DIFFICULTIES OF THE LINDLEY–SAVAGE ARGUMENT

1. INTRODUCTION

From about the year 1971 Allan Birnbaum was much concerned with an influential piece of reasoning which he called the Lindley–Savage argument (L. J. Savage [1961] pp. 180–1, [1962] pp. 171–4, D. V. Lindley [1971] pp. 13–4). Lindley and Savage had applied the argument within the realm of statistical inference, and Birnbaum was specifically concerned to refute this application of it, and not versions of the argument such as that of Raiffa and Schlaifer ([1961] pp. 24–5), which is in explicitly decision-theoretic terms. Birnbaum considered that the argument was essentially decision-theoretic in inspiration, and he obviously felt that to understand why it could not be carried across to statistical inference would give great insight into the nature of that activity. In all this Birnbaum appeared to concede that within decision theory the Lindley–Savage argument, and hence the Bayesian thinking which it supports, were perhaps wholly correct. In this paper I shall try to show that even within decision theory the Lindley–Savage argument, and possible variants of it, are not without substantial difficulties. During the argument we shall consider a version of Birnbaum's own Principle of Conditionality.

2. THE LINDLEY–SAVAGE ARGUMENT

I shall assume that the reader is acquainted with the basic concepts of Wald's decision theory (see for example Ferguson [1967]), and shall confine my attention to decision problems over a finite parameter space of \( n \) points. It is a fundamental presupposition of Wald's theory that the merits or demerits of a decision rule \( d \) depend entirely on its risk function \( R(\theta, d) \). For a finite parameter space \( \Theta = \{\theta_1, \ldots, \theta_n\} \) the risk function can in its turn be identified with a point in \( R^n \) whose Cartesian coordinates are \( R(\theta_i, d) \) for \( i = 1, \ldots, n \).

The Lindley–Savage argument accepts this mise en scène and proceeds to consider the properties of a preference relationship between decision rules (or
equivalently, between risk functions). I shall write \( x \succ y \) for '\( x \) is preferred to \( y \)', and define the *indifference* relation:

\[
x \sim y = \text{df} x \succ y \land y \succ x
\]

I shall assume that '>' is asymmetrical and transitive. Hence indifference is reflexive.

The conclusion that the argument seeks to establish is that there is a fixed vector \( \pi = (\pi_1, \ldots, \pi_n) \) with \( \pi_i > 0 \) for each \( i \), \( \Sigma \pi_i = 1 \), such that:

\[
y \succ x \iff \pi \cdot y < \pi \cdot x
\]

\( \pi \) is the geometric representation of a Bayesian prior distribution, and the inner product \( \pi \cdot x \) is the *Bayes risk* of the decision rule \( x \).

However, it will be convenient to develop the argument in rather different terms. I shall consider for an arbitrary decision rule \( x \) the sets:

\[
E(x) = \{ y : y \sim x \},
\]

\[
P(x) = \{ y : x \succ y \},
\]

\[
N(x) = \{ y : y \succ x \},
\]

Clearly these sets are pairwise disjoint, and their union is the whole of \( \mathbb{R}^n \). The Bayesian conclusion will now be that there is a fixed vector \( \pi \), as above, such that for any \( x \), \( E(x) \) is the hyperplane

\[
\pi \cdot y = \pi \cdot x,
\]

\( P(x) \) is the open half-space on the positive side of \( E(x) \), and \( N(x) \) is the open half-space on the negative side of \( E(x) \).

The basic assumption of the Lindley–Savage argument is

\[
(A1) \quad \text{Convexity of Indifference.} \quad \text{If} \ x \sim y \ \text{and} \ z \sim w, \ \text{then for any} \ \lambda, \ 0 \leq \lambda \leq 1, \ \lambda x + (1 - \lambda)z \sim \lambda y + (1 - \lambda)w.
\]

(In fact there are various ways in which this assumption can be weakened slightly.) We shall also assume that domination implies preference, i.e.

\[
(A0) \quad \text{Admissibility.} \quad \text{If} \ x = (x_1, \ldots, x_n) \ \text{and} \ y = (y_1, \ldots, y_n) \ \text{and we have} \ x_i \leq y_i \ \text{for all} \ i, \ \text{with strict inequality in at least one case, then} \ x \succ y.
\]