Monotone Function Problems and Statistical Selection Procedures

Dieter Schott¹

Accepted: 2 February 2021 / Published online: 22 April 2021
© The Author(s) 2021

Abstract

Monotone function problems are introduced on a very elementary level to reveal the close connection to certain statistical problems. Equations $F(x) = c$ and inequalities $F(x) \geq c$ with monotone increasing functions F are considered. Solution methods are stated. In the following, it is shown how some important problems of statistics, especially also statistical selection problems, can be solved by transformation to monotone function problems.

Keywords Monotone functions · Inequalities · Cumulative distribution functions · Statistical selection procedures

1 Monotone Function Problems

Some statistical methods are based on very simple mathematical principles. Also for practical statisticians, it is worth learning something about the mathematical background of these methods. Here, we consider the monotony principle which is essential for cumulative distribution functions. Applications are given from Sect. 3 on.

We start with some notations. $\mathbb{R}$ is the set of real numbers, $\mathbb{R}_+$ the set of nonnegative real numbers and $\mathbb{N}$ the set of nonnegative integers.

Monotone functions $F(x)$ have a lot of simple and remarkable properties. Here, we concentrate on the solution of equations und inequalities determined by monotone increasing functions. Some elementary statements follow where $I \subseteq \mathbb{R}$ is an interval and $F : I \rightarrow \mathbb{R}$ is a function with range $R(F) \subseteq \mathbb{R}$. Here, I can be a closed interval $[a, b]$, an open interval $]a, b[$ or one of the half open intervals $]a, b]$, $[a, b[$ where $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$. Further, $\overline{M}$ is the closure of the set $M \subseteq \mathbb{R}$.

Part of special issue guest edited by Dieter Rasch, Jürgen Pilz, and Subir Ghosh—Advances in Statistical and Simulation Methods.

✉ Dieter Schott
dieter.schott@hs-wismar.de

¹ Faculty of Engineering Sciences, Hochschule Wismar, Wismar, Germany

Proposition 1 *Let F be strictly monotone increasing and continuous on I . Then the equation $F(x) = c$ has for each $c \in R(F)$ a unique solution $x = x_c$. Since F is invertible in this case, it is $x_c = F^{-1}(c)$. Besides, the inequality $F(x) \geq c$ is fulfilled for all $x \in I$ with $x \geq x_c$.*

A slight generalization follows.

Proposition 2 *Let F be monotone increasing and left continuous on I . Then the equation $F(x) = c$ has for $c \in R(F)$ a minimal solution $x = x_c$. Besides, the inequality $F(x) \geq c$ is fulfilled for all $x \in I$ with $x \geq x_c$.*

Normally the values of $x = x_c$ are determined by zero-problem solvers for $F(x) - c = 0$ implemented in computer packages. But often these solvers need additional assumptions about F or good initial guesses for success.

Sometimes we look for natural n instead of real x satisfying the inequality $F(n) \geq c$. This problem can be solved by a primitive algorithm, called here *counting-up algorithm*. If we start with $n = 0$ and increase n step by step as long as $F(n) \geq c$ is fulfilled for the first time, then this last n is just the minimal natural solution n_c of the inequality. The counting-up algorithm needs no further assumptions about F and can easily be implemented using computer software (as MATLAB used by engineers or R used by statisticians).

An alternative is to use also a zero-problem solver for $F(x) - c = 0$ obtaining $x = x_c$ and rounding off to the next natural $n_c = \lceil x_c \rceil$ which is the smallest integer $n \geq x_c$.

If F is differentiable, then the gradient $F'(0)$ can deliver a first guess about the size of $n = n_c$. But there are functions F whose gradients start largely in 0 and rapidly tend to zero.

Example 1 A first simple application can be found in stochastics concerning cumulative distribution functions (c.d.f.) $F(x) = P(\mathbf{x} \leq x)$ of random variables $\mathbf{x}$. These functions are at least monotone and left continuous. Besides, they are nonnegative and bounded. Hence, the function F in Proposition 2 fits to the case of random $\mathbf{x}$. The solution of $F(x) = c$ ($0 < c < 1$) supplies for continuous $\mathbf{x}$ the c -quantile x_c . The inequality $F(x) \geq c$ means that the inequalities $P(\mathbf{x} \leq x) \geq c$ and $P(\mathbf{x} \geq x) \leq 1 - c$, respectively, have to be fulfilled.

2 Integral Representations of Monotone Functions

A general application concerns problems with monotone functions F generated by integral representations

$$F(x) := \int_{-\infty}^{+\infty} K(x, z) \cdot h(z) \, dz = \int_{C(h)} K(x, z) \cdot h(z) \, dz \quad (1)$$

with fixed kernel K and function h where $x \in D = [d, +\infty[, d \geq -\infty$ and

$$C(h) := \overline{\{z : h(z) \neq 0\}}$$

is the support of h . There is a stochastic interpretation. Namely, if K is considered as a function $K_x(z)$ of z with fixed parameters x , then the values

$$F(x) = \int_{-\infty}^{+\infty} K_x(z) \cdot h(z) \, dz$$

are the expectations $E(K_x(z))$ of the functions $K_x(z)$ depending on the continuous random variable z with the probability density function $h(z)$.

Example 2 Let $X = (x_1, x_2)$ be a two-dimensional continuous random vector where x_1 and x_2 are independently distributed. Further, let $f_{x_1}(x)$, $f_{x_2}(x)$ be the corresponding probability density functions (p.d.f.) and $F_{x_1}(x)$, $F_{x_2}(x)$ be the corresponding cumulative distribution functions (c.d.f.). In the special case $X = x_1 + x_2$, it is known that

$$F_X(x) = \int_{-\infty}^{+\infty} F_{x_1}(x - z) \cdot f_{x_2}(z) \, dz$$

which is of form (1) and corresponds to the convolution of F_{x_1} and f_{x_2} . Correspondingly, the difference $X = x_1 - x_2$ has the c.d.f.

$$F_X(x) = \int_{-\infty}^{+\infty} F_{x_1}(x + z) \cdot f_{x_2}(z) \, dz.$$

This representation is again of form (1). Finally, for the quotient $X = \frac{x_1}{x_2}$ we get

$$F_X(x) := \int_{-\infty}^{+\infty} K(x, z) \cdot f_{x_2}(z) \, dz, \quad K(x, z) = \begin{cases} 1 - F_{x_1}(x \cdot z) : z < 0 \\ F_{x_1}(x \cdot z) : z \geq 0 \end{cases}$$

which also fits into form (1). If you will find other representations in the literature, you will reach at our formulas by simple substitutions.

Convention In the following, we often omit the infinite limits $-\infty$ and $+\infty$ in integrals of form (1).

Proposition 3 *It is supposed that*

- (a) h is integrable, nonnegative and continuous on the support $C(h)$,
- (b) K is continuous on the Cartesian product $D \times C(h)$,
- (c) $0 \leq K(x, z) \leq k$ for all $x \in D$ and $z \in C(h)$,

(d) $K(x, z) \leq K(y, z)$ for all $x, y \in D$ with $x < y$ and all $z \in C(h)$.

Then the function F in (1) is monotone increasing, continuous, nonnegative and bounded.

Proof Obviously $F(x)$ in (1) is defined and continuous for all $x \in D$ considering the assumptions (a) and (b).

Since the integrand in (1) is nonnegative on $D \times \mathbb{R}$ observing (a) and (c), the integral values $F(x)$ are also nonnegative. Besides, we find by (c)

$$F(x) = \int_{C(h)} K(x, z) \cdot h(z) \, dz \leq k \int_{C(h)} h(z) \, dz =: k_1$$

for all x , and therefore, $F(x)$ is bounded. Finally, choosing $x < y$ and taking (c), (d) and the monotony of the integral into account, we get

$$F(x) = \int_{C(h)} K(x, z) \cdot h(z) \, dz \leq \int_{C(h)} K(y, z) \cdot h(z) \, dz = F(y).$$

Therefore, $F(x)$ is monotone increasing.

Remark 1 It is easy to see that the limit $F(+\infty) := \lim_{x \rightarrow +\infty} F(x)$ exists and can be calculated by interchanging limit operation and integration in (1). By the way, we have $F(+\infty) = \sup F$. For $D = \mathbb{R}$ ($d = -\infty$) in (1), we find also $F(-\infty) := \lim_{x \rightarrow -\infty} F(x) = \inf F$.

Remark 2 If $F(x)$ is strictly monotone increasing, then $F(d) < F(+\infty)$ and $F(+\infty)$ is not attained by $F(x)$ (at any x). This is, for example, the case if we modify (d) in Proposition 3 to

(d') $K(x, z) < K(y, z)$ for all x, y with $x < y$ and for all z .

Proposition 3a Let be $D = \mathbb{R}$ ($d = -\infty$) in (1). Beside the assumptions of Proposition 3 it is supposed that

(c') $\lim_{x \rightarrow -\infty} K(x, z) = 0$ for all z , $\lim_{x \rightarrow +\infty} K(x, z) = 0$ for all z .

Then the function F in (1) is a c.d.f. of a continuous random variable.

Proof The assumptions of Proposition 3 guarantee that limit operation and integration can be interchanged. Consequently, we get because of (c')

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1.$$

This shows that all properties of a c.d.f. are fulfilled under the given assumptions.

Remark 3 If $K(x, z)$ is partially differentiable w.r.t. x and the partial derivative $K_x(x, z)$ is continuous, then $F(x)$ is also continuously differentiable. The derivative is obtained by

$$F'(x) = \int K_x(x, z) \cdot h(z) \, dz.$$

Consequently, the gradients for $x = d$ and $x = +\infty$ are

$$F'(d) = \int K_x(d, z) \cdot h(z) \, dz,$$

$$F'(+\infty) = \lim_{x \rightarrow +\infty} \int K_x(x, z) \cdot h(z) \, dz.$$

If $K_x(x, z)$ satisfies analogue conditions as $K(x, z)$ in Proposition 3, then the limit $K_x(+\infty, z)$ exists and we get

$$F'(+\infty) = \int K_x(+\infty, z) \cdot h(z) \, dz.$$

3 Stochastic Applications

Now we discuss further applications of the integral representation (1) in stochastics.

Example 3 We consider the function $F : \mathbb{R} \rightarrow \mathbb{R}_+$ defined by

$$F(x) = \int F_1(x + z) \cdot f_2(z) \, dz \quad (2)$$

where F_1 is a c.d.f. and f_2 is a p.d.f. of possibly different continuous distributions. We have in representation (1)

$$K(x, z) := F_1(x + z), h(z) := f_2(z).$$

We can check the assumptions of Proposition 3. For f_2 and $F_1(x + z)$, conditions (a) and (b) are true. Because of $0 \leq F_1(x + z) \leq 1$ for all x, z , condition (c) is satisfied. Obviously $F_1(x + z)$ fulfils condition (d), since F_1 is monotone increasing. Therefore, the results of Proposition 3 hold. Besides, condition (c') in Proposition 3a is true. Hence, $F(x)$ is a c.d.f. of a certain continuous random variable x satisfying $F(x) = P(x \leq x)$.

By the way, the value

$$F(0) = \int F_1(z) \cdot f_2(z) \, dz$$

describes the *effect* of the first to the second distribution.

Finally, F is differentiable. The derivative is the p.d.f.

$$f(x) = F'(x) = \int f_1(x+z) \cdot f_2(z) \, dz.$$

Example 4 We introduce the function $F : \mathbb{R} \rightarrow \mathbb{R}_+$ given by

$$F(x) = \int F_0(x+z)^l \cdot (1 - F_0(z))^m \cdot f_0(z) \, dz \quad (l, m \in \mathbb{N}) \quad (3)$$

where F_0 is the c.d.f. and f_0 is the p.d.f. of a certain continuous distribution. We can put

$$K(x, z) := F_0(x+z)^l \cdot (1 - F_0(z))^m, \quad h(z) := f_0(z)$$

to get the representation (1). Again the assumptions of Proposition 3 are fulfilled. In particular, we have the estimation

$$0 \leq K(x, z) \leq F_0(x+z)^l \leq 1.$$

Further, it is

$$F(-\infty) = 0, \quad F(0) = J_{l,m} := \int F_0(z)^l \cdot (1 - F_0(z))^m \cdot f_0(z) \, dz \quad (l, m \in \mathbb{N}),$$

$$0 < F(+\infty) = J_{0,m} = \int (1 - F_0(z))^m \cdot f_0(z) \, dz \leq 1.$$

It is easy to show that

$$J_{l,m} = \frac{l! \cdot m!}{(l+m+1)!} \quad (4)$$

which means especially $J_{0,m} = \frac{1}{m+1}$. This holds independently of the concrete c.d.f. F_0 . In [10: p. 61–62] the proof is given for the case that F_0 is the c.d.f. of the standard normal distribution. But the idea is the same as for general F_0 . Because of $F(+\infty) = J_{0,m} < 1$ for $m > 0$, the function F is no c.d.f. in this case.

Further, we get for $l \geq 1$ the derivative

$$F'(x) = l \int F_0(x+z)^{l-1} \cdot (1 - F_0(z))^m \cdot f_0(x+z) \cdot f_0(z) \, dz.$$

This implies

$$F'(0) = l \cdot \int F_0(z)^{l-1} \cdot (1 - F_0(z))^m \cdot f_0^2(z) \, dz.$$

Since the integrand of (3) consists of three factors, we can also split it into the factors

$$\tilde{K}(x, z) := F_0(x + z)^l, \quad \tilde{h}(z) = \tilde{f}_0(z) := (1 - F_0(z))^m \cdot f_0(z).$$

A simple investigation shows that $F_1 = F_0^l$ is again a c.d.f. of a certain distribution and $\tilde{f}_0$ is a positive multiple of a p.d.f. of another distribution f_2 , say $\tilde{f}_0 = c \cdot f_2$ ($c > 0$). Then we get

$$F(x) = c \int F_1(x + z) \cdot f_2(z) \, dz$$

which is only a trivial modification of (2) in Example 3. Hence, F is a multiple of a c.d.f.

By the way, we can write $\tilde{f}_0$ as

$$\tilde{f}_0(z) = \frac{1}{m+1} \cdot \frac{d}{dz} F_2(z) = \frac{1}{m+1} \cdot f_2(z)$$

where

$$F_2(z) = 1 - (1 - F_0(z))^{m+1}$$

is a c.d.f., $f_2(z)$ is the corresponding p.d.f. and $c = \frac{1}{m+1}$.

Now we consider the integral representation

$$F(x) := \int_{-\infty}^{+\infty} F_1(z + v(x)) \cdot dF_2(z) = \int_{C(h)} F_1(z + v(x)) \cdot f_2(z) \, dz \quad (5)$$

where the real functions F_1, F_2 with $F_2' = f_2$ and v occur. Here, the kernel in (1) has the form $K(x, z) = F_1(z + v(x))$ with a generally nonlinear shift function v in the argument of F_1 . If F should be differentiable, then also v has to be differentiable. The wished properties for F have consequences for v . The next result supplies a stochastic statement, where v has additional properties.

Definition The function v is said to be *coercive* iff

$$|v(x)| \rightarrow +\infty \text{ for } |x| \rightarrow +\infty.$$

Hence, a coercive function is unbounded.

Proposition 4 *It is supposed that*

(a) F_1 is the c.d.f. of a continuous random variable,

- (b) f_2 is the p.d.f. of a continuous random variable,
 (c) v is a monotone increasing coercive function with $v(0) = 0$.

Then F in (5) is the c.d.f. of a continuous random variable.

Proof We can proceed similar to Example 3 which is the special case with $v(x) = x$ (see (2)). Namely, because of the coercivity and monotony of v it follows

$$\lim_{x \rightarrow \pm\infty} v(x) = \pm\infty.$$

Remark 4 Considering $v(0) = 0$ it is

$$F(0) = \int F_1(z) \cdot dF_2(z) = \int F_1(z) \cdot f_2(z) dz > 0$$

the effect of F_1 to F_2 analogous to Example 3. If v is continuously differentiable, we get for the derivative

$$F'(x) = v'(x) \cdot \int f_1(z + v(x)) \cdot f_2(z) dz.$$

Example 5 Finally, we regard a generalization of Example 4 with nonlinear shift function v analogously to (5), namely

$$F(x) = \int F_0(z + v(x))^l \cdot (1 - F_0(z))^m \cdot f_0(z) dz \quad (l, m \in \mathbb{N}) \quad (6)$$

with a continuously differentiable monotone increasing coercive function v which fulfils $v(0) = 0$. The representation (3) arises for $v(x) = x$. It is again easy to show that Proposition 3 is true for (6). The values for $F(0)$ and $F(+\infty)$ are the same as in Example 4. The derivative is in this case

$$F'(x) = l \cdot v'(x) \int F_0(z + v(x))^{l-1} \cdot (1 - F_0(z))^m \cdot f_0(z + v(x)) \cdot f_0(z) dz$$

which implies

$$F'(0) = l \cdot v'(0) \int F_0(z)^{l-1} \cdot (1 - F_0(z))^m \cdot f_0^2(z) dz.$$

These results are given in Example 4 for $v(x) = x$.

4 Statistical Selection Procedures

We refer to [6: chapter 11] and consider a set of a populations

$$G = \{A_1, A_2, \dots, A_a\} \quad (a > 1 \text{ integer})$$

which correspond to continuous random variables u_i with c.d.f. F_i ordered by certain scores q_i ($i = 1, \dots, a$) where

$$q_{(1)} \leq q_{(2)} \leq \dots \leq q_{(a)}.$$

We want randomly choose the t best populations according to these scores ($1 \leq t < a$), i.e. the first class subset

$$G_1 = \{A_{(a-t+1)}, \dots, A_{(a)}\},$$

with high probability $c := 1 - \beta$ of correct selection ($0 < \beta < 1$). The $a - t$ remaining populations constitute the second class subset

$$G_2 = \{A_{(1)}, \dots, A_{(a-t)}\}.$$

We follow here the indifference zone formulation assuming a gap between G_1 and G_2 such that for a separation parameter $\delta > 0$ the condition

$$d(G_1, G_2) := q_{(a-t+1)} - q_{(a-t)} \geq \delta$$

holds. Now random samples $u_{i,k}$ ($k = 1, \dots, n$) are produced in the populations A_i of constant size n which are ranked by (estimated) scores $q(u_{i,k})$. If G_{1s} is the set of the t best populations by random selection, we want to realize

$$P_{CS} := P(G_{1s} = G_1 | d(G_1, G_2) \geq \delta) \geq c \quad (7)$$

where CS means correct selection. An idea to exploit Proposition 1 or Proposition 2 is to find a monotone function $F(x) = F(x; a, t, \delta)$ which fulfils $P_{CS} \geq F(n) \geq c$ for the common sample size $x = n$. A natural precondition is

$$\binom{a}{t} \cdot c > 1 \quad (8)$$

since otherwise each selection of t populations is correct. Whether the inequality (7) is fulfilled or not depends on n . The task of this *indifference zone problem* is to find minimal or at least appropriate values n such that (7) is satisfied. Then we can sample realizations G_{1s} of G_{1s} .

In the meantime, there are also interesting studies for two-stage approaches [8] starting with the Gupta subset approach for pre-selection [3] and followed by the Bechhofer approach to reduce the effort of sampling.

5 An Analytical Approach to the Bechhofer Selection Problem

Bechhofer supposed in [1] populations A_i ($i = 1, 2, \dots, a$) with normally distributed random characteristics $u_i \sim N(\mu_i, \sigma^2)$ of known constant variance σ^2 . The F_i are the corresponding c.d.f. If we have drawn random samples of n observations for all u_i , natural scores are the expectation values $q_i = \mu_i$ which can be estimated

by sample means $\bar{u}_i$. Bechhofer found under the indifference zone supposition $\mu_{(a-t+1)} - \mu_{(a-t)} \geq r\sqrt{n}$, $r := \frac{\delta}{\sigma} > 0$ the solution condition

$$\begin{aligned} P_{CS} &= P(\max(\bar{u}_{(1)}, \dots, \bar{u}_{(a-t)}) < \min(\bar{u}_{(a-t+1)}, \dots, \bar{u}_{(a)})) \\ &\geq t \int_{-\infty}^{+\infty} \Phi(z + r\sqrt{n})^{a-t} (1 - \Phi(z))^{t-1} \varphi(z) \, dz =: F(n) \geq c \end{aligned} \quad (9)$$

for a correct selection where $\varphi(z)$ is the p.d.f. and $\Phi(z)$ is the c.d.f. of the standard normal distribution $N(0, 1)$. The inequality $F(n) \geq c$ in (9) can be used to determine (minimal) n . This is the *Bechhofer selection problem* (BSP).

It supplies interesting analytical surplus if we extend the *Bechhofer function* $F: \mathbb{N} \mapsto \mathbb{R}$ to a real function $F: \mathbb{R} \mapsto \mathbb{R}$ replacing n formally by the real variable $x \geq 0$:

$$F(x) = F(x; a, t, r) := t \int_{-\infty}^{+\infty} \Phi(z + r\sqrt{x})^{a-t} (1 - \Phi(z))^{t-1} \varphi(z) \, dz. \quad (10)$$

Here a , t and r are parameters with the given meanings which are supposed to be fix in the following. When approximation by numerical procedures are necessary, we use instead of the improper integral in (10) the proper one with finite integral limits $-N$ and N ($N > 0$ natural number) and denote this cut function by $F_N(x)$. The proofs of the following statements are given in [10], but can also be easily derived from the more general results in this paper.

Proposition 5 *The functions F and F_N are (independently of the parameters a, t, r)*

- (a) *continuous and nonnegative,*
- (b) *strictly monotone increasing and invertible,*
- (c) *smooth for $x > 0$, where derivation can be done by partial derivation after x under the integral,*
- (d) *bounded, where $0 < F_N(x) < F(x) < 1$ for all x ,*
- (e) *almost equal for large N , where $F(x) - F_N(x) \leq 2(1 - \Phi(N)) \rightarrow 0$ for $N \rightarrow +\infty$.*

Proposition 6 *It is*

- (a) $F(0) = \left(\frac{a}{t}\right)^{-1}$, $F(+\infty) := \lim_{x \rightarrow +\infty} F(x) = 1$,
- (b) $F'(0+0) := \lim_{x \rightarrow 0, x > 0} F'(x) = +\infty$, $F'(+\infty) := \lim_{x \rightarrow +\infty} F'(x) = 0$.

Moreover, F can be extended to a c.d.f. of a random variable $\mathbf{x}$, e.g. replacing $\sqrt{x}$ defined on $\mathbb{R}_+$ by $\text{sign}(x) \cdot \sqrt{|x|}$ defined on $\mathbb{R}$. The question is which distribution is behind $\mathbf{x}$. Since F' is the p.d.f. of this distribution, Proposition 6 shows that f is represented by an improper integral having a singularity at 0.

6 Solution Methods for the Bechhofer Selection Problem

Observing the results in Proposition 1, we can state a discrete and a continuous variant to solve BSP with the Bechhofer function (10). Besides, we use results from Sect. 5. Remember that brackets $], [$ describe open intervals.

Proposition 7 *For each constant $c \in \left[\left(\frac{a}{t} \right)^{-1}, 1 \right]$, there is a minimal natural $n = n_c$ such that $F(n) \geq c$ for all $n \geq n_c$.*

Proposition 8 *For each constant $c \in \left[\left(\frac{a}{t} \right)^{-1}, 1 \right]$, there is a unique $x = x_c > 0$ satisfying $F(x) = c$ and $x = F^{-1}(c)$. Besides, $F(x) \geq c$ holds for $x \geq x_c$.*

First we see that the natural precondition (8) of BSP is fulfilled by the interval condition for c . This condition corresponds to $\beta \in \left] 0, 1 - \left(\frac{a}{t} \right)^{-1} \right[$ where β is the risk of wrong selection. Proposition 7 shows that BSP can be solved by the counting-up algorithm mentioned in Sect. 1. The supplied $n = n_c$ is just the minimal sample size of BSP.

Proposition 8 allows to use more sophisticated zero-problem solvers for $F(x) - c = 0$. After determination of the zero $x = x_c$, we round off to the next natural $n = n_c = \lceil x_c \rceil$ to get the minimal sample size.

Although the second algorithm will work often faster, especially for large values of n_c , the time saving is rather minor. Numerical effects in the algorithms can lead to small errors. For example, we can reach at $\tilde{n}_c = n_c + 1$ in bad cases which is really not important in practical applications. To avoid this effect, the cut number N can be increased using F_N and function calculations in the integrand as well as in the integration procedures can be realized with more accuracy.

In [9: p. 20–21], a counting-up algorithm implemented in MATLAB is given. In [7: p. 369–370], the R-function “Bech.snr(a,t,r,beta)” is defined to get the optimal size of samples. Simulation results are, for example, presented in [7–10].

The zero-problem solvers have a considerable advantage if the sample sizes $n_c = n_c(a, t, r)$ are needed for several parameters r . Namely, considering the expression $r\sqrt{x}$ in (10) we get $x_c(a, t, kr) = \frac{1}{k^2}x_c(a, t, r)$ for $k > 0$ which means also $n_c(a, t, kr) = \lceil \frac{1}{k^2}x_c(a, t, r) \rceil$. Hence, we can restrict ourselves to $r = 1$ and generate the x_c for $r \neq 1$ by the formula given above. If a and $a - t$ are large, the asymptotic

$$n_c(a, t, 1) \approx 2 \ln a + K_{t,c}$$

for a constant $K_{t,c}$ depending only on t and c can be proved and also confirmed by simulation (see [10: p. 65]).

7 Non-normal Distributions

Now we study random variables and corresponding distributions with absolutely continuous c.d.f. F such that also the p.d.f. $f = F'$ exists. The following results are well known. Observe that Proposition 9 is a reformulation of a statement in Example 3 concerning the distribution of a difference of random variables using other notations.

Proposition 9 *If x and y are independently distributed (i.d.) with the c.d.f. F_x and F_y , then the difference $d := y - x$ has the c.d.f.*

$$F(d) = P(d \leq d) = \int_{-\infty}^{+\infty} F_y(z + d) \cdot dF_x(z) = \int_{-\infty}^{+\infty} F_y(d + z) \cdot f_x(z) dz.$$

Now we can consider problems $F(d) = c$ and $F(d) \geq c$ applying results of Sect. 1. Observe that the integral representation in Proposition 9 is a special case of (2) in Example 3.

The next step will include several independent random variables with the same c.d.f. They are independently and identically distributed (i.i.d.).

Proposition 10 *Let $x_i (i = 1, \dots, n)$ be i.i.d. with the c.d.f. F and the p.d.f. f .*

Then $P = P_n = \max_{i=1, \dots, n} (x_i)$ has the c.d.f.

$$F_P(x) = P(P \leq x) = F(x)^n$$

and the p.d.f.

$$f_P(x) = n \cdot F(x)^{n-1} f(x).$$

Further, $Q = Q_n = \min_{i=1, \dots, n} (x_i)$ has the c.d.f.

$$F_Q(x) = P(Q \leq x) = 1 - (1 - F(x))^n$$

and the p.d.f.

$$f_Q(x) = n \cdot (1 - F(x))^{n-1} f(x).$$

Now we can combine the results of Proposition 9 and Proposition 10 to get a statement which allows to treat also selection problems with non-normal distributions.

Proposition 11 Let $x_i (i = 1, \dots, m)$ be i.i.d. with the c.d.f. F_1 and the p.d.f. f_1 . Further, let $y_j (j = 1, \dots, l)$ be i.i.d. with the c.d.f. F_2 where also the y_j are independent of the x_i . Then the variable $d := y - x$ with $y = \max_{j=1, \dots, l} (y_j)$ and $x = \min_{i=1, \dots, m} (x_i)$ has the c.d.f.

$$F(d) = m \int_{-\infty}^{+\infty} F_2(z+d)^l (1-F_1(z))^{m-1} f_1(z) dz.$$

Remark 5 Observe that F is monotone and continuous, but need not to be strictly monotone (see the distinction between Proposition 1 and Proposition 2 given in Sect. 1).

Remark 6 If both x_i and y_j are i.i.d. with the common c.d.f. F_0 , then Proposition 11 supplies for $F_0 = F_1 = F_2$ the c.d.f.

$$F(d) = m \int_{-\infty}^{+\infty} F_0(z+d)^l (1-F_0(z))^{m-1} f_0(z) dz \quad (11)$$

for d . A similar representation already occurred in Example 4 (see (3)).

Using formula (4) in Example 4, we get

$$F(0) = m \cdot J_{l,m-1} = \frac{l! \cdot m!}{(l+m)!} = \left(\frac{l+m}{m} \right)^{-1} \in]0, 1[$$

as well as

$$F(+\infty) = m \cdot J_{0,m-1} = m \cdot \frac{(m-1)!}{m!} = 1$$

which is in accordance with the properties of a c.d.f. By the way, putting $l := a - t$ and $m = t$ as in the selection problem for a populations with t first class populations we find $F(0) = \left(\frac{a}{t} \right)^{-1}$ independently of the underlying distribution. In Proposition 6, we stated this result for the Bechhofer function that is the c.d.f. of the Bechhofer selection problem with normally distributed statistical populations and mean value scores (BSP).

8 Indifference Zone Selection Problem for Non-normal Populations

We assume that the a statistical populations can be ranked by transformation according to a common distribution with c.d.f. F_0 and the t best are to be selected with risk β of wrong decision. If

$$F(x) := t \int_{-\infty}^{+\infty} F_0(u + v(x))^{a-t} (1 - F_0(u))^{t-1} dF_0(u) \quad (12)$$

for a certain coercive function v with $v(0) = 0$ satisfies $P_{CS} \geq F(n) \geq c := 1 - \beta$; then, the monotone function approach can be used for selection.

If we specialize to normal distributions given in the Bechhofer approach which means F_i are the c.d.f. of $N(\mu_i, \sigma^2)$, we get $F_i(z) = F_0(u) = \Phi(u)$ and $f_i(z) = \frac{1}{\sigma} \cdot \varphi(u)$ with $z = \mu_i + \sigma \cdot u$ using standardization to $N(0, 1)$. Since the mean value scores of sample size n are distributed as $N(\mu_i, \frac{\sigma^2}{n})$, the gap parameter $d = \delta$ is to be replaced by $d^* = \frac{\delta}{\sigma} \sqrt{n} = r \sqrt{n}$. The function (11) will be transformed for $l := a - t$ and $m = t$ to

$$F_B(n) := F(d^*) = F(r \sqrt{n}) = t \int_{-\infty}^{+\infty} \Phi(u + r \sqrt{n})^{a-t} (1 - \Phi(u))^{t-1} \varphi(u) du.$$

Hence, we get BSP and the Bechhofer function (10) replacing n by x . Here, we have $v(n) := r \sqrt{n}$.

This method can be also applied if we have a populations whose characteristics are not normally distributed. We assume generally different expectations $\mu_i = E(x_i)$, but constant known (or estimated) variance $\sigma^2 = V(x_i)$ with $i = 1, \dots, a$.

Example 6 We consider the class of continuous uniform distributions $U(a, b)$ where $a < b$. For the corresponding random variable x , the p.d.f. is

$$f(x) = \frac{1}{b-a} \text{ for } x \in [a, b] \text{ and } f(x) = 0 \text{ else.}$$

This p.d.f. has the compact (i.e. bounded and closed) support $C(f) = [a, b]$.

Further, it is $\mu = E(x) = \frac{1}{2}(a+b)$ and $\sigma^2 = V(x) = \frac{1}{12}(b-a)^2$. We choose a subfamily of $U(a, b)$ with $b = a + \gamma$ where γ is fixed. This subfamily has constant variance $\sigma^2 = \frac{1}{12}\gamma^2$ and can be related to the standard $U(-\sqrt{3}, \sqrt{3})$ with $\mu = 0$ and $\sigma^2 = 1$. We denote the c.d.f. and p.d.f. of this standard by $F_{x,0}$ and $f_{x,0}$, respectively. Then we get according to (11) by the replacements $l := a - t$, $m = t$ and $d = r \sqrt{x}$ the selection function

$$F(x) = t \int_{-\infty}^{+\infty} F_{x,0}(z + r \sqrt{x})^{a-t} (1 - F_{x,0}(z))^{t-1} f_{x,0}(z) dz, \quad x \geq 0.$$

It can be shown that

$$\left(\frac{a}{t}\right)^{-1} = F(0) < F(x) < 1 \text{ for } 0 < x < \frac{12}{r^2} \text{ and } F(x) = 1 \text{ for } x \geq \frac{12}{r^2}.$$

This has the consequence that for sufficiently large n the selection is made without any risk.

9 Preview to Further Applications

The subset selection procedure of Gupta [3] for populations with unknown estimated variance σ^2 leads instead of the Bechhofer integral (10) to a double integral

$$F(x) := t \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi\left(z + r\sqrt{x}y\right)^{a-t} (1 - \Phi(z))^{t-1} \varphi(z) f_m(y) dz dy$$

given by Gupta and Panchapakesan [4] and Gupta and Huang [5] which defines also a strictly monotone continuous function F . Hence, monotone function methods can be applied for solution.

10 Conclusion

The paper indicates that a lot of problems can be solved in the field of stochastic distribution functions by monotone function methods. A promising idea is to look also in Bayesian approaches [2] for further applications.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Bechhofer RE (1954) A single sample multiple decision procedure for ranking means of normal populations with known variances. *Ann Math Stat* 25:16–39
2. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2014) *Bayesian data analysis*, 3rd edn. Chapman and Hall, Boca Raton
3. Gupta SS (1956) On a decision rule for a problem in ranking means. *Mim. Ser. No 150*, Univ. North Carolina
4. Gupta SS, Panchapakesan S (1970) On a class of subset selection procedures. *Mim. Ser. No 225*, Purdue Univ., West Lafayette

5. Gupta SS, Huang DY (1981) Multiple statistical decision theory: recent developments. Springer, New York
6. Rasch D, Schott D (2018) Mathematical statistics. Wiley, New York
7. Rasch D, Verdooren R, Pilz J (2020) Applied statistics, theory and problem solutions with R. Wiley, New York
8. Rasch D, Yanagida T (2019) An optimal two-stage procedure to select the best out of a normal population. *J Stat Theory Pract* 13:3
9. Schott D (2018) How to get in the top ten? An analysis of the Bechhofer selection problem in statistics. *Wismar Frege Ser* 2:7–22
10. Schott D (2017/18) Some remarks on a statistical selection procedure of Bechhofer for expectations. *Rostock Math Kolloq* 71:57–67

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.