MINIMAX APPROACH TO SOME CONSTRAINT STATISTICAL DECISION MODELS

VOICU BOSCAIU

The paper focuses mainly on the possibility of interpreting and solving some constrained statistical decision models as specific cases of an infinite dimensional programming problem, P. The particularity of P is the lack of any topological or vector structure of the parameter set. Program P covers a broad variety of classical models of statistical decision theory, including testing multiple hypotheses and restricted (constraint) classification, etc.

AMS 2010 Subject Classification: 62C20, 62F03.

Key words: minimax theorem, hypotheses testing, multiple decision, constrained statistical decision.

1. INTRODUCTION

The paper focuses mainly on the possibility of interpreting and solving some constrained statistical decision models as specific cases of an infinite dimensional programming problem. The particularity of the approach is the lack of any topological or vector structure of the parameter set.

A variety of testing statistical hypotheses models can be obtained for appropriate choices of parameter sets: most stringent test of Wald (see [14]) minimax tests (see [10]), weighted tests of Krafft (see [8]), constraints classification (see [1]), (a) symmetrical multiple tests (see [9]), constrained classification in non-mutually exclusive and non-exhaustive classes, etc.

Throughout this paper a minimax program, called "Program P", will be considered. Actually, neither the definition of P is the most general, nor the properties of P are exhaustively studied. We limit ourselves to describing some general results, general enough for covering solution-description of a broad variety of statistical decision models, including above mentioned problems. We focus mainly on the existence of optimal solution of Program P and on some necessary conditions for an n-dimensional decision function for being optimal solution. Some sufficient conditions of optimality will be derived too, but this is not our main issue. Reader can find some detailed proofs of sufficiency for below mentioned Examples 1–4 in References of the paper.

MATH. REPORTS **13(63)**, *3* (2011), 235–263

To solve Program P we chose to use a classic minimax result. It is a Kneser-Fan minimax theorem for concave-convex functions (see [4]), in Terkelsen's reformulation ([13]).

The decision of giving here a detailed proof of Program P, is explained by our intention to offer for the reader interested in statistical decision theory a general enough tool in a (relatively short) standalone text.

We briefly outline two general statistical decision contexts which Program P is useful for.

I. Let consider the n-dimensional, $n \geq 2$, decision (action) set $\{A_1, A_2, \ldots, A_n\}$. The decider makes his choice taking into account the values of an observable variable $\mathbf{X}$ on space X. Decision variable $\mathbf{X}$ is a random variable whose probability distribution P^θ is depending on parameter θ. Moreover, the hypothesis is that a reward function is known. Definitely, the reward (or loss) is $h_j^\theta(x) \in \mathbf{R}$ where $x \in X$ is the observed value of $\mathbf{X}$; A_j is the chosen action; θ is the actual value of parameter. The parameter space Θ is known, but the issue at hand is that neither the actual value of parameter $\theta \in \Theta$ is observable nor a probability distribution on Θ is known.

Having in mind that the decider exclusively bases his choices on the observed values $x \in X$, it is useful to replace action set $\{A_1, A_2, \ldots, A_n\}$ by the vector of multiple decision $\phi(x) := (\phi_1(x), \ldots, \phi_n(x))$, where $\phi_j(x)$ is the conditional probability to decide in favor of action (alternative) j, given the value $x \in X$ of random variable $\mathbf{X}$, $j = 1, \ldots, n$.

In the above framework, an usual conservative optimality criterion is maxi-min reward

$$\min_{\theta \in \Theta} \sum_{j=1}^{n} \int_X \phi_j(x) h_j^\theta(x) \mathrm{d}P^\theta(x) = \max_{\phi}.$$

II. Consider the random variable $\mathbf{X}$ which has a distribution P^θ and a probability density ρ^θ, dependent on the unobservable parameter θ. In one of the following known sets parameter θ could be $\Theta_1, \Theta_2, \ldots, \Theta_n$. The n-dimensional hypothesis set $\{H_1, H_2, \ldots, H_n\}$, is of interest, where hypothesis H_j is $\theta \in \Theta_j$. Decider is interested in making a correct guess of the hypothesis H_j for the actual θ, the only information he can use being the observed value $x \in X$ of $\mathbf{X}$. If the decider establishes $\phi := (\phi_1, \ldots, \phi_n)$, $n \geq 2$, as his strategy, probability of decision in favor of H_j is, for given $\theta \in \Theta_k$,

$$P_\theta\ (\text{decision} = H_j) := E_\theta(\phi_j(\mathbf{X})) = \int_X \phi_j(x) \mathrm{d}P^\theta(x) = \int_X \phi_j(x) \rho^\theta(x) \mathrm{d}\mu(x).$$

If $\theta \in \Theta_j$, then $E_\theta(\phi_j(\mathbf{X}))$ is a probability of a correct decision. To the contrary, if $\theta \notin \Theta_j$, then $E_\theta(\phi_j(\mathbf{X}))$ is a probability of an incorrect decision.

It could be of interest to design decision procedures that control some (linear combination of) probabilities of (in)correct decision.

To start with, we list four examples, particular cases of Program P, cases defined in a typical context of theory of testing (multiple) statistical hypotheses.

Example 1 (Fundamental lemma of Neyman and Pearson, [12], [11]). Find critical function $\psi^* \in \Psi,\ \Psi := \{\psi : X \to [0,1]\}$, such that

$$\int_X \psi^*(x) f_{m+1}(x)\mathrm{d}\mu(x) = \max_{\psi\in\Psi}, \quad \int_X \psi^*(x) f_i(x)\mathrm{d}\mu(x) = c_i, \quad i = 1, \dots, m.$$

Example 2 (Maximin test of level α, [10]). Find critical function $\psi^* \in \Psi$ such that if $0 < \alpha < 1$

$$\inf_{\theta\in\Theta_2} E_\theta(\psi^*(\mathbf{X})) = \max_{\psi\in\Psi}, \quad E_\theta(\psi^*(\mathbf{X})) \le \alpha, \quad (\forall)\ \theta \in \Theta_1.$$

Example 3 (Weighted test *PW*, [8]). Find critical function $\psi^* \in \Psi$ such that

$$e_2 \inf_{\theta\in\Theta_2} E_\theta(\psi^*(\mathbf{X})) - e_1 \inf_{\theta\in\Theta_1} E_\theta(\psi^*(\mathbf{X}))) = \max_{\psi\in\Psi}.$$

Example 4 (Asymmetrical problem PA, [9]). Find multiple decision $\phi^* \in G$, for $1 \le k \le n-1$, $n \ge 2$ such that

$$G := \left\{(\phi_1, \dots, \phi_n) \mid \phi_j : X \to [0,1],\ j = 1, n;\ \sum_{j=1}^n \phi_j(x) = 1\right\},$$

$$\min_{j\le k} \inf_{\theta\in\Theta_j} E_\theta(\phi_j^*(\mathbf{X})) = \max_{\phi\in G},$$

$$E_\theta(\phi_{k+1}^*(\mathbf{X})) \ge 1 - b^\theta, \quad (\forall)\ \theta \in \Theta_{k+1}.$$

Section 7 will be devoted to discussing these examples in detail. But, it must be pointed up, Program P can solve other different models, too.

2. DEFINITIONS

– $(X, \mathfrak{X}, \mu)$, a measurable space $(X, \mathfrak{X})$ with the σ-field $\mathfrak{X}$ and a σ-finite measure μ;

– two bilinear forms, $\langle \cdot\,, \cdot \rangle$ and $[\cdot\,, \cdot]$, unidimensional and multiple, respectively defined by:

$\langle \cdot\,, \cdot \rangle : L^1(X, \mu) \times L^\infty(X, \mu) \to \mathbf{R},\ \langle h, \psi \rangle := \int_X \psi(x) h(x)\mathrm{d}\mu(x)$,

$[\cdot\,, \cdot] : [L^1(X, \mu)]^n \times [L^\infty(X, \mu)]^n \to \mathbf{R},\ [f, \phi] := \sum_{j=1}^n \int_X \phi_j(x) f_j(x)\mathrm{d}\mu(x)$;

– two disjoint sets of indices (parameters or labels) Θ_o and Θ_r: *optimality parameter space* Θ_o and *restriction parameter space* Θ_r, such that the σ-fields $(\Theta_o, \mathfrak{L}_o)$ and $(\Theta_r, \mathfrak{L}_r)$ include the singletons and $\Theta := \Theta_o \times \Theta_r$, $\mathfrak{L} := \mathfrak{L}_o \otimes \mathfrak{L}_r$;

– two given sets of functions indexed by Θ_o and Θ_r; if $s \in \{o, r\}$,

$$\mathcal{S}(\Theta_s) := \big\{ f^\theta := (f_1^\theta,\ f_2^\theta, \ldots, f_n^\theta) \mid f_j^\theta : X \to \mathbf{R},\ \theta \in \Theta_s,$$
$$f_j^{(\cdot)}(\cdot)\ \mathfrak{X} \otimes \mathfrak{L}_s\text{-measurable}, f_j^\theta(\cdot) \in L^1(X, \mu);\ j = 1, \ldots, n \big\};$$

– the *set of randomized multiple-decision functions* F

$$F := \Big\{ \phi = (\phi_1, \ldots, \phi_n) \in [L^\infty(X, \mu)]^n \mid \sum_{j=1}^{n} \phi_j(x) = 1, \phi_j(x) \geq 0,\ j = \overline{1, n} \Big\}, \quad n \geq 2; \tag{1}$$

– the *set of feasible decision functions* $F_r,\ F_r \subset F$:

$$F_r := \big\{ \phi = (\phi_1, \ldots, \phi_n) \in F \mid [f^\theta, \phi] \leq 0,\ (\forall)\ \theta \in \Theta_r \big\}; \tag{2}$$

– $M(\Theta_r)$, $M_+(\Theta_r)$ and $M'_+(\Theta_r)$ are the sets of finite measures, finite positive measures and finite positive measures with finite support on σ-field $(\Theta_r, \mathfrak{L}_r)$;

– $P(\Theta_o)$ and $P'(\Theta_o)$ are the sets of probability measures and probability measures with finite support on σ-field $(\Theta_o, \mathfrak{L}_o)$;

– $\Upsilon := M_+(\Theta_r) \times P(\Theta_o)$, $\Upsilon' := M'_+(\Theta_r) \times P'(\Theta_o)$.

3. STATEMENT AND SOLUTION OF PROGRAM P

Program P. Find the value $V(P)$ defined by

$$V(P) := \inf_{\phi \in F_r} \sup_{\theta \in \Theta_o} [f^\theta, \phi] \tag{3}$$

if the following conditions are satisfied:

C1) $F_r \neq \varnothing$;

C2) there is a dominating function $\tau \in L^1(X, \mu)$ such that

$$|f_j^\theta(\cdot)| < \tau(\cdot),\ \mu\text{-a.e.},\ (\forall)\ \theta \in \Theta_s,\ s \in \{o, r\};\ j = 1, \ldots, n.$$

We say that $\phi^* \in F_r$ *is optimal solution* of P if inf is attained in ϕ^*.

In order to solve program P, a Lagrangian function with respect to the multipliers set $M_+(\Theta_r)$ will be defined for the specified function sets $\mathcal{S}(\Theta_o)$ and $\mathcal{S}(\Theta_r)$

$$L : F \times M_+(\Theta_r) \to \mathbf{R}, \quad L(\phi, \lambda) := \sup_{\theta \in \Theta_o} [f^\theta, \phi] + \int_{\Theta_r} [f^\theta, \phi] \mathrm{d}\lambda(\theta).$$

Below, L will be replaced by function $W(\cdot,\cdot) : F \times \Upsilon \to \mathbf{R}$, which is more appropriate for the subsequent application of the minimax theorem

$$W(\phi, (\lambda, \pi)) := \int_{\Theta_o} [f^\theta, \phi] \, \mathrm{d}\pi(\theta) + \int_{\Theta_r} [f^\theta, \phi] \, \mathrm{d}\lambda(\theta). \tag{4}$$

LEMMA 1. *If conditions* C1 *and* C2 *are verified then* $[f^\theta, \phi]$ *and* $V(P)$ *are finite* (*where* $f^\theta \in \mathcal{S}(\Theta_s)$, $s \in \{o, r\}$, $\phi \in F_r$).

Proof. The result is evident

$$|[f^\theta, \phi]| = \left| \sum_{j=1}^{n} \int_X \phi_j(x) f_j^\theta(x) \mathrm{d}\mu(x) \right| \le \sum_{j=1}^{n} \int_X \phi_j(x) \left| f_j^\theta(x) \right| \mathrm{d}\mu(x) <$$

$$< \sum_{j=1}^{n} \int_X \phi_j(x) \tau(x) \mathrm{d}\mu(x) = \int_X \tau(x) \mathrm{d}\mu(x) < \infty. \quad \square$$

THEOREM 2. *Program P has an optimal solution,* $\phi^* \in F_r$. *Moreover, the following relations are verified:*

$$V(P) = \sup_{\tau \in \Upsilon} \inf_{\phi \in F} W(\phi, \tau) = \inf_{\phi \in F} \sup_{\tau \in \Upsilon} W(\phi, \ \tau) = \sup_{\tau \in \Upsilon} W(\phi^*, \ \tau). \tag{5}$$

If $\phi^* \in F$ *verifies* (5) *then* $\phi^* \in F_r$.

Proof. Theorem 2 is a minimax approach of a common infinite dimensional linear programming. A rather comprehensive proof of the theorem will be given in Appendix. $\square$

THEOREM 3. *Let consider Problem P described by relation* (3).

a) $V(P)$ *is finite and the infimum is attained, namely there exists optimal solution* $\phi^* \in F_r$, *such that*

$$V(P) := \inf_{\phi \in F_r} \sup_{\theta \in \Theta_o} [f^\theta, \phi] = \sup_{\theta \in \Theta_o} [f^\theta, \phi^*]. \tag{6}$$

b) *Let define for specified* $(\lambda, \pi) \in \Upsilon$, *the decision criterion* d_j

$$d_j(x, \lambda, \pi) := \int_{\Theta_r} f_j^\theta(x) \mathrm{d}\lambda(\theta) + \int_{\Theta_o} f_j^\theta(x) \mathrm{d}\pi(\theta), \quad 1 \le j \le n. \tag{7}$$

Then, $V(P)$ *verifies the equality*

$$V(P) = \sup_{(\lambda, \pi) \in \Upsilon} \int_X \min_{j \le n} d_j(x, \lambda, \pi) \mathrm{d}\mu(x). \tag{8}$$

c) *Relation* (8) *remains true even if* sup *is taken in the set of finite support measures,* $\Upsilon' := M'_+(\Theta_r) \times P'(\Theta_o) \subset \Upsilon$.

d) *There exists a sequence* $((\lambda_m, \pi_m))_{m\in N}$, $(\lambda_m, \pi_m) \in \Upsilon'$, *which is the solution of the programming problem*

$$\sup_{(\lambda,\pi)\in\Upsilon} \int_X \min_{j\leq n} d_j(x, \lambda, \pi)\mathrm{d}\mu(x) \tag{9}$$

namely

$$\sup_{(\lambda,\pi)\in\Upsilon} \int_X \min_{j\leq n} d_j(x, \lambda, \pi)\mathrm{d}\mu(x) = \lim_m \int_X \min_{j\leq n} d_j(x, \lambda_m, \pi_m)\mathrm{d}\mu(x). \tag{10}$$

e) *If* ϕ^* *is optimal solution of* P *and if* $((\lambda_m, \pi_m))_{m\in N} \subset \Upsilon'$ *is solution of* (10), *then the following relations are verified for* $1 \leq j \leq n$,

$$\phi_j^*(x)\lim_m \Big[d_j(x, \lambda_m, \pi_m) - \min_{k\leq n} d_k(x, \lambda_m, \pi_m)\Big] = 0, \quad \mu\text{-a.e.}; \tag{11}$$

$$\lim_m \lambda_m \left(\left\{\theta \in \Theta_r \mid [f^\theta, \phi^*] \neq 0\right\}\right) = 0; \tag{12}$$

$$\lim_m \pi_m \left(\left\{\theta \in \Theta_0 \mid [f^\theta,\ \phi^*] \neq \sup_{\theta_0\in\Theta_o} [f^{\theta_0}, \phi^*]\right\}\right) = 0. \tag{13}$$

f) *A sufficient condition for feasible solution* $\phi^* \in F_r$ *to be optimal for* P *is the existence of sequence* $((\lambda_m, \pi_m))_{m\in N} \subset \Upsilon$ *such that the set* $\{\int_{\Theta_r} f_j^\theta(\cdot)\,\mathrm{d}\lambda_m \mid m \in N;\ j = 1, n\}$ *is dominated in* $L^1(X, \mu)$ *and conditions* (11), (12), (13) *are verified.*

Proof. The results will be obtained in twelve steps.

S1. Here and further in the proof we shall use Theorem 2, hence there exist $\phi^* \in F_r$ which verifies (5).

Taking into account (5) and (4) we have:

$$V(P) = \sup_{\tau\in\Upsilon} W(\phi^*,\ \tau) = \sup_{(\lambda,\pi)\in\Upsilon} \left(\int_{\Theta_r} [f^\theta, \phi^*]\ \mathrm{d}\lambda(\theta) + \int_{\Theta_o} [f^\theta, \phi^*]\ \mathrm{d}\pi(\theta)\right) \leq$$
$$\leq \sup_{\pi\in P(\Theta_o)} \int_{\Theta_o} [f^\theta, \phi^*]\ \mathrm{d}\pi(\theta) = \sup_{(\lambda=0,\pi)\in\Upsilon} W(\phi^*,\ (0, \pi)).$$

(The inequality appears due to $[f^\theta, \phi^*] \leq 0$, $\forall\theta \in \Theta_r$.)

But, because of sup definition, the above "$\leq$" must be "$=$", hence (taking also into account (5)) $\phi^* \in F_r$ is an optimal solution of Problem P and the second equality of (6) is true.

S2. Applying Fubini Theorem and using definition (7), we obtain an equivalent relation for W

$$W(\phi,\ (\lambda, \pi)) := \int_{\Theta_r} [f^\theta, \phi]\,\mathrm{d}\lambda + \int_{\Theta_o} [f^\theta, \phi]\,\mathrm{d}\pi =$$

$$= \int_X \sum_{j=1}^{n} \phi_j(x) \left[\int_{\Theta_r} f_j^\theta(x)\, \mathrm{d}\lambda(\theta) + \int_{\Theta_o} f_j^\theta(x)\, \mathrm{d}\pi(\theta) \right] \mathrm{d}\mu(x).$$

We proved

$$W(\phi,\, (\lambda,\pi)) = \int_X \sum_{j=1}^{n} \phi_j(x) d_j(x,\lambda,\pi) \mathrm{d}\mu(x). \tag{14}$$

S3. To obtain the minimum of W in F for specified $(\lambda,\pi) \in \Upsilon$, it is necessary and sufficient to define $\phi \in F$ by:

1) $\phi_k(x) = 0$ if $d_k(x,\lambda,\pi) > \min\{d_j(x,\lambda,\pi) \mid j = \overline{1,n}\}$, μ-a.e.;

2) $\sum_{k\in K} \phi_k(x) = 1$ where K is the subset of integers k, $1 \le k \le n$, such that

$$d_k(x,\lambda,\pi) = \min\{d_j(x,\lambda,\pi) \mid j = \overline{1,n}\}.$$

Hence, for specified $(\lambda,\pi) \in \Upsilon$ inf value of $W(\phi,\, (\lambda,\pi))$ in F is attained and we have

$$\inf_{\phi\in F} W(\phi,\, (\lambda,\pi)) = \int_X \min_{j\le n} d_j(x,\lambda,\pi) \mathrm{d}\mu(x). \tag{15}$$

Equality (8) results on account of (15) and Theorem 2.

S4. For statement c), see on WEB the excellent paper of J.B.G. Frenk, P. Kas, G. Kassay (2004), especially inequalities (25) and related explanations about relations between the different minimax results.

S5. Statement d) derives from (7) and c), taking into account the sup definition and the finiteness of $V(P)$.

S6. Because of statement c) and definition of sup in first part of (5), a sequence $(\tau_m)_{m\in N} := ((\lambda_m,\pi_m))_{m\in N} \subset \Upsilon'$ exists, such that $V(P) = \lim_m \big[\inf_{\phi\in F} W(\phi,\, \tau_m)\big]$. Using (15), (14) and ϕ_j^m defined depending on (λ_m,π_m) by 1) and 2) from S3, we have

$$V(P) = \lim_m \Big[\inf_{\phi\in F} W(\phi, (\lambda_m,\pi_m)) \Big] = \lim_m \int_X \min_{j\le n} d_j(x,\lambda_m,\pi_m) \mathrm{d}\mu(x) =$$

$$= \lim_m \int_X \sum_{j=1}^{n} \phi_j^m(x) d_j(x,\lambda_m,\pi_m) \mathrm{d}\mu(x) \le$$

$$\le \liminf_m \int_X \sum_{j=1}^{n} \phi_j^*(x) d_j(x,\lambda_m,\pi_m) \mathrm{d}\mu(x) = \liminf_m W(\phi^*,\, (\lambda_m,\pi_m)).$$

Hence (taking into account the last term from (5)), we proved $\sup W(\phi^*, \tau) = V(P) \le \liminf_m W(\phi^*, (\lambda_m,\pi_m))$. But, obvious, '$\le$' must be '$=$'. Moreover, a subsequence of $((\lambda_m,\pi_m))_{m\in N}$ exists, such that "lim inf" should be equivalent with “lim”. We shall use that subsequence, without changing notation.

S7. Using the last equalities from S6, finiteness of $V(P)$ and $\sum_{j=1}^{n} \phi_j^* = 1$, we have in turn

$$\text{(i)} \qquad \lim_m \int_X \left[\sum_{j=1}^{n} \phi_j^*(x) d_j(x, \lambda_m, \pi_m) - \min_{j \le n} d_j(x, \lambda_m, \pi_m) \right] \mathrm{d}\mu(x) = 0.$$

$$\lim_m \sum_{j=1}^{n} \int_X \left\{ \phi_j^*(x) \Big[d_j(x, \lambda_m, \pi_m) - \min_{k \le n} d_k(x, \lambda_m, \pi_m) \Big] \right\} \mathrm{d}\mu(x) = 0.$$

All integrands in the above sum are non-negative and therefore each integral must be zero. Namely, for $j = 1, n$,

$$\text{(ii)} \qquad \lim_m \int_X \phi_j^*(x) \Big[d_j(x, \lambda_m, \pi_m) - \min_{k \le n} d_k(x, \lambda_m, \pi_m) \Big] \mathrm{d}\mu(x) = 0.$$

As a consequence, there exists a subsequence of $((\lambda_m, \pi_m))_{m \in N}$ (but we did not change notation) such that, for $j = 1, n$,

$$\text{(16)} \qquad \phi_j^*(x) \lim_m \Big[d_j(x, \lambda_m, \pi_m) - \min_{k \le n} d_k(x, \lambda_m, \pi_m) \Big] = 0, \quad \mu\text{-a.e.}$$

(Taking into account that convergence in mean implies convergence in measure, the subsequence can be sequentially selected-first, a subsequence verifying (16) is extracted for $j = 1$, then, from the extracted subsequence, for $j = 2$ and so on.) We obtained (11).

S8. Let consider the sequence $((\lambda_m, \pi_m))_{m \in N}$ from S6. Using the last equalities from S6 and (4) we have

$$V(P) = \lim_m W\big(\phi^*, (\lambda_m, \pi_m)\big) := \lim_m \int_{\Theta_o} \left[f^\theta, \phi^* \right] \mathrm{d}\pi_m(\theta) + \int_{\Theta_r} \left[f^\theta, \phi^* \right] \mathrm{d}\lambda_m(\theta).$$

Let single out a subsequence of $((\lambda_m, \pi_m))_{m \in N}$ (but we don't change the sequence notation) such that the two terms of the above equality are convergent (because the first integral is bounded and $V(P)$ is finite, such subsequence exists).

Owing to $\phi^* \in F_r$, the integrand of second integral is not positive, hence

$$V(P) = \sup_{\tau \in \Upsilon} W(\phi^*, \tau) \le \lim_m \int_{\Theta_o} [f^\theta, \phi^*] \mathrm{d}\pi_m(\theta).$$

The sup definition implies that the above “$\le$” must be “$=$”, hence

$$\text{(iii)} \qquad \lim_m \int_{\Theta_r} [f^\theta, \phi^*] \, \mathrm{d}\lambda_m(\theta) = 0$$

and

$$\text{(iv)} \qquad \lim_m \int_{\Theta_0} \Big(\sup_{\theta_0 \in \Theta_o} [f^{\theta_0}, \phi^*] - [f^\theta, \phi^*] \Big) \mathrm{d}\pi_m(\theta) = 0.$$

S9. Let suppose (12) is false (i.e., no limit or the limit is not zero) and moreover, that there is not any subsequence of $((\lambda_m, \pi_m))_{m \in N}$ which verifies (12). The meaning of "not equal" in (12) is "less" (indeed, because of $\phi^* \in F_r$, we have $(f^\theta, \phi^*) \leq 0$). Then (irrespective of any assumption on the existence of limit in (12)) there exist $\varepsilon > 0$, $L > 0$ and a subsequence of $((\lambda_m, \pi_m))_{m \in N}$ (but we don't change notation) such that if A is defined by $A := \{\theta \in \Theta_r \mid (f^\theta, \phi^*) < -\varepsilon\}$, then $\lim_m \lambda_m(A) = L$. Taking into account (iii) we obtain a contradiction

$$0 = \lim_m \int_{\Theta_r} [f^\theta, \phi^*] \, \mathrm{d}\lambda_m(\theta) \leq \lim_m \int_A [f^\theta, \phi^*] \, \mathrm{d}\lambda_m(\theta) < -\varepsilon \lim_m \int_A \mathrm{d}\lambda_m(\theta) = -\varepsilon L < 0.$$

As a consequence, (12) must be true. Similarly, if we suppose that (13) is false we obtain a contradiction (using (iv) which has non-positive integrand). We completely proved e).

S10. We shall prove that (12) implies relation (iii) and (13) implies (iv).

Let suppose (iii) is not true. Hence $(\exists)$ $\varepsilon > 0$ such that $(\forall)$ $m \in N$, $(\exists)$ $k_m > m$ such that $|I_{k_m}| > \varepsilon$, where we defined $I_k := \int_{\Theta_r} [f^\theta, \phi^*] \, \mathrm{d}\lambda_k(\theta)$. Thus, because $[f^\theta, \phi^*] \leq 0$, there exists a subsequence of $(I_m)_{m \in N}$ (but we don't change notation) such that $I_m < -\varepsilon$, $(\forall)$ $m \in N$.

Moreover (see Lemma 1), $(\exists)$ $M > 0$ such that $-M < [f^\theta, \phi^*] \leq 0$.

If we define $A := \big\{\theta \in \Theta_r \mid [f^\theta, \phi^*] < 0\big\}$, we have

$$-\varepsilon > I_m = \int_{\Theta_r} [f^\theta, \phi^*] \, \mathrm{d}\lambda_m(\theta) = \int_A [f^\theta, \phi^*] \, \mathrm{d}\lambda_m(\theta) > -M \lambda_m(A).$$

We obtained for all $m \in N$, $\lambda_m \left(\big\{\theta \in \Theta_r \mid [f^\theta, \phi^*] < 0\big\}\right) > \varepsilon / M$, hence (12) is not true.

Similarly, one can prove that relation (13) implies (iv).

S11. We shall prove that relation (11) implies relation (i). Because of the dominating hypothesis of f), $\{d_j(\,\cdot\,, \lambda_m, \pi_m) \mid m \in N,\ j = \overline{1, n}\}$ is dominated and therefore relations (16) imply (ii). Adding relations (ii) we obtain (i).

S12. We shall prove f). Feasible solution ϕ^* verifying (11), (12), (13) also verifies (i), (iii), (iv). We define $V := \sup_{\theta \in \Theta_o} [f^\theta, \phi^*]$.

Subtracting (iii) from (iv), changing the integration order, using relations (7), (i) and then (8) we get

$$V = \lim_m \left(\int_{\Theta_r} [f^\theta, \phi^*] \mathrm{d}\lambda_m(\theta) + \int_{\Theta_0} [f^\theta, \phi^*] \mathrm{d}\pi_m(\theta) \right) =$$

$$= \lim_m \sum_{j=1}^n \int_X \phi_j^*(x) \left[\int_{\Theta_r} f_j^\theta(x) \mathrm{d}\lambda_m(\theta) + \int_{\Theta_o} f_j^\theta(x) \mathrm{d}\pi_m(\theta) \right] \mathrm{d}\mu(x) =$$

$$= \lim_m \sum_{j=1}^n \int_X \phi_j^*(x) d_j(x, \lambda_m, \pi_m) \mathrm{d}\mu(x) = \lim_m \int_X \min_{j \le n} d_j(x, \lambda_m, \pi_m)] \mathrm{d}\mu(x) \le$$

$$\le \sup_{(\lambda,\pi)\in\Upsilon} \int_X \min_{j\le n} d_j(x, \lambda, \pi)] \mathrm{d}\mu(x) = V(P).$$

Hence we proved $V \le V(P)$.

Conversely, $V = \sup_{\theta\in\Theta_o} [f^\theta, \phi^*] \ge \inf_{\phi\in F_r} \sup_{\theta\in\Theta_o} [f^\theta, \phi] = V(P)$.

Therefore, ϕ^* is optimal solution of Problem P. □

COROLLARY 4. *A sufficient condition for feasible solution $\phi^* \in F_r$ to be optimal for P is the existence of $\lambda \in M_+(\Theta_r)$ and $\pi \in P(\Theta_o)$ such that the following conditions are verified*

$$\phi_j^*(x) \left[d_j(x, \lambda, \pi) - \min_{k\le n} d_k(x, \lambda, \pi) \right] = 0; \quad \mu\text{-a.e.} \tag{11'}$$

$$\lambda \left(\{ \theta \in \Theta_r \mid [f^\theta, \phi^*] \ne 0 \} \right) = 0, \tag{12'}$$

$$\pi \left(\left\{ \theta \in \Theta_0 \mid [f^\theta, \phi^*] \ne \sup_{\sigma\in\Theta_o} [f^\sigma, \phi^*] \right\} \right) = 0, \tag{13'}$$

where $d_j(x, \lambda, \pi)$, $j = 1, \dots, n$, were defined by (7).

Proof. The statement results directly from Theorem 3 f) under hypothesis that sequence $((\lambda_m, \pi_m))_{m\in N}$ is independent of m, namely $\lambda_m = \lambda$ and $\pi_m = \pi$ for all $m \in N$. □

4. A MORE GENERAL FORM OF PROGRAM P

Below an application of Program P will be considered.

Program PG. Find value $V(PG)$ defined by

$$\begin{aligned} &V(PG) := \inf_{\phi\in G_r} \sup_{\theta\in\Theta_o} ([g^\theta, \phi] - a^\theta), \\ &G_r := \{\phi \in F \mid [g^\theta, \phi] \le a^\theta,\ (\forall)\ \theta \in \Theta_r\}, \end{aligned} \tag{17}$$

if the following conditions are satisfied:

C1) $G_r \neq \varnothing$;

C2) there is a dominating function $\tau \in L^1(X,\mu)$:

$$|g_j^\theta(\cdot)| \leq \tau(\cdot), \quad \mu\text{-a.e.}, \quad (\forall)\ \theta \in \Theta_o \cup \Theta_r;\ j = 1,\dots,n.$$

C3) $a^{(\cdot)}$ is $(\Theta, \mathfrak{L})$-measurable and there exists a positive real number $A \in \mathbf{R}$ such that $|a^\theta| \leq A$, $(\forall)\ \theta \in \Theta_o \cup \Theta_r$.

Theorem 5. a) *To solve Program PG it is sufficient to solve Program P defined for the following choice of* f_j^θ*'s,* $j = 1,\dots,n$,

$$f_j^\theta(x) := g_j^\theta(x) \ - \ a^\theta \rho(x), \quad (\forall)\ x \in X,\ (\forall)\ \theta \in \Theta_o \cup \Theta_r, \tag{18}$$

where $\rho \in L^1(X,\mu)$ *is any probability* μ*-density function* (*optimal solution of PG is not depending on the choice of* ρ).

b) $V(PG)$ *is finite and verify the equality*

$$V(PG) = \sup_{(\lambda,\pi)\in\Upsilon} \Big\{ \int_X \min_{j\leq n} \Big[\int_{\Theta_r} g_j^\theta(x)\ \mathrm{d}\lambda(\theta) + \int_{\Theta_o} g_j^\theta(x) \mathrm{d}\pi(\theta) \Big] \mathrm{d}\mu(x) - \\ - \int_{\Theta_r} a^\theta\ \mathrm{d}\lambda(\theta) - \int_{\Theta_o} a^\theta \mathrm{d}\pi(\theta) \Big\}.$$

c) *There exists* $\phi^* \in G_r$*, such that infimum is attained in* (17). *Moreover, there exists a sequence* $((\lambda_m, \pi_m))_{m\in N} \subset \Upsilon'$ *such that*

$$\lim_m \lambda_m \Big(\big\{ \theta \in \Theta_r \mid [g^\theta, \phi^*] \neq a^\theta \big\} \Big) = 0, \tag{19}$$

$$\lim_m \pi_m \Big(\big\{ \theta \in \Theta_0 \mid [g^\theta, \phi^*] - a^\theta \neq V(P) \big\} \Big) = 0, \tag{20}$$

$$\phi_j^*(x) \lim_m \Big[d_j(x, \lambda_m, \pi_m) - \min_{k\leq n} d_k(x, \lambda_m, \pi_m) \Big] = 0, \quad \mu\text{-a.e.}, \tag{21}$$

where

$$d_j(x, \lambda, \pi) := \int_{\Theta_r} g_j^\theta(x)\, \mathrm{d}\lambda(\theta) + \int_{\Theta_o} g_j^\theta(x)\, \mathrm{d}\pi(\theta), \quad 1 \leq j \leq n. \tag{22}$$

Proof. S1. The following equalities are evident

$$[g^\theta, \phi] - a^\theta = \sum_{j=1}^n \int_X \phi_j(x) g_j^\theta(x)\, \mathrm{d}\mu(x) - a^\theta \int_X \sum_{j=1}^n \phi_j(x)\ \rho(x)\, \mathrm{d}\mu(x) = \\ = \sum_{j=1}^n \int_X \phi_j(x)(g_j^\theta(x) - a^\theta \rho(x))\, \mathrm{d}\mu(x) = [g^\theta - a^\theta \rho, \phi].$$

Because of conditions C2 and C3 of Program PG, there is a dominating function $\tau' \in L^1(X,\mu)$ for all $f_j^\theta(\cdot) := g_j^\theta(\cdot) - a^\theta\rho(\cdot)$. Hence PG is a P-type program and Theorem 3 is applicable.

S2. Using (22), the decision criterion of PG is

$$d_j^{PG}(x,\lambda,\pi) = \int_{\Theta_r} (g_j^\theta(x) - a^\theta\rho(x))\,\mathrm{d}\lambda(\theta) + \int_{\Theta_o} (g_j^\theta(x) - a^\theta\rho(x))\,\mathrm{d}\pi(\theta) =$$
$$= d_j(x,\lambda,\pi) - \rho(x)\Big[\int_{\Theta_r} a^\theta\,\mathrm{d}\lambda(\theta) + \int_{\Theta_o} a^\theta\,\mathrm{d}\pi(\theta)\Big].$$

S3. b) Direct computation, using S2, (8) and (18) implies

$$V(PG) := \sup_{(\lambda,\pi)\in\Upsilon} \int_X \min_{j\le n} d_j^{PG}(x,\lambda,\pi)\mathrm{d}\mu(x) = \sup_{(\lambda,\pi)\in\Upsilon} \Big\{\int_X \min_{j\le n} d_j(x,\lambda,\pi)\mathrm{d}\mu(x) -$$
$$- \int_X \rho(x)\mathrm{d}\mu(x)\Big[\int_{\Theta_r} a^\theta\,\mathrm{d}\lambda(\theta) + \int_{\Theta_o} a^\theta\,\mathrm{d}\pi(\theta)\Big]\Big\} =$$
$$= \sup_{(\lambda,\pi)\in\Upsilon} \Big\{\int_X \min_{j\le n} d_j(x,\lambda,\pi)\,\mathrm{d}\mu(x) - \int_{\Theta_r} a^\theta\,\mathrm{d}\lambda(\theta) - \int_{\Theta_o} a^\theta\mathrm{d}\pi(\theta)\Big\}.$$

S4. c) All statements derive from Theorem 3. Equalities (19) and (20) derive from (12) and (13). Each decision function $d_j^{PG}(\cdot)$ is depending on a term which contains $\rho(\cdot)$ and all a^θ, but this term is not depending on j. Therefore, (11) and (7) can be replaced by (21) and (22). $\square$

5. FINITE-INDEX-SET CASE OF PROGRAM *P*

Hereafter we consider the finite case of Program P. In addition, two type of restrictions (inequalities and equalities) will be allowed.

Program P'. Find value $V(P')$ defined for $F'_r \neq \varnothing$ by

$$V(P') := \inf_{\phi\in F'_r} \sup_{\theta\in\Theta'_o} [f^\theta,\phi], \tag{23}$$

$$F'_r := \big\{\phi = (\phi_1,\dots,\phi_n) \in F \mid [f^\theta,\phi] \le 0, \ (\forall)\ \theta \in \Theta'_{r1};\ [f^\theta,\phi] = 0,\ (\forall)\ \theta \in \Theta'_{r2}\big\}, \tag{24}$$

where $\mathrm{card}(\Theta'_o) = s_o < \infty$; $\mathrm{card}(\Theta'_{r1}) = s_1 < \infty$; $\mathrm{card}(\Theta'_{r2}) = s_2 < \infty$; $\Theta'_{r1} \cap \Theta'_{r2} = \varnothing$; $\Theta'_o \cap (\Theta'_{r1} \cup \Theta'_{r2}) = \varnothing$.

To avoid unnecessary complication, all superfluous restrictions were excluded from F'_r. Namely: if for given θ inequality $[f^\theta,\phi] \le 0$ is true for all $\phi \in F$ then $\theta \notin \Theta'_{r1}$ and if $[f^\theta,\phi] = 0$ is true for all $\phi \in F$ then $\theta \notin \Theta'_{r2}$.

For the sake of simplicity we may give natural numbers as labels for θ's, namely,

$$\Theta'_{r1} := \{1, 2, \dots, s_1\}; \quad \Theta'_{r2} := \{s_1 + 1,\ s_1 + 2, \dots,\ s_1 + s_2\};$$
$$\Theta'_o := \{s_1 + s_2 + 1,\ s_1 + s_2 + 2, \dots,\ s_1 + s_2 + s_o\}.$$

THEOREM 6. *Let consider: Program P', the sets*

$$\Lambda := \{\lambda \in [0,\infty)^{s_1} \times \mathbf{R}^{s_2} \mid \lambda^\theta \geq 0,\ \forall\theta \in \Theta'_{r1};\ \lambda^\theta \in \mathbf{R},\ \forall\theta \in \Theta'_{r2}\},$$
$$\Pi := \left\{\pi \in [0,1]^{s_o} \mid \sum_{\theta\in\Theta'_o} \pi^\theta \ = 1\right\}, \quad \Gamma := \Lambda \times \Pi \subset \mathbf{R}^{s_1+s_2+s_o}$$

and the decision criterion $d'_j : X \times \Lambda \times \Pi \to R,\ j = 1, \dots, n,$

$$d'_j(x,\lambda,\pi) := \sum_{\theta\in\Theta'_{r1}} \lambda^\theta f^\theta_j(x) + \sum_{\theta\in\Theta'_{r2}} \lambda^\theta f^\theta_j(x) + \sum_{\theta\in\Theta'_o} \pi^\theta f^\theta_j(x). \tag{25}$$

a) *There exists $\phi^* \in F'_r$ such that inf value of P' is attained.*

b) *$V(P')$ is finite and verifies the equality*

$$V(P') = \sup_{(\lambda,\pi)\in\Gamma} \int_X \min_{j\leq n} d'_j(x,\lambda,\pi)\mathrm{d}\mu(x). \tag{26}$$

c) *There exists a sequence $((\lambda_m,\pi_m))_{m\in N},\ (\lambda_m,\pi_m) \in \Gamma$ such that $V(P')$ verifies the equality*

(27)
$$V(P') = \lim_m \int_X \min_{j\leq n} \left[\sum_{\theta\in\Theta'_{r1}} \lambda^\theta_m f^\theta_j(x) + \sum_{\theta\in\Theta'_{r2}} \lambda^\theta_m f^\theta_j(x) + \sum_{\theta\in\Theta'_0} \pi^\theta_m f^\theta_j(x)\right] \mathrm{d}\mu(x).$$

Moreover, ϕ^* verifies the following relations:

$$\phi^*_j(x) \neq 0 \Rightarrow \lim_m \Big[d'_j(x,\lambda_m,\pi_m) - \min_{k\leq n} d'_k(x,\lambda_m,\pi_m)\Big] = 0 \quad \mu\text{-a.e.}; \tag{28}$$

$$(\forall)\ \theta \in \Theta'_{r1} : [f^\theta, \phi^*] \neq 0 \Rightarrow \lim_m \lambda^\theta_m = 0; \tag{29}$$

$$(\forall)\ \theta \in \Theta'_o : [f^\theta, \phi^*] \neq V(P') \Rightarrow \lim_m \pi^\theta_m = 0. \tag{30}$$

Proof. For each $\theta \in \Theta'_{r2}$, restriction $[f^\theta, \phi] = 0$ will be replaced by two inequalities: $[f^\theta, \phi] \leq 0$ and $[-f^\theta, \phi] \leq 0$. Consequently, Theorem 3 is applicable for $\Theta_o = \Theta'_o$ and $\Theta_r = \Theta'_{r1} \cup \Theta^+_{r2} \cup \Theta^-_{r2}$, where index sets Θ^+_{r2} and Θ^-_{r2} correspond to restriction sets $\{[f^\theta, \phi] \leq 0 \mid \theta \in \Theta'_{r2}\}$ and $\{[-f^\theta, \phi] \leq 0 \mid \theta \in \Theta'_{r2}\}$, respectively. All integrals on parameter spaces mentioned in Theorem 3 become sums.

Taking into consideration Theorem 3, for each index $\theta \in \Theta'_{r2}$ we have two positive numbers, $\lambda^{\theta+}$ and $\lambda^{\theta-}$. The coefficient $\lambda^{\theta+}$ multiplies $f_j^\theta(x)$ and the coefficient $\lambda^{\theta-}$ multiplies $-f_j^\theta(x)$. Consequently, if $\theta \in \Theta'_{r2}$, $f_j^\theta(\cdot)$ has the coefficient $\lambda^\theta := \lambda^{\theta+} - \lambda^{\theta-}$ in a type-(7) expression, for $j = 1, 2, \ldots, n$.

Thus, the definition of sequence $((\lambda_m, \pi_m))_{m\in N} \subset \mathbf{R}^{s_1+s_2} \times \mathbf{R}^{s_o}$ comes from the proof of Theorem 3 (step S6) and verifies the relation $V(P') = \lim_m \Big[\inf_{\phi\in F} W(\phi, (\lambda_m, \pi_m)) \Big]$.

Relations (25) and (26) derive now directly from (7) and (8), respectively. Relation (27) is a consequence of sup-definition in (26). Relation (28) derives from (11). Relations (29) and (30) derive from (12) and (13), after simple calculus. □

Theorem 7. a) *Let consider the optimal decision function $\phi^* \in F'_r$ of Program P' and the set $T \subset \mathbf{R}^{s_1+s_2}$:*

$$T := \big\{ \big([f^1, \phi], \ldots, [f^{s_1+s_2}, \phi] \big) \mid \phi \in F \big\}.$$

If $(0, 0, \ldots, 0) \in \mathbf{R}^{s_1+s_2}$ is an interior point of T (relative to $\mathbf{R}^{s_1+s_2}$), then there exist $\lambda \in \Lambda$ and $\pi \in \Pi$ such that the following relations are verified:

$$V(P') = \int_X \min_{j\le n} d'_j(x, \lambda, \pi) \mathrm{d}\mu(x), \tag{31}$$

$$d'_j(x, \lambda, \pi) > \min_{k\le n} d'_k(x, \lambda, \pi) \Rightarrow \phi^*_j(x) = 0; \quad \mu\text{-a.e.}, \tag{32}$$

$$(\forall)\ \theta \in \Theta'_{r1}:\ [f^\theta, \phi^*] \neq 0 \Rightarrow \lambda^\theta = 0, \tag{33}$$

$$(\forall)\ \theta \in \Theta'_o:\ [f^\theta, \phi^*] \neq \min_{\sigma\in\Theta'_o} [f^\sigma, \phi^*] \Rightarrow \pi^\theta = 0. \tag{34}$$

b) *A sufficient condition for a feasible solution $\phi^* \in F'_r$ to be optimal for P' is existence of $\lambda \in \Lambda$ and $\pi \in \Pi$ such that the relations* (32)–(34) *are verified.*

Proof. S1. We shall apply Theorem 6. Let $\phi^* \in F'_r$ be the optimal solution of P' and $((\lambda_m, \pi_m))_{m\in N} \subset \mathbf{R}^{s_1+s_2} \times \mathbf{R}^{s_o}$ be the sequence which verifies the relations (27)–(30) of Theorem 6c.

We must mention that below, in order to avoid unnecessary editing complications, by $((\lambda_m, \pi_m))_{m\in N}$ we refer to the above mentioned sequence, or a specified subsequence of this sequence.

Two cases would be possible:

i) sequence $((\lambda_m, \pi_m))_{m\in N}$ has a convergent subsequence;

ii) sequence $((\lambda_m, \pi_m))_{m\in N}$ does not have any convergent subsequence.

S2. Suppose case i) is true, hence $((\lambda_m, \pi_m))_{m\in N}$ is convergent to (or has a convergent subsequence to) $(\lambda, \pi) \in \mathbf{R}^{s_1+s_2} \times \mathbf{R}^{s_o}$. Relations (32), (33) and (34) derive from (28), (29) and (30), respectively. Relation (31) derives from (27) and Lebesgue Theorem, considering that the set $\{d'_j(\cdot\,, \lambda_m, \pi_m) \mid j = 1, 2, \ldots, n;\ m \in N\}$ is bounded in L^1.

S3. We shall show that case ii) is not our case.

Let suppose case ii) is true, hence sequence $((\lambda_m, \pi_m))_{m\in N}$ does not have any subsequence which is convergent on all its components. Then, a subsequence exists – and will be selected – such that all its components have limit (finite or infinite). Definitely, $\lim\limits_m \pi_m^\theta = \pi^\theta \in [0, 1]$ for $\theta \in \Theta'_o$ and there exists a non-empty index subset $L \neq \varnothing$, $L \subset \Theta'_{r1} \cup \Theta'_{r2}$ such that $\lim\limits_m \lambda_m^\theta \in \{-\infty, \infty\}$ for $\theta \in L$; $\lim\limits_m \lambda_m^\theta = \lambda^\theta \notin \{-\infty, \infty\}$ for $\theta \in \Theta'_{r1} \cup \Theta'_{r2} \setminus L$.

S4. An index $\overline{\theta} \in L$ and a subsequence of the sequence from S3 will be selected such that

$$\lambda_m^{\overline{\theta}} \neq 0, \quad \lambda_m^\theta/\lambda_m^{\overline{\theta}} \in [-1, 1]; \quad \forall m \in N,\ \forall \theta \in \Theta'_{r1} \cup \Theta'_{r2}.$$

(Such a subsequence exists and could be sequentially selected because for any two indices θ, $\sigma \in \Theta'_{r1} \cup \Theta'_{r2}$ at least one of the inequalities $|\lambda_m^\theta| \geq |\lambda_m^\sigma|$, $|\lambda_m^\theta| \leq |\lambda_m^\sigma|$ is true for an infinite number of $m \in N$.)

S5. Now, a subsequence of the sequence from S4 will be selected such that the following limits exist

$$\lim_m (\pi_m^\theta/\lambda_m^{\overline{\theta}}) = 0,\ \forall \theta \in \Theta'_o; \quad \lim_m (\lambda_m^\theta/\lambda_m^{\overline{\theta}}) = \eta^\theta,\ \eta^\theta \in [-1, 1],\ \forall \theta \in \Theta'_{r1} \cup \Theta'_{r2}.$$

S6. Let choose j_0 and suppose $\phi^*_{j_0}(x) \neq 0$.

Relation (28) implies $\lim\limits_m \beta_m(x) = 0$, where

$$\beta_m(x) := \sum_{\theta=1}^{s_1+s_2} f_{j_0}^\theta(x)\ \lambda_m^\theta + \sum_{\theta\in\Theta'_o} f_{j_0}^\theta(x) \pi_m^\theta - \min_{j\leq n} \left[\sum_{\theta=1}^{s_1+s_2} f_j^\theta(x) \lambda_m^\theta + \sum_{\theta\in\Theta'_o} f_j^\theta(x) \pi_m^\theta \right].$$

Let consider an index k such that the minimum from the above relation is attained in $j = k$ for an infinite number of $m \in N$ (at least one such index k exists). For this subsequence we have

$$0 = \lim_m \beta_m(x) = \lim_m \left(\sum_{\theta=1}^{s_1+s_2} \left[f_{j_0}^\theta(x) - f_k^\theta(x) \right] \lambda_m^\theta + \sum_{\theta\in\Theta'_o} \left[f_{j_0}^\theta(x) - f_k^\theta(x) \right] \pi_m^\theta \right) =$$

$$= \sum_{\theta=1, \theta\notin L}^{s_1+s_2} \left[f_{j_0}^\theta(x) - f_k^\theta(x) \right] \lambda^\theta + \sum_{\theta\in\Theta'_o} \left[f_{j_0}^\theta(x) - f_k^\theta(x) \right] \pi^\theta +$$

$$+\lim|\lambda_m^{\overline{\theta}}|\left(\sum_{\theta\in L}\left[f_{j_0}^\theta(x)-f_k^\theta(x)\right]\lambda_m^\theta/|\lambda_m^{\overline{\theta}}|\right).$$

Hence (because of S5), we must have

$$\sum_{\theta\in L}\left[f_{j_0}^\theta(x)-f_k^\theta(x)\right]\eta^\theta=0. \tag{35}$$

S7. Taking into account the definition of index k and of subsequence selected in S6, we have for each j, $j=1,2,\ldots,n$; $m\in N$,

$$\sum_{\theta=1}^{s_1+s_2} f_j^\theta(x)\lambda_m^\theta+\sum_{\theta\in\Theta_o'} f_j^\theta(x)\pi_m^\theta\geq\sum_{\theta=1}^{s_1+s_2} f_k^\theta(x)\lambda_m^\theta+\sum_{\theta\in\Theta_o'} f_k^\theta(x)\pi_m^\theta$$

and therefore

$$|\lambda_m^{\overline{\theta}}|\left(\sum_{\theta\in L}\left[f_j^\theta(x)-f_k^\theta(x)\right]\lambda_m^\theta/|\lambda_m^{\overline{\theta}}|\right)\geq$$

$$\geq-\sum_{\theta=1,\,\theta\notin L}^{s_1+s_2}\left[f_j^\theta(x)-f_k^\theta(x)\right]\lambda_m^\theta-\sum_{\theta\in\Theta_o'}\left[f_j^\theta(x)-f_k^\theta(x)\right]\pi_m^\theta.$$

Because all sums from above are convergent and $\lim_m|\lambda_m^{\overline{\theta}}|=\infty$, it must hold

$$\sum_{\theta\in L}\left[f_j^\theta(x)-f_k^\theta(x)\right]\eta^\theta\geq 0.$$

S8. Computing the difference of above inequality and (35) we obtain

$$\sum_{\theta\in L}\left[f_j^\theta(x)-f_{j_0}^\theta(x)\right]\eta^\theta\geq 0,\quad j=1,2,\ldots,n. \tag{36}$$

Relation (36) depends neither on choice of k nor on definition of subsequence from S6. Hence, we obtain a necessary (but not sufficient!) condition. Definitely, for case stated in S3, there exists $L\subset\Theta_{r1}'\cup\Theta_{r2}'$, $L\neq\varnothing$ and a set of constants $\{\eta^\theta\mid\theta\in L\}$ such that

$$\phi_{j_0}^*(x)\neq 0\Rightarrow\sum_{\theta\in L}\eta^\theta f_{j_0}^\theta(x)=\min_{j\leq n}\sum_{\theta\in L}\eta^\theta f_j^\theta(x). \tag{37}$$

(Of course, $\eta^\theta=0$ could appear for some but not for all $\theta\in L$ because we have $\eta^{\overline{\theta}}=1$.)

S9. Using $\{\eta^\theta\mid\theta\in L\}$ we define $Z:F\to\mathbf{R}$,

$$Z(\phi):=\int_X\sum_{j=1}^n\phi_j(x)\sum_{\theta\in L}\eta^\theta f_j^\theta(x)\mathrm{d}\mu(x).$$

Relation (37) implies

$$Z(\phi^*) \leq Z(\phi), \quad \forall \phi \in F. \tag{38}$$

On the other hand (using the definition of the subsequence from S3 and (29)), the implication $\eta^\theta \neq 0 \Rightarrow [g^\theta, \phi^*] = 0$ is true for $\theta \in L$. Hence, we have

$$\eta^\theta \sum_{j=1}^{n} \int_X \phi_j^*(x) f_j^\theta(x) \mathrm{d}\mu(x) = 0, \quad \forall \theta \in L.$$

Adding all above relations we obtain $Z(\phi^*) = 0$. Thus

$$0 \leq Z(\phi), \quad \forall \phi \in F. \tag{39}$$

S10. Because $(0, 0, \ldots, 0) \in \mathbf{R}^{s_1+s_2}$ is an interior point of T, there exists $\phi^0 \in F$ such that, for each $\theta \in L$, if $\eta^\theta \neq 0$ then $[f^\theta, \phi^0] \neq 0$ and, more, $sgn([f^\theta, \phi^0]) = -sgn(\eta^\theta)$. Namely, we can select $\phi^0 \in F$ such that for all $\theta \in L : \eta^\theta [f^\theta, \phi^0] \leq 0$ and at least one inequality is strict (because $\eta^{\overline{\theta}} = 1$). Adding these inequalities for all $\theta \in L$, we obtain $Z(\phi^0) < 0$, but this contradicts (39). The conclusion is that case ii), can't appear.

S11. Proof of sufficiency.

Relations (32), (33) and (34) are equivalent with

$$\phi_j^*(x)\Big(d_j'(x, \lambda, \pi) - \min_{k \leq n} d_k'(x, \lambda, \pi)\Big) = 0; \quad \mu\text{-a.e.}, \tag{40}$$

$$(\forall)\ \theta \in \Theta_{r1}' : [f^\theta, \phi^*]\lambda^\theta = 0, \tag{41}$$

$$(\forall)\ \theta \in \Theta_o' : \big([f^\theta, \phi^*] - V\big)\pi^\theta = 0, \tag{42}$$

where

$$V := \min_{\theta \in \Theta_o'} [f^\theta, \phi^*]. \tag{43}$$

Adding relations (42) for all $\theta \in \Theta_o'$, taking into account that π is a probability measure, adding relations (41) for all $\theta \in \Theta_{r1}'$ and for all $\theta \in \Theta_{r2}'$, considering in turn (25), (40), equality $\sum\limits_{j=1}^{n} \phi_j(x) = 1$ and (26) we obtain

$$\begin{aligned} V &= \sum_{\theta \in \Theta_0'} \pi^\theta [f^\theta, \phi^*] = \sum_{\theta \in \Theta_0'} \pi^\theta [f^\theta, \phi^*] + \sum_{\theta \in \Theta_1'} \lambda^\theta [f^\theta, \phi^*] + \sum_{\theta \in \Theta_2'} \lambda^\theta [f^\theta, \phi^*] = \\ &= \sum_{j=1}^{n} \int_X \phi_j^*(x) d_j'(x, \lambda, \pi) \mathrm{d}\mu(x) = \int_X \min_{j \leq n} d_j'(x, \lambda, \pi) \mathrm{d}\mu(x) \leq \\ &\leq \sup_{(\lambda, \pi) \in \Upsilon'} \int_X \min_{j \leq n} d_j'(x, \lambda, \pi) \mathrm{d}\mu(x) = V(P'). \end{aligned}$$

Finally, $V \leq V(P')$, (43) and (23) imply

$$V(P') \geq V = \sup_{\theta \in \Theta'_o} [f^{\theta}, \phi^*] \geq \inf_{\phi \in F'_r} \sup_{\theta \in \Theta'_o} [f^{\theta}, \phi] = V(P').$$

We obtained $\sup_{\theta \in \Theta'_o} [f^{\ \theta}, \phi^*] = V(P')$. □

Theorem 7 gives necessary and sufficient condition of optimality for Problem P' for interior points of T, only. The case of boundary points of T were exhaustively solved for Fundamental lemma of Neyman and Pearson (in formulation from Example 1) by Dantzig and Wald ([3]).

6. APPLICATIONS: TWO SPECIAL CASES OF OPTIMALITY CONDITION

Two particular cases of the definition of condition of optimality will be presented further on. More general and complex approaches could be considered by explicitly singling out a loss function definition. Here the loss function concept was avoided. But one must keep in mind that the f^{θ}'s and g^{θ}'s of our present approach could depend on a loss function and on certain appropriate probability density functions.

To obtain more explicit optimality conditions, n non-void and non-overlapping sets Θ_{oj}, $1 \leq j \leq n$, one for each decision alternative, have to be singled out, in order to define a partition of optimality parameter space. Also, the corresponding function sets $\{f_j^{\theta} : X \to \mathbf{R} \mid (\forall)\ \theta \in \Theta_{oj}\}$ have to be specified, $1 \leq j \leq n$.

The restriction parameter space Θ_r and the corresponding set F_r of feasible multiple decisions will not be modified.

Program P_1. Find the value $V(P_1)$:

$$V(P_1) := \inf_{\phi \in F_r} \max_{j \leq k} \sup_{\theta \in \Theta_{oj}} \int_X \phi_j(x) g_j^{\theta}(x) \mathrm{d}\mu(x), \tag{44}$$

where $1 \leq k \leq n$; the set of feasible decision functions (defined by (2)) is non-void, $F_r \neq \varnothing$; there exists a dominating function $\tau \in L^1(X, \mu)$ for all functions g_j^{θ}'s from (44) and all functions f_j^{θ}'s from F_r.

Proposition 8. *The Program* P_1 *is a particular case of Program* P. *All statements of Theorem* 3 *hold if decision criterion* d_j*'s defined by* (7) *are* *replaced by*

$$(7') \qquad \begin{aligned} d_j(x,\lambda,\pi) &:= \int_{\Theta_r} f_j^\theta(x)\mathrm{d}\lambda(\theta) + \int_{\Theta_{oj}} g_j^\theta(x)\mathrm{d}\pi(\theta), \quad 1 \le j \le k; \\ d_j(x,\lambda,\pi) &:= \int_{\Theta_r} f_j^\theta(x)\mathrm{d}\lambda(\theta), \quad k+1 \le j \le n. \end{aligned}$$

(*Of course, if* $k = n$ *the second line of* $(7')$ *disappears.*)

Proof. It is enough to define $\Theta_o := \bigcup_{j \le n} \Theta_{oj}$ and to make the following choice for function family $\mathcal{S}(\Theta_o)$ involved in the optimality condition of P:

$$\begin{aligned} f_j^\theta &:= g_j^\theta \quad \text{if } 1 \le j \le k \text{ and } \theta \in \Theta_{oj}, \\ f_j^\theta &:= 0 \quad \text{if } 1 \le j \le k \text{ and } \theta \in \Theta_o \setminus \Theta_{oj}, \\ f_j^\theta &:= 0 \quad \text{if } k+1 \le j \le n \text{ and } \theta \in \Theta_o. \end{aligned}$$

Using above definitions, relation (3) has structure (44). □

Program P_2. Find value $V(P_2)$:

$$(45) \qquad V(P_2) := \inf_{\phi \in F_r} \sum_{j=1}^{k} \sup_{\theta \in \Theta_{oj}} \int_X \phi_j(x) g_j^\theta(x)\mathrm{d}\mu(x),$$

where $1 \le k \le n$; the set of feasible decision functions is non-void, $F_r \neq \varnothing$; there exists a dominating function $\tau \in L^1(X,\mu)$ for all functions g_j^θ's from (44) and all functions f_j^θ's from F_r.

Proposition 9. *Program* P_2 *is a particular case of Program* P.

Proof. It is enough to define Θ_o depending on $\Theta_{o1}, \Theta_{o2}, \ldots, \Theta_{on}$ and to make the following choice for the family of functions involved in the optimality condition of P:

$$(46) \qquad \begin{aligned} &\Theta_o := \Theta_{o1} \times \Theta_{o2} \times \cdots \times \Theta_{on} \\ &f_j^{\theta_1,\theta_2,\ldots,\theta_n} := g_j^{\theta_j}, \quad (\theta_1,\theta_2,\ldots,\theta_n) \in \Theta_o;\ 1 \le j \le n. \end{aligned}$$

With these definitions, the following equality (easy to be verified but, at a first glance, doubtful) holds

$$(47) \qquad \begin{aligned} \sup_{\theta \in \Theta_o} [f^\theta, \phi] &:= \sup_{(\theta_1,\theta_2,\ldots,\theta_n)} \sum_{j=1}^{n} \int_X \phi_j(x) f_j^{\theta_1,\theta_2,\ldots,\theta_n}(x)\mathrm{d}\mu(x) = \\ &= \sum_{j=1}^{n} \sup_{\theta \in \Theta_{oj}} \int_X \phi_j(x) g_j^\theta(x)\mathrm{d}\mu(x). \end{aligned}$$

Above, f_j^θ was defined for all $\theta \in \Theta_o$ and each j, $1 \le j \le n$.

But, if $k < n$ and if we put

$$g_j^{\theta_j} := 0, \quad (\forall)\ \theta_j \in \Theta_{oj},\ k+1 \le j \le n,$$

then the sum from (47) runs from 1 to k, only. □

PROPOSITION 10. *Let consider Program $P_2'U$, the unrestricted finite case of Program P_2 for $k = n$ (i.e., $\Theta_{oj} = \Theta'_{oj}$, $\mathrm{card}(\Theta'_{oj}) < \infty$ for $1 \le j \le n$ and $F_r = F$)*

$$V(P_2'U) := \inf_{\phi\in F} \sum_{j=1}^{n} \max_{\theta\in\Theta'_{oj}} \int_X \phi_j(x) g_j^{\theta}(x) \mathrm{d}\mu(x). \tag{45'}$$

a) *There exists $\phi^* \in F$ such that* inf *value of $P_2'U$ is attained and there exist n probability measures $\pi_1, \pi_2, \ldots, \pi_n$ on $\Theta'_{o1}, \Theta'_{o2}, \ldots, \Theta'_{on}$ such that the following relations are verified*

$$V(P_2'U) = \int_X \min_{j\le n} d'_j(x) \mathrm{d}\mu(x), \tag{48}$$

$$\Big[d'_j(x) - \min_{k\le n} d'_k(x)\Big] \phi_j^*(x) = 0; \quad \mu\text{-a.e.},\ 1 \le j \le n, \tag{49}$$

and $(\forall)\ \theta_j \in \Theta'_{oj}$, $1 \le j \le n$,

$$\left[\max_{\sigma\in\Theta'_{oj}} \int_X \phi_j^*(x) g_j^{\sigma}(x)\mathrm{d}\mu(x) - \int_X \phi_j^*(x) g_j^{\theta_j}(x)\mathrm{d}\mu(x)\right] \pi_j^{\theta_j} = 0, \tag{50}$$

where

$$d_j{}'(x) := \sum_{\theta\in\Theta'_{oj}} \pi_j^{\theta} g_j^{\theta}(x). \tag{51}$$

b) *A sufficient condition for a feasible solution $\phi^* \in F$ to be optimal for $P_2'U$ is the existence of n probability measures $\pi_1, \pi_2, \ldots, \pi_n$ on $\Theta'_{o1}, \Theta'_{o2}, \ldots, \Theta'_{on}$ such that relations* (49)–(51) *are verified.*

Proof. **S1.** Because $F_r = F$ we can apply Theorem 7 for $P_2'U$ written in form (47), using relation (46) for definition of optimality condition. Thus, the optimal solution is depending on a probability π defined on $\Theta'_o := \Theta'_{o1} \times \Theta'_{o2} \times \cdots \times \Theta'_{on}$.

S2. Relation (25) becomes (51). Indeed, for $j = 1$,

$$d'_1(x) = \sum_{\theta\in\Theta'_o} \pi^{\theta_1,\theta_2,\ldots,\theta_n}\ f_1^{\theta_1,\theta_2,\ldots,\theta_n}(x) =$$

$$= \sum_{\theta\in\Theta'_o} \pi^{\theta_1,\theta_2,\ldots,\theta_n}\ g_1^{\theta_1}(x) = \sum_{\theta\in\Theta'_{o1}} \pi_1^{\theta_1}\ g_1^{\theta_1}(x),$$

where we defined

$$\pi_1^{\theta_1} := \sum_{\theta_2,\dots,\theta_n} \pi^{\theta_1,\,\theta_2,\dots,\theta_n}. \tag{52}$$

Because of $\sum_{\theta\in\Theta'_o} \pi^{\theta} = 1$, we have $\sum_{\theta_1\in\Theta'_{o1}} \pi_1^{\theta_1} = 1$, hence π_1 is a probability. Of course, similar results are true for $j = 2, 3, \dots, n$.

Now, (48) derives from (31) and (49) derives from (32).

S3. For $j = 1$, relation (50) results from (34), using (52) and (47).

Indeed, the following implications are true: $\pi_1^{\theta_1} \neq 0 \Rightarrow$ for each $j \le n$, $(\exists)$ $\theta_j \in \Theta'_{oj}$ such that $\pi^{\theta_1,\theta_2,\dots,\theta_n} \neq 0$ (since (52)) $\Rightarrow \sum_{j=1}^{n} \int_X \phi_j(x) g_j^{\theta_j}(x)\mathrm{d}\mu(x) = \sum_{j=1}^{n} \max_{\sigma\in\Theta'_{oj}} \int_X \phi_j(x) g_j^{\sigma}(x)\mathrm{d}\mu(x)$ (since (34) and (47)) $\Rightarrow \int_X \phi_j(x) g_j^{\theta_j}(x)\mathrm{d}\mu(x) = \max_{\sigma\in\Theta'_{oj}} \int_X \phi_j(x) g_j^{\sigma}(x)\mathrm{d}\mu(x)$ for $j = 1, 2, \dots, n$.

Hence we proved

$$\pi_1^{\theta_1} \neq 0 \;\Rightarrow \int_X \phi_1(x) g_1^{\theta_1}(x)\mathrm{d}\mu(x) = \max_{\sigma\in\Theta'_{o1}} \int_X \phi_1(x) g_1^{\sigma}(x)\mathrm{d}\mu(x).$$

The last implication and (50) for $j = 1$ are equivalent. Of course, similar results hold for any $j \le n$.

S4. *Proof of sufficiency.* $\pi^{\theta_1,\theta_2,\dots,\theta_n} \neq 0 \Rightarrow \pi_j^{\theta_j} \neq 0$ for all $j \le n \Rightarrow \int_X \phi_j(x) g_j^{\theta_j}(x)\mathrm{d}\mu(x) = \max_{\sigma\in\Theta'_{oj}} \int_X \phi_j(x) g_j^{\sigma}(x)\mathrm{d}\mu(x)$ for all $j \le n$ (since (50)) $\Rightarrow \sum_{j=1}^{n} \int_X \phi_j(x) g_j^{\theta_j}(x)\mathrm{d}\mu(x) = \sum_{j=1}^{n} \max_{\sigma\in\Theta'_{oj}} \int_X \phi_j(x) g_j^{\sigma}(x)\mathrm{d}\mu(x)$.

Thus we proved that, in our framework, (50) is equivalent with (34). □

7. EXAMPLES

We shall come back to P-type examples mentioned in Section 1.

A slightly different approach for the set of randomized decision functions will be considered for a two-decision case. In the original formulation of Examples 1–3, authors used the set $\Psi := \{\psi : X \to [0,1]\}$ of critical functions. In spite of its redundancy, we prefer to make use of F defined by (1) for $n = 2$, due to symmetry of approach.

The connection between $(\phi_1, \phi_2) \in F$ and $\psi \in \Psi$ is obvious,

$$\phi_1 := 1 - \psi, \quad \phi_2 := \psi.$$

The following program is a generalization of Example 1.

Program PE_1. Find the value $V(PE_1)$ defined for $p \geq 1$ by

$$V(PE_1) := \max_{\phi \in F_r} \min_{j=1,p} \int_X \phi_2(x) f_{m+j}(x) \mathrm{d}\mu(x),$$

$$F_r := \left\{ (\phi_1, \phi_2) \in F \mid \int_X \phi_2(x) f_i(x) \mathrm{d}\mu(x) = c_i,\ i = 1, \ldots, m \right\}.$$

COROLLARY 11. *Let consider Program* PE_1 *and define the set*

$$T := \left\{ \left(\int_X \phi_2(x) f_1(x) \mathrm{d}\mu(x), \ldots, \int_X \phi_2(x) f_m(x) \mathrm{d}\mu(x) \right) \mid \phi \in F \right\} \subset \mathbf{R}^m.$$

a) *If* $\{f_1, f_2, \ldots, f_{m+p}\} \subset L^1(X, \mu)$ *and if* $(c_1, c_2, \ldots, c_m) \in \mathbf{R}^m$ *is an interior point of* T, *then there exist the optimal decision function* $\phi^* = (\phi_1^*, \phi_2^*) \in F_r$ *and two vectors* $\lambda \in \mathbf{R}^m$ *and* $\pi \in [0,1]^p$, $\sum\limits_{j=1}^{p} \pi_j = 1$ *such that the following relations are verified:*

$$(53) \qquad \begin{aligned} \phi_1^*(x) = 1 \quad &\textit{if} \sum_{j=1}^{p} \pi_j f_{m+j}(x) < \sum_{j=1}^{m} \lambda_j f_j(x), \\ \phi_2^*(x) = 1 \quad &\textit{if} \sum_{j=1}^{p} \pi_j f_{m+j}(x) > \sum_{j=1}^{m} \lambda_j f_j(x). \end{aligned}$$

Moreover, the following implication holds for $k = 1, 2, \ldots, p$,

$$(54) \qquad \pi_k \neq 0 \Rightarrow \min_{j=1,p} \int_X \phi_2^*(x) f_{m+j}(x) \mathrm{d}\mu(x) = \int_X \phi_2^*(x) f_{m+k}(x) \mathrm{d}\mu(x).$$

b) *Sufficient for a feasible* $\phi^* \in F_r$ *to be optimal for* PE_1 *is the existence of* $\lambda \in \mathbf{R}^m$ *and* $\pi \in [0,1]^p$, $\sum\limits_{j=1}^{p} \pi_j = 1$ *such that relations* (53) *and* (54) *hold.*

Proof. By changing the sign of optimality condition and the sign of functions $f_{m+j}(\cdot)$ we obtain a PG-type program. Finally, it is enough to apply Theorem 5 and Theorem 7. Note that, in our case, the finite version of decision criterion (22) are

$$d_1' = 0, \quad d_2' = -\sum_{j=1}^{p} \pi_j f_{m+j}(x) + \sum_{j=1}^{m} \lambda_j f_j(x). \quad \square$$

We rewrite Example 2 in terms of probability density functions.

Program PE_2. Maximin test of level α, $0 < \alpha < 1$.
Find

$$V(PE_2) := \max_{\phi \in G} \inf_{\theta \in \Theta_2} \int_X \phi_2(x) g^{\theta}(x) \mathrm{d}\mu(x),$$

$$G := \left\{ \phi \in F \mid \int_X \phi_2(x) g^\theta(x) \mathrm{d}\mu(x) \le \alpha,\ (\forall)\ \theta \in \Theta_1 \right\}.$$

If we change the sign of optimality condition we obtain a PG-type program:

$$-V(PE_2) := \min_{\phi \in G} \sup_{\theta \in \Theta_2} \int_X \phi_2(x)(-g^\theta(x)) \mathrm{d}\mu(x).$$

COROLLARY 12. a) *Program PE_2 is a PG-type program, hence the optimal decision function $\phi^* \in G$ exists and all statements of Theorem* 5 *hold.*

b) *A sufficient condition for $\phi^* \in G$ to be optimal is the existence of finite positive measure λ and of probability π such that following conditions are verified:*

$$\phi_1^*(x) = 1 \quad \textit{if} \int_{\Theta_2} g^\theta(x) \mathrm{d}\pi(\theta) < \int_{\Theta_1} g^\theta(x) \mathrm{d}\lambda(\theta),$$

$$\phi_2^*(x) = 1 \quad \textit{if} \int_{\Theta_2} g^\theta(x) \mathrm{d}\pi(\theta) > \int_{\Theta_1} g^\theta(x) \mathrm{d}\lambda(\theta);$$

$$\lambda\left(\left\{ \theta \in \Theta_1 \mid \int_X \phi_2^*(x) g^\theta(x) \mathrm{d}\mu(x) \ne \alpha \right\} \right) = 0;$$

$$\pi\left(\left\{ \theta \in \Theta_2 \mid \int_X \phi_2^*(x) g^\theta(x) \mathrm{d}\mu(x) \ne \inf_{\theta \in \Theta_2} \int_X \phi_2^*(x) g^\theta(x) \mathrm{d}\mu(x) \right\} \right) = 0.$$

c) *Let PE_2' be the finite case of PE_2, namely $\Theta_1 := \{1, 2, \dots, m\}$ and $\Theta_2 = \{m+1, m+2, \dots, m+p\}$, and define the set $T \in R^m$,*

$$T := \left\{ \left(\int_X \phi_2(x) g^1(x) \mathrm{d}\mu(x), \dots, \int_X \phi_2(x) g^m(x) \mathrm{d}\mu(x) \right) \mid \phi \in F \right\}.$$

If $(\alpha, \alpha, \dots, \alpha) \in \mathbf{R}^m$ is an interior point of T, then necessary and sufficient for optimality of feasible decision function $\phi^* = (\phi_1^*, \phi_2^*) \in G$ is the existence of two vectors $\lambda \in [0, \infty)^m$ and $\pi \in [0, 1]^p$, $\sum_{j=1}^{p} \pi_j = 1$ such that following relations are verified:

$$\phi_1^*(x) = 1 \quad \textit{if} \sum_{\Theta_2} \pi^\theta g^\theta(x) < \sum_{\Theta_1} \lambda^\theta g^\theta(x),$$

$$\phi_2^*(x) = 1 \quad \textit{if} \sum_{\Theta_2} \pi^\theta g^\theta(x) > \sum_{\Theta_1} \lambda^\theta g^\theta(x);$$

$$\int_X \phi_2^*(x) g^\theta(x) \mathrm{d}\mu(x) \ne \alpha \Rightarrow \lambda^\theta = 0,\ (\forall)\theta \in \Theta_1;$$

$$\int_X \phi_2^*(x) g^\theta(x) \mathrm{d}\mu(x) \ne \inf_{\sigma \in \Theta_2} \int_X \phi_2^*(x) g^\theta(x) \mathrm{d}\mu(x) \Rightarrow \pi^\theta = 0,\ (\forall)\theta \in \Theta_2.$$

(*Because the optimal decision function* $\phi^* \in G$ *exists, necessity implies the existence of the two vectors* $\lambda \in [0,\infty)^m$ *and* $\pi \in [0,1]^p$ *mentioned above.*)

Proof. Point a) is evident. Theorem 5 and Theorem 3f) applied for a constant sequence $((\lambda_m, \pi_m))_{m\in N}$ defined by $\lambda_m = \lambda$, $\pi_m = \pi$ implies the statement b). Note that here the decision criterion (22) are

$$d_1 = 0, \quad d_2 = \int_{\Theta_2} g^\theta(x)\mathrm{d}\pi(\theta) - \int_{\Theta_1} g^\theta(x)\mathrm{d}\lambda(\theta).$$

Theorem 7 implies statement c). □

In the framework defined by F, Example 3 can be rewritten in two equivalent forms:

– Find $Q_1 := \max\limits_{\phi\in F}\Big(e_2 \inf\limits_{\theta\in\Theta_2} E_\theta(\phi_2(\mathbf{X})) + e_1 \inf\limits_{\theta\in\Theta_1} E_\theta(\phi_1(\mathbf{X}))\Big)$.

– Find $Q_2 := \min\limits_{\phi\in F}\Big(e_2 \sup\limits_{\theta\in\Theta_2} E_\theta(\phi_1(\mathbf{X})) + e_1 \sup\limits_{\theta\in\Theta_1} E_\theta(\phi_2(\mathbf{X}))\Big)$.

We rewrite variant Q_2 in terms of probability density functions.

Program PE_3. Weighted test problem PW $(n = 2)$.
Find $V(PE_3)$ defined by

$$\min_{\phi\in F}\left(e_2 \sup_{\theta\in\Theta_2} \int_X \phi_1(x)g^\theta(x)\mathrm{d}\mu(x) + e_1 \sup_{\theta\in\Theta_1} \int_X \phi_2(x)g^\theta(x)\mathrm{d}\mu(x)\right).$$

The following corollary is evident.

Corollary 13. *Program* PE_3 *is a particular case of program* P_2. *Hence all statements of Theorem* 3 *hold. If* $\mathrm{card}(\Theta_1) < \infty$ *and* $\mathrm{card}(\Theta_2) < \infty$, *then all statements of Proposition* 10 *hold.* □

We rewrite Example 4 in terms of probability density functions.

Program PE_4. *Asymmetrical problem* PA.

If for all $\theta \in \Theta_j$, $j = 1, 2, \ldots, n$ $(n \geq 2)$ functions $g_j^\theta : X \to \mathbf{R}$ are probability density functions and $\sup\limits_\theta a^\theta < 1$, the following program is defined for $1 \leq k < i \leq n$:

Find

$$V := \max_{\phi\in G} \min_{j\leq k} \inf_{\theta\in\Theta_j} \int_X \phi_j(x)g_j^\theta(x)\mathrm{d}\mu(x),$$

$$G := \left\{\phi \in F \mid \int_X \phi_j(x)g_j^\theta(x)\mathrm{d}\mu(x) \geq a^\theta,\ (\forall)\ \theta \in \Theta_j;\ k+1 \leq j \leq i\right\}.$$

(If $\sup\limits_{\theta\in\Theta_{k+1}} a^\theta \leq 1$ and $i = k+1$, then always $G \neq \varnothing$ but, if $i > k+1$ it is possible to have $G = \varnothing$ even if $\sup\limits_\theta a^\theta < 1$.)

COROLLARY 14. a) *Asymmetrical problem PA can be expressed as a PG program for* $\Theta_o := \bigcup_{j\leq k} \Theta_j$ *and* $\Theta_r := \bigcup_{k<j\leq i} \Theta_j$.

b) *The optimal decision function* $\phi^* \in G$ *of* PA *exists and all statements of Theorem* 5 *hold, provided that* $G \neq \varnothing$.

c) *A sufficient condition for feasible solution* $\phi^* \in G$ *to be optimal for* PA *is the existence of* $\lambda \in M_+(\Theta_r)$ and $\pi \in P(\Theta_o)$ *such that the following conditions are verified:*

$$(11'') \qquad \phi_j^*(x)\Big[d_j(x,\lambda,\pi) - \max_{s\leq n} d_s(x,\lambda,\pi)\Big] = 0, \quad \mu\text{-a.e. } 1 \leq j \leq n,$$

$$(12'') \qquad \lambda\left(\left\{\theta \in \Theta_j \mid \int_X \phi_j^*(x) g_j^\theta(x) \mathrm{d}\mu(x) \neq a^\theta\right\}\right) = 0, \quad k+1 \leq j \leq i,$$

$$(13'') \qquad \pi\left(\left\{\theta \in \Theta_j \mid \int_X \phi_j^*(x) g_j^\theta(x) \mathrm{d}\mu(x) \neq V\right\}\right) = 0, \quad 1 \leq j \leq k,$$

$$(7'') \qquad \begin{aligned} d_j(x,\lambda,\pi) &:= \int_{\Theta_j} g_j^\theta(x)\mathrm{d}\pi(\theta) && \text{for } 1 \leq j \leq k,\\ d_j(x,\lambda,\pi) &:= \int_{\Theta_j} g_j^\theta(x)\mathrm{d}\lambda(\theta) && \text{for } k+1 \leq j \leq i,\\ d_j(x,\lambda,\pi) &:= 0 && \text{for } i+1 \leq j \leq n, \end{aligned}$$

where

$$V := \min_{j\leq k} \inf_{\theta\in\Theta_j} \int_X \phi_j^*(x) g_j^\theta(x) \mathrm{d}\mu(x).$$

Proof. Statements a) and b) are evident if we change the sign of optimality condition, of restrictions and of $g_j^\theta(\cdot)$' s and we apply Proposition 8.

For statement c) it is enough to use Theorem 5a and Corollary 4. □

8. CONCLUDING REMARK

The interest of the paper focused mainly on P, a (finite or infinite dimensional) programming problem general enough to describe the solution of a broad variety of constrained and unconstrained multiple statistical decision models. We have presented a standalone proof of the existence of the optimal decision function and some necessary conditions for optimality. Generally, the optimal solution does not have a simple form (i.e., basically, necessary and sufficient conditions of optimality are not promising for direct application). Indeed, the optimal solution must to be defined depending on a multidimensional sequence of measures on parameter set. This is the price we pay for the lack of any topological or vector structure of the parameter set. The extremum in the compact set of multiple decision functions is attained but the extremum in the set of measures is not necessarily attained in our case.

We obtained simple (i.e., depending on measures, but not on sequences of measures) *necessary and sufficient* conditions of optimality, for finite dimensional problems, only. Some simple sufficient (but not necessary!) conditions for optimality are also operational for finite parameter set but also for infinite parameter set of P-type decision problems. We gave some examples.

If the parameter space (i.e., the function set indexed or dependent on the parameter set) is finite or has a more complex structure (let's say, a given parametric family of probability density functions equipped with an appropriate topology and having some closedness properties), then the multiple decision function could have a simpler form. Examples: [1] and [11] for the finite case and [2] for an infinite dimensional model.

9. APPENDIX: EXISTENCE OF OPTIMAL SOLUTION OF PROGRAM P

THEOREM 2. *Program P defined by* (3) *has an optimal solution, $\phi^* \in F_r$. Moreover, the following relations are verified:*

$$(A1) \qquad V(P) = \sup_{\tau\in\Upsilon}\ \inf_{\phi\in F} W(\phi,\ \tau) = \inf_{\phi\in F}\ \sup_{\tau\in\Upsilon} W(\phi,\ \tau) = \sup_{\tau\in\Upsilon} W(\phi^*,\ \tau).$$

If $\phi^* \in F$ verifies $(A1)$ then $\phi^* \in F_r$.

The result will be obtained taking into account Lemma A1, Lemma A2 and Lemma A3 which follow.

LEMMA A1. *Program P verifies equality*

$$V(P) = \inf_{\phi\in F}\ \sup_{\lambda\in M_+(\Theta_r)} L(\phi,\lambda),$$

where

$$L(\phi,\lambda) := \sup_{\theta\in\Theta_o} [f^\theta,\phi] + \int_{\Theta_r} [f^\theta,\phi]\ \mathrm{d}\lambda(\theta).$$

Proof. Result will be obtained in two steps, S1 and S2.

S1. If $\phi \in F \setminus F_r$, then $\sup_{\lambda\in M_+(\Theta_r)} L(\phi,\lambda) = +\infty$.

Indeed, $\phi \in F$ and $\phi \notin F_r$ imply the existence of $\theta^* \in \Theta_r$ such that $[f^{\theta^*},\phi] > 0$. For given positive $\alpha > 0$, we define the measure $\lambda^* : (\Theta_r, \mathfrak{L}_r) \to \mathbf{R}$ such that $\forall A \in \mathfrak{L}_r$, $\lambda^*(A \setminus \{\theta^*\}) = 0$ and $\lambda^*(\{\theta^*\}) = \alpha$. Then $L(\phi,\lambda^*) := \sup_{\theta\in\Theta_o} [f^\theta,\phi] + [f^{\theta^*},\phi]\alpha$ is not bounded above.

S2. If $\phi \in F_r$, then $\sup\limits_{\lambda \in M_+(\Theta_r)} L(\phi, \lambda) = \sup\limits_{\theta \in \Theta_o} [f^\theta, \phi]$.

Indeed, for all $\theta \in \Theta_r$ we have $[f^\theta, \phi] \le 0$, hence $\sup\limits_{\lambda \in M_+(\Theta_r)} L(\phi, \lambda)$ is attained for $\lambda = 0$.

Obviously, the infimum value of $\sup\limits_{\lambda \in M_+(\Theta_0)} L(\phi, \lambda)$ in the set F does not involve case S1, hence the infimum equals $V(P)$. □

A more general approach to Lagrangian duality (in the framework of multiplier set defined in the topological dual space of a certain normed linear space and the partial ordering induced by some closed convex cone) may be found in [5], but it is beyond our actual interest.

Lemma A2. *Program P verifies equality*

$$V(P) = \inf_{\phi \in F} \sup_{\tau \in \Upsilon} W(\phi, \tau).$$

Proof. The result derives from Lemma A1 and from the equality

$$\sup_{\theta \in \Theta_o} [f^\theta, \phi] = \sup_{\pi \in P(\Theta_o)} \int_{\Theta_o} [f^\theta, \phi] \, d\pi.$$

(Indeed, for a given function $\varphi : (\Theta, \mathfrak{L}) \to \mathbf{R}$, if the σ-algebra $\mathfrak{L}$ includes all singletons and $P(\Theta)$ is the set of probability measures on $(\Theta, \mathfrak{L})$, then $\sup\limits_{\theta \in \Theta} \varphi(\theta) = \sup\limits_{\pi \in P(\Theta)} \int_\Theta \varphi(\theta) d\pi$.) □

Lemma A3.

$$\inf_{\phi \in F} \sup_{\tau \in \Upsilon} W(\phi, \tau) = \sup_{\tau \in \Upsilon} \inf_{\phi \in F} W(\phi, \tau).$$

Moreover, infimum is attained on both sides, hence inf *may be replaced by* min.

Proof. The result will be obtained in five steps, S1–S5, using Kneser-Fan minimax theorem for concave-convex functions.

S1. $[L^\infty(X, \mu)]^n$ will be endowed with the product weak* topology. It is well known that $L^1(X, \mu)^*$-the topological dual of $L^1(X, \mu)$- is $L^\infty(X, \mu)$. (see [7]).

S2. Hence, for given $h \in L^1(X, \mu)$ and the bilinear form $\langle \cdot, \cdot \rangle$, the function $\langle h, \cdot \rangle : L^\infty(X, \mu) \to \mathbf{R}$ is linear and continuous with respect to the weak* topology.

S3. If a couple of measures $(\lambda, \pi) \in \Upsilon = M_+(\Theta_r) \times P(\Theta_o)$ is fixed, it will be proved that $W(\cdot, (\lambda, \pi)) : [L^\infty(X, \mu)]^n \to \mathbf{R}$ is continuous with respect to product weak* topology.

Let consider T_{r1}, the first term of second integral of W (defined by (4)) and use Fubini Theorem:

$$T_{r1}(\phi_1,\lambda) := \int_{\Theta_r}\int_X \phi_1(x) f_1^\theta(x)\mathrm{d}\mu\mathrm{d}\lambda = \int_X \left[\int_{\Theta_r} f_1^\theta(x)\mathrm{d}\lambda\right]\phi_1(x)\mathrm{d}\mu.$$

Considering condition C2), we have

$$\left|\int_{\Theta_r} f_1^\theta(x)\mathrm{d}\lambda\right| \le \int_{\Theta_r} |f_1^\theta(x)|\mathrm{d}\lambda < \int_{\Theta_r}\tau(x)\mathrm{d}\lambda = \tau(x)\,\lambda(\Theta_r).$$

Hence (for given finite λ), we have $\int_{\Theta_r} f_1^\theta(\cdot)\mathrm{d}\lambda \in L^1(X,\mu)$ and therefore the functional $T_{r1}(\cdot\,,\lambda) : L^\infty(X,\mu) \to \mathbf{R}$ is continuous with respect to the weak* topology on $L^\infty(X,\mu)$. Similar statements are true for all terms of W.

S4. F is a compact set with respect to product weak* topology. (Indeed, $A := \{\phi_0 \in L^\infty(X,\mu) \mid 0 \le \phi_0(x) \le 1\}$ is compact in weak* topology on $[L^\infty(X,\mu)]$, A^n is also compact with respect to product weak* topology and F is closed.)

S5. Thus: 1) $W(\cdot\,,\tau)$ is continuous on $[L^\infty(X,\mu)]^n$ with respect to product weak* topology; 2) W is convex on F and concave on $M^+(\Theta_r)\times P(\Theta_o)$; 3) F is a compact set (with respect to product weak* topology). Namely, all conditions of Kneser-Fan Minimax Theorem are fulfilled. Hence minimax equality holds and the infimum is attained on both sides. □

Comment on minimax theorem.

Above, the following classic result ([4]) was used.

Kneser-Fan minimax theorem for concave-convex functions. *Let X be a compact convex subset of a topological vector space and let Y be a convex subset of a vector space. Let $f : X \times Y \to R$ be a concave-convex function on $X \times Y$ (i.e., f is concave on Y for each $x \in X$ and convex on X for each $y \in Y$) and lower semi-continuous on X for every $y \in Y$. Then*

$$\sup_{y\in Y}\min_{x\in X} f(x,y) = \min_{x\in X}\sup_{y\in Y} f(x,y).$$

Actually, there are numerous related minimax theorems. Example: Frenk and Kassay gave sufficient weaker hypotheses for Fan's result ([6]). We preferred to mention Kneser-Fan Theorem because of its direct applicability to our case. Moreover, the interested reader can find Terkelsen's proof of the theorem. It is a simple, complete and stand alone proof, based on the separation theorem of disjoint convex sets in R^n (see [13], Lemma and Corollary 2). □

Proof of Theorem 2. The first two equalities of ($A1$) derive from Lemma A2 and Lemma A3, directly. Moreover, Lemma A3 implies the existence of $\phi^* \in F$ which verifies the last equality of ($A1$).

Let suppose $\phi^* \in F \setminus F_r$. Then $\sup_{\tau \in \Upsilon} W(\phi^*, \tau) = +\infty$ (see S1 in the proof of Lemma A1), but this contradicts Lemma 1. $\square$

REFERENCES

[1] J.A. Anderson, *Discrimination between k populations with constraints on the probabilities of misclassification.* J. R. Statist. Soc. B **31** (1969), 123–39.

[2] V.S. Borkar, S.K. Mitter and S.R. Venkatesh, *Variations on theme by Neyman and Pearson.* Sankhya **66** (2004), Part 2, 292–305.

[3] G.B. Dantzig and A. Wald, *On the fundamental lemma of Neyman and Pearson.* Ann. Math. Statist. **22** (1951), 87–93.

[4] K. Fan, *Minimax Theorems.* Proc. Natl. Acad. Sci. USA **39** (1953), 42–47.

[5] J.B.G. Frenk, P. Kas and G. Kassay, *On linear programming duality and necessary and sufficient conditions in minimax theory.* Econometric Institute Report E.I. 2004–14.

[6] B.G. Frenk and G. Kassay, *On noncooperative games and minimax theory.* Econometric Institute Report E.I. 2005–22.

[7] L.A. Kantorovici and G.P. Akilov, *Functional Analysis.* Ed. Ştiinţifică şi Enciclopedică, Bucureşti, 1986 (Romanian).

[8] O. Krafft, *Eine symmetrische Behandlung des Testproblems.* Arch. Math. (Basel) **18** (1967), 545–560.

[9] O. Krafft and N. Schmitz, *A symmetrical multiple decision problem and linear programming.* In: R. Henn (Ed.), *Operations Research Verfahren VII*, 126–149, Verlag Anton Hain, Meisenheim, 1970.

[10] O. Krafft and H. Witting, *Optimale Test und ungustigste Verteilungen.* Z. Wahrscheinlichkeitstheorie Verw. Geb. **7** (1967), 289–302.

[11] E.L. Lehmann, *Testing Statistical Hypotheses.* Wiley, New York, 1959.

[12] J. Neyman and E.S. Pearson, *Contributions to the theory of testing statistical hypotheses.* Statist. Res. Memoirs **1** (1936), 1–37.

[13] F. Terkelsen, *Some minimax theorems.* Math. Scand **31** (1972), 405–413.

[14] A. Wald, *On the principles of statistical inference.* Notre Dame Mathematical Lectures, **1**, University of Notre Dame, Indiana, 1942.

Received 16 March 2010

"Gh.Mihoc-C.Iacob" Institute of Mathematical Statistics and Applied Mathematics
Calea 13 Septembrie nr. 13
050711 Bucharest, Romania
vboscaiu@gmail.ro