Gaussian Model to Predict the Risk of Developing Type 2 Diabetes Mellitus in Mexican Population Taking as a Reference Risk Factors

Guerrero-Escamilla Juan Bacilio[1], Lopez-Perez Socrates[1], Jiménez-Proa Erik Gerardo[2]

[1] Academic Area of Sociology and Demography, Institute of Social Sciences and Humanities, Autonomous University of the State of Hidalgo, Pachuca de Soto, Mexico

[2] Academic Area of Gerontology, Institute of Health Sciences, Autonomous University of the State of Hidalgo, Pachuca de Soto, Mexico

Correspondence: Guerrero-Escamilla Juan Bacilio, Academic Area of Sociology and Demography, Institute of Social Sciences and Humanities, Autonomous University of the State of Hidalgo, Pachuca de Soto, Mexico

Abstract

In this research work, an ordinal Gaussian model is constructed, whose objective is to predict the degree of risk of contracting type 2 diabetes mellitus (2DM), taking as reference the risk factors in the Mexican population. It is estimated that the Mexican population has a hereditary susceptibility to develop 2DM, however, the probability increases depending on risk factors; area of residence, background of parents with 2DM, tobacco consumption, alcohol consumption, physical inactivity, body mass index (BMI), and type of feeding, which, despite positively intervening in the appearance of 2DM, they can be modified to obtain the inversely proportional effect. However, the social, economic and political context are important components for the population. Risk factors, as explanatory elements of the prevalence of 2DM, are of the utmost importance to delay or control their early development, as some are factors that can be muffled. For the development of this model, the information published in the National Health and Nutrition Survey (ENSANUT) of 2012 was taken, based on the adult population 20 years of age or older. Among the most outstanding results is the higher prevalence of risk that women have with respect to men, and the fact that age is a fundamental basis for contracting type 2 diabetes mellitus.

Keywords: type 2 diabetes, risk factors, Gaussian model

1. Introduction

Diabetes is a chronic disease that occurs either when the pancreas does not produce enough insulin (It is a hormone responsible for regulating blood sugar) or when the organism does not use effectively the insulin that produces (American Diabetes Association, 2017a).

There are three types of diabetes:

- Type 1 Diabetes (1DM) is characterized by the production of deficient insulin in the organism.
- Type 2 Diabetes (2DM) is when the organism does not use insulin efficiently.
- Gestational diabetes (3DM) is a transient disorder that is produced during the pregnancy and carries a risk of developing diabetes in a specific moment (American Diabetes Association, 2017a).

2DM will be determined due to the interaction of non-modifiable factors (genetic) and environmental (modifiable). The risk increases because of ethnic factors and family history with 2DM. Furthermore, the combination of old age, overweight, obesity, poor nutrition, lack of physical activity, tobacco and alcohol consumption are risk factors which are associated with the development of 2DM (World Diabetes Report, 2016).

2DM is one of the major emergencies of global health in the 21st century because each year strengthens the number of people who are living this condition, this disease can lead to a variety of serious complications over the life (International Diabetes Federation, 2015). Across the world 415 million people with 2DM were registered. It is estimated that there will be 642 million people with DM in 2040 (International Diabetes Federation, 2015). Its complications are the principal causes of death in most countries. 2DM is the most common type and it has increased along with the cultural and social changes. In countries with a high income, there is 91% of the adults who have 2DM. IDF calculates that 193 million people with 2DM are not diagnosed; as a result, they have a higher risk of developing

complications (International Diabetes Federation, 2015).

2DM is viewed as one of the biggest challenges for the public health, not only in developed countries but also in lower and middle-income countries (Barquera, 2003). Mexican population has been influenced through diverse factors over time, transitions lived by inhabitants in our country, have conditioned morbidity and mortality.

The main causes of mortality in Mexico are nontransmissible chronic diseases (NCD) for instance, 2DM and ischemic heart disease. It is important to mention that (NCD) are conditions that mainly affect to the adult population and are characterized by being incurable. These diseases accompany the person for his/her life, becoming a permanent sick person. On the whole, these conditions are asocciated with a level of degeneration of tissues and organs which are usually progressive, until terminal failure of an organ involved evolves (Trindade, Dos Santos, Dalva de Barros, & Silvia, 2014).

In 2000, one of the main causes of mortality in Mexican population was diabetes and was placed in the third place. In 2010 it went up to the second place, and in 2015, it maintained as the second cause of death in Mexico, having 98,521 deaths (Main causes of mortality-INEGI 2000, 2010 & 2015, 2015). DM has become a problem of the public health in Mexico, thus prevalence of this disease has boosted alarmingly in the last decades (Medina, Tolentino-Mayo, López-Ridaura, & Barquera, 2017).

Risk factors considered for 2DM explain that the globalization of the economies, the modifications in the production systems and the uncontrolled urbanization have disturbed the consumer system and grown the exposure to risk factors to develop obesity and NCD such as harmful diets, physical inactivity, and consumption of tobacco and alcohol (Schmidhuber & Prakash, 2007).

2. Problem Statement

The objective of this research is to estimate the prevalence of type 2 diabetes mellitus, taking as a reference the risk factors that affect the health of Mexican population through the construction of probability model, taking into account the National Health and Nutrition Examination Survey (NHNES) 2012.

According to the objective of the phenomenon of 2DM, it will be in function of the related processes with the rapid urbanization, the economic growth and the probable genetic tendencies. The evidence shows a higher influence of elements such as changes in the consumption pattern and reduction of physical activity (Hernández-Ávila, Gutiérrez, & Reynoso-Noverón, 2013). Diabetes is a disease that is originated because of the combination of different factors, for instance, the age, the obesity, the sedentary lifestyle, the inadequate nutrition and the family history (González-Villalpando, Dávila-Cervantes, Zamora-Macorra, Trejo-Valdivia, & González-Villalpando, 2014a). In Mexico, obesity and physical inactivity constitute risk modifiable factors for 2DM (Paternina-de la Ossa, Villaquirán-Hurtado, Jácome-Velasco, Galvis-Fernández, & Granados-Vidal, 2018). For example, tobacco use, unhealthy diets, physical inactivity, harmful use of alcohol are consequences of the rapid urbanization and the lifestyles in the twenty first century and they have been showed in NCD (World Health Organization, 2018). The function of the type 2 diabetes mellitus:

$$P(Y)=F(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9) \tag{1}$$

Where:

- F is the function of equation
- $P(Y)$ is the likelihood of having 2DM (1= if diagnosed; 0= not diagnosed)
- X_1 is the sex of those diagnosed by 2DM (1= woman; 2= men)
- X_2 is the age of those diagnosed by 2DM (> 20 years)
- X_3 is the residential zone of those diagnosed by 2DM (1= rural area; 2= urban area; 3= metropolitan area)
- X_4 is the parents' background of 2DM (1= father; 2= mother; 3= both parents)
- X_5 current tobacco use of those diagnosed by 2DM (1= no tobacco consumed; 2= tobacco consumed)
- X_6 current alcohol use of those diagnosed by 2DM (1= no alcohol consumed; 2= alcohol consumed)
- X_7 is the type of physical activity of those diagnosed by 2DM (1= FA, active; 2= FA, moderate; 3= FA, non-active)
- X_8 is the category of body mass index(BMI) of those diagnosed 2DM (1= low weight; 2= normal weight; 3= overweight; 4= obesity)
- X_9 is the type of food of those diagnosed by 2DM (1= healthy eating; 2= unhealthy eating)

This information was gathered from the National Health and Nutrition Examination Survey (NHNES, 2012).

3. Methodology

Based on the objective, the construction and the development of the Gaussian Model is supported on the operations research methodology which is composed by five stages such as problem statement, construction of the model, model solution, model validation and implementation (Taha, 2012).

1. Problem statement: it was outlined the relation among the 2DM (dependent variable) in Mexico, based on independent variables of sex, age, residence zone and parents' background with 2DM, tobacco consumption, alcohol consumption, physical activity and food.

2. Construction of the model: the mathematical statement was made based on the relationships between the dependent variable and the independent variables.

3. Model solution: seeing that the mathematician statement, it was used different operations which were performed towards the construction of equation that indicated the level of risk to develop 2DM, taking into account the independent variables.

4. Model validation: in this stage, the stated assumptions of statistical inference were fulfilled, which were useful to validate the model.

5. Implementation: starting from the stated assumptions of statistical inference, it was undertaken with parameter interpretation; as well as, the possible scenarios of the phenomenon 2DM, taking as a reference its current dynamics.

4. Model Variability

The current model is ordinal in natural; therefore, it must fulfill with two assumptions. First of all, it should not be co-linearity among the independent variables and second of all, the degree of adjustment. The co-linearity is a problem of the regression analysis that focuses on the model predictors are related to linear combination (Peña, 1987).

The regression model has faced important consequences because if the predictors are found in the linear combination, the influence of each one of them can not been distinguished in the criteria due to they are sneaky one another. In addition, it does not show an explanation about the phenomenon to study, also the forecasts are not reliable since there is another combination of predictors which were introduced in the model and would vary the order and produce predictions in the contradictory criterion. (Peña, 1987).

Approaches based on the correlation of explicative variables, there is a method that consists of calculating the called "factors of inflation of variance" or VIF's defined as:

$$VIF_j = \frac{1}{1-R_j^2} \qquad j=1,2,...9 \qquad (2)$$

Where:

R_J^2 is the coefficient determination of the regression of the umpteenth regressor over the rest.

The criterion of evaluation of co-linearity, hypothesis test is the following:

$$Ho: co - linearity \; vs. \; Ha: there \; is \; no \; co - linearity$$

Having a level of confidence of 0.95, with a level of significance of 0.05 and fulfilling with the assumption of there is not co-linearity among the independent variables ($X1_{woman}$, X2, X3, X4, X6, X7 and X9), the equation 26 explains 19.37% of risk dynamic of developing 2DM. *X5 and X8 are not significant according to the model*.

If VIF is higher to 10 units of each independent variable, it will be co-linearity, which means that Ho is approved.

Table 1. Hypothesis VIF (model 8)

X_1	X_2	X_3	X_4	X_6	X_7	X_9
1.056549*	1.086794	1.041866	1.011181	1.090043	1.014166	1.001383

Source: Own authorship,* units.

According to the criteria taken from table 1, it can be seen that the VIF of each of the independent variables can explain the 2DM variable, since they are below the 10 units. Consequently, Ha is approved, in other words, there is no co-linearity between independent variables.

The deviance of a GLM is defined as the degree of variability of the data that must be compared. The deviance of the invalid model with the residual deviance. This is viewed as a measure as soon as the model keeps the variability of the

data, which would be:

- Deviance of invalid model is of 457.14 units.
- Deviance of the residual is of 368. 61 units.

$$D^2 = \left[\frac{\text{Devianza del modelo nulo} - \text{Devianza de los residuales}}{\text{Devianza del modelo nulo}}\right] * 100 \qquad (3)$$

Substituting in the equation (3), it is obtained that:

$$D^2 = \left[\frac{457.14 - 368.61}{457.14}\right] * 100 = 19.37 \qquad (4)$$

In light of the foregoing, the equation that foresees that 2DM keeps 19.37% of the variability of the data.

$$P\left[E(\widehat{Y})\right] = -0.429 + 0.069 X_1 \text{woman} + 0.009 X_2 + 0.015 - 0.042 X_4 + 0.079 X_6 + 0.049 X_7 + 0.079 X_9 \quad (5)$$

Having a level of confidence of 0.95, with a level of significance of 0.05 and fulfilling with the assumption of there is not co-linearity among the independent variables ($X1_{\text{woman}}$, X2, X3, X4, X6, X7 and X9), the equation explains 19.37% of risk dynamic of developing 2DM.

5. Results

The analysis of the result of the current model is based on the interpretation of the parameters of the following algebraic expression, si $X_1 = \text{woman}$, X_2, X_3, X_4, X_6, X_7 y X_9 remain constant:

$$P\left[E(\widehat{Y})\right] = -0.429 + 0.069 X_1 \text{woman} + 0.009 X_2 + 0.015 X_3 - 0.042 X_4 + 0.079 X_6 + 0.049 X_7 + 0.079 X_9 \quad (6)$$

Replacing:

$$P\left[E(\widehat{Y})\right] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(0) \quad (7)$$

That is to say:

$$P\left[E(\widehat{Y})\right] = -0.429 \qquad (8)$$

In terms of normal distribution:

$$Z_g = \text{normal}\left\{P\left[E(\widehat{Y})\right]\right\} = \text{normal}\{-0.429\} = 0.3339 \qquad (9)$$

The probability to develop 2DM is of 0.3339, that means that if there is no intervention of the explicative variables, any interviewee has a level of risk to develop 2DM of 33.39% , with a level of cofidence of 0.95.

0.3339, is the probability of developing 2DM, where independent variables do not intervene, this probability is interpreted as a genetic parameter, then this factor is read.

The literature explains that there is a genetic factor in Mexicans, as there are variants of the indigenous population that give a greater predisposition to suffer a metabolic disease such as diabetes (Bonilla, 2017). The genetic variant called haplotype is composed of five changes in a gene called SLS16A11. This haplotype was found associated with a predisposition of diabetes in the Mexican population. The haplotype explains 20% of the prevalence of 2DM in the Mexican population (Bonilla, 2017).

If $X_1 \text{Woman} = 1$ and X_2, X_3, X_4, X_6, X_7 and X_9 remain constant, the probability of developing DM2 is 0.3594, that is, if the interviewee is a woman, she has a risk of developing 2DM of 35.94%, with a confidence level of 0.95.

$$P\left[E(\widehat{Y})\right] = -0.429 + 0.069(1) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(0) \quad (10)$$

That is to say:

$$P\left[E(\widehat{Y})\right] = -0.360 \qquad (11)$$

In terms of normal distribution:

$$Z_{\text{woman}} = \text{normal}\left\{P\left[E(\widehat{Y})\right]\right\} = \text{normal}\{-0.360\} = 0.3594 \qquad (12)$$

In starting from the assumption that the probability of risk for the interviewees without focusing on the sex is of 0.3339, and female is of 0.3594, so the risk rate of women with regard to men is of 7.63%.

$$Rr = \frac{Z_{\text{woman}}}{Z_g} = \frac{0.3594}{0.3339} = 1.0763 \qquad (13)$$

As a result, women become susceptible to develop 2DM thus it is stated that: Obesity is the most visible risk factor in women to develop 2DM, in relation to men (Kautzky-Willer, Harreiter, & Pacini, 2016). In women eating disorders are more common than in men, these are mainly characterized by periods in which women eat without control, or the

opposite, they do not eat for fear of gaining weight, and with this the probability of developing 2DM increases (American Diabetes Association, 2017b).

If $X_2 = 1$, and X_1 woman, X_3, X_4, X_6, X_7 and X_9 remain constant, chance of developing 2DM is of 0.3372, in other words, the age is an important factor to develop 2DM. The interviewees will have a a level of risk of 33.72%, with a level of confidence of 0.95.

$$P[E(\widehat{Y})] = -0.429 + 0.069(0) + 0.009(1) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(0) \quad (14)$$

That is to say:

$$P[E(\widehat{Y})] = -0.420 \quad (15)$$

In terms of normal distribution:

$$Z_E = \text{normal}\{P[E(\widehat{Y})]\} = \text{normal}\{-0.420\} = 0.3372 \quad (16)$$

For each elapsing year, the probability of risk of the interviewees to develop 2DM is approximately 0.98% that means of 1%.

$$Rr = \frac{Z_E}{Z_g} = \frac{0.3372}{0.3339} = 1.0098 \quad (16)$$

The age related to 2DM argues that: A higher age, a higher risk to develop 2DM, more heart diseases and strokes (American Diabetes Association, 2017b). 2DM ocurrs in mature age over 45 years old (American Diabetes Association, 2017b). Even though, there are diagnoses of 2DM of an early age associated with obesity and physical inactivity.

If $X_3 = 1, 2,$ or 3. y X_1 woman, X_2, X_4, X_6, X_7 and X_9 remain constant, chance of developing 2DM based on the residence zone, it will be interpreted in the following way:

● If $X_3= 1$(Rural Area) chance of developing 2DM is of 0.3394, in other words, if the interviewee lives in a place of residence of a rural zone, his/her level of risk to develop 2DM will be of 33.94%, with a level of confidence of 0.95.

$$[E(\widehat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(1) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(0) \quad (18)$$

That is to say:

$$[E(\widehat{Y})] = -0.414 \quad (19)$$

In terms of normal distribution:

$$Z_E = \text{normal}\{P[E(\widehat{Y})]\} = \text{normal}\{-0.414\} = 0.3394 \quad (20)$$

● If $X_3 = 2$ (Urban Area) chance of developing 2DM is of 0.3449, that means that if the interviewee lives in a place of residence of a urban zone, his/her level of risk to develop 2DM will be of 34.49%, with a level of confidence of 0.95.

$$P[E(\widehat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(2) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(0) \quad (21)$$

That is to say:

$$P[E(\widehat{Y})] = -0.399 \quad (22)$$

In terms of normal distributuion:

$$Z_E = \text{normal}\{P[E(\widehat{Y})]\} = \text{normal}\{-0.399\} = 0.3449 \quad (23)$$

● If $X_3 = 3$ (Metropolitan Area) chance of developing 2DM is of 0.3504, that means that if the interviewee lives in a place of residence of the metropolitan area, his/her level of risk to develop 2DM will be of 35.04%, with a level of confidence of 0.95.

$$P[E(\widehat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(3) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(0) \quad (24)$$

That is to say:

$$P[E(\widehat{Y})] = -0.384 \quad (25)$$

In terms of normal distribution:

$$Z_E = \text{normal}\{P[E(\widehat{Y})]\} = \text{normal}\{-0.384\} = 0.3504 \quad (26)$$

On the basis of the table 2 and taking as a reference the context of genetic risk, it can be observed that people who live in their place of origin of the metropolitan area, they will have a higher risk of developing 2DM with regard to the other rural and urban areas.

Table 2. Risk rate due to place of residence

Rural Area	$\mathbf{Rr = \dfrac{Z_R}{Z_g} = \dfrac{0.3394}{0.3339} = 1.0164}$	(27)
Urban Area	$Rr = \dfrac{Z_R}{Z_g} = \dfrac{0.3449}{0.3339} = 1.0329$	(28)
Metropolitan Area	$Rr = \dfrac{Z_R}{Z_g} = \dfrac{0.3504}{0.3339} = 1.0491$	(29)

Source: Own authorship.

In the metropolitan area there is a little more risk of developing 2DM, but there is very little discrepancy with the urban area, the difference is only 0.55% (35.04% - 34.49%), it is less than 1%. However, it can be said that in urban areas the domain of refined and indu(27)strialized foods and foods of animal origin with high amounts of fat, salt and sugar are observed. Meanwhile, foods such as corn, beans and chili are still consumed in rural areas [19]. The urbanization effect is an epidemiological transition towards increasing rates of obesity and 2DM [21].

If $X_4 = 1, 2,$ or 3 and X_1 woman, X_2, X_3, X_6, X_7 and X_9 remain constant, chance of developing 2DM on the basis of the residential area and its interpretation would be the folowing:

- If $X_4 = 1$ (Father) chance of developing 2DM is of 0.3482, that means that if the interviewee's father has 2DM, his/her level of risk to develop 2DM will be of 34.82%, with a level of confidence of 0.95.

$$P\big[E(\widehat{Y})\big] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(1) + 0.079(0) + 0.049(0) + 0.079(0) \quad (30)$$

That is to say:

$$P\big[E(\widehat{Y})\big] = -0.471 \quad (31)$$

In terms of normal distribution:

$$Z_E = \text{normal}\big\{P\big[E(\widehat{Y})\big]\big\} = \text{normal}\{-0.471\} = 0.3482 \quad (32)$$

- If $X_4 = 2$ (Mother) chance of developing 2DM is of 0.3631, in other words, if the interviewee's mother has 2DM, his/her level of risk to develop 2DM will be of 36.31%, with a level of confidence of 0.95.

$$P\big[E(\widehat{Y})\big] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(2) + 0.079(0) + 0.049(0) + 0.079(0) \quad (33)$$

That is to say:

$$P\big[E(\widehat{Y})\big] = -0.513 \quad (34)$$

In terms of normal distribution:

$$Z_E = \text{normal}\big\{P\big[E(\widehat{Y})\big]\big\} = \text{normal}\{-0.513\} = 0.3631 \quad (35)$$

- If $X_4 = 3$ (Both parents) chance of developing 2DM is of 0.3782, that means that if the interviewee has both parents with 2DM, his/her level of risk to develop 2DM will be of 37.82%, with a level of confidence of 0.95.

$$P\big[E(\widehat{Y})\big] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(3) + 0.079(0) + 0.049(0) + 0.079(0) \quad (36)$$

That is to say:

$$P\big[E(\widehat{Y})\big] = -0.555 \quad (37)$$

In terms of normal distribution:

$$Z_E = \text{normal}\big\{P\big[E(\widehat{Y})\big]\big\} = \text{normal}\{-0.555\} = 0.3782 \quad (38)$$

On the basis of the table 3 and taking as a reference the context of genetic risk, it can be observed that people who have both parents with 2DM, have a higher risk of developing 2DM with regard to other background of father and mother.

Table 3. Risk rate due to background information on parents with 2DM

Father	$Rr = \dfrac{A_R}{Z_g} = \dfrac{0.3482}{0.3339} = 1.0428$	(39)

Mother
$$Rr = \frac{A_R}{Z_g} = \frac{0.3631}{0.3339} = 1.0874 \tag{40}$$

Both parents
$$Rr = \frac{A_R}{Z_g} = \frac{0.3782}{0.3339} = 1.1326 \tag{41}$$

Source: Own authorship.

In relation to the background information on the parents with 2DM, it is mentioned that: American Diabetes Association (2013) of genetic aspects of diabetes clarifies that diabetes is not a hereditary disease, it is just a pattern which has a difference of some hereditary traits. Nevertheless, there are some people who are more likely to develop diabetes than others. For instance, twin studies have showed results that genetic factors play a relevant role in the emergence of 2DM. That means that genetic factors are not enough because in a test in which twins have identical genes and anyone of them have 2DM, the risk of the other one is 3 of 4 (American Diabetes Association, 2013). Genetic susceptibility or lifestyle factors of someone diagnosed with 2DM, it is difficult to know if his/her situation was due to a family history or a risk factor; it is likely that both. As a result, some studies have demostrated that it is possible to delay or to prevent 2DM when a person does physical exercise and has a healthy body weight (American Diabetes Association, 2013). Developing 2DM might depend on diverse risk factors that include family history of 2DM;although, there is a probability of 11% of having 2DM until 70 years old (American Diabetes Association, 2013). Scientists claimed that the risk of developing 2DM in a child is higher when his/her mother has 2DM (American Diabetes Association, 2013).

If $X_6 = 1$, or 2 and X_1woman, X_2, X_3, X_4, X_7 and X_9 remain constant, chance of developing 2DM with alcohol consumption, it will be interpreted in the following way:

- If $X_6 = 1$ (No alcohol consumed) chance of developing 2DM is of 0.3631, that means that if the interviewee does not consume alcohol, his/her level of risk to develop 2DM will be of 36.31%, with a level of confidence of 0.95.

$$P[E(\hat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(1) + 0.079(0) \tag{42}$$

That is to say:

$$P[E(\hat{Y})] = -0.38$$

In terms of normal distribution:

$$Z_E = normal\{P[E(\hat{Y})]\} = normal\{-0.38\} = 0.3631 \tag{44}$$

- If $X_6 = 2$ (Alcohol consumed) chance of developing 2DM is of 0.3935, in other words, if the interviewee consumes alcohol, his/her level of risk to develop 2DM will be of 39.35%, with a level of confidence of 0.95.

$$P[E(\hat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(2) + 0.079(0) \tag{45}$$

That is to say:

$$P[E(\hat{Y})] = -0.331 \tag{46}$$

In terms of normal distribution :

$$Z_E = normal\{P[E(\hat{Y})]\} = normal\{-0.331\} = 0.3935 \tag{47}$$

On the basis of table 4 and taking as a reference the context of genetic risk, it can be observed that people who consume alcohol, have higher risk to develop 2DM with regard to people who do not consume it.

Table 4. Risk rate due to alcohol consumption

No alcohol consumed
$$Rr = \frac{A_C}{Z_g} = \frac{0.3631}{0.3339} = 1.0874 \tag{48}$$

Alcohol consumed
$$Rr = \frac{A_C}{Z_g} = \frac{0.3935}{0.3339} = 1.1784 \tag{49}$$

Source: Own authorship.

Alcohol consumption in relation to the 2DM, it can be explained that: Alcohol consumption increases or dicreases blood glucose levels, thus alcohol has calories. It is recommended to drink ocassionally and just when the level of sugar in the blood is controlled (Howard, Arnsten, & Gourevitch, 2004). Excessive alcohol consumption increases the risk of pre-diabetes and 2DM in women and men. On one hand, high consumption of alcohol increases the risk of abnormal

glucose regulation in men. On the other hand, associations are more complex because there is a higher risk with a high consumption of alcohol (Cullmann, Hilding, & Östenson, 2012). Alcohol has efects in the genetics of diabetes for people who suffer type 2 diabetes mellitus. It is included aspects such as increasing obesity, inducing pancreatitis, alterations in carbohydrate metabolism and glucose. There are some periods of hypoglycemia when there is a high consumption of alcohol and there are also long fasting periods caused by the intake. However, there is another reason that can produce hyperglycemia, when people consume lower amounts of food or when they are coupled with meals (Dáz-Martńez, et al., 2009).

If X_7=1, 2, or 3 and X_1 woman, X_2, X_3, X_4, X_6 and X_9 remain constant with chance of developing 2DM based on the physical activity(FA), it is interpreted in the following way:

- If $X_7 = 1$ (FA, active) chance of developing 2DM is of 0.3519. That means that if the interviewee has an active physical activity, his/her level of risk to develop 2DM will be of 35.13%, with a level of confidence of 0.95.

$$P[E(\widehat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(1) + 0.079(0) \quad (50)$$

That is to say:

$$P[E(\widehat{Y})] = -0.380 \quad (51)$$

In terms of normal distribution:

$$Z_E = \text{normal}\{P[E(\widehat{Y})]\} = \text{normal}\{-0.38\} = 0.3519 \quad (52)$$

- If $X_7 = 2$ (FA, moderate) chance of developing 2DM is 0.3707, that means that if the interviewee has moderate physical activity, his/her level of risk to develop 2DM will be 37.07%, with a level of confidence of 0.95.

$$P[E(\widehat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(2) + 0.079(0) \quad (53)$$

That is to say es decir:

$$P[E(\widehat{Y})] = -0.331 \quad (54)$$

In terms of normal distribution:

$$Z_E = \text{normal}\{P[E(\widehat{Y})]\} = \text{normal}\{-0.331\} = 0.3707 \quad (55)$$

- If $X_7 = 3$ (FA, non active) chance of developing 2DM is of 0.3897, that means that if the interviewee has non active physical activity, his/her risk to develop 2DM will be 38.97%, with a level of confidence of 0.95.

$$P[E(\widehat{Y})] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(3) + 0.079(0) \quad (56)$$

That is to say:

$$P[E(\widehat{Y})] = -0.282 \quad (57)$$

In terms of normal distribution:

$$Z_E = \text{normal}\{P[E(\widehat{Y})]\} = \text{normal}\{-0.282\} = 0.3897 \quad (58)$$

On the basis of the table 5 and taking as a reference the context of genetic risk, it can be observed that people who have non active physical activity, have a higher risk to develop 2DM with regard to the other two activities which are active and moderate.

Table 5. Risk rate due to physical activity

FA,active	$\text{Rr} = \dfrac{FA_R}{Z_g} = \dfrac{0.3519}{0.3339} = 1.0353$	(59)
FA,moderate	$\text{Rr} = \dfrac{FA_R}{Z_g} = \dfrac{0.3707}{0.3339} = 1.0906$	(60)
FA, non active	$\text{Rr} = \dfrac{FA_R}{Z_g} = \dfrac{0.3897}{0.3339} = 1.1465$	(61)

Source: Own authorship.

Physical inactivity in relation with the 2DM, it is claimed that: Physical inactivity is considered as one of the most important risk factors for mortality in Mexico. It is associated with the emergence and lack of control of diverse chronic diseases for instance, obesity, hypertension, 2DM, dyslipidemias, osteoporosis and cancers (Haskell, et al., 2007). Nonetheless, physical inactivity is a modifiable risk factor to develop 2DM because exercise raises glucose uptake through independent mechanisms of insulin and it also increases the insulin sensitivity (Paternina-de la Ossa, Villaquirán-Hurtado, Jácome-Velasco, Galvis-Fernández, & Granados-Vidal, 2018).

If $X_9 = 1$ or 2. and X_1 woman, X_2, X_3, X_4, X_6 and X_7 remain constant, cause of developing the 2DM based on the type of food (TF), it will be interpreted in the following way:

- If $X_9 = 1$ (healthy eating) chance of developing 2DM is of 0.3631, that means that if the interviewee has a healthy eating, his/her level of risk to develop 2DM will be of 36.31%, with a level of confidence of 0.95.

$$P\big[E(\widehat{Y})\big] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(1)(62) \quad (62)$$

That is to say:

$$P\big[E(\widehat{Y})\big] = -0.35 \quad (63)$$

In terms of normal distribution:

$$Z_E = \text{normal}\big\{P\big[E(\widehat{Y})\big]\big\} = \text{normal}\{-0.35\} = 0.3631 \quad (64)$$

- If $X_9 = 2$ (unhealthy eating) chance of developing 2DM is of 0.3935, that means that if the interviewee has an unhealthy eating, his/her level of risk to develop DM2 will be 39.35%, with a level of confidence of 0.95.

$$P\big[E(\widehat{Y})\big] = -0.429 + 0.069(0) + 0.009(0) + 0.015(0) - 0.042(0) + 0.079(0) + 0.049(0) + 0.079(2) \quad (65)$$

That is to say:

$$P\big[E(\widehat{Y})\big] = -0.271 \quad (66)$$

In terms of normal distribution:

$$Z_E = \text{normal}\big\{P\big[E(\widehat{Y})\big]\big\} = \text{normal}\{-0.271\} = 0.3935 \quad (67)$$

On the basis of the table 6 and taking as a reference the context of genetic risk, it can be observed that people who have an unhealthy eating, show a higher level to develop 2DM with regard to a healthy eating.

Table 6. Risk rate due to type of food

TF, healthy	$Rr = \dfrac{TF_R}{Z_g} = \dfrac{0.3631}{0.3339} = 1.0682$	(68)
TF, unhealthy	$Rr = \dfrac{TF_R}{Z_g} = \dfrac{0.3935}{0.3339} = 1.1576$	(69)

Source: own authorship.

Unhealthy eating in relation to the 2DM, it is explained that: The prevalence of 2DM is associated with the westernized diet, rich in sugar, animal fats[1], refined starches, carbohydrates and meats (Oggioni, Lara, Wells, Soroka, & Siervo, 2014). Worldwide high calories intake or diets high in animal saturated fats, increase the risk of developing diseases associated with metabolic disorders such as obesity and diabetes (Gómez & Latorre, 2010).

6. Argument

2DM is one of the major world health emergencies in the twenty-first century. Each year more people are living this condition which can lead to complications over the life (International Diabetes Federation, 2015). Globally, 415 million people with DM were registered, and there is a projection in 2040 which estimates that there will be 642 million people with DM (International Diabetes Federation, 2015).

2DM and its complications are the main causes of death in most of the countries (International Diabetes Federation, 2015). It is the most common type of diabetes and it has increased along cultural and social changes. In high-income countries, adults develop 2DM until 91%. IDF, estimates that 193 million of people with 2DM are not diagnosed and, as a result; a higher risk to have complications.

Additionally, one of 15 adults have impaired glucose tolerance, this condition is associated with a higher risk to develop 2DM afterwards (International Diabetes Federation, 2015). In 2040, for each hundred thousand men 7.7% will develop 2DM, and for each hundred thousand women 6.4% will develop DM. The rate will increase with regard to that 2015 (International Diabetes Federation, 2015).

2DM is the most prevailing type of diabetes in the world, adults are generally susceptible to develop it, but nowadays is also developed more frequently in children and teenagers. There are people with 2DM who have not been diagnosed yet because of the minor symptoms than those showed in type 1diabetes and also because years may pass before they are

---

[1] Saturated fat has been associated with cholesterol LDL (Low density lipoproteins, since its acronyms in English), it is one of the risk factors in the development of the coronary heart disease, and thus it raises the levels of blood cholesterol (García-Sancho, 2013).

diagnosed (International Diabetes Federation, 2015).

7. Conclusions

Gaussian model was done through risk factors to explain 2DM, it illustrated the behavior of the phenomenon. In light of the foregoing and considering the complexity of the phenomenon of 2DM in function of the risk factors, it was taken as an element Gaussian model, due to the response variable adjusted to the probability model from a Gaussian. When applying the technique of generalized linear models, the results obtained in the construction of Gaussian model were the following. Firstly, gage runs were done through different distributions in which the response variable can be modeled (2DM), these distributions match with the Logit, Probit and Gaussian model, since a nominal and cardinal context, taking as a reference a level of confidence from $(1 - \alpha)$ to 0.95 and a level of significance from (α) to 0.05. The ordinal regression Gaussian model was the one with the best adjustment, it let see that the intercept X_{1woman} (sex), X_2 (age), X_3 (residence zone), X_4 (family history with 2DM), X_6 (today's alcohol consumption), X_7 (physical activity) and X_9 (type of food) are significant within the model, thus P-value of all of them is below from the level of significance ($\alpha = 0.05$). Likelihood to develop 2DM impact in the variables X_{1woman}, X_2, X_3, X_4, X_6, X_7 and X_9. All the parameters are meaningful to predict 2DM, thus they are within the confidence intervals, with a level of confidence of 0.95, and a level of significance of 0.05.

On the basis of these results, it can be concluded that, if there is not intervention of the explicative variables in any of the interviewees, the likelihood to develop 2DM is of 33.39%. If the interviewee is woman, cause of developing 2DM is of 35.94%. If probability of risk to all the interviewees is of 33.39% without taking into account the sex, women's risk is of 39.94%,then the risk rate of women with regard to men is of 7.63%. Cause of developing 2DM is of 33.72%, in other words, age is a factor to develop 2DM. For each elapsing year, the probability of risk of the interviewees to develop 2DM is approximatedly of 0.98%, that means of 1%. People who have their place of origin in the metropolitan zone, show a higher risk to develop 2DM with regard to the rest of the zones such as rural and urban. People who have both parents with 2DM, show higher risk to develop 2DM with regard to the other descendants such as father and mother. People who consume alcohol, have higher risk to develop 2DM with regard to people who do not do it. People have non-active physical activity, have higher risk to develop 2DM with regard to the other two activities such as active and moderate. People who have unhealthy eating, have a higher risk to develop 2DM with regard to healthing eating. The risk factors as explanatory elements of the prevalence of 2DM, are utmost importance to delay or to control the early development of 2DM, because some of them are factors that can be modified.

8. Final Considerations

The survey that was used was the ENSANUT (2012) is self-report information of type 2 diabetes, the results would be more convincing if the data were from blood glucose tests of the observations. The model that was carried out is probabilistic in which there is a 95% confidence level, however, the 5% level of significance is present; that is, according to the model, 95% of respondents should behave based on the equation and 5% go out of that behavior. The work would be improved by adding more variables, which would allow to give greater clarity and specificity to the topic, to use more updated data. This work emphasizes that further studies are carried out in greater depth.

References

American Diabetes Association. (2013). Aspectos genéticos de la diabetes. *ADA*. Retrieved 04 16, 2018, from http://www.diabetes.org/es/informacion-basica-de-la-diabetes/aspectos-genticos-de-la-diabetes.html

American Diabetes Association. (2017a). *Standards of Medical Care in Diabetes*. Retrieved 01 02, 2018, from THE JOURNAL OF CLINICAL AND APPLIED RESEARCH AND EDUCATION: http://care.diabetesjournals.org/content/diacare/suppl/2016/12/15/40.Supplement_1.DC1/DC_40_S1_final.pdf

American Diabetes Association. (2017b). Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care*. https://doi.org/10.2337/diacare.20.7.1183

American Diabetes Association. (2017b). Los Standards of Medical Care in Diabetes. *Diabetes Care*. Retrieved 02 21, 2018, from http://www.redgdps.org/standards-of-medical-care-in-diabetes-2017/

Barquera, S. (2003). Prevención de la diabetes: Un problema mundial. *Salud Pública*, 413-414. https://doi.org/10.1590/S0036-36342003000500013

Bonilla, A. (2017). Variantes geneticas de alto impacto en desarrollo de enfermedades metabolicas. *Cienciamax*. Retrieved Enero 12, 2020, from http://www.cienciamx.com/index.php/ciencia/salud/13413-detectan-variantes-geneticas-alto-impacto-en-desarrollo-de-enfermedades-metabolicas

Cullmann, M., Hilding, A., & Östenson, C. G. (2012). Alcohol consumption and risk of pre-diabetes and type 2 diabetes development in a Swedish population. *Diabetic Medicine*, *29*(4), 441-452. https://doi.org/10.1111/j.1464-5491.2011.03450.x

Díaz Martínez, L. R., Díaz Martínez, A., Hernández-Ávila, C. A., Fernández Varela, H., Solís Torres, C., & Narro Robles, J. (2009). El consumo riesgoso y dañino de alcohol y sus factores predictivos en adolescentes estudiantes del bachillerato. *Salud mental*, *32*(6), 447-458.

Martín, L. G. S. (2013). El diccionario de términos médicos de la Real Academia Nacional de Medicina. *Anales (Reial Acadèmia de Medicina de la Comunitat Valenciana)*, (14), 19-55. https://dialnet.unirioja.es/servlet/articulo?codigo=4730543

Gómez, D., & Latorre, G. (2010). La transición en la epidemiologia y Salud Pública ¿Explicación o condena? *Revista Facultad Nacional de Salud Pública Medellín*. https://doi.org/10.17533/udea.rfnsp

González-Villalpando, C., Dávila-Cervantes, C. A., Zamora-Macorra, M., Trejo-Valdivia, B., & González-Villalpando, M. E. (2014). Risk factors associated to diabetes in Mexican population and phenotype of the individuals who will convert to diabetes. *salud pública de méxico*, *56*(4), 317-322. https://doi.org/10.21149/spm.v56i4.7351

Goryakin, Y., Rocco, L., & Suhrcke, M. (2017). The contribution of urbanization to non-communicable diseases: Evidence from 173 countries from 1980 to 2008. *Economics & Human Biology*, *26*, 151-163. https://doi.org/10.1016/j.ehb.2017.03.004

Haskell, W., Lee, I., Pate, R., Powell, K., Blair, S., & Franklin, B. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc*, 1423-1434. https://doi.org/10.1249/mss.0b013e3180616b27

Hernández-Ávila, M., Gutiérrez, J. P., & Reynoso-Noverón, N. (2013). Diabetes mellitus in Mexico. Status of the epidemic. *Salud pública Méx*, *55*(2). Retrieved 01 16, 2018, from http://webcache.googleusercontent.com/search?q=cache:aEmTlECe1HUJ:www.scielo.org.mx/scielo.php%3Fscript%3Dsci_arttext%26pid%3DS0036363420130008000009+&cd=1&hl=es-419&ct=clnk&gl=mx

Howard, A., Arnsten, J., & Gourevitch, M. (2004). Effect of alcohol consumption on diabetes mellitus: a systematic review. *PubMed*, 211-219. https://doi.org/10.7326/0003-4819-140-6-200403160-00011

International Diabetes Federation. (2015). IDF diabetes atlas - 7th edition. *IDF*. Retrieved 06 21, 2018, from http://www.diabetesatlas.org/

Kautzky-Willer, A., Harreiter, J., & Pacini, G. (2016). Sex and Gender Differences in Risk, Pathophysiology and Complications of Type 2 Diabetes Mellitus. *PubMed*. https://doi.org/10.1210/er.2015-1137

Main causes of mortality-INEGI 2000, 2010 & 2015. (2015). Principales causas de mortalidad por residencia habitual, grupos de edad y sexo del fallecido. *INEGI*. Retrieved 02 23, 2017, from http://www.inegi.org.mx/est/contenidos/proyectos/registros/vitales/mortalidad/tabulados/PC.asp?t=14&c=11817

Medina, C., Tolentino-Mayo, L., López-Ridaura, R., & Barquera, S. (2017). Evidence of increasing sedentarism in Mexico City during the last decade: Sitting time prevalence, trends, and associations with obesity and diabetes. *Open Access Journal*. https://doi.org/10.1371/journal.pone.0188518

NHNES. (2012). *Resultados Principales*. Retrieved 03 17, 2016, from Encuesta Nacional de Salud y Nutrición: https://ensanut.insp.mx/resultados_principales.php

Oggioni, C., Lara, J., Wells, J., Soroka, K., & Siervo, M. (2014). Shifts in population dietary patterns and physical inactivity as determinants of global trends in the prevalence of diabetes: an ecological analysis. *PubMed*, 1105-1111. https://doi.org/10.1016/j.numecd.2014.05.005

Paternina-de la Ossa, A., Villaquirán-Hurtado, A., Jácome-Velasco, S., Galvis-Fernández, B., & Granados-Vidal, Y. (2018). Actividad física en pacientes con diabetes mellitustipo 2 y relación con características sociodemográficas, clínicas y antropométricas. *Universidad y Salud*, 72-81. https://doi.org/10.22267/rus.182001.111

Peña, D. (1987). *Estadística. Modelos y Métodos*. Madrid: Alianza.

Schmidhuber, J., & Prakash, S. (2007). The nutrition transition to 2030. Why developing countries. *Food Economics-Acta Agriculturae Scandinavica, 2*(3), 150-166. https://doi.org/10.1080/16507540500534812

Taha, H. (2012). *Investigación de Operaciones*. Arkansas: Pearson.

Torres, F., & Trápaga, Y. (2001). La alimentación de los mexicanos en la alborada del tercer milenio. *UNAM*.

Trindade, C., Dos Santos, L., Dalva de Barros, M., & Silvia, S. (2014). Factores de riesgo asociados a las enfermedades cardiovasculares en adultos. *Revista Latino-Americana de Enfermagem*, 547-553.

World Diabetes Report. (2016). Diabetes Mundial. *OMS*. Retrieved 08 19, 2018, from http://apps.who.int/iris/bitstream/handle/10665/254649/9789243565255-spa.pdf;jsessionid=A0F8D635E0390E53F B9ADE76E7339088?sequence=1

World Health Organization. (2018). Enfermedades no transmisibles, Diabetes - Datos y cifras. *OMS*. Retrieved 07 02, 2018, from http://www.who.int/es/news-room/fact-sheets/detail/noncommunicable-diseases

Copyrights