

A generalization of Lehmann's Theorem on the comparison of uniform location experiments

Heinz Weisshaupt

University of Vienna, Department of Statistics*
A-1010 Vienna, Universitaetsstrasse 5/3, Austria
e-mail: heinz.weisshaupt@univie.ac.at

October 21, 2010

Abstract

We generalize a result of Lehmann on the comparison of location experiments with uniform distributions on intervals. We compare in this paper a location experiment consisting of uniform distributions on parallelepipeds with a location experiment consisting of uniform distributions on convex bodies. We show that the first experiment can only be more informative than the second one if the convex bodies in the second experiment are themselves parallelepipeds. Further we show that the length of the edges of these parallelepipeds must fulfill a condition similar to the condition on the length of the intervals in Lehmann's result.

Keywords: location experiments, comparison, uniform distribution, convex bodies

AMS Subject Classification: 62B15, 52A20

1. Introduction

We define a statistical experiment \mathbf{F} as a random quantity X for which an indexed family $(P_\vartheta)_{\vartheta \in \Theta}$ of possible distributions of X on a measurable space (Ω, \mathcal{A}) is given. Let $\mathbf{E} = ((Q_\vartheta)_{\vartheta \in \Theta}, (\Omega_2, \mathcal{A}_2))$ be a second statistical experiment with the same index set Θ . We say that the experiment \mathbf{F} is more informative than the experiment \mathbf{E} (or that \mathbf{E} is a randomization of \mathbf{F}), if there exists a Markov kernel $K : \Omega \times \mathcal{A}_2 \mapsto [0, 1]$ with the property that $\mathbb{E}_{P_\vartheta} K(X, A) = Q_\vartheta(A)$ with \mathbb{E}_{P_ϑ} the expectation under the condition that X is P_ϑ distributed. (For the general definition of randomization and the general theory concerned with the comparison of statistical Experiments see [1] and [2].)

*The paper has been written during the time the author spent at the research institute EURANDOM and the mathematics department of the University of Duesseldorf

Let $M \subset \mathbb{R}^m$ be a Borel measurable set of finite Lebesgue measure. We denote by $\mathbf{U}(M)$ the uniform distribution on M , i.e., $\mathbf{U}(M) := \frac{\lambda|_M}{\lambda(M)}$, where λ denotes Lebesgue measure and $\lambda|_M(A) = \lambda(M \cap A)$ for all Borel measurable $A \subset \mathbb{R}^m$. We denote especially by $\mathbf{U}([a, b])$ the uniform distribution on an arbitrary bounded interval $[a, b] \subset \mathbb{R}$. By the term measure we denote a positive or signed Borel measure on \mathbb{R}^m . By the term finite measure we denote a measure of finite total variation. Given two measures μ and ν we denote by $\mu * \nu$ the convolution of μ and ν defined on Borel sets $B \subset \mathbb{R}^m$ by $(\mu * \nu)(B) := (\mu \otimes \nu)(\{(x, y) \in \mathbb{R}^m \times \mathbb{R}^m \text{ such that } x + y \in B\})$ (if it exists). We denote by δ_x the Dirac measure at x . We restate now the result of Lehmann (Theorem 3.1 of [3]) which is the starting point of our investigations.

THEOREM 1 *Let two uniform location experiments*

$$\mathbf{F} = (\mathbf{U}([0, 1]) * \delta_{\vartheta})_{\vartheta \in \mathbb{R}}, (\mathbb{R}, \mathcal{B}) \quad \text{and} \quad \mathbf{E} = (\mathbf{U}([0, r]) * \delta_{\vartheta})_{\vartheta \in \mathbb{R}}, (\mathbb{R}, \mathcal{B})$$

be given. The experiment \mathbf{F} is more informative than the experiment \mathbf{E} if and only if r is an integer ≥ 1 .

Instead of $\mathbf{U}([0, r]) * \delta_{\vartheta}$ or more general $\mathbf{U}(V) * \delta_{\vartheta}$ we also write $\mathbf{U}([0, r] + \vartheta)$ respectively $\mathbf{U}(V + \vartheta)$ since this notation is for the description of differentials of the function $t \rightarrow \mathbf{U}(C + t)$ more convenient.

The purpose of the paper is to extend Theorem 1. This extension compares a location experiment consisting of uniform distributions on parallelepipeds with a location experiment consisting of uniform distributions on convex bodies (convex compact subsets of \mathbb{R}^m which possess nonempty interior). The precise result is displayed by Theorem 2 following the introduction. To establish the result we make use of weak derivatives of measure-valued (set-valued) functions as defined in [4] and [6]. As in [3] we establish the result without using the Theorem of Boll which says that in the case of a location experiment we can always take a convolution kernel instead of a general Markov kernel in the definition of randomization. For simplicity of notation we introduce the stochastic operator S associated with the Markov kernel K by $[S(\mu)](A) := \int K(x, A) d\mu$ for arbitrary measurable sets A and finite signed measures μ , i.e., S is a linear mapping from a space of signed measures to a space of signed measures, which maps probability measures onto probability measures.

The paper is organized as follows. In section 2 we display the result (Theorem 2) of the paper. Further we display Theorem 3 which characterizes parallelepipeds within the space of convex bodies. A proof of this theorem is given in [7]. In section 3 we prove Theorem 2 using Theorem 3. In section 4 we sketch a proof of Theorem 2 which uses the theorem of Boll.

2. The main Theorem

We denote by $\mathbb{N} = \{1, 2, \dots\}$ the set of positive integers. By \bar{V} we denote the closure of a set V . By $\mathbf{1}_B$ we denote the indicator function of B . Given a function f and a measure μ we denote by $f \cdot \mu$ the measure $[f \cdot \mu](A) := \int_A f d\mu$ (if it exists).

DEFINITION 1 *Given a set $\{v_1, \dots, v_n\}$ of linearly independent vectors in \mathbb{R}^m then we call the set $V := \{v = \sum_i \alpha_i v_i \mid 0 \leq \alpha_i < 1\}$ the parallelepiped spanned by $\{v_1, \dots, v_n\}$. A set C is called a (closed) parallelepiped if $C = \bar{V} + \xi$ for some arbitrary $\xi \in \mathbb{R}^m$.*

REMARK 1 *Since the sets $V + j \cdot v_i$ with $j \in \mathbb{N} \cup \{0\}$ are pairwise disjoint we get that for $\ell \in \mathbb{N}$*

$$\mathbf{1}_{(\cup_{j=0}^{\ell} (V+j \cdot v_i))} = \sum_{j=0}^{\ell} \mathbf{1}_{(V+j \cdot v_i)},$$

and, for $0 < h < 1$, we get by the special location of the parallelepipeds $V + j \cdot v_i$ that

$$\mathbf{1}_{V \cup (V+(\ell+h) \cdot v_i)} \cdot \mu \geq \left| \sum_{j=0}^{\ell} \mathbf{1}_{(V+(j+h) \cdot v_i)} \cdot \mu - \sum_{j=0}^{\ell} \mathbf{1}_{(V+j \cdot v_i)} \cdot \mu \right|$$

THEOREM 2 *Let $V \subset \mathbb{R}^m$ be a parallelepiped with nonempty interior, spanned by the vectors v_1, \dots, v_m . Let $C \subset \mathbb{R}^m$ be a convex body. Let $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a bijective linear mapping and let the experiments*

$$\mathbf{F} = ((\mathbf{U}(V + \vartheta))_{\vartheta \in \mathbb{R}^m}, (\mathbb{R}^m, \mathcal{B}^m))$$

and

$$\mathbf{E} = ((\mathbf{U}(C + f(\vartheta)))_{\vartheta \in \mathbb{R}^m}, (\mathbb{R}^m, \mathcal{B}^m))$$

be given. Then \mathbf{F} is more informative than \mathbf{E} if and only if there exists a $\xi \in \mathbb{R}^m$ such that C is a parallelepiped of the form

$$C := \{x \mid x = \sum_{i=1}^m \alpha_i \cdot \gamma_i \cdot f(v_i) \mid 0 \leq \alpha_i \leq 1\} + \xi \quad \text{with } \gamma_i \in \mathbb{N}. \quad (1)$$

REMARK 2 *We will prove Theorem 2 in 3 steps. First we show that if C is of the form described by formula (1), then \mathbf{E} is a randomization of \mathbf{F} . This is done by writing down the Markov kernel without any difficulty.*

In the second step we show that if \mathbf{E} is a randomization of \mathbf{F} then C is of the form

$$C = \{x \mid x = \sum_{i=1}^m \alpha_i \cdot \gamma_i \cdot f(v_i) \mid 0 \leq \alpha_i \leq 1\} + \xi \quad \text{with } \gamma_i \in \mathbb{R}^+ \text{ fixed}. \quad (2)$$

This is the main part of the proof. It uses a characterization of parallelepipeds within the space of convex bodies displayed as Theorem 3.

Finally we show that the γ_i which occur in the formula (2) describing C are positive integers. We derive easily this fact from (2) and Theorem 1 (Theorem 3.1 of [3]). (It is also possible to prove $k_i \in \mathbb{N}$ by differentiation of measure valued functions, but by using Theorem 3.1 of [3] we obtain a very quick and easy proof.)

DEFINITION 2 We denote by \mathcal{C}_c the space of continuous real valued functions with compact support on \mathbb{R}^m . We remark that a finite signed measure μ is determined by the integrals $\int \phi d\mu$ with $\phi \in \mathcal{C}_c$. Let C be a convex set, then we denote by $\lim_{h \rightarrow 0} \frac{\mathbf{U}(C+h \cdot w) - \mathbf{U}(C)}{h}$ the measure τ_w determined by

$$\int \phi d\tau_w = \lim_{h \rightarrow 0} \int \phi d \frac{\mathbf{U}(C + h \cdot w) - \mathbf{U}(C)}{h}.$$

In the proof of Theorem 2 we make use of the following characterization of parallelepipeds, which has been proved in [7].

THEOREM 3 A convex body $C \subset \mathbb{R}^m$ is a parallelepiped if and only if there exists a linearly independent set $\{w_1, \dots, w_m\}$ of vectors $w_i \in \mathbb{R}^m$ such that for any $i \in \{1, \dots, m\}$ there exists an at most countable set $R_i \subset \mathbb{R}$ and an $r_{0i} \in R_i$ such that

$$\left[\lim_{h \rightarrow 0} \frac{\mathbf{U}(C + h \cdot w_i) - \mathbf{U}(C)}{h} * \delta_{\{r \cdot w_i | r \in R_i\}} \right](B) = 0.$$

for any Borel measurable set $B \subseteq \partial C + r_{0i} \cdot w_i$. If the convex body C is a parallelepiped then the vectors w_1, \dots, w_m are parallel to the one dimensional edges of C .

3. Proof of Theorem 2

REMARK 3 Given a finite signed measure μ we denote by $|\mu|$ the measure defined by $|\mu|(B) := \sup\{\mu(A_1) - \mu(A_2) \mid A_1 \cup A_2 = B \text{ and } A_1 \cap A_2 = \emptyset\}$.

Note the following two properties of stochastic operators S :

$$S(|\mu|) \geq |S(\mu)| \text{ for any finite signed measure } \mu \quad (3)$$

and, for bounded functions f, g with $f \geq g$

$$S(f \cdot \mu) \geq S(g \cdot \mu) \text{ for any finite positive measure } \mu. \quad (4)$$

DEFINITION 3 We say that a signed measure μ vanishes on a set $D \subseteq \mathbb{R}^m$ if $\mu(B) = 0$ for all Borel measurable sets $B \subseteq D$

PROOF OF THEOREM 2. First Step: We prove the if part first. Suppose that C fulfills (1), i.e., $C = g \circ f(V) + \xi$ with $g(f(v_i)) = \gamma_i \cdot f(v_i)$ with $\gamma_i \in \mathbb{N}$. We define a Markov kernel $\rho : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \rightarrow [0, 1]$ by

$$\rho(x, B) := \frac{1}{\gamma_1 \cdot \dots \cdot \gamma_m} \sum_{j_1=0}^{\gamma_1-1} \cdots \sum_{j_m=0}^{\gamma_m-1} \delta_{(\sum_{i=1}^m j_i \cdot f(v_i) + \xi)}(B).$$

It is easily seen, that the stochastic operator S defined by $[S(\mu)](B) := \int \rho(x, B) d\mu(x)$ fulfills $S(\mathbf{U}(V + \vartheta)) = \mathbf{U}(C + f(\vartheta))$, which proves the if part.

Second Step: Now we prove the only if part. Let S be a stochastic operator such that

$$S(\mathbf{U}(V + \vartheta)) = \mathbf{U}(C + f(\vartheta)). \quad (5)$$

Let $\mu = \lambda \cdot [\lambda(V)]^{-1}$ and let $\nu = \lambda \cdot [\lambda(C)]^{-1}$. Let $i \in \{1, \dots, m\}$ and $\ell \in 2\mathbb{N}$ be minimal such that

$$C \cap (C + \frac{\ell}{2} \cdot f(v_i)) = \emptyset \quad \text{and} \quad (C + \frac{\ell}{2} \cdot f(v_i)) \cap (C + \ell \cdot f(v_i)) = \emptyset. \quad (6)$$

By convexity of C we obtain from (6) that for $h > 0$

$$C \cap (C + (\ell + h) \cdot f(v_i)) = \emptyset. \quad (7)$$

Let $0 < h < 1$ then

$$\mathbf{1}_{C \cup (C + (\ell + h) \cdot f(v_i))} \cdot \nu = \mathbf{1}_C \cdot \nu + \mathbf{1}_{C + (\ell + h) \cdot f(v_i)} \cdot \nu \quad (8)$$

$$= \mathbf{U}(C) + \mathbf{U}(C + (\ell + h) \cdot f(v_i)) = S(\mathbf{U}(V)) + S(\mathbf{U}(V + (\ell + h) \cdot v_i)) \quad (9)$$

$$= S(\mathbf{1}_V \cdot \mu) + S(\mathbf{1}_{V + (\ell + h) \cdot v_i} \cdot \mu) = S(\mathbf{1}_{V \cup (V + (\ell + h) \cdot v_i)} \cdot \mu) \quad (10)$$

$$\geq S(|\sum_{j=0}^{\ell} [\mathbf{1}_{(V + (j+h) \cdot v_i)} - \mathbf{1}_{(V + j \cdot v_i)}] \cdot \mu|) \quad (11)$$

$$\geq |\sum_{j=0}^{\ell} S([\mathbf{1}_{(V + (j+h) \cdot v_i)} - \mathbf{1}_{(V + j \cdot v_i)}] \cdot \mu)| \quad (12)$$

$$= |\sum_{j=0}^{\ell} S([\mathbf{U}(V + (j+h) \cdot v_i) - \mathbf{U}(V + j \cdot v_i)])| \quad (13)$$

$$= |\sum_{j=0}^{\ell} [\mathbf{U}(C + f((j+h) \cdot v_i)) - \mathbf{U}(C + f(j \cdot v_i))]| \quad (14)$$

The equality sign in equation (8) follows from (7), equality in equation (9) from (5) and the equality sign in (10) from the disjointness of V and $V + (\ell + h)v_i$

(see Remark 1) and the linearity of stochastic operators. The equality signs between the equations (8), (9) and (10) follow from the definitions of ν and μ . The inequality between (10) and (11) follows from (4) and Remark 1. The inequality between (11) and (12) follows from (3). The equalities between (12), (13) and (14) follows from the definition of μ and (5).

From the inequality between (8) and (14) we obtain that for arbitrary $h \in (0, 1)$ the measure

$$\sum_{j=0}^{\ell} [\mathbf{U}(C + f((j+h) \cdot v_i)) - \mathbf{U}(C + f(j \cdot v_i))]$$

vanishes on any set B with $B \cap C \cup \bigcup_{h \in [0,1]} (C + (\ell+h) \cdot f(v_i)) = \emptyset$. So also the measures

$$\sum_{j=0}^{\ell} \frac{\mathbf{U}(C + f((j+h) \cdot v_i)) - \mathbf{U}(C + f(j \cdot v_i))}{h} \quad (15)$$

vanish on such sets B . Since $C \cup \bigcup_{h \in [0,1]} (C + (\ell+h) \cdot f(v_i))$ is closed we get that also the weak limit of (15) which is given by

$$\begin{aligned} & \lim_{h \downarrow 0} \sum_{j=0}^{\ell} \frac{\mathbf{U}(C + f((j+h) \cdot v_i)) - \mathbf{U}(C + f(j \cdot v_i))}{h} \\ &= \lambda(C) \cdot \lim_{h \rightarrow 0} \frac{\mathbf{U}(C + h \cdot f(v_i)) - \mathbf{U}(C)}{h} * \delta_{\{j \cdot f(v_i) | j \in \{0, \dots, \ell\}\}} \end{aligned} \quad (16)$$

vanishes on any set B with $B \cap [C \cup C + \ell \cdot f(v_i)] = \emptyset$.

Thus by (6)

$$\lambda(C) \cdot \lim_{h \rightarrow 0} \frac{\mathbf{U}(C + h \cdot f(v_i)) - \mathbf{U}(C)}{h} * \delta_{\{j \cdot f(v_i) | j \in \{0, \dots, \ell\}\}}$$

vanishes especially on the set $\partial(C + \frac{\ell}{2} \cdot v_i)$. So by Theorem 3 with $R_i = \{0, \dots, \ell\}$, $r_{0i} = \frac{\ell}{2}$ and $w_i = f(v_i)$ we obtain that C is a parallelepiped.

Finally it remains to be proved, that the length of the edge of C parallel to $f(v_i)$ is an integer multiple of $f(v_i)$, i.e., we have to show that $\gamma_i \in \mathbb{N}$. This is done as follows:

By a shift of the experiment \mathbf{F} we may suppose without loss of generality that C possesses a vertex at 0, i.e.,

$$C = \{x \mid x = \sum_{i=1}^m \alpha_i \cdot \gamma_i \cdot f(v_i) \mid 0 \leq \alpha_i \leq 1\} \quad \text{with } \gamma_i \in \mathbb{R}^+. \quad (17)$$

Let T be the stochastic operator which maps δ_{ϑ} for $\vartheta \in \mathbb{R}$ to the uniform distribution \mathbf{U}_i on $V_i + \vartheta v_i$ with V_i the parallelepiped spanned by the linearly independent set $\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m\}$, i.e., $T(\delta_{\vartheta}) = \mathbf{U}(V_i) * \delta_{\vartheta \cdot v_i}$. We denote by C_i the parallelepiped spanned by $\{f(v_1), \dots, f(v_{i-1}), f(v_{i+1}), \dots, f(v_m)\}$. Let

$pr_i : \mathbb{R}^m \rightarrow \mathbb{R}$ be the unique linear mapping with $pr_i(C_i) = 0$ and $pr_i(f(v_i)) = 1$. Let Pr_i be the stochastic operator which maps for any $x \in \mathbb{R}^m$ the Dirac measure δ_x to the Dirac measure $\delta_{pr_i(x)}$. Then the operator $Pr_i \circ S \circ T$ maps the uniform distribution on $[0, 1] + \vartheta$ to the uniform distribution on the interval $[0, \gamma_i] + \vartheta$. Thus an application of Theorem 1 shows that $\gamma_i \in \mathbb{N}$. \square

4. A proof based on the theorem of Boll

We finally sketch a proof of Theorem 2 based on the theorem of Boll. Note that this is only a very short sketch! A complete outline of the proof would be at least as long as the proof given above. The theorem of Boll says that in the case of location experiments on \mathbb{R}^m the Markov kernel has the special form of a convolution kernel, i.e., $Q_{\vartheta} = \nu * P_{\vartheta}$ for some probability measure ν on \mathbb{R}^m . (For a general version of the theorem of Boll see [5] Theorem 55.12. and Theorem 38.15.)

Without loss of generality we replace the linear mapping f by the identity mapping $id: \mathbb{R}^m \rightarrow \mathbb{R}^m$. By the theorem of Boll we obtain that $\mathbf{U}(C) = \mathbf{U}(V) * \tau$ for some probability measure τ . This equation can be rewritten as

$$\mathbf{1}_C(x) = \tau^*(x - \text{int}(V)) \quad \text{with respect to } \lambda \quad \text{a.e.} \quad (18)$$

with τ^* a multiple of τ . The measure τ^* has by boundedness of C compact support which together with (18) shows that τ^* must be discrete. So

$$\tau^* = \sum_{i=1}^n \alpha_i \delta_{a_i} \quad \text{for some } n \in \mathbb{N} \quad (19)$$

From (18) and (19) one obtains that

$$a_j \notin a_i + \text{int}(V) \quad (20)$$

Finally one proves using (18), (19) and (20) that $\bigcup_{i=1}^n \text{int}(V) + a_i \subseteq C \subseteq \bigcup_{i=1}^n \bar{V} + a_i$, where we denote by \bigcup the union if it is taken over a pairwise disjoint family of sets. From this and the convexity of C it follows that C must itself be a parallelepiped with 1-dimensional edges in the direction of the 1-dimensional edges of V and with edge-length integer multiples of the edge-length of V .

Acknowledgment: I would like to thank Georg Pflug for many fruitful discussions and advices concerning the paper.

References

- [1] Le Cam, Lucien Sufficiency and approximate sufficiency *Ann. Math. Stat.* **35**, 1419-1455 (1964).
- [2] Le Cam, Lucien *Asymptotic methods in statistical decision theory*. Springer Series in Statistics. (New York etc.: Springer-Verlag, 1986).
- [3] Lehmann, E.L. Comparing location experiments. *Ann. Stat.* **16**, No.2, 521-533 (1988).

- [4] Pflug, Georg *Optimisation of Stochastic Models*. Kluwer Academic Publishers (Boston, 1996).
- [5] Strasser, Helmut *Mathematical Theory of Statistics*. Walter de Gruyter (Berlin - New York, 1985)
- [6] Weisshaupt, Heinz A measure-valued approach to convex set-valued dynamics. *Set-Valued Anal.* **9**, No.4, 337-373 (2001)
- [7] Weisshaupt, Heinz A characterization of parallelepipeds related to weak derivatives. *Acta Math. Hungar.* **108**