ISMSIT 2017

1st International Symposium on Multidisciplinary Studies and Innovative Technologies Proceedings
(November 2-4, 2017, Tokat, Turkey)
Symposium homepage: www.ismsitconf.org

The Effects of Noise Distributions on Robust Design

Melis Zeybek[1*] and Onur Köksoy[2]

[1,2]*Faculty of Science, Department of Statistics, Ege University, Turkey*
* *melis.zeybek@ege.edu.tr*

Abstract – Taguchi's robust parameter design focuses on reducing the system variability caused by the noise factors. The results obtained under unrealistic assumptions about the noises may mislead the practitioners when it comes to improving quality in robust design. For example, many hydrological data and the multi-path fading of a signal in wireless communication systems are positively skewed and cannot be modelled by any normal distribution. In this study, the case where the quality characteristic is affected by two noise factors is taken into account - one follows a normal distribution and the other one has a gamma distribution. The true effects of noise distributions are investigated and a simulation study is presented to illustrate our findings and quantify the effects of noise distributions in robust design. Additionally, a new density function is proposed.

Keywords – Noise distribution; Gamma Noise; Robust parameter design; Response surface methodology.

I. INTRODUCTION

The term robust parameter design (RPD) was introduced in the early 1980s by Taguchi and has received much attention by many quality engineers in different fields. Although scientists have adopted Taguchi's RPD approach on quality, statistical communities have criticized his experimental methodologies and analysis techniques. Consequently, new methodologies have been proposed. The response surface methodology (RSM), which was first developed by [1], was revisited and popularized in the early 1990s. Following this article, current literature includes several other methods for RSM, for example, [2, 3, 4, 5, 6, 7, 8, 9].

The probability distribution of the experimental data plays an important role in response surface methodology. The belief that 'the normality assumption' is a robust assumption may not be true and there are many cases where it does not apply to real-world problems. The distribution of noise factors has an influence on a response distribution. In fact, [10] and [11] note that fluctuations in noise factors may not be well described by any symmetric distribution in real life problems. In many situations, noise factors have skewed distribution, such as the gamma, beta or exponential. For example, environmental conditions such as streamflow, humidity, temperature, and rainfall, which are general examples of noise factors. Studies on the daily or monthly rainfall have indicated that the real distribution can be defined by a gamma distribution; see, [12] and [13]. If the noise factors are non-normal, then its structure affects the distribution of responses.

In this study, the case where the quality characteristic is affected by two noise factors is taken into account - one follows a normal distribution and the other one has a gamma distribution. The true effects of noise distributions are investigated by a simulation study. Then, a new density function is proposed and the mean and variance models of the distribution are presented.

II. SIMULATION STUDY

Although non-normal noise factors are the reality of a robust design, what is surprising is that the real-life experimental designs under non-normal responses are limited in the current literature. In this context, we concentrated on a simulation study. In the simulation process, an experiment, which is presented by the study [14], as an initial design is defined. In the design, there are two controllable factors (x_1, x_2) and two noise factors (z_1, z_2). Assume that the first noise has normal distribution, on the other hand the second one has gamma. Under these considerations, we simulated the initial design assuming $Z_1 \sim N(0,1)$, $Z_2 \sim G(1,2)$ and $\varepsilon^* \sim N(0,1)$ using the following simulation model,

$$Y^* = 24.472 + 6.89x_1 - 9.11x_2 + 4.94Z_1 + 3.52x_1x_2 + 3.23x_1Z_1 + 1.88x_1Z_2 + \varepsilon^*$$

A large number of potential responses, say 10,000, were randomly generated using JMP. The design of the simulated experiments is constructed and given in Table 1. Figure 1 illustrates the histogram of 10,000 simulated responses for the first experimental run.

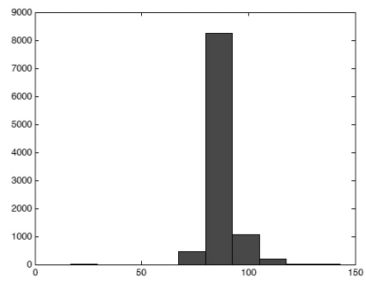

Figure 1. The histogram of 10,000 simulated responses for the first experimental run.

Table 1. The design of simulated experiments.

u	x_1	x_2	z_1	z_2	y
1	0	0	-0.002	2.016	85.991
2	0	0	-0.004	1.978	96.002
3	0	0	0.002	1.986	86.022
4	0	0	0.012	1.988	86.065
5	0	0	-0.008	1.985	85.960
6	-1	-1	-0.008	1.972	83.059
7	-1	-1	0.004	2.005	83.123
8	-1	-1	0.020	1.998	83.284
9	-1	-1	0.004	1.994	83.257
10	-1	1	0.018	1.978	72.741
11	-1	1	0.013	1.967	72.736
12	-1	1	-0.023	2.016	72.839
13	-1	1	0.015	2.011	72.805
14	1	-1	0.006	2.001	62.614
15	1	-1	0.014	2.012	62.605
16	1	-1	-0.005	1.996	62.447
17	1	-1	0.008	2.010	62.630
18	1	1	-0.016	2.015	97.581
19	1	1	-0.004	1.989	97.579
20	1	1	0.002	2.012	97.555
21	1	1	0.008	1.978	97.673

As seen from Figure 1, the simulated responses have a skewed shape. Although the error terms and the first noise distributed normally, the distribution structure of the gamma noise plays an important role in the shape of responses such that the histogram of simulated responses indicates a right-skewed structure.

III. A NEW PROBABILITY DENSITY FUNCTION

Consider the quality of a system with k controllable factors $(x_1, \dots, x_k)$ and two noise factor (Z_1, Z_2), thus, the response can be modelled by the second order response surfaces as follows,

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i{}^2 + \sum_{i<t}^{k} \sum \beta_{it} x_i x_t + \sum_{i=1}^{2} \gamma_i Z_i + \sum_{i=1}^{k} \sum_{j=1}^{2} \gamma_{ij} x_i Z_j + \varepsilon \quad (1)$$

where $\varepsilon \sim NID(0, \sigma_\varepsilon^2)$ is random error and $Z_1 \sim NID(\mu_{z_1}, \sigma_{z_1}^2)$, and $Z_2 \sim Gamma(\theta, \kappa)$.

The conditional distribution of Y given $Z_1 = z_1,\ Z_2 = z_2$ is obtained as,

$$f_Y(y|Z_1 = z_1, Z_2 = z_2) \sim N\left(\mu_x + \sum_{i=1}^{2} \gamma_i Z_i + \sum_{i=1}^{k} \sum_{j=1}^{2} \gamma_{ij} x_i Z_j\, , \sigma_\varepsilon^2\right) \quad (2)$$

where $\mu_x = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i{}^2 + \sum_{i<t}^{k} \sum \beta_{it} x_i x_t$.

Using Equations (1) and (2), the proposed probability density function of the response given by Equation (1) can be defined by,

$$f_Y(y) = \int_0^{+\infty} \int_{-\infty}^{+\infty} . f_Y(y|Z_1 = z_1,\ Z_2 = z_2) f(z_1) f(z_2)\, dz_1\, dz_2$$

$$= \frac{(\sigma^*)^{\theta-1}}{\sqrt{2\pi}\kappa^\theta \left(\gamma_2 + \sum_{i=1}^{k} \gamma_{i2} x_i\right)^\theta} \exp\left(-\frac{\sigma^{*2}}{2\kappa^2 \left(\gamma_2 + \sum_{i=1}^{k} \gamma_{i2} x_i\right)^2}\right) . \exp\left(-\frac{t^2}{4} - \frac{t\sigma^*}{\kappa\left(\gamma_2 + \sum_{i=1}^{k} \gamma_{i2} x_i\right)}\right) D_{-\theta}(-t) \quad (3)$$

for $\theta = 1,2,3, \dots$, where,

$$t = \frac{y - \mu^*}{\sigma^*},$$

$$\mu^* = \mu_x + \mu_{z_1}\left(\gamma_1 + \sum_{i=1}^{k} \gamma_{i1} x_i\right) + \frac{\sigma^*}{\kappa\left(\gamma_2 + \sum_{i=1}^{k} \gamma_{i2} x_i\right)},$$

$$\sigma^* = \sqrt{\sigma_{z_1}^2 \left(\gamma_1 + \sum_{i=1}^{k} \gamma_{i1} x_i\right)^2 + \sigma_\varepsilon^2}$$

and $D_{-\theta}(-t)$ is a parabolic cylinder function. Also, the function given in Equation (3) satisfies the following conditions: $f_Y(y) \geq 0$ and $\int_{-\infty}^{\infty} f_Y(y) dy = 1$ for $\theta = 1,2,3$. As a result, Equation (3) is the density function for a given response defined in Equation (1) under the mentioned assumptions.

The expected value and the variance of Equation (3) are obtained as follows:

$$E(Y) = \mu_x + \mu_{z_1}\left(\gamma_1 + \sum_{i=1}^{k} \gamma_{i1} x_i\right) + \mu_{z_2}\left(\gamma_2 + \sum_{i=1}^{k} \gamma_{i2} x_i\right) \quad (4)$$

$$V(Y) = \sigma^{*2} + \sigma_{z_2}^2 \left(\gamma_2 + \sum_{i=1}^{k} \gamma_{i2} x_i\right)^2 \quad (5)$$

Figure 2 illustrates the graph of the proposed $f_Y(y)$ under $\varepsilon \sim NID(0,4)$ and $E(Y) = 6$ and $V(Y) = 22$. As seen from Figure 3, the most conspicuous property is that the proposed has a right-skewed shape as does the gamma density; however, it also takes negative values. Effective dominance of the right-skewed noise factor on the responses explains the distributional structure of the proposed density, despite the fact that the second noise and random errors are distributed normally.

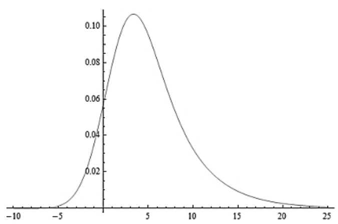

Figure 2. The graph of the proposed $f_Y(y)$ under $Z_1 \sim NID(1,2)$, and $Z_2 \sim Gamma(1,4)$.

IV. RESULTS AND DISCUSSION

We took into account the two noise factor case where one noise has a normal distribution and the other one has have a gamma distribution. A new density function is proposed under the mentioned assumption. Therefore, this study offers a useful reference to practitioners in terms of providing more engineering understanding about the process. This proposed pdf can be used for quality improvement techniques such as in loss function based approaches, or to estimate the distributional properties of the robust design, i.e. system mean and the variance.

REFERENCES

[1] Box, G.E.P. and Wilson, K.B., "On the experimental attainment of optimum conditions", Journal of the Royal Statistical Society, vol. 13, 1-45, 1951.

[2] Vining, G.G. and Myers, R.H., "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach", Journal of Quality Technology, vol. 22, 38-45, 1990.

[3] Del Castillo, E. and Montgomery, D.C., "A Nonlinear Programming Solution to the Dual Response Problem", Journal of Quality Technology, vol. 25, 199-204, 1993.

[4] Lin, D.K.J. and Tu, W., "Dual Response Surface Optimization", Journal of Quality Technology, vol. 27, 34-39, 1995.

[5] Copeland, K.A. and Nelson, P.R., "Dual Response Optimization via Direct Function Minimization", Journal of Quality Technology, vol. 28, 331-336, 1996.

[6] Köksoy, O. and Doganaksoy, N., "Joint Optimization of Mean and Standard Deviation in Response Surface Experimentation", Journal of Quality Technology, vol. 35, 239-252, 2003.

[7] Köksoy, O., "Multiresponse Robust Design: Mean Square Error (MSE criterion)", Applied Mathematics and Computation, vol. 175, 1716-1729, 2006.

[8] Köksoy, O. and Fan, S.S. "An Upside-down Normal Loss Function-based Method for Quality Improvement", Engineering Optimization, vol. 44, 935-945, 2012.

[9] Zeybek M. and Köksoy, O. "Optimization of Correlated Multi-response Quality Engineering by the Upside-down Normal Loss Function", Engineering Optimization, vol. 48, 1419-1431, 2016.

[10] Abraham A.A., Robinson T.J. and Anderson-Cook, C.M., "A graphical approach for assessing optimal operating conditions in robust design", Quality Technology and Quantitative Management, vol. 6, 235-253, 2009.

[11] Brenneman, W.A. and Myers, W.R., "Robust parameter design with categorical noise variables", Journal of Quality Technology, vol. 35, 335–341, 2003.

[12] Hanson L. and Vogel, R., " The probability distribution of daily rainfall in the United States", World Environmental and Water Resources Congress. Honolulu, Hawaii, 2008.

[13] Sharma, M.A. and Singh, J.B., "Use of probability distribution in rainfall analysis", New York Science Journal, vol. 3, 40-49, 2010.

[14] Myers, R.H, Khuri, A.I. and Vining, G., "Response surface alternatives to the Taguchi robust parameter design approach", The American Statistician, 46, 131-139, 1998.