Logistic Model to Predict the Contagion of Covid-19 in Mexico

Juan Bacilio Guerrero Escamilla[1], Sócrates López Pérez[1], Sonia Bass Zavala[1]

[1] Autonomous University of the State of Hidalgo, Hidalgo State, México

Correspondence: Juan Bacilio Guerrero Escamilla, Autonomous University of the State of Hidalgo, Hidalgo State, México

Abstract

The objective of this work is to predict the risk of contracting Covid - 19 in the Mexican population, by means of the construction of a multilogistical model, based on the Ministry of Health applied to patients who presented with the symptoms that encompass the disease. With this probabilistic model, it was possible to estimate the degree of contracting Covid -19, taking as a reference the health conditions of the patients, and with it, medically influence to counteract the disease.

Keywords: covid-19, logistic model, contagion, México

1. Introduction

1.1 Introduce the Problem

In the first report of the World Health Organization (WHO), the emergence of cases of pneumonia was reported on December 31st, 2019, with an unknown etiology and detected in the city of Wuhan, province of Hubei of China. In the course of the next 3 days, the national authorities from China informed to the WHO, a total number of 44 patients, without identifying the causal agent, but with an identification of a new type of coronavirus.

By the year 2020, China shared the genetic sequence, which was used to take part with other countries and to develop specific diagnostic kits (WHO, 2020). From there, Thailand reported the first imported case, as well as, Japan due to the same dates. By the 20th of January 2020, Korea informed to the WHO the first emergence of coronavirus cases, and also it was confirmed by the Wuhan laboratory.

According to the most recent reports by WHO on April 2020, the situation of transmission was considered as a world scene. This setting showed a total of confirmed cases of countries 2,397,216 total of new confirmed cases 83,006, total of death 162,956 and a total of new deaths 5,109. Taking into account this, with the reports of other countries, territories or areas with cases and deaths of COVID-19, and laboratory confirmed, with data starting on 21 April 2020, emitted by the WHO.

For the case of Las Americas (with data taken from WHO) total confirmed cases were showed of 925.291, in 54 affected countries and with 44.775 confirmed deaths. To the specific case of Mexico, we have that in the same date, there were 9,501 cumulative confirmed cases 3,185 active confirmed cases, and 857 deaths.

The specialists considered that from the 7,778 million of habitants of the world,70% of these have probability of being infected, provided that, there is no intervention of the authorities, governments or the same population. From this 70% of infected people Rosen, C. (n.d) in turn of these 80% of them would be asymptomatic (British Medical Journal, April 2020).

The other 20% of the population could show severe symptoms, and taking as a reference this population, it is likely of infection of 14% of them because they would develop more serious symptoms with requirement for hospitalization, the other 6% of them would require to access to intensive cares. From this 6% the half of the 3% would recover intensive cares; although, the other 3% would die. This last would be the fatality rate of the virus SARS-CoV-2.

In light of this situation, analyses must be done to define the settings of pandemic transmission, mortality rate, and the condition that the population has been infected. Nevertheless, this situation is not taken into account the intervention and without any action to impact on the behavior of transmission. As a consequence, diverse analyses and modelling that scientific community must do the necessary to predict the future of spread of the virus. Furthermore, focusing intervention of authorities, calculation of resources, sustained design of public policies (British Medical Journal; Devi Sridhar, Maimuna S. Majumder; April 2020).

1.2 Explore Importance of the Problem

This is a quick response to the scientific community to offer an efficient response with technical tools, that define different settings with certainty in order to decide that model should be followed, and what factors should be included, it is a political option. Nonetheless, in any initial stage there was not enough information to construct these models, subsequently, we should trust on models that adapt these deficiencies of data meanwhile important epidemiological knowledge can be produced.

For Mexico, there is the advantage of the previous behavior of the virus in China and Europe, it helps with more certainty in the estimates of the mortality rate of Covid-19 and how fast it can spread, as well as how different these estimates can be in different populations, as the probabilities of risk were quantified in Germany (Süddeutsche Zeitung; 27. März 2020).

For this reason, the design of these models having regard to the conditions characteristic of each country, in this case Mexico, in respect of its modelling of State, Political Class, implementing public policies, the National Legal System, the National Health System, the Public Healthcare System, resources, quality of care, quality of citizenship, capability in social organization, demographic composition and territorial distribution, as well as, sociodemographic characteristics ,economic composition ,distribution of wealth, levels of poverty, including Territorial Mobility Index. (Gravitational model, Megalopolis; López Pérez and Guerrero Escamilla; Mexico 2015).

For the specific case of Mexico, morbid conditions of the population must be considered to any model of analysis. Thus, the risk level will change based on the comorbidity composition, in other words, there are more diseases as a whole the age.

In this way, the increase of risk would be for people who suffer from Diabetes, because this group develops a viral infection, and this is harder to treat due to

blood glucose levels and possibly, emerging complications of Diabetes (International Diabetes Federation). Making reference of those who are on Hypertension treatment, there are evidences that affirm that Angiotensin Converting Enzyme has some changes when we compare with people that do not have hypertension which can explain this because they develop a higher way, whereby these are the cases that must be implemented health-care measures with intensive care (May Clinic proceedings; Coronavirus COVID-19 Resource Center; April 2020).Focusing on obesity, it has diverse complications in the health process because of apnea and decrease of oxygen levels that get into the cells , until having difficulties due to fat content to ensure attention to intubation ,the respirator fit ,and the mobilization of hospitalization, as well as, other cardiovascular diseases.

1.3 Describe Relevant Scholarship

These models, among the one is being proposed in this present research, has as main purpose to analyze wave of infections to control them, and to minimize dynamics of infections of the number of infected, giving opportunity to Mexico's Hospital System in order to attend cases of severe symptoms, also to implement and to improve health infrastructure or the restructuring of different hospitals to intensive-care nursing. This could reduce mortality rate and recover the more severe cases. In this case, actions leading to the isolation of the population, thus potential transmission of virus from 1.4% to 5.5%, and it should be reduced between 0.9% and 0.5%(Deutschen Gesellschaft für Epidemiologie; Prof. Dr. PH Eva Grill & Prof. Dr. med. Dietrich Rothenbacher, MPH; 23:03:2020).

Emphasizing on strategic objectives that has the WHO in relation to Covid – 19 are the following:

- Interrupting transmission from human being to human being, including the reduction of secondary infections among close contacts.

- Identifying, isolating and treating patients early, including providing optimized care for infected patients.

- Identifying and reducing transmission from the animal source.

- Addressing crucial uncertainties with regard to clinical severity, the degree of transmission and infection, the treatment.

- Communicating critical reporting about risks and events to all communities and countering misinformation.

- Minimizing the social and economic impact through multi-sector partnerships (World; Health Organization; Situation Report 1-92; April 2020).

As a result, this can be achieved through the combination of public health campaigns; for instance, quick identification, diagnosis and case management, identification and monitoring contacts, prevention and infection control in health-care settings, implementing health measures to travelers, raising awareness among the population and communication of risks.

1.4 State Hypotheses and Their Correspondence to Research Design

In light of the foregoing, the main **objective** of this present research work is to predict the degree of risk involved of developing Covi-19, taking into account the Triage respiratory instrument to primary healthcare and which is administered in the Secretary of Health to patients that have access to public hospitals and whose patients have symptoms of coronavirus, for which a multilogistical model is constructed.

The construction of this model is sustained in Operations Research Methodology, which has to fulfill five stages (Thrierauf, 1982):

- First stage (Delimitation of the phenomenon): population, sample and instrument are defined, with the prospect of gathering the necessary information to construct the models.

- Second stage (Formulation and selection of the model):in this section, variables are proposed mathematically to take part in the construction of the models, taking as a reference certain conditions, the selection of the model is done to have a better adjustment.

- Third stage (Estimation of the model): once the model has been identified, an estimation of parameter is done, which will be sustain in the maximum likelihood method.

- Third stage (Validation of the model): once the model has been selected, validation is done through the fulfillment of certain conditions of statistical inference.

- Fourth stage (Analysis of the results): in this last section, parameter interpretation is done.

With the construction of this model, the necessary elements are implemented to foresee the behavior of the phenomenon as of lineal relation that exists between the dependent variable(Y) nominal and the independent variables (X_i , i =1, 2, 3… k).

2. Method

On the basis of the sample size of 7,559 cases and the instrument used by the Secretary of Public Health, the distribution of cases is the following (Figure 1 & 2):
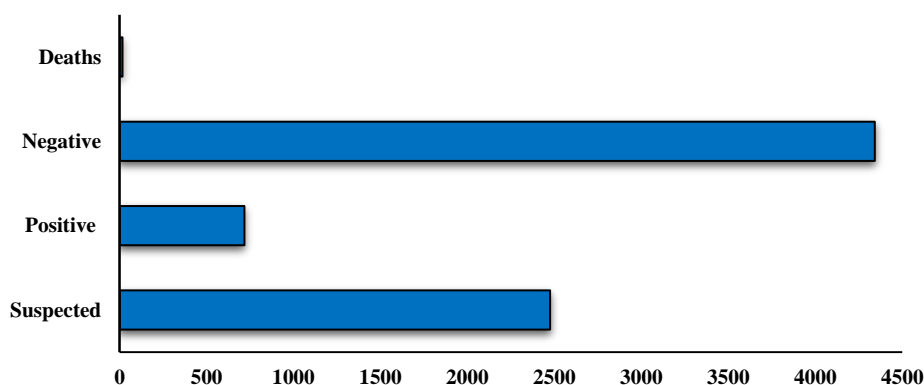

Figure 1. Sampling distribution

Source: Secretary of Health

Respitarory Triage in Primary Care

Date of manufacture

/   /

Objective

To identify patients with signal and symptoms of a respiratory disease with suspicion of COVID - 19 to first contact care in the First and second level, which is intended to reduce the risk of contagion and to reduce waiting times.

General data

Full name

Date

/   /

Male [   ]   Age [   ]

Famale [   ]

Triage

Shortness of breath	No			Yes	
Chest pain	No			Yes	
Inmediate care in the emergency department					

Fever	No		Slight		High	
Sore throat	No		Slight		Serious	
Cough	Sught		Slight		Serious	
Attention area COVID - 19			Si		No	

Others

[ ]	Sore throat and throat pain	[ ]	Fatigue and weakness
[ ]	Conjunctivitis	[ ]	Chills
[ ]	Nasal congestion	[ ]	Sweating
[ ]	Muscular pain	[ ]	Diarrhea, nausea or vomit
[ ]	Joint pain		

Since when have you felt sick? _____

Do you have a desease?

[ ]	Cancer	[ ]	Immunology
[ ]	Cardiovascular	[ ]	Neurological
[ ]	Type 1 and type 2diabetes	[ ]	Obesity
[ ]	Gestational diabetes	[ ]	Pulmonary
[ ]	Pregnancy	[ ]	Immunosuppressive therapy
[ ]	Hematological	[ ]	HIV
[ ]	Hepatic	Other (s) _____

Figure 2. Instruments for its application

Source: Secretariat of Health

2.1 Identify Subsections

Through a descriptive analysis of the variables.In the figure 3 & 4, it can be observed that more than 3,500 interviewees are women, and aproximately 4200 are men.
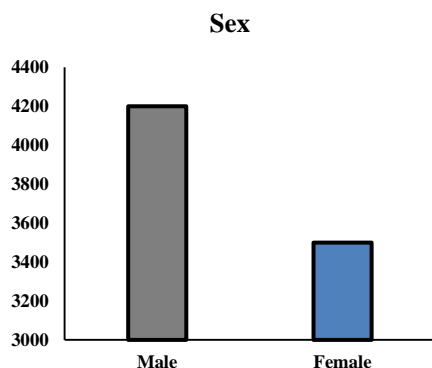
Sex

Figure 3. Variable sex
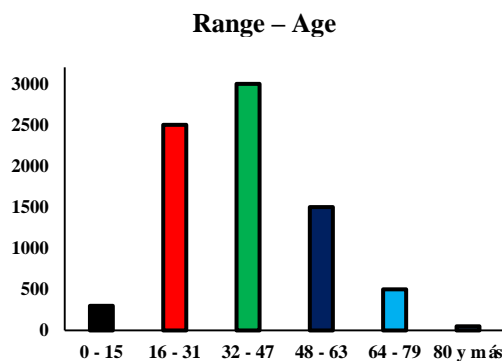Source. Own authorship

Range – Age

Figure 4. Variable age
Source. Own authorship

It can also have been observed that the higher range of the cases is found between 16 to 47 years old, which indicates that the average age of the people is of 40 years old. In the graphs (figure 5), it can be seen that 6% which is equivalent to 435 cases,0.3% which is equivalent to 23 cases, who were positive of cancer and HIV.From the total of the 7,828 interviewees, 12% which is equivalent to 936 cases develop diabetes, beside this, 12% of them have arterial hypertension.

Cancer

HIV

Figure 5. Disease

Source. Own authorship

Diabetes
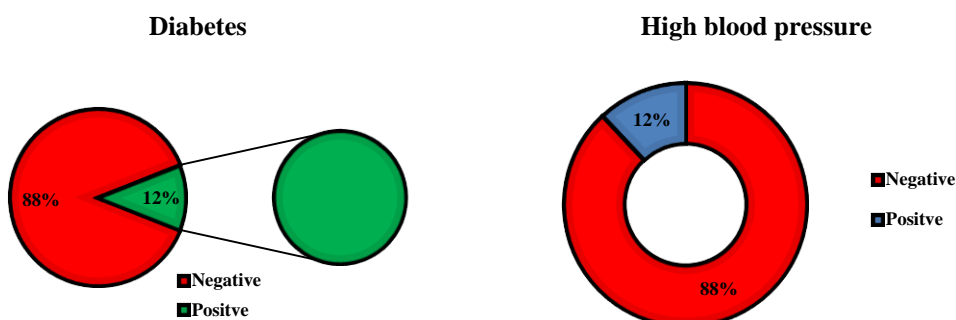

High blood pressure

Figure 6. Disease

Source. Own authorship

In the figure 6, it can be seen that 40% of the interviewees have overweight, 35% have obesity and 25% have normal weight.

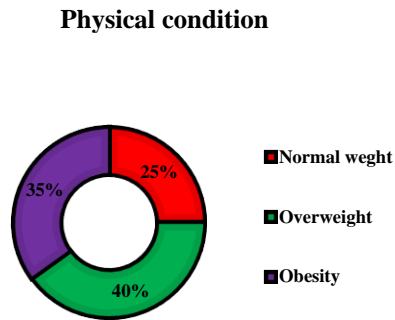**Physical condition**

Transmission

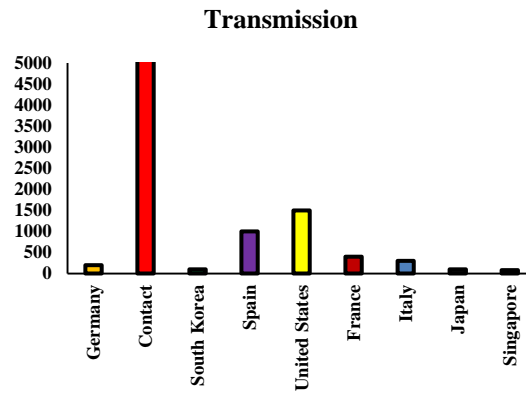Figure 6. Physical condition
Source. Own authorship

Figure 7. Contact point
Source. Own authorship

On one side, based on the figure 7(contact variable), it can be observed that the great majority of infection is found through the contact. In addition to this, other way was the visits done to Spain and United States.

On the other side, 92.3% and 75% of the interviewees have consumed tobacco and alcohol. In the figure 10, it can be observed that the greater transmission is found with physical contact by those who traveled to United States and Spain.
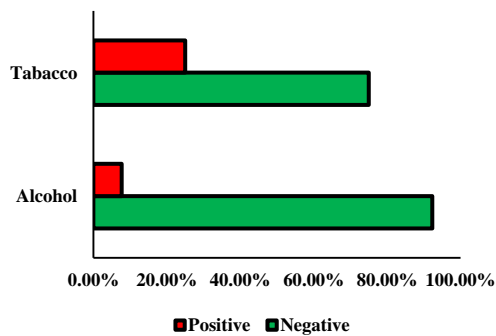

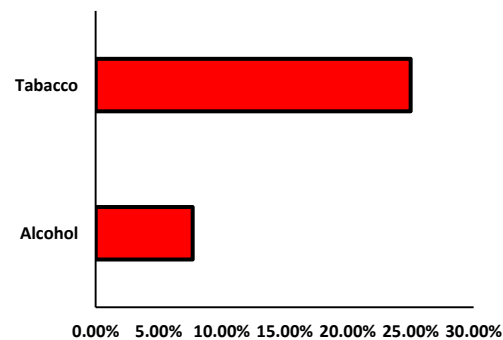Figure 8. Variable of tobacco consumption
Source. Own authorship

Figure 9. Variable of alcohol consumption
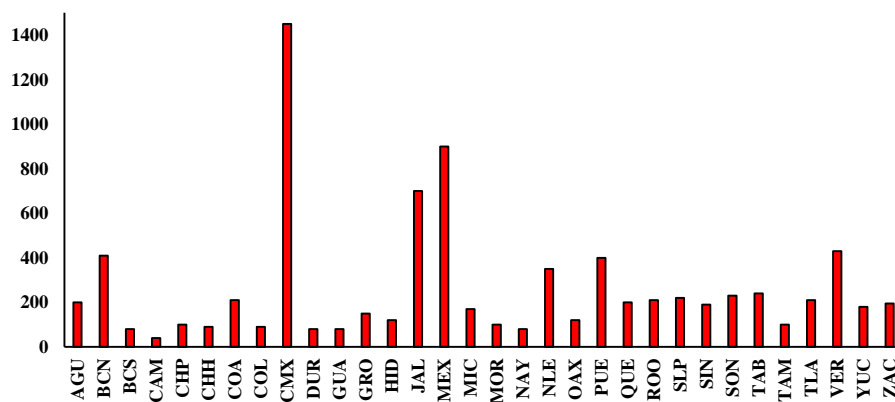Source. Own authorship

Figure 10. Variable of States

Source. Own authorship

Based on this, the following aspect is focusing on the estimate of the probability of the dynamic of the events such as positive, suspected and negative deaths. All of this with a level of confidence of 95%, with a level of significance of 5%.

3. Formulation and Selection of the Model

Taking into consideration the instrument used by the Secretary of Health (Figure 1, the dynamic of Covid-19 is expressed to the following way:

$$Y = f(X_S, X_E, X_F, X_T, X_A, X_{CP}, X_{ED}) \tag{1}$$

Where:

- Y is the expected value of the probability of occurrence of any event.

  - 1 = Deaths

  - 2 = Positive

  - 3 = Suspected

  - 4 = Negative

- Xs is the sex of the patients (M= male and F= female).

- X_E is the possible disease that patients develop.

  - 0 = None   - 1 = Diabetes   - 2 = Hypertension

  - 3 = Cancer   - 4 = HIV

- X_F is the physical condition of the patients.

  - 1 = Normal weight   - 2 = Overweight   - 3 = Obesity

- X_T is the tobacco consumption of patients (1 = smoker & 0 = non-smoker).

- X_A is the alcohol consumption of patients (1= use and 0= no used).

- X_{CP} is the contact point of the patients.

  - 0 = None   - 5 = Germany

  - 1 = South Korea   - 6 = France

  - 2 = China   - 7 = Spain

  - 3 = Japan   - 8 = United States

  - 4 = Singapore   - 9 = Contacts

- X_{ED} is the age of the patients.

Taking as a reference the equation (1), the probability that a Covid-19 is provided considering the Triage instrument that the Secretary of Health calculated, it would be obtained in the following way:

$$P(E) = \hat{\beta}_0 + \hat{\beta}_1 X_S + \hat{\beta}_2 X_E + \hat{\beta}_3 X_F + \hat{\beta}_4 X_T + \hat{\beta}_5 X_A + \hat{\beta}_6 X_{CP} + \hat{\beta}_7 X_{ED} + ei \tag{2}$$

Such that:

- P(E) is the probability that any event happens.

- $\hat{\beta}_i$, i = 0, 1, 2, 3, 4, 5, 6 & 7 are the parameters to estimate for maximum likelihood.

- ei is the margin of error that cannot be explained for model.

In starting from the fact that the probability of Y has more than two possible events with *n* independent trials that allow *k* independent results $E_1, E_2,..., E_K$ whose probabilities are the following (Montoya & Correa, 2017):

$$P(E_1) + P(E_2) + \dots + P(E_k) = \sum_{i=1}^{k} P(E_i) = 1 \tag{3}$$

Where $n_1, n_2, \dots, n_k$ is the number of occurrences of events $E_1, E_{2\dots} E_K$

respectively in *n* trials:

$$n_1 + n_2 + \dots + n_k = \sum_{i=1}^{k} n_i = n \tag{4}$$

The function of probability of $n_1, n_2 \dots n_k$ events is given by:

$$f(n_1, n_{2\dots} n_k) = \frac{n!}{n_1! n_2! \dots n_k!} P(E_1)^{n_1} P(E_2)^{n_2} \dots P(E_k)^{n_k} \tag{5}$$

To $n_i = 0, 1, 2 \dots n$ is subject to $\sum_{i=1}^{k} n_i = n$, where the mean and variance of multinomial distribution is given by:

$$E[n_i] = nP(E_i) \quad (4); \qquad Var[n_i] = nP(E_i)\{1 - P(E_i)\} \tag{6}$$

If Y has more than two possible events $E_1, E_2 \dots E_K$ respectively in n trials, its most appropriate modeling is the multilogistical regression analysis, whereby it is supposed that (Bocco & Herrero, 2009):

$$\ln\left[\frac{Y}{1-Y}\right] = \beta_0 + \sum_{i=1}^{n} \beta_i X_i \; ; \; i = 1, 2, 3, 4, \dots, k \tag{7}$$

Where:

X_i are the dependent variables of the model.

β_i are the parameters to estimate.

The objective of this type of modeling is to estimate the probability that any of the events of an independent variable happens, such that:

$$P[Y = i, X] = P(E_i) = \frac{e^{X_i \beta_i}}{1 + \sum_{j=1}^{g-1} e^{X_i \beta_i}} \; i = 1, \dots, g - 1 \tag{8}$$

Where the parameter estimation is obtained through maximum likelihood, which is given by the following algebraic expression Pando, M. (2004):

$$L = \prod_{i=1}^{n} (P(E_{1i})^{Y_{1i}} * P(E_{2i})^{Y_{2i}} * \dots * P(E_{ni})^{1 - (\Sigma_{i=1}^{n} Y_{ij})}) \tag{9}$$

Therefore, the viability and the adjustment of the model is going to depend on three elements (McCullagh & Nelder, 1983):

- AIC (Akaike's information criterion) that evaluate not only the adjustment of the model of the data but also the complexity of the model:

$$AIC = 2p - 2Lm \tag{10}$$

Where:

2p is the number of estimated parameter

Lm is the plausibility of the actual model

AIC is used to compare models. When AIC is smaller, the adjustment is better.

- The final historic value of its interactions must be in the middle of the residuals of deviance.
- The degree of incidence of the independent variables (X_i, i =1, 2, 3,..., k) over Y, to fulfill this P – Value must be higher level to the level of significance.
- The level of adjustment that the mode has, that means that the level of variability that keeps the model from the model, it is done through deviance $(0 \leq D^2 \leq 1)$.

The multilogistical model supposes that the data of the phenomenon are specific of the case, in which each independent variable has a unique value to each case. The null hypothesis of this type of models is that there is no relation between the independent and dependent variables, in other words, the values of the dependent variable are predicted from a multilogistical equation; even though, there are not close to the real values of the dependent variable (Bocco & Herrero, 2009).

3. Estimation and Validation of the Model

Based on the following algebraic expression (2), the obtained results are the following:

Chart 1. Historical Interactions (model - 1)

Weights: 52 (21 variable)

initial value	10851.91	**inter 10 value**	7674.72
inter 20 value	7211.40	**inter 30 value**	7135.40
inter 40 value	7134.62		
converged			

Source. Own authorship.

Chart 2. Level of significance of the variables (model - 1)

	(Intercept)	X_S	X_E	X_F	X_T	X_A	X_{CP}	X_{ED}
Negative	0.019	0.951	0.000	0.001	0.427	0.972	0.051	0.010
Positive	0.016	0.542	0.000	0.006	0.450	0.994	0.008	0.008
Suspected	0.071	0.478	0.000	0.003	0.411	0.985	0.005	0.000

Source. Own authorship

In the chart 1, it can be observed that the value of the interactions is the double of the final historical interactions; as a result, the model has a good adjustment. Nonetheless, in the chart 2 the variables X_S(sex), X_T(tobacco) and X_A(alcohol) are not meaningful, with a level of significance of 0.05. Therefore, to predict the probability of events of Covid-19, these variables need to be removed from the model.

Chart 3. Historical Interactions (model - 2)

Weights: 28 (15 variable)

initial value	10851.91	**inter 10 value**	7673.89
inter 20 value	7175.66	**inter 30 value**	7135.31
inter 40 value	7135.07		
converged			

Source. Own authorship

Chart 4. Level of significance of the variables (model - 2)

	(Intercept)	X_E	X_F	X_{CP}	X_{ED}
Negative	0.008	0.000	0.071	0.024	0.012
Positive	0.662	0.000	0.085	0.003	0.009
Suspected	0.124	0.000	0.036	0.002	0.000

Source. Own authorship

In the chart 3, it can be observed that the initial value of the interactions of the second model is the double product of the final historic interactions; as a consequence, the model has a good adjustment. Furthermore, in the chart 4 the variables are meaningful to predict the probability of the events of Covid-19. In the light of the foregoing, the calculation of the parameters are the following:

Chart 5. Calculation of parameters(model-2)

	(Intercept)	X_E	X_F	X_{CP}	X_{ED}
Negative	0.131	1.559	2.251	0.897	-0.036
Positive	-2.837	1.713	2.378	1.271	-0.037

Source. Own authorship

The lineal equations of each event are the following (Bacchini, Vázquez, Bianco & Fronti, 2018):

$$\ln\left[\frac{P(Negative)}{P(Death)}\right] = 0.131 + 1.559XE + 2.251XF + 0.897XCP - 0.036XED \tag{11}$$

$$\ln\left[\frac{P(Positive)}{P(Death)}\right] = -2.837 + 1.713XE + 2.378XF + 1.271XCP - 0.036XED \tag{12}$$

$$\ln\left[\frac{P(Suspected)}{P(Death)}\right] = -1.711 + 1.789XE + 2.593XF + 1.336XCP - 0.055XED \tag{13}$$

Substituting the events (Stewart, Redlin & Watson, 2012):

$$Negative = 1.14e^{0.559XE+2.251XF+0.897XCP-0.036XED} \tag{14}$$

$$Positive = 0.06e^{1.713XE+2.378XF+1.271XCP-0.0136XED} \tag{15}$$

$$Suspected = 0.18e^{1.789XE+2.593XF+1.336XCP-0.055XED} \tag{16}$$

As a consequence, the level of confidence is of 0.95 and with a level of significance of 0.05. In the present model keeps 58.75% of the variability of the data, in other words, the present equations explain 58.75% the probability of any of the events happen.

4. Analysis of Results

Taking into account the equations 15,16,17 & 18, the probability that any of the events happen is obtained in the following way: If the average age (X_A) taken from the sample population is of 40 years, the rest of the variables stay constant (X_P, X_{CP} y X_{ED}), the probability of occurrence of each event.

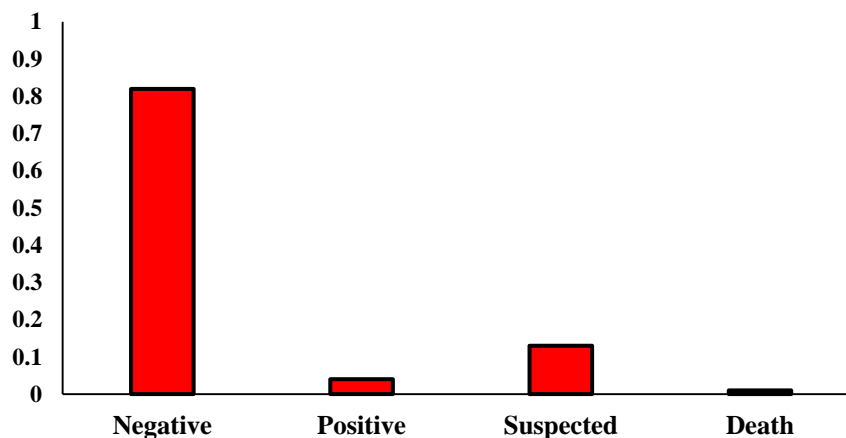Applying an algebraic expression 8, it is obtained that:

Figure 11. Probability of events

Source. Own authorship

Therefore, the higher probability that a person is not infected of Covid-19, is of 0.82, and in case of being positive is of 0.04, provided that the variable stays constant.

In starting from the fact that it is required to predict the probability to develop Covid-19, from a condition (X_D), from a physical condition (X_F) , contact point (X_{CT}) and the age (X_{AG}).The risk will be increased in the following way:

Chart 6. Variables of greater weight

Variables	Increase	P(Positive)
Diseases (XE)	75.00%	0.070
Physical condition (XF)	10.00%	0.044
Contact ponit (XCP)	15.00%	0.046
Age (ED)	12.50%	0.045

Source. Own authorship

In the chart 6, it can be seen that the variable with greater incidence for population to develop Covid-19 are the diseases, and this degree of probability is going to increase inasmuch as the type of condition is identified.

Let us suppose that a person has had all symptoms that encompass Covid-19. She/he is 56 years old, and this person is diabetic, he/she has normal weight, and traveled to United States. What is his/her health condition with regard to Covid-19?

$$\text{Negative} = 1.14e^{0.559(2)+2.251(1)+0.897(8)-0.036(56)} = 5767.70 \tag{17}$$

$$\text{Positive} = 0.06e^{1.713(2)+2.378(1)+1.271(8)-0.0136(56)} = 242071.71 \tag{18}$$

$$\text{Suspected} = 0.18e^{1.789(2)+2.593(1)+1.336(8)-0.055(56)} = 173546.62 \tag{19}$$

Where:

$$P(\text{Event}) = \frac{\text{Event in study}}{1 + \sum_{i=1}^{3} \text{Events}} \tag{20}$$

Chart 7. Calculation of probability

Events	Probability
Negative	0.013
Positive	0.574
Suspected	0.411
Death	0.002

$$P(\text{Negative}) = \frac{5767.70}{1 + 421386.03} = 0.013$$

$$P(\text{Positive}) = \frac{242071.71}{1 + 421386.03} = 0.574$$

$$P(\text{Suspected}) = \frac{173546.62}{1 + 421386.03} = 0.411$$

Source. Own authorship

As it can be observed in chart 7, with the conditions and people's physical data, health and age, it can be classified in suspected with greater tendency of being positive of Covid-19; as a consequence, this person must receive medical treatment.

4. Discussion

It is important to mention that with the development of the probabilism modelling, the risk of developing Covid-19 can be predicted. On the basis of the obtained results, the probability of getting the virus is of 0.04. Nevertheless, this risk has been increased as long as there are different factors: Firstly, the health condition of the patient, it is the greater proportion, thus as long as he/she has a condition; for instance, hypertension, diabetes, cancer or HIV, the risk would be greater. Secondly, it makes reference to the place where the person contracted the virus, if this person traveled abroad, and if he/ she was with an infected person, the risk is higher. Thirdly, the age of the patient, because "a greater age-a greater risk". To conclude, the physical condition because "a greater body mass with regard to height, a greater risk". As it can be seen, the mathematical and statistical models are implemented to describe, to explain and to predict phenomena and processes that encompass specific public health areas; for example, medicine, biology, epidemiology, among others.

References

Bacchini, R., Vázquez, L., Bianco, M., & Fronti, J. (2018). *Introduction to Probability and Statistics*. Argentine. IADCOM - UBA.

Bocco, M., & Herrero, V. (2009). *Multi-logistic model to identify the determinants of joint labor participation modalities in Argentine*. Association of Population Studies of Argentina.

Liu, C., Zhou, Q., Li, Y., Garner, L. V., Watkins, S. P., Carter, L. J., ... & Albaiu, D. (2020). Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. https://doi.org/10.1021/acscentsci.0c00272

Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., ... & Zhao, Y. (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. *Jama*, *323*(11), 1061-1069. https://doi.org/10.1001/jama.2020.1585

Chan, J. F. W., Yuan, S., Kok, K. H., To, K. K. W., Chu, H., Yang, J., ... & Tsoi, H. W. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet*, *395*(10223), 514-523. https://doi.org/10.1016/S0140-6736(20)30154-9

McCullagh, P., & Nelder, J. (1983). *Generalized Linear Models*. United States. Chapman and Hall (Eds.). https://doi.org/10.1007/978-1-4899-3244-0

Sridhar, D., & Majumder, M. S. (2020). Modelling the pandemic BMJ; 369: m1567, 2020 04 21. MEDLINE | ID: covidwho-100009. https://doi.org/10.1007/978-1-4899-3244-0

Montoya, Y., & Correa, J. (2016). *Choice of multinomial distribution from various experts*. Maste Thesis. Colombian National University. http://bdigital.unal.edu.co/53378/

Peeri, N. C., Shrestha, N., Rahman, M. S., Zaki, R., Tan, Z., Bibi, S., ... & Haque, U. (2020). The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned?. *International journal of epidemiology*. https://doi.org/10.1093/ije/dyaa033

Pando, V., & Martń, R., & San Martń Fernández. (2004). Multinomial Logistic Regression. Specialized Social Notebooks in Forest Sciences: 18: 000-000 (2004). Minutes of the Forest Modeling Meeting. University of Valladolid, Spain. pp 323-327.

Süddeutsche Zeitung. (2020). Uhr Kampf gegen Pandemie. Innenministerium dringt auf massive Ausweitung von Corona-Tests. Retrieved from: https://www.sueddeutsche.de/politik/coronavirus-tests-strategie-1.4858950.

Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., ... & Wang, M. H. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International journal of infectious diseases*, *92*, 214-217. https://doi.org/10.1016/j.ijid.2020.01.050

Stewart, J., Redlin, L., & Watson, S. (2012). *Pre-calculus. Mathematics for Calculation.* United States. Cengage – Learning (Eds).

World Health Organization. Situation Report 1-92; April 2020. Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports