



Definite Probabilities from Division of Zero by Itself Perspective

Wangui Patrick Mwangi^{1*}

¹Department of Mathematics and Computer Science, University of Eldoret, Box 1125-30100, Kenya.

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/AJPAS/2020/v6i230155

Editor(s):

(1) Dr. Jiteng Jia, School of Mathematics and Statistics, Xidian University, China.

Reviewers:

(1) Awa Ogbonnaya Dike, Akanu Ibiam Federal Polytechnic, Nigeria.

(2) Zeraoulia Rafik, Batna 2 University, Algeria.

(3) Peter Stallinga, University of the Algarve, Portugal.

Complete Peer review History: <http://www.sdiarticle4.com/review-history/53235>

Received: 14 October 2019

Accepted: 19 December 2019

Published: 22 January 2020

Original Research Article

Abstract

Over the years, the issues surrounding the division of zero by itself remained a mystery until year 2018 when the mystery was solved in numerous ways. Afterwards, the same solutions provided opened many other doors in academic space and one of the applications is in sure probabilities. This research is all about the sure probabilities computed from the zero divided by itself point of view. The solutions obtained in the computations are in harmony with logic and basic knowledge. A wide range of already existing probability distribution functions has been applied in different scenarios to compute the sure probabilities unanimously and new findings have also been encountered along the way. Some of the discrete and continuous probability distribution functions involved are the binomial, hypergeometric, negative binomial, Poisson, normal and exponential among others. It has been found in this work that sure probabilities can be evaluated from the division of zero by itself perspective. Another new finding is that in case of combinatorial, if the numerator is smaller than the denominator, then the solutions tend to zero when knowledge in gamma functions, integrations and factorials is applied. Again, if the case of continuous pdf involves integration and random variable specified in the direction of the parameter, then indirect computation of such probabilities should be applied. Finally, it has been found that the expansion of the domains of some of the parameters in some existing probability distribution functions can be considered and the restriction in conditional probabilities can be revised.

Keywords: Definite probability; division of zero by itself; probability distribution function.

*Corresponding author: E-mail: patrickmwangi72@yahoo.com;

1 Introduction

Generally, probability is the percentage or proportion of times of our expectation of outcomes when it comes to happening of events; which is usually when the experiment has not yet been done [1]. Probability refers to the likelihood of the occurrence of an event and assigns a number between 0 *and* 1 to an event [2]. An example is when we have a coin being tossed once in which there are only 2 possible outcomes, head and tail, each with probability $\frac{1}{2}$ or 50%. Here, there is 50% chances of each event happening (50% chances of getting a tail and 50% chances of getting a head if the coin is tossed). The expected probabilities can change when the actual experiment is conducted. E.g. there may be 10 boys and 10 girls in a class and experiment may show that when some eight students are chosen at random with replacement, there is always 3 boys and 5 girls in the sample. This may lead to probability becoming 0.625 or 62.5% for girls from 50%. On the other hand, a random variable, say Y or X, is a real valued function defined on the sample space, S, of a random experiment; a random experiment is any experiment that can be performed in the same set of conditions repeatedly; a sample space is the set of all possible outcomes of a random experiment or elements of a set describing the outcomes of an experiment of interest and an event is a specific outcome that results when the experiment is performed, distinct outcomes in a set or subsets of a sample space S [1,2,3,4,5]. A sample point can be defined as each outcome in a sample space [6]. Most of the books used in probability and statistics have widely covered most of the axioms involved in probabilities such as the range of a probability value, addition and multiplication rules etc. [7].

2 Definite Probabilities

Definite probabilities are chances that are certain or sure. They are sure probabilities in the sense that one is certain whether an even can or cannot occur/happen. When one is sure that the even under question must occur, then the definite probability is 1. Otherwise, the definite probability is 0 and when such a case occurs, then we have an impossible event [6]. The two limits in chances of events are 0 and 1 and are defined in this work as the definite/sure/certain probabilities or chances.

2.1 Probability distribution functions

There are many probability distributions functions (pdfs) each corresponding to unique characteristics. Each distribution may have one or many parameters. The probability distributions are classified as either continuous or discrete. Examples of probability distributions are Bernoulli, Normal, Gamma, Binomial, Chi-Square, Binomial, Exponential, Negative Binomial, Poisson, Hypergeometric, geometric among others [4,5,8,9]. For the case of discrete pdfs, some of the functions involved in this work are:

The binomial distribution [3,8,10]:

$$f(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0, 1, 2, \dots, n \\ 0, & \text{otherwise} \end{cases}$$

The Bernoulli distribution [3,8,10]:

$$f(x) = \begin{cases} p^x q^{1-x}, & x = 0, 1 \\ 0, & \text{otherwise} \end{cases}$$

The Poisson distribution [2,3,9,10]:

$$f(x) = \begin{cases} \frac{e^{-\lambda} * \lambda^x}{x!}, & x = 0, 1, 2, 3, \dots, \lambda > 0 \\ 0, & \text{elsewhere} \end{cases}$$

The geometric distribution [3,8,10]:

$$f(x) = \begin{cases} pq^x, & x = 0, 1, 2, \dots \\ 0, & \text{otherwise} \end{cases} \quad \text{or} \quad f(x) = \begin{cases} pq^{x-1}, & x = 1, 2, 3, \dots \\ 0, & \text{otherwise} \end{cases}$$

The hypergeometric distribution [3,8,9,10]:

$$f(x) = \begin{cases} \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}}, & x = 0, 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

The negative binomial distribution [3,8,10]:

$$f(k) = \Pr(X_r = k) = \begin{cases} \binom{k-1}{r-1} p^r q^{k-r}, & k = r, r+1, r+2, \dots \\ 0, & \text{elsewhere} \end{cases}$$

The continuous distributions involved in this work are:

The normal distribution [3,8,9,10]:

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}; & \sigma > 0, -\infty < \mu, x < \infty \\ 0; & \text{otherwise} \end{cases}$$

The exponential distribution [3,8]:

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0, \beta > 0 \\ 0, & \text{otherwise} \end{cases}$$

The gamma distribution [3,8,11]:

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy; \quad \alpha > 0$$

There are many other distributions that have not been tackled in this work some of which are used widely in research and are encountered often. Such include the Chi-Square, F, Student t, multinomial, rectangular, bivariate, uniform, beta and Dirichlet distributions among others [12,13]. Given a pdf, it's possible to compute probabilities of specified ranges of the random variables involved [7,14,15]. For examples;

What is $\Pr(X = 3)$ when $p = 0.4$, $q = 1 - p = 0.6$ and $n = 15$ in a binomial distribution?

We can have $\Pr(X = 3) = f(3) = \binom{15}{3} 0.4^3 0.6^{15-3} = 0.0634$

In hypergeometric distribution, what is the probability of finding 4 green balls in a sample of size 12 balls from a bucket that has 20 yellow and 30 green balls?

We have $f(x = 4) = \frac{\binom{50-30}{12-4} \binom{30}{4}}{\binom{50}{12}} = 0.0284$.

Supposing $\sigma = 8$ and $\mu = 20$, what is $Pr(X < 6)$ in a normal distribution? In this case, $\Phi\left(\frac{6-20}{8}\right) = 1 - \Phi(1.75) = 1 - 0.960 = 0.04$.

2.2 Computations involving definite probabilities

The computations in this work widely involve the division of zero by itself, $\frac{0}{0}$, that can be represented in various forms as is indicated in [16] such as 0^0 , $\binom{0}{0}$ and $\frac{0!}{0!(0-0)!}$, whose answer has been determined in many ways to be 1 (i.e. $\frac{0}{0} = 0^0 = \binom{0}{0} = \frac{0!}{0!(0-0)!} = 1$). The work also applies the knowledge in limits.

Case 1:

Consider the case where the probability of a plastic cup becoming the sun is to be determined. Based on this, we can judge that the definite probability is 0, i.e. $p = 0$ while the definite probability of the plastic cup not becoming the sun is 1, i.e. $q = 1$ ($1 - p = 1 - 0 = 1$). By assuming the exponential distribution, then the average waiting time is ∞ i.e. $\beta = \infty$. This means that we can wait forever for the plastic cup to change into the sun. Let X be a random variable representing waiting time. The probability distribution function (pdf) of a random variable, say X , that follows exponential distribution is given by

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0, \beta > 0 \\ 0, & otherwise \end{cases}$$

Suppose we want to use this information to compute $Pr(X > 50)$ where X is the waiting time (continuous variable) for the plastic cup to turn into the sun. This is to say the probability that we will wait for more than 50-unit time for the plastic cup to become the sun (more than 50-unit time will pass before the plastic cup becomes the sun). We can approach this as follows:

$$\begin{aligned} Pr(X > 50) &= 1 - Pr(X < 50) = f(x > 50) = 1 - \int_0^{50} \frac{1}{\beta} e^{-x/\beta} dx = 1 - \int_0^{50} \frac{1}{\infty} e^{-x/\infty} dx \\ &= 1 - \left(\frac{1}{\infty} * \frac{1}{(-1/\infty)} \right) * (e^{-x/\infty})_0^{50} = 1 - \left(0 * \frac{1}{0} \right) * (e^{-50/\infty} - e^{-0/\infty}) \\ &= 1 + \left(\frac{0}{0} \right) * (e^0 - e^0) = 1 + 1(1 - 1) = 1 + 0 = 1. \end{aligned}$$

This means that the probability of waiting for more than 50-unit time for the change to occur is 1 which is in agreement with the definite probability that we must wait forever for the change to take place.

Now suppose that we want $Pr(X < 50)$. This is the probability of waiting for less than 50-unit time before the plastic cup changes into the sun. But from the definite probability information is that this is not possible because we must wait forever ($\beta = \infty$) for the change to happen ($p = 0$). Using the pdf of exponential distribution,

$$\begin{aligned} Pr(X < 50) &= \int_0^{50} \frac{1}{\beta} e^{-x/\beta} dx = \int_0^{50} \frac{1}{\infty} e^{-x/\infty} dx = \left(\frac{1}{\infty} * \frac{1}{(-1/\infty)} \right) * (e^{-x/\infty})_0^{50} \\ &= - \left(0 * \frac{1}{0} \right) * (e^{-50/\infty} - e^{-0/\infty}) = - \left(\frac{0}{0} \right) * (e^0 - e^0) = -(1) * (0) = 0. \end{aligned}$$

This means that the probability of waiting for less than 50-unit time for the change to occur is 0 which is in agreement with the definite probability.

A possible question can be, why not find $Pr(X > 50)$ directly without involving $1 - Pr(X < 50)$? To answer this question, first, one can try to integrate the pdf for exponential distribution with the limits $[50, \infty]$. You will discover that, the integral involves $\frac{\infty}{\infty}$ and this has not yet been solved adequately and convincingly. Secondly, [16] has mentioned that the cases that involve $\frac{0}{0}$ need to be evaluated carefully. The path to take when the scenario is encountered requires one to be very careful because diverse, awkward and contradicting results can be obtained depending on the path followed. To know the direction to take when working with continuous distributions' pdfs involving definite probabilities with integrations, one rule to observe is to ensure that the integral should not be carried out directly when X is specified to be in the same direction as the parameter such as β in this case (i.e. $Pr(X > k)$ and $\beta = \infty$). In this case, X is specified to be in the direction of β .

Case 2:

Consider the case where the probability of a plastic cup not becoming the sun is to be determined. We can easily judge that the definite probability is 1, i.e. $p = 1$ while the definite probability of the plastic cup becoming the sun is 0, i.e. $q = 0$ ($1 - p = 1 - 1 = 0$). By assuming the exponential distribution, then the average waiting time is 0 i.e. $\beta = 0$. This means that we can wait for 0-unit time for the plastic cup not to change into the sun. Here, it means that, just as we start to wait for the plastic cup not to change into sun, the change does not occur. We can use the exponential distribution and violate the condition that $\beta > 0$. We are violating the condition because during the discovery of these functions and distributions till year 2017, a case where $\frac{0}{0}$ occurs could not be defined. However, the $\frac{0}{0} = x$ and $0^0 = x$ have been previously defined in a wide range of methods in [16]. Let X be a random variable representing waiting time. If we are interested in $Pr(X < 50)$, then we have

$$\begin{aligned} Pr(X < 50) &= 1 - Pr(X > 50) = f(x < 50) = 1 - \int_{50}^{\infty} \frac{1}{\beta} e^{-x/\beta} dx = 1 - \int_{50}^{\infty} \frac{1}{0} e^{-x/0} dx \\ &= 1 - \frac{1}{0} * \frac{1}{(-1/0)} * (e^{-x/0})_{50}^{\infty} = 1 - \left(\frac{0}{0}\right) * (e^{-\infty/0} - e^{-50/0}) \\ &= 1 + 1(e^{-\infty} - e^{-\infty}) = 1 + (0 - 0) = 1 + 0 = 1. \end{aligned}$$

This means that the probability of waiting for less than 50-unit time for the change not to occur is 1 which is in agreement with the definite probability.

Now suppose that we want $Pr(X > 50)$. This is the probability of waiting for more that 50-unit time before the plastic cup doesn't change into the sun. But from the definite probability information is that this is not possible because we must wait for 0-unit time ($\beta = 0$) for the change not to happen ($q = 0$). Using the pdf of exponential distribution,

$$\begin{aligned} Pr(X > 50) &= f(x > 50) = \int_{50}^{\infty} \frac{1}{\beta} e^{-x/\beta} dx = 1 - \int_{50}^{\infty} \frac{1}{0} e^{-x/0} dx = \frac{1}{0} * \frac{1}{(-1/0)} * (e^{-x/0})_{50}^{\infty} \\ &= -\left(\frac{0}{0}\right) * (e^{-\infty/0} - e^{-50/0}) = -1(e^{-\infty} - e^{-\infty}) = -(0 - 0) = 0. \end{aligned}$$

This means that the probability of waiting for more than 50-unit time for the change not to occur is 0 which is in agreement with the definite probability.

A possible question can be, why not find $Pr(X < 50)$ directly without involving $1 - Pr(X > 50)$? Again, we answer like in case 1 that X is heading in the same direction as β (i.e. $Pr(X < 50)$ and $\beta = 0$) and

hence, computing it directly leads to awkward and contradicting answers. We therefore need to be careful in evaluating such probabilities with integrations. So, the general rule is to avoid direct computations when X is specified to be in the same direction as the parameter β .

Case 3:

Consider driving a car in the streets of the capital city of your country. Suppose your village is about 300 km away from the capital city and your house is in that village. What is the probability that the car you are driving in the capital city will cause an accident in your house in the village? Here, the definite probability is 0, i.e. $p = 0$ and the definite probability that that car won't cause an accident in your house is 1, i.e. $q = 1 = (1 - p)$. Suppose you are to perform ∞ trials (i.e. drive your car ∞ times where each drive has specified distance within the capital city) in the capital city. Performing the experiment as many times as possible would reveal that you can cause 0 accidents on average. Suppose we use a random variable X that follows the Poisson distribution to represent the number of accidents. Thus, the pdf of X is given by:

$$f(x) = \begin{cases} \frac{e^{-\lambda} * \lambda^x}{x!}, & x = 0, 1, 2, 3, \dots, \quad \lambda > 0 \\ 0, & \text{elsewhere} \end{cases}$$

We may want to find $Pr(X = 0)$ as well as $Pr(X > 0)$. Before we embark on solving the puzzle, we need to define the parameter λ from the information provided. Here, on average, you can cause 0 accidents in your house in the village while driving your car in the capital city. Repeating the experiment for so many number-of-times would always yield $\lambda = 0$. Having defined $\lambda = 0$, then we can apply the Poisson distribution function and violate the condition that $\lambda > 0$, because [16] defined $\frac{0}{0} = x$ and $0^0 = x$. For $Pr(X = 0)$, we have

$$Pr(X = 0) = f(0) = \frac{e^{-\lambda} * \lambda^0}{0!} = \frac{e^{-0} * 0^0}{0!} = \frac{e^0 * 0^0}{0!} = \frac{1 * 0^0}{1} = 0^0 = 1.$$

This is the probability of causing no accident in your house in the village while driving your car in the capital city and is in agreement with the definite probability.

For $Pr(X > 0)$, we can use the formula

$$Pr(X > 0) = 1 - Pr(X = 0) = 1 - 1 = 0.$$

This is the probability of causing an accident (or the probability of causing some accidents) in the house while in capital city driving.

Alternatively, we can use the Poisson pdf to find

$$Pr(X > 0) = Pr(x = 1) + Pr(X = 2) + \dots + Pr(X = \infty)$$

which would give us the probability of causing any number of accidents in the house that is in the village while driving in the capital city that is 300 km away.

$$Pr(X = 1) = f(1) = \frac{e^{-\lambda} * \lambda^1}{1!} = \frac{e^{-0} * 0^1}{1!} = \frac{e^0 * 0^1}{1!} = \frac{1 * 0}{1} = 0.$$

$$Pr(X = 2) = f(2) = \frac{e^{-\lambda} * \lambda^2}{2!} = \frac{e^{-0} * 0^2}{2!} = \frac{e^0 * 0^2}{2!} = \frac{1 * 0}{2} = 0.$$

$$Pr(X = 3) = f(3) = \frac{e^{-\lambda} * \lambda^3}{3!} = \frac{e^{-0} * 0^3}{3!} = \frac{e^0 * 0^3}{3!} = \frac{1 * 0}{6} = 0.$$

$$Pr(X = \infty) = f(\infty) = \frac{e^{-\lambda} * \lambda^{\infty}}{\infty!} = \frac{e^{-0} * 0^{\infty}}{\infty!} = \frac{e^0 * 0^{\infty}}{\infty!} = \frac{1 * 0}{\infty} = 0.$$

Therefore;

$$\Pr(X > 0) = \Pr(x = 1) + \Pr(X = 2) + \dots + \Pr(X = \infty) = 0 + 0 + 0 + \dots + 0 = 0.$$

which in agreement with the definite probability. This generally means that there is no way (it's not possible) you can cause an accident in the village that is about 300 km away from the capital city.

What about a case of $\Pr(X < 3)$? Here we can find the probability by just applying the above computed probabilities as follows:

$$\Pr(X < 3) = \Pr(x = 0) + \Pr(x = 1) + \Pr(X = 2) = 1 + 0 + 0 = 1.$$

The probability of causing less than three accidents in the village when driving in the capital city is 1. Here, we learn that as long as we specify the probability of causing number of accidents less than a given value, then the probability will always be 1 because 0 accidents is always included in that specification and it is the one that carries the probability 1 while the other numbers in the range carry probability 0. I.e. ($\Pr(X < 1) = 1$, $\Pr(X < 2) = 1$, $\Pr(X < 20) = 1$, ... $\Pr(X < 500) = 1$, *etc.* On the other hand, specifying the probability of number of accidents greater than a given value will always give probability 0. I.e. ($\Pr(X > 1) = 0$, $\Pr(X > 2) = 0$, $\Pr(X > 20) = 0$, ... $\Pr(X > 500) = 0$, *etc.*

Case 4:

Consider a case where right 'now' you are observing and recording whether or not you are alive at that time/moment. If you are observing and be able to record 'now', then the definite probability that you are alive at that time is 1, i.e. $p = 1$ while the definite probability that you are not alive at that time when making this observation and recording it 'now' is 0, i.e. $q = 0$. If you make 10 trials, then, this can be the binomial distribution with parameters $p = 1$ and $n = 10$. Let X be a random variable representing the number of recordings made 'now'. In case $n = 1$ (making 1 trial), then this becomes the Bernoulli distribution.

Starting with Bernoulli, the pdf of a random variable X following the Bernoulli distribution is given by:

$$f(x) = \begin{cases} p^x q^{1-x}, & x = 0, 1 \\ 0, & \text{otherwise} \end{cases}$$

If we are interested in $\Pr(X = 0)$ as well as $\Pr(X = 1)$, then we have;

$$\Pr(X = 0) = f(0) = p^0 * q^{1-0} = 1^0 * 0^{1-0} = 1 * 0^1 = 0.$$

This is the probability that you will not make the observation and recording 'now' that you are alive at that time and is in harmony with the definite probability. It also means the probability of making 0 observations and recordings 'now' is 0 when you have made 1 trial.

$$\Pr(X = 1) = f(1) = p^1 * q^{1-1} = 1^1 * 0^{1-1} = 1 * 0^0 = 0^0.$$

This is the probability that you will make the observation and recording 'now' that you are alive at that time and is in harmony with the definite probability. It also means that, the probability of making 1 observation and recording 'now' is 1 when you have made 1 trial. A general explanation will be made in the case of many trials- binomial.

Revisiting case 3, it is known that $p = 0$ is the definite probability of causing an accident in the house while $q = 1$ is the definite probability of not causing any accident in the house. Suppose you make 1 drive/trial. Then for $\Pr(X = 0)$ we get;

$$\Pr(X = 0) = f(0) = p^0 * q^{1-0} = 0^0 * 1^{1-0} = 0^0 * 1^1 = 0^0 = 1.$$

This gives the probability of causing no accident (or the probability of causing 0 accidents) in the house and agrees with the definite probability.

And for $Pr(X = 1)$ we get;

$$Pr(X = 1) = f(1) = p^1 * q^{1-1} = 0^1 * 1^{1-1} = 0 * 1^0 = 0.$$

This gives the probability of causing an accident (or the probability of causing 1 accident) in the house and agrees with the definite probability.

For binomial distribution, $n > 1$, the pdf of random variable X that follows the binomial distribution is given by:

$$f(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0, 1, 2, \dots, n \\ 0, & otherwise \end{cases}$$

For case 4 and $Pr(X = 0)$, we can compute the value as follows:

$$Pr(X = 0) = f(0) = \binom{n}{0} p^0 q^{n-0} = \binom{n}{0} 1^0 0^{n-0} = 1 * 1 * 0^n = 0, \quad \binom{n}{0} \geq 1.$$

This gives the probability that you won't make any observation and recording 'now' that you are alive at that time and agrees with the definite probability. It also means the probability of making 0 observations and recordings 'now' that you are alive at that time when you have made $n > 1$ trials.

For $Pr(X > 0)$, we can use the relation:

$$Pr(X > 0) = 1 - Pr(X = 0) = 1 - 0 = 1.$$

Alternatively,

$$\begin{aligned} Pr(X > 0) &= Pr(X = 1) + Pr(X = 2) + Pr(X = 3) + \dots + Pr(X = n). \\ Pr(X = 1) = f(1) &= \binom{n}{1} p^1 q^{n-1} = \binom{n}{1} 1^1 0^{n-1} = n * 1 * 0^{n-1} = 0, \quad \binom{n}{1} \geq 1. \\ Pr(X = 2) = f(2) &= \binom{n}{2} p^2 q^{n-2} = \binom{n}{2} 1^2 0^{n-2} = \binom{n}{2} * 1 * 0^{n-2} = 0, \quad \binom{n}{2} \geq 1. \\ Pr(X = 3) = f(3) &= \binom{n}{3} p^3 q^{n-3} = \binom{n}{3} 1^3 0^{n-3} = \binom{n}{3} * 1 * 0^{n-3} = 0, \quad \binom{n}{3} \geq 1. \\ &\dots \\ Pr(X = n) = f(n) &= \binom{n}{n} p^n q^{n-n} = \binom{n}{n} 1^n 0^{n-n} = 1 * 1 * 0^0 = 0^0 = 1, \quad \binom{n}{n} \geq 1. \end{aligned}$$

$$Pr(X > 0) = Pr(X = 1) + Pr(X = 2) + Pr(X = 3) + \dots + Pr(X = n) = 0 + 0 + 0 + \dots + 1 = 1.$$

And this gives the probability that you will make all observations and recordings in all 'now' that in all trials you are alive at that time and agrees with the definite probability.

A possible question here can be, why are the probabilities for all the first $n - 1$ trials 0?

The explanation is that this is a sure probability that if you are able to make, say k trials, then you are alive in all the trials hence the definite probability appears in the last trial you make. For example, if you make, say 5 trials, then you are alive in all the trials and if you cannot make the 6th trial, then you are not alive after the 5th trial. It is not possible to make, say only 3 observations and recordings, and still be alive to make a total

of 5 trials. That's the reason why the sure probability accumulates until you make the last trial, and this applies also for the case of Bernoulli where $n = 1$.

Revisiting case 3, the definite probability of the car in the city causing an accident in the village house is 0, i.e. $p = 0$ and the definite probability that that car won't cause an accident in your house is 1, i.e. $q = 1$. Suppose you make $n > 1$ trials. We can have:

$$\Pr(X = 0) = f(0) = \binom{n}{0} p^0 q^{n-0} = \binom{n}{0} 0^0 1^{n-0} = 1 * 0^0 * 1^n = 0^0 = 1, \quad \binom{n}{0} \geq 1$$

This is the probability of causing no accident (or the probability of causing 0 accidents) in the house while busy driving in capital city and confirms the definite probability.

For $\Pr(X > 0)$, we can use

$$\Pr(X > 0) = 1 - \Pr(X = 0) = 1 - 1 = 0.$$

This is the probability of causing any number of accidents in the house.

Alternatively,

$$\begin{aligned} \Pr(X > 0) &= \Pr(X = 1) + \Pr(X = 2) + \Pr(X = 3) + \dots + \Pr(X = n). \\ \Pr(X = 1) &= f(1) = \binom{n}{1} p^1 q^{n-1} = \binom{n}{1} 0^1 1^{n-1} = n * 0 * 1^{n-1} = 0, \quad \binom{n}{1} \geq 1. \\ \Pr(X = 2) &= f(2) = \binom{n}{2} p^2 q^{n-2} = \binom{n}{2} 0^2 1^{n-2} = \binom{n}{2} * 0 * 1^{n-2} = 0, \quad \binom{n}{2} \geq 1. \\ \Pr(X = 3) &= f(3) = \binom{n}{3} p^3 q^{n-3} = \binom{n}{3} 0^3 1^{n-3} = \binom{n}{3} * 0 * 1^{n-3} = 0, \quad \binom{n}{3} \geq 1. \\ &\dots \\ \Pr(X = n) &= f(n) = \binom{n}{n} p^n q^{n-n} = \binom{n}{n} 0^n 1^{n-n} = 1 * 0 * 1^0 = 0, \quad \binom{n}{n} \geq 1. \end{aligned}$$

$$\Pr(X > 0) = \Pr(X = 1) + \Pr(X = 2) + \Pr(X = 3) + \dots + \Pr(X = n) = 0 + 0 + 0 + \dots + 0 = 0.$$

This is the probability of causing any number of accidents in the house and agrees with the definite probability. It translates to the conclusion that it's not possible to cause an accident in the house in the village while in the city driving.

Consider the same case 3 using normal approximation to binomial [2]. We can use the continuous distributions to estimate probabilities involving discrete cases [17]. In such a case, the random variable say X is specified in a range such as $\Pr(6 < X < 15)$ and not a point like $\Pr(X = 6)$. Another condition is that the probability p should approach 0 as sample size n tends to ∞ or n becomes very large. The $p = 0$, $q = 1$ and fixing $n = 100,000$, we can find $\Pr(X > 10)$. We first define $\mu = np$ as well as $\sigma = \sqrt{npq}$ [17]. In this case,

$$\mu = np = 100,000 * 0 = 0 \text{ and } \sigma = \sqrt{npq} = \sqrt{100,000 * 0 * 1} = 0.$$

Suppose we specify in the binomial case that we want to find $\Pr(X \geq 10)$. Then, since p is very small, we can have the approximation of this probability using the normal distribution as follows;

$$\Pr(X > 10) = 1 - \phi\left(\frac{10 - \mu}{\sigma}\right) = 1 - \phi\left(\frac{10 - 0}{0}\right) = 1 - \phi\left(\frac{10}{0}\right) = 1 - \phi(\infty) = 1 - 1 = 0.$$

This is the probability of causing more than 10 accidents in your house in the village while driving your car in the capital city that is about 300 km away. It's in agreement with the definite probability. NB: Using the method of limits, one can verify that as $\frac{10}{0} \rightarrow \infty$.

For $Pr(X < 10)$, we can use the relationship

$$Pr(X < 10) = 1 - Pr(X > 10) = 1 - 0 = 1.$$

This is the probability of causing less than 10 accidents in the village while in the city. Alternatively,

$$Pr(X < 10) = \phi\left(\frac{10 - \mu}{\sigma}\right) = \phi\left(\frac{10 - 0}{0}\right) = \phi\left(\frac{10}{0}\right) = \phi(\infty) = 1.$$

This gives the chance of causing less than 10 accidents in the house in the village while driving in capital city that is about 300 km away. It's in agreement with the definite probability.

Consider case 3 and the geometric distribution where $p = 0$ (chances of causing an accident in the house while driving in the city), $q = 1$ (chances of not causing an accident in the house while driving in the city). We can find $Pr(X > \text{specified value})$ or $Pr(X < \text{specified value})$. But wait! What is the pdf of a random variable X that follows geometric distribution? The pdf is of the form:

$$f(x) = \begin{cases} pq^x, & x = 0, 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

There is nothing wrong with this pdf but we find that the pdf cannot be used whenever $p = 0$. This is because, the multiplication of any value by 0 will always reduce the whole thing to 0. I.e.

$$f(0) = 0 * 1^0 = 0, \quad f(1) = 0 * 1^1 = 0, \quad f(2) = 0 * 1^2 = 0, \quad \dots, \quad f(\infty) = 0 * 1^\infty = 0.$$

This means that there is no way we can obtain the sure probability of 1.

However, if we reverse the situation and have $p = 1$ (probability of not causing an accident in the house in village when driving in the city) and $q = 0$ (probability of causing an accident in the house in the village while driving in the city), here we can apply the geometric distribution pdf as follows:

$$\begin{aligned} Pr(X = 0) = f(0) = 1 * 0^0 = 0^0 = 1, & \quad Pr(X = 1) = f(1) = 1 * 0^1 = 0^1 = 0, \\ Pr(X = 2) = f(2) = 1 * 0^2 = 0^2 = 0, & \quad \dots, \quad Pr(X = \infty) = f(\infty) = 1 * 0^\infty = 0^\infty = 0. \end{aligned}$$

If the pdf was of the form;

$$f(x) = \begin{cases} pq^{x-1}, & x = 1, 2, 3, \dots \\ 0, & \text{otherwise} \end{cases}$$

The we would have;

$$\begin{aligned} Pr(X = 1) = f(1) = 1 * 0^{1-1} = 0^0 = 1, & \quad Pr(X = 2) = f(2) = 1 * 0^{2-1} = 0^1 = 0, \\ Pr(X = 3) = f(3) = 1 * 0^{3-1} = 0^2 = 0, & \quad \dots, \quad Pr(X = \infty) = f(\infty) = 1 * 0^{\infty-1} = 0^\infty = 0. \end{aligned}$$

This means that the probability of being successful in 'causing no accident' in 1st trial/drive is 1 while that of being successful in 'causing no accident' in any other trial/drive is 0. And this make sense because we expect you to be successful in 'causing no accident' in the first trial/drive. Therefore, the probability of causing 0 accidents is 1 while that of causing any other number of accidents is 0. All these agrees with the case when the pdf is of the form $f(x) = pq^x$. For both forms, the probability 1 occurs only in the first case where $q^x = q^{x-1} = 0^0 = 1$ and probability 0 occurs in the rest of all values of q^x and q^{x-1} .

Case 5:

Consider a bag containing, say, 'N' marbles. Suppose we have 'm' marbles that have traits of interest and we sample 'n' marbles from that bag. We can select the marbles randomly without replacement. A sampled

marble is either having the trait of interest or not. Let X denote the number of marbles in the sample that have the traits of interest. Hypergeometric distribution takes care of random variables like X where selection is done without replacement. If the selection is with replacement, then binomial distribution can take care of that. Let us go to the way of hypergeometric distribution where selection is without replacement. A random variable X that follows hypergeometric distribution has pdf given by:

$$f(x) = \begin{cases} \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}}, & x = 0, 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

N is the total number of items in the lot, m is the number of items in the lot with desired traits, n is the sample size while X is the number of items desired from the sample. The conditions to be satisfied are:

$$X \geq 0, X \geq (n - \text{'others or undesired traits'}), n \geq X \text{ and } m \geq X.$$

Consider case 5 such that there are $N = 100$ marbles in total in a bag, $m = 100$ marbles with traits of interest in the bag, $n = 20$ marbles are sampled randomly from the bag and X is the number of marbles in the sample with desired traits. Suppose the trait of interest is the colour 'red' hence there are $m = 100$ 'red' marbles in the bag. But do not forget that the $N = 100$ marbles hence all the marbles in the bag have trait of interest; they are all 'red' (i.e. $N = m = 100$ 'red' marbles). This further means that the sample has 'red' marbles only (i.e. $n = 20$ 'red' marbles). Let's classify marbles that are in the bag with different colour as 'other' marbles. This means that, the 'other' marbles are not red and hence have undesired trait. From the information provided, there are 0 'other' marbles in the bag, 0 'other' marbles can be sampled from the bag and 0 'other' marbles are in the sample. The sure/definite probability can inform us in advance that the probability of finding 0 'other' marbles from the sample is 1 i.e.

$$\Pr(\text{'other'} = 0 \text{ marbles}) = 1.$$

while the sure/definite probability of finding more than 0 'other' marbles from the sample is 0 i.e.

$$\Pr(\text{'other'} > 0 \text{ marbles}) = 0 = 1 - \Pr(\text{'other'} = 0 \text{ marbles})$$

Similarly, the definite probability of having 0 'red' marbles in the sample is 0 i.e.

$$\Pr(\text{'red'} = 0 \text{ marbles}) = 0$$

while the definite probability of having more than 0 'red' marbles from the sample is 1 i.e.

$$\Pr(\text{'red'} > 0 \text{ marbles}) = 1 = 1 - \Pr(\text{'red'} = 0 \text{ marbles})$$

Here, we can start with the marbles in the bag to find the successive probabilities when selection is done without replacement as follows:

$$P_{red} = \frac{\text{number of 'red' marbles left in the bag at any selection}}{\text{total number of marbles in the bag}}$$

hence

$$P_{1st} = \frac{100}{100} = 1, P_{2nd} = \frac{99}{99} = 1, P_{3rd} = \frac{98}{98} = 1, \dots, P_{20th} = \frac{81}{81} = 1 = P_{red}$$

Here, we find that, at any level of selection of a marble from the bag, the probability is constant and is a definite probability. For the case of the 'red' marbles in the sample, we have:

$$P_{red} = \frac{\text{number of 'red' marbles in the sample}}{\text{total number of marbles in the sample}} = \frac{20}{20} = 1$$

Again, here we end up with a definite probability of red marbles in the sample. Both sequential sampling from the bag and the computations using the sampled marbles give the same definite probability as 1. For the ‘other’ marbles, we can use the same idea and procedure to compute probabilities as follows:

$$P_{other} = \frac{\text{number of 'other' marbles left in the bag at any selection}}{\text{total number of marbles in the bag}}$$

hence

$$P_{1st} = \frac{0}{100} = 0, P_{2nd} = \frac{0}{99} = 0, P_{3rd} = \frac{0}{98} = 0, \dots, P_{20th} = \frac{0}{81} = 0 = P_{other}$$

For the case of the ‘other’ marbles in the sample, we have:

$$P_{other} = \frac{\text{number of 'other' marbles in the sample}}{\text{total number of marbles in the sample}} = \frac{0}{20} = 0$$

Both agree that the probability of selecting an ‘other’ marble from the bag as well as finding an ‘other’ marble in the sample is 0.

Before we embark on the application of the hypergeometric distribution, it can be noted that, [16] has shown the solutions to $\frac{0}{0} = \binom{0}{0} = \frac{0!}{0!(0-0)!} = 0^0$ and by definition, $\binom{n}{k} = \binom{n}{n-k} = {}_n C_k = {}^n C_k = \frac{n!}{k!(n-k)!}$ [6,18]. Let’s now use the hypergeometric distribution function to find the $Pr('other' = 0)$. Here,

$$N = 100, \quad m = 0, \quad n = 20 \quad \text{and} \quad X = 0$$

$$Pr('other' = 0) = f(0) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-0}{20-0} \binom{0}{0}}{\binom{100}{20}} = \frac{\binom{100}{20} \binom{0}{0}}{\binom{100}{20}} = \binom{0}{0} = 1$$

Therefore, the chances of having 0 ‘other’ marbles from the sample is 1, agreeing with the definite probability stated initially.

We can use this to find the $Pr('other' > 0)$ as follows:

$$Pr('other' > 0) = f('other' > 0) = 1 - Pr('other' = 0) = 1 - 1 = 0$$

Since there are no ‘other’ marbles in the bag and in the sample, then the chances of getting more than 0 ‘other’ marbles in the selected sample is 0, and this is in agreement with the definite probability.

There is need to note that, if we use some other notations other than $Pr(other > 0)$, then the probability cannot be computed directly (but can be evaluated indirectly), and if one tries to compute, the probability becomes ‘awkward’. I.e. suppose we try

$$Pr('other' = 1) = f(1) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-0}{20-1} \binom{0}{1}}{\binom{100}{20}} = \frac{\binom{100}{19} \binom{0}{1}}{\binom{100}{20}} = 0.2469 * \binom{0}{1}$$

$$Pr('other' = 2) = f(2) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-0}{20-2} \binom{0}{2}}{\binom{100}{20}} = \frac{\binom{100}{18} \binom{0}{2}}{\binom{100}{20}} = 0.0572 * \binom{0}{2}$$

$$Pr('other' = 20) = f(20) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-0}{20-20} \binom{0}{20}}{\binom{100}{20}} = \frac{\binom{100}{0} \binom{0}{20}}{\binom{100}{20}} = (1.8657 * 10^{-21}) * \binom{0}{20}$$

As can be seen from the computations, we cannot find the exact values of $\binom{0}{1}$, $\binom{0}{2}$ and $\binom{0}{20}$. However, indirectly, if we try to solve any of $\binom{0}{1}$, $\binom{0}{2}$, ... and $\binom{0}{20}$ using gamma functions, integrations and factorials [11,15], we find that the value tends to 0 for all cases leading the solutions to tend to 0. Therefore,

$$Pr('other' = 1) = f(1) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-0}{20-1} \binom{0}{1}}{\binom{100}{20}} = \frac{\binom{100}{19} \binom{0}{1}}{\binom{100}{20}} = 0.2469 * \binom{0}{1} = 0.2469 * 0 = 0$$

$$Pr('other' = 2) = f(2) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-0}{20-2} \binom{0}{2}}{\binom{100}{20}} = \frac{\binom{100}{18} \binom{0}{2}}{\binom{100}{20}} = 0.0572 * \binom{0}{2} = 0.0572 * 0 = 0$$

$$Pr('other' = 20) = f(20) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-0}{20-20} \binom{0}{20}}{\binom{100}{20}} = \frac{\binom{100}{0} \binom{0}{20}}{\binom{100}{20}} = (1.8657 * 10^{-21}) * \binom{0}{20} = (1.8657 * 10^{-21}) * 0 = 0$$

This means that the probability of having any number of ‘other’ marbles in the sample apart from 0 ‘other’ marbles is 0. This indeed agrees with the definite probability.

Coming back to the ‘red’ marbles that are in the bag, we might be interested also in determining the probability of finding a specified number of red marbles in the sample. Without difficulties, one can compute the

$$Pr('red' marbles = 20) = 1 - Pr('other' > 0) = 1 - 0 = 1$$

This means that the probability of finding all the sampled marbles as ‘red’ marbles is 1. Similarly,

$$Pr('red' marbles = 0) = 1 - Pr('other' = 0) = 1 - 1 = 0$$

which is the probability of finding no ‘red’ marble in the sample.

Alternatively, we can use the hypergeometric function- where $N = 100, m = 100, n = 20$ and X is specified as number of ‘red’ marbles in the sample- as well as the gamma functions, factorials and integrations [11,15] where necessary to compute some of the probabilities as follows:

$$Pr('red' = 0) = f(0) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-100}{20-0} \binom{100}{0}}{\binom{100}{20}} = \frac{\binom{0}{20} \binom{100}{0}}{\binom{100}{20}} = \binom{0}{20} * (1.8657 * 10^{-21}) = 0 * (1.8657 * 10^{-21}) = 0$$

$$Pr('red' = 20) = f(20) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-100}{20-20} \binom{100}{20}}{\binom{100}{20}} = \frac{\binom{0}{0} \binom{100}{20}}{\binom{100}{20}} = \binom{0}{0} = 1$$

These two agree with the definite probability.

Other probabilities for specified number of ‘red’ marbles can be computed as follows:

$$\begin{aligned} Pr('red' = 1) = f(1) &= \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-100}{20-1} \binom{100}{1}}{\binom{100}{20}} = \frac{\binom{0}{19} \binom{100}{1}}{\binom{100}{20}} = \binom{0}{19} * (1.8657 * 10^{-19}) \\ &= 0 * (1.8657 * 10^{-19}) = 0 \end{aligned}$$

$$\begin{aligned} Pr('red' = 2) = f(2) &= \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-100}{20-2} \binom{100}{2}}{\binom{100}{20}} = \frac{\binom{0}{18} \binom{100}{2}}{\binom{100}{20}} = \binom{0}{18} * (9.2354 * 10^{-18}) \\ &= 0 * (9.2354 * 10^{-18}) = 0 \end{aligned}$$

$$\begin{aligned} Pr('red' = 19) = f(19) &= \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{100-100}{20-19} \binom{100}{19}}{\binom{100}{20}} = \frac{\binom{0}{1} \binom{100}{19}}{\binom{100}{20}} = \binom{0}{1} * 0.2469 \\ &= 0 * 0.2469 = 0. \end{aligned}$$

Here we find that, all the probabilities of finding less than 20 'red' marbles in the sample are 0. Therefore, whenever the number of 'red' marbles is specified as less than what is sampled, then the probability of finding such number of 'red' marbles in the sample is 0. This is because, there is no way we can find fewer 'red' marbles in the sample than what was sampled since all the sampled marbles were from the bag with only 'red' marbles. If we want 'red' marbles that are fewer than what was sampled, then it means that there are other marbles in the bag with different colours. But in our case, the marbles in the bag with 'other' colours were 0. Therefore, if we want to find 'red' marbles in the sample which are less than the number sampled, this would be misleading and the probability of getting such marbles in the sample is always 0. For example, suppose we want $Pr('red' = 6)$. Here, we imply that the remaining $20 - 6 = 14$ marbles are of 'other' colours apart from being 'red'. Similarly, with the 'other' marbles, there is only one possibility of getting 'other' marbles in the sample- this is the possibility of finding 0 'other' marbles in the sample because there were 0 'other' marbles in the bag hence there can only be 0 'other' marbles in the sample. Therefore, specifying the number of 'other' marbles to be greater than 0 means that there were 'other' marbles in the bag initially, which is not true. This is why the probability of finding more than 0 'other' marbles in the sample is always 0, but that of finding 0 'other' marbles in the sample is a sure probability of 1.

Case 6:

Revisit case 3 where the definite probability of causing an accident in the house in the village when driving in capital city that is 300 km away is $p = 0$ while $q = 1$ is the definite probability of causing no accident in the same setting. The probability of getting r^{th} accident in k^{th} drive/trial can be computed in a negative binomial distribution. A random variable X that follows the negative binomial distribution has pdf given by:

$$f(k) = Pr(X_r = k) = \begin{cases} \binom{k-1}{r-1} p^r q^{k-r}, & k = r, r+1, r+2, \dots \\ 0, & elsewhere \end{cases}$$

Computing some probabilities of causing accidents in the house while driving in the capital city, we can use this negative binomial distribution pdf as follows:

- i) Use of the pdf when $p = 0$ and $q = 1$.

Probability of causing 1st accident in 1st drive/trial is

$$Pr(X_1 = 1) = f(1) = \binom{1-1}{1-1} 0^1 1^{1-1} = \binom{0}{0} 0^1 1^0 = 1 * 0 * 1 = 0.$$

Probability of causing 1st accident in 2nd drive/trial is

$$\Pr(X_1 = 2) = f(2) = \binom{2-1}{1-1} 0^1 1^{2-1} = \binom{1}{0} 0^1 1^1 = 1 * 0 * 1 = 0.$$

...

Probability of causing 1st accident in 10th drive/trial is

$$\Pr(X_1 = 10) = f(10) = \binom{10-1}{1-1} 0^1 1^{10-1} = \binom{9}{0} 0^1 1^9 = 1 * 0 * 1 = 0.$$

Probability of causing 10th accident in 10th drive/trial is

$$\Pr(X_{10} = 10) = f(10) = \binom{10-1}{10-1} 0^{10} 1^{10-10} = \binom{9}{9} 0^{10} 1^0 = 1 * 0 * 1 = 0.$$

From the above calculations, we find that the probabilities are 0 for all specified successes. This reflects what was reflected in the geometric distribution case in which if $p = 0$, then all the probabilities computed are also 0 hence the two pdfs are not used when $p = 0$ but can be applied when $p \neq 0$. When $p = 0$, then the whole equation reduces to 0 in all computations.

ii) Use of the pdf when $p = 1$ and $q = 0$.

Suppose $p = 1$ is the definite probability of ‘not causing’ (failure to cause) an accident in the house while $q = 0$ is the definite probability of causing an accident in the house while driving in the capital city that is about 300 km away. We can repeat the computations of probabilities as follows:

Probability of not causing (failure to cause) 1st accident in 1st drive/trial is

$$\Pr(X_1 = 1) = f(1) = \binom{1-1}{1-1} 1^1 0^{1-1} = \binom{0}{0} 1^1 0^0 = 1 * 1 * 0^0 = 1 * 1 * 1 = 1.$$

The probability of 1st failure occurring in 1st trial/drive is 1, which agrees with definite probability.

Probability of not causing 1st accident (1st failure to cause an accident) in 2nd drive/trial is

$$\Pr(X_1 = 2) = f(2) = \binom{2-1}{1-1} 1^1 0^{2-1} = \binom{1}{0} 1^1 0^1 = 1 * 1 * 0 = 0.$$

The probability of 1st failure occurring in 2nd trial/drive is 0. This agrees with definite probability.

...

Probability of not causing 1st accident (1st failure occurring) in 10th drive/trial is

$$\Pr(X_1 = 10) = f(10) = \binom{10-1}{1-1} 1^1 0^{10-1} = \binom{9}{0} 1^1 0^9 = 1 * 1 * 0 = 0.$$

Probability of not causing 5th accident (5th failure occurring) in 10th drive/trial is

$$\Pr(X_5 = 10) = f(10) = \binom{10-1}{5-1} 1^5 0^{10-5} = \binom{9}{4} 1^5 0^5 = 126 * 1 * 0 = 0.$$

Probability of not causing 8th accident (8th failure occurring) in 10th drive/trial is

$$\Pr(X_8 = 10) = f(10) = \binom{10-1}{8-1} 1^8 0^{10-8} = \binom{9}{7} 1^8 0^2 = 36 * 1 * 0 = 0.$$

Probability of not causing 9th accident (9th failure occurring) in 10th drive/trial is

$$\Pr(X_9 = 10) = f(10) = \binom{10-1}{9-1} 1^9 0^{10-9} = \binom{9}{8} 1^9 0^1 = 9 * 1 * 0 = 0.$$

Probability of not causing 10th accident (10th failure occurring) in 10th drive/trial is

$$\Pr(X_{10} = 10) = f(10) = \binom{10-1}{10-1} 1^{10} 0^{10-10} = \binom{9}{9} 1^{10} 0^0 = 1 * 1 * 0^0 = 0^0 = 1.$$

The probability of 10th failure occurring in 10th trial/drive is 1. This agrees with definite probability.

Probability of not causing 6th accident (6th failure occurring) in 6th drive/trial is

$$\Pr(X_6 = 6) = f(6) = \binom{6-1}{6-1} 1^6 0^{6-6} = \binom{5}{5} 1^6 0^0 = 1 * 1 * 0^0 = 0^0 = 1.$$

Probability of not causing 17th accident (17th failure occurring) in 17th drive/trial is

$$\Pr(X_{17} = 17) = f(17) = \binom{17-1}{17-1} 1^{17} 0^{17-17} = \binom{16}{16} 1^{17} 0^0 = 1 * 1 * 0^0 = 0^0 = 1.$$

All these probabilities confirm the definite probabilities stated at the beginning. Looking at these results keenly, we are able to see that all the probabilities corresponding to situations where the number of failures is specified as different from number of trials/drives performed are 0. Otherwise, the probabilities are all 1. This is true because you cannot fail to cause an accident fewer number of times than the number of times you try. For example, you cannot fail to cause an accident for the 1st time when driving the 10th time because you are bound to fail to cause accidents all the times you drive. If you get the 1st failure to cause an accident in 5th trial/drive, then it means that for the first 4 trials, you had caused 4 accidents, which is not true. If you try 10 times to drive, you will fail to cause an accident 10 times, if you try 2 times, you will fail to cause an accident 2 times, etc. So, when driving the 10th time, you are bound to fail to cause an accident the 10th time; when driving the 5th time, you are bound to fail to cause an accident the 5th time, etc. It is not possible to fail to cause an accident for the 1st time in 2nd trial/drive because if it were possible, it would mean that during the 1st trial, you had actually caused an accident, which is not true; it is not possible to fail to cause an accident for the 1st time in 10th trial/drive because if it were possible, it would mean that during the first 9 trials, you had actually caused 9 accidents, which is not true; it is not possible to fail to cause an accident for the 5th time in 10th trial/drive because if it were possible, it would mean that during the first 9 trials, you had already caused 5 accidents, which is not true; it is not possible to fail to cause an accident for the 3rd time in 7th trial/drive because if it were possible, it would mean that during the first 6 trials, you had already caused 4 accidents, which is not true; etc. In general, the probability of failing to cause an accident 'n' times when you make 'n' trials/drives is always 1 while the probability of failing to cause an accident 'k' times when you make 'n' trial/drives where $k \neq n$ or $k < n$ is always 0.

Revisiting case 4, we have 'you' observing and recording 'now', that you are alive at that time and then the definite probability that you are alive at that time is 1, i.e. $p = 1$ while the definite probability that you are not alive at that time when making this observation and recording it 'now' is 0, i.e. $q = 0$.

Making some trials, we can use the negative binomial pdf to compute the probabilities for some specified values of successes and trials. Suppose you make 10 trials. Then,

Probability of making 1st recording in 1st trial is

$$\Pr(X_1 = 1) = f(1) = \binom{1-1}{1-1} 1^1 0^{1-1} = \binom{0}{0} 1^1 0^0 = 1 * 1 * 0^0 = 1 * 1 * 1 = 1.$$

The probability of making 1st recording in 1st trial is 1 which is in agreement with definite probability.

Probability of making 1st recording in 2nd trial is

$$\Pr(X_1 = 2) = f(2) = \binom{2-1}{1-1} 1^1 0^{2-1} = \binom{1}{0} 1^1 0^1 = 1 * 1 * 0 = 0.$$

The probability of making 1st recording in 2nd trial is 0. This agrees with definite probability.

Probability of making 1st recording in 10th trial is

$$\Pr(X_1 = 10) = f(10) = \binom{10-1}{1-1} 1^1 0^{10-1} = \binom{9}{0} 1^1 0^9 = 1 * 1 * 0 = 0.$$

Probability of making 5th recording in 10th trial is

$$\Pr(X_5 = 10) = f(10) = \binom{10-1}{5-1} 1^5 0^{10-5} = \binom{9}{4} 1^5 0^5 = 126 * 1 * 0 = 0.$$

Probability of making 8th recording in 10th trial is

$$\Pr(X_8 = 10) = f(10) = \binom{10-1}{8-1} 1^8 0^{10-8} = \binom{9}{7} 1^8 0^2 = 36 * 1 * 0 = 0.$$

Probability of making 9th recording in 10th trial is

$$\Pr(X_9 = 10) = f(10) = \binom{10-1}{9-1} 1^9 0^{10-9} = \binom{9}{8} 1^9 0^1 = 9 * 1 * 0 = 0.$$

Probability of making 10th recording in 10th trial is

$$\Pr(X_{10} = 10) = f(10) = \binom{10-1}{10-1} 1^{10} 0^{10-10} = \binom{9}{9} 1^{10} 0^0 = 1 * 1 * 0^0 = 1 * 1 * 1 = 1.$$

The probability of making 10th recording in 10th trial is 1. This agrees with definite probability.

Here, we see what had been observed with binomial distribution that it is not possible to observe and record 'now' that you are alive at that time fewer number of times than the trials you make. Therefore, any situation where it is specified that you make fewer number of observations and recordings than the number of trials you make has probability equal to 0 while observations and recordings equal to the trials go with probability 1.

Case 7:

Consider a jug with, say 10, blue marbles only. Suppose that, each time a marble is selected from the jug at random and its colour noted. Whether the selection is with replacement or without replacement, the definite probability of selecting a blue marble from the jug is always 1 (i.e. $p = 1$) while the definite probability of selecting a marble that has other colours apart from colour blue is 0 (i.e. $q = 0$). These sure probabilities can be verified very easily as follows:

a) Selection without replacement

$$P_{blue} = \frac{\text{number of 'blue' marbles left in the jug at any selection}}{\text{total number of marbles in the jug}} = p$$

hence

$$P_{1st} = \frac{10}{10} = 1, P_{2nd} = \frac{9}{9} = 1, P_{3rd} = \frac{8}{8} = 1, \dots, P_{10th} = \frac{1}{1} = 1 = P_{blue} = p$$

b) Selection with replacement

$$P_{blue} = \frac{\text{number of 'blue' marbles in the jug at any selection}}{\text{total number of marbles in the jug}} = p$$

hence

$$P_{1st} = \frac{10}{10} = 1, P_{2nd} = \frac{10}{10} = 1, P_{3rd} = \frac{10}{10} = 1, \dots, P_{10th} = \frac{10}{10} = 1 = P_{blue} = p$$

Letting all the other marbles in the jug that are not blue be labelled as 'others', we have their definite probabilities given as

$$P_{others} = 1 - P_{blue} = 1 - p = 1 - 1 = 0 = q$$

If we use the binomial pdf, we can compute probabilities of interest at any selection/trial we make. We had the binomial pdf given by

$$f(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0, 1, 2, \dots, n \\ 0, & \text{otherwise} \end{cases}$$

Suppose from the jug with 10 blue marbles, 6 marbles are selected at random. Since $p = 1$ and $q = 0$ are the definite probabilities of selecting a blue marble and a marble of other colours respectively, then we can have:

Our sample size is $n = 6$, $p = 1$ and $q = 0$.

Probability of getting 0 blue marble in the sample is

$$Pr(X = 0) = \binom{n}{x} p^x q^{n-x} = \binom{6}{0} 1^0 0^{6-0} = 1 * 1 * 0^6 = 0$$

Probability of getting 1 blue marble in the sample is

$$Pr(X = 1) = \binom{n}{x} p^x q^{n-x} = \binom{6}{1} 1^1 0^{6-1} = 6 * 1 * 0^5 = 0$$

Probability of getting 2 blue marbles in the sample is

$$Pr(X = 2) = \binom{n}{x} p^x q^{n-x} = \binom{6}{2} 1^2 0^{6-2} = 15 * 1 * 0^4 = 0$$

...

Probability of getting 5 blue marbles in the sample is

$$Pr(X = 5) = \binom{n}{x} p^x q^{n-x} = \binom{6}{5} 1^5 0^{6-5} = 6 * 1 * 0^1 = 0$$

Probability of getting 6 blue marbles in the sample is

$$Pr(X = 6) = \binom{n}{x} p^x q^{n-x} = \binom{6}{6} 1^6 0^{6-6} = 1 * 1 * 0^0 = 1 * 1 * 1 = 1$$

The computations with the binomial distribution pdf confirm the definite probabilities. In this case, it's only at $Pr(X = 6) = Pr(X = n)$ we have probability equal to 1. Any other probability like $Pr(X < n)$ gives probabilities equal to 0 always. The reason is that there are only blue marbles in the jug and hence, specifying blue marbles to be less than the number of marbles in the sample means that some other marbles of different colours exist in the jug and hence there is possibility of those other marbles to appear in the sample too, which is not the case. In general,

$$\begin{aligned} \text{for } n = 2, Pr(X = 2) &= 1 \text{ and } Pr(X < 2) = 0 \\ \text{for } n = 3, Pr(X = 3) &= 1 \text{ and } Pr(X < 3) = 0 \\ \text{for } n = 4, Pr(X = 4) &= 1 \text{ and } Pr(X < 4) = 0 \\ &\dots \\ \text{for } n = h, Pr(X = h) &= 1 \text{ and } Pr(X < h) = 0 \end{aligned}$$

In the same case, since we had a jug with 10 marbles that were all of blue colour, and selected say 6 marbles randomly from the jug, we can apply the hypergeometric distribution pdf to compute some probabilities for some specified marbles. In this case, $N = 10, m = 10, n = 6$ and X can be specified while

$$f(x) = \begin{cases} \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}}, & x = 0, 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

We compute some probabilities as follows:

$$\begin{aligned} Pr('blue' = 0) = f(0) &= \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{10-10}{6-0} \binom{10}{0}}{\binom{10}{6}} = \frac{\binom{0}{6} \binom{10}{0}}{\binom{10}{6}} = \binom{0}{6} * (4.7619 * 10^{-3}) \\ &= 0 * (4.7619 * 10^{-3}) = 0 \end{aligned}$$

$$Pr('blue' = 6) = f(6) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{10-10}{6-6} \binom{10}{6}}{\binom{10}{6}} = \frac{\binom{0}{0} \binom{10}{6}}{\binom{10}{6}} = \binom{0}{0} = 1$$

Other probabilities for specified number of 'blue' marbles can be computed as follows:

$$Pr('blue' = 1) = f(1) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{10-10}{6-1} \binom{10}{1}}{\binom{10}{6}} = \frac{\binom{0}{5} \binom{10}{1}}{\binom{10}{6}} = \binom{0}{5} * 0.0476 = 0 * 0.0476 = 0$$

$$Pr('blue' = 2) = f(2) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{10-10}{6-2} \binom{10}{2}}{\binom{10}{6}} = \frac{\binom{0}{5} \binom{10}{2}}{\binom{10}{6}} = \binom{0}{5} * 0.2143 = 0 * 0.2143 = 0$$

$$Pr('blue' = 5) = f(5) = \frac{\binom{N-m}{n-x} \binom{m}{x}}{\binom{N}{n}} = \frac{\binom{10-10}{6-5} \binom{10}{5}}{\binom{10}{6}} = \frac{\binom{0}{1} \binom{10}{5}}{\binom{10}{6}} = \binom{0}{1} * 1.2 = 0 * 1.2 = 0$$

All these probabilities agree with definite probabilities. Such a scenario had been observed in case 5. The probabilities of obtaining all the blue marbles equal to the number of trials are always 1 while anything contrary has probability 0.

Reconsider case 7 and let the jug have 1,000,000 blue marbles only. Suppose 100,000 marbles are selected at random from the jug, $p = 0$ is the definite probability of selecting a 'red' marble from the jug, whether the

selection is with replacement or without replacement. On the other hand, $q = 1$ is the definite probability of selecting (with or without replacement) a marble of other colours apart from 'red'. Suppose that, each time a marble is selected from the jug at random and its colour noted. We had the binomial pdf given by

$$f(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0, 1, 2, \dots, n \\ 0, & \text{otherwise} \end{cases}$$

Probability of getting 0 'red' marble in the sample is

$$Pr(X = 0) = \binom{n}{0} p^0 q^{n-0} = \binom{100000}{0} 0^0 1^{100000-0} = 1 * 0^0 * 1 = 1 * 1 * 1 = 1$$

Probability of getting 1 'red' marble in the sample is

$$Pr(X = 1) = \binom{n}{1} p^1 q^{n-1} = \binom{100000}{1} 0^1 1^{100000-1} = 100000 * 0 * 1 = 0$$

Probability of getting 2 'red' marbles in the sample is

$$Pr(X = 2) = \binom{n}{2} p^2 q^{n-2} = \binom{100000}{2} 0^2 1^{100000-2} = 4999950000 * 0 * 1 = 0$$

...

Probability of getting 100,000 'red' marbles in the sample is

$$Pr(X = 100000) = \binom{n}{x} p^x q^{n-x} = \binom{100000}{100000} 0^{100000} 1^{100000-100000} = 1 * 0 * 1 = 0$$

All these agree with the definite probabilities that since there are 0 'red' marbles in the jug, then the probability of selecting 0 'red' marbles is 1 while that of selecting more than 0 'red' marbles is 0.

Consider using Poisson approximation to this binomial [17] because n is very large and p is very small, where $p = 0$ is the definite probability of selecting a 'red' marble from the jug while $q = 1$ is the definite probability of selecting (with or without replacement) a marble of other colours apart from 'red' from the same jug. In Poisson distribution, we have the pdf as

$$f(x) = \begin{cases} \frac{e^{-\lambda} * \lambda^x}{x!}, & x = 0, 1, 2, 3, \dots, \lambda > 0 \\ 0, & \text{elsewhere} \end{cases}$$

We can use the parameters $n = 100000$ and $p = 0$ from the binomial distribution to estimate the parameter λ in Poisson distribution. The relationship used is $\lambda = np$ [6] [17] and in this case, $\lambda = np = 100000 * 0 = 0$. We can note that, since $E(X) = var(X) = \lambda$ in Poisson distribution, it is only when $p = 0$ (very small value) do we have the equality satisfied when estimated from binomial parameters. I.e. $\lambda = np = npq = 100000 * 0 = 100000 * 0 * 1 = 0$. Using this, we can estimate the probabilities computed using binomial pdf and compare the results as follows:

Probability of getting 0 'red' marble in the sample is

$$Pr(X = 0) = f(0) = \frac{e^{-\lambda} * \lambda^0}{0!} = \frac{e^{-0} * 0^0}{0!} = \frac{e^0 * 0^0}{0!} = \frac{1 * 0^0}{1} = \frac{1 * 1}{1} = 1$$

Probability of getting 1 'red' marble in the sample is

$$Pr(X = 1) = f(1) = \frac{e^{-\lambda} * \lambda^1}{1!} = \frac{e^{-0} * 0^1}{1!} = \frac{e^0 * 0^1}{1!} = \frac{1 * 0}{1} = 0$$

Probability of getting 2 'red' marbles in the sample is

$$Pr(X = 2) = f(2) = \frac{e^{-\lambda} * \lambda^2}{2!} = \frac{e^{-0} * 0^2}{2!} = \frac{e^0 * 0^2}{2!} = \frac{1 * 0}{2} = 0$$

...

Probability of getting 100,000 'red' marbles in the sample is

$$Pr(X = 100000) = f(100000) = \frac{e^{-\lambda} * \lambda^{100000}}{100000!} = \frac{e^{-0} * 0^{100000}}{100000!} = \frac{e^0 * 0^{100000}}{100000!} = \frac{1 * 0}{100000!} = 0$$

All these results confirm the findings gotten using the binomial pdf as well as the definite probabilities. Therefore, it's only when you want 0 'red' marbles from the sample do you have a sure probability, 1. Any other number of 'red' marbles from the sample is an impossible event with a sure probability 0. And the reason? There are no 'red' marbles in the jug and hence no 'red' marbles in the sample.

3 Simple Conditional Probabilities

Conditional probabilities involve computation of chances of events happening based on some past knowledge [1,3,7] such as the happening of another event. E.g. the probability of a person drinking water given that he has eaten some food. In conditional probability, suppose we have two events, A and B, each with probability $P(A)$ and $P(B)$ of happening respectively. Then, the probability of event A happening given that event B has happened, denoted as $Pr(A|B)$, is defined as [1,3,7]:

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(A \text{ and } B)}{Pr(B)}, Pr(B) \neq 0 \text{ or } Pr(B) > 0.$$

This can be converted to the multiplication rule for probabilities by making $Pr(A \cap B)$ subject of the formula as follows:

$$Pr(A \cap B) = Pr(A|B) * Pr(B)$$

Suppose the two events of interest are

- (i) A = Event that you observe and be able to record 'now', that you are alive at that time/moment.
- (ii) B = Event that you cause an accident in your house in the village while driving a car in the capital city that is about 300 km away.

It was seen that $Pr(A) = 1$ while $Pr(B) = 0$ in case 4 and 3 respectively. We may want to compute the probability of event A happening given that event B has occurred, i.e.

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(A \text{ and } B)}{Pr(B)}.$$

Two things to note are:

- (i) that these two events are independent and hence,
 $Pr(A \cap B) = Pr(A \text{ and } B) = Pr(A) * Pr(B)$.

and

- (ii) the conditional probability is equal to probability of A hence,
 $Pr(A \setminus B) = Pr(A) = 1$.

From the usual definition of conditional probabilities, the denominator should not be equal to zero ($Pr(B) > 0$) but we want to violate this condition because the division of zero by itself has a unique solution (i.e. $\frac{0}{0} = 1$). Going with this result, we can compute the

$$Pr(A \setminus B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(A \text{ and } B)}{Pr(B)}; Pr(B) = 0$$

as follows:

$$\begin{aligned} Pr(A \setminus B) &= \frac{Pr(A \text{ and } B)}{Pr(B)} = \left[\frac{Pr(A) * Pr(B)}{Pr(B)} \right] \text{ or } \left[Pr(A) * \frac{Pr(B)}{Pr(B)} \right] = \left[\frac{1 * 0}{0} \right] \text{ or } \left[1 * \frac{0}{0} \right] \\ &= \left[\frac{0}{0} \right] \text{ or } [1 * 1] = 1. \end{aligned}$$

This shows that $Pr(A \setminus B) = 1 = Pr(A)$ and agrees with the rule that $Pr(A \setminus B) = Pr(A)$ when events are independent.

Similarly,

$$\begin{aligned} Pr(B \setminus A) &= \frac{Pr(A \cap B)}{Pr(A)} = \left[\frac{Pr(A) * Pr(B)}{Pr(A)} \right] \text{ or } \left[\frac{Pr(A)}{Pr(A)} * Pr(B) \right] = \left[\frac{1 * 0}{1} \right] \text{ or } \left[\frac{1}{1} * 0 \right] \\ &= \left[\frac{0}{1} \right] \text{ or } [1 * 0] = 0. \end{aligned}$$

Thus $Pr(B \setminus A) = 0 = Pr(B)$ and agrees with the rule that $Pr(B \setminus A) = Pr(B)$ when events are independent.

Again, suppose that

- (i) C = Event that you select a blue marble from a jug that has only red marbles.
(ii) D = Event that you cause an accident in your house in the village while driving a car in the capital city that is about 300 km away.

In this case, $Pr(C) = 0$ and $Pr(D) = 0$. Again, the events C and D are independent and hence

$$Pr(C \setminus D) = \frac{Pr(C \cap D)}{Pr(D)} = \frac{Pr(C \text{ and } D)}{Pr(D)}; Pr(D) = 0.$$

Therefore, the conditional probabilities are as follows:

$$\begin{aligned} Pr(C \setminus D) &= \frac{Pr(C \cap D)}{Pr(D)} = \left[\frac{Pr(C) * Pr(D)}{Pr(D)} \right] \text{ or } \left[Pr(C) * \frac{Pr(D)}{Pr(D)} \right] = \left[\frac{0 * 0}{0} \right] \text{ or } \left[0 * \frac{0}{0} \right] = \left[\frac{0^2}{0^1} \right] \text{ or } [0 * 1] \\ &= [0^{2-1}] \text{ or } [0 * 1] = [0^1] \text{ or } [0 * 1] = 0. \end{aligned}$$

Therefore, $Pr(C \setminus D) = 0 = Pr(C)$ and agrees with the rule that $Pr(C \setminus D) = Pr(C)$ when events are independent.

In a similar way, it can be shown that, $Pr(D \setminus C) = 0 = Pr(D)$.

Let's now turn to a case of dependent events.

Turning to the case of a bag that has, say 10, red marbles only, we see that the $Pr(red) = 1$ and $Pr(Other\ marbles) = 0$, whether selection is with or without replacement, at whatever level of selection. This was demonstrated in case 5. We can let "R" be red marbles and assume the bag has 10 red marbles and 0 blue marbles that are represented by "B", hence

$$Pr(R) = Pr(R_1) = Pr(R_2) = Pr(R_3) = Pr(R_4) = 1 \text{ and } Pr(B) = Pr(B_1) = Pr(B_2) = Pr(B_3) = Pr(B_4) = 0$$

and this is the case for all selections from that bag.

Let events R and B be defined as follows:

- (i) R = event that a red marble is selected.
- (ii) B = event that a blue marble is selected.

Suppose we are interested in selecting 4 marbles from that bag, one marble after the other with or without replacement. If we do it with replacement, then the events R and B are independent. Otherwise, the events are dependent. Going the way of 'without replacement', the events are dependent and the following are the results one can get using a tree diagram or other ways:

$$\begin{aligned} Pr(R_1R_2R_3R_4) &= 1, Pr(R_1R_2R_3B_4) = 0, Pr(R_1R_2B_3R_4) = 0, Pr(R_1R_2B_3B_4) = 0, Pr(R_1B_2R_3R_4) \\ &= 0, Pr(R_1B_2R_3B_4) = 0, Pr(R_1B_2B_3R_4) = 0, Pr(R_1B_2B_3B_4) = 0, Pr(B_1R_2R_3R_4) \\ &= 0, Pr(B_1R_2R_3B_4) = 0, Pr(B_1R_2B_3R_4) = 0, Pr(B_1R_2B_3B_4) = 0, Pr(B_1B_2R_3R_4) \\ &= 0, Pr(B_1B_2R_3B_4) = 0, Pr(B_1B_2B_3R_4) = 0, Pr(B_1B_2B_3B_4) = 0. \end{aligned}$$

Here, the subscripts 1, 2, 3 and 4 indicate the rounds/stages of selection 1st, 2nd, 3rd and 4th selection respectively. These outcomes show that all the intersections involving event B have probability 0 because blue marbles are not present in the bag hence no way to select a blue marble at any stage of selection. The only case that has a probability different from 0 is the case of $(R_1 \cap R_2 \cap R_3 \cap R_4)$ where $Pr(R_1 \cap R_2 \cap R_3 \cap R_4) = 1$. Computing the conditional probabilities in this case needs one to think soberly because some specifications will make sense while others won't. For example: the specifications $Pr(B_4 \setminus R_1)$ and $Pr(B_3 \setminus R_2)$ make a lot of sense while $Pr(R_4 \setminus B_1)$ and $Pr(R_3 \setminus B_2)$ won't make any sense. Why? The case of $Pr(B_4 \setminus R_1)$ means the chances of getting a blue marble in the 4th selection given that you got a red marble in the first round/selection, which is fine because there are red marbles in the bag hence it is possible to have a red marble in the 1st round. However, the case of $Pr(R_4 \setminus B_1)$ is misplaced because it means the chances of getting a red marble in 4th round when you have gotten a blue marble in 1st round. This is not possible because there are no blue marbles in the bag hence one cannot get a blue marble in the 1st round. So, the only computations that are doable and are logical involve computing chances of getting a blue marble (the absent marbles/items/objects) in any round given a red one (the present marbles/items/objects) was gotten in a specified round as well as a red one in a specified round given that a red one was gotten in another round without the chain involving the blue (absent) marble. Examples include:

$$\begin{aligned} Pr(B_4 \setminus R_1) &= \frac{Pr(R_1 \cap B_4)}{Pr(R)} = \frac{Pr(R_1R_2R_3B_4)}{Pr(R)} + \frac{Pr(R_1R_2B_3B_4)}{Pr(R)} + \frac{Pr(R_1B_2R_3B_4)}{Pr(R)} + \frac{Pr(R_1B_2B_3B_4)}{Pr(R)} \\ &= \frac{0}{1} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1} = 0. \end{aligned}$$

Therefore, it's not possible have a blue marble in the 4th round. Others such as $Pr(B_2 \setminus R_1), Pr(B_3 \setminus R_1), Pr(B_7 \setminus R_3), Pr(B_8 \setminus R_6)$, etc. can be computed in a similar manner. For $Pr(R_j \setminus R_k)$ where $j \geq k$ and without involving B, we have examples like:

$$Pr(R_4 \setminus R_1) = Pr(R_3 \setminus R_1) = Pr(R_2 \setminus R_1) = Pr(R_4 \setminus R_2) = Pr(R_3 \setminus R_2) = Pr(R_4 \setminus R_3) = \frac{1}{1} = 1.$$

This makes sense because the chances of getting a red marble in any selection is ever 1 since there are only red marbles in the bag. The other sets of $Pr(R_k \setminus B_m); k > m$ and $Pr(B_k \setminus B_m); k \geq m$ don't make sense hence no need to bother with them.

One thing to note here is that the cases of conditional probabilities present scenarios that need one to be careful in evaluating the equations involving 0 in both numerator and denominator as was noted in [16]. Computing the probabilities in a "hurry" might cause "chaos" and confusion but a careful evaluation that involves critical thinking will always lead to that unique path that leads you "home" and doesn't kill logic.

4 Conclusion

Probability is the measure assigned to the chance of happening of an event. A definite probability is simply a sure or certain probability. The definite probability is either 1 or 0 and, in this work, the probabilities p and q were always assigned 1 or 0 depending on the problem at hand. The definite probabilities were first always provided through logic and healthy reasoning before embarking on computations of the same using various ways such as involvement of pdfs and basic knowledge. The combined equations $x = \frac{0}{0} = 0^0 = \binom{0}{0} = \frac{0!}{0!(0-0)!}$ have been encountered in almost all the computations in this work. Using the solution provided in [16], where the solution is 1, it was possible to evaluate exact solutions in computing sure probabilities using the various probability distribution functions. The solutions provided actually can never contradict the logic behind what we know in reality. For example, if the probability of selecting 0 red marbles from a bag that has only blue marbles is 1 while that of selecting any other number of red marbles from the same bag of blue marbles is 0, then, logic tells us that this is indeed true and computations just reinforce the same. The case of $\binom{0}{k}$ where $k > 0$ was encountered too, and a way out was made using some knowledge in definitions of gamma functions, factorials as well as integrations to arrive at the conclusion that such evaluations have solutions tending to 0, which confirmed that the sure probabilities were never contradicted in the computations. In other cases, violations were made in setting the domains of some parameters such as $\sigma > 0$, $\lambda > 0$. This happened because, [16] solved what could hinder ancient computations when scenarios such as $\sigma = 0$ and $\lambda = 0$ are encountered. In some cases, it was found that when using the exponential pdf, the probabilities of X that are specified to be in the same direction as the parameter β should not be evaluated directly, due to contradictions, numerous and awkward solutions, but indirectly. Indirectly means using the computations involved when X is not in the direction of β and then subtracting the results from 1 to arrive at the desired destinations. In the cases of conditional probabilities, the restriction of the denominator to be a positive value can be revisited to have include the 0 scenario. Therefore, the final comment is that there is a great deal of harmony between this work on definite probabilities and [16] work on division of zero by itself and there is a lot that can be done based on $\frac{0}{0} = 1$ findings.

5 Recommendations

1. Some of the distributions existing to date need some reviewing to ensure they cope and are in line with modern/current/recent findings. Such a case can be the expansion of the domains of parameter λ in Poisson to have $\lambda = 0$, the revision of the restrictions in conditional probabilities such as $Pr(B) > 0$ when $Pr(B)$ is the denominator.
2. Review of the definition of expression $\binom{0}{k}$ where $k > 0$. Such a case was encountered many times in this work and needs to be defined officially from academic angle.

Competing Interests

Author has declared that no competing interests exist.

References

- [1] Terrell GR. *Mathematical statistics: A unified introduction*. Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA; 1999. ISBN: 0-387-98621-9.
- [2] Dekking FM, Lopuhaä, Kraaikamp C, Meester LE. *A modern introduction to probability and statistics. Understanding why and how*. Springer-Verlag London Limited. Springer Science + Business Media, USA; 2005. ISBN-10: 1-85233-896-2, ISBN-13: 978-1-85233-896-1.
- [3] Hogg RV, McKean JW, Craig AT. *Introduction to mathematical statistics*. 6th Edition. Pearson Education International. Pearson Prentice Hall, Pearson, Inc., Upper Saddle River, NJ 07458, USA; 2005a. ISBN: 0-13-122605-3.
Available:<http://lib.stat.cmu.edu/R/CRAN/>
- [4] Hogg RV, McKean JW, Craig AT. *Introduction to mathematical statistics*. 5th Edition. Pearson Education International. Pearson Prentice Hall, Pearson, Inc., Upper Saddle River, NJ 07458, USA; 1995b. ISBN: 7-04-015557-5.
- [5] Hogg RV, Craig AT. *Introduction to mathematical statistics*. 4th Edition. Macmillan Publishing Co., Inc., 866 Third Avenue, New York, New York 10022, USA. Collier Macmillan Publishers, London; 1978. ISBN: 0-02-355210-9.
- [6] Spiegel MR, Schiller J, Srinivasan RA. *Probability and statistics. Shaum's easy outline*. McGraw-Hill Companies, Inc., USA; 2001. ISBN: 0-07-139838-4. e-ISBN: 0-07-138341-7.
- [7] Rohatgi VK, Saleh EAK. *An introduction to probability and statistics*. 2nd Edition. John Wiley & Sons, Inc., USA; 2001.
- [8] Bain LJ. *Introduction to probability and mathematical statistics*. 2nd Edition. Duxbury Thomson Learning. Forest Lodge Road, Pacific Grove, CA 93950 USA, www.duxbury.com 1-800-423-0563; 1992.
- [9] Bickel PK, Doksum KA. *Mathematical statistics: Basic ideas and selected topics*. 2nd Edition, vol. I. Prentice Hall, Inc., Upper Saddle River, New Jersey, USA, 07458; 2001. ISBN: 0-13-850363-X.
- [10] Shao J. *Mathematical statistics: Exercises and solutions*. Springer Science + Business Media, Inc., 233 Spring Street, New York, NY 10013, USA; 2005. ISBN-10: 0-387-24970-2, ISBN-13: 978-0387-24970-4.
- [11] Montgomery DC, Runger GC. *Student workbook with solutions. Applied statistics and probability for engineers*. 3rd Edition. John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, USA; 2003. ISBN: 0-471-42682-2.
- [12] Wilks SS. *Mathematical statistics*. John Wiley & Sons, Inc., New York, USA; 1962.
- [13] Wilks SS. *Mathematical statistics*. Princeton University Press, Princeton, New Jersey; 1947.
- [14] Ross SM. *Introduction to probability and statistics for engineers and scientists*. 3rd Edition. Elsevier Academic Press, 200 Wheeler Road, 6th Floor Burlington, MA 01803, USA; 2004.
- [15] Bernstein S, Bernstein R. *Schaum's outline of theory and problems of elements of statistics: Descriptive statistics and probability*. McGraw-Hill Companies, Inc.; 1999. ISBN: 0-07-005023-6.

- [16] Mwangi WP. Division of zero by itself-division of zero by itself has unique solution. Pure and Applied Mathematics Journal. 2018;7(3):20-36.
DOI: 10.11648/j.pamj.20180703.11
- [17] Papoulis A, Pillai SU. Solutions manual to accompany probability, random variables and stochastic processes. 4th Edition. McGraw-Hill Higher Education, The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY 10020; 2002.
- [18] Schay G. Introduction to probability with statistical applications. Birkhäuser Boston, c/o Springer Science + Business Media LLC, 233 Spring Street, New York, NY 10013, USA; 2007. ISBN-13: 978-0-8176-4497-0, e-ISBN-13: 978-0-8176-4591-5.
Available:www.birkhauser.com

© 2020 Mwangi; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)

<http://www.sdiarticle4.com/review-history/53235>