# LeCam's randomization criterion in the setting of locally convex spaces without lattice structure

Heinz Weisshaupt[†]

22.04.2002

**Abstract:** We prove a general functional analytic version of LeCam's randomization criterion. This theorem can be applied to the case of loss function spaces with unbounded loss functions, the case that the loss function spaces depend also on the parameter of the experiment, to filtered experiments and to continuous linear operators on Hilbert spaces.

*Keywords: Comparison of statistical Experiments, LeCam Theory, randomized decisions, locally convex spaces, separation of convex sets, filtered experiments*
*AMS Subject Classification: 62B15, 46N30, 28A10, 46A03, 46A22*

[†]EURANDOM, TU/e Eindhoven. This paper was partially written at the Department of Statistics and Decision Support Systems at Vienna University

# 1 Introduction

In this paper we prove a functional analytic theorem (Theorem 4) which can be considered as a general Version of LeCam's randomization criterion (Theorem 1). The generalization is four-fold:

1. The theorem applies to *more general loss function spaces*, so that the situation of *unbounded loss functions* can be handled. (This makes it possible to *compare statistical experiments with respect to their moments*, or gives us the possibility to apply the randomization criterion to compare statistical experiments *with respect to stochastic orders* of distributions, since such stochastic orders are often generated by function spaces consisting of unbounded functions (see [7], [8] and the example following Theorem 7).

2. The family of *loss functions* may *depend on the parameter*.

3. The Theorem applies to the case of *filtered decision problems* and we obtain from it results analogous to results of Norberg [3].

4. Our general functional analytic approach frees LeCam's Theorem from the (rather special) structure of $L$- and $M$-Spaces. It makes it possible to prove *results* analogous to the randomization criterion *for Operators on Hilbert spaces* and to reverse the role of the family of stochastic operators and the statistical experiment.

The paper is organized as follows:

In section 2 (The classical randomization criterion) we start with some definitions and considerations concerning stochastic operators and we present a version of the classical randomization criterion for statistical experiments (Theorem 1). In this presentation we follow the lines of LeCam [1], LeCam and Yang [2], Strasser [9], Shiryaev and Spokoiny [4] and Torgersen [10]. All these presentations (approaches) are essentially equivalent by the Kakutani-representation theorem for abstract L-spaces (see Torgersen [10] Theorem 5.7.4 and Schaefer [5] Chapter 5 Section 8.5) and by the equivalence of the description of generalized decision rules (see Torgersen [10] 4.5).

Next we present (also in section 2) a theorem (Theorem 2) which is equivalent with the version of the randomization criterion we stated as Theorem 1. We then prove that the Theorems 1 and 2 are equivalent. The Theorem 2 is then our starting point for all further investigations. We remark, that in contrast to the randomization criterion as usually stated (see the cited literature) in the hypotheses of our Theorem 1 the stochastic operators must be induced by Markov kernels, since otherwise the Theorems 1 and 2 would not be equivalent.

This fact can not be considered as a drawback of our approach, since there is no real reason for more general operators in the hypotheses of the randomization criterion; one rather would wish that the stochastic operator provided by the conclusion of the randomization criterion could also be induced by a Markov kernel, but this is as we know in general not possible. (A way out of this problem can in general only be found by proving a theorem which drops $L$-space structure (Theorem 4) and then applying it in special cases where we are given a compact convex set of Markov operators; but compare with [10] 4.5.8. - 4.5.10.)

In section 3 (The setting of locally convex spaces) we introduce first the notions of locally convex spaces, dual spaces and polar sets needed for the statement and proof of the theorems 3 and 4. Then we state Theorem 3 (and the slightly more general Theorem 3') and show that Theorem 3 implies Theorem 2. Theorem 3 is the first step in the abstraction towards the very general Theorem 4; the theorems 3 and 3' are immediate consequences of Theorem 4. With the proof of Theorem 4 we thus also complete the proof of the theorems 1,2 and 3. The proof of Theorem 4 is based on the separation theorem for convex sets (see Schaefer [5] Chapter 2 Section 9.2).

In section 4 (Applications) we introduce the notion of filtered statistical experiment and derive the compactness of the space of filtered stochastic operators from the compactness of the space of stochastic operators. We then state and prove Theorem 5 which is a filtered version of Theorem 2. Further we state the theorems 6 and 7 which are more general versions of filtered randomization criterions. They can be proved analogous to Theorem 5. Theorem 7 is also interesting in the case that the filtration consists only of one $\sigma$-algebra, since it generalizes the randomization criteria also to the case that the family of loss functions may depend on the parameter.

Also in section 4 we state and derive from Theorem 3' the Theorem 8 which is a version of the randomization criterion for Hilbert spaces. Finally we show in Theorem 9 that from an abstract point of view the role of the space of operators and the experiment can be reversed.

In section 5 we state a theorem for the case of finite experiments or more general finite dimensional L-spaces or arbitrary finite dimensional vector spaces. This theorem is an immediate consequence of Helly's Theorem (see Valentine [11]).

In section 6 (Appendix) we state and prove compactness of the space of stochastic operators. (See also [10] 4.5.13. or compare with [9] 42.3)

## 2 The classical randomization criterion

We start with some definitions:

Let $(\Omega, \mathcal{A})$ be a measurable space. We denote by $ca(\Omega, \mathcal{A})$ the space of bounded countably additive real-valued set-functions on $(\Omega, \mathcal{A})$ (i.e. the space of $\sigma$-additive signed measures of bounded variation on $(\Omega, \mathcal{A})$). By $ba(\Omega, \mathcal{A})$ we denote the space of bounded finitely additive real-valued set-functions on $(\Omega, \mathcal{A})$. We denote by $\|.\|$ the variation norm on $ba(\Omega, \mathcal{A})$ defined by:

$$\|\mu\| := 2 \cdot [\sup_{A \in \mathcal{A}} |\mu(A)|] - |\mu(\Omega)| \ .$$

If we like to mention the $\sigma$-algebra $\mathcal{A}$ explicitly we denote the variation norm by $\|.\|_{\mathcal{A}}$.

Let further $(P_\vartheta)_{\vartheta \in \Theta}$ be a family of probability measures on $(\Omega, \mathcal{A})$. Then we call $E := (\Omega, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ a statistical experiment. We denote by $L(E)$ the L-space of the experiment $E$ which is the vector-space of measures defined by

$$L(E) := \{\mu \in ca(\Omega, \mathcal{A}) \mid \ [\nu \in ca(\Omega, \mathcal{A}) \ \text{ and } \ P_\vartheta \perp \nu \text{ for all } \vartheta \in \Theta] \Longrightarrow \mu \perp \nu\} \ .$$

We denote by $ba^+(\Omega, \mathcal{A})$ the positive cone of $ba(\Omega, \mathcal{A})$ defined by

$$ba^+(\Omega, \mathcal{A}) := \{\mu \in ba(\Omega, \mathcal{A}) \mid \mu(A) \geq 0 \ \text{ for all } \ A \in \mathcal{A}\} \ .$$

We further define the positive cones of $ca(\Omega, \mathcal{A})$ and $L(E)$ by

$$ca^+(\Omega, \mathcal{A}) = ca(\Omega, \mathcal{A}) \cap ba^+(\Omega, \mathcal{A}) \ \text{ and } \ L^+(E) := L(E) \cap ba^+(\Omega, \mathcal{A}) \ .$$

Let $E := (\Omega, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ and $F := (\Omega_2, \mathcal{B}, (Q_\vartheta)_{\vartheta \in \Theta})$ be statistical experiments. We say that a linear mapping $T : L(E) \mapsto L(F)$ is a stochastic operator or a transition (see also [9] 55.2) if

$$T(L^+(E)) \subseteq L^+(F) \ \text{ and } \mu \in L(E)^+ \Rightarrow T(\mu)(\Omega_2) = \mu(\Omega).$$

Analogous we say that a linear mapping $T : L(E) \to ba(\Omega_2, \mathcal{B})$ [respectively $T : ba(\Omega, \mathcal{A}) \to ba(\Omega_2, \mathcal{B})$] is a stochastic operator if $T(L^+(E)) \subseteq ba^+(\Omega_2, \mathcal{B})$ and $\mu \in L^+(E) \Rightarrow T(\mu)(\Omega_2) = \mu(\Omega)$ [respectively $T(ba^+(\Omega, \mathcal{A})) \subseteq ba^+(\Omega_2, \mathcal{B})$ and $\mu \in ba^+(\Omega, \mathcal{A}) \Rightarrow T(\mu)(\Omega_2) = \mu(\Omega)$].

Note that any stochastic operator fulfills $\|T\| \leq 1$ with $\|T\| := \sup\{|T(\mu)| \mid \|\mu\| \leq 1\}$. This fact can be obtained using the Jordan decomposition (see Segal and Kunze [6]) of the finitely [respectively countably] additive measure $\mu$ and the properties of stochastic operators.

We say that a stochastic operator $M_K$ is induced by a Markov-kernel

$$K : \Omega \times \mathcal{B} \mapsto \mathrm{I\!R} \quad \text{if} \quad [M_K(\mu)](B) = \int K(x,B) d\mu(x)\ .$$

So far we have introduced the notion of statistical experiment and some abstract spaces related to this notion. Since LeCam-Theory is concerned with the comparison of statistical experiment based on decisions and losses we have to introduce further the notions of decision space, decision, loss function and loss.

If we observe the outcome $x \in \Omega$ of the experiment $E$ we would like to base a decision $d(x)$ in some measurable space $(\Omega_3, \mathcal{D})$ solely on $x$. So a deterministic decision rule is an $\mathcal{A} - \mathcal{D}$-measurable mapping $d : (\Omega, \mathcal{A}) \to (\Omega_3, \mathcal{D})$ and the space $(\Omega_3, \mathcal{D})$ is called a decision space. More general we will consider randomized decision rules. A randomized decision rule makes a decision not by selecting a single point $d(x) \in \Omega_3$ for $x \in \Omega$, but by instead selecting a probability measure $P_x$ on $(\Omega_3, \mathcal{D})$. Formally such a randomized decision rule is given by a Markov-kernel $K : \Omega \times \mathcal{D} \mapsto \mathrm{I\!R}$ by $P_x(D) = K(x, D)$.

Given a $\mathcal{D}$-measurable function $f : (\Omega_3, \mathcal{D}) \to \mathrm{I\!R}$ and a randomized decision rule $M_K$ given by a Markov-kernel $K$, we say that if

$$R_{f,\vartheta}(M_K) := \int_{x\in\Omega} \int_{\Omega_3} f(y)\ K(x, dy) dP_\vartheta(x)$$

exists, then $R_{f,\vartheta}(M_K)$ is the (expected) loss (also called the risk) of the decision rule $M_K$ given the loss function $f$ and the parameter $\vartheta$.

By further abstraction we define the (expected) loss of a general stochastic operator $T$ to be $R_{f,\vartheta}(T) := \int f\ dT(P_\vartheta)$. We note that any stochastic operator can be weakly approximated by (a net of) decision rules induced by Markov kernels (see [10] 4.5.17. or compare with [9] 42.5) and thus a stochastic operator can be viewed as a generalized decision rule. (We will in the next chapter generalize the notion of decision rule further.)

We are now in the position to state the classical randomization criterion of LeCam.

Theorem 1 :*Let*

$$\begin{aligned} E &:= (\Omega, \mathcal{A}, (P_\vartheta)_{\vartheta\in\Theta}) \quad \textit{and} \\ F &:= (\Omega_2, \mathcal{B}, (Q_\vartheta)_{\vartheta\in\Theta}) \end{aligned}$$

*be statistical experiments indexed by the same set* $\Theta$ *and let* $(\varepsilon_\vartheta)_{\vartheta\in\Theta}$ *be an indexed family of reals* $\geq 0$. *Suppose that for an arbitrary measurable space* $(\Omega_3, \mathcal{D})$

*(the decision space), any stochastic operator $M_K : L(F) \mapsto ba(\Omega_3, \mathcal{D})$ which is induced by a Markov-kernel $K$ (the randomized decision rule) and any parameterized family $(f_\vartheta)_{\vartheta\in\Theta}$ of $\mathcal{D}$-measurable functions $f_\vartheta : \Omega_3 \mapsto [-1+1]$ (the loss functions) there exists a stochastic operator $M : L(E) \mapsto ba(\Omega_3, \mathcal{D})$ such that*

$$\int f_\vartheta \, dM(P_\vartheta) \leq \int f_\vartheta \, dM_K(Q_\vartheta) + \varepsilon_\vartheta \quad \textit{for all} \quad \vartheta \in \Theta \, .$$

*Then there exists a stochastic operator $\widetilde{M} : L(E) \mapsto L(F)$ such that*

$$\|Q_\vartheta - \widetilde{M}(P_\vartheta)\| \leq \varepsilon_\vartheta \quad \textit{for all} \quad \vartheta \in \Theta \, .$$

There exists an equivalent formulation of this theorem which does not involve arbitrary decision spaces, but uses instead $(\Omega_2, \mathcal{B})$ itself as a decision space. We formulate this theorem now:

Theorem 2: *Let $E := (\Omega, \mathcal{A}, (P_\vartheta)_{\vartheta\in\Theta})$ and $F := (\Omega_2, \mathcal{B}, (Q_\vartheta)_{\vartheta\in\Theta})$ be statistical experiments indexed by the same set $\Theta$ and let $(\varepsilon_\vartheta)_{\vartheta\in\Theta}$ be a family of nonnegative real numbers. Suppose that for any parameterized family $(g_\vartheta)_{\vartheta\in\Theta}$ of $\mathcal{B}$-measurable functions $g_\vartheta : \Omega_2 \to [-1, +1]$ there exists a stochastic operator $S : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ such that*

$$\int g_\vartheta \, dS(P_\vartheta) \leq \int g_\vartheta \, dQ_\vartheta + \varepsilon_\vartheta \quad \textit{for all} \quad \vartheta \in \Theta \, .$$

*Then there exists a stochastic operator $\widetilde{S} : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ such that*

$$\|\widetilde{S}(P_\vartheta) - Q_\vartheta\| \leq \varepsilon_\vartheta \quad \textit{for all} \quad \vartheta \in \Theta \, .$$

We prove now that the theorems 1 and 2 imply each other. We do this in both directions by the following general principle: If one wants to show that a theorem $A$ implies a theorem $B$ one simply shows that the hypotheses of $B$ imply the hypotheses of $A$ and that the conclusions of $A$ imply the conclusions of $B$.

We show [Theorem 2 $\Rightarrow$ Theorem 1] first:

Let $(\Omega_3, \mathcal{D}) := (\Omega_2, \mathcal{B})$, and let $M_K$ be the identical imbedding of $L(E)$ into $ba(\Omega_2, \mathcal{B}_2)$. In this special case the hypotheses of Theorem 1 say that:

$$\left\{ \begin{array}{l} \text{For any parametrized family } (f_\vartheta)_{\vartheta\in\Theta} \text{ of } \mathcal{B}\text{-measurable functions} \\ f_\vartheta : \Omega_2 \mapsto [-1+1] \text{ there exists a stochastic operator } M \text{ such that} \\ \\ \displaystyle\int f_\vartheta \, dM(P_\vartheta) \leq \int f_\vartheta \, dQ_\vartheta + \varepsilon_\vartheta \quad \text{for all} \quad \vartheta \in \Theta \, . \end{array} \right.$$

By a change of notation (i.e. $g_\vartheta = f_\vartheta$ and $S = M$) we see that these are exactly the hypotheses of Theorem 2. Thus by assumption of the truth of Theorem 2 we get that there exists a stochastic operator $\widetilde{S} : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ such that

$$\|\widetilde{S}(P_\vartheta) - Q_\vartheta\| \leq \varepsilon_\vartheta \quad \text{for all} \quad \vartheta \in \Theta \; .$$

By [9] 41.7 or [10] 4.5.11. there exists a stochastic operator $T : ba(\Omega_2, \mathcal{B}) \mapsto L(F)$ such that $T \mid_{L(F)} = id_{L(F)}$. Since $T$ (being a stochastic operator) fulfills $\|T\| \leq 1$ and the concatenation of stochastic operators is again a stochastic operator we obtain a stochastic operator $\widetilde{M} : L(E) \mapsto L(F)$ given by $\widetilde{M} = T \circ \widetilde{S}$ such that

$$\|\widetilde{M}(P_\vartheta) - Q_\vartheta\| = \|[\widetilde{T} \circ \widetilde{S}](P_\vartheta) - \widetilde{T}(Q_\vartheta)\| \leq \|T\| \cdot \|\widetilde{S}(P_\vartheta) - Q_\vartheta\| \leq \varepsilon_\vartheta$$

for all $\vartheta \in \Theta$. $\square$

## We prove now [Theorem 1 $\Rightarrow$ Theorem 2].

Let a measurable space $(\Omega_3, \mathcal{D})$, a stochastic operator $M_K : L(F) \mapsto (\Omega_3, \mathcal{D})$ and a parameterized family $(f_\vartheta)_{\vartheta\in\Theta}$ of $\mathcal{D}$-measurable functions $f_\vartheta : \Omega_3 \mapsto [-1, 1]$ be given. Suppose that $M_K$ is induced by a Markov kernel $K : \Omega_2 \times \mathcal{D} \mapsto [0, 1]$. Define functions $g_\vartheta$ for all $\vartheta \in \Theta$ by

$$g_\vartheta(x) := \int f_\vartheta(y) \; K(x, dy)$$

The $g_\vartheta$ are $\mathcal{B}$-measurable functions on $\Omega_2$ with ranges contained in $[-1, +1]$. Thus by the hypotheses of Theorem 2 there exists for our parameterized family $(g_\vartheta)_{\vartheta\in\Theta}$ a stochastic operator $S : L(E) \mapsto (\Omega_2, \mathcal{B})$ such that

$$\int g_\vartheta \; dS(P_\vartheta) \leq \int g_\vartheta \; dQ_\vartheta + \varepsilon_\vartheta \;\; \text{for all} \;\; \vartheta \in \Theta$$

If we let $M := M_K \circ S$ then we obtain for all $\vartheta \in \Theta$

$$\int_{\Omega_3} f_\vartheta \; dM(P_\vartheta) = \int_{\Omega_3} f_\vartheta \; d[M_K \circ S](P_\vartheta) =$$
$$\int_{x\in\Omega_2} \int_{y\in\Omega_3} f_\vartheta(y) \; K(x, dy) \; [dS(P_\vartheta)](x) = \int_{\Omega_2} g_\vartheta \; dS(P_\vartheta) \leq$$
$$\int_{\Omega_2} g_\vartheta \; dQ + \varepsilon_\vartheta = \int_{x\in\Omega_2} \int_{y\in\Omega_3} f_\vartheta(y) \; K(x, dy) \; dQ_\vartheta(x) =$$
$$\int_{\Omega_3} f_\vartheta \; dM_K(Q_\vartheta) + \varepsilon_\vartheta \; .$$

By the arbitrary choice of $(\Omega_3, \mathcal{D})$, $(f_\vartheta)_{\vartheta\in\Theta}$ and $M_K$ we see that the hypotheses of Theorem 1 are fulfilled. Since we assumed Theorem 1 to be true, the conclusion of Theorem 1 also holds. That the conclusion of Theorem 1 implies the conclusion of Theorem 2 is trivial and thus the implication has been shown. $\square$

## 3 The setting of locally convex spaces

In this section we free the randomization criterion from the setting of L-spaces. By this it is possible to consider more general loss function spaces and that the loss function spaces depend on the parameter. (Theorem 4).

To do this it is most comfortable to use the language of the theory of locally convex topological vector spaces. So we will state and prove in this section rather abstract theorems, but we gain by this abstraction a highly flexible theory which can be easily applied to several problems in the next section. Since applicability of the theorems 3 and 4 depends only on compactness properties of the space of mappings (abstract randomization rules) involved, it is within this abstract setting very easy to provide theorems (Theorems 5 to 9) which make our Theory applicable to the case of filtered stochastic operators, more or less general loss function spaces and operators on Hilbert spaces.

Theorem 3 introduces the reader to the general form the randomization criterion takes on in the setting of locally convex spaces. Then it is shown that theorem 2 can be derived from Theorem 3. This shows, that compactness of the space of abstract randomization rules is the property in the heart of the randomization criterion and in how far the order structure of the L-spaces is really involved.

Next we state and prove Theorem 4. Theorem 4 is the most general version of the randomization criterion provided in this paper. It provides the possibility to deal with parameterized families of loss functions. It is immediate that Theorem 3 is a special case of Theorem 4. So altogether we provided a proof of a version (Theorem 1) of the classical randomization criterion of LeCam.

We introduce some notations and definitions first (See also [5]):

Given a topological space $(X, \tau)$ and a subset $Y$ of $X$ we denote by $(Y, \tau)$ the topological space $Y$ endowed with the relative topology $Y$ inherits from $\tau$.

The term vector space always denotes a vector space over the real field.

Given a vector space $W$ and a vector space topology $\tau$ which possesses a basis consisting of convex sets, then we call $(W, \tau)$ a locally convex space (abbreviated l.c.s.). If we do not want to mention $\tau$ explicitly we will also write $W$ instead of $(W, \tau)$ for the topological vector space $(W, \tau)$. We denote by $W'$ the topological dual of $(W, \tau)$. (i.e. the vector space of all $\tau$-continuous linear functionals on $W$).

We denote by $(W, \sigma)$ the vector space $W$ endowed with the weakest topology

$\sigma$ making all the elements of the topological dual $W'$ continuous. Symmetrically we denote by $(W', \sigma')$ the vector space $W'$ endowed with the weakest topology making all the elements of $W$ to continuous linear functionals on $(W', \sigma')$. We note that $(W, \sigma)$ and $(W', \sigma')$ are l.c. spaces. We further note that the topological dual of $(W', \sigma')$ is the space $W$ (See [5] Chapter 4 Section 1.2). Thus the relation between $(W, \sigma)$ and $(W', \sigma')$ is completely symmetric and we write $\langle w', w \rangle$ for the value the functional $w'(.)$ takes on at $w$, or equivalently the functional $w(.)$ takes on at $w'$. We will in general not mention the vector space topology explicitly and will write $W$ instead of $(W, \tau)$ and $W'$ instead of $(W', \sigma')$. We will speak of the l.c.s $W$ and its weak dual $W'$.

Let $M \subseteq W$ or $N \subseteq W'$, then we let

$$M^\circ := \{w' \in W' \mid \langle w', w \rangle \leq 1 \ \text{ if } \ w \in M\} \quad \text{and}$$

$$N^\circ := \{w \in W \mid \langle w', w \rangle \leq 1 \ \text{ if } \ w' \in N\} \ .$$

We call $M^\circ$ the polar of $M$ and $N^\circ$ the polar of $N$. By $N^{\circ\circ}$ we denote the polar of the polar of N (also called the bipolar of N).

Further we denote by $\prod_{\xi \in \Xi} W_\xi$ the product of the topological vector spaces $W_\xi$ and by $\bigoplus_{\xi \in \Xi} W_\xi$ the direct sum of the vector spaces $W_\xi$. By $card(\Xi)$ we denote the cardinality of a set $\Xi$.

Theorem 3: *Let $V$ be a set and let $W$ be a locally convex space. Let $G \subseteq W$ be given such that $G$ is closed, convex and contains $0$. Let $\mathcal{L}$ be a compact convex subset of $(W')^V$ endowed with the product topology. Let $\Theta$ be an index set. Let $(v_\vartheta)_{\vartheta \in \Theta}$ be a family of points in $V$ and let $(w_\vartheta)_{\vartheta \in \Theta}$ be a family of points in $W'$. Let $(\varepsilon_\vartheta)_{\vartheta \in \Theta}$ be a family of reals $\varepsilon_\vartheta \in [0, \infty)$. Suppose that for any finite subset $\Theta_0 \subseteq \Theta$ and for any family $(g_\vartheta)_{\vartheta \in \Theta_0}$ of functionals $g_\vartheta \in G$ there exists an $l \in \mathcal{L}$ such that for all $\vartheta \in \Theta_0$ we have:*

$$\langle \ l(v_\vartheta) - w_\vartheta \ , \ g_\vartheta \ \rangle \ \leq \ \varepsilon_\vartheta \ .$$

*Then there exists an $l \in \mathcal{L}$ such that for all $\vartheta \in \Theta$ we have:*

$$\sup_{g_\vartheta \in G} \langle \ l(v_\vartheta) - w_\vartheta \ , \ g_\vartheta \ \rangle \ \leq \ \varepsilon_\vartheta \ .$$

The theorem remains true if we replace the points $w_\vartheta$ by compact convex sets $K_\vartheta$; i.e. we have

Theorem 3': *Let $V$ be a set and let $W$ be a locally convex space. Let $G \subseteq W$ be given such that $G$ is closed, convex and contains $0$. Let $\mathcal{L}$ be a compact convex subset of $(W')^V$ endowed with the product topology. Let $\Theta$ be an index set and* *let $(v_\vartheta)_{\vartheta\in\Theta}$ be a family of points in $V$ and $(K_\vartheta)_{\vartheta\in\Theta}$ be a family of compact convex subsets of $W'$. Let $(\varepsilon_\vartheta)_{\vartheta\in\Theta}$ be a family of reals $\varepsilon_\vartheta \in [0,\infty)$. Suppose that for any finite subset $\Theta_0 \subseteq \Theta$ and for any family $(g_\vartheta)_{\vartheta\in\Theta_0}$ of functionals $g_\vartheta \in G$ there exists an $l \in \mathcal{L}$ such that for all $\vartheta \in \Theta_0$ we have:*

$$\inf_{w_\vartheta \in K_\vartheta} \langle\, l(v_\vartheta) - w_\vartheta \ ,\ g_\vartheta \,\rangle \ \leq\ \varepsilon_\vartheta \ .$$

*Then there exists an $l \in \mathcal{L}$ such that for all $\vartheta \in \Theta$ we have:*

$$\inf_{w_\vartheta \in K_\vartheta} \big( \sup_{g_\vartheta \in G} \langle\, l(v_\vartheta) - w_\vartheta \ ,\ g_\vartheta \,\rangle \big) \ \leq\ \varepsilon_\vartheta \ .$$

We prove now [Theorem 3 $\Rightarrow$ Theorem 2]

For this let $\mathcal{M}_b$ denote the vector space of bounded $\mathcal{B}$-measurable real valued functions on $(\Omega_2, \mathcal{B})$ endowed with (the topology induced by) the supremum norm. Let $\sigma'$ be the topology which $\mathcal{M}_b$ induces on $ba(\Omega_2, \mathcal{B})$ by the mappings $\mu \mapsto \int m\, d\mu$ with $m \in \mathcal{M}_b$. Then $ba(\Omega, \mathcal{B})$ endowed with $\sigma'$ is the weak dual of the l.c.s. $\mathcal{M}_b$. From Lemma A2 of the appendix we obtain that the space of stochastic operators from $L(E)$ to $ba(\Omega_2, \mathcal{B})$ is a compact convex subset of $(ba(\Omega_2, \mathcal{B}), \sigma')^{L(E)}$. The set $\{m \in \mathcal{M}_b \mid -1 \leq m \leq 1\}$ is a closed convex subset of $\mathcal{M}_b$.

Consider now Theorem 3 in the following special case:

Let $V$ denote the space $L(E)$, let $W$ be the space $\mathcal{M}_b$ endowed with the supremum norm. Let $G := \{m \in \mathcal{M}_b \mid -1 \leq m \leq 1\}$, let $v_\vartheta = P_\vartheta$ and let $w_\vartheta = Q_\vartheta$. Denote by $\mathcal{L}$ the space of stochastic operators from $L(E)$ to $ba(\Omega_2, \mathcal{B})$.

With this agreement of notation we see that in the special case we consider the hypotheses of Theorem 2 imply the hypotheses of Theorem 3. But since the conclusion of Theorem 3 is in this special case equivalent with the conclusion of Theorem 2 we see that Theorem 3 in fact implies Theorem 2. □

Theorem 4: *Let $\Xi$ be a set and let $(W_\xi)_{\xi\in\Xi}$ be a family of locally convex vector spaces. Let $(G_\xi)_{\xi\in\Xi}$ be a family of sets with $G_\xi \subseteq W_\xi$ and let $(\varepsilon_\xi)_{\xi\in\Xi}$ be a family of real numbers $\varepsilon_\xi \in [0,\infty)$. Let $\mathcal{K}, \mathcal{J} \subset \prod_{\xi\in\Xi} W'_\xi$.*

*Suppose that the following hypotheses are fulfilled:*

*(i) $\mathcal{K}$ and $\mathcal{J}$ are compact, convex subset of $\prod_{\xi\in\Xi} W_\xi$.*

*(ii) The sets $G_\xi$ are closed, convex and contain $0$.*

*Then the following hypotheses are equivalent:*

*(iii)* *For any finite $\Xi_0 \subseteq \Xi$ and for any selection $(g_\xi)_{\xi\in\Xi_0}$ of functionals $g_\xi \in G_\xi$ there exists $(k,j) \in \mathcal{K} \times \mathcal{J}$ such that for all $\xi \in \Xi_0$ we have*

$$\langle\, j(\xi) - k(\xi)\ ,\ g_\xi \,\rangle \ \le\ \varepsilon_\xi\ .$$

*(iv)* *For any finite $\Xi_0 \subseteq \Xi$, for any selection $(g_\xi)_{\xi\in\Xi_0}$ of functionals $g_\xi \in G_\xi$ there exists $(k,j) \in \mathcal{K} \times \mathcal{J}$ such that for any set $\{\alpha_\xi \mid \xi \in \Xi_0\}$ of real numbers with $\alpha_\xi \ge 0$ and $\sum_{\xi\in\Xi_0} \alpha_\xi \le 1$ we have*

$$\sum_{\xi\in\Xi_0} \alpha_\xi \cdot \langle\, j(\xi) - k(\xi)\ ,\ g_\xi \,\rangle \ \le\ \sum_{\xi\in\Xi_0} \alpha_\xi \cdot \varepsilon_\xi\ .$$

*(v)* *There exists $(k,j) \in \mathcal{K} \times \mathcal{J}$ such that for all $\xi \in \Xi$*

$$\sup_{g_\xi\in G_\xi} (\langle\, j(\xi) - k(\xi)\ ,\ g_\xi \,\rangle) \ \le\ \varepsilon_\xi\ .$$

## Proof of Theorem 4:

We define for all $\xi \in \Xi$ sets $\widetilde{G_\xi}$ by

$$\widetilde{G_\xi} := \frac{1}{\varepsilon_\xi} G_\xi \qquad \text{if} \qquad \varepsilon_\xi > 0$$
$$\widetilde{G_\xi} := \overline{\bigcup_{n\in\mathbb{N}} n\, G_\xi} \qquad \text{if} \qquad \varepsilon_\xi = 0\ .$$

With these definitions the hypothesis $(iii)$, [resp. $(iv)$ or $(v)$] becomes equivalent with the following hypothesis $(iii')$ [resp. $(iv')$ or $(v')$].

$(iii')$ For any finite $\Xi_0 \subseteq \Xi$ and for any selection $(\widetilde{g_\xi})_{\xi\in\Xi_0}$ of functionals $\widetilde{g_\xi} \in \widetilde{G_\xi}$ there exists $(k,j) \in \mathcal{K} \times \mathcal{J}$ such that for all $\xi \in \Xi_0$ we have

$$\langle\, j(\xi) - k(\xi)\ ,\ \widetilde{g_\xi} \,\rangle \ \le\ 1\ .$$

$(iv')$ For any finite $\Xi_0 \subseteq \Xi$, for any selection $(\widetilde{g_\xi})_{\xi\in\Xi_0}$ of functionals $\widetilde{g_\xi} \in \widetilde{G_\xi}$ there exists $(k,j) \in \mathcal{K} \times \mathcal{J}$ such that for any set $\{\alpha_\xi \mid \xi \in \Xi_0\}$ of real numbers with $\alpha_\xi \ge 0$ and $\sum_{\xi\in\Xi_0} \alpha_\xi \le 1$ we have

$$\sum_{\xi\in\Xi_0} \alpha_\xi \cdot \langle\, j(\xi) - k(\xi)\ ,\ \widetilde{g_\xi} \,\rangle \ \le\ 1\ .$$

$(v')$ There exists $(k,j) \in \mathcal{K} \times \mathcal{J}$ such that for all $\xi \in \Xi$

$$\sup_{\widetilde{g_\xi}\in\widetilde{G_\xi}} (\langle\, j(\xi) - k(\xi)\ ,\ \widetilde{g_\xi} \,\rangle) \ \le\ 1\ .$$

So the statement of Theorem 4 is equivalent with the statement that under the hypotheses $(i)$ - $(ii)$ the hypotheses $(iii')$ - $(v')$ are equivalent.

It is immediate that $(iii)$ implies $(iv)$ [or equivalently that $(iii')$ implies $(iv')$] and also that $(v)$ implies $(iii)$ [or equivalently that $(v')$ implies $(iii')$]. So to prove the theorem it remains only to show that $(iv)$ implies $(v)$. This is done indirect by proving that the negation of $(v')$ implies the negation of $(iv')$.

The negation of (v') is the following statement:

For any $(k, j) \in \mathcal{K} \times \mathcal{J}$ there exists a $\xi \in \Xi$ such that

$$\sup_{\widetilde{g_\xi} \in \widetilde{G_\xi}} (\langle\, j(\xi) - k(\xi)\ ,\ \widetilde{g_\xi}\, \rangle)\ >\ 1\ . \tag{1}$$

By hypothesis $(i)$ we have that.

$$(\mathcal{J} - \mathcal{K}) \text{ is a compact convex subset of } \prod_{\xi \in \Xi} W'_\xi. \tag{2}$$

Since the sets $\widetilde{G_\xi}^\circ$ are closed convex subsets of $W'_\xi$, the set

$$\prod_{\xi \in \Xi} \widetilde{G_\xi}^\circ \text{ is a closed convex subset of } \prod_{\xi \in \Xi} W'_\xi. \tag{3}$$

With the notations introduced (1) can be reformulated as

$$\prod_{\xi \in \Xi} \widetilde{G_\xi}^\circ\ \cap\ (\mathcal{J} - \mathcal{K}) = \emptyset. \tag{4}$$

From (2) - (4) and the separation theorem for convex sets (see [5] chapter II, 9.2) we conclude that there exists a continuous linear functional $f \neq 0$ on $\prod_{\xi \in \Xi} W'_\xi$ and a constant $\gamma$ such that

$$f(\prod_{\xi \in \Xi} \widetilde{G_\xi}^\circ) < \gamma < f(\mathcal{J} - \mathcal{K}) \text{ and} \tag{5}$$

$$\text{since } 0 \in \prod_{\xi \in \Xi} \widetilde{G_\xi}^\circ \text{ we get in addition that} \gamma > 0. \tag{6}$$

The algebraic dual of $\prod_{\xi \in \Xi} W'_\xi$ is $\bigoplus_{\xi \in \Xi} W_\xi$ (see [5] chapter IV, 4.3). Thus (and since $f \neq 0$) our functional $f$ can be represented in the form

$$f((w_\xi)_{\xi\in\Xi}) = \sum_{\xi\in\Xi_0} f_\xi(w_\xi) \tag{7}$$

for some finite nonempty set $\Xi_0 \subseteq \Xi$ and a family of continuous linear functionals $f_\xi \in W_\xi$ with $f_\xi \neq 0$ for $\xi \in \Xi_0$.

Let $\Xi_0' := \{\xi \in \Xi_0 \mid \sup_{g\in\widetilde{G_\xi}^\circ} f_\xi(g) > 0\}$ and define reals $\alpha_\xi$ by

$$\alpha_\xi := \sup_{g\in\widetilde{G_\xi}^\circ} \frac{f_\xi(g)}{\gamma} \quad \text{for} \quad \xi \in \Xi_0'. \tag{8}$$

From (5) - (8) we obtain that

$$\alpha_\xi > 0 \ \text{ for } \ \xi \in \Xi_0' \ \text{ and that } \ \sum_{\xi\in\Xi_0'} \alpha_\xi < 1. \tag{9}$$

Further, if $\Xi_0' \neq \Xi_0$, we define

$$\alpha_\xi := \frac{1 - \sum_{\xi\in\Xi_0'} \alpha_\xi}{card(\Xi_0 \setminus \Xi_0')} \quad \text{for} \quad \xi \in \Xi_0 \setminus \Xi_0' \, . \tag{10}$$

By (9) and (10) we get that

$$\alpha_\xi > 0 \tag{11}$$

for $\xi \in \Xi_0$ and $\sum_{\xi\in\Xi_0} \alpha_\xi \leq 1$.
We define linear functionals $g_\xi$ by

$$g_\xi(.) := \frac{f_\xi(.)}{\gamma \cdot \alpha_\xi}. \tag{12}$$

We obtain from (8) and (12) that $g_\xi \in \widetilde{G_\xi}^{\circ\circ}$ in the case that $\xi \in \Xi_0'$. From the fact that $\xi \in \Xi_0 \setminus \Xi_0'$ implies that $\sup_{g\in\widetilde{G_\xi}^\circ} g_\xi(g) = \sup_{g\in\widetilde{G_\xi}^\circ} \frac{f_\xi(g)}{\gamma\cdot\alpha_\xi} = 0$ we get that $g_\xi \in \widetilde{G_\xi}^{\circ\circ}$ in the case that $\xi \in \Xi_0 \setminus \Xi_0'$. So together with $(ii)$ and the bipolar theorem (see [5] chapt IV, 1.5) we obtain in any case that

$$g_\xi \in \widetilde{G_\xi}^{\circ\circ} = \widetilde{G_\xi} \ \text{ for } \ \xi \in \Xi_0. \tag{13}$$

From (12) and (5) we obtain that

$$\sum_{\xi\in\Xi_0} \alpha_\xi \cdot g_\xi(\mathcal{J}-\mathcal{K}) = \sum_{\xi\in\Xi_0} \frac{f_\xi(\mathcal{J}-\mathcal{K})}{\gamma} > \frac{\gamma}{\gamma} = 1 \ . \tag{14}$$

By (11), (13) and (14) we thus have found that

$$\left.\begin{array}{l}\text{there exists a finite set } \Xi_0 \subseteq \Xi \text{, a selection } (\widetilde{g_\xi})_{\xi\in\Xi_0} \text{ of functionals} \\ \widetilde{g_\xi} \in \widetilde{G_\xi} \text{ and there exists a set } \{\alpha_\xi \mid \xi \in \Xi_0\} \text{ of real numbers with} \\ \alpha_\xi \geq 0 \text{ and } \sum_{\xi\in\Xi_0} \alpha_\xi \leq 1 \text{ such that for any } (k,j) \in \mathcal{K}\times\mathcal{J} \text{ we} \\ \text{have} \\ \qquad\qquad ( \sum_{\xi\in\Xi_0} \alpha_\xi \cdot \langle\, j(\xi) - k(\xi) \ , \ \widetilde{g_\xi} \,\rangle) \ > \ 1 \ . \end{array}\right\} \tag{15}$$

The statement (15) is the negation of $(iv')$ as well as (1) is the negation of $(v')$. Since (15) was concluded from (1) we see that $(iv')$ implies $(v')$. Thus the theorem has been proved. □

Replacing $\mathcal{K}\times\mathcal{J}$ by $\widetilde{\mathcal{K}}$ our proof of Theorem 4 proves also the following theorem 4'.

Theorem 4': *Theorem 4 remains true if we replace the compact convex set* $\mathcal{K}\times\mathcal{J} \subset \prod_{\xi\in\Xi} W'_\xi \times \prod_{\xi\in\Xi} W'_\xi$ *by an arbitrary compact convex set* $\widetilde{\mathcal{K}} \subset \prod_{\xi\in\Xi} W'_\xi \times \prod_{\xi\in\Xi} W'_\xi$.

## We prove now [Theorem 4 ⇒ Theorem 3']

We consider Theorem 4 in the following special case:

Let $\Theta$, $V$, $W$, $G$, $\mathcal{L}$, $(K_\vartheta)_{\vartheta\in\Theta}$, $(\varepsilon_\vartheta)_{\vartheta\in\Theta}$ and $(v_\vartheta)_{\vartheta\in\Theta}$ denote the same mathematical objects as in Theorem 3'. Suppose without loss of generality that $\Theta$ and $V$ are disjoint. Let $\Xi = \Theta \cup V$. Let $G_\xi = G$ for $\xi \in \Theta$ and $G_\xi = \{0\}$ for $\xi \in V$. Let $\varepsilon_\xi \geq$ be arbitrary for $\xi \in V$ and let $W_\xi = W$ for all $\xi \in \Xi$. Let $\mathcal{K} := \prod_{v\in V}\{0\} \times \prod_{\vartheta\in\Theta} K_\vartheta$ and let

$$\mathcal{J} := \{ j \in (W')^\Xi \mid \exists l \in L \text{ s.t. } v \in V, \vartheta \in \Theta \Rightarrow [j(v) = l(v) \text{ and } j(\vartheta) = l(v_\vartheta)] \}$$

In this special case the hypotheses of Theorem 3' are equivalent with the conjunction of the hypotheses $(i)$,$(ii)$ and $(iii)$ of Theorem 4. Thus by Theorem 4 the hypothesis $(v)$ of Theorem 4 also holds in this special case. But the conjunction of the hypotheses $(i)$,$(ii)$ and $(v)$ of Theorem 4 is in this special case equivalent with the conclusion of Theorem 3'. Thus Theorem 3' has been proved. □

# 4 Applications

## The case of filtered experiments

We define now filtered experiments and filtered decision rules and generalize the classical randomization criterion (in the version given by Theorem 2) to the setting of these filtered objects. This makes it possible to apply the randomization criterion to stochastic processes. (See also the preprint of Norberg [3]).

Let a measurable space $(\Omega, \mathcal{A})$, a family $(P_\vartheta)_{\vartheta\in\Theta}$ of probability measures on $(\Omega, \mathcal{A})$ and a family $\mathcal{A}_t$ of sub-$\sigma$-algebras of $\mathcal{A}$ be given. We call $E := (\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\in T}, (P_\vartheta)_{\vartheta\in\Theta})$ a filtered statistical experiment. (Note that $(\mathcal{A}_t)_{t\in T}$ need not be a filtration! But of course $(\mathcal{A}_t)_{t\in T}$ can be any filtration. Therefore we decided to call also such experiments filtered; for the usual definition of filtered experiments see [4] 1.10 and [3].) The L-space $L(E)$ of the filtered experiment $E$ is defined to coincide with the L-space of $(\Omega, \mathcal{A}, (P_t)_{t\in T})$.

Let two filtered statistical experiments $E := (\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\in T}, (P_\vartheta)_{\vartheta\in\Theta})$ and $F := (\Omega_2, \mathcal{B}, (\mathcal{B}_t)_{t\in T}, (Q_\vartheta)_{\vartheta\in\Theta})$ be given. We say that a stochastic operator $S : L(E) \to ba(\Omega_2, \mathcal{B})$ is a $(\mathcal{A}_t)_{t\in T} - (\mathcal{B}_t)_{t\in T}-$filtered stochastic operator if

$$\mu, \nu \in L(E) \Rightarrow [\mu \mid_{\mathcal{A}_t} = \nu \mid_{\mathcal{A}_t} \Rightarrow S(\mu) \mid_{\mathcal{B}_t} = S(\nu) \mid_{\mathcal{B}_t}] \, .$$

Lemma 1: *Let*

$$E := (\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\in T}, (P_\vartheta)_{\vartheta\in\Theta})$$

*and*

$$F := (\Omega_2, \mathcal{B}, (\mathcal{B}_t)_{t\in T}, (Q_\vartheta)_{\vartheta\in\Theta})$$

*be filtered statistical experiments. Then the space $\mathcal{S}_{filt}$ of all $(\mathcal{A}_t)_{t\in T} - (\mathcal{B}_t)_{t\in T}-$ filtered stochastic operators $S : L(E) \to ba(\Omega_2, \mathcal{B})$ is compact (when $ba(\Omega_2, \mathcal{B})$ is endowed with the topology induced by the integrals $I_m(\mu) = \int m \, d\mu$ with $m \in \mathcal{M}_b$ [m a B-measurable bounded real valued function]).*

Proof: Since the space $\mathcal{S}$ of stochastic operators from $L(E)$ to $ba(\Omega_2, \mathcal{B})$ is compact (see Lemma A2 of the Appendix) it suffices to show that $\mathcal{S}_{filt}$ is a closed subspace of $\mathcal{S}$. But this is clear since

$$\mathcal{S}_{filt} = \bigcap_{t\in T} \{S \in \mathcal{S} \mid [\mu \mid_{\mathcal{A}_t} = \nu \mid_{\mathcal{A}_t}] \Rightarrow [m \in \mathcal{M}_b(\Omega_2, \mathcal{B}_t) \Rightarrow I_m(S(\nu)) = I_m(S(\mu))]\}$$

and

$$\{S \in \mathcal{S} \mid [\mu \mid_{\mathcal{A}_t} = \nu \mid_{\mathcal{A}_t}] \Rightarrow [m \in \mathcal{M}_b(\Omega_2, \mathcal{B}_t) \Rightarrow I_m(S(\nu)) = I_m(S(\mu))]\}$$

is for any $t \in T$ a closed subset of $S$. $\square$

We can now derive the filtered version of the classical randomization criterion from Theorem 4 with nearly no additional effort:

Theorem 5: *Let*

$$E := (\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\in T}, (P_\vartheta)_{\vartheta\in\Theta}) \quad \textit{and}$$
$$F := (\Omega_2, \mathcal{B}, (\mathcal{B}_t)_{t\in T}, (Q_\vartheta)_{\vartheta\in\Theta})$$

*be filtered statistical experiments indexed by the same sets* $\Theta$ *and* $T$ *and let* $(\varepsilon_{\vartheta,t})_{\vartheta\in\Theta,t\in T}$ *be a family of nonnegative real numbers. Suppose that for any finite set* $\Theta_0 \times T_0 \subset \Theta \times T$ *and any selection* $(g_{\vartheta,t})_{\vartheta\in\Theta_0,t\in T_0}$ *of functions* $g_{\vartheta,t} : \Omega_2 \mapsto [-1, +1]$ *such that* $g_{\vartheta,t}$ *is* $\mathcal{B}_t$*-measurable there exists a* $(\mathcal{A}_t)_{t\in T} - (\mathcal{B}_t)_{t\in T}$*−filtered stochastic operator* $S : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ *such that*

$$\int g_{\vartheta,t}\, dS(P_\vartheta) \le \int g_{\vartheta,t}\, dQ_\vartheta + \varepsilon_{\vartheta,t} \quad \textit{for all} \quad \vartheta \in \Theta \quad \textit{and} \quad t \in T\,.$$

*Then there exists a* $(\mathcal{A}_t)_{t\in T}-(\mathcal{B}_t)_{t\in T}$*−filtered stochastic operator* $\widetilde{S} : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ *such that*

$$\|\widetilde{S}(P_\vartheta) - Q_\vartheta\|_{\mathcal{B}_t} = \sup_{\substack{g\in\mathcal{M}_b(\Omega_2,\mathcal{B}_t)\\ -1\le g\le 1}} (\int g\, d[\widetilde{S}(P_\vartheta) - Q_\vartheta]) \le \varepsilon_{\vartheta,t}$$

*for all* $\vartheta \in \Theta$ *and* $t \in T$. *(Here* $\mathcal{M}_b(\Omega_2, \mathcal{B}_t)$ *denotes the vector space of bounded,* $\mathcal{B}_t$*-measurable, real valued functions.)*

Proof of Theorem 5:

In analogy with the argument given for deriving Theorem 2 from Theorem 3, we let $\mathcal{M}_b(\Omega_2, \mathcal{B}_t)$ denote the vector space of bounded $\mathcal{B}_t$-measurable real valued functions on $(\Omega_2, \mathcal{B})$ endowed with the supremum norm. Let $\sigma'_t$ be the topology which $\mathcal{M}_b(\Omega_2, \mathcal{B}_t)$ induces on $ba(\Omega_2, \mathcal{B})$ by the mappings $\mu \mapsto \int m\, d\mu$ with $m \in \mathcal{M}_b$. By Lemma 1 we know that the space of $(\mathcal{A}_t)_{t\in T} - (\mathcal{B}_t)_{t\in T}$−filtered stochastic operators $\mathcal{S}_{filt}$ from $L(E)$ to $ba(\Omega_2, \mathcal{B})$ is a compact convex subset of $(ba(\Omega_2, \mathcal{B}, \sigma')^{L(E)}$. The set $\{m \in \mathcal{M}_b(\Omega_2, \mathcal{B}_t), | -1 \le m \le 1\}$ is a closed convex subset of $\mathcal{M}_b(\Omega_2, \mathcal{B})$.

Consider now Theorem 4 in the following special case:

Let $\Xi := (\Theta \cup L(E)) \times T$. For $\xi = (x, t)$ let $W_\xi := \mathcal{M}_b(\Omega_2, \mathcal{B}_t)$, then we have $W'_\xi = (ba(\Omega_2, \mathcal{B}), \sigma'_t)$. Let

$$G_\xi := \begin{cases} \{m \in \mathcal{M}_b(\Omega_2, \mathcal{B}_t) \mid -1 \le m \le 1\} & \text{if} \quad \xi = (\vartheta, t) \in \Theta \times T \\ \{0\} & \text{if} \quad \xi \in L(E) \times T \;\; \text{and} \end{cases}$$

let $\varepsilon_\xi > 0$ be arbitrary if $\xi \in L(E) \times T$.

Let $K = \{k\}$ with $k : \Xi \mapsto ba(\Omega_2, \mathcal{B})$ defined by

$$k(\xi) := \begin{cases} Q_\vartheta & \text{if} \quad \xi = (\vartheta, t) \in \Theta \times T \\ \\ 0 & \text{if} \quad \xi \in L(E) \times T. \end{cases}$$

Let $\mathcal{J} := \mathcal{S}_{filt} \circ h$ with $h : \Xi \to L(E)$ given by

$$h(\xi) := \begin{cases} P_\vartheta & \text{if} \quad \xi = (\vartheta, t) \in \Theta \times T \\ \\ \mu & \text{if} \quad \xi = (\mu, t) \in L(E) \times T. \end{cases}$$

In this special case the hypotheses of Theorem 5 are equivalent with the conjunction of the hypotheses $(i)$,$(ii)$ and $(iii)$ of Theorem 4. Thus by Theorem 4 hypothesis $(v)$ of Theorem 4 also holds in the special case. But the conjunction of the hypotheses $(i)$,$(ii)$ and $(v)$ of Theorem 4 is in this special case equivalent with the conclusion of Theorem 5. Thus Theorem 5 has been proved. □

More general we can use the fact that the sets $G_\xi$ can vary with $\xi$ and can be different from $\{m \in \mathcal{M}_b \mid -1 \leq m \leq +1\}$ to formulate a generalization of the randomization criterion (in the sense of Theorem 2) as follows:

Theorem 6: *Let*

$$\begin{aligned} E &:= (\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\in T}, (P_\vartheta)_{\vartheta\in\Theta}) \quad \textit{and} \\ F &:= (\Omega_2, \mathcal{B}, (\mathcal{B}_t)_{t\in T}, (Q_\vartheta)_{\vartheta\in\Theta}) \end{aligned}$$

*be filtered statistical experiments indexed by the same index sets* $\Theta$ *and* $T$. *Let* $(\varepsilon_{\vartheta,t})_{\vartheta\in\Theta,t\in T}$ *be a family of nonnegative real numbers and let* $(f_{\vartheta,t})_{\vartheta\in\Theta,t\in T}$ *be a family of* $\mathcal{B}$*-measurable functions* $f_{\vartheta,t} : \Omega_2 \to I\!R$. *Suppose that for any finite set* $\Theta_0 \times T_0 \subseteq \Theta \times T$ *and any selection* $(g_{\vartheta,t})_{\vartheta\in\Theta_0,t\in T_0}$ *of functions* $g_{\vartheta,t} \in \{g \in M_b(\Omega_2, \mathcal{B}_t) \mid -f_{\vartheta,t} \leq g \leq f_{\vartheta,t}\}$ *there exists a* $(\mathcal{A}_t)_{t\in T} - (\mathcal{B}_t)_{t\in T}$*−filtered stochastic operator* $S : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ *such that*

$$\int g_{\vartheta,t}\, dS(P_\vartheta) \leq \int g_{\vartheta,t}\, dQ_\vartheta + \varepsilon_{\vartheta,t} \quad \textit{for all} \quad \vartheta \in \Theta \quad \textit{and} \quad t \in T\ .$$

*Then there exists a* $(\mathcal{A}_t)_{t\in T}$*−*$(\mathcal{B}_t)_{t\in T}$*−filtered stochastic operator* $\widetilde{S} : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ *such that*

$$\sup_{\substack{g\in M_b(\Omega_2,\mathcal{B}_t) \\ -f_{\vartheta,t}\leq g\leq f_{\vartheta,t}}} \left(\int g\, d[\widetilde{S}(P_\vartheta) - Q_\vartheta]\right) \leq \varepsilon_{\vartheta,t}$$

*for all* $\vartheta \in \Theta$ *and* $t \in T$.

Remark: Note that even the case $T := \{t_0\}$ and $f_{\vartheta,t_0} = x^2$ for all $\vartheta$ was not covered by the classical randomization criterion. (Compare with Theorem 1 of this paper and with Theorem 3 of LeCam [1].)

More general we formulate:

Theorem 7: *Let*

$$\begin{aligned} E &:= (\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\in T}, (P_\vartheta)_{\vartheta\in\Theta}) \quad \textit{and} \\ F &:= (\Omega_2, \mathcal{B}, (\mathcal{B}_t)_{t\in T}, (Q_\vartheta)_{\vartheta\in\Theta}) \end{aligned}$$

*be filtered statistical experiments indexed by the same index sets* $\Theta$ *and* $T$, *let* $(\varepsilon_{\vartheta,t})_{\vartheta\in\Theta,t\in T}$ *be a family of nonnegative real numbers and let* $(G_{\vartheta,t})_{(\vartheta,t)\in\Theta\times T}$ *be a family of closed convex sets of bounded* $\mathcal{B}$*-measurable functions, such that* $0 \in G_{\vartheta,t} \subseteq M_b(\Omega_2, \mathcal{B}_t)$. *Suppose that for any finite set* $\Theta_0 \times T_0 \subseteq \Theta \times T$ *and any selection* $(g_{\vartheta,t})_{\vartheta\in\Theta_0,t\in T_0}$ *of functions* $g_{\vartheta,t} \in G_{\vartheta,t}$ *there exists a* $(\mathcal{A}_t)_{t\in T}-(\mathcal{B}_t)_{t\in T}$*−filtered stochastic operator* $S : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ *such that*

$$\int g_{\vartheta,t}\, dS(P_\vartheta) \le \int g_{\vartheta,t}\, dQ_\vartheta + \varepsilon_{\vartheta,t} \quad \textit{for all} \quad \vartheta \in \Theta \quad \textit{and} \quad t \in T\ .$$

*Then there exists a* $(\mathcal{A}_t)_{t\in T}-(\mathcal{B}_t)_{t\in T}$*−filtered stochastic operator* $\widetilde{S} : L(E) \mapsto ba(\Omega_2, \mathcal{B})$ *such that*

$$\sup_{g\in G_{\vartheta,t}} \Big(\int g\, d[\widetilde{S}(P_\vartheta) - Q_\vartheta]\Big) \le \varepsilon_{\vartheta,t}$$

*for all* $\vartheta \in \Theta$ *and* $t \in T$.

Remark: Theorem 6 is an immediate consequence of Theorem 7 and Theorem 7 can be proved in a completely analogous way as Theorem 5 has been proved. Even in the non filtered case (i.e. $T = \{t_0\}$) and even if we assume that the loss function space is independent of the parameter (i.e. $G_\vartheta = G$) Theorem 7 still remains very interesting as is shown by the following concrete example on stochastic orders:

Example: *Let* $F$ *denote the family of all measurable monotone increasing functions* $f : I\!R \to [0, 1]$. *Suppose that we are given two families* $(P_\vartheta)_{\vartheta\in\Theta}$ *and* $(Q_\vartheta)_{\vartheta\in\Theta}$ *of probability measures on* $(I\!R, \mathcal{B})$ *(with* $\mathcal{B}$ *the Borel* $\sigma$*-algebra on* $I\!R$*). Suppose further that for any finite set* $\Theta_0 \subset \Theta$ *and any family* $(f_\vartheta)_{\vartheta\in\Theta_0}$ *of functions* $f_\vartheta \in F$ *there exists a stochastic operator* $M_K$ *induced by a Markov kernel* $K$ *such that for all* $\vartheta \in \Theta_0$

$$\int f_\vartheta\, d[M_K(P_\vartheta) - Q_\vartheta] \le 0.$$

*Then there exists a stochastic operator $M$ such that*

$$\sup_{\vartheta\in\Theta}\sup_{f\in F}\int f\, d[M(P_\vartheta)-Q_\vartheta]\le 0;$$

*i.e. there exists a stochastic operator $M$ such that $M(P_\vartheta)\overset{(1)}{\le} Q_\vartheta$ for all $\vartheta\in\Theta$, with $\overset{(1)}{\le}$ the stochastic order defined in Section 1.2 of [7].*

To see that the example holds simply apply Theorem 7 in the case that there is no filtration (i.e. $T=\{t_0\}$), that $G_{\vartheta,t_0}=F$ for all $\vartheta\in\Theta$ and note that $F$ induces the relation $\overset{(1)}{\le}$ in the sense of [7]; i.e. $\mu_1\overset{(1)}{\le}\mu_2 \Leftrightarrow \int f d\mu_1\le\int f d\mu_2$ for all $f\in F$.

## Application to Hilbert spaces

The following theorem is an easy consequence of Theorem 3' but it can not be established using the classical randomization criterion.

Theorem 8: *Let $\mathcal{H}$ be a Hilbert space and let $\Theta$ be a set. Let $(v_\vartheta)_{\vartheta\in\Theta}$ be a indexed family of vectors in $\mathcal{H}$ and let $(K_\vartheta)_{\vartheta\in\Theta}$ be a family of convex norm-closed norm-bounded subsets of $\mathcal{H}$. Let $\mathcal{L}$ be a norm-bounded norm-closed convex family of linear operators $l:\mathcal{H}\to\mathcal{H}$ such that for any finite set $\Theta_0\subseteq\Theta$ and any family $(g_\vartheta)_{\vartheta\in\Theta_0}$ of elements of $\mathcal{H}$ with $\|g_\vartheta\|_2=1$ there exists an $l\in\mathcal{L}$ such that*

$$\inf_{w_\vartheta\in K_\vartheta}\langle\, l(v_\vartheta)-w_\vartheta\ ,\ g_\vartheta\,\rangle\ \le\ \varepsilon_\vartheta\ .$$

*Then there exists an $l\in\mathcal{L}$ such that for all $\vartheta\in\Theta$ we have:*

$$\inf_{w_\vartheta\in K_\vartheta}\|l(v_\vartheta)-w_\vartheta\|_2\ \le\ \varepsilon_\vartheta\ .$$

Proof of Theorem 8: Note that the weak dual of a Hilbert space is the Hilbert space itself endowed with the weak topology induced by itself via the inner product. Note further that the hypothesis $\|g_\vartheta\|_2=1$ is equivalent with the hypothesis $\|g_\vartheta\|_2\le 1$. Note further, that the sets $K_\vartheta$ and the set $\mathcal{L}$ are weakly compact. (This is the theorem of Alaoglu-Bourbaki in the case of the sets $K_\vartheta$ and can in the case of the set $\mathcal{L}$ be concluded using the Tychonoff Theorem analogous to the theorem of Alaoglou-Bourbaki [see [5] Chapter 3 Section 4.3]). Apply now Theorem 3' with $V=W=\mathcal{H}$ and $G=\{\|g_\vartheta\|_2\le 1\}$. $\square$

Finally we show that it is possible to reverse the role of the space of stochastic operators and the experiment $E:=(\Omega,\mathcal{A},(P_\vartheta)_{\vartheta\in\Theta})$ in Theorem 2. To be more precise we state and prove the following Theorem:

Theorem 9: Let $\Theta$ be an index set. Let $\Upsilon \subset ba(\Omega, \mathcal{A})$ be compact and convex. Let $(f_\vartheta)_{\vartheta\in\Theta}$ be a family of functions from $L(\Omega, \mathcal{A}, \Upsilon)$ to $ba(\Omega_2, \mathcal{B})$ and let $(Q_\vartheta)_{\vartheta\in\Theta}$ be a family of elements of $ba(\Omega_2, \mathcal{B})$. Denote by $\mathcal{M}_b(\Omega_2, \mathcal{B})$ the space of bounded measurable real valued functions endowed with the supremum norm.

Suppose that for any finite set $\Theta_0 \subseteq \Theta$ and any family $(g_\vartheta)_{\vartheta\in\Theta}$ of functions $g_\vartheta \in \mathcal{M}_b(\Omega_2, \mathcal{B})$ with $\sup_{x\in\Omega_2} |g_\vartheta| \leq 1$ there exists a $P \in \Upsilon$ such that

$$\int g_\vartheta \; d[f_\vartheta(P)] \leq \int g_\vartheta \; dQ_\vartheta + \varepsilon_\vartheta \; .$$

Then there exists a $P \in \Upsilon$ such that $\forall \vartheta \in \Theta$

$$\|f_\vartheta(P) - Q_\vartheta\| \leq \varepsilon_\vartheta .$$

Remark: The functions $f_\vartheta : L(\Omega, \mathcal{A}, \Upsilon) \mapsto ba(\Omega_2, \mathcal{B})$ can be arbitrary. They can of course be of the form $f_\vartheta = S_\vartheta$ or $f_\vartheta = S_\vartheta - id$, with $id$ the identity on $L(\Omega, \mathcal{A})$ and $(S_\vartheta)_{\vartheta\in\Theta}$ a family of stochastic operators. If we let $f_\vartheta = S_\vartheta - id$ and $Q_\vartheta = 0$ for all $\vartheta \in \Theta$ then we obtain the following corollary:

Corollary: Let $\Upsilon \subset ba(\Omega, \mathcal{A})$ be compact and convex and let $(S_\vartheta)_{\vartheta\in\Theta}$ be a family of stochastic operators from $L(\Omega, \mathcal{A}, \Upsilon)$ to $ba(\Omega_2, \mathcal{B})$. Suppose that for any finite $\Theta_0 \subseteq \Theta$ and any family $(g_\vartheta)_{\vartheta\in\Theta_0}$ with $g_\vartheta \in \mathcal{M}_b(\Omega_2, \mathcal{B})$ and $\sup_{x\in\Omega_2} |g_\vartheta| \leq 1$ there exists a $\mu \in \Upsilon$ with

$$\int g_\vartheta \; d[S_\vartheta(\mu) - \mu] \leq \varepsilon_\vartheta \; .$$

Then there exists a $\mu$ in $\Upsilon$ such that for all $\vartheta \in \Theta$ we have

$$\|S_\vartheta(\mu) - \mu\| \leq \varepsilon_\vartheta \; .$$

Proof of Theorem 9: To obtain Theorem 9 simply apply Theorem 3 with $v_\vartheta = f_\vartheta$, $w_\vartheta = Q_\vartheta$, $\mathcal{L} = \Upsilon$, $G := \{g \in \mathcal{M}_b(\Omega_2, \mathcal{B}) \mid |g| \leq 1\}$ in a completely analogous way as in the derivation of Theorem 2 from Theorem 3.

# 5 The finite dimensional case and Helly's Theorem

We consider now the case that the spaces $L(E)$ and $L(F)$ or two Hilbert spaces or more general two arbitrary vector spaces are finite dimensional. This makes it possible to prove a theorem which is in principle of the form of the theorems

2 - 9, but which involves in its hypotheses only a constant finite number (depending on the dimension of the vector spaces involved) of constraints.

Theorem 10: *Let $\mathcal{L}$ be a compact convex family of linear mappings $l : I\!R^m \to I\!R^n$, let $\Theta$ be an index set, let $(Z_\vartheta)_{\vartheta\in\Theta}$ be a family of convex subsets of $I\!R^n$ and let $(\varepsilon_\vartheta)_{\vartheta\in\Theta}$ be a family of real numbers. Let further $(x_\vartheta)_{\vartheta\in\Theta}$ with $x_\vartheta \in I\!R^m$ and $(y_\vartheta)_{\vartheta\in\Theta}$ with $y_\vartheta \in I\!R^n$ be given. Let $I$ be an index set consisting of $(m \cdot n) + 1$ points. Suppose that for any indexed family $(\vartheta_i)_{i\in I}$ with $\vartheta_i \in \Theta$ and any indexed family $(z_{\vartheta_i})_{i\in I}$ with $z_{\vartheta_i} \in Z_{\vartheta_i}$ there exists an $l \in \mathcal{L}$ such that*

$$\langle\, l(x_{\vartheta_i}) - y_{\vartheta_i}\ ,\ z_{\vartheta_i}\,\rangle \le \varepsilon_{\vartheta_i}.$$

*Then there exists an $l \in \mathcal{L}$ such that for all $\vartheta \in \Theta$ we have*

$$\sup_{z\in Z_\vartheta} \langle\, l(x_\vartheta) - y_\vartheta\ ,\ z\,\rangle \le \varepsilon_\vartheta.$$

Proof of Theorem 10: The sets

$$\mathcal{L}_{\vartheta,z} := \{\ l \in \mathcal{L}\ \text{ such that }\ \langle\, l(x_\vartheta) - y_\vartheta\ ,\ z\,\rangle \le \varepsilon_\vartheta\ \}$$

with $z \in Z_\vartheta$ are compact convex subsets of $\mathcal{L}$ and thus subsets of the $m \cdot n$-dimensional space of all linear mappings from $\mathbb{R}^m$ to $\mathbb{R}^n$. Our Hypothesis says that the intersection of any $(m \cdot n) + 1$ sets $\mathcal{L}_{\vartheta,z}$ wit $z \in Z_\vartheta$ is nonempty. Therefore by Helly's Theorem (see [11] Part VI) the intersection $\bigcap_{\vartheta\in\Theta, z\in Z_\vartheta} \mathcal{L}_{\vartheta,z}$ is nonempty which is precisely the conclusion of our Theorem. $\Box$

# 6 Appendix

Let $\mathcal{M}_b$ denote the vector space of $\mathcal{B}$-measurable real valued functions on $(\Omega_2, \mathcal{B})$. In the following Lemmata we denote by $\sigma'$ the topology on $ba(\Omega_2, \mathcal{B})$ induced by the integrals $I_m : (\Omega_2, \mathcal{B}) \to \mathbb{R}$ defined by $I_m := \int m\, d\mu$. By $E := (\Omega, \mathcal{A}, (P_\vartheta)_{\vartheta\in\Theta})$ we denote a statistical experiment.

Lemma A1 *The set $\{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \le 1\}$ is $\sigma'$-compact.*

Sketch of proof: Let $G := \{m \in \mathcal{M}_b \mid -1 \le m(\omega) \le 1\ \text{ for all }\ \omega \in \Omega_2\}$ and let $I : (\Omega_2, \mathcal{B}) \to \mathbb{R}^G$ be the unique mapping such that $I_m = pr_m \circ I$ for all $m \in \mathcal{G}$. The topology $\sigma'$ is of course the same as the topology induced by the family of integrals $\{I_m \mid m \in G\}$. Thus $(\{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \le 1\}, \sigma')$ is homeomorphic with the set $I(\{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \le 1\}) \subseteq [-1, +1]^G$. Since $[-1, +1]^G$ is by the Tychonoff product theorem compact, it suffices to prove that $I(\{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \le 1\})$ is closed in $[-1, +1]^G$. But this is clear since

$$I(\{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \le 1\}) =$$

$$= \{f \in [-1,+1]^G \quad | \quad -1 \leq pr_{m_1+m_2}(f) = pr_{m_1}(f) + pr_{m_2}(f) \leq 1$$
$$\forall\ m_1, m_2 \in \mathcal{M}_b \ \text{ with } \ -1 \leq m_1, m_2, m_1+m_2 \leq 1\}. \ \square$$

Lemma A2 *The space $\mathcal{S}$ of stochastic operators $S : L(E) \to ba(\Omega_2, \mathcal{B})$ is a compact subset of $(ba(\Omega_2, \mathcal{B}))^{L(E)}$.*

Sketch of proof: We have

$$\mathcal{S} \subseteq \prod_{\mu \in L(E)} \{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \leq 2\|\mu\|\}$$

and we know by the preceding Lemma and the Tychonoff product theorem that $\prod_{\mu \in L(E)} \{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \leq 2\|\mu\|\}$ is compact. It therefore suffices to show that $S$ is closed in $\prod_{\mu \in L(E)} \{\nu \in ba(\Omega_2, \mathcal{B}) \mid \|\nu\| \leq 2\|\mu\|\}$. But this is clear since the space of stochastic operators can be described by

$$\begin{aligned} \mathcal{S} := \{S \quad | \quad & S \in ba(\Omega_2, \mathcal{B})^{L(E)}, S \ \text{ is linear and} \\ & [\mu \in L^+(E) \wedge \|\mu\| = 1] \Longrightarrow \\ & [\forall m \in \mathcal{M}_b (m \geq 0 \Rightarrow I_m(S(\mu)) \geq 0) \ \text{ and } \ I_1(S(\mu)) = 1]\}. \ \square \end{aligned}$$

# References


[1] Le Cam, L. Sufficiency and approximate sufficiency. *Ann. Math. Statist.* **35** (1964) 1419–1455. **34** #6909

[2] Le Cam, Lucien; Yang, Grace Lo *Asymptotics in statistics. Some basic concepts. Second edition.* Springer Series in Statistics. Springer-Verlag, New York, 2000. **2001f**:62019

[3] E. Norberg *Approximate Comparison of Statistical Experiments with Filtered Probability Spaces* Preprint (University of Oslo): http://www.math.uio.no/eprint/stat_report/1999/rept_1999.html

[4] Shiryaev A.N., Spokoiny V. G. *Statistical Experiments and Decisions : Asymptotic Theory* Advanced Series on Statistical Science and Applied Probability Volume **8** World Scientific 2001.

[5] Schaefer, H. H.; Wolff, M. P. *Topological vector spaces. Second edition.* Graduate Texts in Mathematics, **3**. Springer-Verlag, New York, 1999. **2000j**:46001

[6] Segal, Irving E.; Kunze, Ray A. *Integrals and operators. Second revised and enlarged edition.* Grundlehren der Mathematischen Wissenschaften, Band **228**. Springer-Verlag, Berlin-New York, 1978. **58** #6126

[7] Stoyan, Dietrich *Qualitative Eigenschaften und Abschätzungen stochastischer Modelle. (German)* Akademie-Verlag, Berlin, 1977. **56** #13397

[8] Stoyan, Dietrich *Comparison methods for queues and other stochastic models. Translation from the German edited by Daryl J. Daley.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 1983. **85f**:60147

[9] Strasser, Helmut *Mathematical theory of statistics. Statistical experiments and asymptotic decision theory.* de Gruyter Studies in Mathematics, 7. Walter de Gruyter & Co., Berlin, 1985. **87h**:62034

[10] Torgersen, Erik *Comparison of statistical experiments.* Encyclopedia of Mathematics and its Applications, **36**. Cambridge University Press, Cambridge, 1991. **92i**:62007

[11] Valentine, Frederick A. *Convex sets.* McGraw-Hill Series in Higher Mathematics McGraw-Hill Book Co., New York-Toronto-London 1964 **30** #503

Authors Address:

Heinz Weisshaupt
EURANDOM, P.O. Box 513
5600 MB Eindhoven
The Netherlands