Universe Scientific Publishing

# Blood Oxygen Saturation Analysis Model Based on Multiple Normalized Regression

**Chengyu Mu**
**Shenyang Aerospace University, Shenyang 110136, China.**

***Abstract:***Several factors that may have an impact on the saturation of the oxygen saturation, respectively, are age, BMI, gender, and current smoking conditions. For the blood oxygen saturation of 36 patients collected, the data preprocessing, the abnormal data were removed, and the blood oxygen saturation was the dependent variable, with the BMI, age, gender, the current smoking condition is the independent variable, and the stata is normalized. In the case of significant sexual testing of variables, the discovery of the current smoking condition was significant for the oxygen saturation of the blood oxygen, while the bmi and gender in the t test were not different from zero, indicating that the blood oxygen saturation was affected by the BMI, and the influence of gender could be ignored, and the effect of age and current smoking condition on the oxygen saturation was mainly considered.

***Keywords:*** Oxygen Saturation; Normalized Regression

## 1. Introduction

Normally, arterial blood has an oxygen saturation of 98 percent and venous blood 75 percent. Hypoxia Imbalance between oxygen supply and oxygen consumption in the body, that is, tissue cell metabolism in anoxic state. Whether the body is hypoxic depends on whether the oxygen transport and oxygen reserve of each tissue can meet the needs of aerobic metabolism. The harm of hypoxia is related to the degree, occurrence speed and duration of hypoxia. Severe hypoxemia is a common cause of death from anesthesia, accounting for about one-third to two-thirds of deaths from cardiac arrest or severe brain cell damage. Thus, it can be seen that blood oxygen saturation is important for a variety of diseases.

Oxygen saturation is an important indicator to describe cardiovascular diseases. We conducted multiple regression analysis on oxygen saturation of 36 patients with various physical characteristics.

## 2. Evaluation model of factors affecting oxygen saturation based on multiple linear regression

### 2.1 Data preprocessing

For Annex I, we take data on (+3, -3), according to reference, due to the love of measurement errors and the limitations of measurement instrument accuracy. Where means the sample mean and means the standard deviation. Because the data volume is large enough, based on the central limit law and the law of large numbers, we take the sample mean as the population mean and the sample variance as the population variance. The sample mean and sample variance of some patients are shown in Table 1:

Table 1 part of sample mean and sample variance

| Patients with the serial number | Sample mean | Sample variance $S^2$ |
|---|---|---|
| 010217A | 98.24 | 0.53 |
| 010217B | 98.16 | 0.54 |
| 010217C | 94.41 | 1.65 |

Due to the large number of patient data, only some relevant data of patients are displayed. Please refer to the attachment for the remaining data.

### 2.2 Introduction of factors affecting blood oxygen saturation

Blood oxygen saturation is the percentage of oxygen-bound hemoglobin capacity in the blood as a percentage of all binding hemoglobin capacity, that is, the oxygen concentration in the blood, which is an important physiological parameter in the respiratory cycle .Arterial blood has an oxygen saturation of about 98% and venous blood of about 75%.The process of human metabolism is the process of biological oxidation, and the oxygen needed in the process of metabolism, is to enter the blood through the respiratory system, and the hemoglobin in the blood red blood cells, combined into oxygenated hemoglobin, and then transported to the human body to all parts of the tissue cells. The ability of the blood to carry oxygen is measured by oxygen saturation.

Through consulting literature ， we know that hypoxia has a huge impact on the body. For example, there are effects on CNS, liver and kidney function. In a hypoxic state, compensatory heart rate increases, cardiac beat and cardiac output increases, and the circulatory system compensates for oxygen deficiency in a highly dynamic state. Redistribution of blood flow and selective dilation of brain and coronary vessels are also produced to ensure adequate blood supply. However, in severe hypoxic conditions, ATP synthesis is reduced due to subendocardial lactic acid accumulation, resulting in myocardial inhibition, leading to bradycardia, pre-phase contraction, decreased blood pressure and cardiac output, as well as ventricular fibrillation and other arrhythmias and even cardiac arrest. Therefore, we can see that the study of oxygen saturation is of high physiological and medical significance.

In order to establish a typical model to describe blood oxygen saturation, the following fields that may influence blood oxygen saturation are selected for study.

(1) BMI:

Body mass index is a standard commonly used in the international community to measure the degree of body fat and thin and whether a healthy standard. Abnormal BMI may be due to obesity or weight loss. The normal medical BMI ranges from 18.5 to 24.5, with a BMI below 18.5 considered too light and a BMI above 24.5 considered overweight, suggesting that this level of obesity may have an impact on oxygen saturation.

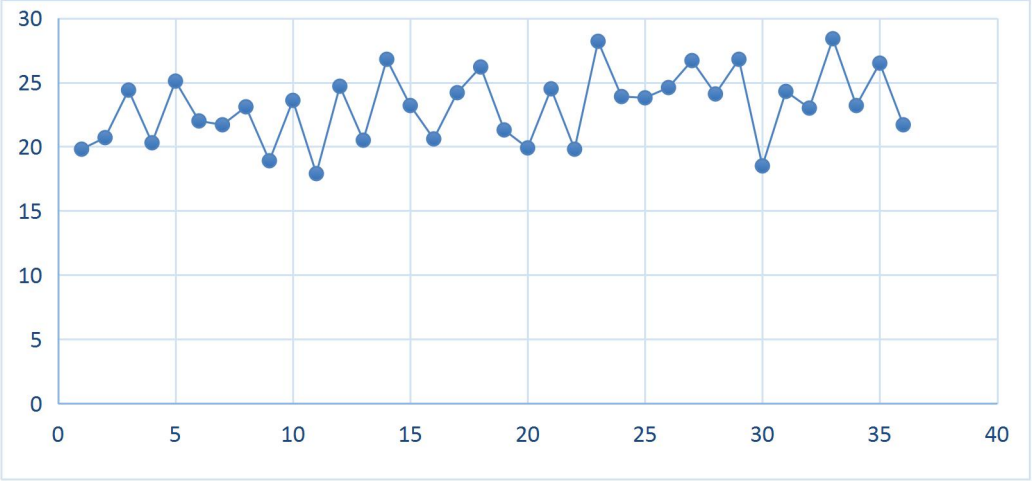


Figure 1 BMI of the patients given in the sample

By analyzing the BMI of the patients given in the sample, it was found that one patient was too thin, 15 patients were overweight, and the remaining patients were all normal.

(2) Gender

Does gender have any effect on human health? According to the lancet journals again , because men and women there is biggish difference in (1) physical structure, men have only one x chromosome from the mother, and on the y chromosome specific genes may make some disease increases, the probability for women, on the x chromosome genetic diversity more abundant, it may increase disease incidence, at the same time can also bring benefits. There are also large differences in lifestyle habits (men tend to smoke and drink, while women pay more attention to health and hygiene).

Thus, gender may indeed have an effect on oxygen saturation in both sexes.

(3) Age

With the increase of age, the water content of human body decreases significantly. At the same time, the activity of

various enzymes of human body decreases, energy demand decreases, and cell metabolism rate slows down, which is likely to lead to the reduction of oxygen demand of some cells, thus affecting blood oxygen saturation

Patients were roughly divided into three sections according to age, respectively enantiodromia the young (19,36), the middle-aged (36,53), and the elderly (53,70). The right side of the table represents the percentage of the total number of patients at a certain age, while the left side represents the total number of patients at a certain age.It was found that young patients accounted for the largest proportion, about 60% of the total number of patients, while the elderly patients accounted for the least, about 17%.

(4) Current smoking status

According to literature review, smoking is a major risk factor for many cardiovascular and cerebrovascular diseases. As high-density lipoprotein cholesterol (HDL-C) can stimulate the production of prostaglandin (PGI2) in vascular endothelial cells, PGI2 is the most effective substance for vascular dilation and inhibition of platelet aggregation. Smoking will damage vascular endothelial cells, resulting in decreased serum HDL-C, elevated cholesterol, and decreased PGI2 level, leading to contraction of peripheral blood vessels and coronary arteries, thickening of the wall, narrowing of the lumen and slowing of blood flow, resulting in myocardial hypoxia. In addition, carbon monoxide in smoke combines with hemoglobin to form carbon oxygen hemoglobin, which affects the oxygen-carrying capacity of red blood cells. Therefore, whether a patient smokes or not has a significant impact on blood oxygen saturation.

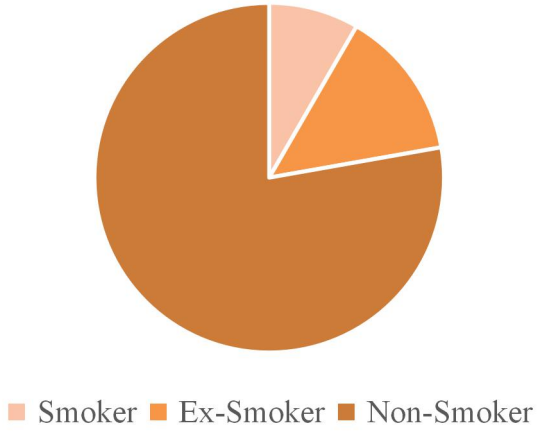


Figure 2 Statistical chart of current smoking status of patients

As can be seen from the figure, among the 36 patients, only 3 were smokers, 5 were smokers, and the vast majority were never smokers (28).

## 2.3 Evaluation model of factors influencing blood oxygen saturation based on multiple linear regression

Since gender and dummy variables are qualitative variables, they are not suitable for mathematical modeling, so it is necessary to introduce dummy variables to convert them into quantitative variables. A dummy variable is an artificial variable used to reflect the attributes of a quality. It is a quantized independent variable, often with the value of 0 or 1. The introduction of dummy variables will make the linear regression model more complex, but the description of the problem is clearer and the quantization of qualitative variables is completed.

After the introduction of dummy variables, the regression equation of blood oxygen saturation can be expressed as:

$$Oximetry = c_0 + c_1 BMI + c_2 Age + c_3 Gender + c_4 Smoking1 + c_5 Smoking2 + \mu_i$$

Oximetry is the virtual variables introduced by blood oxygen saturation, Gender and Smoking, $c_0$ is the constant term of the regression equation, $c_i$ is the regression coefficient of the regression equation, $\mu_i$ is the perturbation term of the regression equation, and the equation should follow the normal distribution when there is no endogen.

Take Gender as an example to explain it:

Suppose that the I sample is male, and Gender=0 means that the I sample is female, then the expectations of Oximetry in these two states are discussed separately.

$$E(Oximetry|Gender = 1 Control\ other\ independent\ variables\ constant) = c_3 \times 1 + c$$

$$E(Oximetry|Gender = 0 Control\ other\ independent\ variables\ constant) = c_3 \times 0 + c$$

Under constant control of other variables, all independent variables except the target variable can be represented by constant C, then the difference of expected values:

$$E(Oximetry|Gender = 1 Control\ other\ independent\ variables\ constant) -$$

$$(Oximetry|Gender = 0 Control\ other\ independent\ variables\ constant) = c_3$$

Therefore, $c_3$ can be interpreted as the difference value of male and female oxygen saturation under the control of other independent variables remaining unchanged, which makes a reasonable explanation for the qualitative variable.

After pretreatment of the data and the addition of dummy variables, multiple linear regression with Oximetry as the dependent variable and BMI, Age, Gender and Smoking as the independent variables was carried out, and the results were as follows:

Table 2 Multiple linear regression analysis of blood oxygen saturation table of variance

| source | ss | df | MS |
|---|---|---|---|
| model | 17.98 | 5 | 3.40 |
| residual | 36.42 | 30 | 1.21 |
| total | 54.40 | 35 | 1.55 |

| |
|---|
| Number of obs = 36 |
| F(5,30) = 2.96 |
| Prob > F = 0.0273 |
| R-squared = 0.3305 |
| Adj R-squared = 0.2190 |
| Root MSE = 1.102 |

First of all, we can see that *P=0.0273<0.05* indicated that the regression model passed the joint significance test.

At the same time, the goodness of fit of the regression model $R^2 = 0.219$, and the sum of squared errors SSE=54.40, if the model is used as a prediction model, the prediction results will have a large error. However , the regression models we are familiar with can usually be divided into two types, namely, explanatory regression and predictive regression.

(1) Predictive regression requires the regression equation to express the dependent variable to the maximum extent, which requires a high degree of goodness of fit R^2, otherwise it will lead to a huge deviation in the prediction results.

(2) Explanatory regression pays more attention to the statistical significance of each independent variable and the overall significance of the model, focusing on the causal relationship between the established research independent variable and dependent variable.

Table 3 blood oxygen saturation multiple linear variance coefficient of expansion

| Variable | VIF |
|---|---|
| smoking2 | 1.95 |
| smoking3 | 1.88 |
| Age | 1.50 |
| BMI | 1.47 |
| gender1 | 1.21 |
| Mean VIF | 1.60 |

Variance coefficient of expansion (VIF) is an important index to measure the multicollinearity severity of multiple linear regression models.It represents the ratio of the variance of the estimator of the regression coefficient to the variance of the assumed independent variable with no linear correlation.

$$VIF = \frac{1}{1 - R_i^2}$$

It is generally believed that VIF<5 means that there is no multicollinearity among all independent variables. For example, each independent variable VIF in the table is less than 5, indicating that there is no multicollinearity in this regression model.

Table 4 Analysis table of multivariate linear regression coefficient of blood oxygen saturation

| Variable | Coef | Std.Err | t | P>\|t\| |
|---|---|---|---|---|
| BMI | 0.054 | 0.820 | 0.66 | 0.513 |
| Age | -0.041 | 0.014 | -2.87 | 0.007 |
| smoking1 | | | (omitted) | |
| smoking2 | -0.457 | 0.617 | -0.74 | 0.465 |
| smoknig3 | -2.068 | 0.911 | -2.27 | 0.031 |
| gender1 | 0.347 | 0.404 | 0.86 | 0.398 |
| gender2 | | | (omitted) | |
| cons | 98.170 | 1.877 | 52.3 | 0 |

To eliminate multicollinearity for regression results, we will introduce the classification of the virtual variable number minus one, each independent variable corresponding t inspection P value, BMI and Gender were greater than 0.05, shows that the independent variable on the dependent variable and there is no difference between zero, that won't affect, and smoking and Age are all passed the test of significance, Oxygen saturation was only associated with age and current smoking status, but not with gender or obesity. At the same time, we also noticed that for the parameter smoking2 (who has quit smoking), there was no difference between the blood oxygen saturation and zero, indicating that only the current smoking status had an impact on the blood oxygen saturation by smoking, while there was no difference between the influence of never smoking and quitting smoking on the blood oxygen saturation relative to smoking.

# 3. Conclusion

For multiple factors that may affect the oxygen saturation: age, BMI, gender, standardized regression:

(1)Blood oxygen saturation is affected by a small amount of sex and bmi

(2)Blood oxygen saturation is significantly affected by age and current smoking conditions

# References

[1] Rodríguez-Molinero Alejandro, Narvaiza Leire, Ruiz Jorge, Gálvez-Barrón César. Normal respiratory rate and peripheral blood oxygen saturation in the elderly population.[J]. Journal of the American Geriatrics Society, 2013, 61(12).

[2] Oxygen saturation. [J]. Emergency nurse : the journal of the RCN Accident and Emergency Nursing Association, 2004, 12(2).

[3] Kapur Vishesh K, Wilsdon Anthony G, Au David, Avdalovic Mark, Enright Paul, Fan Vincent S, Hansel Nadia N, Heckbert Susan R, Jiang Rui, Krishnan Jerry A, Mukamal Kenneth, Yende Sachin, Barr R Graham. Obesity is associated with a lower resting oxygen saturation in the ambulatory elderly: results from the cardiovascular health study.[J]. Respiratory care, 2013, 58(5).