ICIC Express Letters ICIC International ©2017 ISSN 1881-803X
Volume **11**, Number **9**, September **2017** pp. 1385–1392

A DIFFERENT SOLUTION FOR THE BINOMIAL DISTRIBUTION WITH APPLICATION TO ACCOUNTING

Delia Imelda Rivera Loredo, Arnulfo Luévanos Rojas*
Sandra López Chavarría and Manuel Medina Elizondo

Facultad de Contaduría y Administración
Universidad Autónoma de Coahuila
Blvd. Revolución 151 Ote. CP 27000, Torreón, Coahuila, México
*Corresponding author: arnulfol_2007@hotmail.com

Received March 2017; accepted May 2017

Abstract. *This paper proposes a different solution for the binomial probability distribution by means of the Derive software with application to an accounting example, which is the main part of this research. The classical solution is obtained through tables that present several books. The use of tables has the following drawbacks: 1) decimals are restricted for several digits; 2) in some cases the probability is located between two values and generally these values are interpolated; 3) at other times the probability is not located in tables. Therefore, the solution proposed in this paper is more accurate and also any value can be obtained using the software.*
Keywords: A different solution, Binomial probability distribution, Application to accounting

1. **Introduction.** In probability theory and statistics, the binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent experiments, each asking a yes-no question, and each with its own Boolean-valued outcome: a random variable containing single bit of information: success/yes/true/one (with probability p) or failure/no/false/zero (with probability $q = 1 - p$). A single success/failure experiment is also called a Bernoulli trial or Bernoulli experiment and a sequence of outcomes is called a Bernoulli process; for a single trial, i.e., $n = 1$, the binomial distribution is a Bernoulli distribution. The binomial distribution is the basis for the popular binomial test of statistical significance [1-4].

Newell and MacFarlane presented an application of the binomial distribution in which the distribution is used to detect differences between the sensory properties of food products. Included is a BASIC computer program listing used to generate triangle and duo-trio test results [5].

Ling discussed a different type of binomial distribution of order k. The mean, the variance and the moment generating function of this distribution are derived [6].

Makri and Philippou consider the Bernoulli case the number of l-overlapping success runs of length k in trials and two exact formulas are derived for its probability distribution function in terms of multinomial and binomial coefficients respectively [7].

Moteria and Ghodasara discussed a novel approach of a cluster to find out similarity or dissimilarity in binary variables using binomial probability distribution [8].

Moteria and Ghodasara presented a novel approach to find out similarity or dissimilarity in categorical variables by contingency table and ratio of mismatches using binomial probability distribution [9].

Hong presents a simple derivation for an exact formula with a closed-form expression for the cumulative distribution function (cdf) of the Poisson binomial distribution. The derivation uses the discrete Fourier transform of the characteristic function of the distribution. Numerical studies were conducted to study the accuracy of the developed algorithm and approximation methods [10].

Yuen et al. introduced a discrete-time self-exciting threshold binomial model to price derivative securities, and the key idea is to incorporate the regime switching effect in a discrete-time binomial model for an asset's prices via the "self-exciting" threshold principle. This model provides a simple structure for pricing options in a changing economic environment [11].

Bao et al. presented an inflated-parameter binomial distribution, which is a generalization of the classical binomial distribution, and also shows that there exists exactly one renewal process such that the number of renewals has an inflated-parameter binomial distribution [12].

Shmueli et al. in the year 2005 introduced the COM-Poisson-binomial distribution, but they did not study the mathematical properties of this family of distributions. Borges et al. discussed some properties and an asymptotic approximation of it by the COM-Poisson distribution. Moreover, three datasets are also considered [13].

Etukudo and Usen proposed that a theoretical idea of a six-sided die binomial experiment, in the theory of probability distributions is examined in this paper. This idea is extended and generalized in the postulate of an n-sided die binomial experiment in which there are n finite possible outcomes of such die toss, for m-trials with equal probabilities p and $1-p$ of success and failure, respectively, for each trial [14].

The binomial distribution may be applicable in the following scenarios [15]:

- The number of heads/tails in a sequence of coin flips,
- Vote counts for two different candidates in an election,
- The number of male/female employees in a company,
- The number of accounts that are in compliance or not in compliance with an accounting procedure,
- The number of successful sales calls,
- The number of defective products in a production cycle,
- The number of days in a month your company's computer network experiences a problem.

There is a set of assumptions which, if valid, would lead to a binomial distribution. These are [10]:

- A set of n experiments or trials are conducted;
- Each trial could result in either a success or a failure;
- The probability p of success is the same for all trials;
- The outcomes of different trials are independent;
- We are interested in the total number of successes in these n trials.

This paper proposes to use the DERIVE software to obtain the solution for the binomial probability distribution applying it to an accounting problem. The classical solution is obtained through tables that present several books. The use of tables has the following drawbacks: 1) decimals are restricted for several digits; 2) in some cases the probability is located between two values and generally these values are interpolated or take the nearest of these two; 3) at other times the probability is not located in tables.

The paper is organized as follows. Section 2 describes the general concepts of the binomial distribution that shows the equations of the binomial probability distribution function and the cumulative binomial probability distribution function, also the tables and figures showing the behavior of these functions. Section 3 presents an application for a company in the accounting department that has past due portfolio. Conclusion (Section 4) completes the paper.

2. **Binomial Distribution.** The binomial distribution is obtained by making "n" tests of Bernoulli independent of each other, in which case it has the following characteristics.

1) n: number of independent replicates of the Bernoulli experiment.

2) All tests must have a constant probability of success "p" and a constant probability of failure "$q = 1 - p$".

3) X: the number of successes in the n tests; $n - X$: failures number.

The binomial probability distribution function has the following form:

$$f(x) = C_x^n p^x (1-p)^{n-x} \quad (1)$$

where $x = 0, 1, 2, 3, \ldots, n$.

The cumulative binomial probability distribution function is:

$$F(x) = P(X \leq x) = \sum_0^x C_x^n p^x (1-p)^{n-x} \quad (2)$$

The arithmetic mean "μ" of a random variable is:

$$\mu = np \quad (3)$$

The variance "σ" is obtained as follows:

$$\sigma = npq = np(1-p) \quad (4)$$

The binomial distribution is symmetric when $p = 0.5$; otherwise, it is asymmetrical to the left or to the right, depending on whether the value of p is less than or greater than 0.5.

The binomial probability distribution function $f(x)$ (for $n = 20$) is presented in the following equation:

$$f(x) = C_x^{20} p^x (1-p)^{20-x} = \frac{20!}{(20-x)!x!} p^x (1-p)^{20-x} \quad (5)$$

where $x = 0, 1, 2, 3, \ldots, 20$.

The cumulative binomial probability distribution function $F(x)$ (for $n = 20$) is presented in the following equation:

$$F(x) = P(X \leq x) = \sum_0^x C_x^{20} p^x (1-p)^{20-x} = \sum_0^x \frac{20!}{(20-x)!x!} p^x (1-p)^{20-x} \quad (6)$$

Table 1 presents the values of the binomial probability distribution function $f(x)$ for $n = 20$, and $p = 0.1$, 0.3, 0.5, 0.7 and 0.9.

Figure 1 shows the binomial probability distribution function $f(x)$ for $n = 20$, and $p = 0.1$, 0.3, 0.5, 0.7 and 0.9.

Table 2 presents the values of the cumulative binomial probability distribution function $F(x)$ for $n = 20$, and $p = 0.1$, 0.3, 0.5, 0.7 and 0.9.

Figure 2 shows the cumulative binomial probability distribution function $F(x)$ for $n = 20$, and $p = 0.1$, 0.3, 0.5, 0.7 and 0.9.

3. **Application.** It is known that 20% of a company's portfolio is past due, and a random sample of 15 accounts is taken. You want to know the following:

1) What is the probability that there are four or fewer overdue accounts?

2) What is the probability that there are fewer than four overdue accounts?

3) What is the probability that there are more than two overdue accounts?

4) What is the probability that there are more than two but less than five overdue accounts?

5) What is the probability that there are exactly 3 overdue accounts?

6) What is the probability that there are no overdue accounts?

7) What is the arithmetic mean of overdue accounts?

TABLE 1. Binomial probability distribution function

x	$f(x)$				
	p				
	0.1	0.3	0.5	0.7	0.9
0	0.12158	0.00080	9.54×10^{-7}	3.49×10^{-11}	1.00×10^{-20}
1	0.27017	0.00684	1.91×10^{-5}	1.63×10^{-9}	1.80×10^{-18}
2	0.28518	0.02785	0.00018	3.61×10^{-8}	1.54×10^{-16}
3	0.19012	0.07160	0.00109	5.05×10^{-7}	8.31×10^{-15}
4	0.08978	0.13042	0.00462	5.01×10^{-6}	3.18×10^{-13}
5	0.03192	0.17886	0.01479	3.74×10^{-5}	9.15×10^{-12}
6	0.00887	0.19164	0.03696	0.00022	2.06×10^{-10}
7	0.00197	0.16426	0.07393	0.00102	3.71×10^{-9}
8	0.00036	0.11440	0.12013	0.00386	5.42×10^{-8}
9	5.23×10^{-5}	0.06537	0.16018	0.01201	6.50×10^{-7}
10	6.44×10^{-6}	0.03082	0.17620	0.03082	6.44×10^{-6}
11	6.50×10^{-7}	0.01201	0.16018	0.06537	5.23×10^{-5}
12	5.42×10^{-8}	0.00386	0.12013	0.11440	0.00036
13	3.71×10^{-9}	0.00102	0.07393	0.16426	0.00197
14	2.06×10^{-10}	0.00022	0.03696	0.19164	0.00887
15	9.15×10^{-12}	3.74×10^{-5}	0.01479	0.17886	0.03192
16	3.18×10^{-13}	5.01×10^{-6}	0.00462	0.13042	0.08978
17	8.31×10^{-15}	5.05×10^{-7}	0.00109	0.07160	0.19012
18	1.54×10^{-16}	3.61×10^{-8}	0.00018	0.02785	0.28518
19	1.80×10^{-18}	1.63×10^{-9}	1.91×10^{-5}	0.00684	0.27017
20	1.00×10^{-20}	3.49×10^{-11}	9.54×10^{-7}	0.00080	0.12158

8) What is the standard deviation for the number of overdue accounts?

3.1. Classical solution (Tables).

X: number of overdue accounts.

Success: overdue accounts.

Probability of success: $p = 0.2$

Number of tests $n = 15$

1) The probability of having four or fewer overdue accounts is obtained by $p(X \leq 4)$ from the tables of binomial distributions, $b \sim (15,\ 0.2)$, for $n = 15$ in the intersection $x = 4$ and $p = 0.2$, and we obtain $p(X \leq 4) = 0.836$ (see Table [16]).

2) The probability that there are fewer than four overdue accounts is obtained by $p(X < 4) = p(X \leq 3)$ from the tables of binomial distributions, $b \sim (15, 0.2)$, for $n = 15$ in the intersection $x = 3$ and $p = 0.2$, and we obtain $p(X < 4) = p(X \leq 3) = 0.648$ (see Table [16]).

3) The probability of having more than two overdue accounts is obtained by $p(X > 2) = 1 - p(X \leq 2)$ from the tables of binomial distributions, $b \sim (15, 0.2)$, for $n = 15$ in the intersection $x = 2$ and $p = 0.2$, and we obtain $p(X > 2) = 1 - p(X \leq 2) = 1 - 0.398 = 0.602$ (see Table [16]).

4) The probability of having more than two but less than five overdue accounts is obtained by $p(2 < X < 5) = p(2 < X \leq 4) = p(X \leq 4) - p(X \leq 2)$ from the tables of

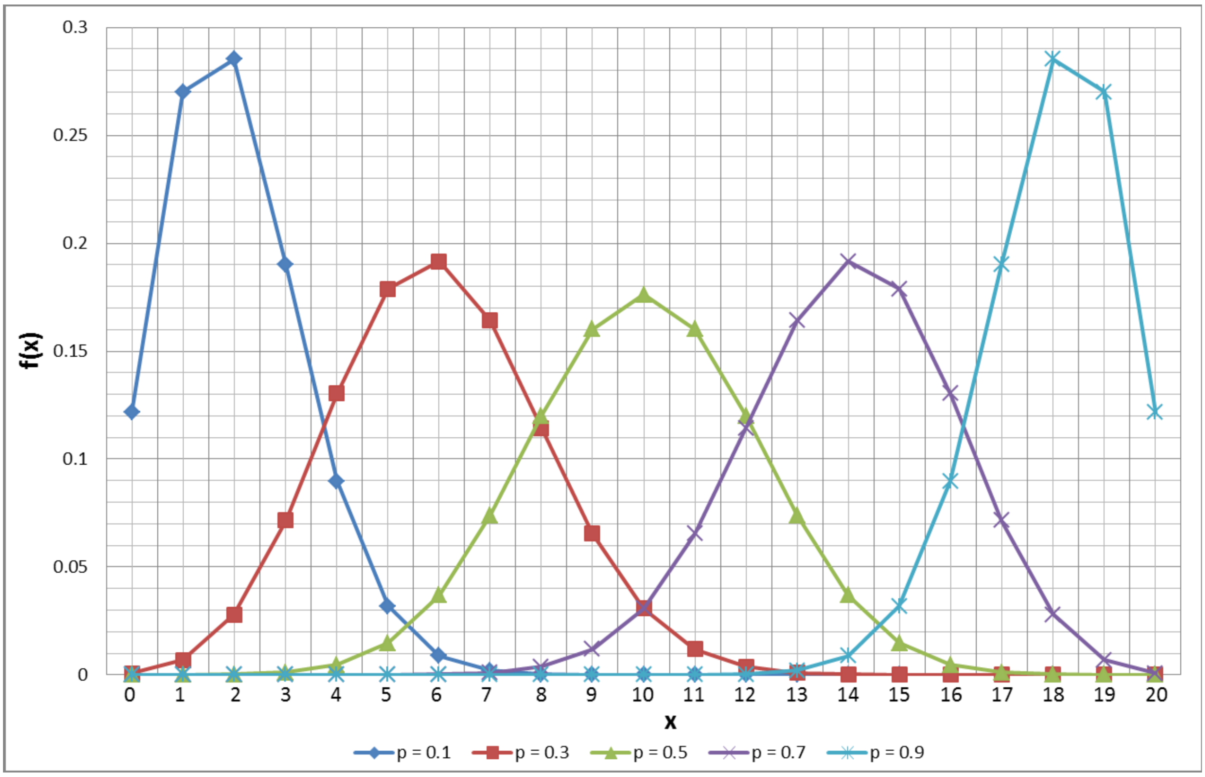

Figure 1. Binomial probability distribution function for $p = 0.1$, 0.3, 0.5, 0.7 and 0.9

Table 2. Cumulative binomial probability distribution function

x	$F(x)$				
	p				
	0.1	0.3	0.5	0.7	0.9
0	0.12157	0.00080	0.00000	0.00000	0.00000
1	0.39174	0.00764	0.00002	0.00000	0.00000
2	0.67692	0.03548	0.00020	0.00000	0.00000
3	0.86704	0.10708	0.00129	0.00000	0.00000
4	0.95682	0.23750	0.00591	0.00001	0.00000
5	0.98874	0.41637	0.02069	0.00004	0.00000
6	0.99761	0.60800	0.05766	0.00026	0.00000
7	0.99958	0.77227	0.13158	0.00128	0.00000
8	0.99994	0.88666	0.25172	0.00514	0.00000
9	1.00000	0.95203	0.41190	0.01714	0.00000
10	1.00000	0.98285	0.58809	0.04796	0.00001
11	1.00000	0.99486	0.74827	0.11333	0.00006
12	1.00000	0.99872	0.86841	0.22772	0.00042
13	1.00000	0.99973	0.94234	0.39199	0.00239
14	1.00000	0.99995	0.97930	0.58362	0.01125
15	1.00000	1.00000	0.99409	0.76249	0.04317
16	1.00000	1.00000	0.99871	0.89291	0.13295
17	1.00000	1.00000	0.99979	0.96451	0.32307
18	1.00000	1.00000	0.99997	0.99236	0.60825
19	1.00000	1.00000	1.00000	0.99920	0.87842
20	1.00000	1.00000	1.00000	1.00000	1.00000

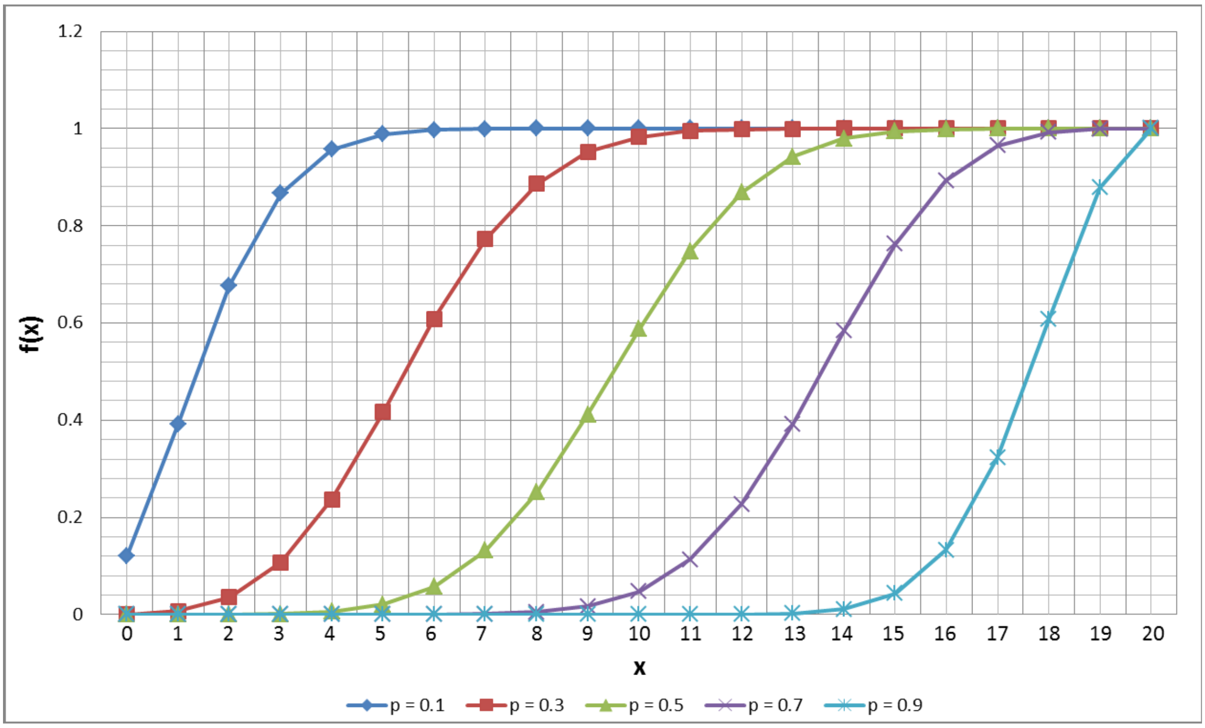

FIGURE 2. Cumulative binomial probability distribution function for $p =$ 0.1, 0.3, 0.5, 0.7 and 0.9

binomial distributions, $b \sim (15, 0.2)$, for $n = 15$ in the intersection $x = 4$ and $p = 0.2$; and for $x = 2$ and $p = 0.2$, and we obtain $p(X \leq 4) = 0.836$; $p(X \leq 2) = 0.398$, and $p(2 < X < 5) = p(2 < X \leq 4) = p(X \leq 4) - p(X \leq 2) = 0.836 - 0.398 = 0.438$ (see Table [16]).

5) The probability of exactly 3 overdue accounts is given by $p(X = 3) = p(X \leq 3) - p(X \leq 2)$ from the tables of binomial distributions, $b \sim (15, 0.2)$, for $n = 15$ in the intersection $x = 3$ and $p = 0.2$; and for $x = 2$ and $p = 0.2$, and we obtain $p(X \leq 3) = 0.648$; $p(X \leq 2) = 0.398$, and $p(X = 3) = p(X \leq 3) - p(X \leq 2) = 0.648 - 0.398 = 0.250$ (see Table [16]).

6) The probability of non-overdue accounts is given by $p(X = 0)$ from the tables of binomial distributions, $b \sim (15, 0.2)$, for $n = 15$ in the intersection $x = 0$ and $p = 0.2$, and we obtain $p(X \leq 0) = 0.035$ (see Table [16]).

7) The arithmetic mean of overdue accounts is obtained by $\mu = n = 15(0.2) = 3$.

8) The standard deviation for the number of overdue accounts is obtained. First we calculate the variance $\sigma^2 = np(1-p) = 15(0.2)(1-0.2) = 2.4$. Then the standard deviation is $\sigma = \sqrt{np(1-p)} = \sqrt{15(0.2)(1-0.2)} = 1.549$.

3.2. **Proposed solution (Derive software).** Another way to get the solution is by using software as in this case the results were calculated using the Derive software. The results are shown in Figure 3.

1) The probability of $p(X \leq 4) = 0.8357662760$ (see Equation #5 in Figure 3).

2) The probability of $p(X < 4) = p(X \leq 3) = 0.6481621045$ (see Equation #7 in Figure 3).

3) The probability of $p(X > 2) = 1 - p(X \leq 2) = 1 - 0.3980232092 = 0.6019767908$ (see Equation #9 in Figure 3).

4) The probability of $p(2 < X < 5) = p(2 < X \leq 4) = p(X \leq 4) - p(X \leq 2) = 0.8357662760 - 0.3980232092 = 0.4377430668$ (see Equations #5 and #9 in Figure 3).

5) The probability of $p(X = 3) = p(X \leq 3) - p(X \leq 2) = 0.6481621045 - 0.3980232092 = 0.2501388953$ (see Equations #7 and #9 in Figure 3).

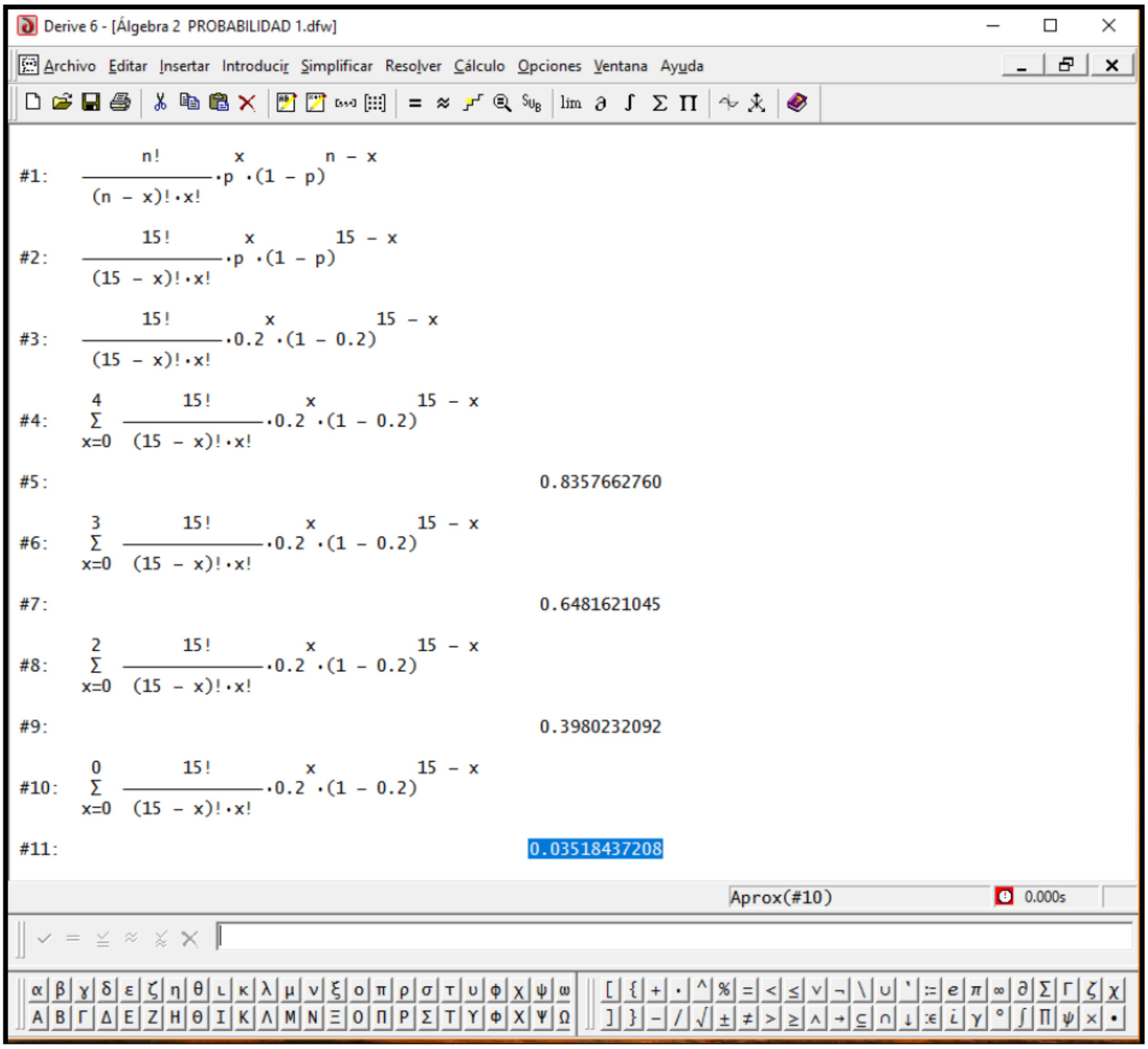

FIGURE 3. Input data and problem results

6) The probability of $p(X \leq 0) = 0.03518437208$ (see Equation #11 in Figure 3).

7) The arithmetic mean of overdue accounts is obtained by $\mu = n = 15(0.2) = 3$.

8) The standard deviation for the number of overdue accounts is obtained. First we calculate the variance $\sigma^2 = np(1-p) = 15(0.2)(1-0.2) = 2.4$. And the standard deviation is $\sigma = \sqrt{np(1-p)} = \sqrt{15(0.2)(1-0.2)} = 1.549193338$.

4. Conclusions.

This paper shows a solution using the Derive software to obtain the binomial probability distribution function and the cumulative binomial probability distribution function with application example to accounting. The application is presented for a company in the accounting department that has past due portfolio.

The classical solution is obtained through tables that present several books. The use of tables has the following drawbacks: 1) decimals are restricted for several digits; 2) in some cases the probability is located between two values and generally these values are interpolated; 3) at other times the probability is not located in tables.

Besides the efficiency and accuracy of the proposed solution presented in this investigation, the significant advantages are: 1) the decimals are not restricted for several digits; 2) the probability found is more accurate and also any value can be obtained by the proposed solution.

Therefore, the solution proposed in this paper is more accurate and also any value can be obtained using the software.

REFERENCES

[1] R. L. Moy, *Study Guide for Statistics for Business and Financial Economics*, St. John's University, New York, 1999.

[2] F. J. Martín Pliego, *Introducción a la Estadística Económica y Empresarial*, Thomson, Madrid, 2004.

[3] T. K. Tiemann, *Introductory Business Statistics*, The Global Text Project is Funded by the Jacobs Foundation, Zurich, Switzerland, 2010.

[4] J. M. Sarabia Alegría and M. Pascual Sáez, *Curso básico de estadística para los grados en economía y administración y dirección de empresas*, Universidad de Cantabria, Santander, Cantabria, España, 2012.

[5] G. J. Newell and J. D. MacFarlane, An interesting application of the binomial distribution, *The Australian Mathematics Teacher*, vol.40, no.1, pp.1-13, 1984.

[6] K. D. Ling, On binomial distributions of order k, *Statistics & Probability Letters*, vol.6, pp.247-250, 1988.

[7] F. S. Makri and A. N. Philippou, On binomial and circular binomial distributions of order k for l-overlapping success runs of length k, *Statistical Papers*, vol.46, no.3, pp.411-432, 2005.

[8] P. M. Moteria and Y. R. Ghodasara, Application of binomial probability distribution in cluster analysis of similar binary variables, *International Journal of Emerging Technology and Advanced Engineering*, vol.2, no.4, pp.84-86, 2012.

[9] P. M. Moteria and Y. R. Ghodasara, Application of binomial probability distribution in cluster analysis of similar categorical variables, *International Journal of Emerging Technology and Advanced Engineering*, vol.2, no.6, pp.133-135, 2012.

[10] Y. Hong, On computing the distribution function for the Poisson binomial distribution, *Computational Statistics & Data Analysis*, vol.59, pp.41-51, 2013.

[11] F. L. Yuen, T. K. Siu and H. Yang, Option valuation by a self-exciting threshold binomial model, *Mathematical and Computer Modelling*, vol.58, nos.1-2, pp.28-37, 2013.

[12] Z. Bao, L. Song and H. Liu, A note on the inflated-parameter binomial distribution, *Statistics & Probability Letters*, vol.83, no.8, pp.1911-1914, 2013.

[13] P. Borges, J. Rodrigues, N. Balakrishnan and J. Bazán, A COM-Poisson type generalization of the binomial distribution and its properties and applications, *Statistics & Probability Letters*, vol.87, pp.158-166, 2014.

[14] I. Etukudo and J. Usen, The generalized die binomial experiment, *American Scientific Research Journal for Engineering, Technology, & Sciences (ASRJETS)*, vol.15, no.1, pp.160-166, 2016.

[15] *http://www.utdallas.edu/scniu/OPRE-6301/documents/Important_Probability_Distributions.pdf.*

[16] *https://www.joomag.com/magazine/te-adoro-tesoroescondido/0598927001443040788?page=154.*