Metrika (2006) 63: 207–213
DOI 10.1007/s00184-005-0012-0

ORIGINAL ARTICLE

Heike Trautmann · Claus Weihs

On the distribution of the desirability index using Harrington's desirability function

Received: 31 October 2003 / Accepted: 9 November 2004 / Published online: 28 October 2005

Abstract The concept of desirability is a means for complexity reduction of multivariate quality optimization. This paper provides a theoretical breakthrough regarding desirability indices, which application fields were formerly limited primarily by the lack of its distribution. Focussed are the distributions of Harrington's desirability functions and different types of the desirability index.

Keywords Desirability index · Desirability Function · Multicriteria optimization · Double lognormal distribution · Geometric mean

1 Introduction

The concept of desirability, introduced by Harrington (1965), is a method for multicriteria optimization in industrial quality management. Via desirability functions (DFs), which allow for comparing different scales of the quality measures (QMs) by mapping them to [0, 1], and the desirability index (DI) the multivariate optimization problem is converted into a univariate one. Based on design of experiment methods optimal levels of process influencing factors can be determined that optimize all often competing QMs simultaneously. Modifications came up either in terms of more flexible DFs (e.g. Derringer and Suich 1980; Castillo et al. 1996) or in terms of different DIs (e.g. Kim and Lin 2000). Although the Harrington and Derringer-Suich approaches for the DI became widely accepted (e.g. Averill et al. 2001; Ben-Gal and Bukchin 2002; Tuyev and Fedchenko 1997; Wu et al. 2000), due to the lack of its distribution other applications like process control or confidence statements were not accessible. Now, we eventually succeeded in setting up this distribution for Harrington's DFs (Trautmann 2004a), which is outlined in the following.

H. Trautmann (✉) · C. Weihs
Department of Computer-Aided Statistics, University of Dortmund, Dortmund Germany
E-mail: trautmann@statistik.uni-dortmund.de

2 Harrington's desirability functions and the desirability index

Harrington introduced two types of DFs which transform the QMs onto (0, 1]. One aims at maximization of the QM (one-sided specification) whereas the other one reflects a target value problem (two-sided specification). Concerning the latter the transformation requires two specification limits (LSL, USL) for a QM Y symmetrically around the target value, which are associated with a desirability of $1/e$. Then the DF d is defined as

$$d(Y') = e^{-|Y'|^n} \quad \text{with} \quad Y' = \frac{2Y - (USL + LSL)}{USL - LSL}. \tag{1}$$

The parameter $n > 0$ is to be chosen so that the resulting kurtosis of the function adequately meets the expert's preferences. The one-sided DF uses a special form of the Gompertz-curve, where the kurtosis of the function is determined by the solution (b_0, b_1) of a system of two linear equations that require two values of Y and related values of d:

$$d(Y') = e^{-e^{-Y'}} \quad \text{with} \quad Y' = -[\ln(-\ln d)] = b_0 + b_1 Y. \tag{2}$$

The DI combines k individual desirability functions d_i into one overall quality value by

$$D := (\prod_{i=1}^{k} d_i)^{1/k} \text{ or as a modification of Harrington's approach} \tag{3}$$

$$D := \min_{i=1,\dots,k} d_i \quad \text{(Kim and Lin 2000).} \tag{4}$$

3 Distribution of the desirability function

The analysis of the distribution of the DF is based on the assumption of a QM $Y_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Focussing on the one-sided specification using (2) it is obvious that

$$Z_i := e^{-Y_i'} \sim \text{Lognormal}(\tilde{\mu}_i, \tilde{\sigma}_i^2) \text{ with } \tilde{\mu}_i = -(b_{0i} + b_{1i} \cdot \mu_i) \text{ and } \tilde{\sigma}_i^2 = (b_{1i})^2. \tag{5}$$

Thus the application of the density transformation theorem (DTT) (Mood et al. 1974, p 200) to the random variable (RV) $D_i := \exp(-Z_i)$ leads to the following theorem

Theorem 1 (One-sided specification) *Given a QM $Y_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ Harrington's DF (2) has the double lognormal distribution (DLN) (Holland and Ahsanullah 1989):*

$$f_{D_i}(d_i) = -\frac{1}{\sqrt{2\pi} \cdot \tilde{\sigma}_i \cdot \log(d_i) \cdot d_i} \cdot \exp\left[-\frac{1}{2\tilde{\sigma}_i^2} \cdot (\log(-\log(d_i)) - \tilde{\mu}_i)^2\right], \tag{6}$$

$$F_{D_i}(d_i) = 1 - \Phi[(\log(-\log(d_i)) - \tilde{\mu}_i)/\tilde{\sigma}_i], \quad 0 < d_i < 1; \tag{7}$$

$$\Phi(x) := \textit{Distribution function of } \mathcal{N}(0, 1) \textit{ in } x. \tag{8}$$

The two-sided specification is analyzed analogously. Based on (1)

$$X_i := |Y_i'| \sim FoldedNormal(\tilde{\mu}_i, \tilde{\sigma}_i^2) \text{ (Johnson et al. 1994, p 170) with} \tag{9}$$
$$\tilde{\mu}_i = \tfrac{2}{USL_i - LSL_i} \cdot \mu_i - \tfrac{USL_i + LSL_i}{USL_i - LSL_i} \text{ and } \tilde{\sigma}_i^2 = (\tfrac{2}{USL_i - LSL_i})^2 \cdot \sigma_i^2.$$

By the DTT the density results, the distribution function is then calculated by integration.

Theorem 2 (Two-sided specification) *Given a QM $Y_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ Harrington's two-sided DF (1) has the following distribution:*

$$\begin{aligned} f_{D_i}(d_i) = {} & \frac{1}{\sqrt{2\pi} \cdot \tilde{\sigma}_i \cdot d_i \cdot n_i} \\ & \cdot (-\log(d_i))^{1/n_i - 1} \cdot \left[\exp(-((-\log(d_i))^{1/n_i} - \tilde{\mu}_i)^2/2\tilde{\sigma}_i^2)\right. \\ & \left. + \exp(-((-\log(d_i))^{1/n_i} + \tilde{\mu}_i)^2/2\tilde{\sigma}_i^2)\right], \end{aligned} \tag{10}$$

$$\begin{aligned} F_{D_i}(d_i) = {} & 2 - \Phi\left[\frac{((-\log(d_i))^{1/n_i} - \tilde{\mu}_i)}{\tilde{\sigma}_i}\right] \\ & - \Phi\left[\frac{((-\log(d_i))^{1/n_i} + \tilde{\mu}_i)}{\tilde{\sigma}_i}\right]. \end{aligned} \tag{11}$$

4 Distribution of the desirability index

The Distribution of the DI is analyzed separately for one-sided and two-sided DFs combined either by (3) or (4). When considering the geometric mean in the two-sided case the distribution is derived by rewriting the DI as

$$D := (\prod_{i=1}^{k} d_i)^{1/k} = (\prod_{i=1}^{k} e^{-|Y_i'|^{n_i}})^{1/k} = (e^{-\sum_{i=1}^{k} |Y_i'|^{n_i}})^{1/k}. \tag{12}$$

A general analytical form of the distribution unfortunately was not possible but for $k = 2$ and $n_i = 1$ recalling (9) the problem can be traced back to the distribution of $Z := |Y_1'| + |Y_2'|$, the sum of two independent folded-normal RVs, which was provided in Weber and Weihs (2003) resp. Trautmann (2004a). The distribution of the DI then again results from the DTT applied to the RV $D := \exp(-Z)^{1/2}$.

Theorem 3 (Two-sided DF, DI Geometric mean) *Given two independent QMs Y_i $(i = 1, 2)$ with two-sided DFs d_i (1), $n_i = 1\,\forall i$, and DF densities (10) the DI*

D (3) has the following density function:

$$
\begin{aligned}
f_{\mathcal{D}}(D) = {} & \frac{\sqrt{2}}{2D\sqrt{\pi(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)}} \cdot \left(\exp\left(-\frac{(-2\log(D)-\tilde{\mu_1}-\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) \right. \\
& \times \operatorname{erf}\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_2}^2-\tilde{\mu_1}\tilde{\sigma_2}^2+\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \\
& + \exp\left(-\frac{(-2\log(D)-\tilde{\mu_1}+\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) erf\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_2}^2-\tilde{\mu_1}\tilde{\sigma_2}^2-\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \\
& + \exp\left(-\frac{(-2\log(D)+\tilde{\mu_1}-\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) erf\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_2}^2+\tilde{\mu_1}\tilde{\sigma_2}^2+\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \\
& + \exp\left(-\frac{(-2\log(D)+\tilde{\mu_1}+\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) erf\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_2}^2+\tilde{\mu_1}\tilde{\sigma_2}^2-\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \\
& + \exp\left(-\frac{(-2\log(D)-\tilde{\mu_1}-\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) erf\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_1}^2+\tilde{\mu_1}\tilde{\sigma_2}^2-\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \\
& + \exp\left(-\frac{(-2\log(D)-\tilde{\mu_1}+\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) erf\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_1}^2+\tilde{\mu_1}\tilde{\sigma_2}^2+\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \\
& + \exp\left(-\frac{(-2\log(D)+\tilde{\mu_1}-\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) erf\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_1}^2-\tilde{\mu_1}\tilde{\sigma_2}^2-\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \\
& \left. + \exp\left(-\frac{(-2\log(D)+\tilde{\mu_1}+\tilde{\mu_2})^2}{2(\tilde{\sigma_2}^2+\tilde{\sigma_1}^2)} \right) erf\left(\frac{((-2\log(\mathrm{D}))\tilde{\sigma_1}^2-\tilde{\mu_1}\tilde{\sigma_2}^2+\tilde{\mu_2}\tilde{\sigma_1}^2)}{\tilde{\sigma_2}\tilde{\sigma_1}\sqrt{2}\sqrt{\tilde{\sigma_2}^2+\tilde{\sigma_1}^2}} \right) \right);
\end{aligned}
$$

$\mathrm{erf(x)} = 2 \cdot \Phi(\sqrt{2}x) - 1$ (Gaussian error function).

For one-sided DFs combined by the geometric mean an approximation of the distribution is the most suitable solution. The DI is rewritten as

$$D := \left(\prod_{i=1}^{k} d_i\right)^{1/k} = \left(\prod_{i=1}^{k} e^{-e^{-Y_i'}}\right)^{1/k} = \left(e^{-\sum_{i=1}^{k} e^{-Y_i'}}\right)^{1/k} \tag{13}$$

including the sum of lognormal RVs (see (5)), whose distribution can only be approximated. The approaches of Wilkinson, Farley as well as Schwartz and Yeh [see Schwartz and Yeh (1982) for a review] are widely known, which assume stability of the lognormal distribution under summation. For our problem Schwartz and Yeh's method is the most appropriate one as it has the widest application range; it is even valid for correlated QMs. It is outlined for two RVs, the multidimensional case is straightforward in an iterative way. For two lognormal RVs X_i with parameters $\tilde{\mu}_i$ and $\tilde{\sigma}_i^2$ $(i = 1, 2)$ the parameters of the approximating lognormal distribution are

$$\mu^* = \tilde{\mu}_1 + G_1(\sigma_w, \mu_w), \tag{14}$$

$$\sigma^{*2} = \tilde{\sigma}_1^2 - G_1^2(\sigma_w, \mu_w) - 2\rho^2 G_3(\sigma_w, \mu_w) + G_2(\sigma_w, \mu_w) \text{ with} \tag{15}$$

$$\rho = -\tilde{\sigma}_1/\sigma_W, \;\; \mu_W = \tilde{\mu}_2 - \tilde{\mu}_1,$$

$$\sigma_w^2 = \tilde{\sigma}_2^2 + \tilde{\sigma}_1^2 - \rho_{12}\tilde{\sigma}_1\tilde{\sigma}_2, \;\; \rho_{12} = \mathrm{corr}(\mathrm{X}_1, \mathrm{X}_2).$$

As the exact expressions for G_1, G_2, G_3 require an intense computational effort the authors propose polynomial approximations of the form (Schwartz and Yeh 1982)

$$\log_{10} G_i(\sigma_w, \mu_w) = \sum_{j=0}^{4} \sum_{k=0}^{4} A_{jk}(i) \sigma_w^{j/2} |\mu_w|^{k/2}, \qquad i = 1, \dots, 3. \tag{16}$$

Using $D := (e^{-\sum_{i=1}^{k} e^{-Y_i'}})^{1/k}$ with $\sum_{i=1}^{k} e^{-Y_i'} \overset{\text{approx.}}{\sim}$ Lognormal (μ^*, σ^{*2}), and therefore $Z := e^{-\sum_{i=1}^{k} e^{-Y_i'}} \approx \mathcal{DLN}(\mu^*, \sigma^{*2})$ analogous to Theorem 1, the density function of the DI results from transforming the RV Z into $D := Z^{1/k}$:

Theorem 4 (One-sided DF, DI geometric mean) *Given k QMs Y_i $(i = 1, \dots, k)$ with one-sided DFs d_i (2) and DF densities (6) the DI D (3) has the following approximative distribution:*

$$f_{\mathcal{D}}(D) \approx -\frac{1}{\sqrt{2\pi} \cdot \sigma^* \cdot \log(D) \cdot D} \cdot \exp\left[-\frac{1}{2\sigma^{*2}}(\log(-k \cdot \log(D)) - \mu^*)^2\right]$$

with μ^* and σ^{*2} as defined in (14) und (15),

$$F_{\mathcal{D}}(D) \approx 1 - \Phi\left[\frac{\log(k) + \log(-\log(D)) - \mu^*}{\sigma^*}\right].$$

Instead of the geometric mean also the minimum of the DFs can be applied as a DI. Subject to Mood et al. (1974) (p. 184) the distribution function of $D = min(d_1, \dots, d_k)$ is determined as $F_{\mathcal{D}}(D) = 1 - \prod_{i=1}^{k}[1 - F_{D_i}(D)]$ for independent d_i, which together with Theorem 1 and 2 leads to the following distribution functions of the DI.

Theorem 5 (Distribution functions, DI minimum DF) *If $Y_1, \dots, Y_k$ are k independent QMs with DFs d_i as in (2) and (1) as well as with densities $f_{D_i}(d_i)$ as in Theorems 1 and 2 then*

$$F_{\mathcal{D}}(D) = 1 - \prod_{i=1}^{k} \Phi\left[\frac{(\log(-\log(D)) - \tilde{\mu}_i)}{\tilde{\sigma}_i}\right] \quad (\textit{One-sided DFs}),$$

$$F_{\mathcal{D}}(D) = 1 - \prod_{i=1}^{k} \left(-1 + \Phi\left[\frac{((-\log(D))^{1/n_i} - \tilde{\mu}_i)}{\tilde{\sigma}_i}\right] + \Phi\left[\frac{((-\log(D))^{1/n_i} + \tilde{\mu}_i)}{\tilde{\sigma}_i}\right]\right)$$

$$(\textit{Two-sided DFs}).$$

By means of differentiating the distribution functions $F_{\mathcal{D}}(D)$ and the induction method the related density functions can be determined.

Theorem 6 (Density functions, DI minimum DF) *If $Y_1, \dots, Y_k$ are k independent QMs with DFs d_i as in (2) and (1) as well as with densities $f_{D_i}(d_i)$ as in*

Theorems 1 and 2 then

$$
\begin{aligned}
f_{\mathcal{D}}(D) &= -\frac{1}{D\log(D)} \\
&\quad \cdot \sum_{i=1}^{k}\left[\frac{1}{\tilde{\sigma}_i}\phi\left(\frac{\log(-\log(D))-\tilde{\mu}_i}{\tilde{\sigma}_i}\right)\right. \\
&\quad \left.\cdot \prod_{\substack{j=1\\ j\neq i}}^{k}\Phi\left(\frac{\log(-\log(D))-\tilde{\mu}_j}{\tilde{\sigma}_j}\right)\right] \text{(One-sided DFs)}, \\
f_{\mathcal{D}}(D) &= -\sum_{i=1}^{k}\left(\frac{(-\log(D))^{1/n_i}}{n_i\cdot D\cdot\log(D)\tilde{\sigma}_i}\left(\phi\left[\frac{(-\log(D))^{1/n_i}-\tilde{\mu}_i}{\tilde{\sigma}_i}\right]\right.\right. \\
&\quad \left. +\phi\left[\frac{(-\log(D))^{1/n_i}+\tilde{\mu}_i}{\tilde{\sigma}_i}\right]\right) \\
&\quad \cdot\prod_{\substack{j=1\\ j\neq i}}^{k}\left(-1+\Phi\left[\frac{(-\log(D))^{1/n_j}-\tilde{\mu}_j}{\tilde{\sigma}_j}\right]\right. \\
&\quad \left.\left. +\Phi\left[\frac{(-\log(D))^{1/n_j}+\tilde{\mu}_j}{\tilde{\sigma}_j}\right]\right)\right) \quad \text{(Two-sided DFs)}, \\
\phi(x) &:= \text{Density of } \mathcal{N}(0,1) \text{in } x.
\end{aligned}
$$

5 Summary and application fields

The distributions of different types of the DI as well as Harrington's DFs are derived, which intensively will provide assistance in various application fields. One is the development of an improved optimization procedure of the DI, which allows the inclusion of the uncertainty of the optimized DI. As Steuer (2000) stated, "researchers were satisfied with approximative solutions as unbiased results would have required analytical expressions for the distribution of Desirability Indices". Furthermore the access to process control now is ensured. When a process was designed with the objective of reaching the optimal value of the DI the DI therefore is the most appropriate measure to monitor this optimality over time. Trautmann (2004b) sets up control charts for the DI and introduces an innovative procedure for the analysis of out-of-control signals. In addition the uncertainty of the optimum influence factor levels can be analyzed (Trautmann and Weihs 2004). For further research the extension of these approaches to the Derringer–Suich type of DFs would be attractive, which are frequently used due to their great flexibility.

Acknowledgements The work presented has been supported by the *Graduate School of Production Engineering and Logistics* at the University of Dortmund.

References

Averill AF, Ingram JM, Nolan RF (2001) Optimising the selection of metal cleaning processes or cleaning agents using desirability functions. Trans Inst Metal Finishing 79(4):123–128

Ben-Gal I, Bukchin J (2002) The ergonomic design of workstations using virtual manufacturing and response surface methodology. IEEE Trans 34:375–391

Castillo ED, Montgomery DC, McCarville DR (1996) Modified desirability functions for multiple response optimization. J Qual Technol 28(3):337–344

Derringer GC, Suich D (1980) Simultaneous optimization of several response variables. J Qual Technol 12(4):214–219

Harrington J (1965) The desirability function. Ind Qual Control 21(10):494–498

Holland B, Ahsanullah M (1989) Further results on the distribution of Meinhold and Singpurwalla. Am Stat 43(4):216–219

Johnson NL, Kotz S, Balakrishnan N (1994) Continuous univariate distributions. vol 1, 2nd edn, Wiley, New York

Kim K-J, Lin DKJ (2000) Simultaneous optimization of mechanical properties of steel by maximizing desirability functions. Appl Stat 49(3):311–326

Mood AM, Graybill FA, Boes DC (1974) Introduction to the theory of statistics. 3rd edn, McGraw-Hill, Singapore

Schwartz SC, Yeh YS (1982) On the distribution function and moments of power sums with log-normal components. Bell Syst Tech J 61(7):1441–1462

Steuer D (2000) An improved optimisation procedure for desirability indices. Technical Report 27/00, Statistics Department, Dortmund University, Germany

Trautmann H (2004a) Qualitätskontrolle in der Industrie anhand von Kontrollkarten für Wünschbarkeitsindizes – Anwendungsfeld Lagerverwaltung. Dissertation, http://hdl.handle.net/2003/2794

Trautmann H (2004b) The desirability index as an instrument for multivariate process control. Technical Report 43/04, SFB 475, Dortmund University, Germany

Trautmann H, Weihs C (2004) Uncertainty of the optimum influence factor levels in multicriteria optimization using the concept of desirability. Technical Report 23/04, SFB 475, Dortmund University, Germany

Tuyev VV, Fedchenko PP (1997) The study of soil humus changes with the harrington "desirability" function. Mapp Sci Remote Sens 34(4):264–272

Weber H, Weihs C (2003) On the distribution of the desirability index using harrington's desirability function. Technical Report 2003/3, Statistics Department, Dortmund University, Germany

Wu W, Cui GH, Lu B (2000) Optimization of multiple variables: application of central composite design and overall desirability. Chin Pharm J - Beijing - 35(8):530–532