

A METHOD FOR ROBUST DESIGN OF PRODUCTS OR PROCESSES WITH  
CATEGORICAL RESPONSE

A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES  
OF  
MIDDLE EAST TECHNICAL UNIVERSITY

BY

SERKAN ERDURAL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR  
THE DEGREE OF MASTER OF SCIENCE  
IN  
INDUSTRIAL ENGINEERING

DECEMBER 2006

Approval of the Graduate School of Natural and Applied Science

\_\_\_\_\_  
Prof. Dr. Canan Özgen  
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science

\_\_\_\_\_  
Prof. Dr. Çağlar Güven  
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

\_\_\_\_\_  
Dr. Özlem İlk  
Co-Supervisor

\_\_\_\_\_  
Prof. Dr. Gülser Köksal  
Supervisor

Examining Committee Members

Prof. Dr. Nur Evin Özdemirel (METU, IE) \_\_\_\_\_

Prof. Dr. Gülser Köksal (METU, IE) \_\_\_\_\_

Doç.Dr. Murat Caner Testik (HACETTEPE, IE) \_\_\_\_\_

Dr. Özlem İlk (METU, STAT) \_\_\_\_\_

Dr. Özlem Türker (ÇANKAYA, IE) \_\_\_\_\_

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

**Serkan Erdural**

## **ABSTRACT**

### **A METHOD FOR ROBUST DESIGN OF PRODUCTS OR PROCESSES WITH CATEGORICAL RESPONSE**

Erdural, Serkan

M.S., Department of Industrial Engineering

Supervisor : Prof. Dr. Gülser Köksal

Co-Supervisor: Dr. Özlem İlk

December 2006, 104 pages

In industrial processes decreasing variation is very important while achieving the targets. For manufacturers, finding out optimal settings of product and process parameters that are capable of producing desired results under great conditions is crucial. In most cases, the quality response is measured on a continuous scale. However, in some cases, the desired quality response may be qualitative (categorical). There are many effective methods to design robust products/process through industrial experimentation when the response variable is continuous. But methods proposed so far in the literature for robust design with categorical response variables have various limitations.

This study offers a simple and effective method for the analysis of categorical response data for robust product or process design. This method handles both location and dispersion effects to explore robust settings in an effective way. The method is illustrated on two cases: A foam molding process design and an iron-casting process design.

**Keywords:** Robust Design, Categorical Response, Taguchi Parameter Design Approach, Response Surface Method, Generalized Linear Models

## ÖZ

### KATEGORİK ÇIKTISI OLAN ÜRÜN VEYA SÜREÇLERİN ROBUST TASARIMI İÇİN BİR METOT

Erdural, Serkan

Yüksek Lisans, Endüstri Mühendisliği Bölümü

Tez Yöneticisi : Prof. Dr. Gülser Köksal

Yardımcı Tez Yöneticisi: Dr. Özlem İlk

Aralık 2006, 104 sayfa

Endüstriyel süreçlerde hedeflenene ulaşırken değişkenliği azaltmak oldukça önemlidir. Üreticiler için, istenilen hedefleri mükemmel derecede gerçekleştirecek optimum ürün ve süreç parametreleri belirlemek çok kritiktir. Bir çok durumda, kalite verileri nicel (sürekli) olarak ölçülür. Ancak bazı durumlarda kalite verileri nitel (kategorik) olabilir. Süreç çıktı değişkenlerinin sürekli skalada olduğu durumlarda endüstriyel deneylerdeki robust ürün/süreç tasarımı için birçok etkili

metotlar mevcuttur. Fakat literatürde kategorik çıktı deęişkenleri ile robust tasarım için çeşitli kısıtlamalar vardır.

Bu çalışma, çıktı deęişkeni kategorik olan ürün ya da süreçlerin robust parametre tasarımı için basit ve etkili bir metot sunmaktadır. Bu metot, lokasyon ve varyasyon etkilerini tanımlayarak, robust parametre ayarlarının etkili bir yolla bulunmasını sağlamaktadır. Metodun uygulanması iki farklı vaka analizinde gösterilmiştir: Köpük kalıplama deney tasarımı ve demir döküm süreci tasarımı.

**Anahtar Kelimeler:** Robust Tasarım, Kategorik çıktı, Taguchi Parametre Tasarım Yaklaşımı, Cevap Yüzeyi Metodu, Genelleştirmiş Lineer Modeller

*To my Beril*



## ACKNOWLEDGEMENTS

First of all, I would like to sincerely express my thanks and gratitude to Prof. Dr. Gülser Köksal for her valuable guidance, supervision and endless understanding throughout the research.

I am also grateful to Dr. Özlem İlk for her endless understanding, valuable criticism on categorical data analyses and suggestions in all phases of this study.

Appreciation is also expressed to Erhan İşkol for his support in supplying data and his precious help on real life application.

My special thanks are due to my bosses, Akın Polat, Birol Cömert and Tümer Arıtürk for their understanding and support.

I want to express my sincere gratitude to my parents Müzeyyen and Hüseyin Erdural for their understanding, help, patience and being a constant source of emotional support.

Finally, I want to express my deepest love to my wife Beril Erdural for her heartfelt support, her excellent sense of humor and patience.

## TABLE OF CONTENTS

PLAGIARISM .....	iii
ABSTRACT .....	iv
ÖZ .....	vi
ACKNOWLEDGEMENTS .....	ix
TABLE OF CONTENTS .....	x
LIST OF TABLES .....	xii
LIST OF FIGURES .....	xiii
CHAPTERS	
1. INTRODUCTION .....	1
2. LITERATURE SURVEY AND BACKGROUND .....	3
2.1. Robust (Parameter) Design .....	4
2.1.1. Taguchi's Robust Design Method .....	4
2.1.1.1. History of Taguchi Method .....	4
2.1.1.2. Taguchi's Parameter Design Approach .....	5
2.1.1.3. Quality Loss Function .....	7
2.1.1.4. Design of Experiments Techniques in Taguchi's Approach ...	9
2.1.1.5. Signal-to-Noise Ratios (SNR) .....	10
2.1.1.6. Taguchi's 2-step Optimization Procedure .....	11
2.1.2. Response Surface Method .....	12
2.1.2.1. Dual Response Optimization .....	13
2.1.2.2. Multiple Response Optimization with Desirability Functions ..	15
2.2. The Usage of Categorical Data in Industrial Processes .....	19
2.2.1. Techniques For The Robust Analysis Of Categorical Data .....	19
2.2.1.1. Accumulation Analysis .....	20
2.2.1.2. Scoring Methods .....	23

2.2.1.3. Generalized Linear Models .....	25
2.2.1.4. Bayesian Approach for Categorical Data .....	38
3. THE PROPOSED METHOD.....	40
3.1. Description .....	40
3.2. Illustrative Case Study-1 .....	44
3.3 Illustrative Case Study-2 .....	60
4. DISCUSSION OF NOMINAL DATA CASE .....	75
5. CONCLUSION AND FURTHER STUDIES .....	77
REFERENCES .....	79
APPENDICES	
A. Post-Etch Window Size Example .....	83
B. Minitab Logistic Regression Output for Foam-Molding Experiment .....	86
C. Residual Plots for Foam-Molding Experiment.....	87
D. GLM model for SNR in case study 1.....	90
E. Accumulation Analysis (AA) solution for case study 1 .....	91
F. Variance Estimation for desired Probability in case study 1 .....	94
G. Confidence Interval (CI) Estimation for $\hat{P}(Y=j)$ .....	96
H. Minitab Logistic Regression Output for Case Study 2 .....	98
I. Residual Plots for N-Gas Defect Problem.....	100
J. Classification Table for N-Gas Defect Problem For $c=0.02$ .....	101

## LIST OF TABLES

1. Some commonly used ortogonal arrays .....	9
2. Categorical response variable types.....	19
3. Typical link functions for distributions.....	27
4. Classification table.....	37
5. Optimization models for different problem types .....	44
6. Foam Molding experiment factors and levels .....	45
7. Foam Molding experiment design and frequencies for good (0), OK (I) and poor (II).....	46
8. Parameter estimates from logistic regression.....	47
9. Estimates of event probabilities of categories, expected categories and variance for foam molding experiment data .....	51
10. Estimates of event probabilities of categories, expected categories, variance and SNR values for foam molding experiment data.....	53
11. ANOVA table for SNR values .....	54
12. Comparison of the proposed approach to the accumulation analysis and the bayesian approach.....	60
13. Nitrogen gas defect problem factors and levels .....	61
14. The logistic regression results for N-gas defect data.....	62
15. Prediction error rate values for case study 2 .....	67
16. Comparison of real life results and proposed method .....	71

## LIST OF FIGURES

1. Block diagram representation of a simple parameter design problem.....	7
2. The shape of the desirability functions based on different problem types, s and t values .....	18
3. The Goodness of fit test results for the model.....	49
4. Deviance Residual plots vs. index for category 0 .....	50
5. Deviance Residual plots vs. Factor A for category 0 .....	51
6. The effect plot of Factors A, B and G for SNR.....	54
7. The effect plot of Factors C, E and F for $\hat{E} [Y]$ .....	55
8. The Desirability optimization sheet in Excel Solver .....	58
9. Desirability response surfaces for Factor A and Factor B when C=-1, E=1, F=-1 and G=-1.....	59
10. The Goodness of fit test results for the model 2.....	64
11. Deviance Residual plots vs. index for case 2.....	65
12. Deviance Residual plots vs. Factor A for case 2 .....	66
13. The Desirability optimization sheet in Excel Solver for case 2 .....	70
14. Histogram of percentage response.....	73

## **CHAPTER I**

### **INTRODUCTION**

Variation in the performance of a product degrades the product's quality and most of the time decreasing variation is very important for achieving targets in products and/or processes. Designing products and processes that are insensitive to sources of variation provide a serious competitive advantage to producers. This approach is called robust design (also known as parameter design or optimization). The goal of parameter design studies is to choose the settings of the control factors so that the performance of a system (product or process) is insensitive to variation in uncontrollable "noise" factors. This is done through a designed experiment where the effect of control (design) factors on both location and dispersion are investigated (Phadke, 1989). This is a very cost-effective approach introduced by Genichi Taguchi (Taguchi, 1986) and extended and successfully applied by many practitioners.

A quality response of products or processes is typically measured quantitatively on interval or ratio scale. However, in some cases, the response has to be measured qualitatively (categorical) using a nominal or ordinal scale. Quality characteristics that require expert judgment or comparison to a standard or go/no-go decisions can be an example for qualitative response. Manufacturers generally find it easy and cheap to assess quality on an ordinal scale such as "very good", "good" and "bad", or "pass" and "fail" or sometimes it is impossible to measure in a quantitative scale.

There are many effective methods to obtain robustness by industrial experiments when response variable is continuous. But when the response variable is

categorical, the approaches developed by analysis of both location and dispersion effects are limited and generally conducting the analysis is hard to apply for engineers.

This study offers a relatively simple and effective method for the analysis of categorical response data for robust product or process design. The main basis of method is to find the robust settings of the design variables when the response variable is categorical (binary, ordinal or nominal) by handling both location and dispersion effects to explore robust settings in an effective way. The main advantages of the proposed method compared to the existing methods are: (1) the method clearly identifies location (mean) effects and dispersion (variance) effects and (2) the application of the method is relatively simple and the method can be applied by popular software packages. We also apply the proposed method to two different problems to illustrate the use of the proposed method and discuss the results.

We begin with a review of parameter design and then we consider the response surface method for robust design in Section 2. Also we make a review and provide some background information for existing methods available for robust analyzing of categorical data in Section 2. In Section 3, we introduce the proposed method that for robust design of products or processes with categorical response. In Section 4 we discuss the application of the proposed method for nominal data and finally, the study concludes in Section 5 with a summary of the study, advantages and disadvantages of the proposed method and discussion for the future work.

## **CHAPTER II**

### **LITERATURE SURVEY AND BACKGROUND**

Many approaches have been proposed to find robust parameter levels when the response variables are measured on continuous scale (Robinson et al., 2004, Phadke, 1989). To investigate the relationship between the response and the product or process parameters (control factors) under the influence of uncontrollable (noise) factors, conventional experimental design techniques are used. The collected data through experimentation are typically used to estimate the quality characteristic's (or the response variable's) mean and variance for given parameter levels. Then, a search is performed to find the parameter levels that yield the minimum variance and bring the mean to the target.

Taguchi's approach (Taguchi and Wu, 1980) gives fundamental tools for practitioners by suggesting them to choose the levels that maximize the signal-to-noise ratios (SNR). By utilizing SNR, Taguchi proposed a two step procedure that utilizes SNR to minimization of expected quality loss to the customers. But many authors criticized Taguchi's approach because of inefficiency of the method, the requirement for large experimental runs and inadequacy of SNR (Nair et al, 1992) and suggested some extensions on the method or the use of separate models for the system mean and variance for a better understanding of the process. By following these suggestions, Response Surface Methodology (RSM) becomes an alternative approach to the robust design problem. (Robinson et al., 2004)



## **2.1 Robust (Parameter) Design**

Robust design is an engineering methodology for improving productivity during research and development so that high-quality products can be produced quickly and at low cost (Phadke, 1989). The idea behind robust design is to improve the quality of a product by minimizing the effects of variation without eliminating the causes (since they are too difficult or too expensive to control). This method is an off-line quality control method that is applied at both the product and process design stage to improve product quality by making products insensitive to environmental conditions and component variations. Briefly robust design is a design that has minimum sensitivity to variations in uncontrollable factors.

There are two basic designing approaches in a robust design experiment: the product array and the combined array. The product array separately fractionates the control factors and the noise factors and forms the cross product by running the noise array with every control factor combination in the control array. The combined array fractionates all of the factors in a single factorial design.

There are many parameter design methods but mainly Taguchi's techniques and Response Surface Methodology (RSM) is represented in this study.

### **2.1.1. Taguchi's Robust Design Method**

#### **2.1.1.1. History of Taguchi Method**

Since 1960, Taguchi methods have been used for improving the quality of Japanese products with great success. During the 1980's, many companies finally realized that the old methods for quality assurance were not competitive with the Japanese methods. The old methods for quality assurance relied heavily upon inspecting products as they produced the production line and rejecting those products

that did not fall within a certain acceptance range. However, according to Taguchi, a product can not be improved by large amount of inspection; quality must be designed into a product from the start. Nowadays, Taguchi's parameter design approaches are widely used to improve product quality and design robustness and most of the companies are adapting this philosophy to their products and processes.

### **2.1.1.2.Taguchi's Parameter Design Approach**

During the design stage of the products and processes, off-line quality control techniques are applied to reduce the variation by product and process engineers. After that, in production stage, process and production engineers apply on-line quality control techniques to make adjustments on the average response. Within the Taguchi philosophy both quality improvement methods are considered; however, the main goal is building quality into the product during the design stage (i.e., off-line).

To achieve desirable product quality in design stage, a three-stage process is suggested by Taguchi (Kackar 1985):

- Ø System design
- Ø Parameter design
- Ø Tolerance design.

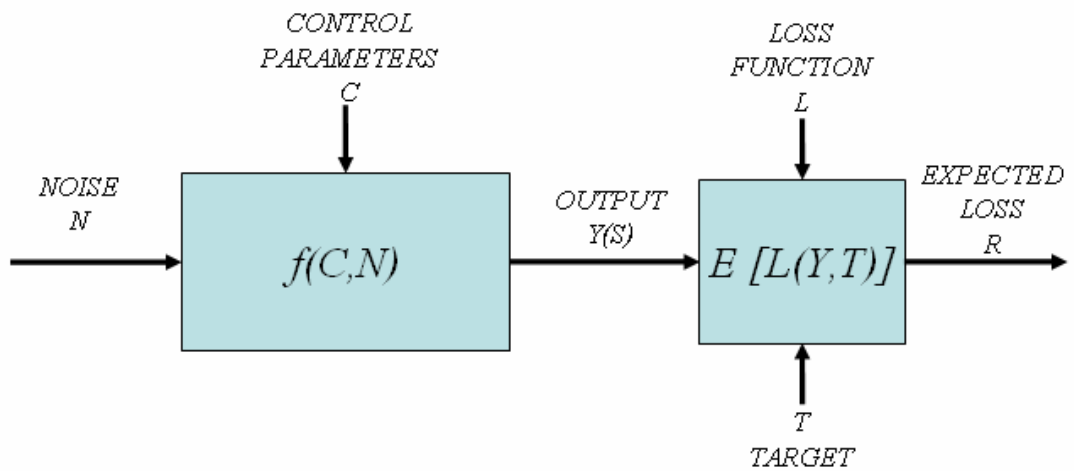
*System design* is the process of applying scientific and engineering knowledge to produce a basic functional prototype design, which defines the initial settings of the product or process design characteristics.

*Tolerance design*, which is the final part of design stage, is a method for determining tolerances that maximizes the product quality and minimizes the sum of product manufacturing and lifetime costs.

*Parameter design*, which is the main part of Taguchi's approach is an investigation conducted to identify settings that minimize the performance variation. A product or a process can perform its intended function at many settings of its design characteristics. However, variation in the performance characteristics may change with different settings. This variation increase both product manufacturing and lifetime costs. Therefore, identifying optimal parameter settings that will reduce the variation is important and parameter design aims to reduce the cost of variation by optimizing the process parameter settings.

In parameter design the system variables are experimentally analyzed to determine how the product or process reacts to uncontrollable "noise" in the system. Parameter design is related to finding the appropriate design factor levels to make the system less sensitive to variations in uncontrollable noise factors, i.e., to make the system robust. In this way the product performs better, reducing the loss to the customer.

In parameter design, there are two types of factors that affect a product's functional characteristic: *control factors* and *noise factors*. Control factors are those factors which can easily be controlled such as material choice, cycle time, or mold temperature in an injection molding process. Noise factors are factors that are difficult or impossible or too expensive to control. There are three types of noise factors: outer noise, inner noise, and between product noise. Noise factors are primarily reasons that cause a product's performance to deviate from its target value. Hence, parameter design seeks to identify settings of the control factors which make the product insensitive to variations in the noise factors, i.e., make the product more robust. Figure 1 is a block diagram representation of a simple parameter design problem. For a given setting of control parameters  $C$ , noise  $N$  produces a characteristic output  $Y$ . The output is determined by the transfer function,  $f(C,N)$ .



**Figure 1.** A block diagram representation of a simple parameter design problem (adapted from Leon et al., 1987)

The performance of the output is measured by determining how it deviates from the target. In parameter design this deviation is expressed by total expected loss which is calculated with the quality loss function as it can be seen in Figure 1.

### 2.1.1.3. Quality Loss Function

The quality loss function is a continuous function that is defined in terms of the deviation of a design parameter from an ideal or target value.

Taguchi introduced the loss function first that was expressed in terms of the quadratic relationship:

$$L = k (y - T)^2 \quad [2.1]$$

where  $y$  is the performance output,  $L$  is the loss occurred from a particular parameter  $y$ ,  $T$  is the nominal (target) value of the parameter specification,  $k$  is a constant that depends on the cost at the specification limits. The constant  $k$  can be determined by dividing the cost of failure (defect), by the square of the lower or higher tolerance values. This function penalizes the deviation of a parameter from the specification value that contributes to decreasing the performance of the product, resulting in a loss to the customer. The key point here is that a product engineer has a good understanding of what the nominal size of the specification is. The usual lower and upper limits for the tolerance of a given design parameter are changed to a continuous function that presents any parameter value other than the nominal as a loss.

The loss function given in Equation [2.1] is referred to as “nominal is best,” but there are also expressions for cases when higher or lower values of parameters are better (Phadke, 1989).

If a large number of parts are considered, say  $N$ , the average loss per part is equal to the summation of the losses given by Equation [2.1] for each part, divided by the total  $N$ . The average quality loss results from deviation around the average value of  $y$  from the target and the mean square deviation of  $y$  around its own mean. The average quality loss can be expressed as:

$$L = k[\sigma^2 + (\mu - T)^2] \quad [2.2]$$

where  $\mu$  is the average value of  $y$  for the set of parts, and  $\sigma^2$  is the variance around the average.

To minimize loss, the traditional approach is to monitor the process variables during production and adjust the process to reduce manufacturing imperfections so that response parameters fall within the specified tolerances. This method adds cost to the manufacturing process and does not improve the quality of the product. Using

Taguchi's approach the average response has to be adjusted, and the variance must be reduced in order to minimize loss.

#### 2.1.1.4. Design of Experiments Techniques in Taguchi's Approach

Design of experiments techniques, specifically Orthogonal Arrays (OAs), are employed in Taguchi's approach to systematically vary and test the different levels of each of the control factors. Commonly used OAs include the L4, L9, L12, L18, and L27, several of which are listed in Table 1. A complete listing of OAs can be found in textbooks such as Phadke (1989). The columns in the OA indicate the factor and its corresponding levels, and each row in the OA constitutes an experimental run which is performed at the given factor settings. Selecting the number of factors and their levels is the main part in planning robust design experiments.

**Table 1.** Some commonly used orthogonal arrays (adapted from Phadke 1989)

Factors			
Run	A	B	C
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

(a)  $L_4(2^3)$  array

Factors							
Run	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

(b)  $L_8(2^7)$  array

Factors				
Run	A	B	C	D
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

(c)  $L_9(3^4)$  array

To implement robust design, Taguchi techniques use an “inner array” and “outer array” approach. The “inner array” consists of the OA that contains the control factor settings; the “outer array” consists of the OA that contains the noise

factors and their settings which are under investigation. The combination of the “inner array” and “outer array” constitutes the “product array” or “complete parameter design layout.” The product array is used to systematically test various combinations of the control factor settings over all combinations of noise factors after which the mean response and standard deviation may be approximated for each run using the following equations. (Shoemaker et al, 1991)

Mean response:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad [2.3]$$

Standard deviation:

$$S = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}} \quad [2.4]$$

### 2.1.1.5. Signal-to-Noise Ratios (SNR)

The mean and the variance are combined into a single performance measure known as the signal-to-noise ratio (SNR). The preferred parameter settings are then determined through analysis of the SNR where the factor levels that maximize the appropriate SNR are chosen. Taguchi classifies parameter-design problems into different categories depending on the goal of the researcher and a different SNR is defined for each category. The categories are the smaller the better, the larger the better, or nominal the best for the quality characteristic  $y$  (Phadke, 1989).

- Smaller the better (for making the system response as small as possible):

$$SNR = -10 \log\left(\frac{1}{n} \sum_{i=1}^n y_i^2\right) \quad [2.5]$$

- Larger the better (for making the system response as large as possible):

$$SNR = -10 \log \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{y_i^2} \right) \quad [2.6]$$

- Nominal the best (for reducing variability around a target):

When variance is linked to mean:

$$\begin{aligned} SNR &= 10 \log(\bar{y}^2 / s^2) \quad \text{or} \\ SNR &= 10 \log \left[ (\bar{y}^2 / s^2) - (1/n) \right] \end{aligned} \quad [2.7]$$

When variance is not linked to mean:

$$SNR = -10 \log s^2 \quad [2.8]$$

These SNRs are derived from the quadratic loss functions and are expressed in a decibel (dB) scale. Once all of the SNRs have been computed for each run of an experiment, graphical approaches, such as main effect plots, and analysis of variance (ANOVA) are used to analyze the data. In the graphical approach, the SNR and average responses are plotted for each factor against each of its levels. The standard ANOVA is used to identify the control factors that affect the SNR. Those factors that do not affect the SNR are referred to as ‘adjustment factors’ and are used to adjust the average performance on target.

#### **2.1.1.6. Taguchi’s 2-step Optimization Procedure**

To accomplish the objective of minimal expected squared-error loss Taguchi proposes the following two-step optimization procedure:



1. Calculate and model the SNRs and find the non-adjustment factor settings which maximize the SNR,
2. Shift mean response to the target by changing the adjustment factor(s).

Using this information, the control factors can also be grouped as follows.

1. Factors that affect both the variation and the average performance of the product.
2. Factors that affect the variation only.
3. Factors that affect the average only.
4. Factors that do not affect either the variance or the average.

Factors in the first and second groups can be utilized to reduce the variations in the system, making it more robust. Factors in the third group are then used to adjust the average to the target value. Lastly, factors in the fourth group are set to the most economical level. Finally, confirmation tests should be run at the “optimal” product settings to verify that the predicted performance is actually realized.

### **2.1.2. Response Surface Method**

Response surface methodology (RSM) is a collection of mathematical and statistical techniques that are useful for modeling and analysis in applications where one or more response variables are influenced by a set of quantitative experimental variables or factors and the objective is to optimize the response variable(s). (Montgomery, 2001)

The main aim of RSM is modeling a relationship between the quantitative factors and the response variable(s). RSM is often employed to find factor settings (operating conditions) that produce the "best" response variable(s) and that satisfy

process and operating specifications. RSM is also useful for identifying new operating conditions that produce desired improvement in product quality over the quality achieved by current conditions.

In this study our problem is to find robust settings by obtaining the desired mean response with minimum variance. Since we will deal with two response surfaces, in the next sections this RSM problem will be called dual response optimization problem.

### **2.1.2.1. Dual Response Optimization**

Dual response optimization is a special case of multi-objective RSM especially used for modeling two objectives: mean and variance. The dual response optimization provides a reasonable basis to achieve the basic goals of Taguchi's robust design philosophy without resorting to combining information about both mean and variability into a SNR. (Köksoy and Muluk, 2004).

The dual response surface (DRS) approach, that became popular in 1990s, suggests that the process characteristic and its process variability form a DRS, and two separate models are established for the response and its variance. This approach allows the use of all regression tools to define the two response surfaces. In practice, the two separate models give the analyst a more scientific understanding of the total process, and thus allow them to see what levels of the control factors can lead to satisfactory values of the response as well as the variance (Vining and Myers, 1990).

Like other optimization work in RSM, the dual response optimization problem also consists of the following three stages. The first stage is to build an optimal experiment so that the information among the responses and the control factors can be obtained efficiently. The second stage is to build two models based on

the data from the experiment, one for the process characteristic and the other for the process variance. The last stage is to search for the optimal operating condition throughout the region of interest ( $x \in R$ ) under certain optimization constraints based on the process or product. A typical DRS optimization model is represented in [2.9].

$$\begin{aligned} & \text{Min} \{ \hat{y}_m, \hat{y}_s \} \\ & \text{where} \\ & \hat{y}_m = b_0 + x'b + x'Bx \quad [2.9] \\ & \hat{y}_s = c_0 + x'c + x'Cx \\ & x \in R \end{aligned}$$

where  $\mathbf{x}$  is a  $k \times 1$  vector of control or design variables,  $\mathbf{B}$  and  $\mathbf{C}$  are  $k \times k$  matrices containing the estimated coefficients of the second-order terms of each response model,  $\mathbf{b}$  and  $\mathbf{c}$  are  $k \times 1$  vectors containing the estimated coefficients of the first-order terms of each response model,  $b_0$  and  $c_0$  are intercepts,  $\mathbf{R}$  defines the experimental region of interest ( $\mathbf{x} \in \mathbf{R}$ ).

There are many optimization algorithms employed for the multi response problem such as Quasi-Newton Method, General Reduced Gradient (GRG), Genetic Algorithms and NIMBUS (Köksoy and Muluk, 2004). Especially GRG algorithm is a widely used local optimization algorithm in DRS problems (Del Castillo and Montgomery, 1993). One of the advantages of GRG is that it can be applied by the widely used software packages such as Microsoft Excel and Minitab. The main disadvantage of GRG method is that the results sometimes depend on the initial settings and these initial values should be chosen carefully for finding the global optimum values.

For the multi-response model solution with GRG algorithm the usage of desirability functions are recommended (Del Castillo and Montgomery, 1993 and Tang and Xu, 2002). Usage of desirability functions is very common and most of the software packages such as Minitab, JMP and Statistica use these functions for multi-

response optimization. In the next section the detailed information is given about these functions.

### 2.1.2.2. Multiple Response Optimization with Desirability Functions

The desirability function approach is one of the most widely used methods in industry for the optimization of multiple response processes. It is based on the idea that the "quality" of a product or process that has multiple quality characteristics, with one of them outside of some "desired" limits, is completely unacceptable. The method finds operating conditions that provide the "most desirable" response values (Derringer and Suich, 1980).

The desirability approach is a popular method that assigns a "score" to a set of responses and chooses factor settings that maximize that score. For each response  $Y_i(x)$ , a desirability function  $d_i(Y_i)$  assigns numbers between 0 and 1 to the possible values of  $Y_i$ , with  $d_i(Y_i) = 0$  representing a completely undesirable value of  $Y_i$  and  $d_i(Y_i) = 1$  representing a completely desirable or ideal response value. The individual desirabilities are then combined using the geometric mean, which gives the *overall desirability D*:

$$D = (d_1(Y_1) * d_2(Y_2) * \dots * d_k(Y_k))^{1/k} \quad [2.10]$$

with  $k$  denoting the number of responses. Notice that if any response  $Y_i$  is completely undesirable ( $d_i(Y_i) = 0$ ), then the overall desirability is zero.

Depending on whether a particular response  $Y_i$  is to be maximized, minimized, or assigned a target value, different desirability functions  $d_i(Y_i)$  can be used. A useful class of desirability functions was proposed by Derringer and Suich (1980). Let  $L_i$ ,  $U_i$  and  $T_i$  be the lower, upper, and target values, respectively, that are desired for response  $Y_i$ , with  $L_i \leq T_i \leq U_i$ .

In the next sections, different types of desirability functions are provided for the particular cases of the response such as nominal the best, smaller the better and larger the better.

### Desirability function for "target is best"

If a response is of the "target is best" kind, then its individual desirability function is

$$d_i(\hat{Y}_i) = \begin{cases} 0 & \text{if } \hat{Y}_i(x) < L_i \\ \left( \frac{\hat{Y}_i(x) - L_i}{T_i - L_i} \right)^s & \text{if } L_i \leq \hat{Y}_i(x) \leq T_i \\ \left( \frac{\hat{Y}_i(x) - U_i}{T_i - U_i} \right)^t & \text{if } T_i \leq \hat{Y}_i(x) \leq U_i \\ 0 & \text{if } \hat{Y}_i(x) > U_i \end{cases} \quad [2.11]$$

with the exponents  $s$  and  $t$  determining how important it is to hit the target value. For  $s = t = 1$ , the desirability function increases linearly towards  $T_i$ ; for  $s < 1$ ,  $t < 1$ , the function is convex, and for  $s > 1$ ,  $t > 1$ , the function is concave. Figure 2 shows the shape of desirability function based on different problem types,  $s$  and  $t$  values.

### Desirability function for maximizing a response

If a response is to be maximized instead, the individual desirability is defined as

$$d_i(\hat{Y}_i) = \begin{cases} 0 & \text{if } \hat{Y}_i(x) < L_i \\ \left( \frac{\hat{Y}_i(x) - L_i}{T_i - L_i} \right)^s & \text{if } L_i \leq \hat{Y}_i(x) \leq T_i \\ 1.0 & \text{if } \hat{Y}_i(x) > T_i \end{cases} \quad [2.12]$$

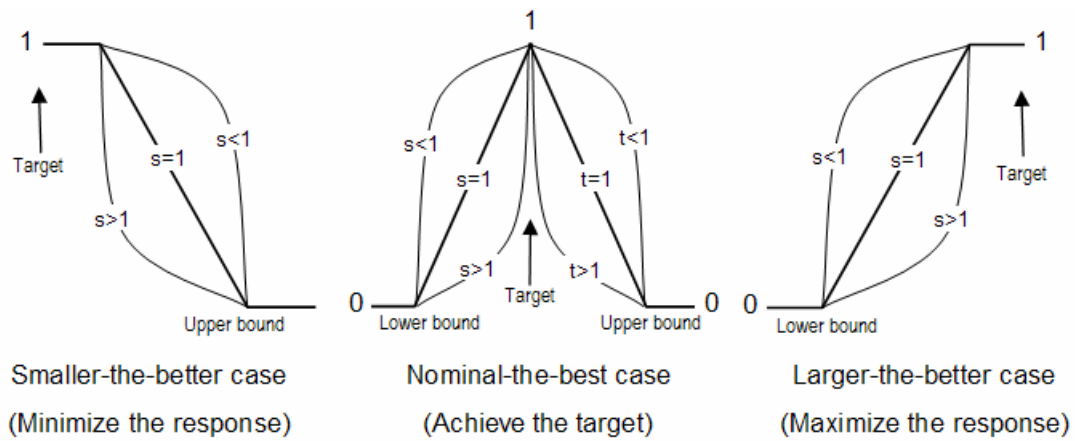
with  $T_i$  in this case interpreted as a large enough value for the response

### **Desirability function for minimizing a response**

Finally, if we want to minimize a response, we could use

$$d_i(\hat{Y}_i) = \begin{cases} 1.0 & \text{if } \hat{Y}_i(x) < T_i \\ \left( \frac{\hat{Y}_i(x) - U_i}{T_i - U_i} \right)^s & \text{if } T_i \leq \hat{Y}_i(x) \leq U_i \\ 0 & \text{if } \hat{Y}_i(x) > U_i \end{cases} \quad [2.13]$$

with  $T_i$  denoting a small enough value for the response.



**Figure 2.** The shape of the desirability functions based on different problem types,  $s$  and  $t$  values

### Desirability approach steps

The desirability approach consists of the following steps:

1. Conduct experiments and fit response models for all  $k$  responses;
2. Define individual desirability functions for each response;
3. *Maximize* the overall desirability  $D$  with respect to the controllable factors.

Tang and Xu (2002) defines some approaches to optimize a dual response system by using desirability functions in Excel solver and shows practical solutions on some examples.

## 2.2. The Usage of Categorical Data in Industrial Processes

Industrial and other experimental results often appear as categorical data. Generally it is easier and cheaper to assess quality on categorical measure instead of measuring quantitatively. Moreover, sometimes it is impossible to measure in quantitative scale because of practical and technical difficulties. Therefore qualitative measurements are sometimes more critical in some processes.

There are three types of categorical responses generally observed in industrial processes: binary, ordinal (ordered) and nominal. Table 2 gives some definitions and examples for these types.

Standard textbooks on experimental design and analysis do not give special attention to these problems. In next sections, some of the methods for modeling ordinal categorical data are briefly presented and the main characteristics and pitfalls of these models are discussed.

**Table 2.** Categorical response variable types

Variable type	Number of categories	Characteristics	Examples
Binary	2	two levels	success, failure yes, no
Ordinal	3 or more	natural ordering of the levels	very bad, bad, good high, medium, low
Nominal	3 or more	no natural ordering of the levels	blue, black, red, yellow Fiat, Toyota, Opel



### 2.2.1. Techniques Available For The Robust Analysis Of Categorical Data

Techniques available for the robust design analysis of categorical data include Accumulation Analysis (AA), scoring methods, Generalized Linear models and Bayesian Analysis (BA).

#### 2.2.1.1. Accumulation Analysis

Accumulation Analysis (AA) was introduced by Taguchi in 1974 for analyzing ordered categorical data from industrial experiments. It is an ANOVA-like approach using cumulative frequencies of the response categories and the name accumulation analysis comes from accumulating these frequencies to the last category. In spite of criticism presented in the next section, AA is still referred to as one of attribute data analysis methods in some books (Ross, 1996) and software packages (Statistica, 2005).

Nair (1986) defines the method on an example. Consider a one-factor experiment with factor  $A$  with  $I$  levels and assume that there are an equal number,  $n$ , of observations at each level. The observations are classified into one of  $K$  ordered categories, and  $Y_{ik}$  denotes the observed frequency in category  $k$  at level  $i$  ( $k = 1, \dots, K$  and  $i = 1, \dots, I$ ).

Lets denote the cumulative frequencies by

$$C_{ik} = \sum_{j=1}^k Y_{ij} \quad [2.14]$$

and their averages across levels by

$$C_k = \sum_{i=1}^I \frac{C_{ik}}{I} \quad [2.15]$$

Then the sum of squares (SS) for factor A is given by

$$SS_A = n \sum_{k=1}^{K-1} \sum_{i=1}^I \frac{(C_{ik} - C_k)^2}{[C_k(n - C_k)]} \quad [2.16]$$

This is obtained as follows. From the cumulative frequencies in the  $k$ th column, we get the sum of squares for factor A as

$$SS_{A,k} = \sum_{i=1}^I (C_{ik} - C_k)^2 \quad [2.17]$$

Since these have different expectations under the null hypothesis that factor A has no effect, they are standardized before being combined to get the single sum of squares in [2.16]. AA suggests using  $(I-1)(K-1)$  degrees of freedom (df) for  $SS_A$ .

Since the cumulative frequency  $C_{ik}$  in the  $(i, k)$ th cell is made up of  $C_{ik}$  ones and  $(n - C_{ik})$  zeros, the within sum of square is proportional to  $[C_{ik}(n - C_{ik})]$ . By combining these, the sum of squares for error term will be:

$$SS_e = n \sum_{k=1}^{K-1} \sum_{i=1}^I \frac{[C_{ik}(n - C_{ik})]}{[C_k(n - C_k)]} \quad [2.18]$$

The expectation of  $SS_e$  is  $[n(n-1)I^2(K-1)/(In-1)] \approx I(n-1)(K-1)$ . Because of this, Taguchi suggested taking  $I(n-1)(K-1)$  as df for  $SS_e$  and using the statistic

$$F_A = \frac{MS_A}{MS_e} \quad [2.19]$$

to test for the effect of factor  $A$  ( $MS$  denotes mean square). The preceding definitions extend in a straightforward way to multifactor situations.

AA's simplicity and similarity to ANOVA is appealing. Unfortunately, it does not possess ANOVA's property of independent sums of squares. Noticing that  $SSe = \text{constant} - SS_A$ , Nair (1986) and Box and Jones (1986) pointed out the undesirable property that  $SSe$  depends on the effect of factor  $A$ . Consequently, they proposed using only the numerator sum of squares  $SS_A$  from [2.16].

Furthermore, Nair (1986) noted that in the multifactor setting, the modified AA statistic eliminates one way in which the AA statistic for a factor depends on the other factors. For example, this can be easily seen in the two-factor main effects setting, where the distribution of the original AA statistic for factor  $A$  depends on factor  $B$  since  $SSe = \text{constant} - SS_A - SS_B$ .

Based on the heavy criticism by Nair (1986), Hamada and Wu (1986, 1990) and Box and Jones (1986), the main pitfalls of the method are:

- § The cumulative frequencies do not satisfy the necessary model assumptions
- § Factor effects become dependent
- § Sometimes it detects spurious factor effects
- § It detects a mixture of location and dispersion effects

In view of these results and the findings of earlier research, it is concluded that accumulation analysis is inefficient, as well as unnecessarily complicated consequently should not be taught or recommended.

Accumulation Analysis and Scoring methods are applied on Post-Etch Window size experiment (Nair, 86), the results of these two method are presented in Appendix-1.

### 2.2.1.2.Scoring Methods

Another method for analyzing ordered categories is assigning scores on the ordered categories and performing ANOVA on these scores. Although this so-called scored ANOVA approach is simple, shortcomings of this method are: (1) the scored categories are not continuous and the normal distribution assumptions may be unreasonable, (2) the results depend completely on the scores assigned (See Nair (1986) and Hamada and Wu (1990) for further discussion).

Nair (1986) showed that in the single factor setting the AA statistic can be expressed as a weighted sum of score statistics. Since the scores are linear, quadratic, etc., the first two components have been interpreted as tests for location and dispersion effects. Then he proposed using these two components and simpler alternatives to test for location and dispersion effects. Nair suggested using simpler alternatives  $SS(l)$  and  $SS(d)$  with data-based scores  $l$  and  $d$ .

The first set of scores that Nair proposed for determining location effects ( $l$ ) for observations fall into category  $k$  can be computed as follows:

$$l_k = \hat{\tau}_k / \left[ \sum_{j=1}^k q_j \bar{\tau}_j^2 \right]^{1/2} \quad [2.20]$$

where  $q_j$  is the overall proportion of observations in  $k^{\text{th}}$  category and  $\tau_k$  is

$$t_k = \sum_{j=1}^{k-1} q_j + q_k / 2 \quad [2.21]$$

which is proportional to the mid-rank category k.

Similarly, the second set of scores for determining location effects ( $l$ ), for observations that fall into category k, can be computed as follows:

$$d_k = e_k / \left[ \sum_{j=1}^K q_j e_j^2 \right]^{1/2} \quad [2.22]$$

where  $q_j$  is again the overall proportion of observations in  $k^{\text{th}}$  category and  $e_k$  is

$$e_k = l_k (l_k - \sum_{j=1}^K q_j l_j^3) - 1 \quad [2.23]$$

Based on location ( $l$ ) and dispersion ( $d$ ) scores, Nair proposed to calculate new type of sum of squares for each factor in the model. These sum of squares can be calculated by the formulas below:

$$SS(l) = \sum_{i=1}^I (l^T (y_i - y)) ^2 / n \quad [2.24]$$

$$SS(d) = \sum_{i=1}^I (d^T (y_i - y)) ^2 / n \quad [2.25]$$

where  $SS(l)$  is the sum of squares that will be used to detect location for every factor and  $SS(d)$  will be used for dispersion.

The simple alternative of scoring the categories converts the categorical data into numerical data so that usual methods like ANOVA can be employed. The method of scoring the categories is simple and particularly effective if the scores are reasonably chosen. So that if these scores are not reasonable measures for the categories, the results can be misleading. This highlights the importance of knowing the relative weights of the categories in a practical situation (Hamada and Wu, 1990). As Nair (1986) mentioned, the main attraction of this technique is its simplicity, which suggests that other approaches could be more accurate.

Although the above two methods, AA and Scoring Methods, can achieve parameter optimization, they can only determine the optimum categorical level setting. They cannot determine the optimum continuous setting values of control factors.

### **2.2.1.3. Generalized Linear Models**

The Generalized Linear Models (GLZ), which is an extension of the General Linear Models (GLM), can be used for response variables that follow any probability distribution in the exponential family of distributions such as the Normal, Binomial, Poisson, Multinomial, and Gamma. Hypothesis tests applied to the GLZ's don't require normality of the residuals and they also don't require homogeneity of variances like ordinary regression does. Therefore, GLZ's can be used when response variables follow distributions other than the Normal distribution, and when variances are not constant.

In GLZ, the dependent variable values are predicted from a linear combination of predictor variables, which are "connected" to the dependent variable via various link functions. The variety of this link functions provide GLZs a main advantage that the distribution of the dependent or response variable can be non-

normal, and does not have to be continuous, i.e., it can be binomial, ordinal multinomial or multinomial. In fact, the GLM for a single dependent variable can be considered a special case of the GLZ, because in GLM dependent variable values are expected to be normally distributed and the link function is a simple identity function.

To illustrate more clearly, in GLM a response variable  $Y$  is linearly associated with values on the  $X$  variables by

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + e$$

where  $\varepsilon$  stands for the error that cannot be accounted for by the predictors. Note that,  $\varepsilon$  is assumed to be from normal distribution with zero mean and constant variance. On the other hand, the relationship in GLZ is assumed to be

$$Y = g(b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k) + e$$

where  $\varepsilon$  is the error term and  $g(X)$  is a function whose inverse function of  $g(X)$ ,  $g^{-1}(X)$ , is called the link function. Hence,

$$g^{-1}(\hat{Y}) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

where  $\hat{Y}$  is the expected value of  $Y$ . If link function is identity function, the GLZ becomes GLM. As mentioned before, the superiority of GLZ comes from this link function. The variety of link functions in GLZ modeling allows both linear and non-linear relationships between response variable and independent variables. Various link functions can be chosen (McCullagh and Nelder, 1989), depending on the assumed distribution of the response variable values which are presented in Table 3.

**Table 3.** Typical link functions for distributions (adapted from McCullagh and Nelder, 1989)

Probability Distribution	Link Function ( $g^{-1}(z)$ )
Normal	Identity ( $g^{-1}(z) = z$ )
Binomial, Ordinal Multinomial	Logit ( $g^{-1}(z) = \log(z/(1-z))$ ), Probit ( $g^{-1}(z) = F^{-1}(z)$ ), Complementary log-log ( $g^{-1}(z) = \log(-\log(1-z))$ ) where $F^{-1}$ is inverse normal distribution
Multinomial	Generalized logit link ( $g^{-1}(z_1/z_2, \dots, z_c) = \log(z_1/(1-z_1-\dots-z_c))$ ) for $c+1$ categories
Poisson	Log ( $g^{-1}(z) = \log(z)$ )
Gamma	Reciprocal ( $g^{-1}(z) = 1/z$ )

When the dependent response variable is categorical and the independent variables are discrete or continuous, the link function can transform the probabilistic scale (0,1) to the continuous infinity scale  $(-\infty, +\infty)$ . That is why GLZ are used for categorical response.

### Estimation and Testing in GLZ

The values of the parameters ( $b_j$ ) in the GLZ can be obtained by maximum likelihood estimation (MLE) through iterative computational procedures. Mainly, iterative methods like the Newton-Raphson and Fisher-Scoring method are used for MLE in GLZ (Agresti, 1996).

Tests for the significance of the effects in the model can be performed via the Wald Statistics, the likelihood ratio (LR), or score statistics. Detailed descriptions of these are given in logistic regression part.



### 2.2.1.3.1. Logistic Regression

Logistic regression, which is a special and most commonly used type of GLZ, is especially used for analyzing categorical response variables such as binary, ordinal and nominal response. It uses a link function (logit, probit or complementary log-log) and estimation of the parameters is carried out via MLE as used in GLZ (McCullagh, 1980).

The dependent variable in logistic regression is usually binary, that is, the dependent variable can take the value of 1 with a probability of success  $\theta$ , or the value of 0 with probability of failure  $1-\theta$ . This type of variable is assumed to have a Bernoulli distribution, which is a special case of Binomial distribution. Applications of logistic regression have also been extended to cases where the dependent variable takes more than two cases (ordinal or nominal). In this case dependent variable is assumed to have a multinomial distribution.

Similar to GLZ, the independent or predictor variables in logistic regression can take discrete or continuous form. That is, logistic regression makes no assumption about the distribution of the independent variables. They do not have to be normally distributed, linearly related or have equal variance within each group.

The relationship between the predictor and response variables is not a linear function in logistic regression. For binary response it takes the following form:

$$E(Y) = P(Y = 1) = p = \frac{e^{(b_0 + b_1x_1 + b_2x_2 + \dots + b_jx_j)}}{1 + e^{(b_0 + b_1x_1 + b_2x_2 + \dots + b_jx_j)}} \quad [2.26]$$

where  $\beta_0$  is the constant of the equation, and  $\beta_k$  is the coefficient of the predictor variable  $x_k$  and  $\pi$  is the probability of the event (category 1).

An alternative form of the logistic regression equation is:

$$\log it[p] = \log \left[ \frac{p}{1-p} \right] = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_j x_j \quad [2.27]$$

The expression “ $\pi / (1 - \pi)$ ” is called odds. Fixing all other factors, to see the effect of one unit increase in  $x_k$ , consider  $x_k^* = x_k + 1$ . Odds ratio ( $\theta$ ), which is the ratio of two odds under two different conditions, is useful in interpreting the relationship between a predictor and response. Let  $\pi^*$  be the probability of event when  $x_k^*$  is observed, while  $\pi$  be the probability of event under the condition  $x_k$ . The odds ratio is then,

$$q = \frac{\frac{p^*}{1-p^*}}{\frac{p}{1-p}} = \frac{\frac{e^{b_0 + b_1 x_1 + \dots + b_k x_k^* + \dots + b_j x_j}}{1 + e^{b_0 + b_1 x_1 + \dots + b_k x_k^* + \dots + b_j x_j}}}{\frac{e^{b_0 + b_1 x_1 + \dots + b_k x_k + \dots + b_j x_j}}{1 + e^{b_0 + b_1 x_1 + \dots + b_k x_k + \dots + b_j x_j}}} = e^{b_k (x_k^* - x_k)} = e^{b_k (x_k + 1 - x_k)} = e^{b_k} \quad [2.28]$$

Hence,  $\beta_k$  is the log-odds ratio of success when  $x_k$  (reference level) is increased by 1 unit while keeping other factors ( $x_i, \forall i \neq k$ ) constant.

The odds ratio can be any nonnegative number. The odds ratio = 1 serves as the baseline for comparison.  $\theta = 1$  indicates there is no association between the response and the predictor. If  $\theta > 1$ , the odds of success are higher for the reference level of the factor (or for higher levels of a continuous predictor). If  $\theta < 1$ , the odds of success are less for the reference level of the factor (or for higher levels of a continuous predictor). Values farther from 1 represent stronger degrees of association.

There are two main uses of logistic regression. The first is the prediction of group membership. Since logistic regression calculates the probability of success over the probability of failure, the results of the analysis are in the form of odds ratio. For example, logistic regression is often used in industrial studies where the result of the analysis is the probability of defects after controlling for other associated risks. Logistic regression also provides knowledge of the relationships and strengths among the variables (e.g., an increase in density also increases the probability of seeing a defect ).

### **Logistic Regression Model Diagnostics**

#### Significance of Coefficients:

*Wald Test* is used to determine the significance of terms in the logistic regression models. It is used to test the statistical significance of each coefficient ( $b_j$ ) in the model ( $j= 0, 1, \dots, J$ ). It test the hypothesis for this purpose below.

$$H_0: \beta_j = 0$$

$$H_a: \beta_j \neq 0$$

Wald test calculates a Z statistic, which is:

$$z = \frac{\hat{b}}{SE(\hat{b})} \quad [2.29]$$

The  $Z^2$  value is asymptotically chi-square distribution and alpha values are calculated from chi-square distribution for determining the significance of factors.

### Model Adequacy:

The adequacy of the model is tested by the null hypothesis that all the coefficients associated with predictors are equal to zero versus at least one coefficient is not zero.

This hypothesis test is below:

$H_0$ : Coefficients of all variables are zero

$H_a$ : Coefficients of some variables differ from zero

G statistics is used to test this hypothesis. G statistics, sometimes called partial deviance, compare the deviance(D) of the model with the full set of predictors to that of a reduced model with one or more of the predictors missing. The partial deviance(G) is defined as (Hosmer and Lemeshow, 1989):

$$G = D(\text{Reduced Model}) - D(\text{Full Model})$$

The concept of deviance is the logistic regression analog of SSE in ordinary least square regression. The principle is that the observed and predicted values of the response variable are compared with and without some or all of the predictor variables.

The deviance is defined as:

$$D = -2 \sum_{i=1}^n [Y_i \ln(\hat{p}_i) + (1 - Y_i) \ln(1 - \hat{p}_i)]$$

For instance, consider the model with  $p-1$  predictors. Let  $D \equiv D(X_0, X_1, \dots, X_{p-1})$  denote *the* deviance for the full model.  $G \equiv D(X_0, X_1, \dots, X_{q-1}) - D(X_0, X_1, \dots, X_{p-1})$  denotes the partial deviance for the model that has been reduced by the final  $p-q-1$  variables.

Under the null hypothesis,  $G \sim \chi^2_{p-q}$  (Hosmer and Lemeshow, 1989). By using  $G$  value, a  $p$ -value is computed from the chi-square distribution with  $p-q$  degrees of freedom to test the fit of the logistic model.

### **Residual Plots**

After fitting a GLZ model to the data, to check the adequacy of the respective model, residuals are generally computed. Basically, there are two main type of residuals; Pearson and Deviance.

#### Pearson residuals

Pearson residuals and its standardized version is one type of residual which are elements of the Pearson chi-square. Pearson residuals are defined to be the standardized difference between the observed frequency and the predicted frequency. They measure the difference between the actual and predicted observation the relative deviations between the observed and fitted values and can be used to detect fitting problems in factor patterns. The formula is

$$r_i = \frac{Y_i - \hat{p}_i}{\sqrt{\hat{p}_i(1 - \hat{p}_i)}} \quad [2.30]$$

where  $Y_i$  is the actual observed value and  $\hat{p}_i$  is the predicted value for factor-level combination  $i$ .

### Deviance residuals

Deviance residual is another type of residual. It is a measure of how well the observation is predicted by the model. Observations that are poorly fit by the model have high deviance residuals. Since logistic regression uses the maximum likelihood principle, the goal in logistic regression is to minimize the sum of the deviance residuals. Therefore, this residual is parallel to the raw residual in ordinary regression, where the goal is to minimize the sum of squared residuals.

$$dev_i = \pm \left\{ -2 \left[ Y_i \log(\hat{p}_i) + (1 - Y_i) \log(1 - \hat{p}_i) \right] \right\}^{1/2} \quad [2.31]$$

Plotting deviance residuals against factors and checking these graphs for patterns or outliers helps to check the model fit.

These two measures, Pearson residual and deviance residual are considered to be the basic building blocks for logistic regression diagnostics. They are generally inspected to see any outlier or influential point. A good way of looking at them is to graph them against predictors (factors) and against case (index) numbers.

### Goodness of Fit Tests:

Goodness of fit tests measure how well the model fits the data. In these tests the hypothesis below is tested:

$$H_0 : \hat{E}[Y] = \frac{e^{b_0 + b_1 x_1 + \dots + b_k x_k + \dots + b_j x_j}}{1 + e^{b_0 + b_1 x_1 + \dots + b_k x_k + \dots + b_j x_j}}$$
$$H_a : \hat{E}[Y] \neq \frac{e^{b_0 + b_1 x_1 + \dots + b_k x_k + \dots + b_j x_j}}{1 + e^{b_0 + b_1 x_1 + \dots + b_k x_k + \dots + b_j x_j}}$$

Typically Pearson Statistics, Deviance Test or Hosmer-Lemeshow statistics are employed to test this hypothesis.

### Deviance Test

Under the hypothesis that the logistic model is correct, the deviance statistic, mentioned in previous sections, is asymptotically distributed as  $\chi^2_{n-p}$ . A p-value is computed from the chi-square distribution with  $n-p$  degrees of freedom to test the fit of the logistic model. If the Deviance goodness-of-fit test statistic is greater than 0.05, we fail to reject the null hypothesis implying that the model's estimates fit the data at an acceptable level.

### Chi-Square Test (Pearson)

The sum of the squared Pearson residuals, which are mentioned at previous sections, follows (asymptotically) a chi-square distribution, i.e.  $\sum_{i=1}^n r_i^2 \sim \chi^2_{n-p}$ . Similar to Deviance test a p-value is computed from the chi-square distribution with  $n-p$  degrees of freedom to test the fit of the logistic model.

### Hosmer-Lemeshow statistic

When applying *Pearson* statistic and the *Deviance* test, the degree of freedom (*df*) is determined by determining the number of distinct covariate patterns,  $I$ , i.e.  $df = I - p$ , where  $p$  is the number of covariates in the model. However, for large  $I$ , such as  $I \gg n$ , Hosmer and Lemeshow (1989) argue that this modification is inadequate. Their recommendation is to group observations into classes.

The Hosmer-Lemeshow statistic is used for binary logistic regression. It is the chi-square goodness-of-fit statistic from a 2 x (the number of groups) table. Typically using 10 classes is recommended, but comment that fewer may be required for smaller  $n$ . Hosmer and Lemeshow suggest using a minimum of six groups.

To calculate Hosmer-Lemeshow test statistic, the following steps are applied:

1. Order the fitted values
2. Group the fitted values in to  $c$  classes ( $c$  is between 6 and 10) of roughly equal size
3. Calculate the observed number in each group

$$O_{j(1)} = \sum_{k=1}^{n_j} Y_k, \quad j = 1, 2, \dots, c.$$

$$O_{j(0)} = \sum_{k=1}^{n_j} (1 - Y_k) = n_j - O_{j(1)}, \quad j = 1, 2, \dots, c.$$

4. Under the hypothesis that the logistic fit is correct, the expected counts for each cell are:

$$E_{j(1)} = \sum_{k=1}^{n_j} \hat{p}_k, \quad j = 1, 2, \dots, c.$$

$$E_{j(0)} = \sum_{k=1}^{n_j} (1 - \hat{p}_k) = n_j - E_{j(1)}, \quad j = 1, 2, \dots, c.$$

5. The test statistic has the form:

$$\sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} = \sum_{j=1}^c \sum_{l=0}^1 \frac{(O_{j(l)} - E_{j(l)})^2}{E_{j(l)}}$$



Under the hypothesis that the model is correct, the test statistic is asymptotically distributed as  $\chi^2_{c-2}$ . A p-value is computed from the chi-square distribution with  $c-2$  degrees of freedom to test the fit of the logistic model. If the Hosmer-Lemeshow goodness-of-fit test statistic is greater than 0.05, we fail to reject the null hypothesis that there is no difference between observed and model-predicted values, implying that the model's estimates fit the data at an acceptable level.

#### Classification Tables for Model Fit:

A classification table cross-classifies the binary response with a prediction of whether  $y=0$  or  $1$ . The prediction is  $\hat{y}=1$  when  $\hat{p}_i > \pi_0$  and  $\hat{y}=0$  when  $\hat{p}_i \leq \pi_0$ , for some cutoff point  $\pi_0$ , where  $\hat{p}_i$  is the probability of seeing event (1) for factor combination  $i$  and cutoff point is a value in  $[0, 1]$ , below which the outcome 0 is predicted and above which the outcome 1 is predicted. Determining the cutoff point is sometimes difficult. The choice of cutoff is arbitrary and generally  $\pi_0=0.5$  is used if it is equally likely in the population of interest that outcomes 0 and 1 will occur and the cost of incorrectly predicting 0 and 1 are approximately the same. If occurrence of outcomes 0 and 1 are not equally likely, a starting cutoff point is determined by estimating proportion of number of events in the total number of observations. Different cutoff points around this starting point are evaluated and the cutoff for which the proportion of incorrect predictions (prediction error rate) is the lowest is the one to be employed (Neter et al, 1996).

Table 4 gives a representation of a classification table.

**Table 4.** Classification table

		Prediction	
		$\hat{y}=0$	$\hat{y}=1$
Observation	$y=0$	Correct	<b>Incorrect</b>
	$y=1$	<b>Incorrect</b>	Correct

One of the ways that can be used to measure the model fit in classification table is calculating the prediction error rate which is formulated in [2.32]. High prediction error rate value means bad model fit.

$$\text{Prediction Error Rate} = \frac{\text{Number of incorrectly classified}}{\text{Total observation}} \quad [2.32]$$

### **Ordinal Logistic Regression**

Ordinal logistic regression is an extension of binary logistic regression to model ordinal categorical data. Let's have an ordinal categorical response with  $n$  categories. Ordinal logistic proceeds by comparing the effect of covariates on the possibility of being in each of  $n - 1$  categories compared with a "baseline category",  $k$ :

$$\log\left(\frac{p_j}{p_k}\right) = a_j + b_1x_1 + \dots + b_jx_j \quad [2.33]$$

where  $\alpha_j$  are the constants of the equation for  $n-1$  categories and  $\beta_j$  is the coefficient of the predictor variable  $x_j$ . By setting the  $k^{\text{th}}$  category as baseline, this is conceptually equivalent to fitting  $n - 1$  separate binary logistic models, comparing category 1 with category  $k$ , category 2 with  $k$ , and so on. In practice the many software packages usually estimates a simultaneous model which is more statistically sophisticated. For each independent variable, we have  $n - 1$  parameter estimates, each estimating the effect of a one-unit change in this variable on the log odds of being in category  $j$  rather than another category. (Agresti, 2002).

Logistic regression is the most widely used technique in GLZ and it has a large application area for analyzing categorical data (Agresti, 2002). This method generally analyzes location effects, but gives less information about dispersion. In proposed method we apply logistic regression for model fitting and offer a different method detecting dispersion.

#### **2.2.1.4. Bayesian Approach for Categorical Data**

Chipman and Hamada (1996) applied Bayesian Analysis for ordered categorical data. It is a comprehensive technique to analyze both location and dispersion effects. By using Gibbs Sampling algorithm, they sample from the posterior distribution of the factor coefficients of the generalized linear models. It has many advantages, but the main disadvantages of the approach are that it needs

complex computer applications and expert knowledge to determine the priors. It is therefore difficult to implement by the practitioners.

## CHAPTER III

### THE PROPOSED METHOD

#### 3.1. Description

Let  $Y$  be a categorical response variable with  $J$  categories. Let also  $X_1, \dots, X_k$  be the variables that the product or process designer controls to obtain the desired category response consistently.

The proposed method is aimed to find the robust settings of the design variables when the response variable is categorical (binary, ordinal or nominal). The main advantages of the proposed method compared to the existing methods are:

- ✓ The method clearly identifies location (mean) effects and dispersion (variance) effects.
- ✓ The application of the method is relatively simple and the method can be applied by popular software packages

The proposed method for analysis of categorical data for robust product or process design is given as following steps.

#### **Step 1. Create a suitable experimental design and collect data**

By considering the control factors ( $X_1, \dots, X_k$ ) and noise factors, that are causing variability beyond designers control, create a suitable experimental design by using fractional factorial and orthogonal array techniques or others. Conduct the experiments and collect the results.

**Step 2. Fit a categorical regression model and calculate event probabilities for each category**

By using categorical regression, fit a model that estimates the event probabilities for each category and each factor level combination,  $\hat{P}(Y_i=j)$ . Event probability is the probability that the factor level combination has the desired response category estimated by Equation [3.1].

$$\hat{P}(Y_i \leq j) = \hat{p} = \frac{e^{(\hat{b}_j + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \dots + \hat{b}_k x_k)}}{1 + e^{(\hat{b}_j + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \dots + \hat{b}_k x_k)}} \quad [3.1]$$

where,

$j=1, \dots, J$

$\hat{P}(Y_i \leq j)$  = the probability to see  $j^{th}$  category or below on the response for factor level combination  $i, i=1, \dots, I$ .

$b_k$  = coefficients for the factors,  $k=1, \dots, K$

$b_j$  = intercept for the  $j^{th}$  category

By the way,  $\hat{P}(Y_i=j) = \hat{P}(Y_i \leq j) - \hat{P}(Y_i \leq j-1)$ .

The regression model should be like in Equation [3.2] below.

$$\text{link}[P(Y \leq j)] = b_{j0} + b_1 x_1 + b_2 x_2 + \dots + b_k x_k \quad [3.2]$$

The reason for selecting categorical regression method (or GLZ) here is, GLZ models like Logistic regression can be applied by popular software packages (Minitab, Statistica, SPSS, JMP, etc.) and it is one of the most widely used methods for analyzing categorical data (Agresti, 2002).

### Step 3. Calculate expected category for each factor combination

By using control factor level combinations in step 1 and estimated event probabilities for each category in step 2, calculate the expected category and the variation for each factor level combination in the experiment by the formula in [3.3].

$$\begin{aligned}\hat{m}_i &= \hat{E}[Y_i] = \sum_{j=1}^J j\hat{P}(Y_i = j) \\ \hat{S}_i^2 &= \hat{V}(Y_i) = \hat{E}[Y_i^2] - (\hat{E}[Y_i])^2 = \left[ \sum_{j=1}^J j^2 \hat{P}(Y_i = j) \right] - \hat{m}_i^2\end{aligned}\quad [3.3]$$

where,

$$i = 1, 2, \dots, m$$

$\hat{m}_i$  = Estimator of the mean category for factor level combination i

$\hat{S}_i^2$  = Estimator of the variance of the categories for factor level combination i

j = category number (j=0,1,...,J)

$\hat{P}(Y_i = j)$  = Estimator of the probability that j<sup>th</sup> category is observed at i<sup>th</sup> factor level combination.

### Step4 . Find the optimal factor settings for desired mean and minimum variance

In order to obtain optimal factor settings that will provide the desired mean and minimum variance, two different approaches can be applied: Signal to Noise Ratios (SNR) and Response Surface Optimization (RSM).

#### a) Find the optimal factor setting that maximize the SNR

Calculate Taguchi's Signal to Noise ratios using  $\hat{E}[Y_i]$  and  $\hat{V}[Y_i]$  calculated in step 3 for each factor level combination. The SNR for a smaller the better problem is given below.

$$SNR_i = -10 \log_{10} (\hat{m}_i^2 + \hat{S}_i^2) \quad [3.4]$$

The SNR for a larger-the-better problem is:

$$SNR_i = -10 \log_{10} \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{\hat{m}_i^2} \right) \quad [3.5]$$

By using ANOVA or GLM, find the values of  $X_1, \dots, X_k$  that maximize SNR in order to achieve robust parameter design.

For the nominal-the-best problem, the SNR formula below in equation [3.6] is used. Other than smaller-the-better and larger-the-better case, a two-step optimization procedure is applied for nominal-the-best case.

$$SNR_i = 10 \log_{10} \frac{\hat{m}_i^2}{\hat{S}_i^2} \quad [3.6]$$

The two-step optimization of nominal-the-best problems:

1. By ignoring the mean find the factor level combinations that will maximize the SNR.
2. Without changing SNR, adjust the mean on target by using adjustment factors which have no effect on SNR.

#### **b) Find optimal factor settings by using RSM**

Obtain empirical models,  $f_1$  and  $f_2$ , for the mean,  $\hat{m}_i$ , and variance,  $\hat{S}_i^2$ , respectively. By using response surface optimization techniques mentioned in Section 2.1.2, build and solve an optimization model for the desired mean and minimum variance including model constraints (such as factor level boundaries). The optimization models for different problem types are given in Table 5.



**Table 5.** Optimization models for different problem types

<u>Problem Type</u>	<u>Mathematical Model</u>
Smaller-the-best	$\begin{aligned} \text{Min } \hat{E}(Y) &= f_1(x_1, x_2, \dots, x_k) \\ \text{Min } \hat{V}(Y) &= f_2(x_1, x_2, \dots, x_k) \\ \text{s.t.} \\ l_i &\leq x_i \leq u_i \\ \hat{E}(Y) &\geq 0 \\ \hat{V}(Y) &\geq 0 \\ i &= 1, \dots, k \end{aligned}$
Larger-the-best	$\begin{aligned} \text{Max } \hat{E}(Y) &= f_1(x_1, x_2, \dots, x_k) \\ \text{Min } \hat{V}(Y) &= f_2(x_1, x_2, \dots, x_k) \\ \text{s.t.} \\ l_i &\leq x_i \leq u_i \\ \hat{E}(Y) &\geq 0 \\ \hat{V}(Y) &\geq 0 \\ i &= 1, \dots, k \end{aligned}$
Nominal-the-best	$\begin{aligned} \text{Min } [\hat{E}(Y) - \text{Target}]^2 &= [f_1(x_1, x_2, \dots, x_k) - \text{Target}]^2 \\ \text{Min } \hat{V}(Y) &= f_2(x_1, x_2, \dots, x_k) \\ \text{s.t.} \\ l_i &\leq x_i \leq u_i \\ \hat{V}(Y) &\geq 0 \\ i &= 1, \dots, k \end{aligned}$

### 3.2. Illustrative Case Study-1

In this section, an example data set is used to illustrate the use of the proposed method. The proposed method and accumulation analysis methods are applied to this data set and the results are compared with the bayesian approach solution of Chipman and Hamada (1996).

### Step 1: A Foam Molding Experiment

The “Foam Molding Experiment” data, originally analyzed by Chipman and Hamada (1996), arise from an experiment to reduce voids in a urethane-foam product. The response consists of three levels (good, acceptable, poor), and all the design variables are at two levels,  $\{-1,1\}$ . The design is a fractionated eight-run control array crossed with a four-run noise array. The factors are shown in Table 6, and the data are given in Table 7. At each level of the control and noise factors, 10 parts are classified into one of the three categories, yielding a total of  $8 \times 4 \times 10 = 320$  observations.

**Table 6.** Foam Molding experiment factors and levels

<b>Control Factors</b>	<b>Noise Factors</b>
A: Shot weight (185/250)	H: Shift (second/third)
B: Mold temperature (70°F/1 20°F)	I: Shell quality (good/bad)
C: Foam block (use/do not use)	
D: RTV insert (use/do not use)	
E: Vent shell (vented/unvented)	
F: Spray wax viscosity (2:1/4: 1)	
G: Tool elevation (level/elevated)	

**Table 7.** Foam Molding experiment design and frequencies for good (0), OK (I) and poor (II)

							H	-1			1			-1			1		
							I	-1			-1			1			1		
A	B	C	D	E	F	G	0 I II			0 I II			0 I II			0 I II			
-1	-1	-1	-1	-1	-1	-1	3	6	1	6	4	0	1	4	5	0	10	0	
-1	-1	-1	1	1	1	1	0	3	7	3	4	3	0	6	4	0	7	3	
-1	1	1	-1	-1	1	1	0	0	10	0	1	9	0	0	10	0	0	10	
-1	1	1	1	1	-1	-1	0	0	10	0	10	0	0	3	7	0	9	1	
1	-1	1	-1	1	-1	1	3	5	2	3	7	0	3	5	2	1	6	3	
1	-1	1	1	-1	1	-1	2	8	0	4	5	1	0	5	5	1	5	4	
1	1	-1	-1	1	1	-1	2	7	1	2	5	3	2	7	1	1	6	3	
1	1	-1	1	-1	-1	1	0	4	6	1	7	2	0	4	6	0	3	7	

### Step 2: Analysis of the Problem by Applying the Proposed Method

The data in Table 7 was modeled by using ordinal logistic regression method and Minitab statistical program. The design is highly fractioned (1/16) with resolution III, therefore computation of 2-way or higher interactions is not possible. Noise factors H, I are also not included directly in the model since they are not controllable. They are considered within the replications. Logistic regression results are presented in Table 8. The full logistic regression output can be seen in Appendix-2.

Logistic regression applies Wald test to factor levels and Z values are obtained from the test. The  $Z^2$  values have asymptotical Chi-square distribution, so p values are calculated from Chi-square table. The analysis results show that factors A,B,C,E,F and G were statistically significant at  $\alpha$  being 0.013 at maximum. Table 8 also shows the estimated regression coefficients for the problem.

**Table 8.** Parameter estimates from logistic regression

Predictor	Coefficient	SE Coef	Z (Wald)	P
Const(1)	-2.596110	0.21163	-12.27	0.000
Const(2)	0.360502	0.14465	2.49	0.013
A	0.693708	0.13973	4.96	0.000
B	-0.912559	0.14309	-6.38	0.000
C	-0.488463	0.13809	-3.54	0.000
E	0.523686	0.13832	3.79	0.000
F	-0.513168	0.13881	-3.7	0.000
G	-0.768099	0.14058	-5.46	0.000

By choosing logit link and substituting the estimated coefficients in the logit function in equation [3.2], logistic regression equations for probability of the categories are given as:

$$\text{Logit}[\hat{P}(Y=0)] = -2.59611 + 0.693708 A - 0.912559 B - 0.488463 C + 0.523686 E - 0.513168 F - 0.768099 G$$

$$\text{Logit}[\hat{P}(Y \leq 1)] = 0.360502 + 0.693708 A - 0.912559 B - 0.488463 C + 0.523686 E - 0.513168 F - 0.768099 G$$

[3.7]

Here logit link is preferred because the transformation and representation of logit link is much easier compared with the other link functions listed in Table 3. The event probabilities of each response category for each experiment trial are calculated by transforming the equations [3.7] to equation format in [3.1]. For instance for the first factor level combination that is A=-1, B=-1, C=-1, E=-1, F=-1 and G=-1, the estimated event probabilities for each category are:

$$\begin{aligned} \hat{P}(Y_1=0) &= \exp(-2.59611 + 0.693708*(-1) - 0.912559*(-1) - 0.488463*(-1) + \\ &\quad 0.523686*(-1) - 0.513168*(-1) - 0.768099*(-1)) / (1 + \exp(-2.59611 + \\ &\quad 0.693708*(-1) - 0.912559*(-1) - 0.488463*(-1) + 0.523686*(-1) - \\ &\quad 0.513168*(-1) - 0.768099*(-1))) \\ &= 0.244 \end{aligned}$$

$$\begin{aligned} \hat{P}(Y_1=1) &= \exp(0.360502 + 0.693708*(-1) - 0.912559*(-1) - 0.488463*(-1) + \\ &\quad 0.523686*(-1) - 0.513168*(-1) - 0.768099*(-1)) / (1 + \exp(0.360502 + \\ &\quad 0.693708*(-1) - 0.912559*(-1) - 0.488463*(-1) + 0.523686*(-1) - \\ &\quad 0.513168*(-1) - 0.768099*(-1))) - \hat{P}(Y_1=0) \\ &= 0.617 \end{aligned}$$

$$\begin{aligned} \hat{P}(Y_1=2) &= 1 - \hat{P}(Y_1=0) - \hat{P}(Y_1=1) \\ &= 0.139 \end{aligned}$$

### Model Checking

Figure 3 is the output of goodness of fit results obtained from Minitab software. Log-Likelihood value, which has to be maximized in order to estimate the most suitable coefficients, is calculated from the maximum likelihood iterations. G statistic tests the null hypothesis that all the coefficients associated with predictors are equal to zero versus at least one coefficient is not zero. This hypotheses test is below:

$H_0$ : Coefficients of all variables are zero

$H_a$ : Coefficients of some variables differ from zero

P-value calculated from G statistics is very close to zero, indicating the adequacy of the model.

```
Log-Likelihood = -255,068  
Test that all slopes are zero: G = 110,834, DF = 7, P-Value = 0,000
```

```
Goodness-of-Fit Tests  
Method      Chi-Square  DF    P  
Pearson      4,14240    7     0,763  
Deviance     6,35541    7     0,499
```

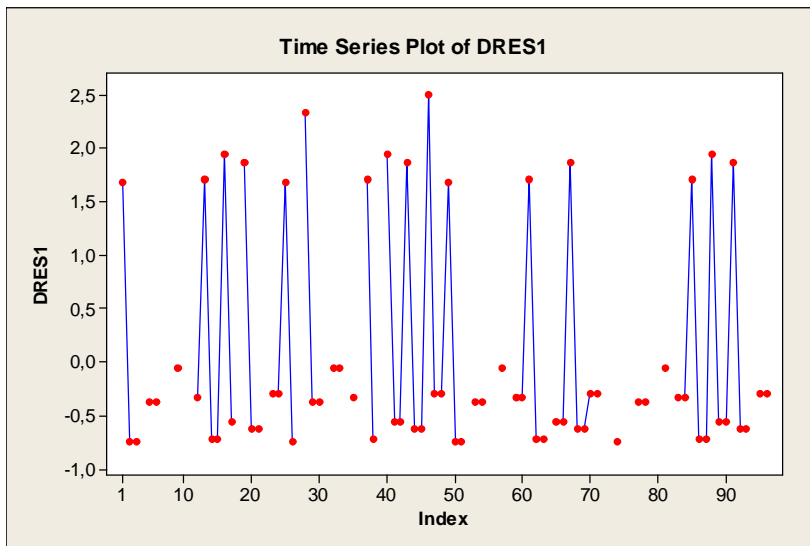
**Figure 3.** The Goodness of fit test results for the model

Also some goodness of fit test results are shown above which represent overall model adequacy. The p-value for the Pearson test is 0.763, and the p-value for the deviance test is 0.499, indicating that there is insufficient evidence to claim that the model does not fit the data adequately. If the p-value is less than your selected  $\alpha$ -level, the test rejects the null hypothesis that the model fits the data adequately.

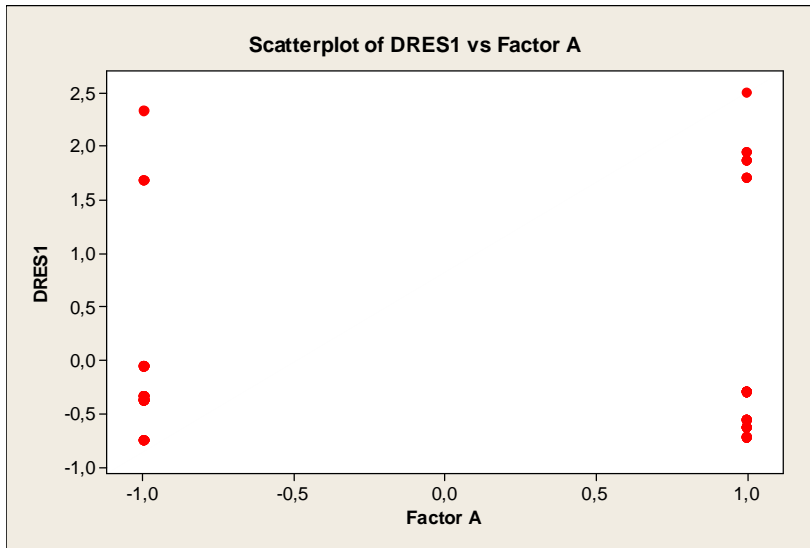
### Residual Check

Figure 4 shows deviance residuals vs. experiment trials (index) graph for category 0, which is the first category (good). The aim of checking these graphs is to catch a specific pattern or any outliers. From Figure 4 it can be seen that there is no

evidence for any pattern or outliers. Figure 5 is another type of residual plot which represents the scatter plot of deviance residuals vs. covariates (factors). The aim for checking these graphs is again to catch any patterns or outliers that will show the model inadequacy. Figure 4 and 5 only show the plot for factor A, the residual plots for the remaining factors can be seen in Appendix-3. Residual plots for category 1 are also presented in Appendix-3.



**Figure 4.** Deviance Residual plots vs. index for category 0



**Figure 5.** Deviance Residual plots vs. Factor A for category 0

### Step 3: Calculation of Expected Categories for Each Factor Combination

By using equations in [3.3], estimates for the expected category and the deviation from this category are obtained for each experiment factor level combination. Again for the first combination:

$$\begin{aligned} \hat{E}[Y_1] &= 0 \cdot \hat{P}(Y_1=0) + 1 \cdot \hat{P}(Y_1=1) + 2 \cdot \hat{P}(Y_1=2) \\ &= 0 \cdot 0.244 + 1 \cdot 0.617 + 2 \cdot 0.139 \\ &= 0.8949 \end{aligned}$$

$$\begin{aligned} \hat{V}(Y_1) &= 0^2 \cdot \hat{P}(Y_1=0) + 1^2 \cdot \hat{P}(Y_1=1) + 2^2 \cdot \hat{P}(Y_1=2) - (\hat{E}[Y_1])^2 \\ &= 0 \cdot 0.244 + 1 \cdot 0.617 + 4 \cdot 0.139 - (0.8949)^2 \\ &= 0.3717 \end{aligned}$$

All these results for estimated event probabilities, expected mean category and variance for all factor level combinations are presented in Table 9.



**Table 9.** Estimates of event probabilities of categories, expected categories and variance for foam molding experiment data

Trial	Factors						$\hat{P}(Y_i=j)$			$\hat{E}[Y_i]$	$\hat{V}[Y_i]$
	A	B	C	E	F	G	j=0	j=1	j=2		
1	-1	-1	-1	-1	-1	-1	0.244	0.617	0.139	0.8949	0.3717
2	-1	-1	-1	1	1	1	0.066	0.511	0.423	1.3568	0.3619
3	-1	1	1	-1	1	1	0.002	0.027	0.972	1.9703	0.0319
4	-1	1	1	1	-1	-1	0.053	0.465	0.482	1.4295	0.3507
5	1	-1	1	1	-1	1	0.23	0.622	0.148	0.9186	0.3716
6	1	-1	1	-1	1	-1	0.148	0.622	0.23	1.0814	0.3716
7	1	1	-1	1	1	-1	0.175	0.628	0.196	1.021	0.3714
8	1	1	-1	-1	-1	1	0.043	0.42	0.537	1.4942	0.3358

#### Step-4a: Parameter Optimization by SNR

In this problem, achieving the smallest category is desired, therefore it is a smaller the better problem. Reducing variance is the second goal, so a parameter optimization that will both reduce the mean and the variance is needed. Taguchi's Signal to Noise ratio (SNR) in equation [3.4] is a suitable choice to solve the problem. The SNR for the first factor level combination is:

$$SNR_1 = -10 * \log ((0.8949)^2 + 0.3717) = -0.691$$

Similarly all SNR values are calculated for each factor level combination and they are listed in Table 10.

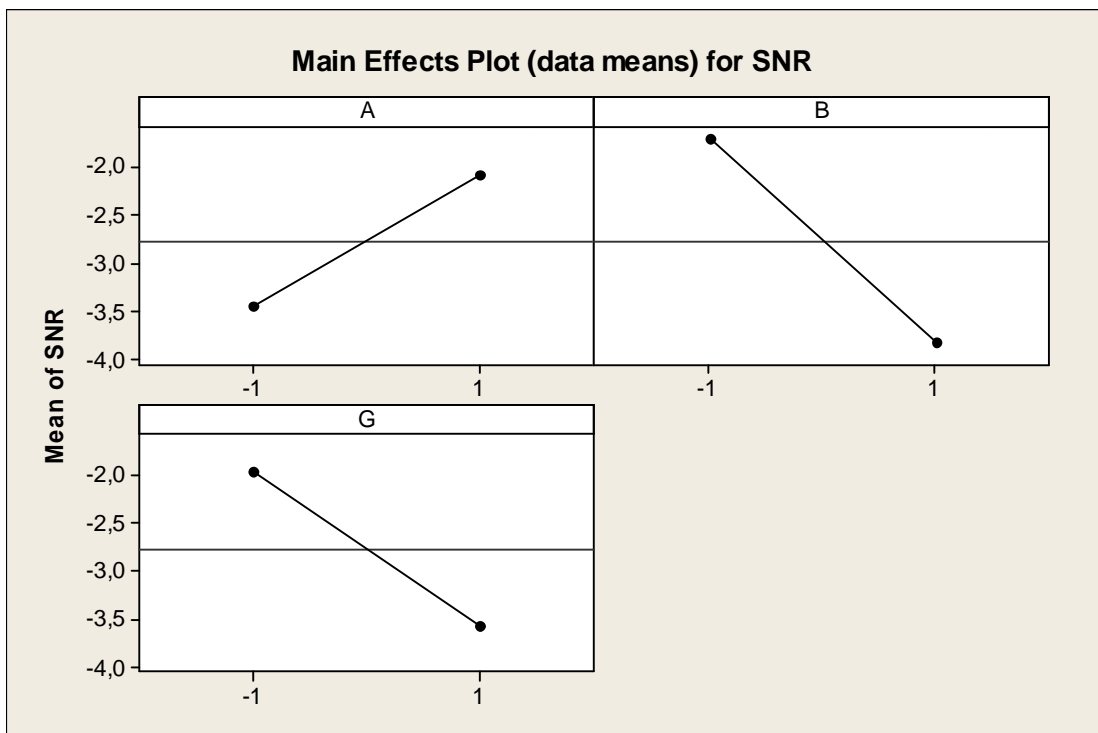
**Table 10.** Estimates of event probabilities of categories, expected categories, variance and SNR values for foam molding experiment data

<b>Trial</b>	<b>Factors</b>						$\hat{E}[Y_i]$	$\hat{V}(Y_i)$	<b>SNR<sub>i</sub></b>
<i>i</i>	<b>A</b>	<b>B</b>	<b>C</b>	<b>E</b>	<b>F</b>	<b>G</b>			
1	-1	-1	-1	-1	-1	-1	0.8949	0.3717	-0.691
2	-1	-1	-1	1	1	1	1.3568	0.3619	-3.43
3	-1	1	1	-1	1	1	1.9703	0.0319	-5.926
4	-1	1	1	1	-1	-1	1.4295	0.3507	-3.791
5	1	-1	1	1	-1	1	0.9186	0.3716	-0.847
6	1	-1	1	-1	1	-1	1.0814	0.3716	-1.878
7	1	1	-1	1	1	-1	1.021	0.3714	-1.504
8	1	1	-1	-1	-1	1	1.4942	0.3358	-4.096

In order to obtain smaller mean and variance, the SNR ratio values should be maximized. An ANOVA model is fitted to obtain optimum factor levels that will maximize SNR values. This model is fitted by using factor combinations and SNR values are in Table 10. ANOVA is applied by choosing  $\alpha=0.2$  and the final results are shown in Table 11. The full ANOVA output can be seen in Appendix-4. From Table 11, at  $\alpha=0.2$  level, it can be seen that factors A, B and G were found to be significant so these factors were examined to achieve maximum SNR. Figure 6 shows the main effects of these factors. Factor levels for (A, B, G) should be selected (+1, -1, -1) respectively to achieve maximum SNR, which are expected to provide best combinations for robustness.

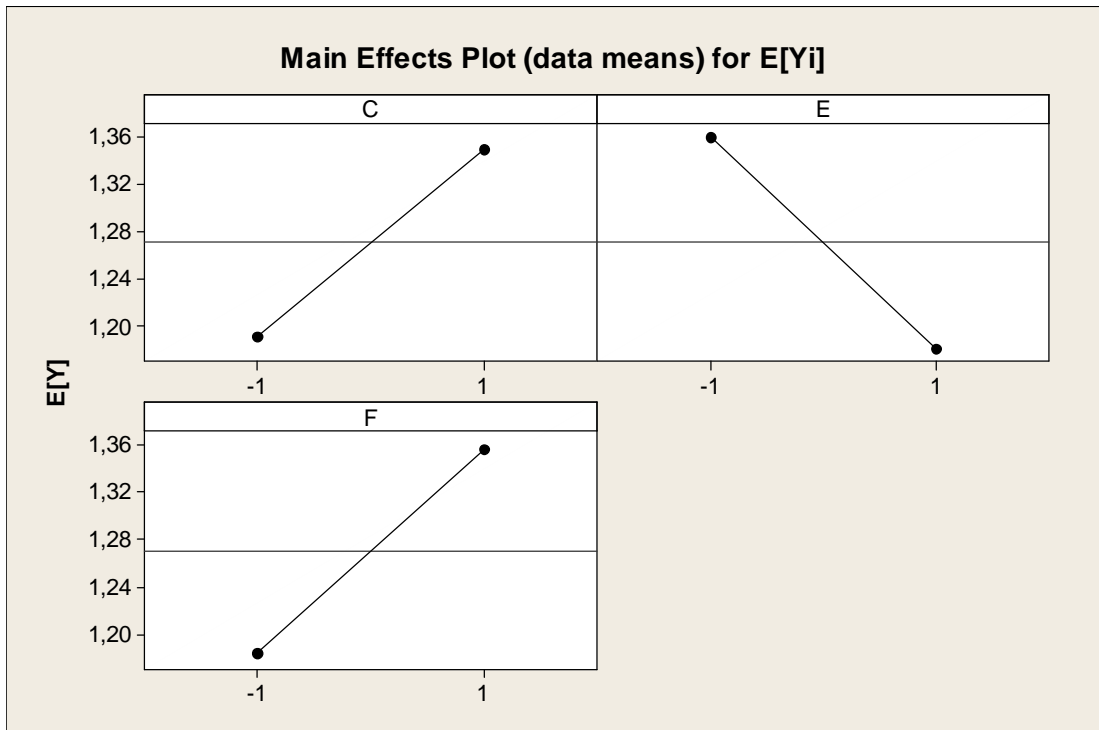
**Table 11.** ANOVA table for SNR values

Source	DF	SeqSS	AdjSS	AdjMS	F	P
A	1	3.799	3.799	3.799	2.68	0.177
B	1	8.97	8.97	8.97	6.33	0.066
G	1	5.176	5.176	5.176	3.65	0.129
Error	4	5.67	5.67	1.418		
Total	7	23.615				



**Figure 6.** The effect plot of Factors A, B and G for SNR

Factors C, E and F turn out to be insignificant in maximizing SNR and we are free to use any settings for these factors. On the other hand, in the logistic regression results factors C, E and F have significant effects on  $\hat{E}[Y]$  since they have effects on  $\hat{P}(Y=j)$ . This situation can be seen in the effect plots in Figure 7. These factors can be considered as adjustment factors, therefore (-1, +1, -1) levels of these factors can be chosen respectively for estimating the smallest mean value. In this manner, the optimum factor levels to achieve minimum mean and variance should be chosen as (A,B,C,E,F,G) = (+1, -1, -1, +1, -1, -1). These combinations predict  $\hat{P}(\text{Good}) = 0.79$ . The deviation on this probability can also be predicted as  $\hat{V}(\text{Pr}(\text{Good})) = 0.0069$ . The calculation for estimating the variation of probability is given in Appendix-6. %95 confidence intervals for  $\hat{P}(\text{Good})$  are (0.585, 0.909). The calculation for estimating the confidence bounds for probability is also given in Appendix-7.



**Figure 7.** The effect plot of Factors C, E and F for  $\hat{E}[Y]$

#### Step-4b: Parameter Optimization by Response Optimization

The surface functions for mean and variance are obtained by using Equation [3.3] and [3.7]. The mean function for this problem is:

$$\begin{aligned}
 \hat{E}[Y] &= 0 * \hat{P}(Y=0) + 1 * \hat{P}(Y=1) + 2 * \hat{P}(Y=2) \\
 &= 1 * [\exp(0.360502 + 0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - 0.513168*F - \\
 &0.768099*G)) / (1 + \exp(0.360502 + 0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - \\
 &0.513168*F - 0.768099*G)) - \exp(-2.59611 + 0.693708*A - 0.912559*B - 0.488463*C + \\
 &0.523686*E - 0.513168*F - 0.768099*G)) / (1 + \exp(-2.59611 + 0.693708*A - 0.912559*B - \\
 &0.488463*C + 0.523686*E - 0.513168*F - 0.768099*G))] + 2 * [1 - \exp(0.360502 + \\
 &0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - 0.513168*F - 0.768099*G)) / (1 + \\
 &\exp(0.360502 + 0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - 0.513168*F - \\
 &0.768099*G))]
 \end{aligned}$$

[3.8]

And the variance function is:

$$\begin{aligned}
 \hat{V}(Y) &= 0^2 * \hat{P}(Y=0) + 1^2 * \hat{P}(Y=1) + 2^2 * \hat{P}(Y=2) - (\hat{E}[Y])^2 \\
 &= 1 * [\exp(0.360502 + 0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - 0.513168*F - \\
 &0.768099*G)) / (1 + \exp(0.360502 + 0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - \\
 &0.513168*F - 0.768099*G)) - \exp(-2.59611 + 0.693708*A - 0.912559*B - 0.488463*C + \\
 &0.523686*E - 0.513168*F - 0.768099*G)) / (1 + \exp(-2.59611 + 0.693708*A - 0.912559*B - \\
 &0.488463*C + 0.523686*E - 0.513168*F - 0.768099*G))] + 4 * [1 - \exp(0.360502 + \\
 &0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - 0.513168*F - 0.768099*G)) / (1 + \\
 &\exp(0.360502 + 0.693708*A - 0.912559*B - 0.488463*C + 0.523686*E - 0.513168*F - \\
 &0.768099*G))] - \hat{E}[Y]^2
 \end{aligned}$$

[3.9]

By response surface optimization techniques mentioned in section 2.1.2 and optimizations models in Table 5, a multi-response optimization model is built for the

desired mean and minimum variance including model constraints (factor level boundaries) below.

$$\begin{aligned}
 & \text{Min } E(Y), \text{ Var}(Y) \\
 & \quad \underline{s.t.} \\
 & \quad -1 \leq x_i \leq 1 \\
 & \quad E(Y), \text{ Var}(Y) \geq 0
 \end{aligned}
 \tag{3.10}$$

The solution to the multi-response model in [3.10] can be obtained by using desirability functions mentioned in section 2.1.2.2. For mean surface [3.8] and variance surface [3.9] the individual desirability functions are formulated by substituting the problem inputs for the smaller the better problem in [2.11] below.

$$d_1(\hat{m}) = \begin{cases} 1.0 & \text{if } \hat{m} < 0.01 \\ \left( \frac{\hat{m} - 0.5}{0.01 - 0.5} \right)^s & \text{if } 0.01 \leq \hat{m} \leq 0.5 \\ 0 & \text{if } \hat{m} > 0.5 \end{cases}$$

$$d_2(\hat{S}^2) = \begin{cases} 1.0 & \text{if } \hat{S}^2 < 0.01 \\ \left( \frac{\hat{S}^2 - 0.4}{0.01 - 0.4} \right)^s & \text{if } 0.01 \leq \hat{S}^2 \leq 0.4 \\ 0 & \text{if } \hat{S}^2 > 0.4 \end{cases}$$

where  $d_1(\hat{m})$  is the desirability function for the mean,  $d_2(\hat{S}^2)$  is the desirability function for the variance. In the problem the exponent term,  $s$  is equal to 1 so that the function will increase or decrease linearly. The upper and lower values,  $U_i$  and  $L_i$ , are also determined by considering the problem requirements.

The overall desirability is:

$$D = (d_1(\hat{\mathbf{m}}) * d_2(\hat{\mathcal{S}}^2))^{1/2}$$

For optimization of the overall desirability MS Excel Solver is used. Excel solver uses GRG (General Reduced Gradient) NLP algorithm to get the optimal solution. The Solver sheet for the problem can be seen in Figure 8. The optimum solution for the problem, which will maximize  $D$ , is  $\hat{E}(Y) = 0.22$  and  $\hat{V}(Y) = 0.20$ . This solution estimates the probability of seeing category 0 as  $\hat{P}(\text{Good}) = 0.78$ .

From the optimization results, it can be concluded that all factor levels are found to be on bounds. Factors C, E, F and G are qualitative so we can not exceed the bounds but for Factor A and B better solution can be obtained by exceeding the bounds. This situation is presented in the response surface plot in Figure 9.

Factors						Invidual Desirabilities	
A	B	C	E	F	G	d1	d2
1	-1	-1	1	-1	-1	0,5561482	0,5036331

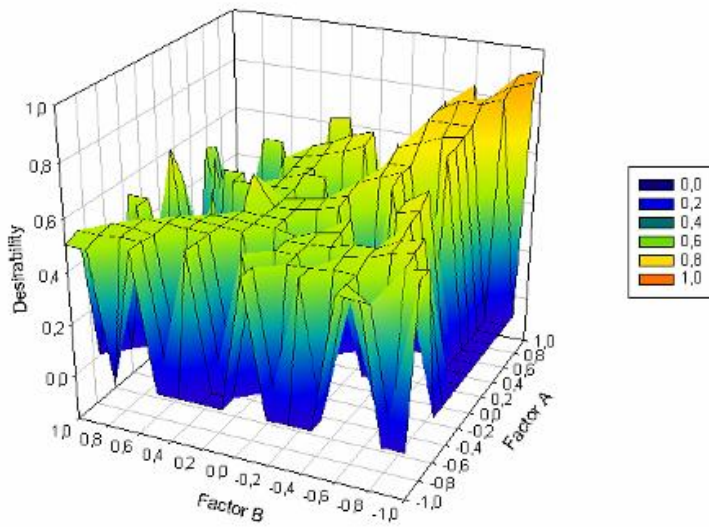
Mean	Variance	Category Probabilities		
E(Yi)	V(Yi)	P(Yi=0)	P(Yi=1)	P(Yi=2)
0,227487	0,203583	0,786436	0,199641	0,013923

Overall Desirability
0,529239674

**Figure 8.** The Desirability optimization sheet in Excel Solver

Response Surface Plot Desirability vs. Factor A, Factor B



**Figure 9.** Desirability response surfaces for Factor A and Factor B when  $C=-1$ ,  $E=1$ ,  $F=-1$  and  $G=-1$

### **Comparison of the Results of Proposed Method and the Existing Methods**

Table 12 gives a brief comparison of the results of the proposed method, Accumulation Analysis (AA) and Bayesian Approach (BA). AA calculations and results for the foam molding data are shown in Appendix-5. AA provides the worst estimate for the mean and no information about the variance compared with the other methods. Also it detects spurious effect of factor D. AA provides a mixture of location and dispersion information; therefore it is impossible to detect these effects separately. The results of Bayesian Approach have already been shown to be realistic and discussed by Chipman and Hamada (1996) in detail. Proposed method gives both location and dispersion effects and provides good estimates for both the mean and the variance compared with the Bayesian approach solution.



**Table 12.** Comparison of the proposed approach to the accumulation analysis and the Bayesian approach

Approach	Significant Factors & Levels	P(Good)	V(P(Good))
Accumulation Analysis	A+, B-, C-, D-, E+, F-, G- (location & dispersion effects)	0.45	-
Bayesian Approach	A+, B-, C-, E+, F-, G- (location & dispersion effects)	0.80	0.0076
Proposed Approach	A+, B-, G- (location&dispersion effects) C-, E+, F- (location effects)	0.79	0.0069

### 3.3. Illustrative Case Study-2

#### Step 1: Nitrogen Gas Defect Problem

In this section we apply the proposed approach to a real life dataset. The dataset is taken from an iron-casting company which has performed a Six Sigma project for reducing Nitrogen(N) gas defects on the cast iron parts (İşkol et al, 2005). The response is binary: N-gas defective ( $Y=1$ ) or non-defective ( $Y=0$ ). Seven factors affecting the response are measured. These factors are listed in Table 13. Due to confidentiality, the factor names are not displayed explicitly and coded by letters. The data was collected from real time production for a six month period consisting of 12417 observations. The dataset is consisting of real observations, it is not an experimental design and this caused wider range of levels for each factor. As it can be seen from Table 13, the standardized levels for factors range from -2.4 to +3.9. ANOVA like procedures are not suitable for these data since factor levels are not categorical but continuous.

**Table 13.** Nitrogen gas defect problem factors and levels

<b>Factor</b>	<b>Levels</b>
A	-1.7 – 3
B	-1.4 – 2.74
C	-1.68 – 3.6
D	-1.68 – 3.9
E	-2.4 – 3.13
F	-1.94 – 2.47
G	-1.32 – 2.08

### **Step 2: Analysis of Data by Applying Proposed Method**

The data for nitrogen gas defect problem are modeled by binary logistic regression using Minitab statistical program. All main factors and two-way interactions are included in the model. By backward elimination procedure based on checking p values calculated by Z statistics (Wald Test), the fitted model is simplified and insignificant factors, which have high p values, are eliminated from the model. Logistic regression results of the simplified model are presented in Table 14. The full logistic regression Output can be seen in Appendix-8.

**Table 14.** The logistic regression results for N-gas defect data

Predictor	Coefficient	SE Coef	Z	P
Constant	-4.32801	0.0995319	-43.48	0.000
A	0.16207	0.16076	1.01	0.313
B	0.632905	0.143828	4.4	0.000
E	-0.0543835	0.126899	-0.43	0.668
F	-0.503251	0.217173	-2.32	0.020
G	-0.437856	0.231717	-1.89	0.059
A*F	0.36527	0.104453	3.5	0.000
E*F	-0.338806	0.0925527	-3.66	0.000

From logistic regression results, Factor B and F are obviously significant, and factor interactions A\*F and E\*F are significant, too. Factor A and E are insignificant but because their interactions have low p values, they can not be eliminated from the model. There is an interesting case for factor G; its p value is 0.06 higher than 0.05 but the experts of the process recommends keeping factor G in the model because its effect is proven at previous studies in the process. So Factor G was kept in the model by expert opinion.

Similar to case study 1, logistic regression equations for probabilities of the categories are formulated below by choosing logit link and substituting the estimated coefficients in the logit function in equation [3.2]. Since we have two categories, only the regression function for probability of detecting a defect,  $P(Y=1)$ , is formulated. Here logit link is preferred because of the same reason in case study 1.

$$\text{Logit}[\hat{P}(Y=1)] = -4.32801 + 0.16207 A + 0.63291 B - 0.05438 E - 0.50325 F - 0.43786 G + 0.36527 A*F - 0.33881 E*F$$

[3.11]

The event probabilities of each response category for each experiment trial were calculated by transforming the equations [3.11] to equation format in [3.1]. For

instance for the factor level combination A=-1, B=1, E=1, F=-1 and G=1 the estimated event probabilities for each category are:

$$\begin{aligned}\hat{P}(Y=1) &= \exp(-4.32801 + 0.16207*(-1) + 0.63291*1 - 0.05438*1 - 0.50325*(-1) - 0.43786*1 + 0.36527*(-1)*(-1) - 0.33881*1*(-1)) / (1 + \exp(-4.32801 + 0.16207*(-1) + 0.63291*1 - 0.05438*1 - 0.50325*(-1) - 0.43786*1 + 0.36527*(-1)*(-1) - 0.33881*1*(-1))) \\ &= 0.041\end{aligned}$$

$$\hat{P}(Y=0) = 1 - \hat{P}(Y=1) = 0.959$$

### Step-3: Calculation of expected categories for each factor combination

By using equations in [3.3], estimates of the expected category and the variance from this category of each experiment factor level combination can be obtained. Again for the factor level combination used in step 2:

$$\begin{aligned}\hat{E}[Y] &= 0 * \hat{P}(Y=0) + 1 * \hat{P}(Y=1) \\ &= 0 * 0.959 + 1 * 0.041 \\ &= 0.041\end{aligned}$$

$$\begin{aligned}\hat{V}[Y] &= 0^2 * \hat{P}(Y=0) + 1^2 * \hat{P}(Y=1) - (\hat{E}[Y])^2 \\ &= 0 * 0.959 + 1 * 0.041 - (0.041)^2 \\ &= 0.039\end{aligned}$$

## Model Checking

The G statistics, which tests the null hypothesis that all the coefficients associated with predictors are equal to zero versus at least one coefficient is not zero, is obtained from Minitab software and shown in Figure 10. P-value calculated from G statistics is very close to zero, which is indicating the adequacy of the model.

```
Log-Likelihood = -1041,585  
Test that all slopes are zero: G = 119,480, DF = 7, P-Value = 0,000
```

### Goodness-of-Fit Tests

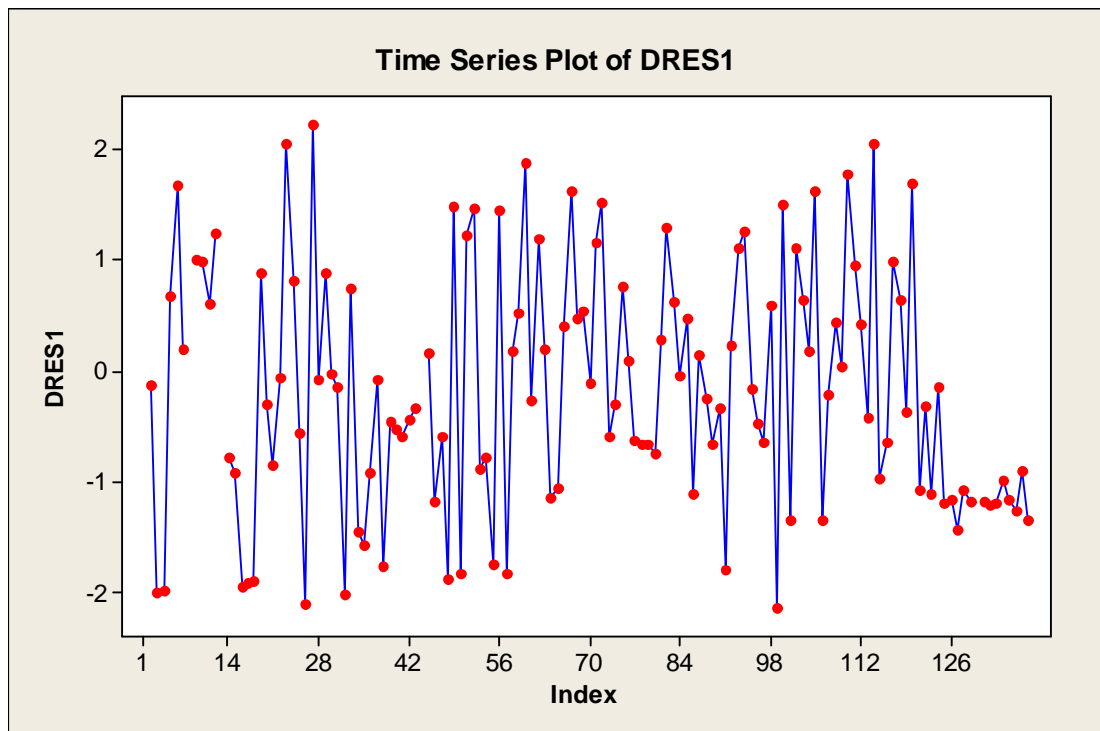
Method	Chi-Square	DF	P
Pearson	152,946	125	0,045
Deviance	162,017	125	0,014
Hosmer-Lemeshow	3,986	8	0,858

**Figure 10.** The Goodness of fit test results for the model 2

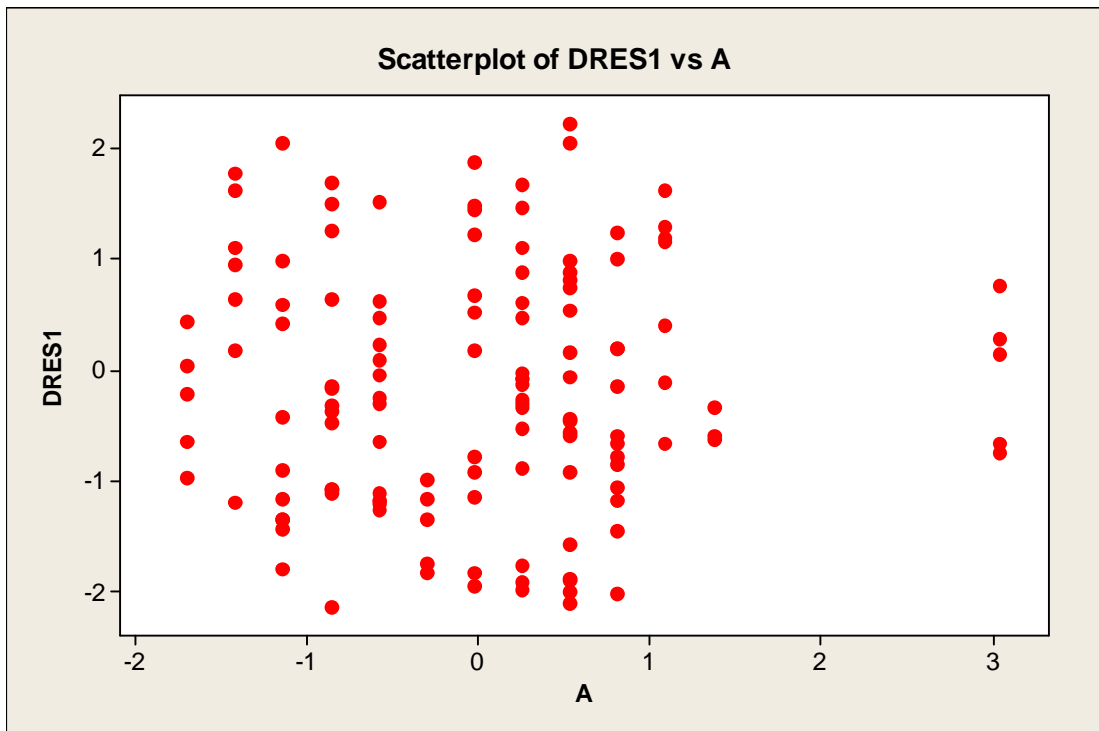
From Goodness of fit test results obtained in Minitab software which are also presented in Figure 10, it can be seen that the p-value for the Pearson test and the p-value for the deviance test is low, indicating poor fit. The main reason for that may be related with the argument of Hosmer-Lemeshow which is mentioned in Logistic regression background section. Hosmer-Lemeshow argued that for large number of covariate patterns ( $I$ ), such as  $I \gg n$ , the degree of freedom assumption for deviance and Pearson tests may be inadequate. In this case their recommendation is to use Hosmer-Lemeshow test. Hosmer-Lemeshow test results in Figure 10, gives a p value

of 0.858 which is high enough to indicate that there is insufficient evidence to claim model inadequacy.

Figure 11 shows Deviance Residual plots vs. experiment trials (index) graph. From Figure 11 it can be seen that there is no evidence for a trend or outliers. Figure 12 is another type of residual plots which represents the scatter plot of deviance residuals vs. covariates (factors). Figure 12 only shows the plot for factor A, the residual plots for the remaining factors can be seen in Appendix-9.



**Figure 11.** Deviance Residual plots vs. index for case 2



**Figure 12.** Deviance Residual plots vs. Factor A for case 2

Appendix-10 is the classification table for the model. Since occurrence of outcomes 0 and 1 are not equally likely in the dataset, a starting cutoff point is determined as 0.017 by estimating proportion of number of events in the total number of observations (220/12574). Different cutoff points around 0.017 are selected and estimated prediction error rates for each cutoff are presented in Table 15. From Table 15, it can be seen that prediction error rate becomes smaller when cutoff becomes larger, but cutoff point is selected as 0.023 for the problem since it gives relatively better result compared with the other alternatives and it is close to our starting point, 0.017.

**Table 15.** Prediction error rate values for case study 2

<b>Cutoff point</b>	<b>Prediction Error Rate</b>
c=0.014	0.6
c=0.017	0.43
c=0.020	0.26
c=0.023	0.18
c=0.5	0.017

By selecting the cutoff point as 0.023, the prediction error rate is 0.18 which is not high enough to claim bad fit.

#### **Step-4a: Parameter Optimization by SNR**

The factor level values in the dataset range non-uniformly and there are excessive number of factor levels observed for each factor. ANOVA like procedures are inadequate for analyzing this type of data, therefore SNR solution alternative of proposed method couldn't be applied for the problem.

#### **Step 4b: Parameter Optimization by Response Optimization**

The surface functions for the mean and the variance are obtained by using equations [3.3] and [3.7].



So the mean function is:

$$\begin{aligned} \hat{E}[Y] &= 0 * \hat{P}(Y=0) + 1 * \hat{P}(Y=1) \\ &= 0 * [1 - \exp(-4.32801 + 0.16207 A + 0.63291 B - 0.05438 E - 0.50325 F - 0.43786 G + \\ &0.36527 A * F - 0.33881 E * F) / (1 + \exp(-4.32801 + 0.16207 A + 0.63291 B - 0.05438 E + \\ &0.50325 F - 0.43786 G + 0.36527 A * F - 0.33881 E * F))] + 1 * [\exp(-4.32801 + 0.16207 A - \\ &0.63291 B - 0.05438 E - 0.50325 F - 0.43786 G + 0.36527 A * F - 0.33881 E * F) / (1 + \exp(- \\ &4.32801 + 0.16207 A + 0.63291 B - 0.05438 E - 0.50325 F - 0.43786 G + 0.36527 A * F - \\ &0.33881 E * F))] \end{aligned} \quad [3.12]$$

And the variance function is:

$$\begin{aligned} \hat{V}(\hat{Y}) &= 0^2 * \hat{P}(Y=0) + 1^2 * \hat{P}(Y=1) - (\hat{E}[Y])^2 \\ &= 0 * [1 - \exp(-4.32801 + 0.16207 A + 0.63291 B - 0.05438 E - 0.50325 F - 0.43786 G + \\ &0.36527 A * F - 0.33881 E * F) / (1 + \exp(-4.32801 + 0.16207 A + 0.63291 B - 0.05438 E - \\ &0.50325 F - 0.43786 G + 0.36527 A * F - 0.33881 E * F))] + 1 * [\exp(-4.32801 + 0.16207 A \\ &+ 0.63291 B - 0.05438 E - 0.50325 F - 0.43786 G + 0.36527 A * F - 0.33881 E * F) / (1 + \exp(- \\ &4.32801 + 0.16207 A + 0.63291 B - 0.05438 E - 0.50325 F - 0.43786 G + 0.36527 A * F - \\ &0.33881 E * F))] - \hat{E}[Y]^2 \end{aligned} \quad [3.13]$$

By response surface optimization techniques mentioned in section 2.1.2 and optimizations models in Table 5, a multi-response optimization model is built for the desired mean and minimum variance including model constraints (factor level boundaries) below. Factor level boundaries are determined by the process owners by considering economic conditions.

$$\text{Min } E(Y), \text{ Var}(Y)$$

s.t.

$$-0.86 \leq A \leq 0.25$$

$$-1.19 \leq B \leq -0.03$$

$$-1.6 \leq E \leq 0.37$$

$$-1.68 \leq F \leq -0.97$$

$$1.53 \leq G \leq 1.97$$

$$E(Y), \text{ Var}(Y) \geq 0$$

[3.14]

The solution to the multi-response model in [3.14] can be obtained by using desirability functions mentioned in section 2.1.2.2. For mean surface [3.12] and variance surface [3.13], similar to case study 1, the individual desirability functions are formulated by substituting the problem inputs for the smaller the better problem in [2.11] below.

$$d_1(\hat{m}) = \begin{cases} 1.0 & \text{if } \hat{m} < 1 * 10^{-13} \\ \left( \frac{\hat{m} - 0.8}{1 * 10^{-13} - 0.8} \right)^s & \text{if } 1 * 10^{-13} \leq \hat{m} \leq 0.8 \\ 0 & \text{if } \hat{m} > 0.8 \end{cases}$$

$$d_2(\hat{S}^2) = \begin{cases} 1.0 & \text{if } \hat{S}^2 < 1 * 10^{-13} \\ \left( \frac{\hat{S}^2 - 0.4}{1 * 10^{-13} - 0.4} \right)^s & \text{if } 1 * 10^{-13} \leq \hat{S}^2 \leq 0.4 \\ 0 & \text{if } \hat{S}^2 > 0.4 \end{cases}$$

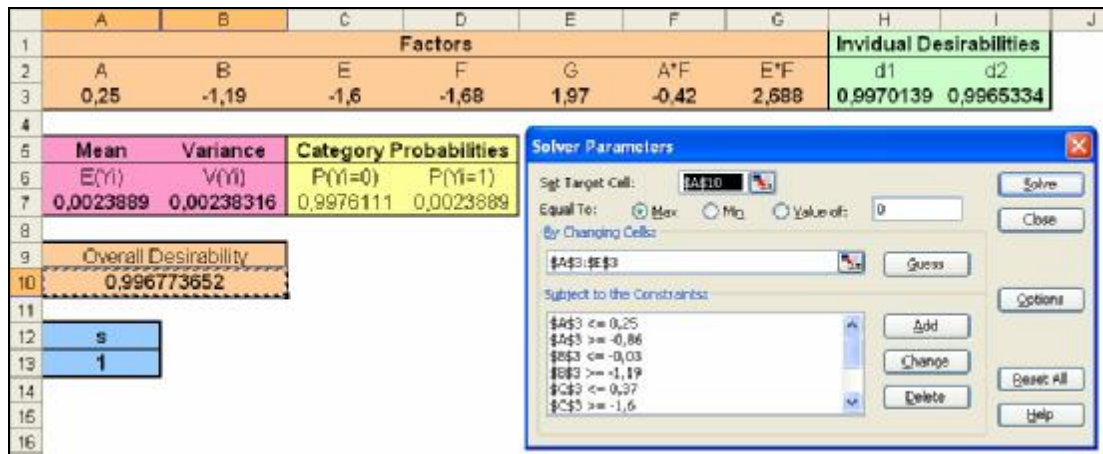
where  $d_1(\hat{m})$  is the desirability function for mean,  $d_2(\hat{S}^2)$  is the desirability function for variance. In the problem the exponent term,  $s$  is selected as 1 so that the function

will increase or decrease linearly. The upper and lower values,  $U_i$  and  $L_i$ , are also determined by considering the problem requirements.

The overall desirability is:

$$D = (d_1(\hat{m}) * d_2(\hat{S}^2))^{1/2}$$

The optimization of the overall desirability is performed in MS Excel Solver by GRG algorithm. The Solver sheet for the problem can be seen in Figure 13. The optimum solution for the problem, which will maximize  $D$ , is  $\hat{E}(Y) = 0.002$  and  $\hat{V}(Y) = 0.002$ . This solution estimates the probability of seeing category 0,  $\hat{P}(\text{Good}) = 0.997$ .



**Figure 13.** The Desirability optimization sheet in Excel Solver for case 2

From the optimization results, it can be concluded that all factor levels are chosen at the bounds. To obtain minimum mean category and variance, Factors A and G should be set at their upper limits and Factors B, E and F should be set at their lower limits. Figure 13 shows the results when these settings are used.

### Comparison of the Results of Proposed Method and the Real Life Results

The results obtained by the proposed method are compared with the six sigma project results presented in İşkol et al (2005). Table 16 gives the results of this comparison.

**Table 16.** Comparison of real life economical results and proposed method

	Real life economical factor levels	Proposed Method factor levels
<b>Factor A</b>	low	high
<b>Factor B</b>	low	low
<b>Factor C</b>	low	insignificant
<b>Factor D</b>	insignificant	insignificant
<b>Factor E</b>	low	low
<b>Factor F</b>	low	low
<b>Factor G</b>	high	high

From Table 16, it is seen that the levels for factors A, C and E are different in proposed method from real life economical factor levels. The reason for this difference can be explained as follows:

- ✓ In the process there are several factors other than seven factors considered in the proposed method. The applications in İşkol et al (2005) are performed with these additional factors and also additional data.
- ✓ Factor A is set at lower levels in the process because it is costly to increase it. Actually, in the analysis main effect of factor A is not significant but A\*F is significant. Therefore it can be set in upper economical bound, which is 0.25 in the process, to reduce the cost.
- ✓ Factor C is insignificant for N-gas defect problem, but higher levels of this factor can cause other defects in the process, so it is set at low levels in real life.

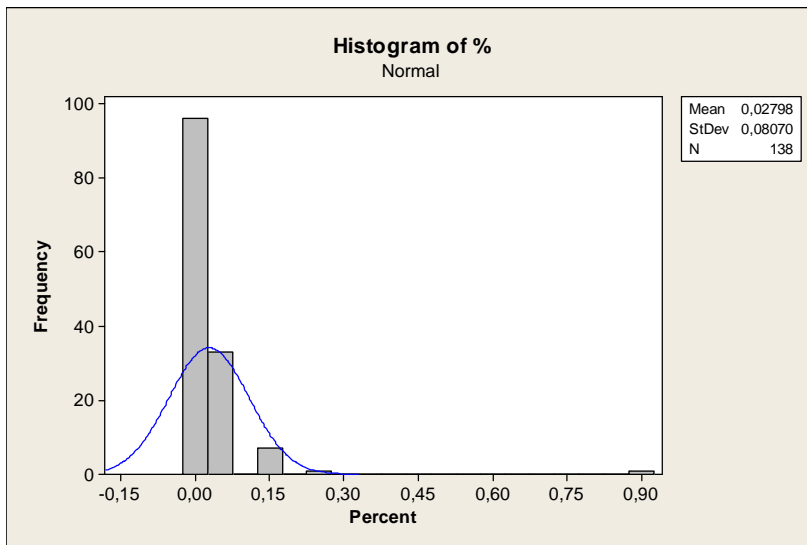
Because of the reasons above real life application results differ from the proposed method for some factors. Others agree with the proposed method.

### **Comparison of the Results of Proposed Method and OLS Regression Results**

Ordinary Least Square (OLS) regression can sometimes be used for binary response problems in industry. OLS regression can be applied to continuous response and can not be applied to binary response directly but people transform binary data to continuous by calculating the ratio or percentage of the events. This is generally a theoretically wrong application because (1) percent values are not continuous in infinity scale, (2) percent values does not satisfy model assumptions. But especially in industry some practitioners use OLS regression for binary responses because of its simplicity.

In our problem we tried to compare the results of the proposed method and OLS regression. But we found that it is unsuitable to apply OLS regression to the

dataset because the response (percentage) is highly non-normal as seen in Figure 14, although many different transformation techniques are used.



**Figure 14.** Histogram of percentage response

### **Discussion of Binary Problem**

Binary problem is actually a special case of the proposed method. In the third step of the proposed method when we calculate the mean category and variance for the binary problem as follows:

$$\hat{E}[Y] = 0 * \hat{P}(Y=0) + 1 * \hat{P}(Y=1) = \hat{P}(Y=1)$$

$$\hat{V}[Y] = 0^2 * \hat{P}(Y=0) + 1^2 * \hat{P}(Y=1) - (\hat{E}[Y])^2 = \hat{P}(Y=1) - [\hat{P}(Y=1)]^2$$

Above it can be seen that for the binary problem the estimator for variance is dependent to the estimator. Whatever our aim is to reduce or increase the mean, the variance will approach to zero if the mean comes closer to bounds 0 or 1. Therefore estimating two different functions for mean and variance by proposed method is not necessary for binary case. The solution directly obtained by Logistic Regression will be almost the same with the proposed method results.

## CHAPTER IV

### DISCUSSION OF NOMINAL DATA CASE

In two illustrations, we apply the proposed method to different type of categorical response: ordinal and binary. In the third step of our proposed method, in equation [3.3], we assume a natural ordering between categories and give a weight for each category, i.e. category numbers. For ordinal and binary response this assumption is applicable and the proposed method performs well. But for the nominal response types, there is no such a natural ordering and we have to make some other assumptions to apply the proposed method. In this section, we proposed two alternative treatments for the case of nominal data.

#### 4.1. Binary Treatment

In this treatment, we group the response categories into two groups; preferred and not-preferred. The preferred group is coded as 1 and the not-preferred group is coded as 0.

Let  $Y = \{Y_1, \dots, Y_J\}$  is a nominal categorical response with  $J$  distinct categories and  $Y_k$  is preferred as the response, where  $1 \leq k \leq J$ . By applying the treatment,  $Y_k$  is coded as 1 and the remaining  $n-1$  categories other than  $Y_k$  that are not preferred are coded as 0. Then our problem becomes a binary problem with two responses,  $Y = \{0, 1\}$ . By this way, we can apply proposed method to this binary data like in case study 2.

By Equation [3.3] the mean and variance functions can be obtained:

$$\hat{E}(Y) = 0 * \hat{P}(Y=0) + 1 * \hat{P}(Y=1) = \hat{P}(Y=1) = \hat{P}(\text{Preferred})$$



$$\begin{aligned}\hat{V}(Y) &= \hat{E}(Y^2) - [\hat{E}(Y)]^2 = 0^2 * \hat{P}(Y=0) + 1^2 * \hat{P}(Y=1) - [\hat{P}(Y=1)]^2 \\ &= \hat{P}(Y=1) * (1 - \hat{P}(Y=1)) = \hat{P}(\text{Preferred}) * (1 - \hat{P}(\text{Preferred}))\end{aligned}$$

If we desire to find expected number of preferred response,  $Y_k$ , in  $n$  observations and its variance:

$$\hat{E}(\# \text{ of preferred}) = n * \hat{E}(Y) = n * \hat{P}(\text{Preferred})$$

$$\hat{V}(\# \text{ of preferred}) = n * \hat{V}(Y) = n * \hat{P}(\text{Preferred}) * (1 - \hat{P}(\text{Preferred}))$$

which are expected mean and variance for multi-nominal distribution (Scheaffer, 1995).

#### 4.2. Ordinal Treatment

Another way to apply the proposed method to nominal data is converting the nominal categories to ordered categories by ranking by preference. This treatment is explained on an example below.

Let  $Y = \{\text{Company A, Company B, Company C}\}$  be a nominal response with 3 categories. In the problem, company B is preferred the most, company C is less preferred and company A is the least preferred. We can rank the categories by preference, such as company B is 1, company C is 2 and company A is 3, and the proposed approach can be applied on these ordinal categories.

## CHAPTER V

### CONCLUSION AND FURTHER STUDIES

In this study we propose a method which can be used to find the robust settings of the parameters when the response variable(s) is categorical. In four steps the method can be applied and it provides two different alternatives, Signal-to-Noise ratio and Response Surface Method, for finding optimum robust settings. The method has clear advantages since it identifies both location effects (mean) and dispersion (variance) for optimization, the application of the method is simple, advanced statistical background and complex assumptions is not required and it can be applied by popular software packages.

We applied the proposed method into two different datasets for illustration. In the first illustrative case study, we applied the method on a problem which has ordinal categorical response and has been studied by some authors before. Our proposed method and Accumulation Analysis methods are applied to this data set and the results are compared with the comprehensive Bayesian Approach solution of Chipman and Hamada (1996). Accumulation Analysis gives poor results but our proposed method gives both location and dispersion effects and provides good estimates for both mean and variance compared with the solution with bayesian technique. By this study it can be concluded that proposed method is well applicable for designed experiments with ordinal categorical response.

In the second illustrative study we applied our proposed method on a binary response problem. The data was provided from daily production observations of a real life process. Due to non-uniform and excess number of factor levels ANOVA like procedures are inadequate for analyzing this data, therefore SNR ratio solution alternative of proposed method couldn't be applied for this kind of data.

Also because of the same reason, we couldn't apply Accumulation Analysis method to this dataset. Here we can conclude that SNR solution alternative of our method is not applicable when the data has a non-uniform experimental design. We applied RSM to this data set and compared the results with a previous study results in this process. The results obtained by our proposed method make sense for the process and the optimum results are approved by the process experts that they are useful to obtain desired mean and minimum variance.

In two illustrative cases, the proposed method is applied on single response. If there is more than one response in a problem, different multi-objective optimization algorithms should be used or more than two desirability functions should be considered. This situation can be studied on future works. Also, multiple response case with mixtures of continuous and categorical data is even more complex and the proposed method may be expanded for these types of problems.

The proposed method uses logistic regression for model fitting and Maximum Likelihood Estimation (MLE) is usually used for parameter estimation in logistic regression. With small sample sizes (less than 10), MLE results can be heavily biased and the large sample optimality properties do not apply. In our study, the datasets in two different cases have large sample sizes and we don't need to check the estimation procedure. But for small samples, there may be some problems and bias in estimation of the model. This situation should be discussed in a future study.

The proposed model assumes there is no dependency between multiple responses. Most of the practitioners assume this but sometimes this situation can be observed in real life processes. This problem can also be studied in the future.

## REFERENCES

Agresti, A., 1996, "An Introduction to Categorical Data Analysis" John Wiley and Sons, Inc.

Agresti, A., 2002, "Categorical Data Analysis", John Wiley & Sons Inc. Hoboken New Jersey

Box, G.E.P., 1985 "Discussion of Off-Line Quality Control, Parameter Design, and the Taguchi Method", *Journal of Quality Technology*, Vol. 17, p.198-206.

Box, G. and Jones, S., 1986, "Discussion of "Testing in Industrial Experiments with Ordered Categorical Data by V.N. Nair", *Technometrics*, Vol. 28, p.295-301

Chipman H., Hamada M., 1996, "Bayesian Analysis of Ordered Categorical Data From Industrial Experiments.", *Technometrics* , Vol. 38(1), p.1-10

Del Castillo, E. and Montgomery, D.C., 1993, "A Nonlinear Programming Solution to the Dual Response Problem", *Journal of Quality Technology* Vol. 25, p. 199–204.

Derringer, G., and Suich, R., 1980, "Simultaneous Optimization of Several Response Variables", *Journal of Quality Technology*, Vol. 12(4), p.214-219.

Hamada M., Wu C.F.J. , 1986, "Should Accumulation Analysis and Related Methods be Used for Industrial Experiments? Discussion of Testing in Industrial Experiments with Ordered Categorical Data by V.N. Nair", *Techometrics*, Vol. 28, p.302-306.

Hamada M., Wu C.F.J., 1990, "A Critical Look at Accumulation Analysis and Related Methods", *Technometrics*, Vol. 32, p.119-162.

Hosmer, D., Lemeshow S., 1989, "Applied Logistic Regression", NY: Wiley & Sons.

İşkol M.E., Erkul İ., Karagüzel B., Gürsoy K., Aral B., Gülbay B., 2005, "6 Sigma Applications at Casting Process", *Altı Sigma Forum*, Vol. 5, p 24-32

Kacker R.N., 1985, "Off-Line Quality Control, Parameter Design, and the Taguchi Method", *Journal of Quality Technology*, Vol. 17(4), p.176-188

Köksoy, O., Muluk, F.Z., 2004, "Solution To The Taguchi's Problem With Correlated Responses", *G.U. Journal of Science*, Vol. 17(1), p. 59-70.

Köksoy, O., Hocaoglu, G., 2005, "Multi-Objective Optimization Solutions to the Taguchi's Problem", *G.U. Journal of Science*, Vol. 18(4), p. 613-626

Leon R.L., Shoemaker A.C., Kacker R.N., 1987, "Performance measures independent of adjustment", *Technometrics*, Vol. 29, p. 253–265.

McCullagh P., 1980, "Regression Models for Ordinal Data", *Journal of the Royal Statistical Society B*, Vol. 42, p.109-142

McCullagh, P. and Nelder J.A., 1989, "Generalized Linear Models", Chapman and Hall: London.

Microsoft Excel Solver (2003), Microsoft Corp., U.S.,  
<http://office.microsoft.com>.

Minitab (2005), Minitab Statistical Software, Minitab Ltd. UK.  
[www.minitab.co.uk](http://www.minitab.co.uk).

Montgomery D.C., 2001, "Introduction to Statistical Quality Control", 4th edition, John Wiley & Sons Inc.

Nair, V.N., 1986, "Testing Industrial Experiments with Ordered Categorical Data", *Technometrics*, Vol. 28(4), p. 283-291.

Nair, V. N., Abraham, B., MacKay, J., Box, G., Kacker, R., Lorenzen, T., Lucas, J., Myers, R., Vining, G., Nelder, J., Phadke, M., Sacks, J., Welch, W., Shoemaker, A., Tsui, K., Taguchi, S., and Wu, C. F. J., 1992, "Taguchi's parameter design: A panel discussion.", *Technometrics*, Vol. 34, p. 127-161

Nair, V. N. Taam, W., Ye, K.Q., 2002, "Analysis of Functional Responses From Robust Design Studies", *Journal of Quality Technology*, Vol. 34(4), p.355-370

Neter, J., Kutner, M.H., Nachtsheim, C.J. Wasserman, W., 1996, "Applied Linear Statistical Models", McGraw-Hill Irwin, Fourth Edition,

Phadke M.S., 1989, "Quality Engineering Using Robust Design.", Prentice-Hall, Englewood Cliffs, New Jersey, USA

Robinson T. J., Borrer C. M. and Myers C. M., 2004, "Robust Parameter Design:A Review", *Qual. Reliab. Engng. Int.*, Vol. 20, p. 81–101

Ross P.J., 1996, "Taguchi Techniques for Quality Engineering", McGraw-Hill Inc. 2<sup>nd</sup> edition

Scheaffer, R.L., McClave, J.T., 1995, "Probability and Statistics for Engineers", Duxbury Press, 4<sup>th</sup> Edition, California

Shoemaker, A.C., Tsui, K.L., Wu, C.F.J., 1991, "Economical experimentation methods for robust parameter design", *Technometrics*, Vol. 33, p. 415–427.

Statistica 2005, Statsoft Inc. UK. , [www.statsoft.com](http://www.statsoft.com).

Taguchi, G. and Wu, Y., 1980, "Introduction to off-line quality control.", Nagoya, Japan: Central Japan Quality Association.

Taguchi, G., 1986, "Introduction to quality engineering", White Plains, New York: UNIPUB/Krauss International.

Tang, L.C., Xu, K., 2002, "A Unified Approach for Dual Response Surface Optimization", *Journal of Quality Technology*, Vol. 34(4), p. 437-447.

Vining, G.G., Myers, R.H., 1990, "Combining Taguchi and response surface philosophies: A dual response approach", *Journal of Quality Technology*, Vol. 22, p. 38-45.

## APPENDIX A

### Post-Etch Window Size Example

FACTORS								OUTPUT				
A	BD	C	E	F	G	H	I	I	II	III	IV	V
1	1	1	1	1	1	1	1	10	0	0	0	0
1	1	2	2	2	2	2	2	0	3	3	2	2
1	1	3	3	3	3	3	3	1	0	0	9	0
1	2	1	1	2	2	3	3	10	0	0	0	0
1	2	2	2	3	3	1	1	10	0	0	0	0
1	2	3	3	1	1	2	2	5	3	2	0	0
1	3	1	2	1	3	2	3	10	0	0	0	0
1	3	2	3	2	1	3	1	5	0	0	5	0
1	3	3	1	3	2	1	2	0	1	4	5	0
2	1	1	3	3	2	2	1	2	5	3	0	0
2	1	2	1	1	3	3	2	1	1	2	6	0
2	1	3	2	2	1	1	3	1	0	1	3	5
2	2	1	2	3	1	3	2	5	0	3	2	0
2	2	2	3	1	2	1	3	6	3	1	0	0
2	2	3	1	2	3	2	1	10	0	0	0	0
2	3	1	3	2	3	1	2	10	0	0	0	0
2	3	2	1	3	1	2	3	0	0	4	3	3
2	3	3	2	1	2	3	1	0	0	0	0	10

- A: Mask Dimension
- BD: viscosity x Bake Temp
- C: Spin speed
- E: Bake time
- F: Aperture
- G: Exposure Time
- H: Developing Time
- I: Plasma etch time



OUTPUT:

- I-Window not open
- II-(0, 2.25)
- III- (2.25,2.75)
- IV- (2.75,3.25)
- V- (>3.25)

**1. Accumulation Analysis (Taguchi) Results:**

<b>ANOVA Table</b>				
<b>Source</b>	<b>df</b>	<b>S</b>	<b>MS</b>	<b>F</b>
A	4	26,65526	6,663816	15,67433
BD	8	112,3241	14,04051	33,02546
C	8	125,5418	15,69272	36,91172
E	8	36,97849	4,622311	10,8724
F	8	27,90145	3,487682	8,20357
G	8	42,29724	5,287156	12,43621
H	8	45,58938	5,698672	13,40416
I	8	23,81925	2,977406	7,003322
E2	656	278,8931	0,425142	
<b>TOTAL</b>	<b>716</b>		<b>720</b>	

BD, C, G,H and A seems to be most significant factors.

All factors seems to be significant because of small MSE. This is caused by large df of error.

**2. Results of Scoring Method by Nair (1986)**

<b>ANOVA Table</b>					
<b>Source</b>	<b>df</b>	<b>SS(l)</b>	<b>SS(d)</b>	<b>MS(l)</b>	<b>MS(d)</b>
A	1	7,496716	1,300572	7,496716	1,300572
BD	2	38,28807	8,491404	19,14404	4,245702
C	2	43,29349	5,610754	21,64674	2,805377
E	2	5,043237	17,05333	2,521618	8,526666
F	2	6,068595	14,05101	3,034297	7,025506
G	2	14,25637	2,101915	7,128187	1,050957
H	2	12,2564	7,9501	6,1282	3,97505
I	2	3,626939	18,59202	1,813469	9,29601

BD,C has strong location effects.  
E, F and I has strong dispersion effects.

Accumulation analysis gives a mixture of effects but scoring methods separates location and dispersion and provides better information for selecting robust settings.

## APPENDIX B

### Minitab Logistic Regression Output for Foam-Molding Experiment

#### Results for: Foam Molding

#### Ordinal Logistic Regression: Response versus A; B; C; E; F; G

Link Function: Logit

#### Response Information

Variable	Value	Count
Response	1	38
	2	156
	3	126
	Total	320

Frequency: Freq

\* NOTE \* 71 cases were used

\* NOTE \* 25 cases contained missing values or was a case with zero frequency.

#### Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds	95% CI	
					Ratio	Lower	Upper
Const(1)	-2,59611	0,211630	-12,27	0,000			
Const(2)	0,360502	0,144654	2,49	0,013			
A	0,693708	0,139729	4,96	0,000	2,00	1,52	2,63
B	-0,912559	0,143088	-6,38	0,000	0,40	0,30	0,53
C	-0,488463	0,138092	-3,54	0,000	0,61	0,47	0,80
E	0,523686	0,138324	3,79	0,000	1,69	1,29	2,21
F	-0,513168	0,138814	-3,70	0,000	0,60	0,46	0,79
G	-0,768099	0,140581	-5,46	0,000	0,46	0,35	0,61

Log-Likelihood = -255,082

Test that all slopes are zero: G = 110,806, DF = 6, P-Value = 0,000

#### Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	4,21124	8	0,838
Deviance	6,38399	8	0,604

#### Measures of Association:

(Between the Response Variable and Predicted Probabilities)

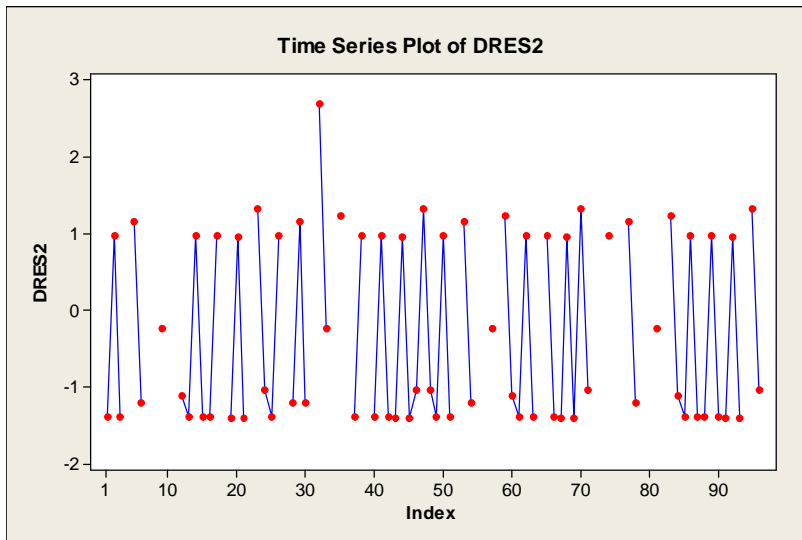
Pairs	Number	Percent	Summary Measures
Concordant	21423	70,5	Somers' D 0,51
Discordant	5849	19,3	Goodman-Kruskal Gamma 0,57
Ties	3100	10,2	Kendall's Tau-a 0,31
Total	30372	100,0	

## APPENDIX C

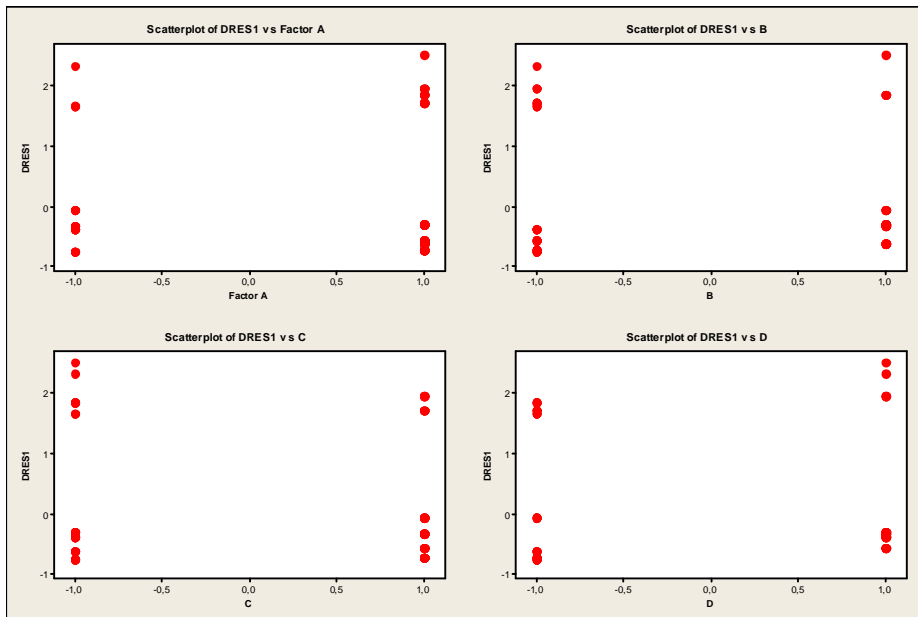
### Residual Plots for Foam-Molding Experiment



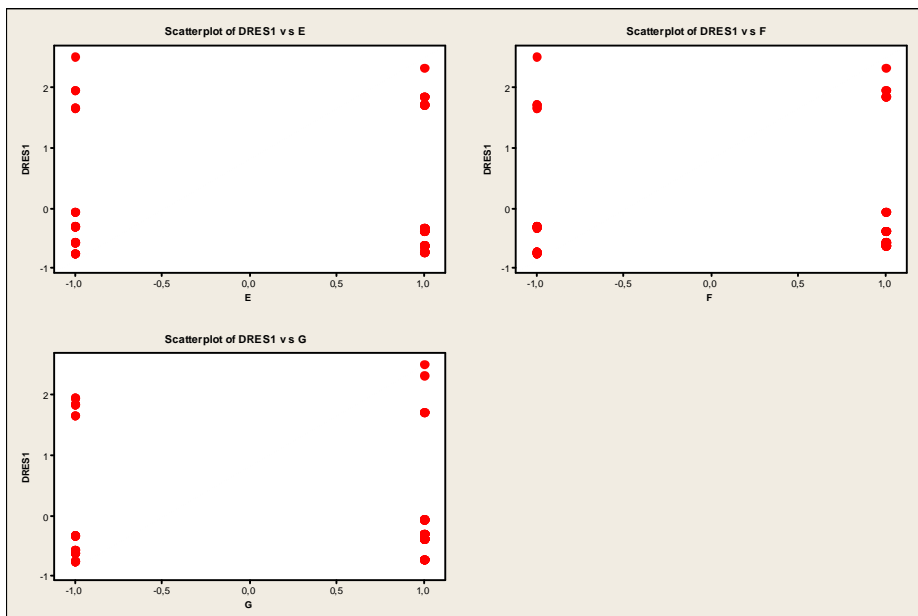
Deviance Residuals vs. index plot for category 0



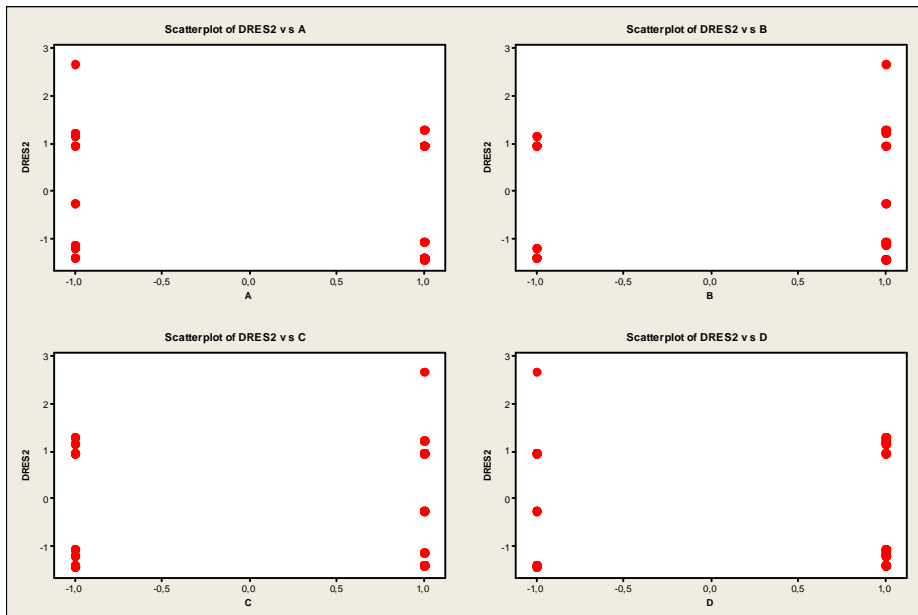
Deviance Residuals vs. index plot for category 1



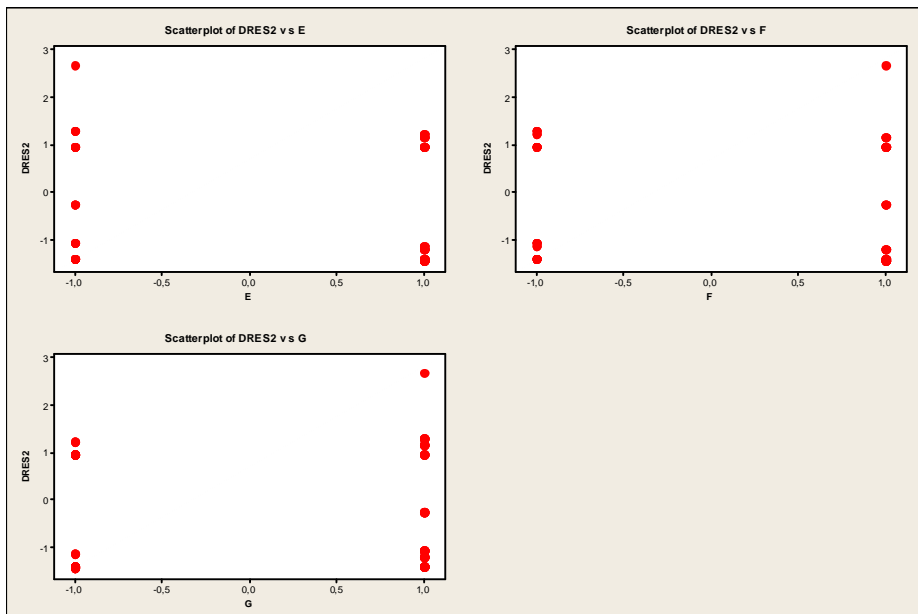
Deviance Residuals vs. Factors A,B,C,D plot for category 0



Deviance Residuals vs. Factors E,F,G plot for category 0



Deviance Residuals vs. Factors A,B,C,D plot for category 1



Deviance Residuals vs. Factors E,F,G plot for category 1

## APPENDIX D

### GLM model for SNR in case study 1

#### Full Model:

##### General Linear Model: SNR versus A; B; C; E; F; G

Factor	Type	Levels	Values
A	fixed	2	-1; 1
B	fixed	2	-1; 1
C	fixed	2	-1; 1
E	fixed	2	-1; 1
F	fixed	2	-1; 1
G	fixed	2	-1; 1

Analysis of Variance for SNR, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
A	1	3,799	3,799	3,799	1,70	0,416
B	1	8,970	8,970	8,970	4,02	0,295
C	1	0,925	0,925	0,925	0,41	0,636
E	1	1,139	1,139	1,139	0,51	0,605
F	1	1,372	1,372	1,372	0,61	0,577
G	1	5,176	5,176	5,176	2,32	0,370
Error	1	2,233	2,233	2,233		
Total	7	23,615				

S = 1,49447    R-Sq = 90,54%    R-Sq(adj) = 33,80%

#### Reduced Model:

##### General Linear Model: SNR versus A; B; G

Factor	Type	Levels	Values
A	fixed	2	-1; 1
B	fixed	2	-1; 1
G	fixed	2	-1; 1

Analysis of Variance for SNR, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
A	1	3,799	3,799	3,799	2,68	0,177
B	1	8,970	8,970	8,970	6,33	0,066
G	1	5,176	5,176	5,176	3,65	0,129
Error	4	5,670	5,670	1,418		
Total	7	23,615				

S = 1,19061    R-Sq = 75,99%    R-Sq(adj) = 57,98%

## APPENDIX E

### Accumulation Analysis (AA) solution for case study 1

Foam Molding Experiment

Control Factors: A, B, C, D, E, F, G

Noise Factors: H, I

Response categories: Good(I), OK(II), Poor(III)

$2_{III}^{7-4}$  1/16 fractional experiment

4 replicates (considering noise factors- outer array)

#### Accumulation Analysis

FACTORS							OUTPUT			CUMULATIVE FREQ		
A	B	C	D	E	F	G	1	2	3	I	II	III
-1	-1	-1	-1	-1	-1	-1	3	6	1	3	9	10
-1	-1	-1	1	1	1	1	0	3	7	0	3	10
-1	1	1	-1	-1	1	1	0	0	10	0	0	10
-1	1	1	1	1	-1	-1	0	0	10	0	0	10
1	-1	1	-1	1	-1	1	3	5	2	3	8	10
1	-1	1	1	-1	1	-1	2	8	0	2	10	10
1	1	-1	-1	1	1	-1	2	7	1	2	9	10
1	1	-1	1	-1	-1	1	0	4	6	0	4	10
-1	-1	-1	-1	-1	-1	-1	6	4	0	6	10	10
-1	-1	-1	1	1	1	1	3	4	3	3	7	10
-1	1	1	-1	-1	1	1	0	1	9	0	1	10
-1	1	1	1	1	-1	-1	0	10	0	0	10	10
1	-1	1	-1	1	-1	1	3	7	0	3	10	10
1	-1	1	1	-1	1	-1	4	5	1	4	9	10
1	1	-1	-1	1	1	-1	2	5	3	2	7	10
1	1	-1	1	-1	-1	1	1	7	2	1	8	10
-1	-1	-1	-1	-1	-1	-1	1	4	5	1	5	10
-1	-1	-1	1	1	1	1	0	6	4	0	6	10
-1	1	1	-1	-1	1	1	0	0	10	0	0	10
-1	1	1	1	1	-1	-1	0	3	7	0	3	10
1	-1	1	-1	1	-1	1	3	5	2	3	8	10
1	-1	1	1	-1	1	-1	0	5	5	0	5	10
1	1	-1	-1	1	1	-1	2	7	1	2	9	10
1	1	-1	1	-1	-1	1	0	4	6	0	4	10
-1	-1	-1	-1	-1	-1	-1	0	10	0	0	10	10
-1	-1	-1	1	1	1	1	0	7	3	0	7	10
-1	1	1	-1	-1	1	1	0	0	10	0	0	10

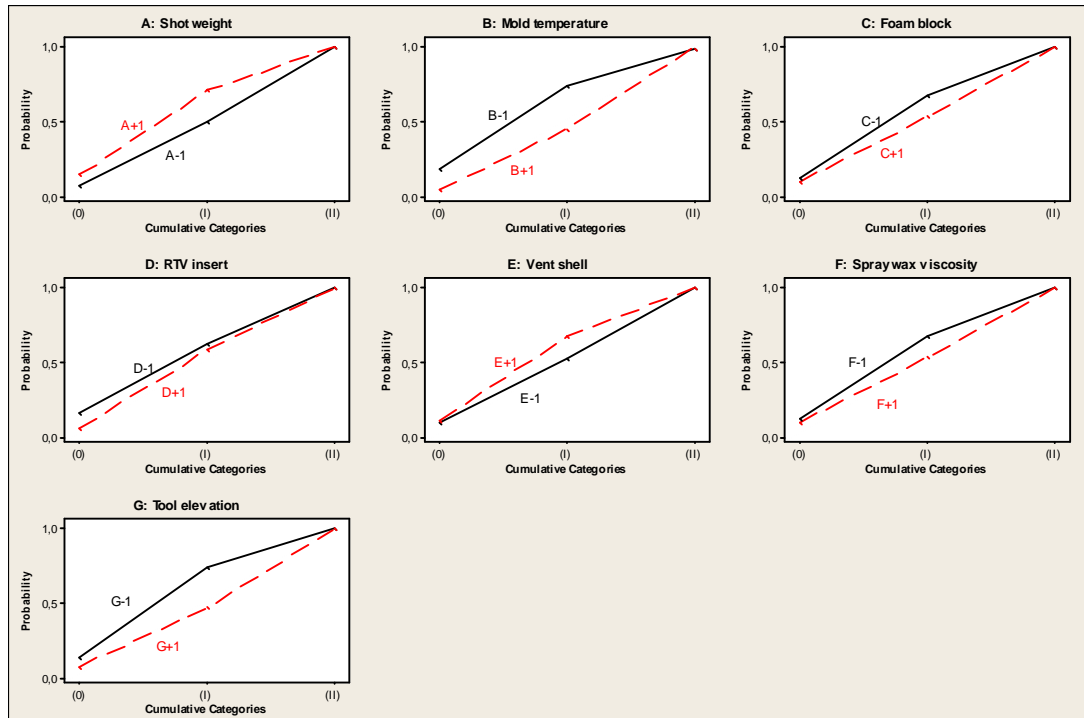


-1	1	1	1	1	-1	-1	0	9	1	0	9	10
1	-1	1	-1	1	-1	1	1	6	3	1	7	10
1	-1	1	1	-1	1	-1	1	5	4	1	6	10
1	1	-1	-1	1	1	-1	1	6	3	1	7	10
1	1	-1	1	-1	-1	1	0	3	7	0	3	10

### ANOVA Table

Source	df	S	MS	F	p
A	2	19,43348	9,716739	11,62386	0,000
B	2	42,15402	21,07701	25,21383	0,000
C	2	6,813905	3,406953	4,07564	0,017
D	2	8,115925	4,057962	4,854425	0,008
E	2	8,969063	4,484532	5,364717	0,005
F	2	6,813905	3,406953	4,07564	0,017
G	2	26,07897	13,03949	15,59877	0,000
ERROR	624	521,6207	0,835931		
TOTAL	638	640			

All factors are significant.



Probability plots of cumulative categories for factors A, B, C, D, E, F, G

Using Omega transform for probability p below:

$$w(p) = 10 \log_{10} \frac{p}{1-p}$$

Predicted omega value for category 0:

$$\omega_{A+B-C-D-E+F-G}(0) = \omega_{\mu(0)} + [\omega_{A+(0)} - \omega_{\mu(0)}] + [\omega_{B-(0)} - \omega_{\mu(0)}] + [\omega_{C-(0)} - \omega_{\mu(0)}] + [\omega_{D-(0)} - \omega_{\mu(0)}] + [\omega_{E+(0)} - \omega_{\mu(0)}] + [\omega_{F-(0)} - \omega_{\mu(0)}] + [\omega_{G-(0)} - \omega_{\mu(0)}]$$

$$\begin{aligned} &= -8,70 + [-7,32 + 8,70] + [-6,37 + 8,70] + [-8,21 + 8,70] + [-6,92 + 8,70] + [-8,51 + 8,70] + [-8,21 + 8,70] + [-7,53 + 8,70] \\ &= -0,848 \text{ dB} \end{aligned}$$

Then by inverse omega transform, the predicted probability for category 0 (good) is 0.45. Predicted category for the cumulative category (I) is predicted to be 0.96 in the same manner. Prediction is obviously 1.0 for category (II).

## APPENDIX F

### Variance Estimation for desired Probability in case study 1

In our problem we have the following model:

$$\text{logit}(\pi_i) = \beta_{i0} + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_5 X_{i5} + \beta_6 X_{i6} + \beta_7 X_{i7}$$

By Delta Method, the variance formula for  $\pi_i$  can be shown to be:

$$\text{Var}(\pi) = G \text{cov}(\pi) G^T$$

where

$$\begin{aligned} G &= [ \partial \pi / \partial \beta_0, \partial \pi / \partial \beta_1, \partial \pi / \partial \beta_2, \partial \pi / \partial \beta_3, \partial \pi / \partial \beta_5, \partial \pi / \partial \beta_6, \partial \pi / \partial \beta_7 ] \\ &= [ \pi (1 - \pi), \pi (1 - \pi) X_1, \pi (1 - \pi) X_2, \pi (1 - \pi) X_3, \pi (1 - \pi) X_5, \pi (1 - \pi) X_6, \pi (1 - \pi) X_7 ] \\ &= \pi (1 - \pi) [ 1, X_1, X_2, X_3, X_5, X_6, X_7 ] \end{aligned}$$

$\text{cov}(\pi)$  is the variance-covariance matrix of the parameters ( $\beta_j$ ),  $j=0, 1, 2, 3, 5, 6, 7$

In our estimated model:

For the factor level combination

$$[1, X_1, X_2, X_3, X_5, X_6, X_7] = [1, 1, 1, -1, -1, 1, -1], \hat{p} = 0,786435448976803$$

Then substituting  $\hat{p}$  and X to the equation for G,

$$\hat{G} = (0,167954734, 0,167954734, -0,167954734, -0,167954734, 0,167954734, -0,167954734, -0,167954734)$$

Variance-covariance matrix is calculated by Minitab Software:

$$\hat{\text{cov}}(\hat{p}) = \begin{pmatrix} 0,044787 & -0,0111 & 0,013185 & 0,009388 & -0,00942 & 0,010243 & 0,011662 \\ -0,0111 & 0,019524 & -0,00762 & -0,00682 & 0,00712 & -0,00694 & -0,00734 \\ 0,013185 & -0,00762 & 0,020474 & 0,006966 & -0,00707 & 0,007561 & 0,007815 \\ 0,009388 & -0,00682 & 0,006966 & 0,019069 & -0,0068 & 0,006894 & 0,007171 \\ -0,00942 & 0,00712 & -0,00707 & -0,0068 & 0,019134 & -0,00688 & -0,00693 \\ 0,010243 & -0,00694 & 0,007561 & 0,006894 & -0,00688 & 0,019269 & 0,006958 \\ 0,011662 & -0,00734 & 0,007815 & 0,007171 & -0,00693 & 0,006958 & 0,019763 \end{pmatrix}$$

Finally, we can find  $\hat{Var}(\hat{\rho})$  by substituting  $\hat{G}$  and  $\hat{cov}(\hat{\rho})$  to the equation for  $\hat{Var}(\hat{\rho})$ :

$$\hat{Var}(\hat{\rho}) = 0.006934$$

## APPENDIX G

### Confidence Interval (CI) Estimation for $\hat{P}$ (Y=j)

Since  $\hat{P}$  (Y=j) values are not normally distributed, we make a transformation for  $\hat{P}$ .

Let  $q = \ln(p / (1 - p))$ ,

The transformed CI =  $(L^*, U^*)$  will be:

$$L^* = \hat{q} - Z_{\alpha/2} \sqrt{\hat{Var}(\hat{q})}$$

$$U^* = \hat{q} + Z_{\alpha/2} \sqrt{\hat{Var}(\hat{q})}$$

Where,

$$Var(q) = \left( \frac{\partial q}{\partial p} \right)^2 Var(p) = \frac{1}{p^2(1-p)^2} Var(p) \text{ by Delta Method.}$$

Since in our problem,  $\hat{P} = 0.79$  and  $\hat{Var}(\hat{p}) = 0.0069$ ,

$$\hat{q} = \ln(0.79 / (1 - 0.79)) = 1.325$$

$$\hat{Var}(\hat{q}) = 1/(0.79^2(1-0.79^2))*0.0069 = 0.029$$

By choosing  $\alpha=0.05$

$$L^* = 1.325 - 1.96\sqrt{0.029} = 0.343$$

$$U^* = 1.325 + 1.96\sqrt{0.029} = 2.306$$

By backward transformation % 95 CI for  $\hat{P}$  is:

$$(L, U) = \left( \frac{e^{L^*}}{1 + e^{L^*}}, \frac{e^{U^*}}{1 + e^{U^*}} \right) = \left( \frac{e^{0.343}}{1 + e^{0.343}}, \frac{e^{2.306}}{1 + e^{2.306}} \right) = (0.585, 0.909)$$

## APPENDIX H

### Minitab Logistic Regression Output for Case Study 2

#### Binary Logistic Regression: N-defect; Total\_casted versus A; B; E; F; G

Link Function: Logit

##### Response Information

Variable	Value	Count
N-defect	Success	219
	Failure	12198
Total_casted	Total	12417

\* NOTE \* 135 cases were used  
 \* NOTE \* 3 cases contained missing values

##### Logistic Regression Table

Predictor	Coef	SE_Coef	Z	P	Odds	95% CI	
					Ratio	Lower	Upper
Constant	-4,32801	0,0995319	-43,48	0,000			
A	0,162070	0,160760	1,01	0,313	1,18	0,86	1,61
B	0,632905	0,143828	4,40	0,000	1,88	1,42	2,50
E	-0,0543835	0,126899	-0,43	0,668	0,95	0,74	1,21
F	-0,503251	0,217173	-2,32	0,020	0,60	0,39	0,93
G	-0,437856	0,231717	-1,89	0,059	0,65	0,41	1,02
A*F	0,365270	0,104453	3,50	0,000	1,44	1,17	1,77
E*F	-0,338806	0,0925527	-3,66	0,000	0,71	0,59	0,85

Log-Likelihood = -1041,585

Test that all slopes are zero: G = 119,480, DF = 7, P-Value = 0,000

##### Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	152,946	125	0,045
Deviance	162,017	125	0,014
Hosmer-Lemeshow	3,986	8	0,858

Table of Observed and Expected Frequencies:  
 (See Hosmer-Lemeshow Test for the Pearson Chi-Square Statistic)

Value Group	Success		Failure		Total
	Observed	Expected	Observed	Expected	
1	6	6,8	1311	1310,2	1317
2	11	9,2	1260	1261,8	1271
3	12	13,2	1238	1236,8	1250
4	19	17,1	1236	1237,9	1255
5	26	20,3	1283	1288,7	1309
6	25	25,3	1463	1462,7	1488
7	18	23,8	1241	1235,2	1259

8	27	29,1	1306	1303,9	1333
9	37	36,6	1216	1216,4	1253
10	38	37,7	644	644,3	682

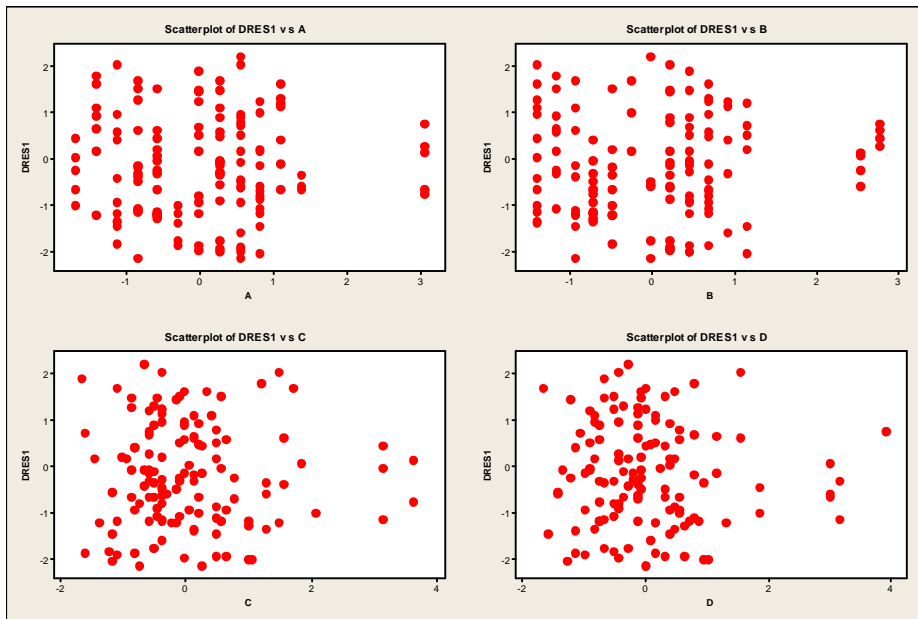
Measures of Association:  
 (Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	1657005	62,0	Somers' D	0,32
Discordant	802307	30,0	Goodman-Kruskal Gamma	0,35
Ties	212050	7,9	Kendall's Tau-a	0,01
Total	2671362	100,0		

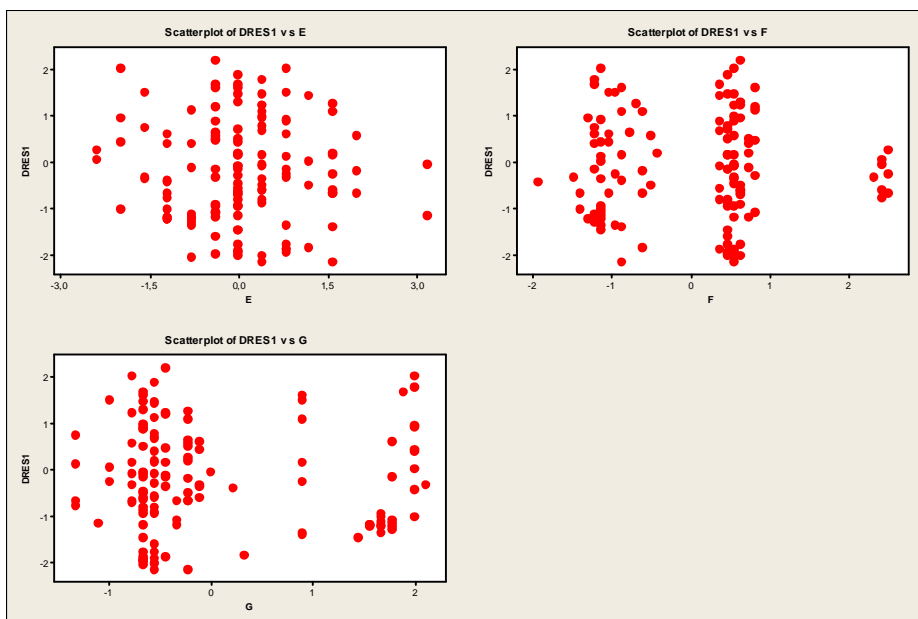


## APPENDIX I

### Residual Plots for N-Gas Defect Problem



Deviance Residuals vs. Factors A,B,C,D plot



Deviance Residuals vs. Factors E,F,G plot

## APPENDIX J

**Classification Table for N-Gas Defect Problem For  $c=0.02$**

Observation No	Total No of Observations	Observation N-defect	Prediction $P^{\wedge}(Y=1)$	Correct	Incorrect
1	113	1	0,019	112	1
2	110	2	0,020	108	2
3	103	0	0,019	103	0
4	110	0	0,018	110	0
5	117	3	0,017	114	3
6	115	4	0,013	111	4
7	30	0	0,016	30	0
8	80	2	0,016	78	2
9	115	3	0,014	112	3
10	110	4	0,022	4	106
11	115	4	0,025	4	111
12	110	5	0,025	5	105
13	28	0	0,019	28	0
14	79	1	0,025	1	78
15	110	1	0,020	1	109
16	110	0	0,017	110	0
17	105	0	0,017	105	0
18	112	0	0,016	112	0
19	111	3	0,016	108	3
20	110	2	0,022	2	108
21	110	1	0,019	109	1
22	110	2	0,019	108	2
23	108	5	0,016	103	5
24	110	3	0,017	107	3
25	109	1	0,015	108	1
26	160	0	0,014	160	0
27	100	5	0,016	95	5
28	115	2	0,018	113	2
29	108	4	0,023	4	104
30	108	1	0,018	107	1
31	109	2	0,020	2	107
32	105	1	0,043	1	104
33	165	7	0,032	7	158
34	170	3	0,036	3	167
35	170	2	0,030	2	168
36	104	1	0,021	1	103
37	110	2	0,019	108	2

38	110	0	0,014	110	0
39	110	1	0,014	109	1
40	110	1	0,015	109	1
41	110	1	0,016	109	1
42	110	2	0,024	2	108
43	100	1	0,014	99	1
44	180	4	0,018	176	4
45	104	2	0,017	102	2
46	104	1	0,026	1	103
47	96	1	0,018	95	1
48	100	0	0,017	100	0
49	109	4	0,016	105	4
50	110	0	0,015	110	0
51	106	4	0,019	102	4
52	69	3	0,016	66	3
53	110	1	0,020	109	1
54	110	1	0,018	109	1
55	130	0	0,012	130	0
56	90	3	0,013	87	3
57	110	0	0,015	110	0
58	110	2	0,016	108	2
59	69	2	0,020	67	2
60	110	5	0,017	105	5
61	145	2	0,017	143	2
62	130	8	0,040	8	122
63	168	6	0,033	6	162
64	116	1	0,023	1	115
65	97	1	0,025	1	96
66	104	5	0,040	5	99
67	110	7	0,033	7	103
68	110	3	0,021	3	107
69	119	4	0,026	4	115
70	96	3	0,033	3	93
71	109	7	0,041	7	102
72	8	1	0,018	7	1
73	8	0	0,022	0	8
74	8	0	0,006	8	0
75	8	1	0,054	1	7
76	8	1	0,115	1	7
77	8	0	0,025	0	8
78	8	0	0,028	0	8
79	8	0	0,027	0	8
80	8	1	0,230	1	7
81	8	7	0,840	7	1
82	7	1	0,029	1	6
83	8	2	0,163	2	6

84	8	0	0,000	8	0
85	8	1	0,077	1	7
86	8	0	0,074	0	8
87	8	1	0,108	1	7
88	8	0	0,004	8	0
89	7	0	0,031	0	7
90	8	0	0,007	8	0
91	110	0	0,015	110	0
92	112	2	0,015	110	2
93	110	3	0,013	107	3
94	109	3	0,012	106	3
95	110	2	0,020	2	108
96	110	1	0,014	109	1
97	110	1	0,016	109	1
98	110	2	0,012	108	2
99	110	0	0,020	0	110
100	110	3	0,010	107	3
101	104	0	0,009	104	0
102	105	2	0,008	103	2
103	110	2	0,011	108	2
104	147	2	0,012	145	2
105	70	2	0,007	68	2
106	98	0	0,009	98	0
107	105	1	0,012	104	1
108	112	1	0,006	111	1
109	105	1	0,009	104	1
110	105	3	0,008	102	3
111	110	2	0,009	108	2
112	109	1	0,006	108	1
113	16	0	0,005	16	0
114	110	2	0,003	108	2
115	94	0	0,005	94	0
116	16	0	0,013	16	0
117	108	1	0,003	107	1
118	96	1	0,005	95	1
119	8	0	0,009	8	0
120	70	2	0,007	68	2
121	102	0	0,006	102	0
122	8	0	0,006	8	0
123	105	0	0,006	105	0
124	110	1	0,010	109	1
125	110	0	0,007	110	0
126	110	0	0,006	110	0
127	113	0	0,009	113	0
128	94	0	0,006	94	0
129	104	0	0,007	104	0

130	16	0	0,005	16	0
131	100	0	0,007	100	0
132	110	0	0,007	110	0
133	104	0	0,007	104	0
134	104	0	0,005	104	0
135	104	0	0,006	104	0
136	125	0	0,006	125	0
137	49	0	0,008	49	0
138	122	0	0,007	122	0
			<b>TOTAL</b>	<b>9290</b>	<b>3284</b>