

# Probability Distributions Related to Modeling Epidemic Spread Data “COVID-19 Status and Developments”

Bachioua Lahcene

Department of Basic Sciences Prep. Year, P.O. Box 2440, University of Ha'il, Ha'il, Saudi Arabia.

**How to cite this paper:** Bachioua Lahcene. (2021) Probability Distributions Related to Modeling Epidemic Spread Data “COVID-19 Status and Developments”. *Journal of Applied Mathematics and Computation*, 5(2), 134-144. DOI: 10.26855/jamc.2021.06.008

**Received:** April 20, 2021

**Accepted:** May 18, 2021

**Published:** June 4, 2021

\***Corresponding author:** Bachioua Lahcene, Department of Basic Sciences Prep. Year, P.O. Box 2440, University of Ha'il, Ha'il, Saudi Arabia.  
**Email:** drbachioua@gmail.com

## Abstract

This paper introduces some elements of mathematical modeling to show suspected, recovered and deceased COVID-19 patients, and the chains adopted in deploying operations around the world, and it is still not definitively known when people acquire long-term immunity, but the formulation of proposed models for probability distributions allows for the definition of the finite difference scheme. Non-standard, some graphical results have been presented to some carefully selected countries. The results showed that health safety plans and isolation of infected and suspected humans, in general, is the only way so far that can reduce the risk of the spread of this epidemic in the near future, and also through statistical analysis using fitted models that revealed a high and unstable exponential growth of the number of confirmed cases. And deaths and cases that responded to treatment based on the results of experimental COVID-19 predictions, and it is expected that the number of infected cases and daily deaths will stabilize after the measures taken by most countries, and this situation will continue until the largest number of people are vaccinated in order to obtain herd immunity, and control the causes. As the epidemic spreads like human gatherings and contact, the results of this work will be useful to practitioners in various fields of theoretical and applied sciences.

## Keywords

Probability Distributions, Normal Distribution, Lognormal Distribution, Rayleigh Distribution, Exponential and Gamma Distribution, Generating, Algorithm, Modeling, Epidemic, Estimation Methods, COVID-19

## 1. Introduction

The history of statistics in the modern sense dates from the mid-17th century, with the term probability and statistics itself coined in 1749 in German, although there have been changes to the interpretation of the word over time. A continuous random variable is one that is measured on a continuous scale. Abraham de Moivre's (1667-1754) first work in probability density function and cumulative distribution function of continuous random variables and described the normal distribution in 1733. In 1812, Pierre de Laplace (1749-1827) introduced a host of new ideas and mathematical techniques in his book, “Théorie Analytique des Probabilités” [1].

Before Laplace, probability theory was solely concerned with developing a mathematical analysis of games of chance. Laplace applied probabilistic ideas to many scientific and practical problems. The theory of errors, actuarial mathematics, and statistical mechanics are examples of some of the important applications of probability theory developed in the 19th

century [1].

Statisticians are constantly developing many statistical methods now for use in medical studies in pathology, which are concerned with studying a disease and knowing the degree of its spread in certain areas, the rate of infection and the rate of cure.

The models describe the various probability distributions related to the epidemiological spread data represented by the adopted model, and its predictions mainly related to the random mean, making it necessary to analyse the spread phenomena using approved statistical methods which in most cases can be satisfactorily describing changes in the propagation parameters of the phenomenon in time and space by means of known statistical distributions; Which is known for the basic characteristics of the probability distributions most used in statistical studies of the spread of epidemics by using the statistical information related to the prevalence modelling from the use of available data to predict the performance parameters of the estimated systems of the general model using a step-by-step procedure in order to approximate the complementary cumulative distribution by means of complementary cumulative distribution.

The data accumulated from the experiment showed that the information on modelling the spread of epidemics using average values describes the changes of the spread parameters in time and space in a satisfactory description by means of known statistical distributions, which are sufficient to characterize the performance of the epidemiological model systems and must also take into account the changes in time and place and the procedural precautions taken.

The dynamic behaviour of the phenomenon of the spread of epidemics and their causes plays a set of interfering elements in the processes of injuries and plays a decisive role in analysing the reliability of preventive systems, and in choosing system parameters that define the range of changes in the phenomenon of the spread of the epidemic in order to be able to determine the parameters of the adopted model to track changes in this situation, such as the pattern of protection ratio from interventions, prevention measures and treatment method.

To describe the performance of prevention and treatment systems, it is often sufficient to note the time series of fluctuations in the spread of the epidemic, the results of prevention and treatment, and characterize these fluctuations as a random process.

The data modelling associated with ophthalmology, the variability of its prevalence, and the prediction of the performance of the model systems for prevention and treatment change from one country to another also require knowledge of the mechanisms of interaction between the procedures adopted in each country, and the composition of the results of the adopted model and its predictive state change a lot according to the place and time, and therefore the modelling of interactions is required.

Mathematics provides mathematical models useful for understanding the behaviour of infection as it enters society and for investigating conditions that will be eliminated or persistent. Now, COVID-19 is a major concern for researchers, governments and all people due to the high prevalence of infection and the large number of deaths that have occurred. In December 2019, COVID-19 was first reported in Wuhan, China, and it is an infectious disease caused by the newly discovered Coronavirus. The virus that causes COVID-19 is primarily transmitted through droplets generated when an infected person coughs sneezes or exhales. These drops are too heavy to be suspended in the air and to fall quickly onto floors or surfaces. The number of confirmed cases of COVID-19 has reached nearly four million in 187 countries, and nearly 295,000 people have died from this virus [2].

The ophthalmology uses statistical methods in order to distinguish the various physical parameters for each case that describe through the parameters that have been evaluated, that determine the behaviour of the epidemic signal and the interactions between infection processes and the data of the cases that have been treated, and that link these parameters between them, and may provide background probability distributions. Common statistical methods for predicting prevalence and prevention outcomes resulting from the approved data, and in determining the results associated with them. Stigler shows how statistics arose from the interplay of mathematical concepts and the needs of several applied sciences [2]. He addresses many intriguing questions: How did scientists learn to combine measurements made under different conditions? And how were they led to use probability theory to measure the accuracy of the result? Why were statistical methods used successfully in several applied sciences?

## 2. The Probability Distribution Model

During an infectious disease outbreak, biases in the data and complexities of the underlying dynamics pose significant challenges in mathematically modelling the outbreak and designing policy [2]. The random processes are often described by the probability density function, and in other cases in terms of the cumulative distribution function, and the probability density function  $p(x)$  for the random variable  $X$ , so that the probability value for the random variable representing the phenomenon to take the value of  $x$  in the infinitesimal interval Between  $x$  and  $x + dx$ ,  $p(x)$  is  $dx$ , and the result is given in terms of the cumulative distribution function denoted by the symbol  $F(x)$ , and represents the probability that the variable will take a value less than  $x$ , and the relationship between the probability density function and the cumulative distribution

function is shown, and as follows:

$$p(x) = \frac{d}{dx}[F(x)] \quad \text{or} \quad F(x) = \int_c^x p(t) \, dt \quad (1)$$

Where:  $c$  is the minimum value that  $t$  can take.

The situation is similar for probability, except instead of calling it a mass density, we call it a probability density. The function that describes the probability density is called a probability density function (pdf), denoted by  $f_X(x)$ .

Thinking carefully about the sand in the aquarium, one can see that probability density, *i.e.*, the probability per unit change of the random variable can be expressed mathematically as;

$$f_X(x) = \lim_{\Delta x \rightarrow 0} \frac{P[x \leq X \leq x + \Delta x]}{\Delta x} \quad (2)$$

Probability density functions can take a variety of shapes. These are in fact two important properties of all pdfs;

1.  $f_X(x) \geq 0$
2.  $\int_{-\infty}^{+\infty} f_X(x) \, dx = 1$

The first one says that the “sand mass per unit length”, *i.e.*, the probability per unit random variable, must be non-negative; this is a restatement of the first axiom. The second one says that the total “mass of sand”, *i.e.*, the total probability, must be 1; this is a restatement of the second axiom.

Probability evaluation using pdfs is just like mass evaluation using the mass density function. To see this in a more rigorous fashion, let's return to the definition of a pdf in equation (2). Multiplying both sides by  $\Delta x$ , we obtain;

$$\lim_{\Delta x \rightarrow 0} f_X(x) \Delta x = \lim_{\Delta x \rightarrow 0} P[x \leq X \leq x + \Delta x] \quad \text{or} \quad f_X(x) dx = \lim_{\Delta x \rightarrow 0} P[x \leq X \leq x + \Delta x]$$

We can lift this requirement, letting  $\Delta x$  be as large as we like, by summing the corresponding products on the left hand side through integration;

$$\int_x^{x+\Delta x} f_X(x) dx = P[x \leq X \leq x + \Delta x] \quad (3)$$

Letting  $x = a$  and  $x + \Delta x = b$  be any two real numbers, then  $P[a \leq X \leq b] = \int_a^b f_X(x) dx$ .

Since the interval  $a \leq X \leq b$  corresponds to an event, equation indicates that the probability of this event may be computed as the area under the curve of the PDF in the interval. As a consequence of the fact that for continuous random variable,  $P[X = x] = 0$ , the interval of integration may be closed or open, *i.e.*;

$$P[a \leq X \leq b] = P[a < X < b] = P[a \leq X < b] = P[a < X \leq b]$$

This is not the case if the random variable is of the mixed type, *i.e.*, if it contains some discontinuities so that it is partially discrete. In this case, we use;

$$P[a < X \leq b] = \int_a^b f_X(x) dx \quad (4)$$

We are frequently interested in the case where the lower bound is  $-\infty$  and the upper bound is finite, say  $b$ , because this is just the probability that the random variable is less than or equal to  $b$ , *i.e.*;

$$P[-\infty \leq X \leq b] = P[X \leq b] \quad (5)$$

### 3. Types of Models of Probability Distributions

There are models that have a mathematical background that the researcher uses in order to predict results in specific situations when the conditions necessary to calculate the probability of each outcome are met, predict the average of long-term results, and estimate the variance in the results of random variables that represent the target phenomenon without the need to obtain all results.

The actual random variables that he is interested in studying, and for generalizing the results of the limited analysis to broader behaviors, which is a guide used to fit a random variable in order to generalize the behavior of the studied phenomenon, and in this way it helps the researcher to save time and money and reach more reliable results.

There are many models of probability distributions for different situations, and the main point is that the researcher must select the appropriate model that fits with the available data, facilitate the initial modeling process for this data, explain what he is trying to understand, and study using different assumptions so that the results of the model have the correct probabilities, and test the model and checking it against the real data of the random variable, but there are different types of random variables, the most important of which are:

### 3.1. The Extended Normal Distribution Model

This distribution applies to data that can be represented by a continuous variable with any algebraic sign. The probability density is of the type:

$$p(x) = e^{-T(x)} \tag{5}$$

Where  $T(x)$  is a non-negative quadratic polynomial, so if we use as the parameters of the mean  $m$  and the standard deviation  $\sigma$ , then  $p(x)$  is written in the following usual way:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] \tag{6}$$

Subsequently:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2\right] dt = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{x-m}{\sigma\sqrt{2}}\right)\right] \tag{7}$$

Where  $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ . A normal distribution is encountered when the studied magnitude values are the result

of the additive effect of several random causes, each of which is of relatively little significance. The connected lines represent the functions  $p(x)$  and  $F(x)$  where  $m$  is equal to zero and  $\sigma$  equals the integer one in the case of the tabular standard distribution, which expresses the correspondence between  $x$ , and  $F(x)$  for some approximate values of  $x$  or  $F(x)$ , as well as the cumulative normal distribution.  $F(x)$  in abbreviated form is for the same condition. In complex cases, the stretched distribution can be used with a greater number of parameters, and mixture models can be adopted [3].

### 3.2. The Extended Lognormal Distribution Model

It is the distribution of a positive variable in which the logarithm has a normal distribution. Then, we can directly write the probability density function and the cumulative distribution function as follows:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{x} \exp\left[-\frac{1}{2}\left(\frac{\ln x - m}{\sigma}\right)^2\right] \tag{8}$$

Subsequently:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{1}{t} \exp\left[-\frac{1}{2}\left(\frac{\ln t - m}{\sigma}\right)^2\right] dt = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{\ln x - m}{\sigma\sqrt{2}}\right)\right] \tag{9}$$

However, in these two equations  $m$  and  $\sigma$  are the mean and standard deviation of the logarithm of the variable  $x$  and not of the variable  $x$  itself. In comparison with the normal distribution, it can be considered that the logarithmic normal distribution means that the numerical values of the variable are the result of several causes of little single importance, but with a multiplier effect. In particular, the mean value, the median value, and the most probable value (usually called the mode) are not the same, and the characteristic quantities of the numerical variable  $x$  are as follows:

1. Most probable value:  $\exp(m - \sigma^2)$
2. Average value:  $\exp(m)$
3. Average value  $\exp\left(m + \frac{\sigma^2}{2}\right)$

4. RMS value  $\exp(m + \sigma^2)$
5. Standard deviation:  $\exp\left(m + \frac{\sigma^2}{2}\right) \sqrt{\exp(\sigma^2) - 1}$

In the advanced cases, the stretched distribution can be used with a greater number of parameters, and mixed models can be adopted [4].

### 3.3. The Extended Rayleigh Distribution Model

The Rayleigh distribution applies to a positive continuous variable, and it is related to the normal distribution considering that there is a two-dimensional normal distribution of two independent variables  $y$  and  $z$  with a mean of zero and the same standard deviation  $\sigma$  and therefore the random variable  $x$  is as follows:

$$x = \sqrt{y^2 + z^2} \quad (10)$$

This random variable has a Rayleigh distribution. The most probable value of  $x$  is. The Rayleigh distribution represents the vector length distribution, which is the sum of a large number of isotropic vectors whose phases have a uniform distribution. The probability density function and the cumulative distribution function are given by:

$$p(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (11)$$

Subsequently:

$$F(x) = 1 - \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (12)$$

The distinct values of the variable are as most probable value  $\frac{b}{\sqrt{2}}$ ; middle value:  $\frac{b}{2} \ln 2 \approx 0.833b$ , average value  $\frac{b}{2} \sqrt{\pi} \approx 0.886b$ , RMS value:  $b$ , and Standard deviation  $b \sqrt{1 - \pi/4} \approx 0.463b$ .

The Rayleigh distribution is often used close to the origin, that is, for low values of  $x$ . In this case, it will be:

$$F(x) \approx \frac{x^2}{b^2} \quad (13)$$

In the advanced cases, the stretched distribution can be used with a greater number of parameters, and mixed models can be adopted [5].

### 3.4. The Extended Exponential and Gamma Distribution Model

In contrast to the previous distributions that are related to the generalized Gaussian distribution, the distribution of the generalized exponential distribution stems from a generalized gamma distribution, which constitutes a generalization to it. It is applied to an infinite positive variable, and the probability density function is

$$p(x) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\alpha x} \quad (13)$$

$\Gamma$  is the Euler function of the second degree. This distribution depends on two parameters  $\alpha$  and  $\nu$ . However,  $\alpha$  is only a parameter of the variable  $x$ . The characteristic values of the variable are the Median value  $\frac{\nu}{\alpha}$ , RMS value  $\frac{\sqrt{\nu(1+\nu)}}{\alpha}$ , and standard deviation  $\frac{\sqrt{\nu}}{\alpha}$ . The cumulative distribution of the gamma function complementary to small values shows that there is little chance that the variable  $x$  is much greater than zero. This explains in particular the use of the gamma distribution to represent the rates of increase in the spread of the epidemic since the total percentage of the time of the acceleration of the spread of the epidemic generally ranges between 2 and 10%. In the advanced cases, In complex cases, the stretched distribution can be used with a greater number of parameters, and mixture models can also be adopted [6], [7].

### 4. Estimation Algorithm for Generating Predictive Values

It is noted from the approved probability models that there is a correlation between them, and for this it is possible to think of building an algorithm that starts from the first step and then moves to a later step aimed at the last to approximate the complementary cumulative distribution by means of a complementary cumulative distribution according to an algorithm, and in order to clarify the idea We build the algorithm on the log-normal distribution, as follows [8]:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{1}{t} \exp\left[-\frac{1}{2}\left(\frac{\ln t - m}{\sigma}\right)^2\right] dt = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{\ln x - m}{\sigma\sqrt{2}}\right)\right]$$

Alternatively, likewise:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln x - m}{\sigma}} \exp\left[-\frac{t^2}{2}\right] dt$$

Likewise, a lognormal complementary cumulative distribution is defined as follows:

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_x^{\infty} \frac{1}{t} \exp\left[-\frac{1}{2}\left(\frac{\ln t - m}{\sigma}\right)^2\right] dt = \frac{1}{2} \operatorname{erfc}\left(\frac{\ln x - m}{\sigma\sqrt{2}}\right) = \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{\ln x - m}{\sigma\sqrt{2}}\right)\right]$$

Alternatively, likewise:

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln x - m}{\sigma}}^{\infty} \exp\left[-\frac{t^2}{2}\right] dt = Q\left(\frac{\ln x - m}{\sigma}\right)$$

Where  $Q(\cdot)$  is the integral of the natural complementary cumulative probability [9]. Estimate the parameters  $m$  and  $\sigma$  can from a set of  $n$  pairs of  $(x_i, G_i)$  as follows:

The natural logarithmic parameters  $m$  and  $\sigma$  are estimated as follows:

**Step 1:** Create a set of  $n$  pairs of  $(x_i, G_i)$  where  $G_i$  is the probability that  $x_i$  exceeds.

**Step 2:** Convert the set of  $n$  pairs from  $(x_i, G_i)$  to  $(\ln x_i, Z_i)$  where:

$$Z_i = \sqrt{2} \operatorname{erfc}^{-1}(2G_i) = \sqrt{2} \operatorname{erf}^{-1}(1 - 2G_i) \quad \text{Equivalent} \quad Z_i = Q^{-1}(G_i)$$

$$Z_i = \sqrt{2} \operatorname{erfc}^{-1}(2G_i) = \sqrt{2} \operatorname{erf}^{-1}(1 - 2G_i) \quad \text{Equivalent} \quad Z_i = Q^{-1}(G_i)$$

**Step 3:** Determine the two variables  $m$  and  $\sigma$  grant them during the least squares matching procedure of the linear function:

$$\ln x_i = \sigma Z_i + m$$

As follows:

$$\sigma = \frac{n \sum_{i=1}^n Z_i \ln x_i - \sum_{i=1}^n Z_i \sum_{i=1}^n \ln x_i}{n \sum_{i=1}^n Z_i^2 - \left[\sum_{i=1}^n Z_i\right]^2}$$

$$m = \frac{\sum_{i=1}^n \ln x_i - \sigma \sum_{i=1}^n Z_i}{n}$$

### 5. Mathematics Models Designed to Track the Spread of COVID-19:

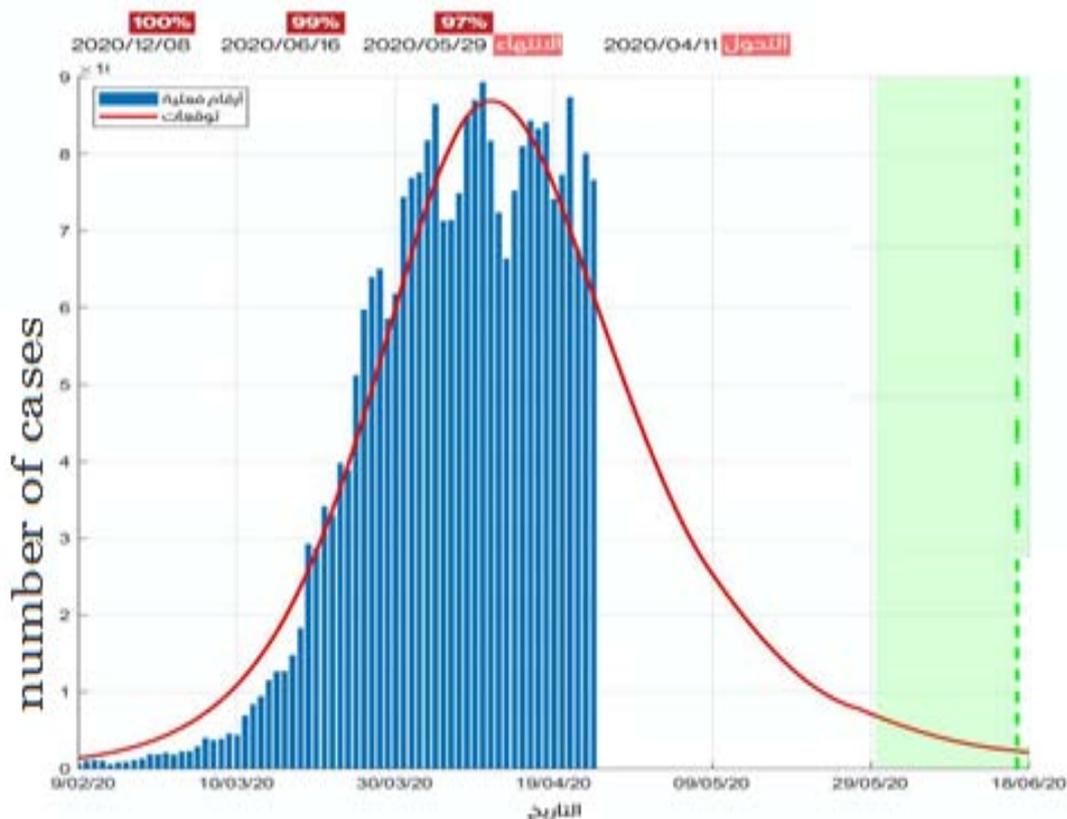
Mathematical models designed to track the spread of infectious diseases have been clarified, and with great reliance on available data to predict the probability of the spread of diseases around the world, which is caused by the COVID-19 [10], in the first days of the spread of a new virus, when reliable data are still scarce, researchers resort to Mathematical

models that can predict the path of spread and the consequences thereof through computer models that adopt programmatic statistical equations, and through which they calculate the probability of disease transmission from infected people to others, and from one region to another, and through modern computing capabilities, researchers can quickly include multiple inputs in these models, such as the ability of a specific disease to transmit from one person to another, and the movement patterns of air and land travelers among those likely to have the disease.

During the second semester of the spread of the COVID-19 epidemic, statistician Sally Cripps from the “University of Washington” built a prediction model that they called HIME [10]. The aforementioned model relied on predicting what the epidemic curve would look like when new foci appear in new areas in the United States of America, and the American model relied on a few variables, including social measures of divergence, assuming that the epidemic conditions in some states will be similar with conditions in the states that record new infections. The model was also fed mainly from Chinese data, especially since China imposed harsh measures and applied social distancing strictly and for long periods.

Sally Cripps explicitly acknowledged that his results were imprecise, as the analysis found that some of the model’s predictions were underestimated, indicating that she feared that lack of accuracy would undermine public confidence in science. Krebs pointed out that “it is not possible to rely on this model or other in major political decisions”.

The following are a number of graphs that represent the cases of the spread of the epidemic, taken from the website (<https://covid19.who.int/table>). The model requires a set of variables, including data on disease prevalence, number of infected people, and time of endemicity of the epidemic in each country. These models show the differential links between the various variables that affect and are affected by the developments in the data of this epidemic, and the output of the model equation appears with a result that predicts the stability of the spread process, and its subsequent contraction until the end of the epidemic and the possibilities of controlling it. The attached graphics show the chain of spread, and the expectations have shown by each case and in each country referred to regarding ending the COVID-19 epidemic in a number of countries (see Figure 1):



**Figure 1. The number of cases in the world.**

The rapid spread of the epidemic in the United States of America was very fast, and it caused loss of life despite the economic and social capabilities that this country possesses, and it ranked first in the number of injuries and the number of deaths resulting from that (see Figure 2).

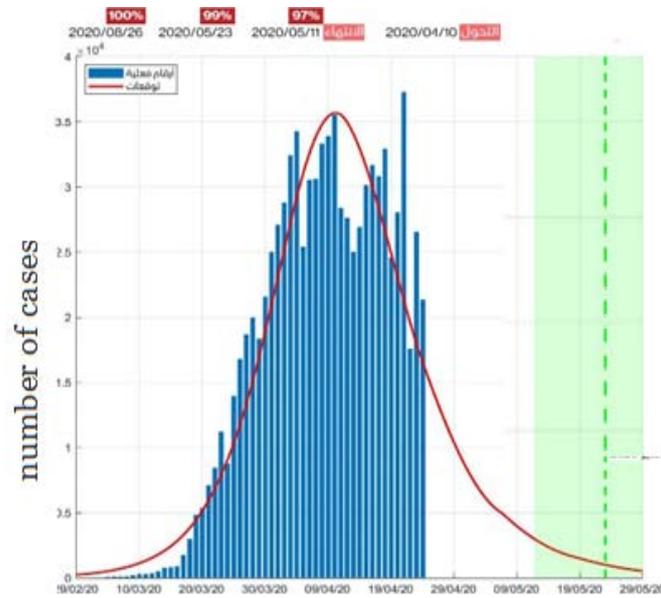


Figure 2. The number of cases in the USA.

It is also noticed that the epidemic in Britain spread rapidly, especially after the emergence of a new, modified series, and it caused loss of life, and a faster spread appeared, despite the approved health protocols, whether it was the closure of gatherings between individuals, or the precautions for social distancing, and it occupied the first ranks in the number of deaths per population (see Figure 3).

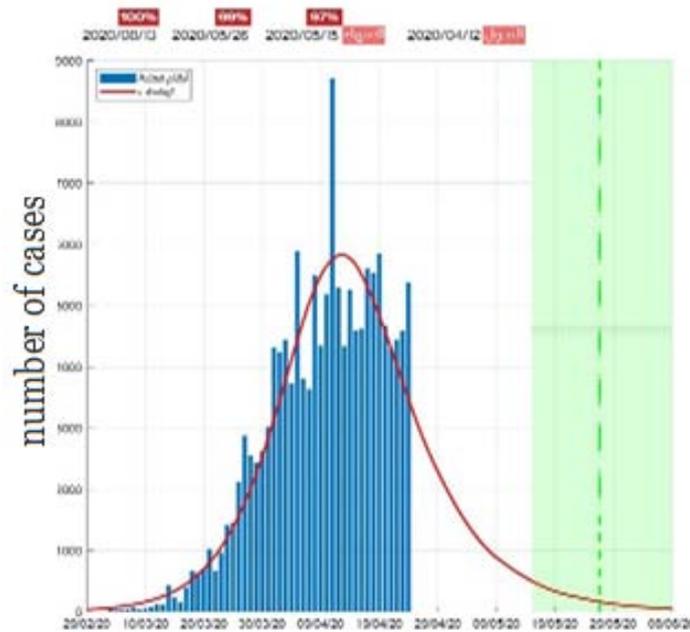


Figure 3. The number of cases in the United Kingdom.

It is also noticed that the epidemic in China accepted severe precautionary measures, and this allowed the state to control the spread of the epidemic despite the population, and the epidemic spread rapidly in the beginning, and it was controlled during the first three months after the announcement of the first cases that appeared in the first epidemic focus discovered, and he returned the curve to its natural shape despite the emergence of regressive jumps of the spread of the epidemic in special cases after the emergence of a new chain, and the closure of gatherings between individuals, or the precautions for social distancing and government strictness had positive results at the expense of the natural movement between the population (see Figure 4).

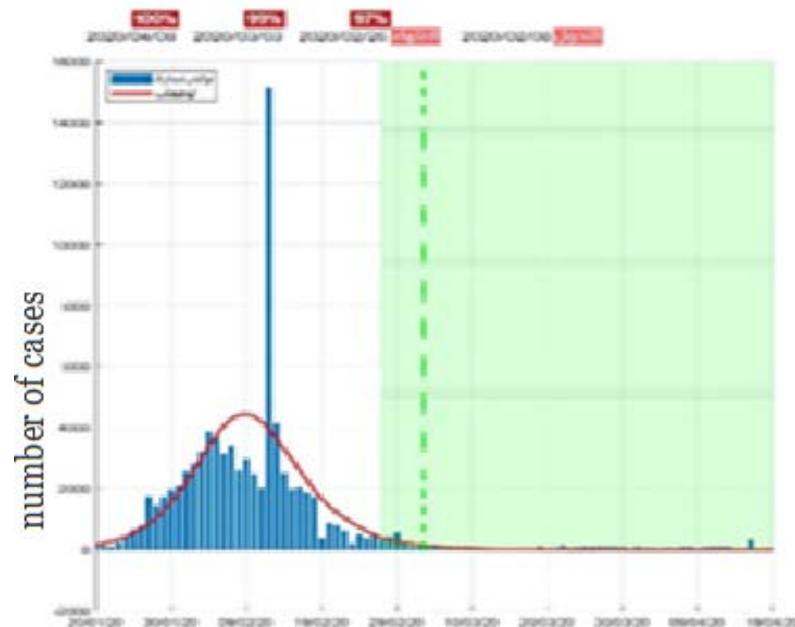


Figure 4. The number of cases in the China.

It is noticed that the epidemic has spread in the three Arab countries (KSA, Egypt, and UEA). The epidemic has spread at a close pace despite the differences between these countries and the strict precautions that have been adopted regarding the movement of movement and movement to these countries, and the announced figures indicate that the epidemic has accepted severe precautionary measures. This allowed these countries to have reasonable results for other countries in the world, and the epidemic spread rapidly in the beginning, and it was controlled during the first six months, and the closure of gatherings between individuals, or the precautions for social distancing and governmental strictness had positive results (see Figures 5, 6 and 7).

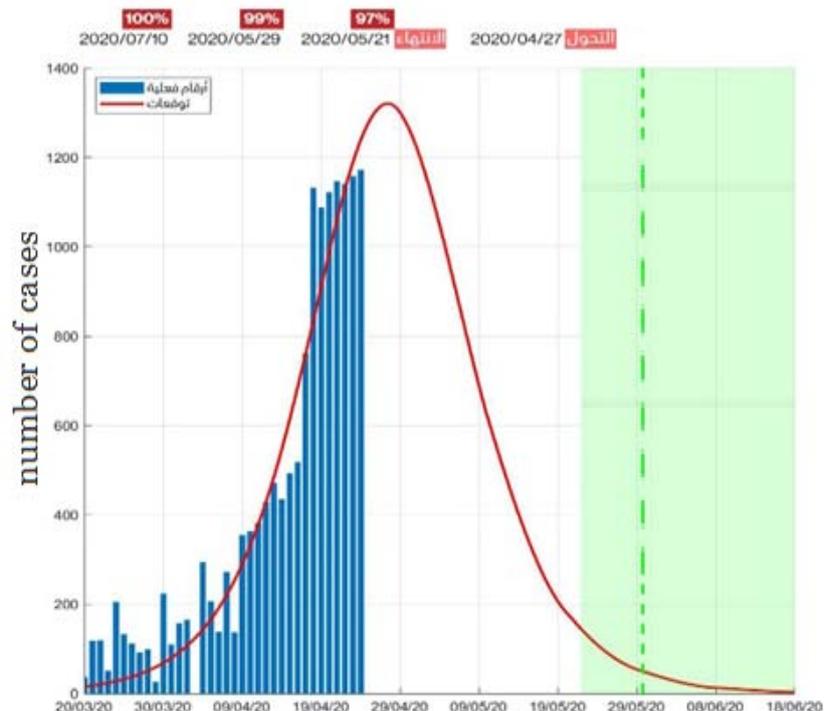


Figure 5. The number of cases in the KSA.



Figure 6. The number of cases in the Egypt.



Figure 7. The number of cases in the UEA.

Although the model that was built could change its inputs and update the basic data using the new data, the researchers were not able to reach the results of the dates of the end of the pandemic in each country, and for this it can be observed through the selected figures for the seven countries as a sample of the world, it is clear that there are differences in The spread of the epidemic, and also the consequences of it, whether the number of deaths, or the speed of spread, and this is due to the type of political systems, and the possibilities available in each country, and the method of prediction and its guesses is always confirmed according to the accuracy of this mathematical model.

## 6. Findings and Recommendations

The World Health Organization indicates that the epidemic is “Is far from finished”, and humans may have to live with some restrictions adopted for a period of two years or more to keep the COVID-19 under control, especially after the emergence of new strains of the COVID-19 to the announcement of new closures in various Countries around the world, and these measures have so far been able to limit the spread of infection, and expectations indicate that closure and separation measures will continue despite the emergence of vaccines, and large experiments on vaccines will continue and

improve, and people must continue to wear masks, follow hygiene rules and maintain spacing.

A team of epidemiologists expects the spread of the new COVID-19, and experts have highlighted that the virus will spread for up to two years until herd immunity develops among the population of societies, and they base their expectations on these projections on models provided by research centers such as those at Washington University and the Imperial College in the capital, Britain, London, whose figures showed that the virus will cause the death of millions.

That is why everyone should prepare for a large-scale transmission of infection and not be complacent, otherwise, the virus may come back, and the solution available so far is strict preventive measures that must be taken by most countries of the world, starting with the closure of regions and cities and ending with the suspension of studies and preventing gatherings, and continuing with vaccinations for the largest number possible from the segments of society, especially the vulnerable groups, until reaching the herd immunity for society, until the community reaches the herd immunity, and continuing to develop mathematical models that predict the map of the spread of the Corona of all kinds and the consequences of it.

## References

- [1] Stigler, Stephen M. (1990). *The History of Statistics: The Measurement of Uncertainty before 1900*. Academic Trade, booksellers, U.S. and Canada.
- [2] Zeb Anwar, Alzahrani Ebraheem, Erturk Vedat Suat, and Zaman Gul. (2020). *Mathematical Model for Covid-19 Disease 2019 (COVID-19) Containing Isolation Class*. *BioMed Research International*, 145(2020), 16-30.
- [3] Bachioua Lahcene. (2018). *On Extended Normal Distribution Model with Application in Health Care*. *International Journal of Statistics in Medical Research*, N(7), pp. 88-95.
- [4] Bachioua Lahcene, (2019). *On Extended Normal Inverse Gaussian Distribution: Theory, Methodology, Properties and Applications*. *American Journal of Applied Mathematics and Statistics*, Vol. (7), No. (6), 224-230.
- [5] Bachioua Lahcene. (2018). *On Recent Modifications of Extended Rayleigh Distribution and its Applications*. *JP Journal of Fundamental and Applied Statistics*, Vol. (12), Issue (1): pp. 1-13.
- [6] Bachioua Lahcene. (2020). *On Extended Exponential Distribution: Properties and Applications In Tracking the Pandemic Covid-19*. *SunText Review of Medical and Clinical Research*, Vol. (1), No. (2), pp. 1-8.
- [7] Bachioua Lahcene. (2021). *A New Extended-Gamma Family of Distributions: Properties and Applications*. *Journal of Applied Mathematics and Computation*, 5(1), pp. 9-17.
- [8] Bachioua Lahcene. (2020). *Extended Lognormal Distribution: Properties and Applications*. *World Scientific News*, 145(2020), 16-30.
- [9] Bachioua Lahcene. (2018). *On Recent Modifications of Extended Weibull Families Distributions and Its Applications*. *Asian Journal of Fuzzy and Applied Mathematics*, Vol. (06), No. (01), February, pp. 1-11.
- [10] Overtonak, Christopher E., Stagea, Helena B. (2020). *Using Statistics and Mathematical Modelling to Understand Infectious Disease Outbreaks: COVID-19 as an Example*. *Infectious Disease Modelling*, Vol (5), pp. 409-441.