About the distribution of Derringer-Suich type desirabilities

Detlef Steuer
Helmut-Schmidt-Universität Hamburg
Universität der Bundeswehr
steuer@hsu-hh.de
http://fawn.unibw-hamburg.de/steuer.html

11.3.2004

Contents

- (Short) Introduction to MCO and desirabilities
- Definition of Derringer-Suich desirabilities
- Current practice using desirabilities
- Distribution of desirabilities
- Improving practice with "realistic desirabilities"
- Concluding remarks

The MCO Problem

- Example: design of a new cookie.
- Objectives: Not too hard to eat, but hard enough to not fall apart before consumption.
- Problem: “Antagonistic” objectives must be optimised simultaneously.
- Natural aim: Trying to optimise overall **quality**.
- Need for compromise, need for expert knowledge!
- In general: Best object $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ must be identified. Each component y_i is of type (TV) or (LB).

Desirabilities

- Introduced by Harrington 1965
- Idea: First transform all quality measures to a unitless scale
- Then “oranges and apples” can be compared, i.e. calculate some mean value for the different measures.
- Transformation is chosen according to expert knowledge, calculation of mean allows some weighting of variables.

Derringer-Suich desirabillity functions

Derringer-Suich (1980) improved on Harrington using more flexible functions

DS-desirabilities for target value problems

$$d_{DS}^{TV}(y) := \begin{cases} 0, & \text{für} \quad y < l \\ (\frac{y-l}{t-l})^{\beta_l}, & \text{für} \quad l \leq y \leq t \\ (\frac{u-y}{u-t})^{\beta_r}, & \text{für} \quad t < z \leq u \\ 0, & \text{für} \quad u < y \end{cases}$$

DS-desirabilities for "the-larger-the-better" type problems

$$d_{DS}^{LB}(y) := \begin{cases} 0, & \text{für} \quad y < l \\ (\frac{y-l}{t-l})^{\beta_l}, & \text{für} \quad l \leq y \leq t \\ 1, & \text{für} \quad t < y \end{cases}$$

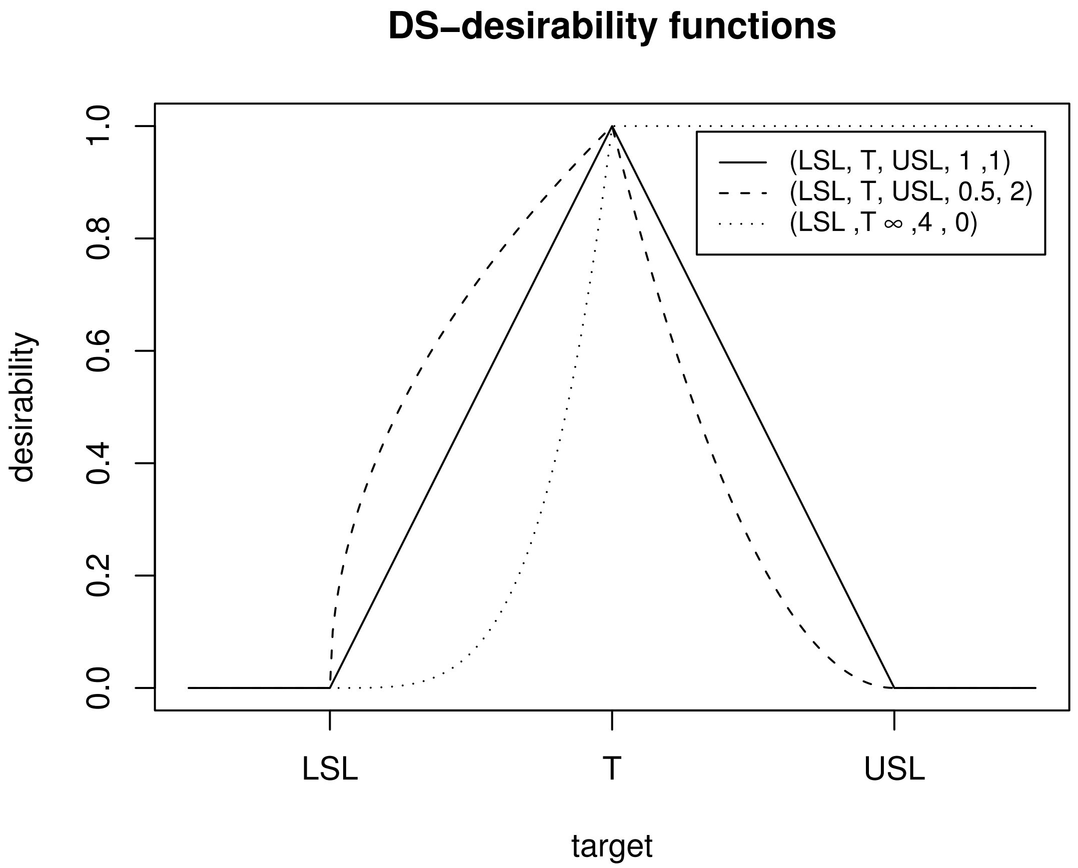
Unified notation:
A Quintupel $(l, t, u, \beta_l, \beta_r)$ defines a DS-desirability with:

$$\begin{aligned} \text{If } l < t < u \in \mathbb{R}, \beta_l, \beta_r \in \mathbb{R}^+, \text{then} \quad & (l, t, u, \beta_l, \beta_r)(y) := d_{DS}^{TV}(y). \\ \text{If } u = \infty, \text{then also} \beta_r = 0 \quad \text{and} \quad & (l, t, \infty, \beta_l, 0)(y) := d_{DS}^{LB}(y). \end{aligned}$$

Desirability Index Two different ways:

$$\begin{aligned} \text{(geometric mean)}\, Q(y) \quad &:= \quad (\prod_{i=1}^{m} d_i(y_i))^{\frac{1}{m}} \quad \text{or} \\ \text{(maximin)}\, Q(y) \quad &:= \quad \min_{i=1,\ldots,m} d_i(y_i) \end{aligned}$$

Desirability functions of the DS-type for some parameters



The right question?

Current Practice. Is this what we want?

- Product $y = (y_1, \ldots, y_m)$ to be optimised, y depends on factor settings $x = (x_1, \ldots, x_k)$, $y = f(x) + \epsilon$, ϵ multivariate normal with diagonal covarince.
- Define desirabilities $d_i(y_i)$ for each component.
- Performing experiments according some DOE
- Fit linear or quadratic response curves $\hat{f}_i, i = 1, \ldots, m$ for the components.
- Perform numerical optimisation for $Q(\widehat{f(x)})$ over region of operability $\mathbb{O}$ and estimate best factor setting $\widehat{x_{opt}}$. (idealized desirabilities)

No!

This way we ignore the **error** and the **non-linerarity** of the desirabilities.

Today's practice gives:

$$\widehat{x_{opt}} := \max_{x \in \mathbb{O}} Q(E(Y|x))$$

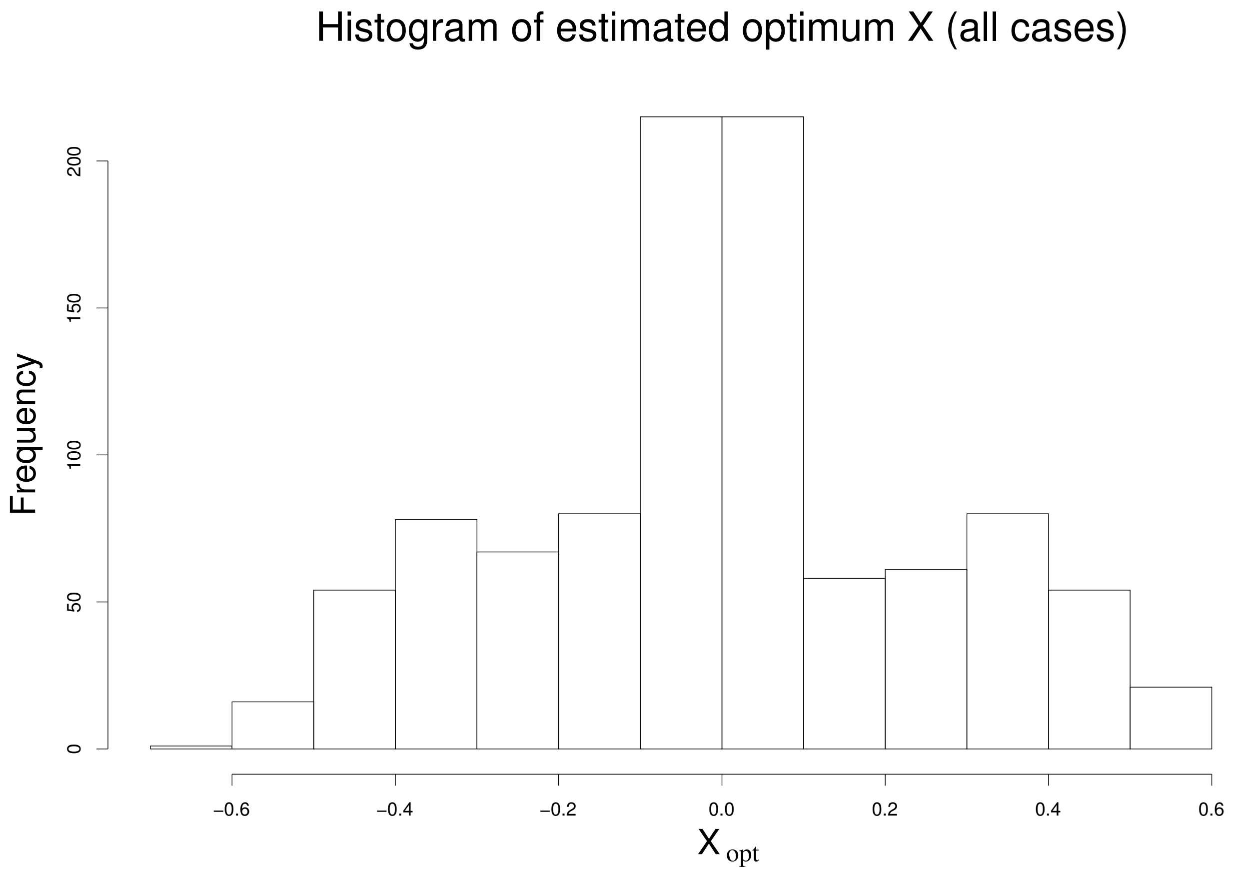
Simplified and probably wrong solution!

We really search:

$$\widehat{x_{opt}} := \max_{x \in \mathbb{O}} E(Q(Y|x))!$$

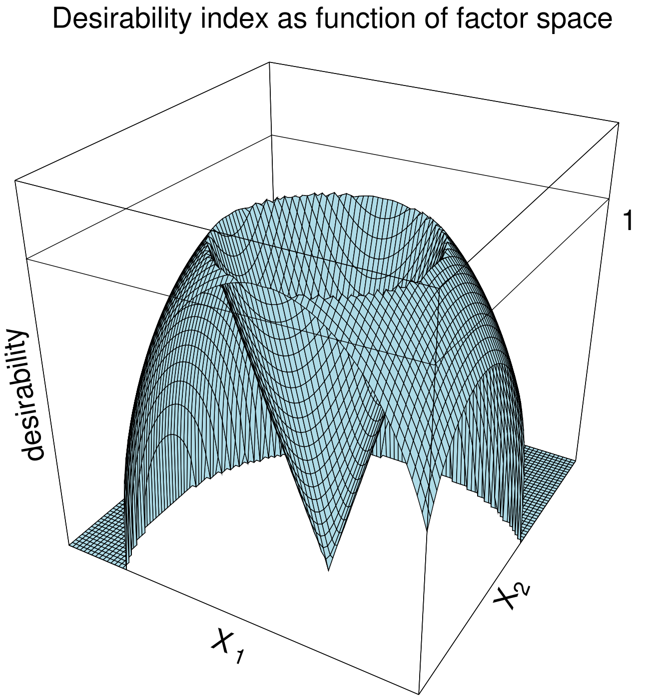
Distribution of X_{opt} may be multimodal

Model: $y = x^2 + \epsilon, d = (-1, 0, 1, 0.1, 1), \epsilon \sim N(0, 0.1), f$ fully quadratic



$Q(y)$ may have complicated structure

Model: $Y = x_1^2 + x_2^2, \quad d = (-1, 0, 4, 1/2, 1/2)$

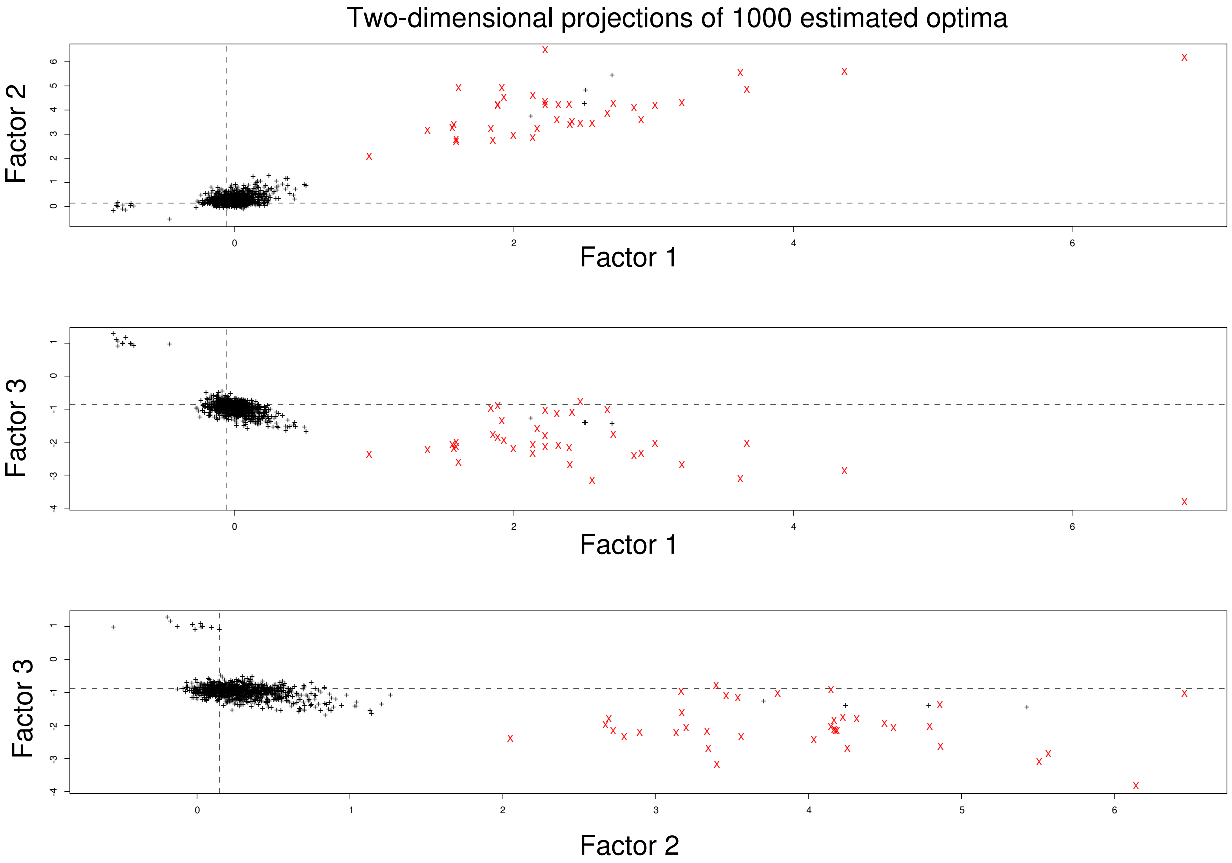


Why ignoring the error can be harmful

Simulation: Repeat optimisation from Derringer/Suich

- Four objectives y_1 to y_4, two of type (TV), two of type (LB), three controllable variables x_1 to x_3.
- Central-composite design with 20 experiments
- Second-order models $\hat{f}_i$ including all interactions.
- Repeatedly generate data using the estimated model.
- Each time record $\widehat{x_{opt}}$. What happens?

Why ignoring the error can be harmful



About 5% of the estimated $\widehat{x_{opt}}$ have true desirability **0**!

Incorporating the error
Distribution of desirability function values

Each d_i and Q are random variates!

Model: $Y = f(x) + \epsilon, \epsilon \sim N(0, \sigma^2)$

Random desirability defined as

$$d(x,\epsilon) = \begin{cases} \frac{f(x)+\epsilon-l}{t-l} & \text{for } l \leq f(x)+\epsilon < t; \\ \frac{u-f(x)-\epsilon}{u-t} & \text{for } t \leq f(x)+\epsilon < u; \\ 0 & \text{else.} \end{cases}$$

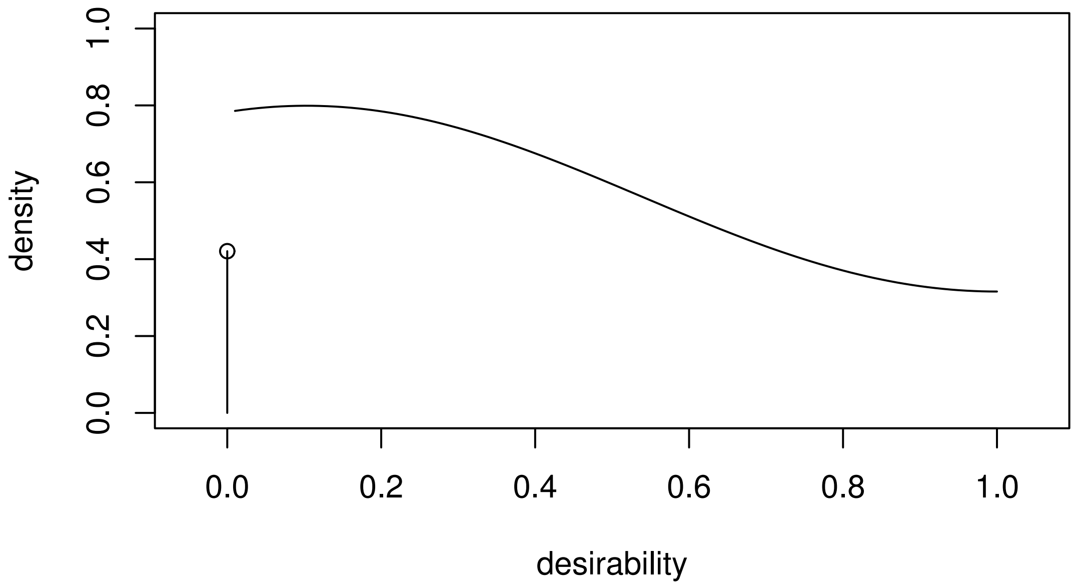
Distribution of $d(x,\epsilon)$ of type $(l,t,u,1,1)$

$$F_{d(x,\epsilon)}(d) = \begin{cases} 0 & \text{for } d < 0; \\ \Phi\left(\frac{l + d \cdot (t-l) - f(x)}{\sigma}\right) + & \\ \quad 1 - \Phi(\frac{u - d \cdot (u-t) - f(x)}{\sigma}) & \text{for } d \in [0,1]; \\ 1 & \text{for } d > 1. \end{cases}$$

Density $f_{d(x,\epsilon)}$ for d of type $(l,t,r,1,1)$ and fixed x:

$$f_{d(x,\epsilon)}(d) = \begin{cases} 0 & \text{for } d \notin [0,1); \\ \Phi(\frac{l-f(x)}{\sigma}) + 1 - \Phi(\frac{u-f(x)}{\sigma}) & \text{for } d = 0 \text{ (failure rate)}; \\ \frac{t-l}{\sigma}\varphi(\frac{l+d\cdot(t-l)-f(x)}{\sigma}) + \frac{u-t}{\sigma}\varphi(\frac{u-d\cdot(u-t)-f(x)}{\sigma}) & \text{for } d \in (0,1). \end{cases}$$

Density for a desirability $(-1, 0, 1, 1, 1)$, with $f(x) = y + \epsilon, \epsilon \sim N(0, 0.5), E(f(x)) = 0.9$



Expected value for a given x and d of type $(l, t, r, 1, 1)$

$$E(d(x,\epsilon)) = G_l \frac{f(x)-l}{t-l} + g_l \frac{\sigma}{t-l} + G_r \frac{u-f(x)}{u-t} - g_r \frac{\sigma}{u-t} \text{ with}$$

$$\begin{aligned}
G_l &= \Phi\left(\frac{t-f(x)}{\sigma}\right) - \Phi\left(\frac{l-f(x)}{\sigma}\right), \\
g_l &= \varphi\left(\frac{l-f(x)}{\sigma}\right) - \varphi\left(\frac{t-f(x)}{\sigma}\right), \\
G_r &= \Phi\left(\frac{u-f(x)}{\sigma}\right) - \Phi\left(\frac{t-f(x)}{\sigma}\right), \\
g_r &= \varphi\left(\frac{t-f(x)}{\sigma}\right) - \varphi\left(\frac{u-f(x)}{\sigma}\right).
\end{aligned}$$

Is this any better?

- Define *realistic* desirabilities using the now know distribution functions and optimise $Q^{real}(y) = (\prod_1^m E(d_i(x,\epsilon)))^{\frac{1}{m}}$
- Repeat optimisation from Derringer/Suich using Q^{real}
- Result using realistic instead of idealised desirabilities:
  Different estimation for $\widehat{x_{opt}}$:
  $\widehat{x_{opt}^{ideal}} = (-0.05, 0.145, -0.868)$ (Derringer/Suich)
  $\widehat{x_{opt}^{real}} = (0.13, 0.50, -1.08)$ (realistic desirabilities)
- **It's better! True** 10% relative improvement for values of Q:

$$Q(\widehat{x_{opt}^{real}}) = 0.44 \quad \text{and} \quad Q(\widehat{x_{opt}^{ideal}}) = 0.40$$

Better approach, but not simpler

- Optimisation is now looking for a "best" *random variate.*
- Order of distributions is needed.
- Concentrate on key features of distribution: Expected value? Failure rate? Mode?
- Only "some" factor setting can be compared:
  If $F(d(x_1, \epsilon)) >_{st} F(d(x_2, \epsilon))$ (stochastically larger) then factor setting x_1 is better than setting x_2.
- **New MCO!**

Remarks

- For analytical expressions for the distributions the weights $\beta_l = \beta_r = 1$ are crucial. For other exponents simulations are possible.
- The distribution of Q for the geometric must be simulated, because F_d are not in a family of stable distributions.
- For the *maximin* approach results for extrem value statistics may be used.
- The distribution of x_{opt} seems completely out of reach for simple analytical formulation. Finding x_{opt} is a callibration problem of the many-to-many type and Q^{-1} is a not very nicely behaving function. Nevertheless simulations are possible.