International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 5
ISSN: 2454-4248
104 – 109

Application of the Variance Function of the Difference Between two estimated responses in regulating Blood Sugar Level in a Diabetic patient using Herbal Formula

Karanjah Anthony N.
School of Science
Maasai Mara University
Narok, Kenya.
karanjah@mmarau.ac.ke;

F. Njui
School of Mathematics
University of Nairobi
Nairobi, Kenya.
francis.njui@uonbi.ac.ke

Pokhariyal G.P
School of Mathematics
University of Nairobi
Nairobi, Kenya.
pokhariyal@uonbi.ac.ke

***Abstract*—** In this paper response surface methodology has been employed in investigating effectiveness of herbal formula (a mix of six herbs) in regulating the blood sugar level of a diabetic to within acceptable levels. In the experimentation phase, observations are made to investigate the effectiveness for particular level of concentration at regulating the blood sugar level with time.

The most feasible of all the identified points of equal yield has been identified as one in which the variance function is minimal. In this investigative research we use the variance function of the difference between two points to provide reliable advice on the range around which the dosage is desirable and time required to effectively regulate the blood sugar level to within acceptable range.

In the set up the herbal formula extract has been shown to have successfully regulated the blood sugar level in a diabetic to 11.3898 mMol/L. This is possible by effecting a treatment of herbal formula at a concentration of 66.1125 mg/dl and the effectiveness is within 175.4580 minutes upon treatment.

***Keywords*-** *Diabetes, Herbal-Formula, Response surface, Treatment, Variance function.*

*****

1. Introduction

Response surface methodology as a statistical technique is useful in modeling and analysis of problems in which response of interest is influenced by several variables for which the objective is to optimize the response. This is achieved by locating feasible treatment combinations for which the mean response is optimized. In any treatment arrangement, we seek a treatment or treatment combination that can be used to either reverse a condition, eradicate it or arrest it in order to minimize suffering or to help the patient bear a condition with less pain.

This research employs response surface methodology to investigate effectiveness of herbal medicine in reducing the blood sugar level of a diabetic patient to a level that is acceptable. In this setup, observations are made to investigate effectiveness for particular dosage at reducing the blood sugar level with time. The variance function of the difference between two points is used to provide reliable advice on the range around which the dosage is desirable and time required, to accomplish it.

The experimental set up was based on herbal drug (referred as Herbal Formula) extract sourced from a leading herbalist in Kenya. The herbal formula used, is a mix of six herbs, whose botanical name are: Momodica foetida, Utica Masaica, Cinamon Species, Azandracta Indica, Moringa Oliefera and Gymnema Sylvestre.

In the set-up 24, albino rats were involved with the treatment procedure of four groups all induced with diabetes as per laboratory standards and procedures. After the recording of the Fasting Blood glucose the treatments were carried on the animals as per grouping and dosage levels effected using the Herbal Formula with varied concentrations of 125 mg/Kg, 250 mg/Kg and 500 mg/Kg. The control group was treated with conventional drug Metformin 500mg/Kg.

The readings of the Oral Glucose Tolerance Test (OGTT) was undertaken at $t = 30$, $t = 60$ and $t = 90$ minutes, which availed the data in use.

2. The Model

The desire of any pharmaceutical process is to develop a formulation which is acceptable or effective in shortest time possible and at the same time using minimum number of man-hours and raw materials. Traditionally, pharmaceutical formulations are developed by changing one variable at a time by trial and error method which is time consuming. Further it requires a lot of imaginative efforts, Saeed Ghanbarzadeh et al (2013). Moreover, it may be difficult to develop an ideal formulation using this classical technique, since the joint effects of independent variables are not considered. It is therefore very essential to understand the complexity of pharmaceutical formulations by using a collection of mathematical and statistical technique which quantifies the functional relationship between a number of measured

 ISSN: 2454-4248
104 – 109

response variables and several explanatory factors to obtain an optimal response by using a series of tests.

In a treatment arrangement, we seek treatment or treatment combination which may be used to either reverse a condition, eradicate or arrest a condition in order to minimize suffering or to help the patient bear a condition with less pain, Karanjah et al (2015). We seek to regulate the blood sugar level y of a diabetic patient at particular time to a condusive level according to medical standards. We make observations so as to find out how effective a particular dosage and concentration of the herbal formula in question is at regulating the blood sugar level with time. In this investigation we will determine;

(i). Time taken x_1, to regulate the blood sugar level to within acceptable standard, and

(ii). The best possible level of concentration x_2, of the identified herbal Formula.

The variance function of the difference between two points is used to provide reliable advice (on the basis of the findings) on the range around which the dosage is desirable and time required.

In Response Surface Methods, the true response function f is unknown or complex, we therefore need to approximate the function. In order to develop a proper approximation for f, we model the data by starting with a low-order polynomial in some small region. If the response can be defined by a linear function of independent variables, then the approximating function is a first-order model. With reference to any d^{th} order polynomial regression model the general design for n observations is given as

$$y_i = f'(x_i)\beta + \epsilon_i, \quad i = 1,2,3,\ldots n. \qquad (1)$$

Specifically, a multiple-regression model with k independent variables is of the form

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i, \quad i = 1,2,3,\ldots n.$$
$$= \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \epsilon_i, \quad j = 1,2,3,\ldots,k \qquad (2$$

where parameter β_j measures expected change in response y per unit increase in xi when other independent variables are assumed constant. The i^{th} observation and j^{th} level of independent variable is denoted by x_{ij}.

In case curvature exists, then a higher order model is to be fitted to the data to explore the nature of the response surface. The data generated in the experimental set up indicated curvature and hence a second order model is explored. The second order model is of the form;

$$y_i = \beta_0 + \sum_{j=1}^{q} \beta_j x_j + \sum_{j=1}^{q} \beta_{jj} x_{jj}^2 + \sum_i \sum_{j<i} \beta_{ij} x_i x_j + \epsilon_i) \qquad (3)$$

$$= \beta_0 + x'_i \beta_i + x'_i \beta_i x_i + \epsilon_i \qquad (4)$$

Specifically with two predictor variables, equation (4) is of the form,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2 + \beta_{12} x_{i1} x_{i2} + \epsilon_i \qquad (5)$$

where β_{12} is the interaction effect coefficient between time and concentration.

3. Parameter Estimates

The regression coefficients are determined using the method of least squares. The same results can be obtained easily by use of Design of Experiment (DoE) for the regression analysis. The second order model parameter estimates the degrees of freedom, the corresponding standard error of the estimates as well as the 95% confidence interval of the parameters generated by the design of experiment software are as in the following table.

Table 1 Parameter Estimates for the Quadratic model

Factor	Estimate	d.f	s. error	95% LCL	95% UCL
Intercept	13.0718	1	1.96	9.09	17.06
Time x_1	-2.2593	1	1.06	-4.41	0.11
Conc. x_1	-3.6393	1	1.23	-6.14	-1.14
$x_1 x_2$	0.8157	1	-.81	-0.84	2.47
x_1^2	-0.1326	1	.93	0.41	4.20
x_2^2	2.3041	1	1.81	-3.81	3.54

The corresponding regression equation is given as

$$\hat{y}_i = 13.0718 \pm 2.2593 x_1 - 3.6393 x_2 - 0.1326 x_1^2 + 2.3041 x_2^2 + 0.8157 x_1 x_2 \qquad (6)$$

where the variables x_1 is *time* and the x_2 is *concentration*. The regression coefficients are interpreted as follows:

(i). $\beta_0 = 13.7018$, is the mean response, which is the blood sugar level in the absence of the factors of interest.

(ii). $\beta_1 = -2.2593$, shows that a unit change in x_1 causes a decrease in blood sugar level by 2.2593 units.

(iii). $\beta_2 = -3.6393$, indicates a unit change in x_2 causes a decrease in blood sugar level by 3.6393 units.

(iv). $\beta_{11} = -0.1362$, indicates a unit change in x_1^2 causes a decrease in blood sugar level by -0.1362 units.

(v). $\beta_{22} = 2.3041$, indicates a unit change in x_2^2 causes an increase in blood sugar level by 2.3041 units.

(vi). $\beta_{12} = 0.8157$, indicates a unit change in $x_1 x_2$ causes an increase in blood sugar level by 0.8157 units.

The ANOVA table generated from the regression equation is as follows:

 ISSN: 2454-4248
104 – 109

Table 2 ANOVA for Response Surface Quadratic model

Source	S.S	d.f	Mean S.S	F-value	p-Value
Model	392.20	5	78.44	2.7369	0.0035
Time x_1	130.53	1	130.53	4.55	0.0401
Conc. x_1	251.71	1	251.71	8.78	0.0055
x_1x_2	28.86	1	28.86	1.01	0.3227
x_1^2	0.15	1	0.15	0.00054	0.9420
x_2^2	174.65	1	174.65	6.09	0.0188
Residual	974.51	34	28.66		
Lack of Fit	27.62	3	9.21	0.30	0.8241
Pure Error	946.89	31	30.54		
Total	1366.71	39			

4. Tests of hypotheses

4.1 Model Adequacy

The hypothesis for the overall performance of the model is stated as follows

$H_0: \beta_1 = \beta_2 = \beta_{11} = \beta_{22} = \beta_{12} = 0$, against

$H_1: \beta_j \neq 0$, for at least one j.

The computed Fc value of 2.7369 in the above ANOVA table compared to $F0.05,(5,34)$ implies that the model is significant, with only a 3.50% chance that an F-value this large could occur due to error (noise).

Coefficient of multiple determination $R^2 = 0.2870$ shows that 28.70 % of the variation in the blood sugar level is accounted for by the model.

4.2 Individual parameter estimates

The hypotheses to be tested for the three parameters can be stated in general as;

$H_0: \beta_j = 0$, against

$H_1: \beta_j \neq 0$, for $j = 1, 2, 3, 11, 22, 12$

Using the values provided in the ANOVA table, we find that $p - value$ values which are less than 0.0500 imply that model terms are significant. In this case x_1, x_2 and x_2^2 are significant model terms. This is an indication that the linear terms in the second order model are all significant and the quadratic term associated with concentration is significant.

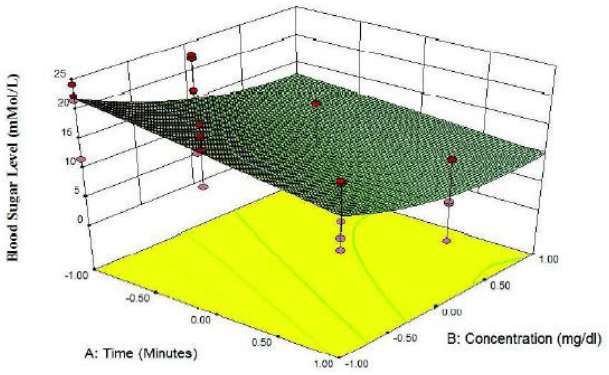
This is an improvement when compared to the first order model. The Lack of Fit F-value of 0.30 implies the Lack of Fit is not significant relative to the pure error. This indicates that there is an 82.41% chance that a 'Lack of Fit F-value' this large could occur due to noise.

5. Analysis of the Stationary Point of the Second-Order Model

When there is a curvature in the response surface the first-order model is not sufficient.

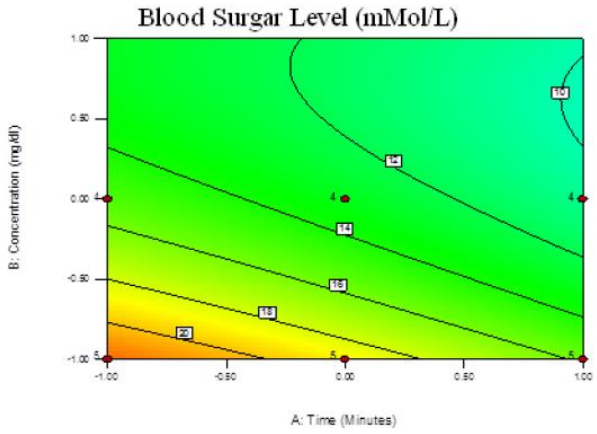
Thus, a second-order model becomes useful in approximating a portion of the true response surface with parabolic curvature. Using a statistical software (Design of Experiments-DoE) in analysis of a quadratic response, we get the following three dimension plots for the two continuous factors *time* and *concentration*;

Figure 1. Three d-surface plot View



The accompanying contour plot for the three dimension view is as follows:

Figure 2 Contour plot



Both the three dimension plot and its accompanying contour plot above reveals a trough that suggests minimum function throughout this region described by the factor combinations.

Hence the need to explore the design space by use of a second-order model which is flexible as it takes a variety of functional forms and approximates the response surface locally which is a good estimation of the true response surface. .

Using these results we conclude that response surface may be better visualized by the second-order model. Graphical visualization of contour plots helps in understanding the second-order response surface. It also helps characterize the shape of the surface. These observations helps us locate the optimum response, Karanjah et al (2008). With this

 ISSN: 2454-4248
104 – 109

information we now determine the optimum setting and recommend them for the effective management of average sugar level in a diabetic patient using the herbal formula.

Using the fit of the second-order models, we illustrate quadratic response surfaces such as minimum, maximum, ridge, and saddle point. In the case that an optimum exits, then this point is a stationary point which can result in any of the aforementioned four possibilities.

The stationary point in response surface models is the combination of design variables, where the surface is at either a maximum or a minimum in all directions. If the stationary point is a maximum in some direction and minimum in another direction, then the stationary point is a saddle point.

The stationary point is evaluated by use of matrix algebra for which the fitted second order model (5) in matrix form is expressed as follows:

$$\hat{y} = \hat{\beta}_0 + x'b + x'Bx. \quad (7)$$

The derivative of $\hat{y}$ with respect to the vector $\mathbf{x}$ will be,

$$\frac{\delta\hat{y}}{\delta x} = b + 2Bx, \quad (8)$$

$$x_s = -\frac{1}{2}B'b, \quad (9)$$

where $\boldsymbol{b}$ is a $(q \times 1)$ vector of the first-order regression coefficients and $\boldsymbol{B}$ is a $(q \times q)$ symmetric matrix with the main diagonal elements being the quadratic coefficients ($\hat{\beta}_{ii}$), while off diagonal elements are one-half the mixed quadratic coefficients ($\beta_{ij;i\neq j}$), Montgomery (2005).

As a result, the estimated response value for the fitted model at the identified stationary point is obtained as:

$$B = \begin{pmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 \\ \hat{\beta}_{21}/2 & \hat{\beta}_{22} \end{pmatrix} = \begin{pmatrix} -0.1376 & 0.4079 \\ 0.4079 & 2.3040 \end{pmatrix} \quad (10)$$

while,

$$b = \begin{pmatrix} -2.2593 \\ -3.6393 \end{pmatrix} \quad (11)$$

The results used above are for a maximum condition, modifying equation (9) for a minimum condition by negation in order to achieve our desired results, Karanjah et al (2015), the stationary point solution is as follows:

$$x_{s3} = -\frac{1}{2}\begin{pmatrix} -0.1376 & 0.4079 \\ 0.4079 & 2.3040 \end{pmatrix}^{-1}\begin{pmatrix} -2.2593 \\ -3.6393 \end{pmatrix}$$

$$= \begin{pmatrix} 3.8486 \\ -1.4711 \end{pmatrix} \quad (12)$$

Decoding back to the natural variables from the coding concept adopted earlier, the stationary point will be found as described below. With respect to $time$, we have;

$$x_1 = \frac{X_1 - 60}{30}$$
$$X_1 = x_1 \times 30 + 60$$
$$= 3.8486 \times 30 + 60$$
$$= 175.4580 \quad (13)$$

The implication is that the time taken to reduce the blood sugar level to within acceptable range is 175.4580 minutes.

In regard to the second variable $concentration$ we have;

$$x_2 = \frac{X_2 - 250}{125}$$
$$X_2 = x_2 \times 125 + 250$$
$$= -1.4711 \times 125 + 250$$
$$= 66.1125 \quad (14)$$

Thus 66.1125 mg/dl of the herbal formula is to be used to regulate the blood sugar level to within the acceptable range.

On the basis of these result, the estimated response value at the stationary point is given as

$$\hat{y}^* = \hat{\beta}_0 + \frac{1}{2}x'_s b, \quad (15)$$

retransforming from adopted transformation

$$\hat{y}^* = \frac{1}{(Y+k)^{0.5}}, \quad (16)$$

where the constant $k = -3$, this yields

$$\hat{Y} = 13.0718 \times \frac{1}{2}(3.8486 \quad -1.4711)\begin{pmatrix} -2.2593 \\ -3.6393 \end{pmatrix}$$
$$= (11.3896) \quad (17)$$

This is the estimated minimum blood sugar level for the given predictor variables.

6. Variance of Function

6.1 Variance of Estimated Response

Generally we use the variance function of the fitted model to evaluate competing designs, the most suitable design is one which has the smallest possible variance. The variance of estimated response ($\hat{y}$) at a point on the sphere of radius ρ where, $y = X'\hat{\beta}$ is

$$Var(\hat{y}) = X' Var(\hat{\beta})X = \sigma^2 f'(x)(X'X)^{-1}f(x) \quad (18)$$

where σ^2 is assumed to be unknown but constant while $x_i, i = 1, 2, \ldots,$ k are taken to be non-stochastic. The prediction variance of the estimated response at a point say $\mathbf{x}$ is given by,

$$Var(\hat{y}(x)) = N^{-1}\mathbf{x}'_t(X'X)^{-1}\mathbf{x}_t\sigma^2, \quad (19)$$

where $\mathbf{x_t}$ is the vector of co-ordinates of a point in the design space expanded to model form.

Using scaled predicted variance (SPV) arrived at by multiplying (18) by the design size and then dividing through by the process variance σ^2, that is

$$V_t = \frac{Var(\hat{y}(x))}{\sigma^2} = \mathbf{x}'_t(X'X)^{-1}\mathbf{x}_t \quad (20)$$

One may undertake comparisons among designs of various sizes while eliminating the need to know the value of σ^2.

Using the observations vector for the stationary point given in equation (12)

$$x'_{s3} = (3.8486 \quad -1.4711).$$

for which we construct the observation vector for the quadratic arrangement as

$$\mathbf{x}_t = (1 \; 3.8486 \; -1.4711 \; 14.8117 \; 2.1641 \; -5.6617.) \quad (21)$$

Using (20), the variance is expressed using

 ISSN: 2454-4248
104 – 109

$$X'X = \begin{pmatrix} 40 & 1 & 11 & 27 & 67 & 2 \\ 1 & 27 & 2 & 1 & 4 & 8 \\ 11 & 2 & 67 & 8 & 89 & 4 \\ 27 & 1 & 8 & 27 & 46 & 2 \\ 67 & 4 & 89 & 46 & 23 & 8 \\ 2 & 8 & 4 & 2 & 8 & 46 \end{pmatrix}.$$

With this matrix, we get the variance V_t as,

$$V_t = (1\ 3.8486\ -1.4711\ 14.8117\ 2.1641\ -5.6617.) \times \begin{pmatrix} 40 & 1 & 11 & 27 & 67 & 2 \\ 1 & 27 & 2 & 1 & 4 & 8 \\ 11 & 2 & 67 & 8 & 89 & 4 \\ 27 & 1 & 8 & 27 & 46 & 2 \\ 67 & 4 & 89 & 46 & 23 & 8 \\ 2 & 8 & 4 & 2 & 8 & 46 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 3.8489 \\ -1.4711 \\ 14.8117 \\ 2.1641 \\ -5.6617 \end{pmatrix}$$

$$= 24.8193 \quad (22)$$

Our task now is to show that V_t is minimum by comparison with variance of any other point on the same response surface. If we select any other point different from the stationary point in the neighbourhood of this stationary point, say $\mathbf{x}_h$ where

$$\mathbf{x}'_h = (1,\ \ 4,\ -2,\ 16,\ \ 4,\ -8). \quad (23)$$

Computing the variance function for this estimated response using this vector $\mathbf{x}_h$ we get,

$$V_{\mathbf{x}_h} = (1\ \ 4\ \ -2\ \ 16\ \ 4\ \ -8) \times \begin{pmatrix} 40 & 1 & 11 & 27 & 67 & 2 \\ 1 & 27 & 2 & 1 & 4 & 8 \\ 11 & 2 & 67 & 8 & 89 & 4 \\ 27 & 1 & 8 & 27 & 46 & 2 \\ 67 & 4 & 89 & 46 & 223 & 8 \\ 2 & 8 & 4 & 2 & 8 & 46 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 4 \\ -2 \\ 16 \\ 4 \\ -8 \end{pmatrix}$$

$$= 30.3998 \quad (24)$$

Comparison of the results of equation (22) and (24), it is clear that the variance of the estimated response arising from the vector in equation (21) generated from the estimated response in (12), is smaller compared to that generated by the vector of (23) on the same response surface. Thus V_t is minimum.

6.2 Variances Function of the Difference between two Estimated Responses

Suppose that $\mathbf{z}$ and $\mathbf{x}$ are two row vectors of formed from a row of X, from two distinct points identified on two estimated response surfaces of different radii. Let,

$$V\ [\hat{y}(z)\ -\hat{y}(x)] \quad (25)$$

denote the variance of the difference between the two estimated responses at the points $\mathbf{z}$ and $\mathbf{x}$. This variance simplifies to

$$V_c = [z - x]'(X'X)^{-1}[z - x]\sigma 2. \quad (26)$$

When the design is rotatable, then $\boldsymbol{X'X}$ has a specific form and the variance of equation (26) is invariant under orthogonal rotations in the predictor space, Herzberg (1967).

Taking the points described in (26) as $z = \mathbf{x}_r$ and $\mathbf{x} = \mathbf{x}_q$ of (21) and (23) respectively, then the variance in (26) is computed for $c = (\mathbf{z} - \mathbf{x})$ as follows;

$$c = (0, -0.1514, 0.5289, -1.1883, -1.8359, 2.3383) \quad (27)$$

$$V_c = (0\ \ -0.1514\ \ 0.5289\ \ -1.1883\ \ -1.8359\ \ 2.3383) \times \begin{pmatrix} 40 & 1 & 11 & 27 & 67 & 2 \\ 1 & 27 & 2 & 1 & 4 & 8 \\ 11 & 2 & 67 & 8 & 89 & 4 \\ 27 & 1 & 8 & 27 & 46 & 2 \\ 67 & 4 & 89 & 46 & 23 & 8 \\ 2 & 8 & 4 & 2 & 8 & 46 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ -0.1514 \\ 0.5289 \\ -1.1883 \\ -1.8359 \\ 2.3383 \end{pmatrix}$$

$$= 0.4840 \quad (28)$$

This is the variance of the difference between two estimated responses for a second order model from the test run involving a herbal formula extract.

7. Conclusion

This research has successfully undertaken the analysis of stationary points by providing a method of finding the stationary point relating to a minimum. The values generated for the two test runs can be used to give a setting for the utilization by herbalists as a starting point in working the concentration of herbal medicine and with which they have the knowledge of the time it takes to reduce the blood sugar level for a diabetic patient in the course of treatment.

This gives the predictive aspect by the herbalists to be able to ascertain the concentration to use as well as the time it takes to achieve desired results.

From the experimentation set up herbal formula extract from a mix of six herbs have been shown in this research to have successfully regulated the blood sugar level in a diabetic to 11.3898 mMol/L. This is possible by effecting a treatment of herbal formula at a concentration of 66.1125 mg/dl and the effectiveness is within 175.4580 minutes upon treatment.

Equation (28) yields the variance function of the difference between two estimated responses for the herbal formula extract. The comparison of the results of the variance function of the difference between two estimated responses shows that the first is smaller than the second. Therefore, we can use the variance function of the difference between two estimated response to be able to map the range over which we can vary the factors of interest, *time* and *concentration* of herbal formula to achieve blood sugar level that is within the acceptable range.

Consider points close together in the factor space, an optimal design with regard to rotatable design in two dimensions from this approach will be chosen on the basis of minimum variance function criterion as emphasized by Herzberg (1967), Box and Draper (1980), and, Huda and Mukerjee (1984).

Research funded by the National Commission for Science, Technology and Innovation (NACOSTI), Kenya.

 ISSN: 2454-4248
104 – 109

8. References

[1] Box, G. E. P. and Draper N. R. **The variance function of the difference between two estimated responses**. *Journal of the Royal Statistical Society, B 42 (1980), 79-82*

[2] Herzberg A. M. (1967). **The behaviour of the variance function of the difference between two estimated responses**. *Journal of the Royal Statistical Society, B 29, 174-179.*

[3] Huda S and Mukerjee R, **Minimizing the maximum variance of the difference between two estimated responses,** *Biometrika, 71, (1984), 381-385.*

[4] Montgomery Douglas, C. (2005). **Design and Analysis of Experiments response surface methods and design**. John Willey and Sons Inc. New Jersey.

[5] Karanjah, A, Njui, F. and Pokhariyal, G. P. (2015). **The variance function of the difference between two estimated responses in regulating Blood sugar Level in a Diabetic patient using herbal Medicine**. *International Journal of Engineering Studies and Technical Approach, Vol. 01, July, Pages 18 - 28. ISSN 2395-0900.*

[6] Karanjah, A, Njui, F. and Pokhariyal, G. P. (2008). **The variance function of the difference between two estimated responses for a fourth order rotatable design in two dimensions**. *Far East Journal of Theoretical Statistics, Vol. 26, Iss. 2, Pages 177 - 191.*

[7] Saeed Ghanbarzadeh, Hadi Valizadeh and Parvin Zakeri-Milani. (2013). **Application of Response Surface Methodology in Development of Sirolimus Liposomes Prepared by Thin Film Hydration Technique**. *Bioimpacts, 3(2), 75-81.*