Quality Loss Functions for Nonnegative Variables and Their Applications

V. Roshan Joseph[1]

Georgia Institute of Technology

Atlanta, GA 30332-0205

To appear in *Journal of Quality Technology*

Abstract

Loss functions play a fundamental role in every quality engineering method. A new set of loss functions is proposed based on Taguchi's societal loss concept. Its applications to robust parameter design are discussed in detail. The loss functions are shown to posses some interesting properties and lead to theoretical results that cannot be handled with other loss functions.

Introduction

Quality is an abstract concept and is very difficult to have a precise definition. Crosby (1979) views quality as "conformance to requirements". Does he mean to say that a product can be treated as a quality product only if it meets the requirements and inferior otherwise? Conceptually it is more appealing to consider that the product has best quality when it exactly meets the requirements and that it suffers a loss of quality when it deviates from the requirements. This is inherent in Taguchi's (1986) definition of quality, who states that "quality is the loss a product causes to society after being shipped, other than any losses caused by its intrinsic functions". Furthermore Taguchi quantifies the deviations from the requirements in terms of monetary units by using a quadratic loss function given by

$$L(Y) = c(Y - T)^2, \tag{1}$$

[1]Dr. Joseph is an Assistant Professor in the School of Industrial and Systems Engineering. He is a member of ASQ. His email address is roshan@isye.gatech.edu

where Y is the quality characteristic, T its target value, and c a cost-related numerical constant. The quadratic loss function in Equation (1) can be justified as an approximation to the true loss function using Taylor series expansion. The quality loss functions were a radical change from the then prevailing concept of evaluating quality in terms of specifications. By converting the abstract concept of quality into loss functions, we can develop quantitative methods for quality control and improvement.

The quality characteristic in the above discussion is known as a nominal-the-best (NTB) characteristic, because the quality is evaluated with respect to a nominal value T. Two other types of quality characteristics are smaller-the-better (STB) and larger-the-better (LTB). An STB characteristic is a nonnegative variable with target $T = 0$. Thus its loss function is given by

$$L(Y) = cY^2. \tag{2}$$

An LTB characteristic is a nonnegative variable with target equal to a large value (infinity). Here we cannot put $T = \infty$ in Equation (1). To circumvent this problem, Taguchi uses the trick that, if Y is LTB, then $1/Y$ should be STB and therefore its loss function is given by

$$L(Y) = c/Y^2. \tag{3}$$

Although the quadratic loss function in (1) is a reasonable choice for many quality characteristics, there are several situations where it is not appropriate. In the case of a time watch, any deviations from the actual time, slow or fast, will result in a loss and the quadratic loss function can be used to adequately model the loss. Whereas in the case of painting thickness on a car body, there is no target specified by the customer and the manufacturer will have to decide upon a target based on losses incurred by the painting thickness to the society. Quadratic loss function is not the natural choice here.

Consider a nonnegative NTB characteristic. Because Y is restricted in the interval $[0, \infty]$, we may think 0 and ∞ as equally bad to the society. Thus a function that assigns ∞ to the loss when $Y = \infty$ should also assign ∞ to the loss when $Y = 0$. The quadratic loss function in Equation (1) does not posses this property. This may not cause much problem if the distribution of Y is concentrated away from 0, but not otherwise. Because most of the real physical variables are nonnegative in nature the above point demands special attention and careful treatment.

Some of the quality characteristics such as the neck size of a ready-made shirt, the amount of soft drink in a bottle, drug dose, etc. have asymmetric loss with respect to their target value. Quadratic loss function as defined in Equation (1) cannot be used to model such losses. Ironically, it is impossible to have a symmetric loss function for nonnegative variables, because there is no mid point in $[0, \infty]$.

The estimation of the cost-coefficient c in Equation (1) is not a trivial task. The usual technique is to guess the loss at some value of Y and then solve for the c from Equation (1). This difficulty is mainly because of not having any physical interpretation for c. The quadratic loss function is only a mathematical approximation to the true loss function. If a loss function can be derived directly from the definition of quality, then the parameters in that loss function will have some interpretation and will be easier to estimate.

Loss functions are widely used in statistics, economics, and other disciplines. Taguchi is probably the first to apply them in the area of quality. Since then several authors have proposed alternatives to Taguchi's quality loss functions. Spiring (1993) has proposed an inverted normal density function to model losses with finite maximum. Further work in this direction can be found in Sun, Laramee, and Ramberg (1996) and Spiring and Yeung (1998). Moorhead and Wu (1998) discusses a class of asymmetric quality loss functions.

The article is organized as follows. We first derive a new set of quality loss functions for nonnegative variables based on Taguchi's definition of quality. The loss function is then compared with the quadratic loss function. Some examples where the new loss functions seem more appropriate are presented. Its applications to quality engineering particularly to robust parameter design, is discussed. A multivariate extension of the quality loss function is proposed. Some concluding remarks are given at the end.

Loss Functions

We will use Taguchi's societal loss concept to derive some quality loss functions. Here the society includes the manufacturer, customer, environment, and all others who directly or indirectly come in contact with the product. Consider a nonnegative quality characteristic (Y) of the product taking values in $[0, \infty]$. If Y is physically restricted in $[a, b], 0 < a < b < \infty$, then we should transform it to, say, $(Y - a)/(b - Y)$ so that the transformed variable takes values in $[0, \infty]$. Now define the three types of characteristics as follows:

- Y is an STB characteristic if all the members of the society wants Y to be as small as possible.
- Y is an LTB characteristic if all the members of the society wants Y to be as large as possible.
- Y is an NTB characteristic if some of the members of the society want Y to be as small as possible and some others want it to be as large as possible.

Note that in the definition of NTB there are two groups of members with conflicting interests. If a member wants Y to be neither large nor small, then that member is assigned to both groups. We will see that a single target will be arrived as a compromise choice between these two groups. The NTB case will be illustrated with some real examples in a later section.

Consider an STB characteristic. Let I be the set of members in the society and $L_i(Y)$ be the loss caused by the product to member i. From the definition of STB characteristic it is clear that $L_i(Y)$ is an increasing function of Y for all $i \in I$. There are infinite number of increasing functions, the simplest of which is the linear function $L_i(Y) = C_i Y$. Assume that *the loss is additive* for each member. Then the total loss imparted to the society is given by

$$L(Y) = \sum_{i \in I} L_i(Y) = \sum_{i \in I} C_i Y = cY. \tag{4}$$

Similarly for an LTB characteristic choose $L_i(Y) = C_i/Y$. Then

$$L(Y) = \sum_{i \in I} L_i(Y) = \sum_{i \in I} C_i/Y = c/Y. \tag{5}$$

This choice of loss function is more arbitrary, but the derived results are good to justify it. For an NTB characteristic, let I_1 and I_2 represent the two groups ($I = I_1 \cup I_2$), where Y is STB for the members in I_1 and LTB for the members in I_2. Then

$$\begin{aligned} L(Y) &= \sum_{i \in I_1} L_i(Y) + \sum_{i \in I_2} L_i(Y) \\ &= \sum_{i \in I_1} C_i Y + \sum_{i \in I_2} C_i / Y \\ &= c_1 Y + c_2 / Y. \end{aligned} \tag{6}$$

Note that (4) and (5) are special cases of (6) with $c_2 = 0$ and $c_1 = 0$ respectively. Therefore in the future we can study the loss function in (6) without any reference to the type of characteristic. Thus these loss functions are more congruent than Taguchi's loss functions.

Define the target (T) of a quality characteristic as

$$T = arg \min_{Y \geq 0} L(Y).$$

Differentiating $L(Y)$ in Equation 6 with respect to Y and equating to zero we get a unique solution for $Y \in [0, \infty]$,

$$T = \sqrt{c_2/c_1}. \tag{7}$$

The minimum loss is $L(T) = 2\sqrt{c_1 c_2}$. For an NTB characteristic $L(T) > 0$. This would seem to suggest that it is better not to produce the product, because it imparts a loss to society whatever the value of Y is! This is a confusion that is created by working with the loss caused by the product instead of the utility of the product. If we assume the utility function to be $U(Y) = K - L(Y)$, with the constant $K \geq L(T)$, then this problem will be avoided. As a matter of convention, we may define the loss function as

$$L(Y) = c_0 + c_1 Y + c_2/Y, \tag{8}$$

where c_0 is chosen so as to force the minimum loss to zero.

Here the choice of loss functions is mainly driven by mathematical simplicity. The true loss functions could be different and complicated. So it is prudent to consider a more general form of the loss function,

$$L(Y) = c_0 + c_1 Y^{\alpha_1} + c_2/Y^{\alpha_2}, \tag{9}$$

where α_1 and α_2 are positive constants. The α_1 and α_2 should be selected based on the technical knowledge of the product. In the absence of any such knowledge we may have to content with (8). The choice $\alpha_1 = \alpha_2 = 2$ is also of interest as it leads to Taguchi's loss functions for STB and LTB.

Comparison to Quadratic Loss Functions

When the target for the quality characteristic is specified before hand, we can impute the cost ratio using (7) as $c_2/c_1 = T^2$. Substituting for c_2 in (8) we get

$$\begin{aligned} L(Y) &= -2\sqrt{c_1^2 T^2} + c_1 Y + c_1 T^2/Y \\ &= c_1 T\left(\frac{Y}{T} + \frac{T}{Y} - 2\right) \qquad (10) \\ &= c_1 (Y-T)^2/Y. \qquad (11) \end{aligned}$$

Note the resemblance of (11) with the quadratic loss function in Equation (1). Expanding (11) as follows,

$$L(Y) = \frac{c_1(Y-T)^2}{T\left[1+(Y-T)/T\right]} = \frac{c_1}{T}\left[(Y-T)^2 - \frac{(Y-T)^3}{T} + \frac{(Y-T)^4}{T^2} - \cdots\right] \qquad (12)$$

for $|Y - T| < T$. Thus, when T is large and the distribution of Y is concentrated around T, the behavior of $L(Y)$ in Equation (11) is close to the quadratic loss function in Equation (1) with $c_1 = cT$. Now consider a quadratic loss function with a logarithmic transformation of Y.

$$L_{log}(Y) = c(\log Y - \log T)^2 \qquad (13)$$

Using Taylor's series expansion

$$L_{log}(Y) = \frac{c}{T^2}\left[(Y-T)^2 - \frac{(Y-T)^3}{T} + \frac{11}{12}\frac{(Y-T)^4}{T^2} - \cdots\right]. \qquad (14)$$

Comparing (12) and (14) we see that (13) gives a much better approximation to (11) with $c_1 = c/T$. These loss functions are plotted in Figure 1 with $c_1 = 1/T$. As can be seen that the quadratic loss function on $\log Y$ and the loss function in Equation (11) are almost indistinguishable in the range $[.5T, 2T]$. This should not be too surprising because the quadratic loss function is meant for unrestricted variables. The log function will transform the nonnegative variables to unrestricted variables and therefore the quadratic loss function on $\log Y$ should be comparable to a loss function such as (11) that is designed for nonnegative variables.

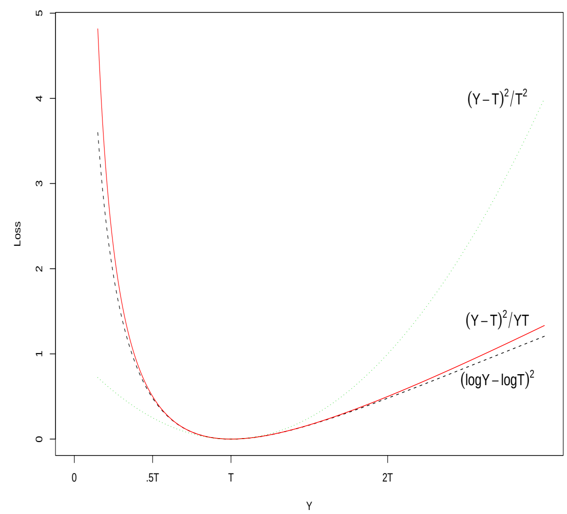

Figure 1: Comparison of loss functions

Some Examples

Consider the solder mask thickness (Y) of a printed circuit board (PCB). The solder mask prevents solder from being deposited on specific areas of the PCB during the hot air leveling process and also functions as an insulator for the circuits. As the thickness of the mask is reduced, the reliability of the PCB is reduced and therefore a customer will want the thickness to be large. But as the thickness increases the amount of ink required in the solder mask coating process increases, thereby increasing the manufacturing cost. This cost increases linearly with the thickness. Thus the loss function for the thickness can be written as in Equation (8), where c_2/Y is an approximation to the customer's loss and $c_1 Y$ is the manufacturer's loss. Note that the manufacturer will eventually pass his loss to the customer through product pricing, but we are not concerned with that because the total loss to the society remains the same.

Consider the design of a beam with rectangular cross section having width Y_1 and thickness Y_2. The bending stress developed in the beam by a certain load is inversely proportional to the moment of inertia $Y_1Y_2^3/12$. Therefore to avoid the bending failure we will want $Y_1Y_2^3$ to be as large as possible. The cost of the beam increases linearly with its area Y_1Y_2. Therefore the loss function is

$$L(Y_1, Y_2) = c_0 + c_1Y_1Y_2 + \frac{c_2}{Y_1Y_2^3}. \tag{15}$$

Then for a given thickness $L(Y_1) = c_0^{'} + c_1^{'}Y_1 + c_2^{'}/Y_1$ and for a given width $L(Y_2) = c_0^{''} + c_1^{''}Y_2 + c_2^{''}/Y_2^3$.

As another example, consider the amount of milk (Y) filled in a bottle. For the customer, the amount of milk is an LTB characteristic. But as Y increases the cost of the product increases linearly. Therefore we can use

$$L(Y) = c_0 + c_1Y + \frac{c_2}{Y} = c_1T\left(\frac{Y}{T} + \frac{T}{Y} - 2\right)$$

to model the loss, where T is the target for Y. It is a well known fact that the loss of under-filling is more than that of over-filling. The above loss function captures this behavior. The asymmetric loss can also be modeled using a skewed quadratic loss or absolute loss function (Taguchi, 1986; Moorhead and Wu, 1998), but with these modifications the mathematical simplicity possessed by the quadratic loss function is lost.

In the above examples the material cost played a prominent role. This is not related to the loss caused by the product "after being shipped". Therefore this approach is a slight departure from Taguchi's definition of quality. In our formulation, we include all the losses to the society relating to the quality characteristic of the product from its manufacturing to the end of its life cycle.

Estimation of Loss Function

It is usually difficult to accurately estimate the cost coefficients c_1 and c_2 in Equation (8). The commonly used approach is to guess the loss at some values of Y and solve for c_1 and c_2. Suppose L_1 is the loss at $Y = y_1$ and L_2 is the loss at $Y = y_2$. Then, we obtain

$$c_1 = \frac{(\sqrt{L_1y_1} - \sqrt{L_2y_2})^2}{(y_1 - y_2)^2} \text{ and } c_2 = \frac{(y_2\sqrt{L_1y_1} - y_1\sqrt{L_2y_2})^2}{(y_1 - y_2)^2}.$$

If one can assume a reasonable target for Y, then we can set $y_2 = T$ and $L_2 = 0$ in the above formulas to obtain $c_1 = L_1 y_1/(y_1 - T)^2$ and $c_2 = c_1 T^2$. In some cases c_1 or c_2 posses some physical interpretation and therefore can be directly estimated. We illustrate such a case with the solder mask thickness example discussed earlier. In this example we know that c_1 is related to the material cost. Suppose the cost of the ink is USD 100/litre. If the area of the PCB is 500 cm^2, then c_1 can be estimated as $500 \times 100/10^3 \times 1/10^4 = 0.005$ USD/microns. Direct estimation of c_2 is difficult. Suppose a thickness of 15 microns is reasonable. Then as in Equation (11) the loss function can be written as $L(Y) = .005(Y-15)^2/Y$. If a quadratic loss function is used instead, then $L(Y) = c(Y - 15)^2$. Here the estimation of c is very difficult because it does not have any physical interpretation as c_1 has.

Applications to Quality Engineering

The quality engineering activities are aimed at making the quality loss zero. Because Y is random, $L(Y)$ is also random and therefore we could state this mathematically as to make $L(Y) = 0$ with probability 1. A necessary and sufficient condition for this is to make $E[L(Y)] = 0$ as can be seen from the Markov's inequality: for all $\epsilon > 0$, $P\{L(Y) > \epsilon\} \leq E[L(Y)]/\epsilon$. For the quadratic loss function in (1), $E[L(Y)] = c[(E(Y) - T)^2 + Var(Y)]$. Thus to make the expected loss zero, we will aim at reducing variation to zero while keeping the mean at target. Because of the wide spread applications of quadratic loss functions, variation reduction became synonymous to quality improvement. With the new loss function in Equation (8) the statistics for control and optimization will be related to $E(Y)$ and $E(1/Y)$ rather than $E(Y)$ and $Var(Y)$. In the next section we will explain in detail the impact of the new loss function on robust parameter design.

Screening of products with respect to a specification limit $[T_1, T_2]$ is a technique that is different from other quality engineering methods. Screening changes the shape of the loss function whereas the other techniques change the distribution of Y. Suppose r_1 and r_2 are the rejection/rework costs when Y is below T_1 and above T_2 respectively. If r_1 and r_2 are independent of Y, then the loss function after screening is given by

$$L_a(Y) = \begin{cases} r_1 & , Y < T_1 \\ c_0 + c_1 Y + c_2/Y & , T_1 \le Y \le T_2 \\ r_2 & , Y > T_2 \end{cases} .$$

This is shown in Figure 2.

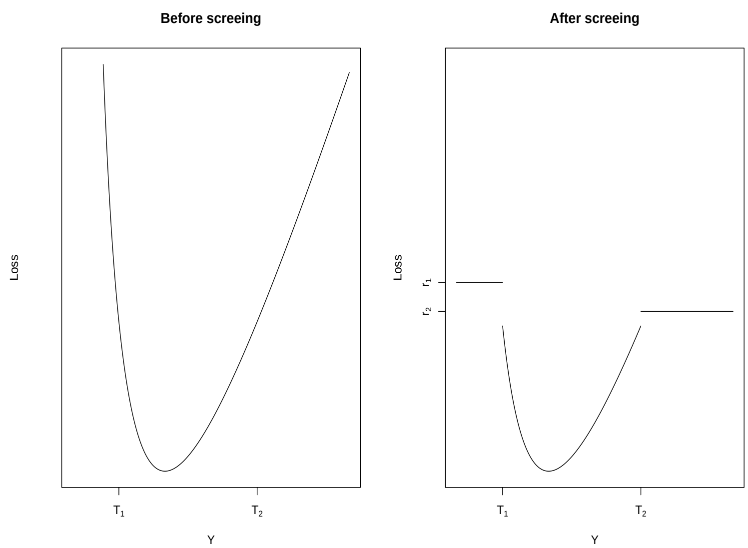

Figure 2: Effect of screening

Robust Parameter Design

In this section we investigate the applications of the new loss functions to robust parameter design. Robust parameter design is one of the most important tools for quality improvement. The basic idea is to find a setting for the control factors in the system to make the output insensitive to the noise variation.

STB and LTB Problems

The STB and LTB characteristics arise occasionally in robust parameter design experiments. For examples, Quinlan (1985) reported a case study on speedometer cable casing for reducing post-extrusion shrinkage, Byrne and Taguchi (1987) described a case study on assembly of elastometric connector to a nylon tube for maximizing the pull-off force, etc.

Let $\mathbf{X}$ be the set of control factors. The objective is to find an $\mathbf{X} \in D$ to minimize $E[L(Y)]$, where D is the decision region of $\mathbf{X}$. From Equation (4), the performance measure to minimize for an STB characteristic is given by

$$PM = E(Y). \tag{16}$$

The PM can be estimated by $\overline{Y}$. If Taguchi's loss function is used, then $E(L) = c[E^2(Y) + Var(Y)]$. Thus both mean and variance are simultaneously minimized. Box(1988) argued that the STB and LTB problems are mainly concerned with "location" and the analysis of $\overline{Y}$ would suffice to solve the problem. A two-step procedure of first minimizing $E(Y)$ and then $Var(Y)$ was suggested by Tsui and Li (1994). In many real cases minimizing the mean of a nonnegative variable will also minimize the variance and therefore it is unlikely that these procedures will lead to conflicting results.

Similarly from Equation (5), the performance measure to minimize for an LTB characteristic is given by

$$PM = E(1/Y), \tag{17}$$

which can be estimated by $\overline{(1/Y)}$. It is important to note that minimizing $\overline{(1/Y)}$ and maximizing $\overline{Y}$ are different. As argued before, in most real cases the former will lead to a smaller variance of Y (since $Var(Y) \propto Var(1/Y)$) while the latter can lead to a setting that increases variance of Y.

Multiple Target Systems

Multiple target systems arise when the response can take several targets specified by the customer. The different targets are achieved by changing a factor in the system, known as signal factor. For example, in the braking system of an automobile different stopping distances can be achieved by changing the force on the pedal, different plating thicknesses in an electro-plating process can be achieved by changing the plating time, etc. The robust parameter design of multiple target systems is also known as dynamic parameter design, see Miller and Wu (1996) and Joseph and Wu (2002a,b) for a detailed discussion. (We will study the NTB problem as a special case in the next section.) Let M be the signal factor and $\mathbf{Z}$

the set of noise factors. Consider the following model,

$$Y = \beta(\mathbf{X}, \mathbf{Z})M. \tag{18}$$

Under the assumption that the costs are independent of the control factor setting, for the purpose of optimization, the loss function in Equation (10) can be taken as

$$L(Y) = \frac{Y}{T} + \frac{T}{Y}.$$

We will now derive the performance measure independent of adjustment (PerMIA). See Leon, Shoemaker, and Kacker (1987) and Leon and Wu (1992) for details on PerMIA and Joseph and Wu (2002a,b) for its extension to dynamic parameter design. Note that we did not consider PerMIAs in STB and LTB problems because no adjustment parameter exists in those problems.

For a given customer intent T, the expected loss is

$$\begin{aligned} E_Z(L|T) &= \frac{1}{T}E_Z(Y|T) + TE_Z(1/Y|T) \\ &= \frac{M}{T}E_Z(\beta) + \frac{T}{M}E_Z(1/\beta). \end{aligned}$$

We can set the signal factor to minimize this loss. Solving M from

$$\frac{1}{T}E_Z(\beta) - \frac{T}{M^2}E_Z(1/\beta) = 0$$

we get

$$M^* = T\sqrt{\frac{E_Z(1/\beta)}{E_Z(\beta)}}. \tag{19}$$

Thus depending on the customer intent T, the signal factor will be adjusted based on (19). Then the expected loss at this optimal setting is

$$E_Z(L^*|T) = 2\sqrt{E_Z(\beta)E_Z(1/\beta)}.$$

Thus by smoothing,

$$E(L^*) = E_T[E_Z(L^*|T)] = 2\sqrt{E_Z(\beta)E_Z(1/\beta)}.$$

Minimizing $E(L^*)$ is equivalent to minimizing the performance measure

$$PM(\mathbf{X}) = E_Z(\beta)E_Z(1/\beta). \tag{20}$$

The procedure can be summarized as

1. Find $\mathbf{X}^* \in D$ to minimize $PM(\mathbf{X})$ in Equation (20).
2. Adjust M depending on T as $M = T\sqrt{\frac{E_Z(1/\beta(\mathbf{X}^*,\mathbf{Z}))}{E_Z(\beta(\mathbf{X}^*,\mathbf{Z}))}}$.

The above optimization procedure can be implemented using a response modeling approach or by using a performance measure modeling approach. See Wu and Hamada (2000, Chapter 10 and 11) for details. If the noise factors have random levels in the experiment, PM can be estimated as

$$\hat{PM} = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i}{M_i}\,\frac{1}{n}\sum_{i=1}^{n}\frac{M_i}{Y_i}, \tag{21}$$

where the average is taken over the noise levels. By Jensen's inequality $E(1/\beta) \geq 1/E(\beta)$, which implies that $\log PM(\mathbf{X})$ is always nonnegative. Therefore when using linear models, we may fit $\log\log \hat{PM}$ in terms of $\mathbf{X}$.

We will now compare this approach with that of the quadratic loss function. Using PerMIA theory it is easy to show that the performance measure to maximize is equivalent to

$$SN(\mathbf{X}) = \frac{E_Z^2(\beta)}{Var_Z(\beta)}, \tag{22}$$

and the optimal signal setting is given by

$$M^* = T\frac{E_Z(\beta)}{E_Z^2(\beta) + Var_Z(\beta)}. \tag{23}$$

The SN in (22) can be considered as a signal-to-noise ratio. Using Taylor's series expansion

$$PM(\mathbf{X}) = E_Z(\beta)\left(\frac{1}{E_Z(\beta)} + \frac{Var_Z(\beta)}{E_Z^3(\beta)} + \cdots\right) = 1 + \frac{Var_Z(\beta)}{E_Z^2(\beta)} + \cdots \approx 1 + \frac{1}{SN(\mathbf{X})}.$$

Thus minimizing PM is approximately equivalent to maximizing SN. After the optimal signal setting in (23) the mean of Y will be

$$E_Z(Y|T) = T\frac{E_Z^2(\beta)}{E_Z^2(\beta) + Var_Z(\beta)} \leq T.$$

Thus the mean is adjusted to a value lower than the target. This shrinkage property was observed by Leon, Shomaker, and Kacker (1987), Box (1988), and Leon and Wu (1992) in the case of NTB characteristics. Under the signal setting given by (19) the mean of Y will be

$$E_Z(Y|T) = T\sqrt{E_Z(\beta)E_Z(1/\beta)} \geq T.$$

Thus the mean adjustment is in the opposite direction. But when the variation is small, both will be close to T. Interestingly both the performance measures PM and SN remain the same even under an unbiased adjustment strategy $M^* = T/E_Z(\beta)$.

NTB Problem

The NTB problem can be considered as a special case of multiple target systems with the customer intent taking a single target. Suppose there exists a signal factor (also known as scaling factor in NTB problems) outside the set of experimental factors. For example, in the famous Ina tile experiment by Taguchi (1986), the mould dimension can be used as a signal factor to achieve the desired tile dimension, even though the mould dimension is not a factor in the experiment. Suppose all the experiments were carried out at a fixed $M = M_0$. Then from (20) we obtain the PerMIA as

$$\begin{aligned} PM(\mathbf{X}) &= E_Z(\beta M_0)E_Z(\frac{1}{\beta M_0}) \\ &= E_Z(Y)E_Z(1/Y), \end{aligned} \tag{24}$$

and its sample analog is

$$\hat{PM} = \frac{1}{n}\sum_{i=1}^{n} Y_i \, \frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i}. \tag{25}$$

From (16), (17), and (24) we see that

$$\log PM_{NTB} = \log PM_{STB} + \log PM_{LTB}.$$

This is an interesting result because, although the loss functions are additive, *the performance measures are log-additive.*

Thus to solve an NTB problem using performance measure modeling approach, model $\eta = \log\log \hat{PM}$ and

$$\nu = \log\sqrt{\frac{\frac{1}{n}\sum_{i=1}^{n} Y_i}{\frac{1}{n}\sum_{i=1}^{n} 1/Y_i}}$$

in terms of $\mathbf{X}$. Now the two-step optimization can be stated as

1. Find $\mathbf{X}^* \in D$ to minimize $\eta(\mathbf{X})$.

2. Adjust M to $M_0 T \exp(-\nu(\mathbf{X}^*))$.

The definition of ν and the adjustment step are based on (19).

If there exists no signal factor outside the set of experimental factors then we have two options. The first option is to still use (25), hoping that a signal factor is present among the experimental factors. We can try to identify the signal factor from $\mathbf{X}$ using $\nu(\mathbf{X})$ because signal factor should have a large effect on it but less effect on $\eta(\mathbf{X})$. Then the identified signal factor can be adjusted so that $\nu(\mathbf{X})$ is equal to $\log T$. This approach may fail if there is no such factor. The second option is to find the control factor setting by directly minimizing the expected loss

$$\frac{1}{n}\sum_{i=1}^{n}\frac{Y_i}{T} + \frac{1}{n}\sum_{i=1}^{n}\frac{T}{Y_i}.$$

We have already noticed that the loss function in Equation (11) is close to a quadratic loss function with a log-transform on Y. Interestingly the results of robust parameter design is exact if Y follows a log-normal distribution, which is evident from the following two equations:

$$E\left[\log Y\right] = \log\sqrt{\frac{E(Y)}{E(1/Y)}} \quad \text{and} \quad Var\left[\log Y\right] = \log\left[E(Y)E(1/Y)\right].$$

Noise Factor Compounding

Noise factor compounding is a technique introduced by Taguchi (1986) to reduce the number of runs in robust parameter design experiments. This technique can be justified as follows. Under the loss function in Equation (9), the $P\hat{M}$ in Equation (25) becomes

$$P\hat{M} = \left(\frac{1}{n}\sum_{i=1}^{n} Y_i^{\alpha_1}\right)^{1/\alpha_1} \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i^{\alpha_2}}\right)^{1/\alpha_2}.$$

Suppose large values of α_1 and α_2 are meaningful. Then taking the limit $\alpha_1, \alpha_2 \to \infty$ we get

$$P\hat{M} = \frac{\max_i Y_i}{\min_i Y_i}.$$

Thus we need only the maximum and minimum values of Y to compute the performance measure. If we know the effect of noise factors before hand, we can compound them into two levels

$$N_1 = \text{noise settings to minimize } Y,$$

$$N_2 = \text{noise settings to maximize } Y.$$

Thus the performance measure can be computed from the two compounded noise levels instead of repeating the experiment at n noise levels. Noise factor compounding is a useful technique in robust parameter design experiments which cannot be mathematically justified using quadratic loss functions. Some theoretical properties and limitations of noise factor compounding technique are studied in Hou (2002).

Multivariate Loss Function

Often the quality of a product is expressed using more than one characteristics $Y_1, \cdots, Y_p$. We will call two characteristics "non-interacting" with respect to loss, if the loss caused by one characteristic is independent of the other. Thus if $Y_1, \cdots, Y_p$ are non-interacting, then the total loss will be

$$L(Y_1, \cdots, Y_p) = \sum_{i=1}^{p} L_i(Y_i) = c_o + \sum_{i=1}^{p} c_{1i} Y_i + \sum_{i=1}^{p} c_{2i}/Y_i. \tag{26}$$

In the beam example considered before the two variables width and thickness "interact" with respect to loss and therefore the total loss cannot be given by Equation (26). A more general loss function than Equation (26) is

$$L(Y_1, \cdots, Y_p) = \sum_{i_1} \cdots \sum_{i_p} c_{i_1, \cdots, i_p} Y_1^{i_1} \cdots Y_p^{i_p}, \tag{27}$$

where $i_1, \cdots, i_p \in \{0, 1, -1\}$. As a special case if Y_1 is STB and Y_2 is LTB, then the loss function reduces to

$$L(Y_1, Y_2) = c_{1,0} Y_1 + c_{1,-1} Y_1/Y_2 + c_{0,-1} 1/Y_2.$$

For a real example, consider the image transfer process in PCB manufacturing. By exposing to ultra violet rays using a negative type photo-tool, the circuitry portion of the dry film photo-resist gets polymerized but the non-circuitry portion remains intact. If the circuitry portion is not hardened enough it will be washed-off during developing and leads to "opens" in the circuits. Therefore the degree of polymerization on the circuitry portion of photo resist (Y_1) is an LTB characteristic. Y_1 can be increased by exposing for more time. But then the non-circuitry portion will also get exposed, which will lead to "shorts" in the circuits. The degree of polymerization on the non-circuitry portion of the photo-resist (Y_2) is thus an STB characteristic. Although Y_1 and Y_2 will be positively correlated they can be considered to be non-interacting with respect to the loss and therefore the loss is given by

$$L(Y_1, Y_2) = c_1/Y_1 + c_2 Y_2.$$

It is generally true that a process should not be attempted to optimize with respect to only STB or only LTB type characteristics. It should be optimized using a combination of both STB and LTB characteristics or by using an NTB characteristic. In the above example if only Y_1 were considered for optimization, then we might succeed in solving the opens problem but might end up in a situation with full of shorts. See Joseph and Wu (2002) for a real application of the above loss function.

The multivariate generalization of quadratic loss functions and its applications to robust parameter design are discussed in Pignatiello (1993) and Tsui (1999). Its applications to multivariate process control are studied in Tsui and Woodall (1993). These multivariate quadratic loss functions consider only up to the second order interactions between the characteristics. The loss function in Equation (27) include up to the p^{th} order interactions.

Conclusions

In this article we have introduced a new set of loss functions for nonnegative variables. Some examples are presented to demonstrate the advantages of these loss functions over the quadratic loss functions. The new loss functions under some modeling assumptions lead to simple performance measures for robust parameter design. They also lead to some theoretical results that cannot be handled effectively by other loss functions. No attempt is made here to claim that these loss functions are uniformly better than other loss functions. They are presented as an alternative choice that a quality practitioner should keep in mind and use depending on the problem in hand.

Acknowledgments

This research is a part of author's Ph.D. thesis at the University of Michigan, done under the supervision of Professor C. F. Jeff Wu. The research was supported by NSF DMS 0072489.

References

Box, G. (1988). "Signal-to-Noise Ratios, Performance Criteria, and Transformations" (with discussion). *Technometrics* 30, pp. 1-40.

Byrne, D. M. and Taguchi, S. (1987). "The Taguchi Approach to Parameter Design". *Quality Progress* December, pp. 19-26.

Crosby, P. B. (1979). *Quality is Free.* McGraw-Hill, New York, NY.

Hou, X. S. (2002). "On the Use of Compound Noise Factor in Parameter Design Experiments". *Applied Stochastic Models in Buisiness and Industry*, to appear.

Joseph, V. R., and Wu, C. F. J. (2002a). "Robust Parameter Design of Multiple Target Systems". *Technometrics* 44, pp. 338-346.

Joseph, V. R., and Wu, C. F. J. (2002b). "Performance Measures in Dynamic Parameter Design". *Journal of Japanese Quality Engineering Society* 10, pp. 82-86.

Joseph, V. R., and Wu, C. F. J. (2002c). "Failure Amplification Method: An Information Maximization Approach to Categorical Response Optimization". *Technical Report* 387, Department of Statistics, University of Michigan, Ann Arbor.

Leon, R. V., Shoemaker, A. C., and Kacker, R. N. (1987). "Performance Measures Independent of Adjustment: An Explanation and Extension of Taguchi's Signal-to-Noise Ratios"(with discussion). *Technometrics* 29, pp. 253-285.

Leon, R. V. and Wu, C. F. J. (1992). "A Theory of Performance Measures in Parameter Design". *Statistica Sinica* 2, pp. 335-358.

Miller, A., and Wu, C. F. J. (1996). "Parameter Design for Signal-Response Systems: A Different Look at Taguchi's Dynamic Parameter Design". *Statistical Science* 11, pp. 122-136.

Moorhead, P. R. and Wu, C. F. J. (1998). "Cost-Driven Parameter Design". *Technometrics* 40, pp. 111-119.

Pignatiello, J. J., Jr. (1993). "Strategies for Robust Multiresponse Quality Engineering". *IIE Transactions* 25, pp. 5-15.

Quinlan, J. (1985). "Product Improvement by Application of Taguchi Methods". *American Supplier Institute News*, American Supplier Institute, Dearborn, MI, pp. 11-16.

Spiring, F. A. (1993). "The Reflected Normal Loss Function". *Canadian Journal of Statistics* 24, pp. 241-249.

Spiring, F. A. and Yeung, A. S. (1998). "A General Class of Loss Functions with Industrial Applications". *Journal of Quality Technology* 30, pp. 152-162.

Sun, F.; Laramee, J.; and Ramberg, J. (1996). "On Spiring's Inverted Normal Loss Function". *Canadian Journal of Statistics* 24, pp. 241-249.

Taguchi, G. (1986). *Introduction to Quality Engineering: Designing Quality into Products and Processes.* Asian Productivity Organization, UNIPUB/Kraus International, White Plains, NY.

Tsui, K. L. and Li, A. (1994). "Analysis of Smaller-and Larger-the Better Robust Design Experiments". *International Journal of Industrial Engineering* 1, pp. 193-202.

Tsui, K. L. (1999). "Robust Design Optimization for Multiple Characteristic Problems". *International Journal of Production Research* 37, pp. 433-445.

Tsui, K. L. and Woodall, W. H. (1993). "Multivariate Control Charts Based on Loss Functions". *Sequential Analysis* 12, pp. 79-92.

Wu, C.F.J., and Hamada, M. (2000). *Experiments: Planning, Analysis, and Parameter Design Optimization.* John Wiley & Sons, New York.

Key Words: *Quality Engineering, Robust Parameter Design, Multiple Target Systems, Performance Measure.*