The Generalized β-Method in Taguchi Experiments

J. Kunert and F. Lehmkuhl

ABSTRACT: In the presence of variability control factors in Taguchi experiments, then the original β-method (Logothetis, 1990) is liable to lead to wrong transformations. We propose a generalization of the β-method which should lead to correct transformations, even if there is a variability control factor which also influences the mean.

KEYWORDS: Design factor, target control factor, transformation, variability control factor

1 Introduction

In Taguchi experiments the design factors are separated into variability control factors, target control factors and neutral factors (see, e.g. Logothetis and Wynn, 1989, p.244). The variability control factors are the most important factors, they are used to reduce the variance of a product. It was pointed out by Box (1988) that correct identification of these factors requires appropriate transformation of the response y. Then the variability control factors influence the variance of the transformed variable

$$u = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{if} \quad \lambda \neq 0 \ , \\ \log y, & \text{if} \quad \lambda = 0 \ , \end{cases}$$

where λ is the parameter of transformation, which has to be determined from the data. Box (1988) proposed to identify the parameter λ with the help of the λ-plot.

Logothetis (1990) proposed an alternative method to identify the appropriate transformation which is called β-method. The present paper examines the β-method in some detail. We show with the help of an artifical example that the parameter λ estimated from this method can be systematically wrong. This has also been observed by Engel (1992) who proposed a generalization of the β-method. We point out some problems with Engel's generalization. This is done with the help of the same artifical example. Finally, we propose a new generalization of the β-method, which avoids these problems.

2 Description of the model

We assume that the observed variable y is a transformation of an ideal variable z which follows a simple model

$$z = \mu + \sum_{j \in T} x_j \alpha_j + \sum_{j \in V} x_j \alpha_j + \prod_{j \in V} e^{x_j \delta_j} \epsilon, \tag{1.1}$$

where T is the set of all target control factors and V is the set of all variability control factors. The number x_j denotes the setting of the factor j. We assume that we have a two-level design, then without loss of generality x_j is either -1 or +1. The parameter α_j gives the effect of factor j on the mean of z, while δ_j is the effect on the variance of z. The parameter μ is the overall mean and ϵ is a random variable with mean 0 and variance σ^2. For each target control factor j we assume that $\alpha_j \neq 0$, while for each variability control factor j we assume that $\delta_j \neq 0$. It is possible, however, that a variability factor has no influence on the mean, i.e. that $\alpha_j = 0$.
Finally, we assume that there are more target control factors than variability control factors, i.e. $|T| > |V|$ unless both are zero.
Things would be fairly easy if we could observe z directly. However, we cannot observe z but can only observe y, where we assume that y is a transformation of z of the form

$$y = \begin{cases} (\lambda z + 1)^{1/\lambda}, & \text{if} \quad \lambda \neq 0 \ , \\ e^z \ , & \text{if} \quad \lambda = 0 \ . \end{cases}$$

Note that this is the usual family considered for the Box-Cox transformations. Using first-order Taylor approximation, we approximate the expectation and variance of y to be

$$\mathrm{E}y \approx \begin{cases} (\lambda \mathrm{E}z + 1)^{1/\lambda}, & \text{if} \quad \lambda \neq 0 \ , \\ e^{\mathrm{E}z}, & \text{if} \quad \lambda = 0 \ , \end{cases}$$

$$\mathrm{Var}y \approx \left\{ \begin{array}{l} (\lambda \mathrm{E}z + 1)^{2\frac{1-\lambda}{\lambda}} \mathrm{Var}z \\ (e^{\mathrm{E}z})^2 \mathrm{Var}z \end{array} \right\} \approx (\mathrm{E}y)^{2(1-\lambda)} \mathrm{Var}z.$$

It follows that the logarithm of the square root of $\mathrm{Var}y$ is a linear function of the logarithm of $\mathrm{E}y$, more precisely

$$\log(\sqrt{\mathrm{Var}y}) \approx (1-\lambda)\log(\mathrm{E}y) + \log(\sigma) + \sum_{j \in V} x_j \delta_j . \tag{1.2}$$

The slope $\beta = 1 - \lambda$ of this linear relation can be used to retransform the observations y to become the nonobservable z, namely

$$z = \begin{cases} \frac{y^{1-\beta}-1}{1-\beta}, & \text{if} \quad \beta \neq 1 \\ \log y, & \text{if} \quad \beta = 1 \ . \end{cases}$$

The β-method (and all generalizations of the β-method) estimates β from the empirical means m and variances s^2 observed in an experiment.
For simplicity, we assume that these data were derived in an n-run orthogonal design of the design factors (Taguchi's Inner Array), where each factor is observed at two levels. For each combination of the design factors in the experiment we have repeated observations. These may either be produced by observing a factorial experiment of some noise factors (Taguchi's Outer Array) or by simply observing repeatedly.
Then for the i-th $(1 \leq i \leq n)$ combination of the design factors, we observe the empirical mean m_i and the empirical variance s_i^2 of y.
The original β-method simply regresses the $\log s_i$ on the $\log m_i$, using the estimated slope $\hat{\beta}_0$ of the regression as an estimate for β. More precisely, we estimate β_0 in the model

$$\log s_i = \nu + \beta_0(\log m_i) + error. \tag{1.3}$$

It was first observed by Engel (1993) that this method is not consistent as it neglects the term $\sum_{j \in V} x_j \delta_j$ from equ. (1.2). Hence, even if we were able to estimate $\mathrm{Var} y$ and $\mathrm{E} y$ without error, the original β-method would in general give a $\hat{\beta}_0$ which is not equal to the true β.

3 Engel's (1992) generalization of the β-method

As an example to see the non-consistency of the original β-method, consider the case that there is only one variability control factor (factor 1, say) and there are two target control factors (say factors 2 and 3). Further assume that in model (1.1) we have

$$\mu = 100, \alpha_1 = 10 \ , \alpha_2 = 2 \ , \alpha_3 = 1 \ , \delta_1 = 10 \ , \sigma = 1$$

and that the observed variable y is

$$y = e^z$$

i.e. $\lambda = 0$ and $\beta = 1$. Hence, the appropriate transformation would be to take the logarithm of y.
We assume that the design is a fractional factorial design where the interaction between factors 2 and 3 is not confounded with factor 1. Then we will observe a set $I(1)$ of $n/8$ runs i such that

$$x_1 = 1 \ , x_2 = 1 \ , x_3 = 1 \quad \text{for all} \quad i \in I(1).$$

Since in our example all other factors are neutral factors, we have for all factor combinations in $I(1)$ that the expectation of y is the same, namely

$$\mathrm{E} y \approx e^{100 + x_1 10 + x_2 2 + x_3 1} = e^{113}.$$

It follows that $\log(\mathrm{E}y) = 100 + x_1 10 + x_2 2 + x_3 1 = 113$. Similarly, for all factor combinations $i \in I(1)$ we have from equ. (1.2) that

$$\log(\sqrt{\mathrm{Var}y}) \approx 0 + 10\, x_1 + \log(\mathrm{E}y) \approx 100 + 20x_1 + 2\, x_2 + 1\, x_3 = 123.$$

There are 8 combinations of factors 1, 2 and 3, all of which appear in $n/8$ runs of the inner array. Ordering the runs of the inner array in such a way that the $n/8$ runs with the same setting of factors 1,2 and 3 come in a row, then the columns of the design matrix which correspond to the mean and factors 1, 2 and 3 can be written as

$$X = \begin{bmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & -1 & +1 \\ 1 & +1 & -1 & -1 \\ 1 & -1 & +1 & +1 \\ 1 & -1 & +1 & -1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \otimes 1_{n/8},$$

where $\otimes$ is the Kronecker symbol and $1_{n/8}$ is the $(n/8)$-vector of ones. The vector of logarithms of the expected values (in the same ordering) equals M, while the vector of the logarithms of the square roots of the variances equals S, where

$$M = \begin{bmatrix} 113 \\ 111 \\ 109 \\ 107 \\ 93 \\ 91 \\ 89 \\ 87 \end{bmatrix} \otimes 1_{n/8}, \text{ and } \quad S = \begin{bmatrix} 123 \\ 121 \\ 119 \\ 117 \\ 83 \\ 81 \\ 79 \\ 77 \end{bmatrix} \otimes 1_{n/8}.$$

Assume we can estimate the expectations and variances without error, that is assume that the vector of the observed $\log(m_i)$ equals M, while the vector of the observed $\log(s_i)$ equals S. Plotting $\log(s_i)$ against $\log(m_i)$ (mean variance plot), we then get Figure 1, which also displays the regression line derived from the original β-method. This line has slope 1.95 which means that the original β-method makes us transform the observations to

$$u^{(1)} = \frac{y^{-0.95} - 1}{-0.95} = \frac{e^{-0.95z} - 1}{-0.95}.$$

For this transformed variable $u^{(1)}$ we can approximate the variance by

$$\mathrm{Var}u^{(1)} \approx (e^{-0.95\mathrm{E}z})^2 \mathrm{Var}z.$$

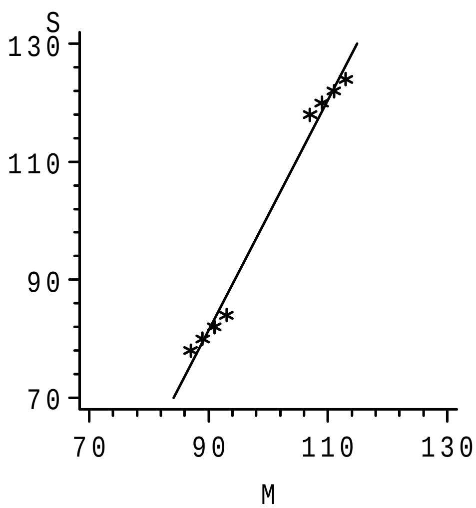

FIGURE 1. The β-plot for the example

Engel (1992) proposed to use an iterative version of the β-method. This method, in principle, works as follows:

- As a first step do the original β-method and transform the observations to the $u^{(1)}$ proposed by $\hat{\beta}_0$, the slope of the regression line.
- For the transformed variable $u^{(1)}$ identify all factors which have an influence on the variance. Denote the set of all such factors by U.
- As a second step determine the estimate $\hat{\beta}_E$ of the slope in the model

$$\log s = \nu + \beta_E(\log m) + \sum_{j \in U} x_j \gamma_j + error \tag{1.4}$$

- Transform the observations to

$$u^{(2)} = \frac{y^{1-\hat{\beta}_E} - 1}{1 - \hat{\beta}_E}$$

and determine target control factors and variability control factors in the model for $u^{(2)}$.

It can be shown that (at least if the estimated s_i and m_i are without error) the true variability control factors are within U. However, the effect of the true variability control factor on the variance of $u^{(1)}$ can be smaller than the effect of the target control factors. If the true variability control factors are in U then Engel (1992) claims that his procedure is consistent. However this is not generally true.

To see this, we return to our artificial example. For this example the set U contains all three factors 1, 2 and 3. Hence, in step 2 of Engel's procedure, we estimate β_E from the model

$$\log s = \nu + x_1\gamma_1 + x_2\gamma_2 + x_3\gamma_3 + \beta_E(\log m) + error.$$

With the vectors and matrices introduced above, this can be written in vector notation as

$$S = X \begin{bmatrix} \nu \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} + M\beta_E + error.$$

Note that (due to the fact that we estimate without error in our example) M is in the column span of X. Hence, β_E is not estimable and Engel's procedure cannot be done!

Of course, if there is error in the estimation of the mean, then the vector of observed logarithms of the mean would not be in the column span of X. However, this would not help: after correcting for the γ_j of all factors with an effect on the mean, then all that is left of the vector of means is noise. So we can do step 2 of Engel's procedure only if not all factors with an effect on the mean are in U. If, however, we would omit one factor from U in our example, then we surely would omit factor 1, which has the smallest

apparent effect on the variance. In that case, however, we get another wrong transformation $\hat{\beta}_E = 2$.
It should be pointed out, however, that in many cases the variability control factor will have a large effect on the variance of the $u^{(1)}$. The surprising fact that this effect of the variability control factor almost disappears in our example comes from the large effect of the variability control factor on the mean. So our example was extreme. However, if the effect of the true variability control factor on the variance of the $u^{(1)}$ is large compared to the effect of the target control factors, this means that the original β-method works quite well already.
Finally, it should be pointed out that in our example the vector containing $\log(\mathrm{E}y)$ is in the column span of X only because the true β is 1 and therefore $\log(\mathrm{E}y) \approx \mathrm{E}z$. For $\beta \neq 1$ we have

$$\log(\mathrm{E}y) \approx \frac{1}{1-\beta}\log\{(1-\beta)Ez + 1\}$$
$$\approx \frac{1}{1-\beta}\log\{(1-\beta)(\mu + \sum_{j\in T\cup V} x_j\alpha_j) + 1\},$$

and $\log(\mathrm{E}y)$ is not a linear function of the $x_j, j \in T\cup V$. However, this does not help very much. At least for large μ, $\log(\mathrm{E}y)$ can be further approximated by

$$\log(\mathrm{E}y) \approx \frac{1}{1-\beta}\log((1-\beta)\mu + 1) + \sum_{j\in T\cup V} x_j \frac{\alpha_j}{(1-\beta)\mu + 1}.$$

This approximation is a linear combination of the $x_j, \quad j \in T\cup V$.

4 A new generalization of the β-method

In what follows we assume that not all variability control factors have the same δ_j. If there are variability control factors, then most likely there will be one $j^\star$ which has the largest δ_j. We think it is important to find a transformation for which this $j^\star$ has a large effect on the variance of the transformed variable u.
Assume that there is just one variability control factor. Then from equ. (1.2) there are two different lines in the mean variance plot, both having the same slope β but different intersections with the vertical axis. If there is more than one variability control factor, then there are more than two lines, for every combination of the variability control factors which appears in the design there is one. Fortunately, as can be seen from equ. (1.2), all have the same slope. For every variability control factor j we have two groups of lines, one in which factor j is at level $+1$ and one in which this

factor is at level -1. The mean difference between the groups depends on the size of δ_j, the effect of the factor j on the variance.
For every design factor j fit the following model in the mean variance plot:

$$\log s = \nu + x_j\gamma_j + \beta_j \log m + error. \tag{1.5}$$

Then the R^2 of the fit depends on the size of the γ_j of the factors which are neglected. For the factor $j^\star$ with the largest effect on the variability, the R^2 will be largest.
Hence, we transform the variables with the parameter $\hat{\beta}_{j(R)}$, where $j(R)$ is the factor for which equation (1.5) gives the largest R^2.
It is evident, that this method leads to the correct transformation in our artifical example. In fact, whenever there is at most one variability control factor, then the new generalization of the β-method is consistent.
In the present paper we do not go into technical details of the performance of this generalized β-method. These will be reported elsewhere, see Lehmkuhl (1998). However, we want to point out, why we fit only two parallel lines, i.e. use only one variability control factor.
The first reason was already present in the example. With too many parameters in the model the estimates from the model get instable, in fact we can have nonestimability.
The second reason is that if we allow for too many variability control factors, then the model can jump into the wrong direction. To see this, consider our artifical example again. If we allow for two factors j_1 and j_2 in equation (1.5), we might fit factors 2 and 3, and get $\hat{\beta} = 2$ with an R^2 of 1! Such a wrong transformation can, however, only get such a high R^2 if we fit more than one factor in equation (1.5).

References

Box, G.E.P. (1988). Signal-to-Noise Ratios, Performance Criteria and Transformations. *Technometrics* **30**, 1 - 17

Engel, J. (1992). Modelling Variation in Industrial Experiments. *Applied Statistics* **41**, 579 - 593

Lehmkuhl, F. (1998). Dissertation, University of Dortmund (in preparation)

Logothetis, N. (1990). Box-Cox Transformations and the Taguchi-Method. *Applied Statistics*, **39**, 31 - 48

Logothetis, N. and Wynn, H.P. (1989). *Quality Through Design*, Oxford Science Publications, Clarendon Press, Oxford