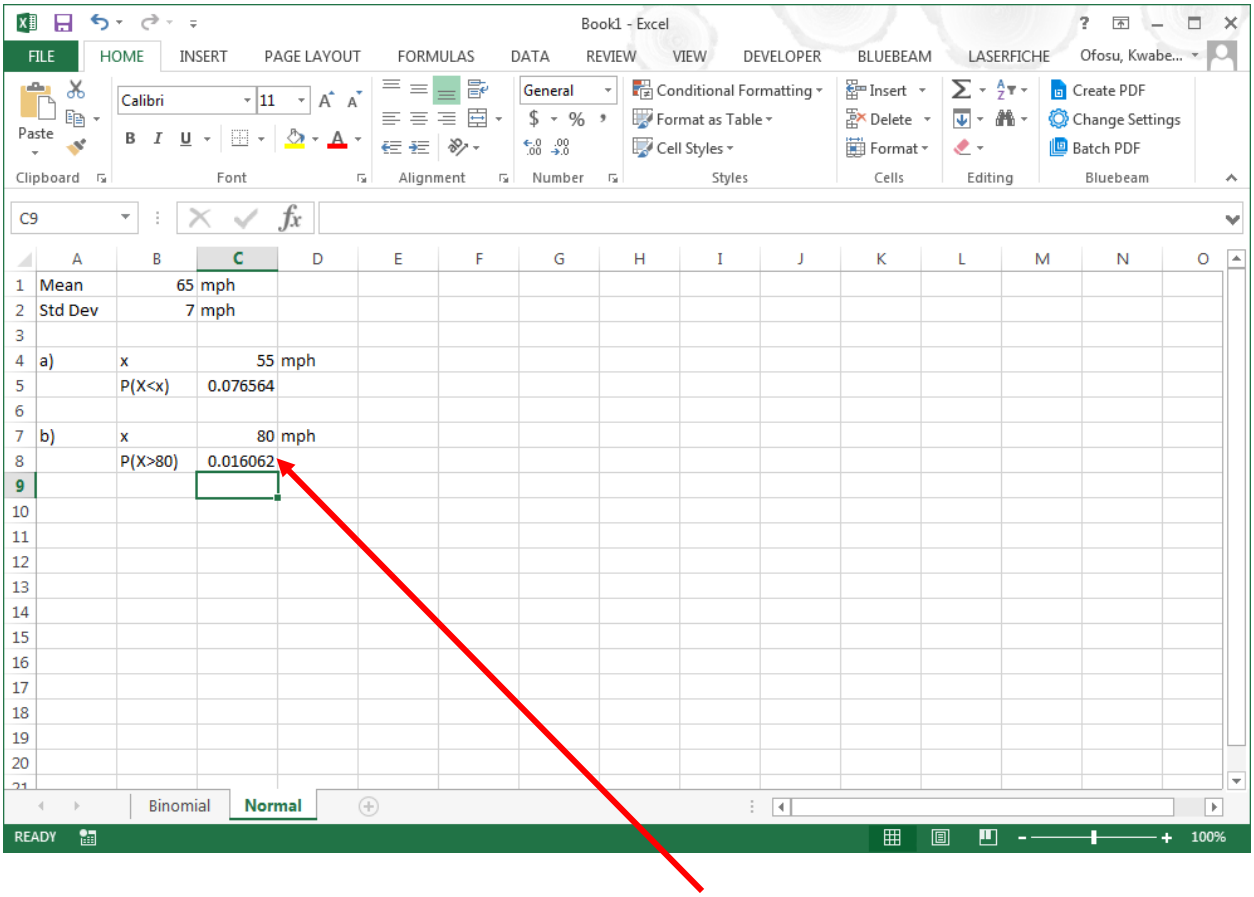
Point-and click to enter the arguments in the order shown in the tooltip, separating each argument with a comma (“,”).



Close the parenthesis.
Hit **Enter** on your keyboard.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

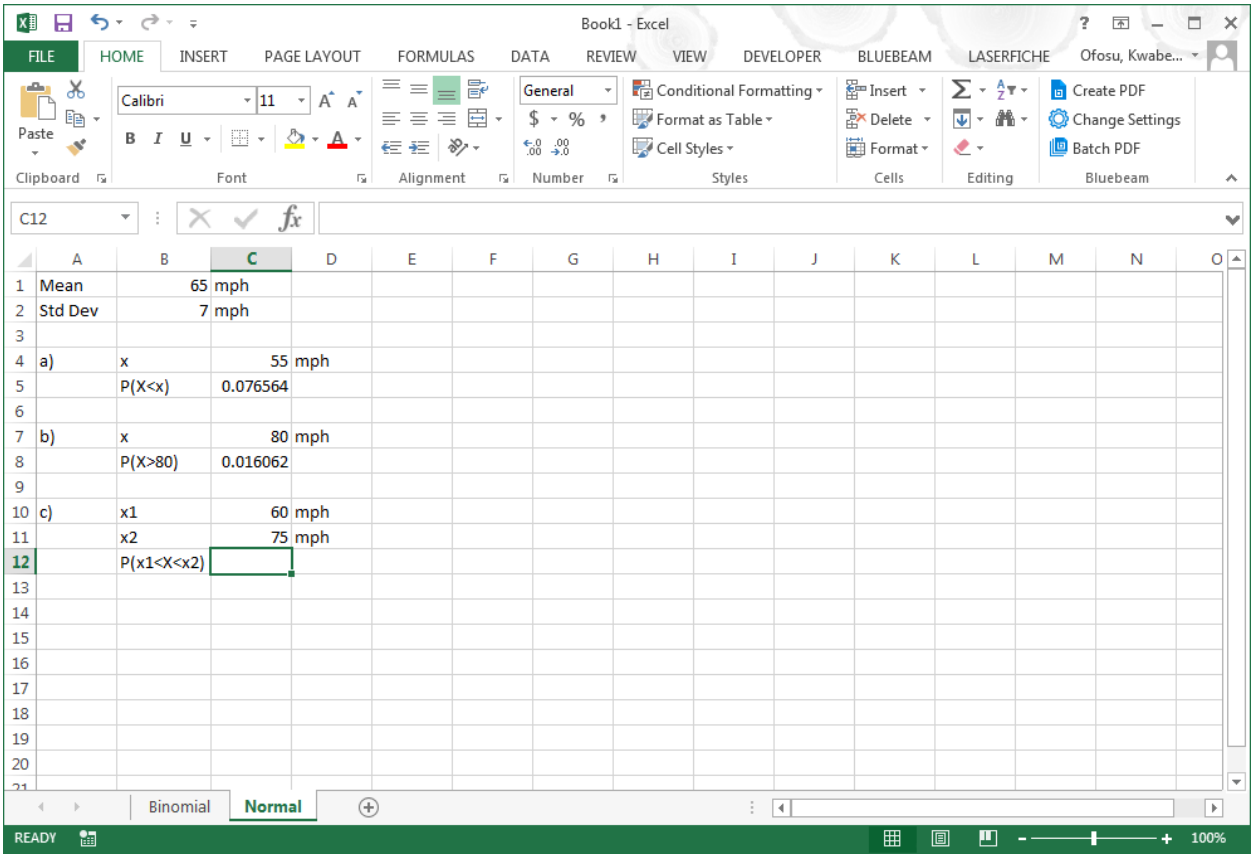
The result is displayed in the current cell.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

c) $P(60 \leq X \leq 75)$

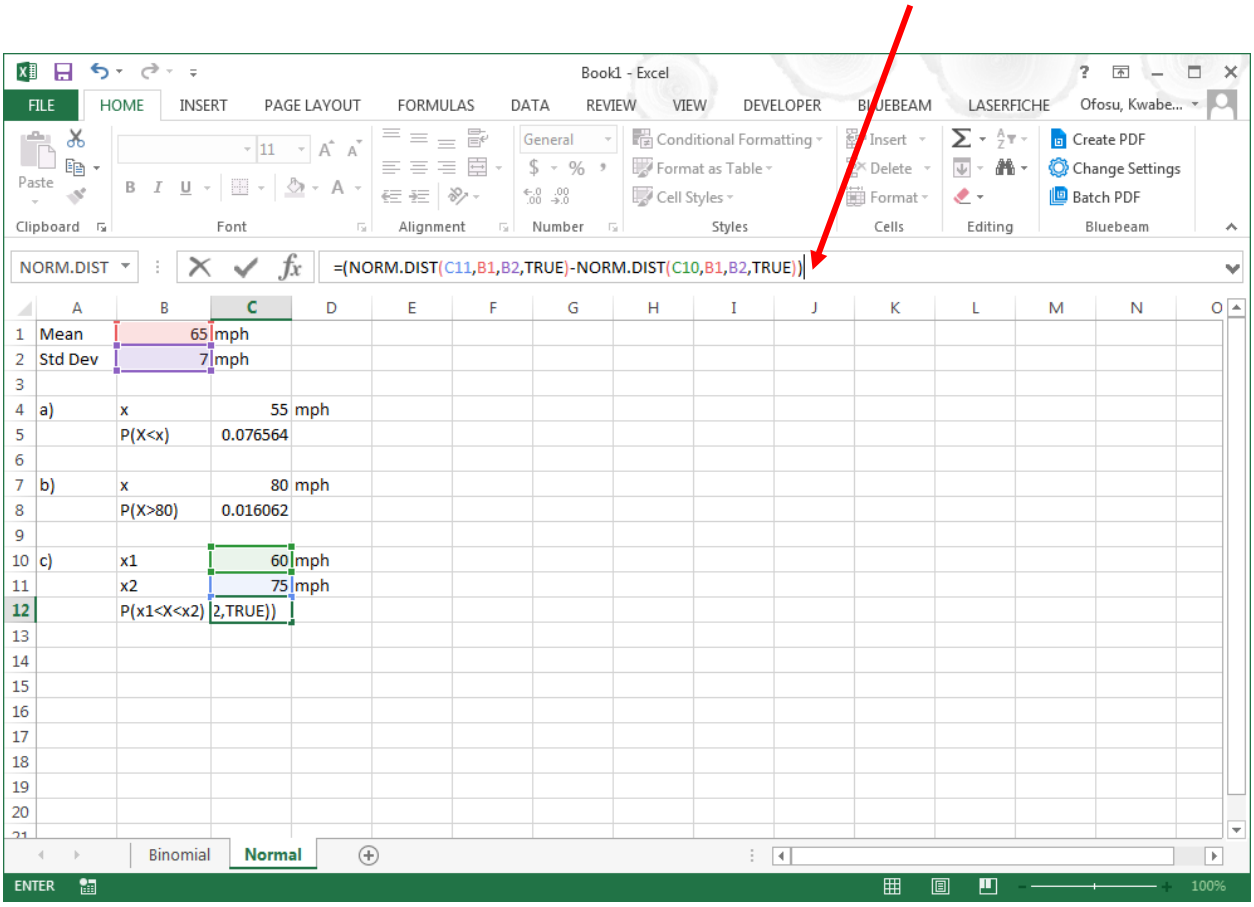
Update the spreadsheet as follows:



	A	B	C	D
1	Mean	65	mph	
2	Std Dev	7	mph	
3				
4	a)	x	55	mph
5		P(X<x)	0.076564	
6				
7	b)	x	80	mph
8		P(X>80)	0.016062	
9				
10	c)	x1	60	mph
11		x2	75	mph
12		P(x1<X<x2)		

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

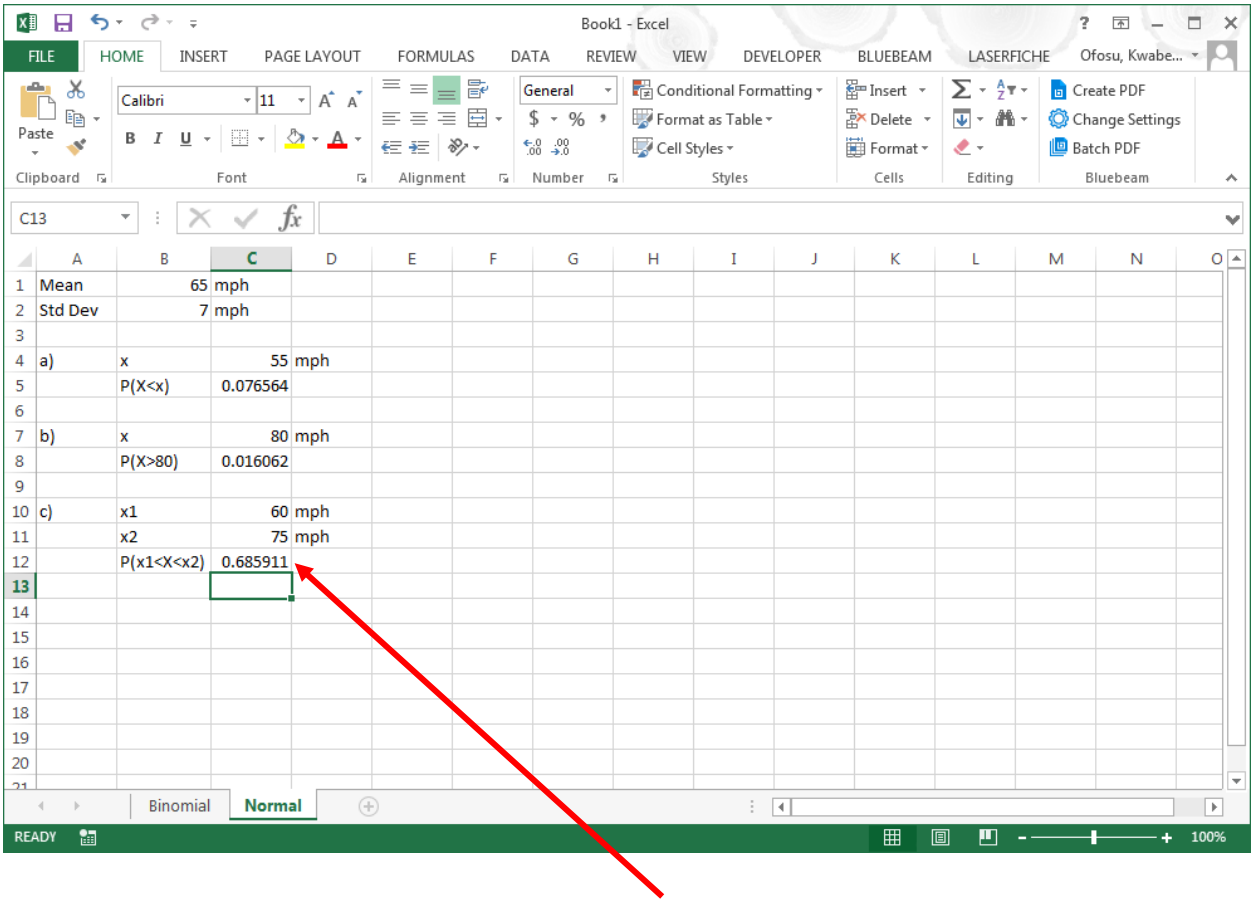
In the **Formula Bar** type the following:
(Pay attention to the parentheses.
Use the tooltips as they appear.
Use point-and-click as needed versus typing in all entries.)



Hit **Enter** on the keyboard.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

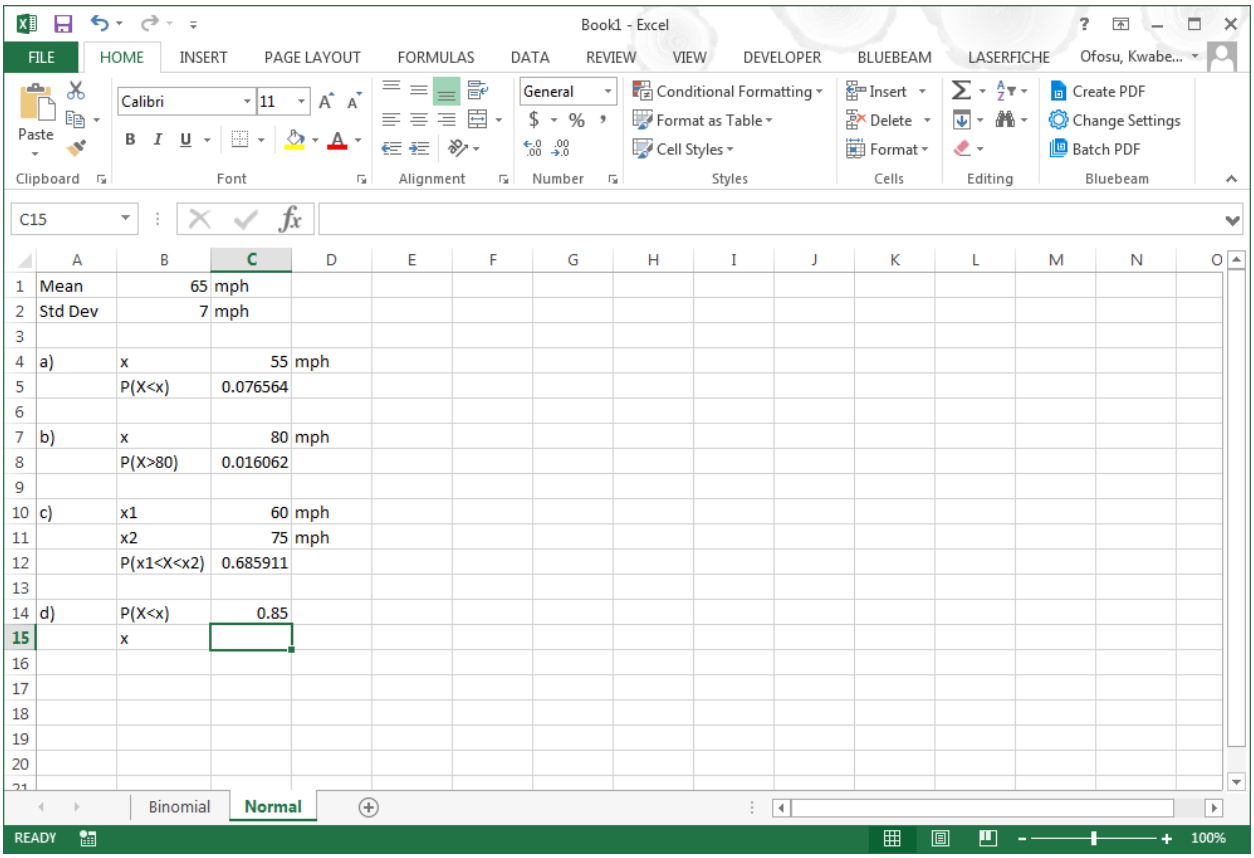
The result displays in the current cell.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

d) The 85th percentile speed is the x value that has a cumulative probability of 0.85. This can be calculated by the inverse cumulative density function (also called the quantile function, or the percent point function).

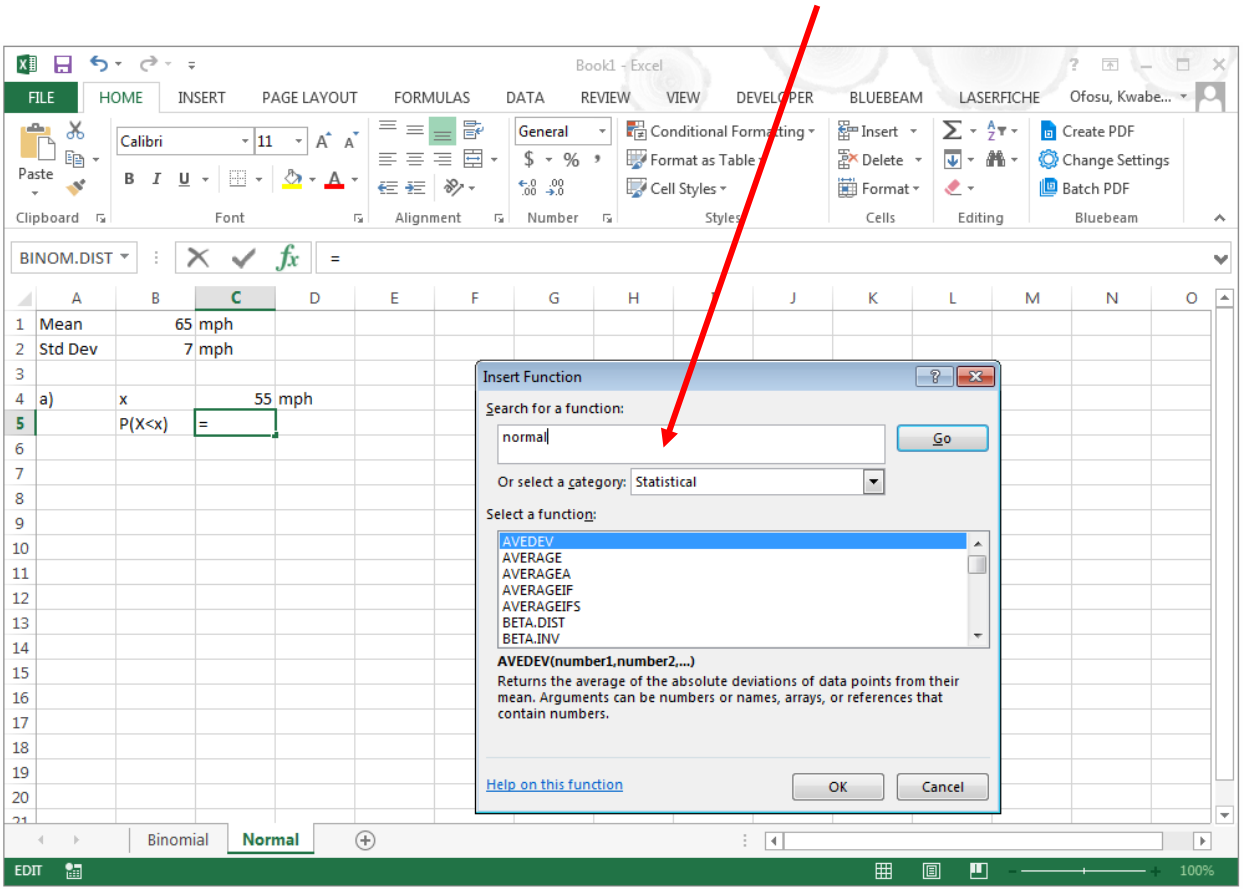
Update the spreadsheet as follows:



	A	B	C	D
1	Mean	65	mph	
2	Std Dev	7	mph	
3				
4	a)	x	55	mph
5		P(X<x)	0.076564	
6				
7	b)	x	80	mph
8		P(X>80)	0.016062	
9				
10	c)	x1	60	mph
11		x2	75	mph
12		P(x1<X<x2)	0.685911	
13				
14	d)	P(X<x)	0.85	
15		x		
16				
17				
18				
19				
20				

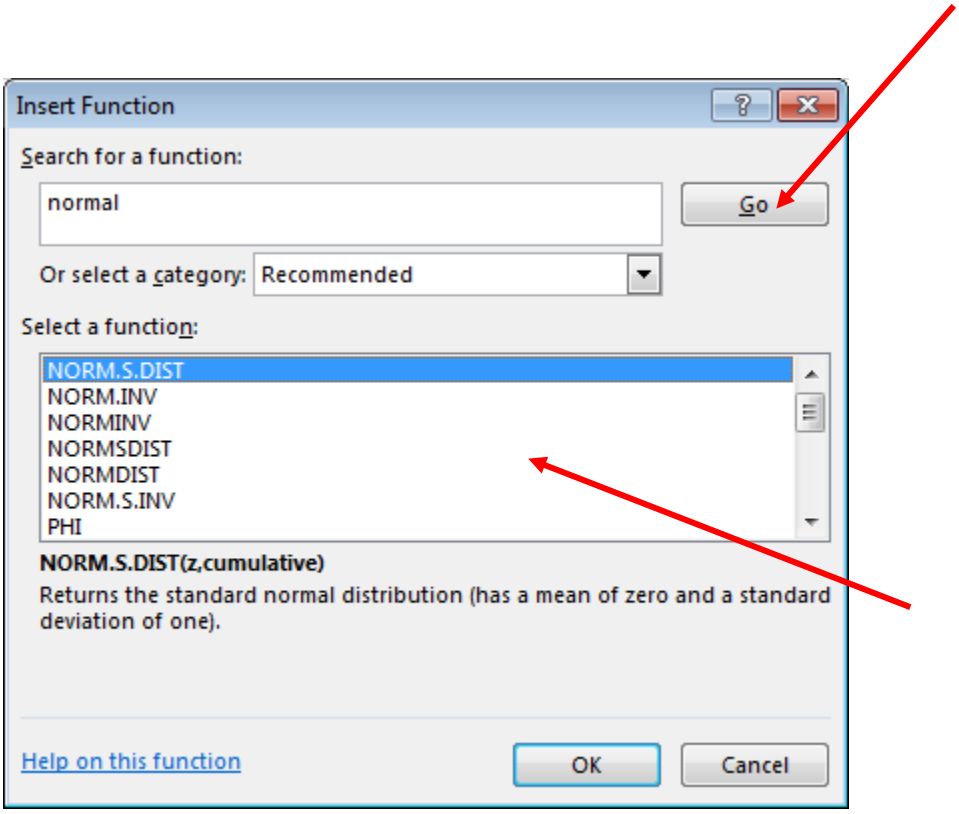
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Click on **Insert Function**.
Under **Or select a category**, select **Statistical**.
Under **Search for a function**, type “Normal”.



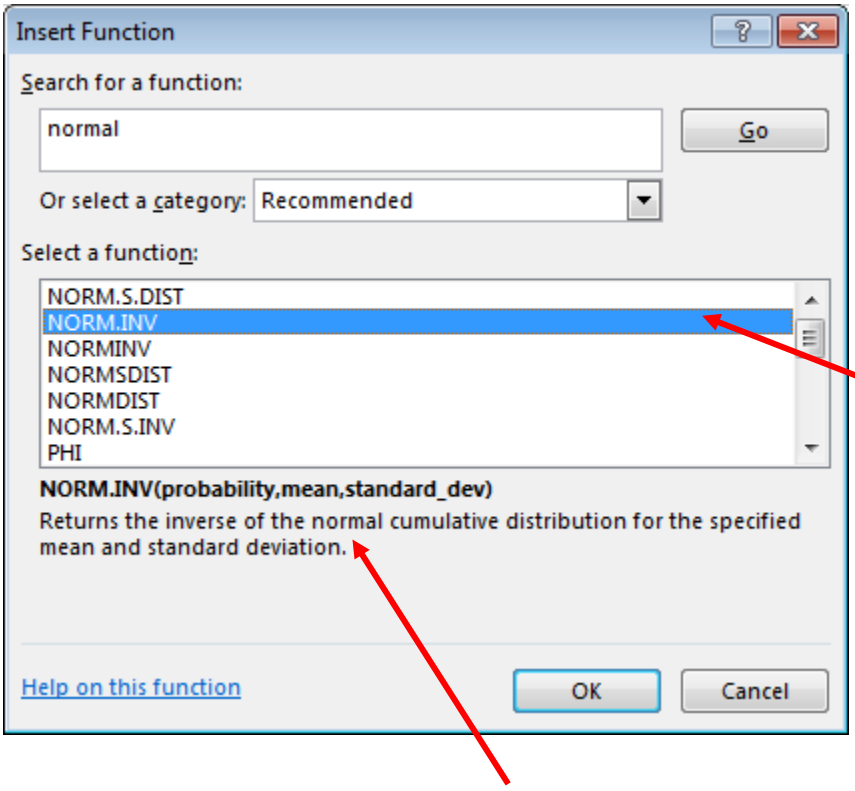
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Hit **Go**.
The list narrows down.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

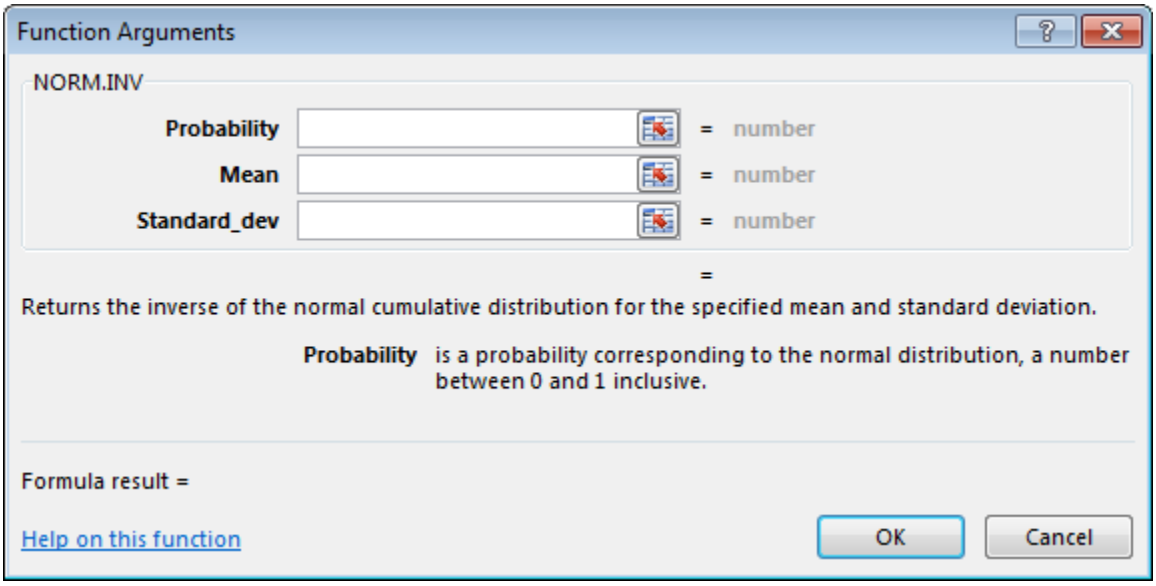
Scroll to NORM.INV
Select NORM.INV.
Review the description.
Confirm this is the function that is needed.



Hit **OK**.

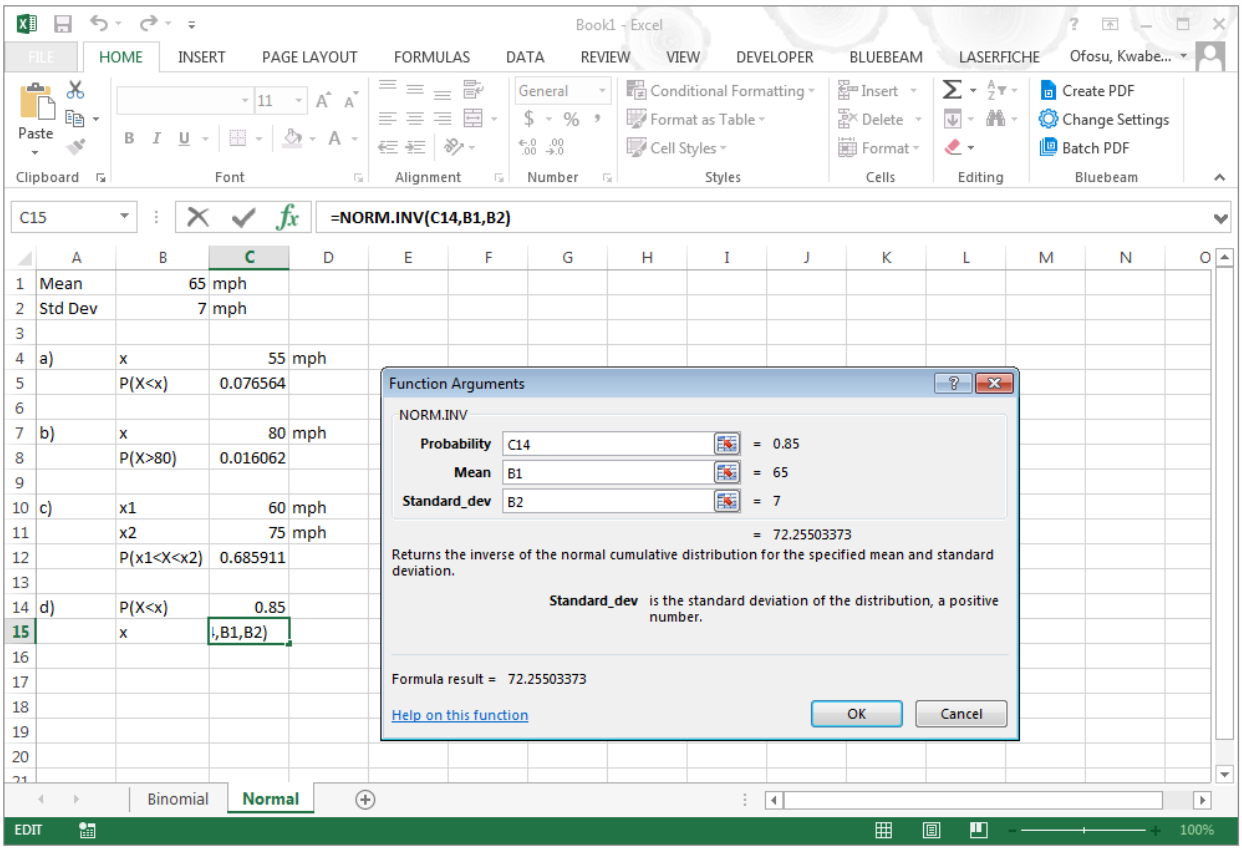
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The **Function Arguments** window opens.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

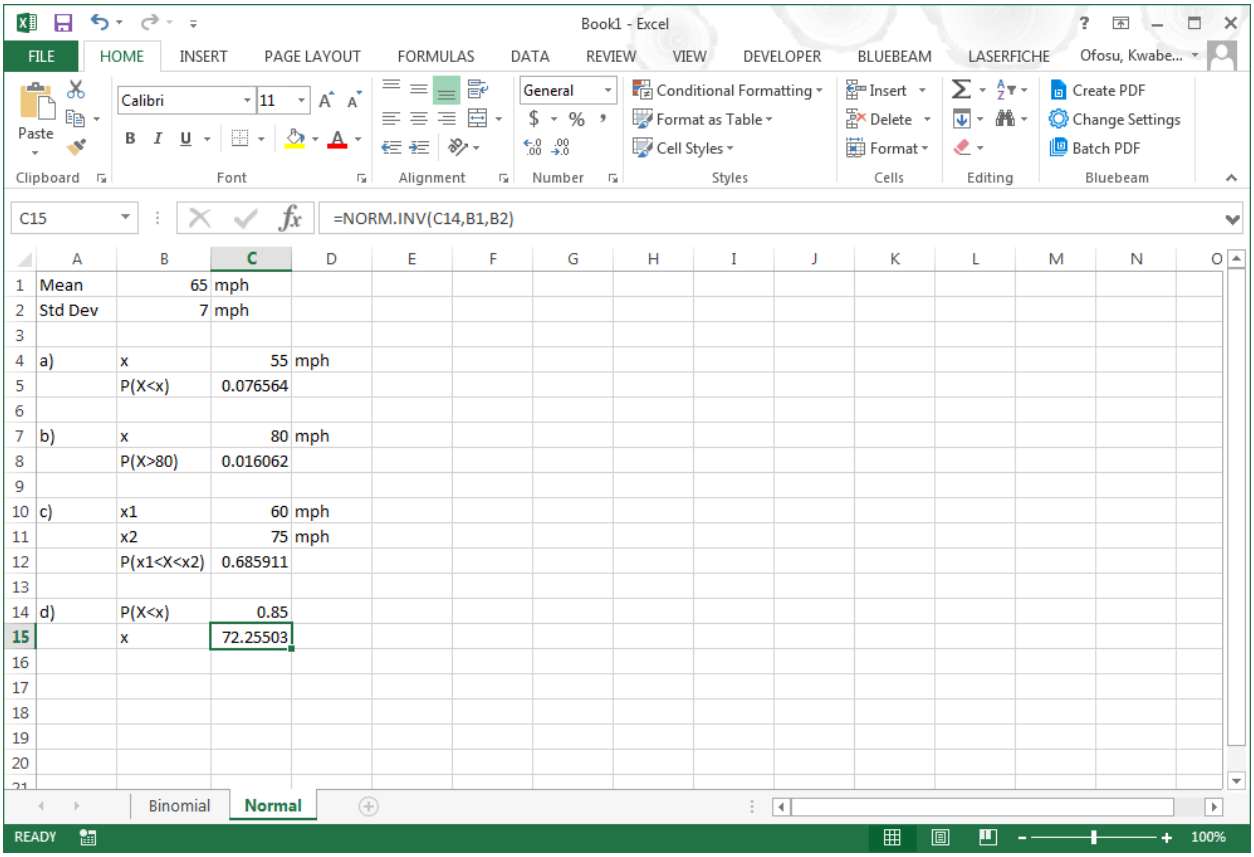
Complete the arguments entries as follows.
(Use point-and-click) as needed.



Review the **Formula result**.
Hit **OK**.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The result displays.



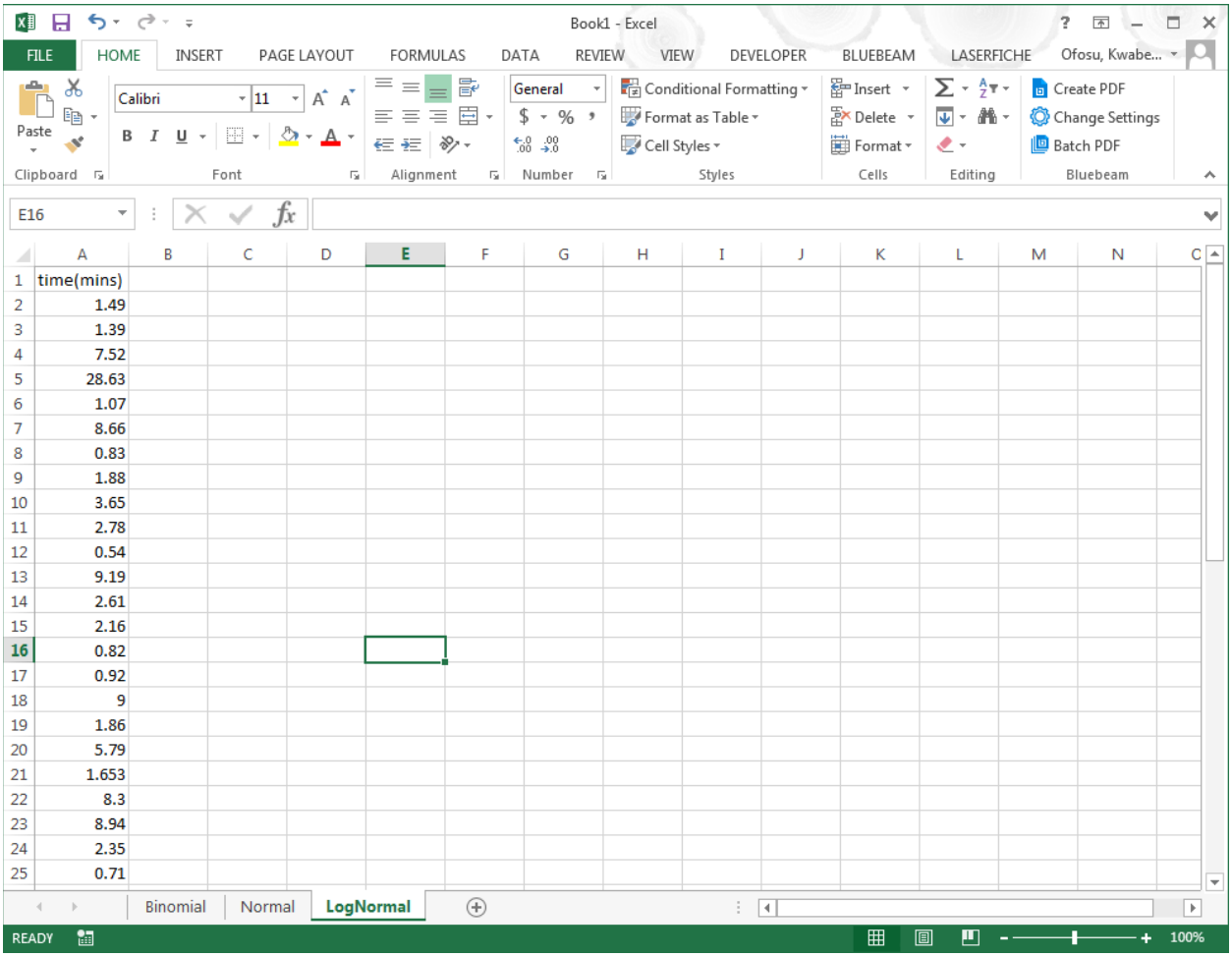
	A	B	C	D
1	Mean	65	mph	
2	Std Dev	7	mph	
3				
4	a)	x	55	mph
5		P(X<x)	0.076564	
6				
7	b)	x	80	mph
8		P(X>80)	0.016062	
9				
10	c)	x1	60	mph
11		x2	75	mph
12		P(x1<X<x2)	0.685911	
13				
14	d)	P(X<x)	0.85	
15		x	72.25503	

Complete.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Example:

An engineer collected the following data for the completion time (in minutes) of an industrial process. (The data can be found in the *Excel* file provided with your Suncam course materials, on the spreadsheet named “LogNormal”).

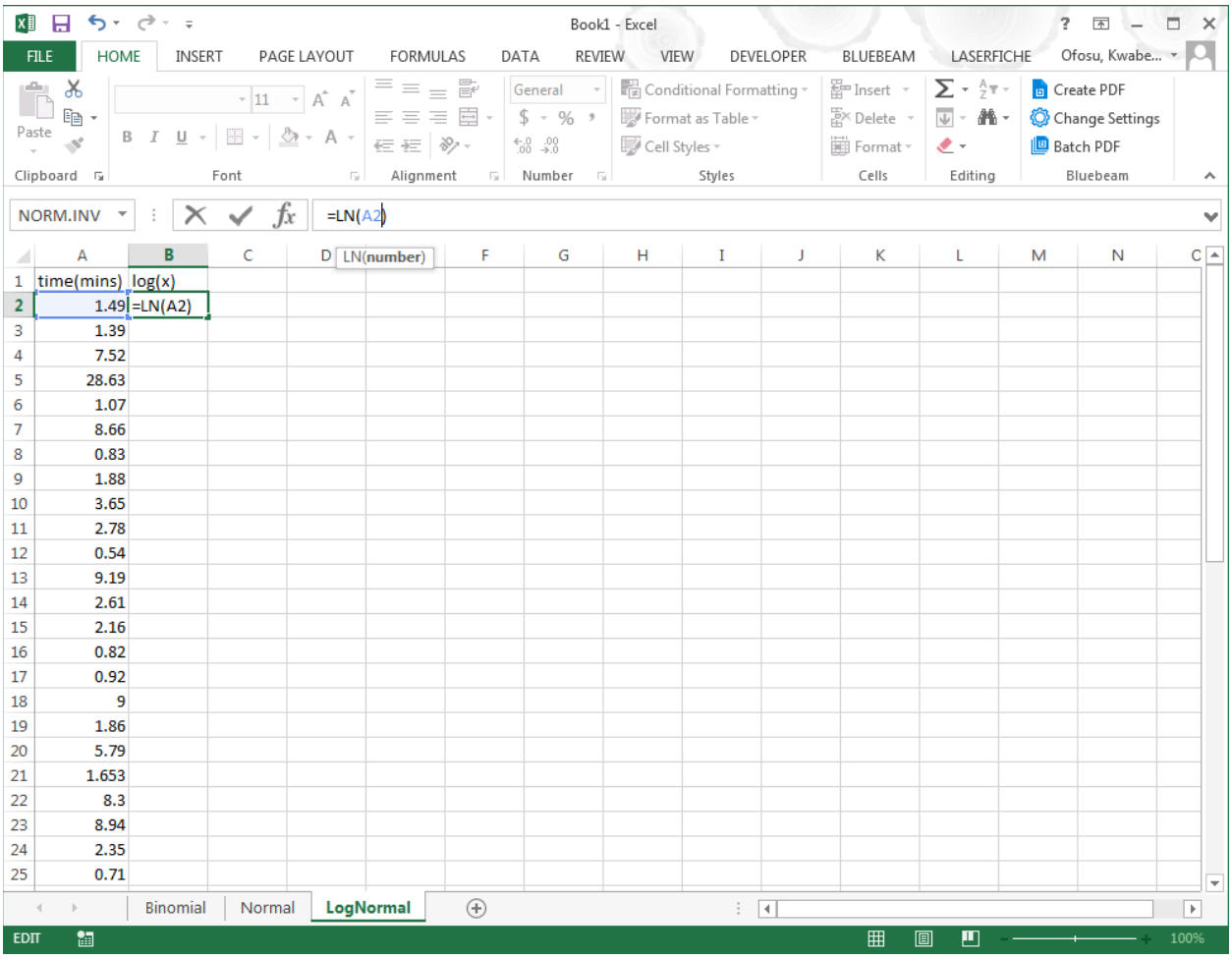


	A
1	time(mins)
2	1.49
3	1.39
4	7.52
5	28.63
6	1.07
7	8.66
8	0.83
9	1.88
10	3.65
11	2.78
12	0.54
13	9.19
14	2.61
15	2.16
16	0.82
17	0.92
18	9
19	1.86
20	5.79
21	1.653
22	8.3
23	8.94
24	2.35
25	0.71

Prior knowledge and experience suggests the Log-Normal distribution is the appropriate distribution to describe this data.

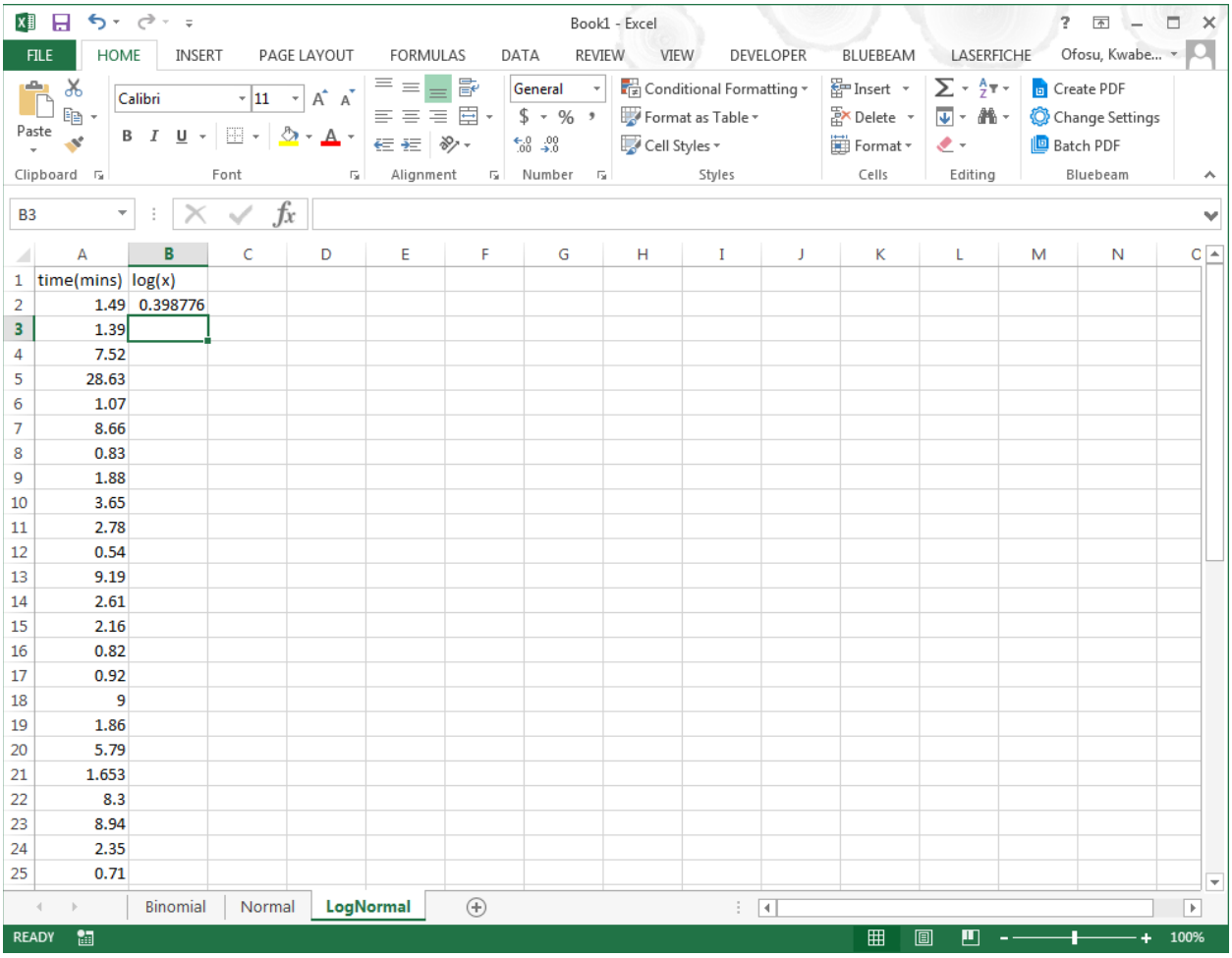
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

In order to work with the Log-Normal distribution, each value must be converted to its natural logarithm value.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

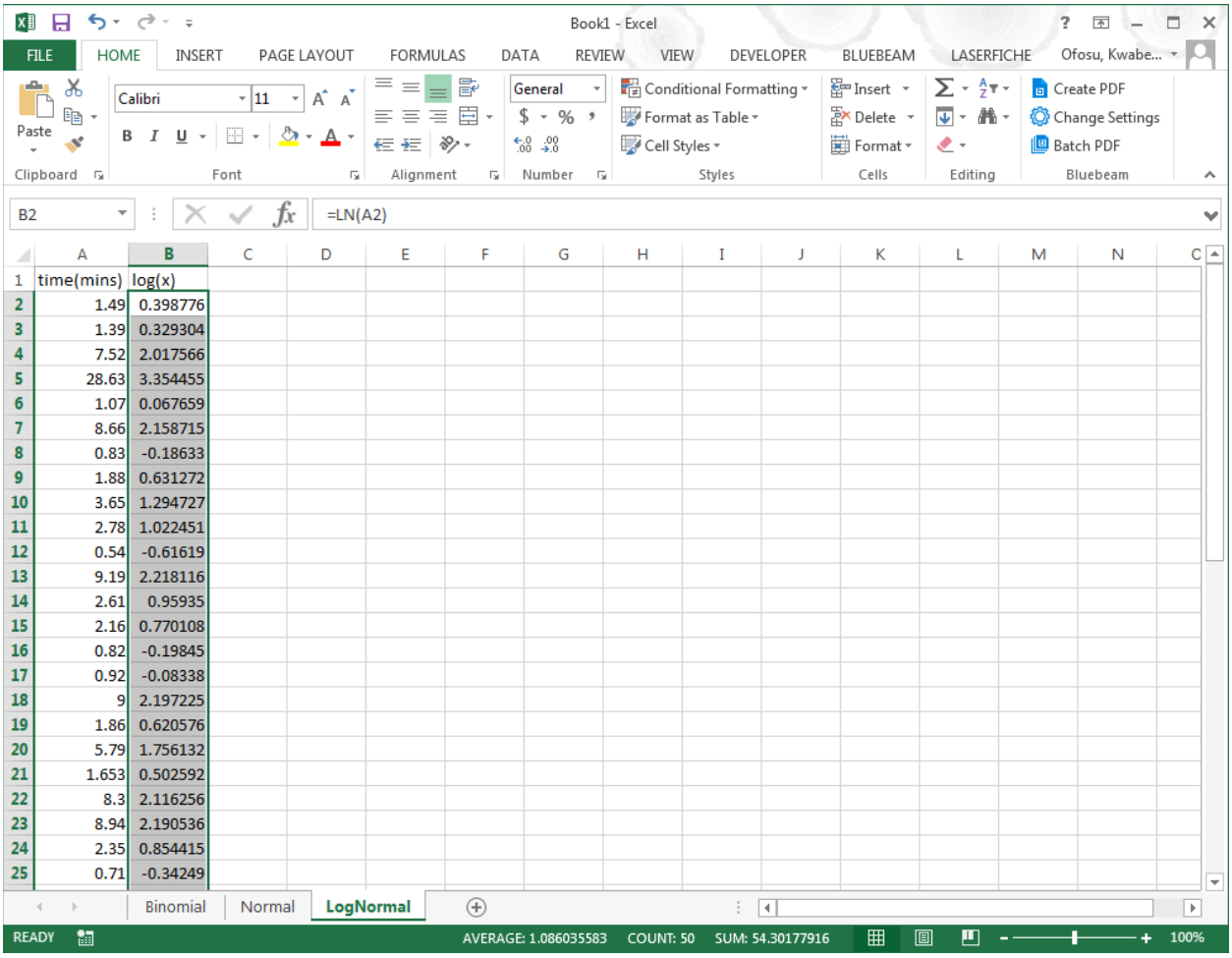
Which yields,



	A	B
1	time(mins)	log(x)
2	1.49	0.398776
3	1.39	
4	7.52	
5	28.63	
6	1.07	
7	8.66	
8	0.83	
9	1.88	
10	3.65	
11	2.78	
12	0.54	
13	9.19	
14	2.61	
15	2.16	
16	0.82	
17	0.92	
18	9	
19	1.86	
20	5.79	
21	1.653	
22	8.3	
23	8.94	
24	2.35	
25	0.71	

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

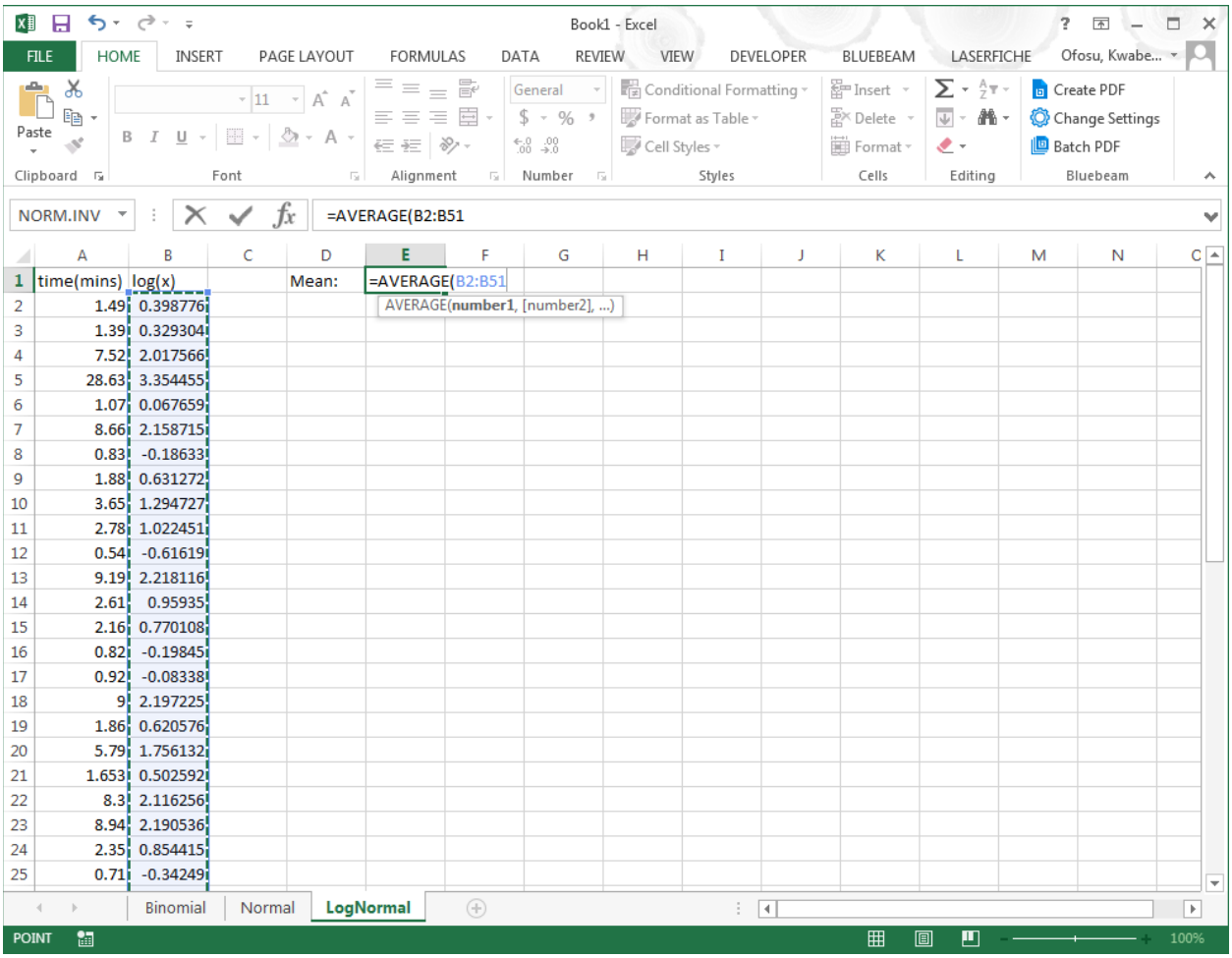
And then replicate for the other values.



	A	B
1	time(mins)	log(x)
2	1.49	0.398776
3	1.39	0.329304
4	7.52	2.017566
5	28.63	3.354455
6	1.07	0.067659
7	8.66	2.158715
8	0.83	-0.18633
9	1.88	0.631272
10	3.65	1.294727
11	2.78	1.022451
12	0.54	-0.61619
13	9.19	2.218116
14	2.61	0.95935
15	2.16	0.770108
16	0.82	-0.19845
17	0.92	-0.08338
18	9	2.197225
19	1.86	0.620576
20	5.79	1.756132
21	1.653	0.502592
22	8.3	2.116256
23	8.94	2.190536
24	2.35	0.854415
25	0.71	-0.34249

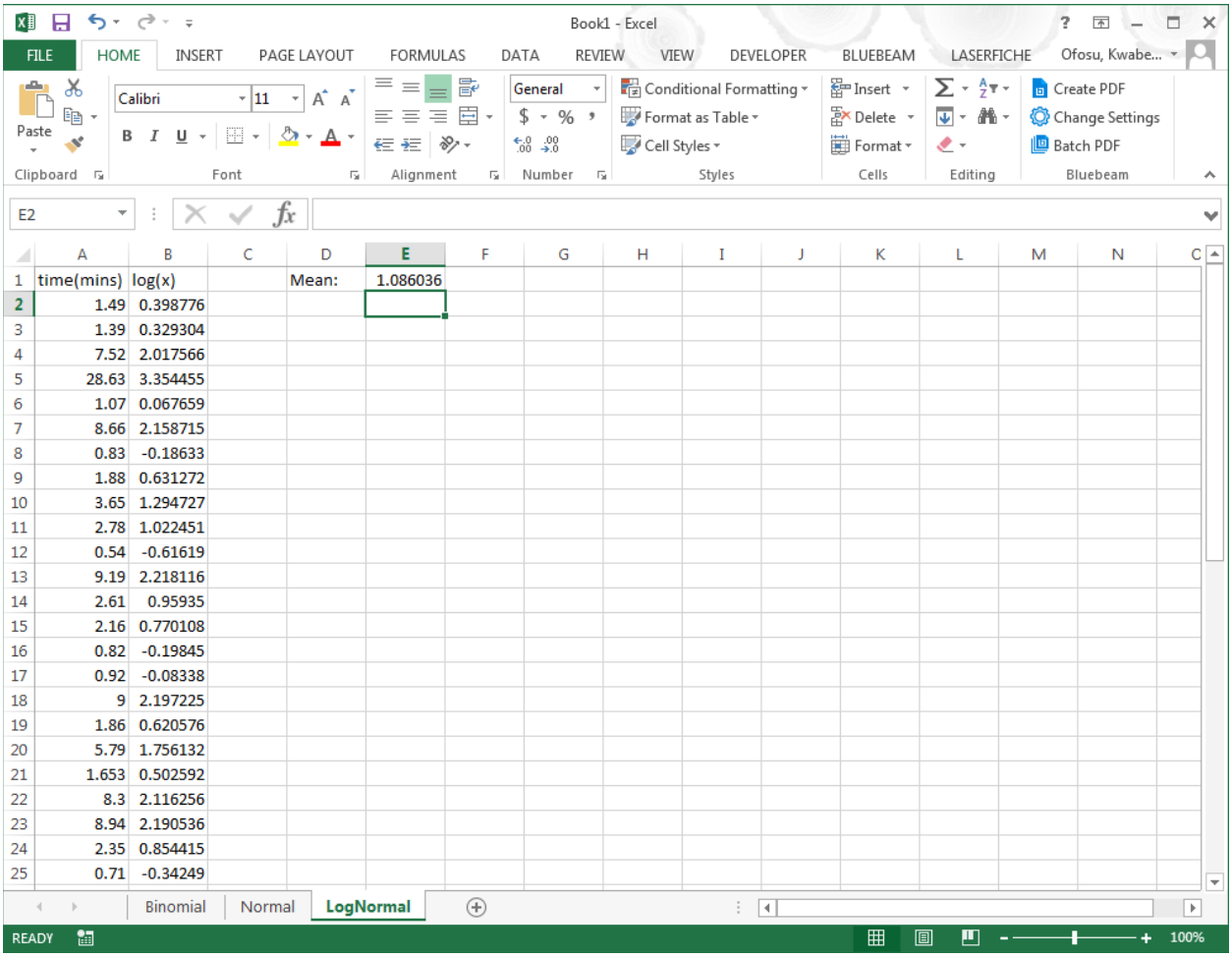
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Compute the mean of the log (x) values.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

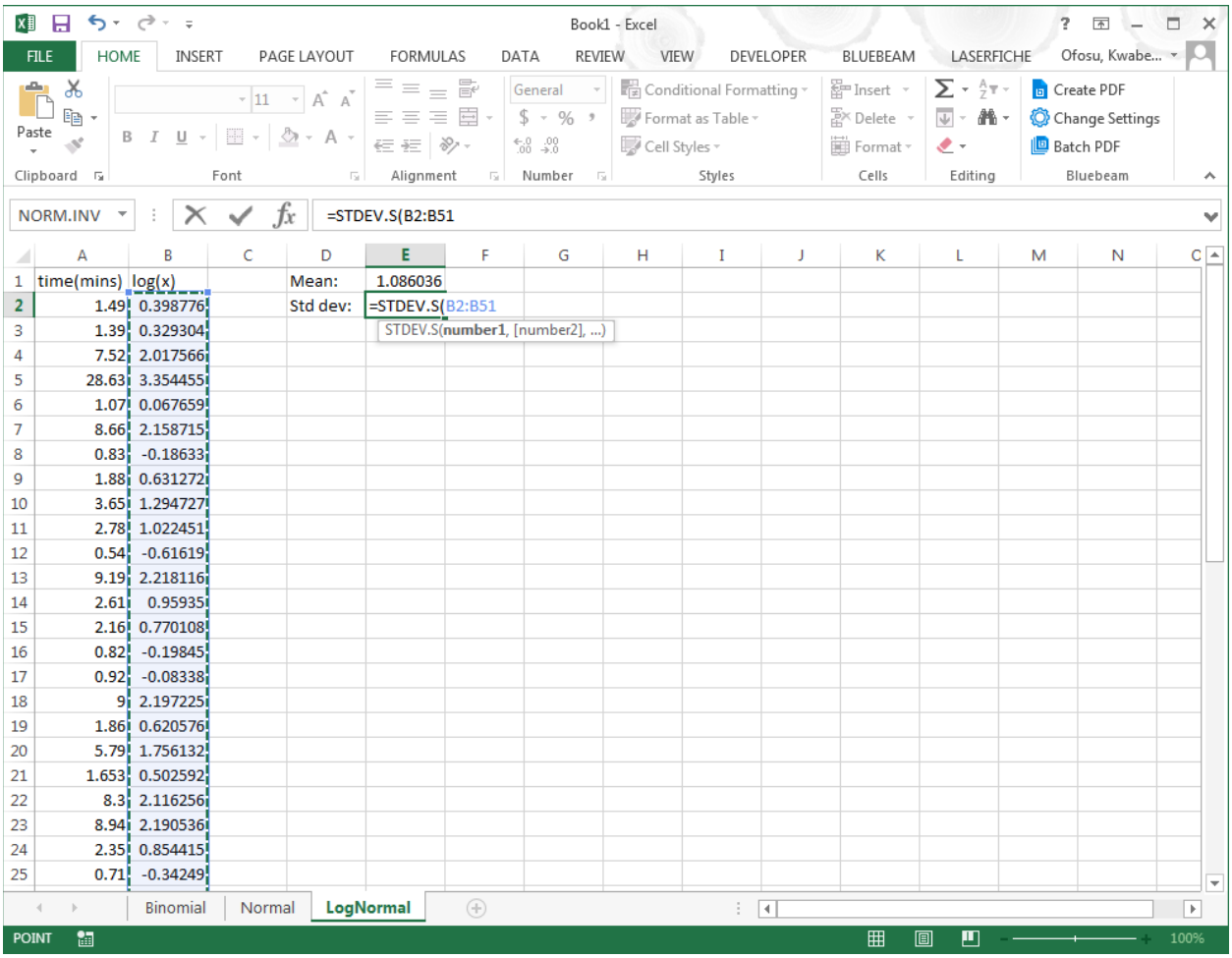
The result is as follows.



	A	B	C	D	E
1	time(mins)	log(x)		Mean:	1.086036
2	1.49	0.398776			
3	1.39	0.329304			
4	7.52	2.017566			
5	28.63	3.354455			
6	1.07	0.067659			
7	8.66	2.158715			
8	0.83	-0.18633			
9	1.88	0.631272			
10	3.65	1.294727			
11	2.78	1.022451			
12	0.54	-0.61619			
13	9.19	2.218116			
14	2.61	0.95935			
15	2.16	0.770108			
16	0.82	-0.19845			
17	0.92	-0.08338			
18	9	2.197225			
19	1.86	0.620576			
20	5.79	1.756132			
21	1.653	0.502592			
22	8.3	2.116256			
23	8.94	2.190536			
24	2.35	0.854415			
25	0.71	-0.34249			

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

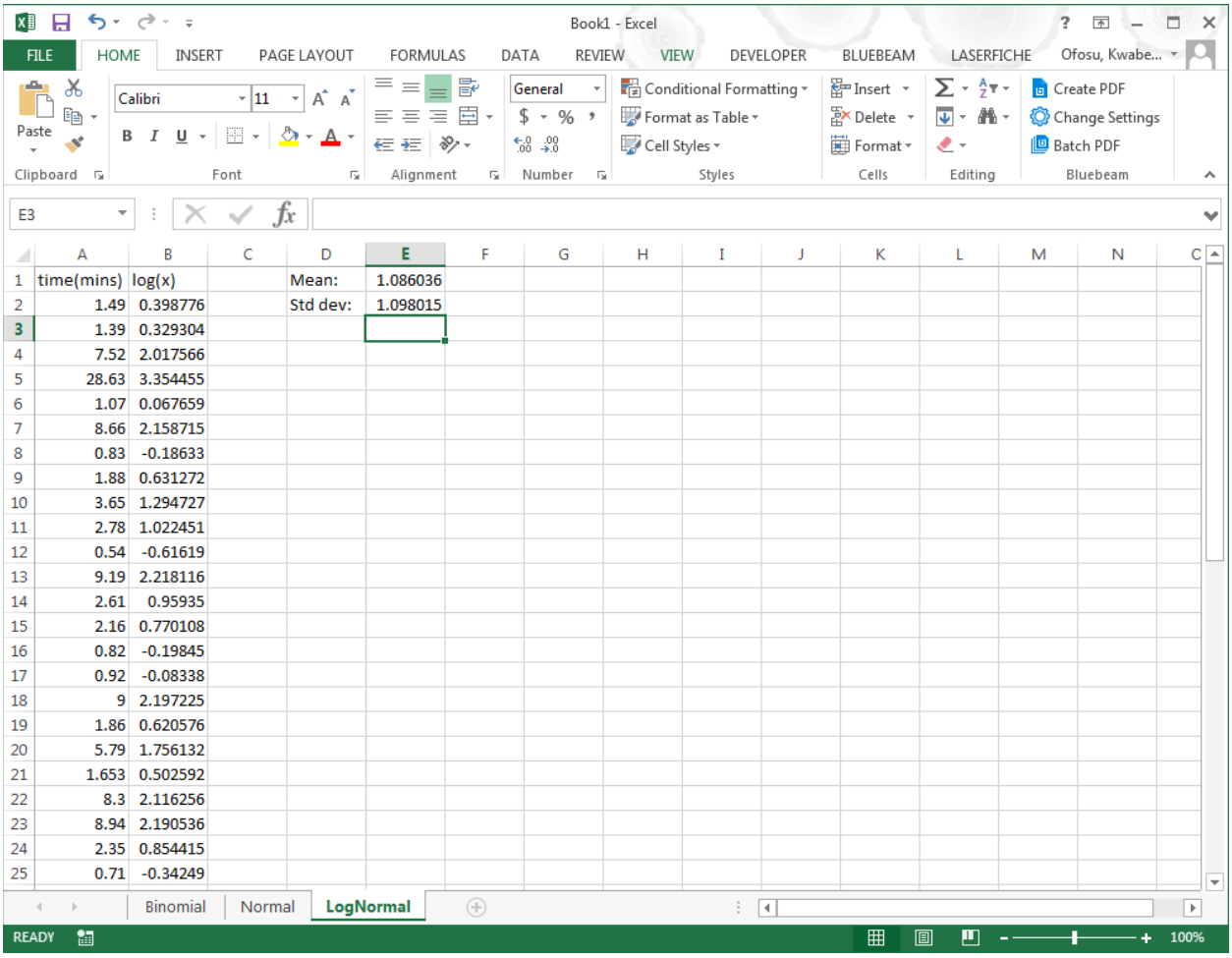
Compute the standard deviation of the log (x) values.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	C
1	time(mins)	log(x)		Mean:	1.086036										
2	1.49	0.398776		Std dev:	=STDEV.S(B2:B51										
3	1.39	0.329304													
4	7.52	2.017566													
5	28.63	3.354455													
6	1.07	0.067659													
7	8.66	2.158715													
8	0.83	-0.18633													
9	1.88	0.631272													
10	3.65	1.294727													
11	2.78	1.022451													
12	0.54	-0.61619													
13	9.19	2.218116													
14	2.61	0.95935													
15	2.16	0.770108													
16	0.82	-0.19845													
17	0.92	-0.08338													
18	9	2.197225													
19	1.86	0.620576													
20	5.79	1.756132													
21	1.653	0.502592													
22	8.3	2.116256													
23	8.94	2.190536													
24	2.35	0.854415													
25	0.71	-0.34249													

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

The result is as follows.



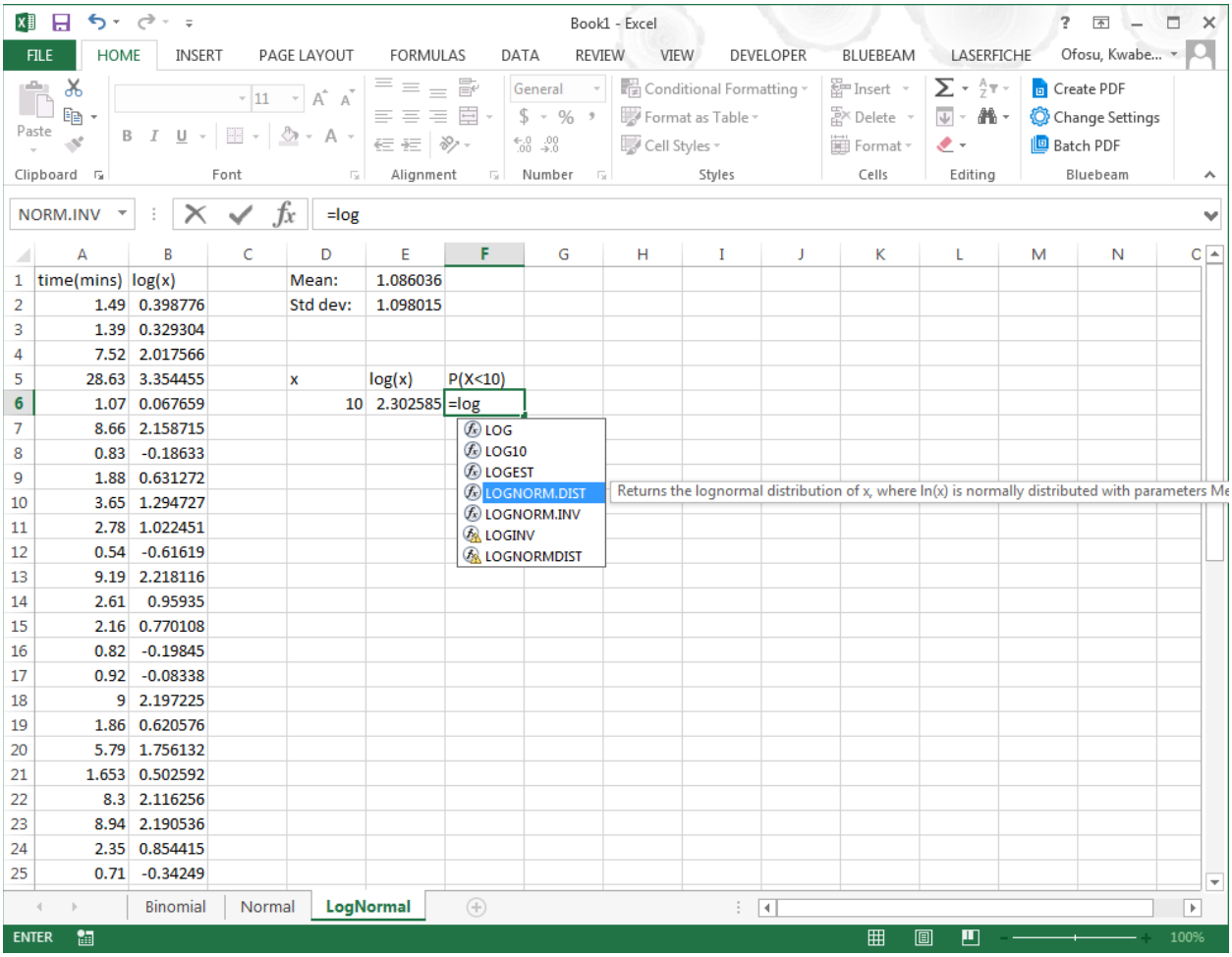
	A	B	C	D	E
1	time(mins)	log(x)		Mean:	1.086036
2	1.49	0.398776		Std dev:	1.098015
3	1.39	0.329304			
4	7.52	2.017566			
5	28.63	3.354455			
6	1.07	0.067659			
7	8.66	2.158715			
8	0.83	-0.18633			
9	1.88	0.631272			
10	3.65	1.294727			
11	2.78	1.022451			
12	0.54	-0.61619			
13	9.19	2.218116			
14	2.61	0.95935			
15	2.16	0.770108			
16	0.82	-0.19845			
17	0.92	-0.08338			
18	9	2.197225			
19	1.86	0.620576			
20	5.79	1.756132			
21	1.653	0.502592			
22	8.3	2.116256			
23	8.94	2.190536			
24	2.35	0.854415			
25	0.71	-0.34249			

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Conduct a probability calculation.

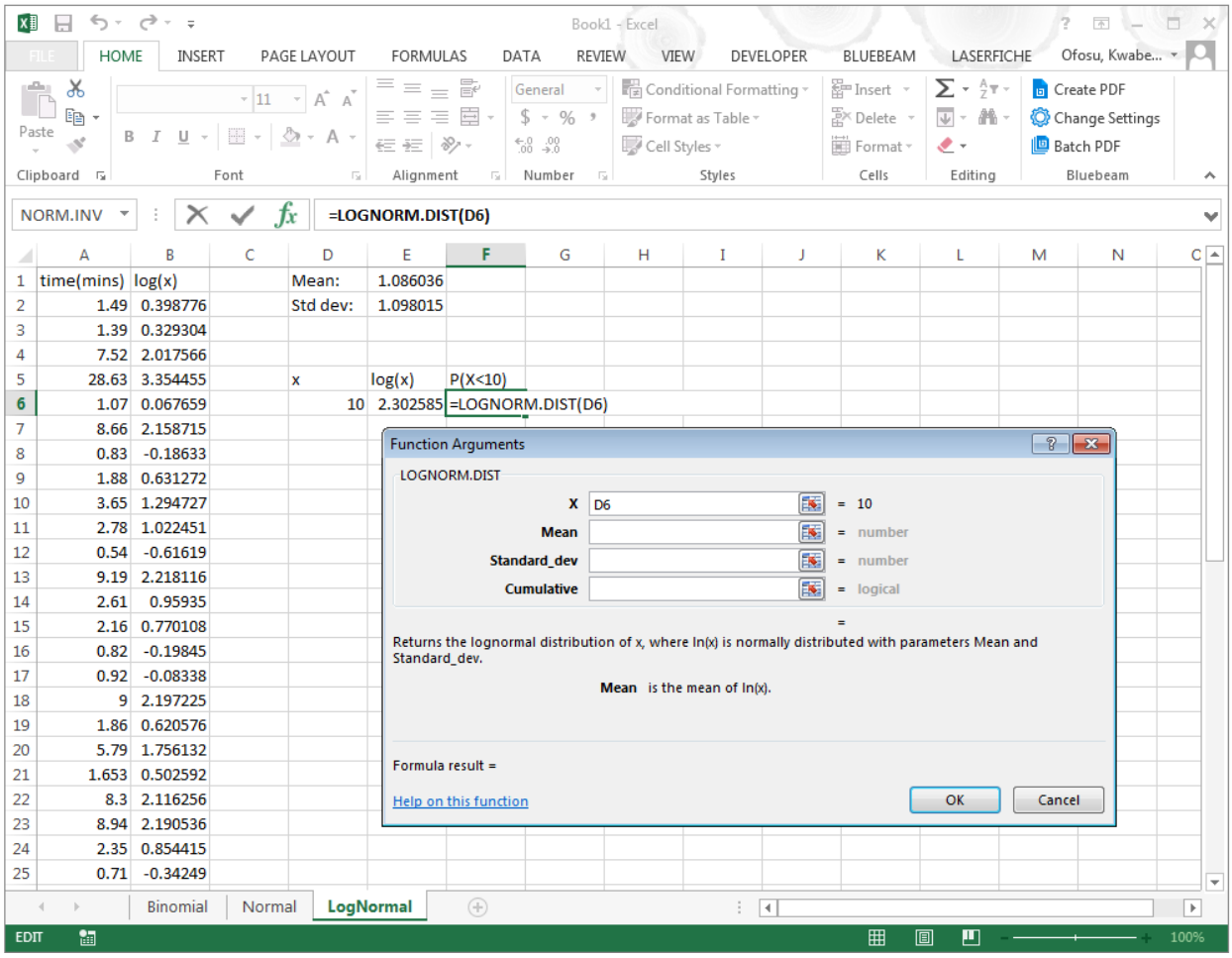
$$P(X < 10) = P(\log(x)) < \log(10)$$

$$P(\log(x)) < 2.303$$



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

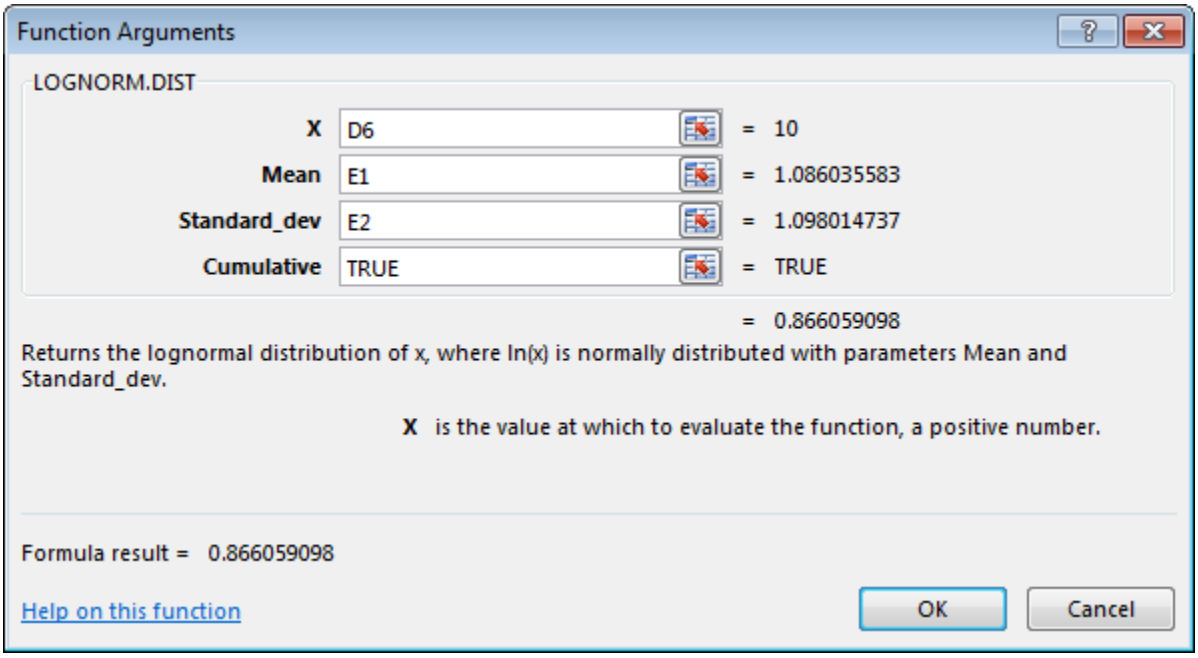
Or, alternately one may look up the Log-Normal function via the **Insert Function** icon.



	A	B	C	D	E	F
1	time(mins)	log(x)		Mean:	1.086036	
2	1.49	0.398776		Std dev:	1.098015	
3	1.39	0.329304				
4	7.52	2.017566				
5	28.63	3.354455		x	log(x)	P(X<10)
6	1.07	0.067659		10	2.302585	=LOGNORM.DIST(D6)
7	8.66	2.158715				
8	0.83	-0.18633				
9	1.88	0.631272				
10	3.65	1.294727				
11	2.78	1.022451				
12	0.54	-0.61619				
13	9.19	2.218116				
14	2.61	0.95935				
15	2.16	0.770108				
16	0.82	-0.19845				
17	0.92	-0.08338				
18	9	2.197225				
19	1.86	0.620576				
20	5.79	1.756132				
21	1.653	0.502592				
22	8.3	2.116256				
23	8.94	2.190536				
24	2.35	0.854415				
25	0.71	-0.34249				

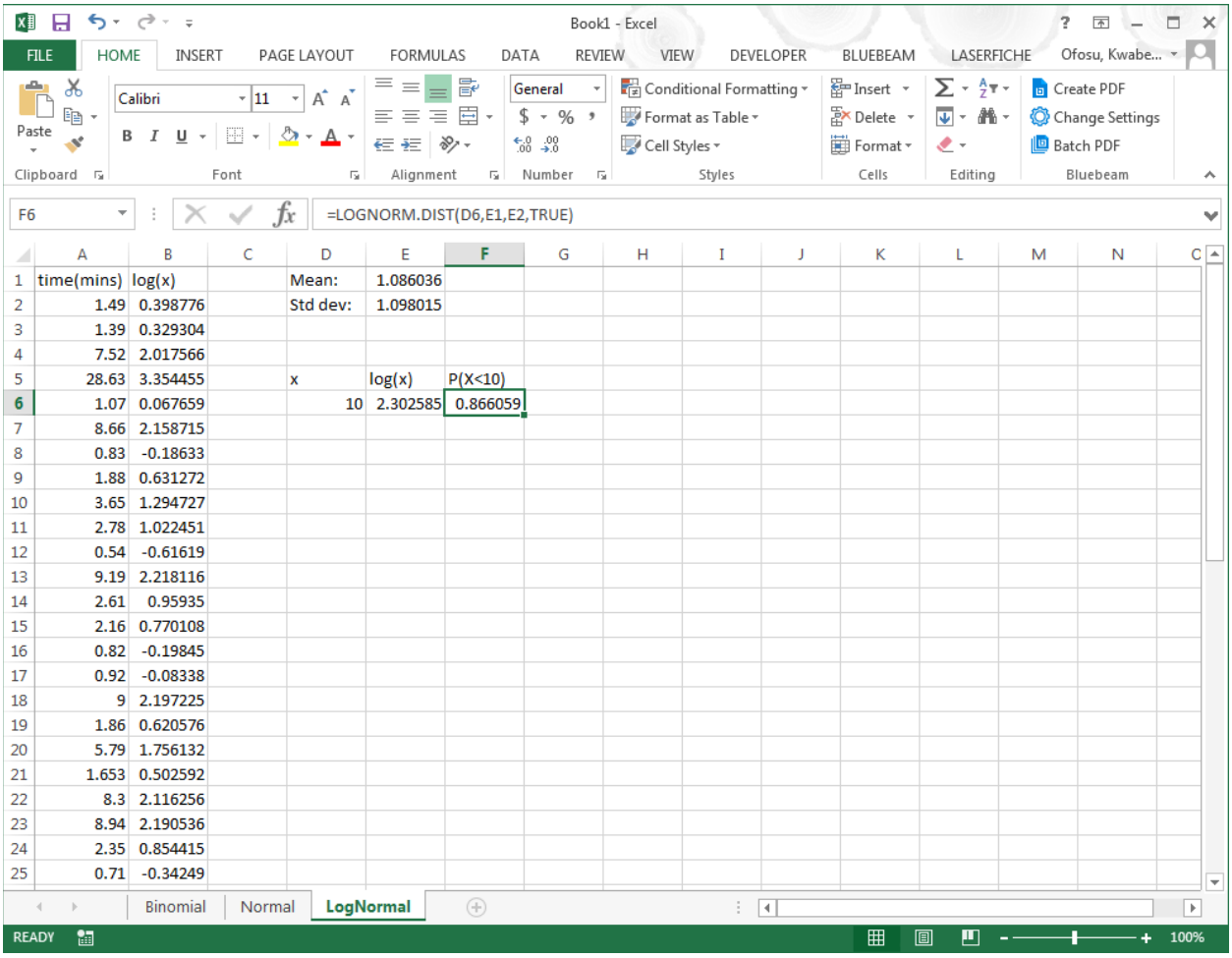
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Completing the arguments,



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

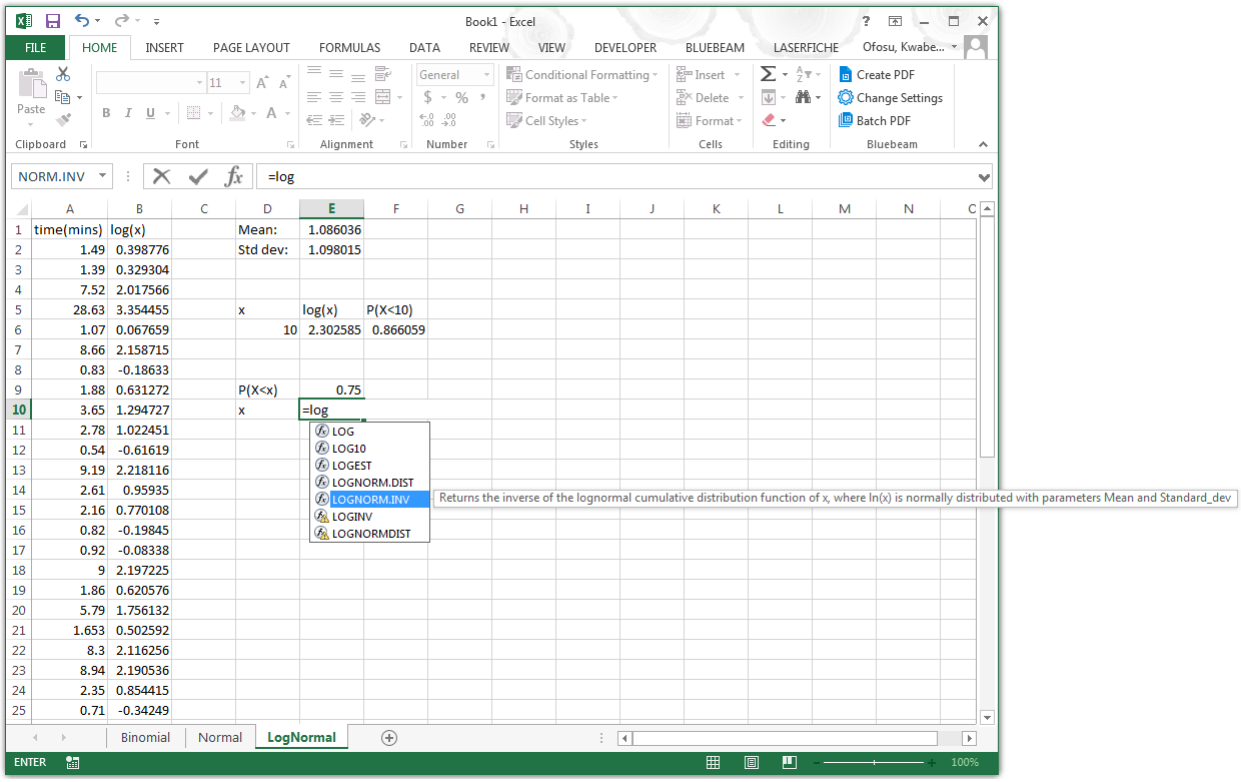
The result is displayed accordingly.



	A	B	C	D	E	F
1	time(mins)	log(x)		Mean:	1.086036	
2	1.49	0.398776		Std dev:	1.098015	
3	1.39	0.329304				
4	7.52	2.017566				
5	28.63	3.354455		x	log(x)	P(X<10)
6	1.07	0.067659		10	2.302585	0.866059
7	8.66	2.158715				
8	0.83	-0.18633				
9	1.88	0.631272				
10	3.65	1.294727				
11	2.78	1.022451				
12	0.54	-0.61619				
13	9.19	2.218116				
14	2.61	0.95935				
15	2.16	0.770108				
16	0.82	-0.19845				
17	0.92	-0.08338				
18	9	2.197225				
19	1.86	0.620576				
20	5.79	1.756132				
21	1.653	0.502592				
22	8.3	2.116256				
23	8.94	2.190536				
24	2.35	0.854415				
25	0.71	-0.34249				

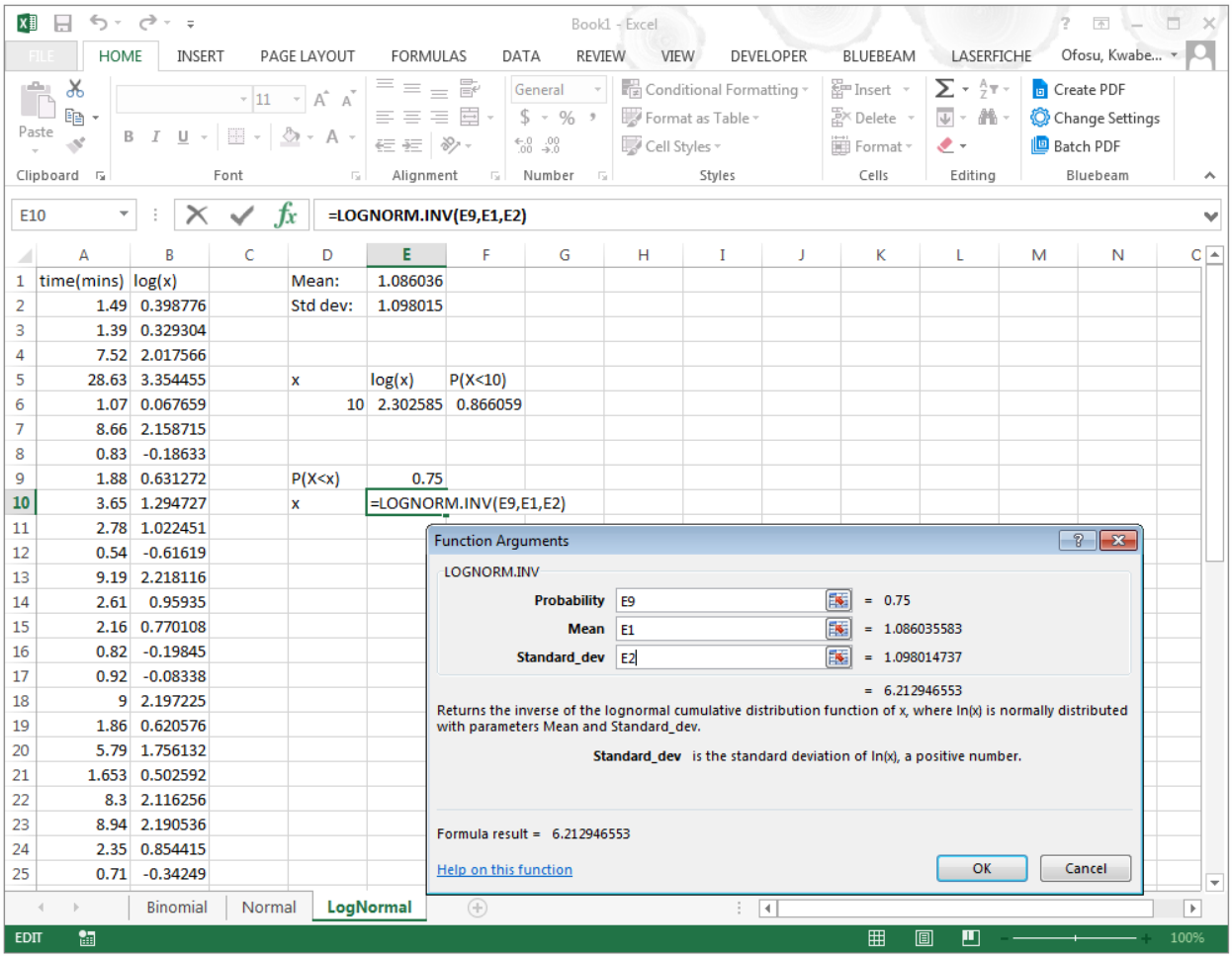
 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Next, use the inverse cumulative density function (the quantile function) to determine the value that has a cumulative probability of 0.75, otherwise known as the 75th percentile.



 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

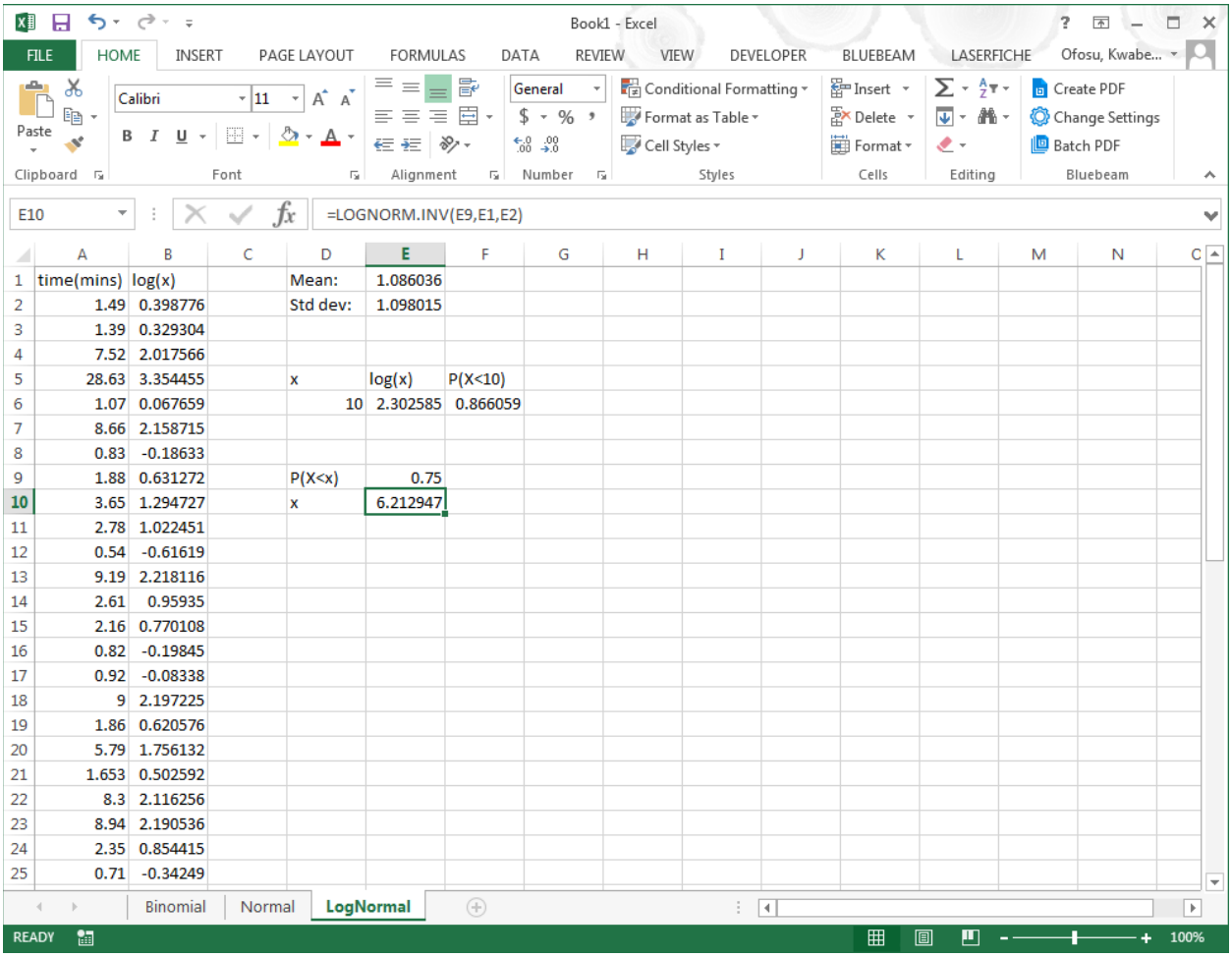
Going through the **Inset Function** icon,



	A	B	C	D	E	F
1	time(mins)	log(x)		Mean:	1.086036	
2	1.49	0.398776		Std dev:	1.098015	
3	1.39	0.329304				
4	7.52	2.017566				
5	28.63	3.354455		x	log(x)	P(X<10)
6	1.07	0.067659		10	2.302585	0.866059
7	8.66	2.158715				
8	0.83	-0.18633				
9	1.88	0.631272		P(X<x)	0.75	
10	3.65	1.294727		x	=LOGNORM.INV(E9,E1,E2)	
11	2.78	1.022451				
12	0.54	-0.61619				
13	9.19	2.218116				
14	2.61	0.95935				
15	2.16	0.770108				
16	0.82	-0.19845				
17	0.92	-0.08338				
18	9	2.197225				
19	1.86	0.620576				
20	5.79	1.756132				
21	1.653	0.502592				
22	8.3	2.116256				
23	8.94	2.190536				
24	2.35	0.854415				
25	0.71	-0.34249				

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

Yielding the result,



	A	B	C	D	E	F
1	time(mins)	log(x)		Mean:	1.086036	
2	1.49	0.398776		Std dev:	1.098015	
3	1.39	0.329304				
4	7.52	2.017566				
5	28.63	3.354455		x	log(x)	P(X<10)
6	1.07	0.067659		10	2.302585	0.866059
7	8.66	2.158715				
8	0.83	-0.18633				
9	1.88	0.631272		P(X<x)	0.75	
10	3.65	1.294727		x	6.212947	
11	2.78	1.022451				
12	0.54	-0.61619				
13	9.19	2.218116				
14	2.61	0.95935				
15	2.16	0.770108				
16	0.82	-0.19845				
17	0.92	-0.08338				
18	9	2.197225				
19	1.86	0.620576				
20	5.79	1.756132				
21	1.653	0.502592				
22	8.3	2.116256				
23	8.94	2.190536				
24	2.35	0.854415				
25	0.71	-0.34249				

Successful completion.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

5. CONCLUSION

This course presented an overview of Monte Carlo simulation, and a review of random variables and statistical distributions, and how they can be implemented in *Excel*.

This course began with an introduction to simulation problems encountered by engineers in various fields. The fundamental principles of simulation as well as the structure of simulation problems were then presented. A review of random variables followed, in which the discrete random variables and the continuous random variables were presented in detail. Upon completion of this course, participants will have gained skills in statistical distributions, and will be able to apply theses skills in the simulation and modeling of a real engineering system.

This course has enabled participants to identify professional situations where the innovative application of techniques learned in this course are relevant and will be of benefit to their productivity, efficiency, and the quality of their work product. Practitioners are strongly encouraged to look out for situations in their domains of expertise where random variables, simulation and systems modeling are applicable and will be of benefit to their work product and to their organization.

A successful application of engineering methods in *Excel* requires a careful and meticulous approach, and can only be mastered and retained by practice and repetition. It has been my utmost pleasure presenting this topic to you. Thank you.

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE

REFERENCES

Johnson, R. A. (2011). *Miller and Freund's Probability and Statistics for Engineers* (Eighth ed.).

Larson, R., & Hostetler, R. (2007). *Precalculus.* Houghton Mifflin Company.

May, A. D. (1990). *Traffic Flow Fundamentals.*

Meeker, W. Q., & Escobar, L. A. (1998). *Statistical Methods for Reliability Data.* John Wiley & Sons.

The National Instititute of Standards and Testing. (2013). NIST/SEMATECH e-Handbook of Statistical Methods. Retrieved November 2019, from http://www.itl.nist.gov/div898/handbook/

The National Institute of Standards and Testing (NIST). (2011). Probability Distributions. Retrieved December 2019, from https://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/probdist.htm

Wolfram MathWorld. (2019). Statistical Distributions. Retrieved November 2019, from http://mathworld.wolfram.com/topics/StatisticalDistributions.html

Images were all drawn/ prepared by Kwabena. Ofosu

 Copyright© 2020 Kwabena Ofosu, Ph.D., P.E., PTOE