COHERENT ODDS AND BAYESIAN ODDS

by

John N. Quiring

Technical Report No. 179

August 1972

University of Minnesota
Minneapolis, Minnesota

ABSTRACT

Coherent Odds and Bayesian Odds

Let $X = \{x\}$ be a sample space, $\Theta = \{\theta\}$ be a parameter space, and $p(x, \theta)$ be a probability mass function in x for all θ. A statement of the form $P\{\theta \in C|x\} = \alpha$ is called a "probability assertion." Testing the validity of probability assertions is viewed as a game between two players, Peter and Paul, under the supervision of a master of ceremonies. The latter selects $\theta_o \in \Theta$, draws x randomly according to $p(x, \theta_o)$, and reveals x (but not θ_o) to Peter and Paul. Peter chooses $N = N(x)$ subsets $C_1(x), \ldots, C_N(x)$ of Θ and asserts the probability $\alpha_i(x)$ of θ_o belonging to $C_i(x)$ for each i. Paul tests the validity of the assertions by specifying a strategy s which determines bets placed on or against the events $\theta_o \in C_1(x), \ldots, \theta_o \in C_N(x)$ at odds determined by the probabilities $\alpha_i(x)$. After the bets are placed θ_o is disclosed and Paul receives a payoff based on the amount of his wager, the odds, and the truth or falsity of the events. $G_s^\alpha(\theta)$ denotes the expected payoff to Paul when Peter uses strategy α, Paul strategy s, and θ is true.

Cornfield (Biometrics 25 (1969) 617-658) and Freedman and Purves (Annals of Mathematical Statistics 40 (1969) 1177-1186) have considered problems of this kind when X and Θ are both finite and Peter must choose all subsets of Θ. Cornfield defines α to be coherent if there exists no s such that $G_s^\alpha(\theta) \geq 0$ for all θ and $G_s^\alpha(\theta) > 0$ for some θ. The main results of Cornfield and of Freedman and Purves show that α is coherent if and only if it is "π-Bayes" for some π; that is, the probabilities $\alpha_i(x)$ correspond to posterior probabilities for some prior distribution π.

The present thesis gives generalizations to countably infinite (and

continuous) spaces X and Θ. Here π can be either proper or improper, and it is shown that a slightly stricter definition of coherence is needed to obtain the following theorem: If X and Θ are countable and if for some $1 < p < \infty$, $p(x, \cdot) \in \ell^p$ for all x, and if α is coherent, then α is π-Bayes for some $\pi \in \ell^q$, $q = p/(p-1)$. Cases when $p = 1$ or ∞ are also considered. The proofs of these results use separation theorems of functional analysis.

If Peter is required to assert the probabilities of all possible subsets of Θ given x and if the class of strategies which Paul can use is suitably restricted, then it is shown that coherence again implies that α is π-Bayes. Moreover, α must agree with a countably additive set function on all possible subsets of Θ for each x.

Generalizations to continuous X and Θ spaces are given for the case when $N(x)$ is identically equal to one. These results are related to the theory of confidence intervals and to Buehler (Annals of Mathematical Statistics 30 (1959) 845-863) and Wallace (Annals of Mathematical Statistics 30 (1959) 864-876).

ACKNOWLEDGEMENTS

With deep gratitude I acknowledge the support of others.

For his never ending patience and encouragement I thank Professor R. J. Buehler under whose capable direction this thesis was written. He suggested the topic and provided numerous improvements to the dissertation during the many hours he spent in consultation. His help has been invaluable.

I am grateful to Professors D. C. Heath and W. D. Sudderth for their excellent mathematical advice, and enthusiasm they generated toward the subject matter.

I also thank the U.S. Bureau of Mines Twin Cities Research Center for the opportunity of part time employment and the National Science Foundation for their partial support under research grant GP-24183.

My thanks to Mr. Gerry DuChaine and Mrs. Linda Gardiner for typing and preparing the final draft. Without a doubt, Mr. DuChaine's typing abilities expedited the completion of this thesis.

TABLE OF CONTENTS

CHAPTER I

Introduction and Summary

1.1 Introduction.

Statistical techniques which specify the degree of uncertainty of an inference about a set or sets of parameter values can be divided into two major categories, Bayesian and non-Bayesian. Through the use of a prior the Bayesian assigns a posterior distribution over the parameter space Θ. Non-Bayesian procedures include confidence regions which look like a partial specification of a probability distribution. Other assignments not fitting a confidence interval scheme can be imagined.

Following Buehler (1959) we examine the implications of viewing inference as a game between two players. Peter makes inferences while Paul questions their validity by placing bets for which he receives a payoff proportional to the degree of uncertainty of the inference and its truth or falsity.

Cornfield (1969) also studied such a game for finite sample and parameter spaces in light of coherence, a property involving expected payoffs to Paul. Peter's assertions or inferences, as defined by Cornfield, are coherent if there does not exist any betting scheme such that the expected payoff to Paul is non-negative for all θ and positive for some θ. (In the body of the thesis we call this strict coherence.) If Peter's assertions are consistent with posterior probabilities corresponding to some (proper or improper) prior measure π, then they will be called Bayes with respect to (w.r.t.) π.

In the same year Freedman and Purves (1969) introduced a game for finite spaces phrased in terms of odds. Apart from difficulties posed by infinite odds, the Cornfield, Freedman and Purves papers show that Peter's assertions are coherent if and only if they are Bayes w.r.t. some π.

Basicly this thesis is an extension of this type of result to countably infinite or continuous spaces. One consideration for the infinite spaces is the use of improper priors; another is a suitable definition of coherence. As we shall see, a straightforward extension of coherence for finite to infinite spaces is not strong enough to imply a Bayesian solution.

1.2 Summary

The papers of Cornfield (1969) and Freedman and Purves (1969) require Peter to post either probabilities or odds for all proper subsets of the parameter space given an observed sample point. Often in statistical inference, for an observed sample point, assertions are required for a single subset of the parameter space. This subset usually varies with the observation as in confidence interval applications, but conceivably it could be fixed.

For finite spaces, Chapter II formulates a Peter-Paul game (in terms of odds) in which Peter, given a sample point, may be required to make probability assertions on fewer than all subsets, possibly only one. We have a slight extension of previous results by showing that, for finite odds, Peter is coherent if and only if for some π he is Bayes w.r.t. π. By giving a slightly different twist to conventions introduced by Freedman and Purves (1969), we extend this result to include cases where the odds may be infinite. From the development of Chapter

II we will see that games phrased in terms of probability assertions (as is done in later chapters) appear to be better suited for studying statistical inference techniques than their counterparts based on odds.

Having introduced a Peter-Paul game in finite spaces, we consider some mathematical results which will be needed when studying games for more general spaces. The basic concept used in Chapter II to show that coherence implies a Bayesian solution is the existence of a positive vector (all elements of the vector are greater than zero) orthogonal to a subspace containing no semi-positive vector. Chapter III provides analogous tools which will be applied in countably infinite models (Chapter IV) or continuous models (Chapter VI).

Conditions for the existence of an orthogonal vector in more general spaces are not as simple as those for the finite dimensional spaces. The development begins with a normed linear space which contains a linear manifold and a convex cone. The manifold and the cone are said to have an empty expanded intersection if there is a positive distance between the manifold and the cone after the cone has been translated by any member of itself. Under this assumption we show the existence of a continuous linear functional which separates the manifold and the cone. Later this is specialized to either ℓ^p (for countably infinite models) or L^p (for continuous models), where the cone is the set of all semi-positive vectors. The representation of continuous linear functionals in these spaces gives the existence of an orthogonal vector in ℓ^q (for ℓ^p spaces) or function in L^q (for L^p spaces) where $q = p/(p - 1)$; $p \in (1, \infty)$. Results are also stated for $p = 1$ or ∞.

Having this at our disposal Chapter IV introduces two models for countably infinite spaces phrased in terms of probability assertions. Model I deals with the case where Peter is required to make a finite number of assertions for each observed sample point while Model II handles a countable number. We modify the definition of coherence to mean that the space of all expected payoffs to Paul and the semi-positive vectors have an empty expanded intersection. If for some $p \in [1, \infty)$, $f(x, \theta) \in \ell^p$ as a vector in θ for all x (f is the probability mass function used in calculating the expected payoff function), then the modified version of coherence in Model I implies that Peter's probability assertions are Bayes w.r.t. π for some $\pi \in \ell^q$.

To study cases where $p = \infty$, we consider a Bayes' solution based on a finitely additive set function which plays the role of a prior and state an analogous result.

In Model II coherence implies not only a Bayesian solution but also that Peter's assignments must agree with a countably additive probability measure on the parameter space. This is true for all $p \in [1, \infty]$.

Now that we have mentioned that coherence implies a Bayes' solution, what can be said about the converse? It is not surprising that a Bayes' solution w.r.t. a proper $(\Sigma\, \pi(\theta) = 1)$ prior will be coherent. In Chapter V we present sufficient conditions which show, that in Model I, this is also true for some Bayesian solutions based on an improper prior in ℓ^q. However, many questions with regard to the converse remain open.

Chapter VI introduces a model for continuous spaces in which Peter is required to make only one probability assertion for an observed sample point. Statistical interpretation of this game has implications in confidence interval theory. Because of measure theoretic aspects, coherence is modified to imply that there is no almost everywhere semi-positive (i.e. no function which is almost everywhere non-negative and positive on a set of positive measure) expected payoff function available to Paul.

For densities satisfying very mild assumptions, we show that if Peter is coherent, then his probability assertions are Bayes w.r.t. some prior π in L^p for suitable p. Once again the choice of p depends on the density used to calculate the expected payoff. We briefly discuss the converse problem and move on to compare our model and its properties with those of Buehler (1959) and Wallace (1959). These papers are concerned with the conditional behavior of confidence intervals.

The basic results of this thesis are (i) the conditions under which there is a vector orthogonal to an infinite dimensional linear manifold, and (ii) the application of this to show that coherence requires Peter to be consistent with Bayes' rule in the various models that we consider. From the contrapositive of these results we see that if Peter is not Bayes from some π, then there is a betting strategy for Paul which yields an expected payoff which is never negative, and for some value or values of θ is positive.

CHAPTER II

Finite Discrete Parameter and Sample Spaces

2.1 Introduction.

The material in this chapter has its origin in the papers of Cornfield (1969) and Freedman and Purves (1969). Essentially, they consider a game between two players; the first one posts odds (Freedman and Purves) or probabilities (Cornfield) on a random event while the second player places stakes or bets on the event. The outcome of the event is determined and the second player receives a payoff which depends on the stakes, the odds or probabilities, and the outcome. We call such a game a Peter-Paul game; the characters are from Buehler (1959), who describes a similar model. Since the Peter-Paul game of this chapter is based on odds as in Freedman and Purves (1969) while the theorems are stated in terms of coherence, a property described by Cornfield (1969), we have a blend of these papers.

As in Cornfield (1969) and Freedman and Purves (1969), this chapter deals with finite sample and parameter spaces where the mathematics does not obscure the spirit of the Peter-Paul game. Hopefully, this will serve to introduce the reader to our notation and familiarize him with some concepts common to later chapters, even though we switch to games phrased in terms of probabilities. From the development we will see that the possibility of infinite odds introduces difficulties which appear to be more of a nuisance than of value. We will base our approach to handling infinite odds on a convention of Freedman and Purves (1969). Although we do not state

Peter-Paul games in terms of odds for more general spaces due to the difficulties encountered with infinite odds, our method of handling infinite odds in finite spaces should easily extend to more general spaces.

Section 2.2 describes the finite space Peter-Paul game based on odds. Section 2.3 studies games where the odds are finite and presents an aspect of the game called coherence (as defined by Cornfield). Theorem 2.2 of this section is similar to Theorem 1 of Freedman and Purves (1969) or Section 6 of Cornfield (1969). The conclusion of Theorem 2.2 holds even when Peter is required to post odds on only one event, and for this reason, it can be considered a slight extension of a result of Freedman and Purves (1969) and Cornfield (1969). Section 2.4 presents a parallel theorem when the odds may be infinite.

2.2 Description of a Peter-Paul Game for Finite Spaces.

Freedman and Purves (1969) suggest a model similar to the following Peter-Paul game. Let Θ and X be fixed finite sets and let $p(x, \theta)$ be a probability mass function in x for every θ. To avoid trivial details we assume that for each $x \in X$

$$p(x, \theta) > 0 \quad \text{for some} \quad \theta \in \Theta . \tag{2.1}$$

Let $\mathcal{G}$ be a collection of proper subsets of Θ which may or may not be closed under complementation. Consider three persons who participate in a game: a master of ceremonies, Peter,and Paul.

The master of ceremonies chooses and fixes the sets in $\mathcal{G}$. Then he selects $\theta \in \Theta$ and draws $x \in X$ according to the p.m.f. based on θ. The x is announced to Peter and Paul, but θ is retained by the master of ceremonies until the odds are posted and bets are placed.

Peter is required to post odds $\lambda(x, A)$ where $0 \leq \lambda(x, A) < \infty$ for all $A \in \mathcal{G}$. Given that, x was revealed, $\lambda(x, A)$ is the odds against the event that the master of ceremonies chose $\theta \in A$. We consider the set of odds which Peter specifies to be a strategy and denote it by λ. If the probability $P(\theta \in A|x) \neq 0$ is known and the odds are consistent with these probabilities, then $\lambda(x, A) = P(\theta \in \widetilde{A}|x)/P(\theta \in A|x)$. It is noted that if $\lambda(x, A) = (\lambda(x,\widetilde{A}))^{-1}$, then $\lambda(x, A)$ is consistent with the probability $P(\theta \in A|x) = (\lambda(x,A)+1)^{-1}$. In contrast to Freedman and Purves (1969) our assumptions do not include the relationship $\lambda(x, A) = (\lambda(x, A) + 1)^{-1}$ and, in fact, this equality is a consequence of the Peter-Paul game and coherence.

Paul's role in the game is to test the validity of Peter's odds. Hence, with the knowledge of x and Peter's odds, Paul places bets or stakes $s(x, A)$ for all $A \in \mathcal{G}$. The $s(x, A)$ are stakes on the event $\theta \in A$ given x and are assumed to be finite. We denote Paul's strategy by s. After the odds and stakes are determined, the master of ceremonies reveals θ so that Peter and Paul can settle up. For each $A \in \mathcal{G}$, Paul gains

$$(2.2) \qquad \begin{array}{ll} \lambda(x, A)s(x, A) & \text{if } \theta \in A \\ -\, s(x, A) & \text{if } \theta \in \widetilde{A}\,. \end{array}$$

For the chosen x, the payoff to Paul is the sum over $A \in \mathcal{G}$. Since the values in (2.2) are finite, the payoff is well defined.

Note that this game allows Paul to use both negative and positive stakes, which in effect, permits him to bet on or against any A. If $\mathcal{G}$ is closed under complementation and if $\lambda(x, A) = (\lambda(x, A) + 1)^{-1}$, then for any strategy s containing negative stakes there is a strategy s' consisting of non-negative stakes with the same payoff to Paul. Since our assumptions do not include that $\mathcal{A}$ is closed under complements and the equality of $\lambda(x, A)$ and $(\lambda(x, A) + 1)^{-1}$, it is possible that there does not exist a strategy s' containing only non-negative stakes which has a payoff identical to a strategy s having negative stakes. Thus, this Peter-Paul game differs from one having only betting strategies which contain non-negative bets. Later we will see the mathematical convenience behind allowing non-negative stakes. Next we define an expected payoff to Paul.

Definition 2.1.

$G_s^\lambda(\theta)$ will denote the expected payoff to Paul when Peter's strategy is λ and Paul's is s. It is computed as

$$G_s^\lambda(\theta) = \sum_{A\in\mathcal{A}} \sum_{x\in X} \{[\lambda(x, A)+1]I_A(\theta)-1\}s(x, A)p(x, \theta). \tag{2.3}$$

Since the right hand side of (2.3) is a finite sum of terms which are finite, $G_s^\lambda(\theta)$ is well defined.

2.3 Finite X and Θ Spaces with Finite Odds.

Now let us focus our attention on the behavior of $G_s^\lambda(\theta)$ as a function in θ for a fixed strategy λ as Paul varies s. To accommodate this, we present a series of definitions found in Gale (1960, p. 43). Let $\underset{\sim}{\nu}$ be a vector in R^m or R^∞; that is, $\underset{\sim}{\nu} = (\nu_1, \nu_2, \ldots, \nu_m)$ or $\underset{\sim}{\nu} = (\nu_1, \nu_2, \ldots)$ where ν_i are in $(-\infty, \infty)$.

Definition 2.2.

$\underset{\sim}{\nu}$ is non-negative (written $\underset{\sim}{\nu} \geqq \underset{\sim}{0}$) if and only if $\nu_i \geq 0$ for all i.

$\underset{\sim}{\nu}$ is positive (written $\underset{\sim}{\nu} > \underset{\sim}{0}$) if and only if $\nu_i > 0$ for all i.

$\underset{\sim}{\nu}$ is <u>semi-positive</u> (written $\underset{\sim}{\nu} \geq \underset{\sim}{0}$) if and only if $\underset{\sim}{\nu} \geq \underset{\sim}{0}$ and $\underset{\sim}{\nu} \neq \underset{\sim}{0}$.

Cornfield (1969) studied a property of $G_s^\lambda(\theta)$ called coherence; our version is essentially the same.

<u>Definition 2.3</u>.

When Θ is a finite space, Peter's strategy λ is called <u>strictly</u> [weakly] <u>coherent</u> if there does not exist a strategy s such that $G_s^\lambda(\underset{\sim}{\theta}) \geq \underset{\sim}{0}$ [$G_s^\lambda(\underset{\sim}{\theta}) > \underset{\sim}{0}$].

The following is an interpretation of strict coherence. Assume that λ is strictly coherent. If for any strategy s there exists a θ in Θ such that $G_s^\lambda(\theta) > 0$, then there is a θ' in Θ such that $G_s^\lambda(\theta') < 0$.

Cornfield (1969) and Freedman and Purves (1969) examine the case when X and Θ are finite, and $\mathcal{G}$ is either the collection of all subsets or all proper subsets of Θ. In terms of coherence their results roughly state that Peter's λ is strictly coherent if and only if λ is determined by a Bayes' procedure. This chapter extends the theory to cases where $\mathcal{G}$ may contain fewer than all the proper subsets of Θ, possibly only one. In this situation the Peter-Paul game can be given a hypothesis testing interpretation. The development of this theory uses a result commonly found in linear programming, which is presented for the sake of completeness.

<u>Theorem 2.1</u> (Gale (1960), p. 48-49) .

Let D be a m by n matrix; $\underset{\sim}{w} \in R^m$ and $\underset{\sim}{s} \in R^n$. Exactly one of the following alternatives holds. Either there exists $\underset{\sim}{w} > \underset{\sim}{0}$ [$\underset{\sim}{w} \geq \underset{\sim}{0}$] such that

(2.4) $\underset{\sim}{w}D = \underset{\sim}{0}$

or there exists $\underset{\sim}{s} \in R^n$ such that

(2.5) $D\underset{\sim}{s} \geq \underset{\sim}{0}$ $[D\underset{\sim}{s} > \underset{\sim}{0}]$.

We interpret a portion of Theorem 2.1 as follows. Let $\mathcal{H}$ be the linear space spanned by the column vectors of D, and let $\mathcal{P}$ be the set of all semi-positive vectors in R^m. If there is no vector in $\mathcal{H}$ and in $\mathcal{P}$, then there exists a vector ν such that $\underset{\sim}{\nu} > \underset{\sim}{0}$ and $\underset{\sim}{\nu}$ is orthogonal to $\mathcal{H}$. In finite dimensional spaces this is quite clear and the proof is easy. A picture for R^2 is shown in Figure 1. Application of this theorem (and its analogues for ℓ^p

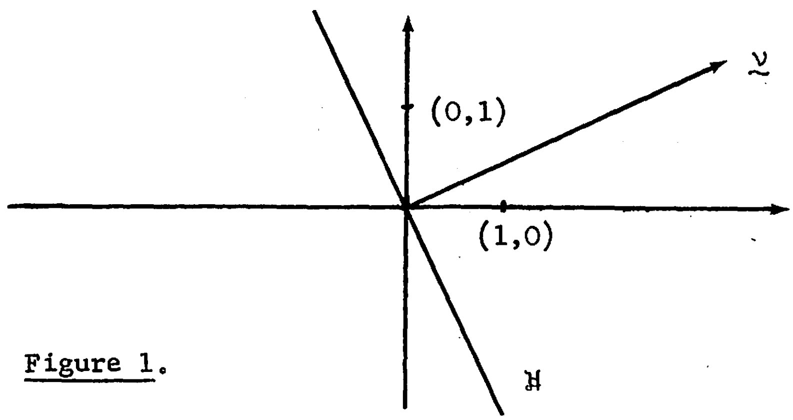

Figure 1.

and L^p spaces developed in Chapter III) is the key in solving a necessary condition for coherence. We pause for one more definition before applying Theorem 2.1 to the Peter-Paul game.

Definition 2.4.

Peter's λ is proper Bayes with respect to (w.r.t.) π if and only if there exists a function π such that $\pi: \Theta \to [0, M)$, where $M < \infty$, and for all $x \in X$ and $A \in \mathcal{A}$

(2.6) $$\lambda(x, A) = \Big(\sum_{\theta \in \tilde{A}} p(x, \theta)\pi(\theta)\Big)/\Big(\sum_{\theta \in A} p(x, \theta)\pi(\theta)\Big).$$

For notational simplicity, suppose $\{x_i\}_{i=1}^{m}$, $\{\theta_j\}_{j=1}^{p}$, and $\{A_k\}_{k=1}^{n}$ are enumerations of X, Θ, and $\mathcal{A}$ respectively. To shorten the notation let

$$
\begin{aligned}
(2.7)\qquad & p(x_i, \theta_j) = p(i, \theta_j) = p(i, j) \\
& \pi(\theta_j) = \pi_j \\
& \lambda(x_i, A_k) = \lambda(i, A_k) = \lambda(i, k) \\
& s(x_i, A_k) = S(i, A_k) = S(i, k).
\end{aligned}
$$

Theorem 2.2.

Assume X, Θ, and $\mathcal{A}$ are finite where $\mathcal{A}$ contains n subsets of Θ, and assume all the odds in Peter's λ are finite. Then λ is strictly [weakly] coherent if and only if λ is proper Bayes w.r.t. $\underset{\sim}{\pi} > \underset{\sim}{0}$ [$\underset{\sim}{\pi} \geq \underset{\sim}{0}$].

Proof:

Let D be a p by $m \cdot n$ matrix with column vectors $d_{ik}(\underset{\sim}{\theta})$, for $i = 1,\ldots, m$; $k = 1,\ldots, n$, where the elements of $d_{ik}(\underset{\sim}{\theta})$ are

$$(2.8)\qquad d_{ik}(\theta) = [(\lambda(i, k) + 1)I_{A_k}(\theta) - 1]p(i, \theta).$$

Any strategy s can be represented as a $n \cdot m$ by 1 vector $\underset{\sim}{s}$ having elements $s(i, k)$. Then the payoff to Paul can be written

$$(2.9)\qquad G_s^{\lambda}(\underset{\sim}{\theta}) = D_{p\times n\cdot m}\, s_{n\cdot m\times 1}\ .$$

Assume λ is strictly coherent. Then there does not exist any s such that $\underset{\sim}{D}\underset{\sim}{s} \geq \underset{\sim}{0}$. By Theorem 2.1 there exists $\underset{\sim}{\pi} > \underset{\sim}{0}$ such that $\underset{\sim}{\pi}'D = \underset{\sim}{0}$. This implies

$$(2.10)\qquad \sum_{j=1}^{p} d_{ik}(j)\pi_j = 0 \ \text{ for all } i \text{ and } k,$$

which implies that λ is proper Bayes by rewriting (2.10) as (2.6).

Assume λ is proper Bayes w.r.t. $\underset{\sim}{\pi} > \underset{\sim}{0}$. From (2.6) a simple computation gives $\underset{\sim}{\pi}' d_{ik}(\underset{\sim}{\theta}) = 0$ for all i and k; that is, $\underset{\sim}{\pi}' D = \underset{\sim}{0}$. By Theorem 2.1 there does not exist any $\underset{\sim}{s}$ such that $\underset{\sim}{Ds} \geq \underset{\sim}{0}$. This implies λ is strictly coherent. A similar proof holds for weakly coherent. □

Theorem 2.2 proves that strict coherence implies λ is proper Bayes w.r.t. $\underset{\approx}{\pi} > \underset{\sim}{0}$. From equation (2.6) we have $\lambda(x, A) = (\lambda(x, \widetilde{A}))^{-1}$ and thus $\lambda(x, A)$ is consistent with the probability $P(\theta \in A|x) = (\lambda(x, A) + 1)^{-1}$. If $\mathfrak{A}$ contains all the proper subsets of Θ, then an argument of Freedman and Purves (1969) shows that the prior π is unique. If $\mathfrak{A}$ contains fewer than all proper subsets, however, then π may not be uniquely determined. This occurs when the column space of D has dimension less than $p - 1$. Here is an example.

Example 2.1.

Let $\Theta = \{1, 2, 3, 4, 5, 6\}$, $A = \{1, 2\}$ and $X = \{1, 2\}$.

Values of $p(x, \theta)$

$p(x, \theta)$	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
x_1	.8	.2	.6	.4	.9	.1
x_2	.2	.8	.4	.6	.1	.9

Let $\pi' = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$. Then the Bayes' odds w.r.t. π are

$$\lambda(1, A) = 2 \qquad \lambda(1, \widetilde{A}) = \tfrac{1}{2}$$
$$\lambda(2, A) = 2 \qquad \lambda(1, \widetilde{A}) = \tfrac{1}{2}\ ,$$

and these odds are strictly coherent by Theorem 2.2. Now the Bayes' odds w.r.t. $\pi' = (1/6, 1/6, 1/9, 1/9, 2/9, 2/9)$ are equal to odds w.r.t. π, and thus, π is not uniquely defined. In fact, any prior of the form $(\pi_1, \pi_1, \pi_2, \pi_2, \pi_3, \pi_3)$ where

$$\frac{\pi_2 + \pi_3}{\pi_1} = 2 \quad \text{and} \quad \pi_1 + \pi_2 + \pi_3 = 1/2$$

will produce the above odds.

2.4 Finite X and Θ Spaces With Infinite Odds.

In this section Theorem 2.2 is strengthened to include infinite odds. To avoid infinite values in the payoff function, we use a convention from Freedman and Purves (1969). If for any $x \in X$ and $A \in \mathcal{A}$ $\lambda(x, A) = \infty$, then Paul can ask for any $M(x, A)$ where $0 \leq M(x, A) < \infty$ provided $\theta \in A$ and $s(x, A) > 0$. That is, if $\lambda(x, A) = \infty$, then Paul gains

$$\begin{array}{lll} M(x, A) & \text{if} & s(x, A) > 0 \text{ and } \theta \in A \\ 0 & \text{if} & s(x, A) = 0 \text{ and } \theta \in A \\ -s(x, A) & \text{if} & \theta \in \tilde{A} \end{array} \tag{2.11}$$

where $M(x, A)$ is predetermined by Paul as part of his strategy s. We will also restrict Paul to non-negative stakes whenever $\lambda(x, A) = \infty$.

Definition 2.5.

Paul's stakes s satisfies Restriction P if $s(x, A) \geq 0$ whenever $\lambda(x, A) = \infty$.

A justification for assuming Restriction P is the following. If the event $\theta \in A$ given x is impossible and $\theta \in \tilde{A}$ given x is sure, then the odds $\lambda(x, A) = \infty$ and $\lambda(x, \tilde{A}) = 0$ are consistent with probabilities $P(\theta \in A|x) = 0$ and $P(\theta \in \tilde{A}|x) = 1$. Any amount that Paul bets on A he is sure to lose. Thus, if Paul is allowed to bet a negative stake on A, he guarantees a positive expected payoff. This would not be fair to Peter!

Our pursuit of a result similar to Theorem 2.2 for infinite odds can take two courses. One of these modifies the definition of strict coherence to include the closure of the space of all expected payoffs as Paul varies s against a fixed λ. That is, alter strict coherence to mean that the closure of the space of payoffs contains no semi-positive vector. The new definition would not affect the result of Theorem 2.2 since the space of payoffs under the assumptions is a hyperplane in R^p. Since hyperplanes are closed, Theorem 2.2 would not change if we modified the definition of strict coherence.

An alternate approach is to add the restriction that $\mathcal{A}$ is closed under complementation. After three easy lemmas, an extended result may be stated and proved. Since one of the purposes of this chapter is to acquaint the reader with the spirit of the material without complexity, we choose to present this method. Thus, for the remainder of this chapter we assume $\mathcal{A}$ is closed under complementation.

Although the convention of Freedman and Purves (1969) avoids infinite terms in the formation of the expected payoff (2.3), it destroys the linearity of the space of expected payoffs. The net effect is that $G_s^\lambda(\underset{\sim}{\theta})$ cannot be expressed as in (2.9). A procedure for skirting this problem first shows, that if λ is strictly coherent and $\lambda(x_0, \widetilde{A}) = 0$, then $p(x_0, \theta) = 0$ for all $\theta \in A$. Next it shows, if $\lambda(x_0, A) = \infty$, then $p(x_0, \theta) = 0$ for all $\theta \in A$. Finally, the expected payoff is broken into two parts; one part comes from the finite odds and the other from the infinite odds. The infinite part is shown to be identically zero or semi-negative and combined with Theorem 2.2 to give the desired result.

Lemma 2.1.

Assume λ is strictly coherent and s satisfies Restriction P. If $\lambda(x_0, \tilde{A}) = 0$, then

$$p(x_0, \theta) = 0 \quad \text{for all} \quad \theta \in A. \tag{2.12}$$

Proof:

Let Paul specify s in which all the stakes are zero except

$$s(x_0, \tilde{A}) = -1. \tag{2.13}$$

Then the expected payoff to Paul is

$$G_s^\lambda(\theta) = \begin{cases} p(x_0, \theta) & \theta \in A \\ 0 & \theta \in \tilde{A}. \end{cases} \tag{2.14}$$

Thus, if $p(x_0, \theta) > 0$ for some $\theta \in A$ then $G_s^\lambda(\theta)$ is semi-positive, and the conclusion follows by contraposition. □

Lemma 2.2.

Assume λ is strictly coherent and s satisfies Restriction P. If $\lambda(x_0, A) = \infty$, then $\lambda(x_0, \tilde{A}) = 0$.

Proof:

Assume $\lambda(x_0, \tilde{A}) > 0$. Let Paul specify s with all elements zero except

$$\begin{aligned} s(x_0, A) &= \tfrac{1}{2} \,;\; M(x_0, A) = (\lambda(x_0, \tilde{A}))^{-1} + 1 \\ s(x_0, \tilde{A}) &= (\lambda(x_0, \tilde{A}))^{-1}. \end{aligned} \tag{2.15}$$

A simple calculation shows

$$(2.16)\qquad G_s^\lambda(\theta) = \begin{cases} p(x_0, \theta) & \theta \in A \\ \frac{1}{2}\, p(x_0, \theta) & \theta \in \tilde{A}. \end{cases}$$

Since $p(x_0, \theta) > 0$ for some $\theta \in \Theta$, $G_s^\lambda(\theta)$ is semi-positive which is a contradiction to strict coherence. Thus $\lambda(x_0, A) = 0$. □

<u>Lemma 2.3</u>.

Assume λ is strictly coherent and s satisfies Restriction P. If $\lambda(x_0, A) = \infty$, then

$$(2.17)\qquad \begin{aligned} &p(x_0, \theta) = 0 && \text{for all } \theta \in A \\ &p(x_0, \theta) > 0 && \text{for some } \theta \in \tilde{A}. \end{aligned}$$

The proof follows directly from Lemmas 2.1 and 2.2.

With these lemmas established we state and prove the necessary conditions for strict coherence. The technique in the proof breaks up the expected payoff into components which depend on finite and infinite odds.

<u>Theorem 2.3</u>.

If λ is strictly coherent and s satisfies Restriction P, then λ is proper Bayes w.r.t. $\underset{\sim}{\pi} > \underset{\sim}{0}$.

<u>Proof</u>:

As in the proof of Theorem 2.2 we associate a matrix D containing column vectors formed from all the odds which are finite. Corresponding to infinite odds $\lambda(x_0, A_0) = \infty$, let $\nu(x_0, A_0, \underset{\sim}{\theta})$ be a vector with elements

$$(2.18)\qquad \begin{aligned} &M(x_0, A_0)p(x_0, \theta) && \text{if } \theta \in A_0 \text{ and } s(x_0, A_0) > 0 \\ &0 && \text{if } \theta \in A_0 \text{ and } s(x_0, A_0) = 0 \\ &-s(x_0, A_0)p(x_0, \theta) && \text{if } \theta \in \tilde{A}_0. \end{aligned}$$

Then the expected payoff can be written as

$$(2.19) \qquad G_s^\lambda(\underset{\sim}{\theta}) = D\underset{\sim}{s} + \Sigma\, \nu(x_0, A_0, \underset{\sim}{\theta})$$

where the sum is taken over all x_0 and A_0 such that $\lambda(x_0, A_0) = \infty$. Since s satisfies Restriction P, by Lemma 2.3 and (2.17) we see that $\nu(x_0, A_0, \theta) \leq 0$ for all $\theta \in \Theta$. Thus, Paul suffers needless loss unless $s(x_0, A_0) = 0$ whenever $\lambda(x_0, A_0) = \infty$. If $s(x_0, A_0) = 0$, we have $\nu(x_0, A_0, \underset{\sim}{\theta}) = \underset{\sim}{0}$, and strict coherence implies that $D\underset{\sim}{s}$ cannot be semi-positive for any s. As in the proof of Theorem 2.2, there exists $\underset{\sim}{\pi} > \underset{\sim}{0}$ such that all finite odds are proper Bayes w.r.t. π. If we adopt the usual definition $R/0 = \infty$ where $R \in (0, \infty)$, then by Lemma 2.3 and (2.6) the infinite odds are proper Bayes w.r.t. π. □

<u>Theorem 2.4.</u>

If λ is proper Bayes w.r.t. $\underset{\sim}{\pi} > \underset{\sim}{0}$ and s satisfies Restriction P then

$$(2.20) \qquad \sum_{\theta} G_s^\lambda(\theta)\pi(\theta) \leq 0 \quad \text{for all strategies } s,$$

and thus, λ is strictly coherent.

<u>Proof:</u>

As in the proof of Theorem 2.3 we have $\underset{\sim}{\pi}'D = \underset{\sim}{0}$, where the columns of D are formed from finite odds in λ. If $\lambda(x_0, A_0) = \infty$ then (2.17) holds, and an easy computation shows $\Sigma\, \pi(\theta)\cdot\nu(x_0, A_0, \theta) \leq 0$. Since

$$(2.21) \qquad \sum_{\Theta} G_s^\lambda(\theta)\pi(\theta) = \pi' D\underset{\sim}{s} + \Sigma' \sum_{\Theta} \pi(\theta)\cdot\nu(x_0, A_0, \theta)$$

where Σ' denotes a sum over x_0 and A_0 such that $\lambda(x_0, A_0) = \infty$, (2.20) follows. The remainder of the proof is clear. □

This completes the extension of Theorem 2.2 as promised at the beginning of this section. When the results are phrased in terms of coherence, one can see that the presence of infinite odds only introduces mechanical difficulties that are a nuisance. On the other hand, if one is interested in characterizing the space of payoffs available to Paul, then the presence of infinite odds is relevant. That is, under assumptions of coherence and finite odds, the space of payoffs available to Paul is a hyperplane in R^p. With infinite odds, the space of payoffs becomes a half space determined by a hyperplane which is no longer a linear space. Thus, the main mathematical inconvenience in dealing with infinite odds is the loss of the linear space of payoff functions. This concludes our treatment of finite sample and parameter space Peter-Paul games phased in terms of odds.

CHAPTER III

Mathematical Background

3.1 Introduction.

Before proceeding to more complex sample and parameter spaces, we pause to present some mathematical theory which will facilitate the study of more general Peter-Paul games. The type of theory needed for countably infinite spaces is the existence of a positive vector which is orthogonal to a subspace containing no semi-positive vector. This represents an extension of Theorem 2.1, which applies to finite dimensional vector spaces.

For normed linear spaces, such as ℓ^p, one approach in generalizing Theorem 2.1 might be to apply a separating hyperplane theorem such as Theorem 8 of Dunford and Schwartz (1958, p. 417). That is, if P is the set of all semi-positive vectors and M is a linear manifold such that $M \cap P = \varphi$, then a separating hyperplane theorem would show the existence of a non-zero continuous linear functional F such that $F(M) = 0$ and $F(P) > 0$. Then we could utilize a representation theorem of continuous linear functionals to get the desired results. There is a snag in the methodology, however, in that the hypothesis of separating hyperplane theorems require P to have a non-empty interior. Unfortunately P has an empty interior, and as a result, these theorems do not apply.

The theory of this chapter solves the problem by using the fact that P is not only a convex set but a convex cone. We also assume that the distance between certain sequences of vectors in P and M does not go to zero (an exact statement of this is Definition 3.1). With these ideas we develop the desired result.

Section 3.2 focuses attention on a real normed linear space, say $\mathcal{Y}$. The principal result of this section is the existence of a continuous linear functional separating disjoint sets M and N, where M is a manifold in $\mathcal{Y}$ and N is a convex cone in $\mathcal{Y}$. Sections 3.3 and 3.4 specialize this to the normed linear spaces ℓ^p and L^p where N is the set of semi-positive vectors or functions. The purpose of this development is to show that for linear manifolds which do not contain any semi-positive vector or function, there exists an orthogonal positive vector or function. Those who are not interested in the proof and development of such a result may find a formal statement of it in Theorem 3.5 (for ℓ^p) or Theorem 3.7 (for L^p), and skip to the next chaper where we continue with Peter-Paul games.

3.2 <u>Separation Theorems for Real Normed Linear Spaces</u>.

Let $\mathcal{Y}$ be a real normed linear space with norm given by $\|y\|$ for $y \in \mathcal{Y}$. A subset $Y \subset \mathcal{Y}$ is said to be a convex cone if Y is convex and $\alpha Y \subset Y$ for all real $\alpha > 0$. Suppose N is a non-empty convex cone in $\mathcal{Y}$, and M is a linear manifold in $\mathcal{Y}$. By applying the norm $\|\cdot\|$ we could state assumptions in terms of the closure of M or N. For instance, we could assume that N and the closure of M are disjoint (note M and the closure of N are disjoint only in trivial cases). However, in our study of convex cones and linear manifolds, we will need a slightly stronger assumption which is expressed in the next definition.

<u>Definition 3.1</u>.

A non-empty convex cone N and a manifold M are said to have an <u>empty expanded intersection</u> [though n_0] if for all $n_0 \in N$ [if for $n_0 \in N$]

$$(3.1) \qquad \inf_{m \in M} \inf_{n \in N} \|n_0 + n - m\| > 0 .$$

Equivalently $d(N + n_0, M) > 0$ where d denotes distance.

Notice that if N and M have an empty expanded intersection, then the intersection of N and the closure of M is empty. This can be seen by taking the infimum over n in the inequality $\|n_0 + n - m\| \leq \|n_0 - m\| + \|n\|$. An empty expanded intersection also implies that $\{0\}$ is not a member of N. Let $K = M - N$ and define a function g on $\mathcal{Y}$ by

$$(3.2) \qquad g(y) = \inf_{k \in K} \|k - y\| \quad \text{for all} \quad y \in \mathcal{Y}.$$

Lemma 3.1 shows g to be a sublinear functional on $\mathcal{Y}$.

Lemma 3.1.

Assume M is a linear manifold in $\mathcal{Y}$, and N is a non-empty convex cone in $\mathcal{Y}$ such that M and N have an empty expanded intersection through n_0. If $g(x)$ is defined by (3.2) where $K = M - N$, then $g(x)$ has the following properties.

$$(3.3) \qquad g(y) \geq 0 \qquad y \in \mathcal{Y}$$

$$(3.4) \qquad g(k) = 0 \qquad \text{for all} \quad k \in K$$

$$(3.5) \qquad g(0) = 0$$

$$(3.6) \qquad g(y) + g(y') \geq g(y+y') \quad \text{for} \quad y, y' \in \mathcal{Y}$$

$$(3.7) \qquad g(\alpha y) = \alpha g(y) \quad \text{for} \quad \alpha \geq 0, \ y \in \mathcal{Y}$$

$$(3.8a) \qquad g(n) \geq 0 \qquad \text{for all} \quad n \in N$$

$$(3.8b) \qquad g(n_0) > 0.$$

Proof (3.5).

Assume $g(0) = \varepsilon > 0$ and choose $n \in N$ such that $\|n\| < \varepsilon/2$. However, $-n$ is in K and

$$(3.9) \qquad \|0 + n\| < \varepsilon/2 .$$

Now, (3.9) contradicts $g(0) = \varepsilon$. Hence $g(0) = 0$.

Proof (3.6).

Case 1. $y \in K$, $y' \in K$.

Since y and y' are in K,

$$(3.10) \qquad y = m - n \quad \text{where} \quad m \in M,\ n \in N$$

and

$$(3.11) \qquad y' = m' - n' \quad \text{where} \quad m' \in M,\ n' \in N.$$

This implies

$$(3.12) \qquad y + y' = m + m' - n - n' = m'' - n'' \quad \text{where} \quad m'' \in M,\ n'' \in N.$$

Hence, since $y + y' \in K$ by (3.4) we have

$$(3.13) \qquad g(y) + g(y') = 0 = g(y + y').$$

Case 2. $y \notin K$, $y' \in K$ and $y + y' \notin K$.

Since $y' + k \in K$ when $k \in K$, we have

$$(3.14) \qquad g(y + y') = \inf_{k \in K} \|y + y' - k\|$$

$$(3.15) \qquad \leq \inf_{\substack{k' = y' + k \\ k \in K}} \|y + y' - k'\|$$

$$(3.16) \qquad = \inf_{k \in K} \|y - k\|$$

$$(3.17) \qquad = g(y).$$

Since $g(y') = 0$, (3.4) and (3.14) through (3.17) imply (3.6).

Case 3. $y \notin K$, $y' \notin K$ and $y + y' \notin K$.

For each $k' \in K$ we have

$$(3.18) \qquad \inf_{k \in K} \|y + y' - k\| \leq \inf_{\substack{k''=k'+k \\ k \in K}} \|y + y' - k''\|$$

$$(3.19) \qquad \leq \|y - k'\| + \inf_{k \in K} \|y' - k\|$$

which implies

$$(3.20) \qquad g(y + y') \leq \|y - k'\| + g(y').$$

Taking the inf over $k' \in K$ in (3.20) gives (3.6). The remaining cases are trivial since $y + y' \in K$. This completes the proof of (3.6).

Proof (3.7).

Case 1. $y = 0$ or $\alpha = 0$.

In this case we have $\alpha y = 0$. By (3.5) we have

$$(3.21) \qquad \alpha g(y) = g(\alpha y) = 0.$$

Case 2. $y \in K$ $\alpha > 0$.

Since $\alpha y \in K$, (3.4) implies (3.21).

Case 3. $y \notin K$ $\alpha > 0$.

Since $\alpha y \notin K$ we have

$$(3.22) \qquad g(\alpha y) = \inf_{k \in K} \|\alpha y - k\| = \inf_{k \in K} \|\alpha y - \alpha k\|$$

$$(3.23) \qquad = \alpha \inf_{k \in K} \|y - k\| = \alpha g(y).$$

This completes the proof of (3.7).

Proof (3.8b).

For every $k \in K$ we have

$$(3.24) \qquad \|n_0 - k\| = \|n_0 - (m-n)\| = \|n_0 + n - m\|$$

for some $m \in M$ and $n \in N$. Taking the infimum over $n \in N$ gives

$$\|n_0 - k\| \geq \inf_{n \in N} \|n_0 + n - m\|, \tag{3.25}$$

and taking the infimum over $m \in M$ in (3.25) further implies

$$\|n_0 - k\| \geq \inf_{m \in M} \inf_{n \in N} \|n_0 + n - m\|. \tag{3.26}$$

Since (3.26) is true for every $k \in K$, we have

$$\inf_{k \in K} \|n_0 - k\| \geq \inf_{m \in M} \inf_{n \in N} \|n_0 - m + n\|, \tag{3.27}$$

and the right hand side of (3.27) is greater than zero since the expanded intersection of M and N through n_0 is empty. □

For the sake of completeness we present Theorem 3.1 and Lemma 3.2 from Dunford and Schwartz (1958). Theorem 3.1 is commonly referred to as the Hahn-Banach Theorem.

Theorem 3.1. (Dunford and Schwartz, p. 62).

Let the real function p on the real linear space Y satisfy

$$p(x + y) \leq p(x) + p(y),\ p(\alpha y) = \alpha p(y) \quad \text{for} \quad \alpha \geq 0,\ x,\ y \in Y. \tag{3.28}$$

Let f be a real linear functional on a subspace S of Y with

$$f(y) \leq p(y) \qquad y \in S. \tag{3.29}$$

Then there is a real linear functional F on Y for which

$$F(y) = f(y) \quad \text{for} \quad y \in S;\ F(y) \leq p(y) \quad \text{for} \quad y \in Y. \tag{3.30}$$

Definition 3.2. (Ibid., p. 411).

If Y is a vector space, and S and T are subsets of Y, a functional f on Y is said to separate S and T if there exists a real constant c with $f(S) \geq c$ and $f(T) \leq c$ where $f(S) = \{x : x = f(s) \text{ for some } s \in S\}$.

Lemma 3.2. (Ibid., p. 412).

The linear functional f separates the subsets S and T of Y if and only if it separates the subsets $S - T$ and $\{0\}$ of Y.

Lemmas 3.1 and 3.2 will be combined with the Hahn-Banach Theorem to prove the existence of a linear functional which separates the previously defined subsets M and N. Theorem 3.2 is similar to separating theorems such as Theorem 8 of Dunford and Schwartz (1958) which involves the separation of two disjoint convex sets--one of which has a non-empty interior. Theorem 3.2, however, does not require that either of the sets have a non-empty interior, but rather that a non-empty convex cone has an empty expanded intersection with a linear manifold.

Theorem 3.2. (Basic Separation Theorem).

Assume M is a linear manifold contained in $\mathcal{Y}$, and N is a non-empty convex cone in $\mathcal{Y}$ such that the expanded intersection of M and N through n_0 is empty. Then there exists a non-zero linear functional F on $\mathcal{Y}$ such that

$$F(n_0) > 0 \tag{3.31}$$

$$F(n) \geq 0 \quad \text{for all } n \in N \tag{3.32}$$

$$F(m) = 0 \quad \text{for all } m \in M. \tag{3.33}$$

Proof:

Let

$$f_0(an_0) = \begin{cases} g(an_0) & \text{if } a \geq 0 \\ -g(-an_0) & \text{if } a < 0 \end{cases} \tag{3.34}$$

where g is defined in (3.2).

Let H be the subspace consisting of real multiples of n_0. Then f_0 is defined on H, and we will show that f_0 is a linear functional on H.

Case 1. $n = an_0$ where $a \geq 0$.

If $\alpha \geq 0$ then

$$(3.35) \qquad f_0(\alpha n) = f_0(\alpha a n_0) = g(\alpha a n_0) = \alpha g(an_0) = \alpha f_0(an_0) = \alpha f_0(n).$$

If $\alpha < 0$ then

$$(3.36) \qquad f_0(\alpha n) = f_0(\alpha a n_0) = -g(-\alpha a n_0) = \alpha g(an_0) = \alpha f_0(an_0) = \alpha f_0(n).$$

Case 2. $n = an_0$ where $a < 0$.

If $\alpha \geq 0$ then

$$(3.37) \qquad f_0(\alpha n) = f_0(\alpha a n_0) = -g(-\alpha a n_0) = -\alpha g(-an_0) = \alpha f_0(an_0) = \alpha f_0(n).$$

If $\alpha < 0$ then

$$(3.38) \qquad f_0(\alpha n) = f_0(\alpha a n_0) = g(\alpha a n_0) = -\alpha g(-an_0) = \alpha f_0(an_0) = \alpha f_0(n).$$

Thus $f_0(\alpha n) = \alpha f_0(n)$ for all real α and $n \in H$.

Let $n = an_0$ and $n' = a'n_0$. Then

$$(3.39) \qquad n + n' = (a + a')n_0$$

and

$$(3.40) \qquad f_0(n) + f_0(n') = f_0(an_0) + f_0(a'n_0) = af_0(n_0) + a'f_0(n_0)$$
$$= (a + a')f_0(n_0) = f_0(n + n').$$

Thus f_0 is a linear functional defined on H which is a one-dimensional subspace of $\mathfrak{y}$ consisting of real multiples of n_0. Moreover, $f_0(n) \leq g(n)$ for all $n \in H$. Thus the conditions of Theorem 3.1 are satisfied; hence, there exists a real linear functional

F on $\mathcal{Y}$ such that

(3.41) $\quad F(n) = f_0(n)$ for $n \in H$

(3.42) $\quad F(y) \leq g(y)$ for $y \in \mathcal{Y}$.

Thus, for $K = M - N$

(3.43) $\quad F(K) \leq 0$

and

(3.44) $\quad F(0) = 0$.

Since F separates $M - N$ and zero, by Lemma 3.2 F separates M and N. (3.33) follows since M is a subspace. (3.8b) implies $g(n_0) > 0$, and (3.34) and (3.41) show $F(n_0) > 0$; thus (3.31) holds. By combining the separation of M and N with (3.44) and (3.31), (3.32) follows. □

Theorem 3.2 proves the existence of a linear functional which separates M and N. However, we need a continuous linear functional which separates M and N. The next theorem shows the existence of a continuous linear functional with the desired properties.

Theorem 3.3.

A necessary and sufficient condition for the existence of a non-zero continuous linear functional F satisfying (3.31) through (3.33) is that M and N have an empty expanded intersection through n_0.

Proof:

The Basic Separation Theorem proves the existence of F satisfying (3.31), (3.32), and (3.33). We shall show F to be continuous at zero. First, we get bounds on F in terms of g as defined by (3.2).

(3.42) implies

$$F(-y) \leq g(-y), \tag{3.45}$$

and with the homogeneity of F, we have

$$|F(y)| \leq \max\{g(y), g(-y)\} \text{ for } y \in \mathcal{Y}. \tag{3.46}$$

But for $y \in \mathcal{Y}$ we have

$$g(y) \leq \|y\| + \inf_{k \in K} \|k\| = \|y\|. \tag{3.47}$$

Combining (3.46) and (3.47) gives

$$|F(y)| \leq \|y\| \text{ for all } y \in \mathcal{Y}. \tag{3.48}$$

Thus, for $\varepsilon > 0$, $|F(y) - F(0)| = |F(y)| < \varepsilon$ whenever $\|y\| \leq \varepsilon/2$, and so F is continuous at zero. By Lemma II.1.3 of Dunford and Schwartz (1958), F is continuous. □

To show necessity assume

$$\inf_{n,m} \|n_0 + n - m\| = 0. \tag{3.49}$$

By (3.31) to (3.33) and the linearity of F, for all n, m we have

$$F(n_0 + n - m) \geq F(n_0) = \varepsilon > 0. \tag{3.50}$$

This contradicts the continuity of F since $F(0) = 0$, and thus (3.49) cannot hold. □

The preceding theorem guarantees the existence of a continuous linear functional which separates M and N. If N has a countable base it will be useful to know that there exists a continuous linear functional such that $F(M) = 0$ and $F(N) > 0$, i.e., F strictly separates M and N. This is the goal of the next theorem.

Theorem 3.4.

Assume M is a linear manifold contained in $\mathcal{Y}$, and N is a non-empty convex cone in $\mathcal{Y}$. Assume N has a countable base $\{e_i\}$

(that is, there exists a sequence $\{e_i\}$ where $e_i \in N$ such that for all $n \in N$, $n = \sum_{i=1}^{\infty} c_i e_i$ for some $c_i \geq 0$ for $i = 1,2,\ldots$).
The expanded intersection of M and N is empty if and only if (iff) there exists a continuous linear functional F such that

(3.51) $\quad F(n) > 0$ for all $n \in N$

(3.52) $\quad F(m) = 0$ for all $m \in M$.

<u>Proof:</u>

By repeatedly applying Theorem 3.3 there exists a sequence $\{F_i\}_{i=1}^{\infty}$ of non-zero continuous linear functionals on $\mathcal{Y}$ such that for each $i = 1,2,3,\ldots$

(3.53) $\quad F_i(e_i) > 0$

(3.54) $\quad F_i(n) \geq 0$ for all $n \in N$

(3.55) $\quad F_i(m) = 0$ for all $m \in M$.

The norm of F_i is defined by

(3.56) $$\|F_i\| = \sup_{\|y\| \leq 1; y \in \mathcal{Y}} |F_i(y)|,$$

and since F_i is continuous, by Theorem 3.1A of Taylor (1961, p. 85), $\|F_i\|$ is finite. Thus, without loss of generality we can assume

(3.57) $\quad \|F_i\| = 1/i^2$.

Furthermore, F_i is a member of the Banach space of all continuous linear operators on $\mathcal{Y}$ into the real line (see Dunford and Schwartz (1958, p. 61, Lemma 8)). Notice

(3.58) $$\sum_{i=1}^{\infty} \|F_i\| = \sum_{i=1}^{\infty} 1/i^2 < \infty,$$

and thus by Theorem 3.13-C of Taylor (1961, p. 100), F defined by

$$(3.59) \qquad F = \sum_{i=1}^{\infty} F_i$$

is a continuous linear functional on $\mathcal{Y}$. Clearly (3.51) and (3.52) hold for this F. Necessity follows as in the previous proof. □

This concludes our development of theorems which hold for real normed linear spaces. Next we specialize this to ℓ^p spaces.

3.3 Separation Theorems for ℓ^p Space.

For $p \in [1, \infty)$ let ℓ^p be the set of infinite sequences ν such that $\sum_{i=1}^{\infty} |\nu_i|^p < \infty$. The norm for ℓ^p is given by

$$(3.60) \qquad \|\nu\|_p = \{\sum_{i=1}^{\infty} |\nu_i|^p\}^{1/p} \qquad \nu \in \ell^p.$$

ℓ^∞ will denote the set of infinite sequences ν such that $\sup|\nu_i|$ is finite. The norm for ℓ^∞ is

$$(3.61) \qquad \|\nu\|_\infty = \sup_{i=1,2,\ldots} |\nu_i| \qquad \nu \in \ell^\infty.$$

ℓ^p spaces are normed linear spaces (see Taylor (1961)), and thus, the theory of Section 3.2 applies. In this section, attention is focused on the form of a continuous linear functional which separates the set of all semi-positive vectors from a linear manifold in a ℓ^p space, under the assumption that the linear manifold has an empty expanded intersection with the semi-positive vectors. First, we consider a result which follows as a corollary to Theorem 3.4.

Corollary 3.1.

Let M be a linear manifold in ℓ^q, $q \in [1, \infty]$ and let P be the set of all semi-positive vectors in ℓ^q. The expanded intersection of M and P is empty iff there exists a continuous linear functional F such that

(3.62) $\quad F(m) = 0 \quad$ for all $m \in M$

(3.63) $\quad F(p) > 0 \quad$ for all $p \in P$.

Proof:

Clearly P is a non-empty convex cone. Let e_i be the vector with i^{th} element equal to one with all other elements zero. Thus for any $p \in P$, $p = \Sigma p_i e_i$ and $p_i \geq 0$. Now apply Theorem 3.4. □

Next, we consider the representation of F on ℓ^p. First, we examine the case when $p \in [1, \infty)$, and then $p = \infty$.

Definition 3.3.

$< y, y' >$ will denote the inner product between $y \in \ell^p$ and $y' \in \ell^q$ where p and q are determined by

$$(3.64) \qquad q = \begin{cases} p/(p-1) & \text{if } p \in (1, \infty) \\ 1 & \text{if } p = \infty \\ \infty & \text{if } p = 1 . \end{cases}$$

By Hölder's inequality (see Royden (1963), p. 97) $< y, y' >$ is well defined and finite for $y \in \ell^p$ and $y' \in \ell^q$. Theorem 3.5 is a version of Theorem 3.4 for ℓ^p spaces where the convex cone is either the set of all semi-positive or positive vectors.

Theorem 3.5.

Let M be a linear manifold in ℓ^p where $p \in [1, \infty)$, and let P be the set of all semi-positive [positive] vectors in ℓ^p. The expanded intersection of P and M is empty iff there exists a vector $\nu \in \ell^q$, where q is defined by (3.64), such that $\nu > 0$ $[\nu \geq 0]$ and $< \nu, m > = 0$ for all $m \in M$. ($\nu > 0$ $[\nu \geq 0]$ denotes that ν is semi-positive [positive].)

Proof:

Assume P is the set of all semi-positive vectors. From Corollary 3.1 there exists a continuous linear function F such that (3.62) and (3.63) hold. By Theorem 4.32-A of Taylor (1961, p. 194), there exists $\nu \in \ell^q$ such that

$$F(x) = < \nu, x > \text{ for all } x \in \ell^p. \tag{3.65}$$

Now $e_i \in P$ and so by (3.63) and (3.65)

$$< \nu, e_i > \quad > \quad 0 \tag{3.66}$$

which implies $\nu_i > 0$. This holds for all $i = 1,2,...$ and so $\nu > 0$. $< \nu, m > = 0$ for all $m \in M$ as a consequence of (3.62) and (3.65). This completes the proof when P is the set of semi-positive vectors.

Next, assume P is the set of positive vectors. As above we represent a continuous linear functional F, for which (3.62) and (3.63) holds, as an inner product with ν. Now e_i is in the closure of P, and so by combining the continuity of F with (3.63), we see that $< \nu, e_i >$ is greater than or equal to zero. This holds for all $i = 1,2,...,$ and thus $\nu \geq 0$. The necessity is clear. □

Theorem 3.6.

Let M be a linear manifold in ℓ^∞ and P be the set of all semi-positive [positive] vectors in ℓ^∞. Let J denote the set of all positive integers. The expanded intersection of P and M is empty iff there exists a bounded finitely additive set function μ defined on J such that

(3.67) $\int_J m_j \mu(dj) = 0$ for all $(m_1, m_2, \ldots) \in M$

$\mu(j) > 0$ $[\mu(j) \geq 0]$ for all $j \in J$.

Proof:

By Corollary 3.1 there exists a continuous linear functional F such that (3.62) and (3.63) hold. By applying Corollary 3 of Dunford and Schwartz (1958, p. 259), we have

(3.68) $$F(\nu) = \int_J \nu_j \mu(dj) \qquad \nu \in \ell^\infty$$

where μ is a bounded finitely additive set function on J. By (3.63) and (3.68) we have $\mu(i) = F(e_i) > 0$. Applying (3.68) to (3.62) completes the proof when P is the set of semi-positive vectors. For the positive vectors, use a proof similar to the second part of the proof of Theorem 3.5. □

To see the result of weakening the empty expanded intersection assumption to $P \cap \overline{M} = \varphi$ where $\overline{M}$ is the closure of M in ℓ^P, let us consider the next example.

Example 2.1.

Let M be spanned by the vectors

$m_1 = (1, 1, -1/2, 0, 0, 0, 0, 0, \ldots)$

$m_2 = (0, 0, 1/2, 1, -1/3, 0, 0, 0, \ldots)$

$m_3 = (0, 0, 0, 0, 1/3, 1, -1/4, 0, \ldots)$

etc. Clearly $M \subset \ell^1$ and M contains no semi-positive vector; furthermore, neither does the closure of M and $\overline{M} \cap P = \varphi$. We see that an orthogonal vector is (1, 0, 2, 0, 3, etc.) which is not in ℓ^∞ nor is it positive.

12. Royden, H. L. (1963). Real Analysis. The Macmillan Co., New York.

13. Rubin, Herman (1971). Discussion of Buehler (1971).

14. Taylor, Angus E. (1958). Introduction to Functional Analysis. John Wiley and Sons, Inc., New York.

15. Wallace, David L. (1959). Conditional confidence level properties. Annals of Mathematical Statistics **30** 864-876.

REFERENCES

1. Buehler, Robert J. (1959). Some validity criteria for statistical inferences. Annals of Mathematical Statistics 30 845-863.

2. Buehler, Robert J. (1971). Measuring Information and Uncertainty. In Godambe and Sprott (1971), 330-341.

3. Cornfield, Jerome (1969). The Bayesian outlook and its applications. Biometrics 25 617-658.

4. Dunford, Nelson and Schwartz, Jacob T. (1958). Linear Operators Part I: General Theory. Interscience Publishers, Inc., New York.

5. Fraser, Donald A. S. (1971). Events, Information Processing, and the Structured Model. In Godambe and Sprott (1971), 32-55.

6. Freedman, David A. and Purves, Roger A. (1969). Bayes' method for bookies. Annals of Mathematical Statistics 40 1177-1186.

7. Gale, David (1960). The Theory of Linear Economic Models. McGraw-Hill Book Company, Inc., New York.

8. Godambe, V. P. and Sprott, D. A. (eds.)(1971). Foundations of Statistical Inference. Holt, Rinehart and Winston of Canada, Ltd., Toronto and Montreal.

9. Heath, David C. and Sudderth, William D. (1972). On a theorem of de Finetti, oddsmaking,and game theory. Annals of Mathematical Statistics (to appear).

10. Lindley, D. V. (1971). Bayesian Statistics, A Review. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

11. Loève, Michel (1955). Probability Theory. D. Van Nostrand Co., Inc., Princeton, New Jersey.

$$s_+(x) = I_{[0,\ M]}(s(x)) \cdot s(x)/M$$

$$s_-(x) = -\ I_{[-M,\ 0]}(s(x)) \cdot s(x)/M.$$

For ξ defined in Lemma 6.1 we have

$$\begin{aligned} M^{-1}\ G_s^c(\theta) &= \int \xi(x,\theta)\ s_+(x)\ f(x,\theta)\ d\mu(x) \\ &\quad - \int \xi(x,\theta)\ s_-(x)\ f(x,\theta)\ d\mu(x) \\ &= I_{(0,\ 1]}(E_\theta(s+)) \cdot (P^{(s+)}C_{\cdot\theta} - \alpha) \\ &\quad - I_{[-1,\ 0)}(E_\theta(s-)) \cdot (P^{(s-)}C_{\cdot\theta} - \alpha). \end{aligned} \tag{6.27}$$

Since M, $E_\theta(s+)$, and/or $E_\theta(s-)$ are positive for a non-trivial s, strict coherence implies $S_3(\alpha)$ a.e. $[\lambda]$. On the other hand even if c is $S_3(\alpha)$ a.e. $[\lambda]$, this does not imply that the linear combination (6.27) could not be semi-positive for some s+ and s-. Thus, strict coherence is stronger than $S_3(\alpha)$ a.e. $[\lambda]$.

It is interesting to note that the proof of Theorem 1 of Wallace only shows that if c is proper Bayes w.r.t. $\pi > 0$, then c is $S_3(\alpha)$ a.e. $[\lambda]$ and not $S_3(\alpha)$ as is claimed. Our Theorem 6.1 is somewhat like a converse to Theorem 1 or Wallace.

Irrespective of a.e. considerations it is easily seen that weak coherence implies S_0, S_1, and S_2. We close this chapter with a quote from Wallace.

> Buehler's examples, combined with the examples and results of this paper, seem to indicate the need for properties intermediate to advance probability and strong advance probability, and even suggest that strong advance probability may be so strong and rare as to be of little value.

There are, however, mathematical advantages in viewing the problem in terms of payoff functions $G_s^c(\theta)$ since they form a linear vector space for sup-bounded stake functions. Suppose we let $V(\theta, s) = P_\theta^{(s)}(C_{\cdot\theta}) - \alpha$. Then we note that $kV(\theta, s)$ is not equal to $V(\theta, ks)$ for all choices of real k because $V(\theta, ks)$ is not defined for $k < 0$ or for $ks > 1$. Thus, the set $V(\theta, s)$ does not form a linear vector space, and this limits the scope of mathematics that applies.

At first S_3 and strict coherence appear equivalent since they both have the flavor of avoiding semi-positive payoffs. Likewise one might consider the selection of Wallace as a generalization of the pure strategies of Buehler. However, both Buehler's and our scheme allow Paul the opportunity to bet on or against the confidence interval at odds proportional to the confidence level. In contrast, Wallace's selection only allows Paul to vary the amount bet on the confidence interval. Intuitively, this means that if the odds are too low, an expected payoff will never become positive just by varying the amount of a positive stake. Let us study the relationship between S_3 and strict coherence in detail.

Suppose we modify the definition of $S_3(\alpha)$ as follows.

(viii)'. c has property $S_3(\alpha)$ a.e. $[\lambda]$ if, for every selection stake function s for which $G_s^c(\theta) \leq 0$ for all θ or else $G_s^c(\theta) \geq 0$ for all θ, equality holds a.e. $[\lambda]$.

For an arbitrary stake function s with sup-bound M, define two selection stake functions as

$$0 \leq \sup_{\theta \in \Theta} G_s^c(\theta).$$

(vi). c has property $S_1(\alpha)$ if, for every selection stake function s

$$\inf_{\theta \in \Theta} G_s^c(\theta) \leq 0 \leq \sup_{\theta \in \Theta} G_s^c(\theta).$$

(vii). c has property $S_2(\alpha)$ if, for every selection stake function s, there exists parameter values θ_1, θ_2, such that

$$G_s^c(\theta_1) \leq 0 \leq G_s^c(\theta_2).$$

(viii). c has property $S_3(\alpha)$ if, for every selection stake function s for which $G_s^c(\theta) \leq 0$ for all θ or else $G_s^c(\theta) \geq 0$ for all θ, equality holds for all θ.

It is known that exact confidence is equivalent to weak exactness, and earlier in this section we discussed the equivalent of weak exactness in Peter-Paul games. If a weakly coherent confidence procedure c is weakly exact, then it will have the property of advance probability. It is conjectured that a continuous version of Example 5.3 would give a c having advance probability.

Since Wallace only considers selections for which $E_\theta(s) > 0$, by (6.25) we see that $G_s^c(\theta) \gtreqless 0$ is equivalent to $P_\theta^{(s)}(C_{\cdot\theta}) \gtreqless \alpha$ respectively. This fact was used in presenting our translated versions of properties S_0, S_1, S_2 and S_3. Thus, if one's concern is the semi-positiveness of $P_s^{(s)}(C_{\cdot\theta}) - \alpha$, it is equivalent to consider either betting a fixed amount, say one dollar, with probability $s(x)$ ($s(x) \in [0, 1]$) or placing a bet of $s(x)$ dollars with probability one.

Then if $E_\theta(s) > 0$, we have

$$(6.25)\qquad G_s^c(\theta)/E_s(\theta) = \int [I_{C_{x\cdot}}(\theta) - \alpha(x)]\, s(x)\, f(x, \theta)\, d\mu(x)/E_\theta(s)$$

$$= \int I_{C_{x\cdot}}(\theta)\, f^{(s)}(x, \theta)\, d\mu(x) - \alpha \int f^{(s)}(x, \theta)\, d\mu(x)$$

$$= P_\theta^{(s)}(C_{\cdot\theta}) - \alpha$$

where

$$(6.26)\qquad f^{(s)}(x, \theta) = f(x, \theta)\, s(x)/E_\theta(s)$$

and $P_\theta^{(s)}$ denotes probability w.r.t. $f^{(s)}(x, \theta)$.

Now we can translate several of his performance properties of a confidence procedure c assuming $\alpha(x) \equiv \alpha$.

(i). c has property $c(\alpha)$, called exact confidence α, if for all $\theta \in \Theta$, $G_s^c(\theta) = 0$ whenever $s(x) \equiv 1$.

(ii). c has property $\underline{c}(\alpha)$ called lower confidence α if for $s(x) \equiv 1$ we have

$$\inf_{\theta \in \Theta} G_s^c(\theta) = 0.$$

(iii). c has property $\overline{c}(\alpha)$ called upper confidence α if for $s(x) \equiv 1$

$$\sup_{\theta \in \Theta} G_s^c(\theta) = 0.$$

(iv). c has advance probability α if it has exact confidence α, and if, for any selection stake function for which $G_s^c(\theta) = q$ for all $\theta \in \Theta$, $q = 0$.

(v). c has property $S_0(\alpha)$ if, for every selection stake function s

We observe that the rule R is equivalent to c, the U to θ, and Paul bets consistently that A is false by letting $s(x) \equiv -1$. Thus the translation: c is weakly exact if $G_s^c(\theta) \equiv 0$ whenever $s(x)$ is either identically plus or minus one and $\alpha(x) \equiv \alpha$. As Buehler points out, weak exactness is equivalent to writing $P_\theta\{C_{\cdot\theta}\} = \alpha$ for all θ. This follows by rewriting the expected payoff (6.4).

We translate his concept of semirelevant and relevant subsets. Let $\varepsilon > 0$ be independent of θ. Then a pure stake function induces (by 6.23) subsets which are called

$$(6.24) \qquad \text{semirelevant} \begin{cases} \text{if } G_s^c(\theta) > 0 & \text{for all } \theta \\ \text{or if } G_s^c(\theta) < 0 & \end{cases}$$

$$\text{relevant} \begin{cases} \text{if } G_s^c(\theta) \geq \varepsilon & \text{for all } \theta \\ \text{or if } G_s^c(\theta) \leq -\varepsilon & . \end{cases}$$

If we considered (6.24) to hold a.e. $[\lambda]$ instead of for all θ, then semirelevant subsets would be tantamount to pure stake functions with positive expected payoffs. Irrespective of a.e. considerations, by (6.24) and Definition 6.3, weak coherence (and hence strict coherence) implies there exists no semirelevant subsets. This follows since weak coherence requires that in the class of all sup-bounded strategies of Paul (not only pure ones), there exists no positve payoff. In this thesis we have not studied a counter-part to relevant subsets.

In the same year Wallace (1959) addressed a similar topic. To establish the relation between his paper and ours, suppose $\alpha(x) \equiv \alpha$ for $\alpha \in (0, 1)$ and consider stake functions s which are bounded by zero and one. We will call such an s a <u>selection stake function</u>.

Although our model comes from Buehler, our concept of coherence stems from a generalization of Cornfield (1969) and Freedman and Purves (1969). As previously discussed, they apply a Peter-Paul game to finite discrete sample and parameter spaces where Peter is required to state his odds or probabilities of all possible subsets of the parameter space for a given sample observation. In Cornfield's discrete model, coherence implies that Paul can notachieve a semi-positive payoff. Thus, our definition of coherence for the confidence interval model is both a generalization (from discrete to continuous spaces) and a restriction (from all possible subsets to one) of Cornfield (1969). Let us examine the similarities between coherence and Buehler's criteria.

Suppose we restrict Paul to stake functions (referred to as pure) which take values of either -1, 0, or $+1$ for each x. The sets C^+ and C^- of Buehler are

$$(6.23) \qquad C^+ = \{x: s(x) = 1\} \text{ and } C^- = \{x: s(x) = -1\}.$$

He considers an expected gain to Paul; by inspecting (6.4) we see that it corresponds to $G_s^c(\theta)$ for pure stake functions where Peter's strategy c determines the event A in a probability assertion for each x. Buehler proposed a criterion of weak exactness as:

> If the model is adequately specified, one should in principle be able to calculate the expected gain to Paul. For any fixed experimental conditions K the expected gain would be a function of (i) the state of nature U, (ii) Peter's rule R, and (iii) Paul's strategy S. Different criteria for the sensibility of Peter's rule might be put forward in terms of this expected gain. For example, I propose the following. Suppose Paul's strategy is to bet consistently that A is false, regardless of the observations. Thenif Paul's expected gain is zero for all U, Peter's rule R will be defined to be weakly exact.

class (p), and that c is (q) weak Bayes w.r.t. $\pi > 0$ $[\pi \geq 0]$. If either

$$(6.20) \qquad \int_X \alpha(x)[1 - \alpha(x)] \int_\Theta f(x, \theta)\, \pi(\theta)\, d\lambda(\theta)\, d\mu(x) < \infty,$$

$$(6.21) \qquad \int_X \alpha(x)[1 - \alpha(x)] \; \| f(x, \underset{\sim}{\theta}) \|_p \, d\mu(x) < \infty,$$

or

$$(6.22) \qquad \int_X \{\int_{\theta \notin C_{x\cdot}} |\alpha(x)\, f(x,\theta)|^p d\lambda(\theta) + \int_{\theta \in C_{x\cdot}} |[1-\alpha(x)]\, f(x,\theta)|^p d\lambda(\theta)\}^{1/p} d\mu(x) < \infty$$

holds, then c is strictly [weakly] coherent.

The above sufficient conditions are rather inadequate because they fail to hold whenever $\alpha(x) \equiv \delta$ $(0 < \delta < 1)$ and $\pi \in L^q$ where $q > 1$. These sufficient conditions are all aimed at showing π is orthogonal to $\mathcal{Q}^c$ which implies coherence. There may be other ways to demonstrate coherence, such as a continuous version of Example 5.3. We leave this as an open question.

6.5 Comparison with Other Models.

Buehler (1959) and Wallace (1959) introduced papers dealing with the conditional confidence levels of confidence procedures. This section discusses the relationship between the above papers and our confidence interval model. A correspondence is made between Buehler's semirelevant subsets, stake functions, and Wallace's concept of selection.

We have fashioned our model after one proposed by Buehler (1959) where, in a continuous sample and parameter space, testing the validity of confidence intervals takes the form of a game between two players. We will see that some of the criteria that Buehler proposes have direct counterparts in our model while others do not.

6.4 Sufficient Conditions for Coherent Bayes.

Unfortunately necessary and sufficient conditions for Bayesian strategies to be coherent have not yet been established. From the previous section we know that coherence for class (p) densities implies a Bayes' solution based on a prior in L^q. In this section we will state theorems which provide sufficient conditions for some Bayes' solutions to be coherent. Most of these results are analogous to those of Chapter V for discrete spaces, and since the proofs are similar, they will be omitted. The first lemma states that whenever π is orthogonal to $\mathcal{H}_p^c$, the empty expanded intersection condition is satisfied.

Lemma 6.3.

Assume there exists $\pi \in L^q$ such that $\pi > 0$ [$\pi \geq 0$], and π is orthogonal to $\mathcal{H}_p^c$. Then the expanded intersection of $\mathcal{H}_p^c$ and P is empty where P is the set of semi-positive [positive] functions in L^p.

In particular for a class (p) density, π is orthogonal to $\mathcal{H}_p^c$ whenever c is (q) weak Bayes w.r.t. π, and we have the next result.

Theorem 6.3.

If c is proper Bayes w.r.t. $\pi > 0$ [$\pi \geq 0$], then c is strictly [weakly] coherent.

By combining continuous analogues of Lemma 5.2 and 5.3, and Theorems 5.2 and 5.3, we have a series of sufficient conditions for coherence of (p) weak Bayes' strategies.

Theorem 6.4.

Assume the density of the confidence interval specification is

Proof:

(q) weak Bayes w.r.t. π follows by applying Theorem 3.7 as in the previous proof. To see that (6.12) must be well defined on a set of x with positive μ measure, let $A_\pi = \{x: \int f\,\pi\,d\lambda > 0\}$. Since $f\,\pi \geq 0$, by the Iterated Integrals Theorem of Loeve (1955, p. 136) we have

$$(6.18) \qquad \int\int f\,\pi\,d\lambda\,d\mu = \int \pi \int f\,d\mu\,d\lambda = \int \pi\,d\lambda > 0.$$

This implies $\mu(A_\pi)$ can not be zero.

To study the cases when $p = \infty$, consider the next definition.

Definition 6.6.

Let π be a bounded finitely additive set function on $\mathcal{B}$. c is said to be BA Bayes w.r.t. π if $\pi(B) \geq 0$ for all $B \in \mathcal{B}$, and

$$(6.19) \qquad \int_{c_x} f(x,\theta)\,\pi(d\theta) = \alpha(x) \int_{\Theta} f(x,\theta)\,\pi(d\theta) \text{ a.e. } [\mu].$$

By applying Theorem 3.10 in a way similar to the proof of Theorem 6.1, one could prove an analogous result for class (∞) densities provided (6.16) is true. Previously (6.16) held since the interchange of the order of integration was with respect to two σ-additive measures, μ and λ. Now the interchange would involve a σ-additive measure μ and a finitely additive set function π, and the validity of (6.16) is in question. We leave the verification of (6.16) as an open question and state the analogous result as a conjecture.

Conjecture:

Assume the f in the specification is a tractable class (∞) density. If c is strictly coherent, then c is BA Bayes.

to (6.15) allows the interchange of the order of summation giving

$$(6.16)\qquad \int_{\Theta} \pi(\theta) \int_{A} \xi(x, \theta)\, f(x, \theta)\, d\mu(x)\, d\lambda(\theta)$$

$$= \int_{A} \int_{\Theta} \xi(x, \theta)\, f(x, \theta)\, \pi(\theta)\, d\lambda(\theta)\, d\mu(x)$$

Since π is orthogonal to $\mathcal{H}_p^c$, the left hand side of (6.16), and hence, the right hand side equals zero. This occurs for <u>all</u> measurable sets $A \subset A_{ij}$, and thus we conclude that

$$(6.17)\qquad \int_{\Theta} \xi(x, \theta)\, f(x, \theta)\, \pi(\theta)\, d\lambda(\theta) = 0$$

a.e. $[\mu]$ where μ is restricted to A_{ij}.

Note that π does not depend on A_{ij} so that (6.17) holds a.e.$[\mu]$ for all A_{ij}, $i = 1, 2, \ldots; j = 1, 2, \ldots$ Since $X = \sum A_{ij}$ we conclude (6.17) holds a.e.$[\mu]$. Clearly (6.17) can be rewritten as (6.12) which implies c is (q) weak Bayes w.r.t. π.

Define B_i' by $B_i' = \{\theta: \pi(\theta) > 0\} \cap B_i$. Because of (6.14), $\lambda(B_i')$ is greater than zero for all i. The B_i have been chosen so that for each x there exists some i such that $f(x, \theta) > 0$ for all θ in B_i. Thus for $x \in B_i$, $\int_{\Theta} f\pi d\lambda \geq \int_{B_i'} f\pi d\lambda > 0$, and (6.12) is well defined.

As you can see, the last part of the proof assures us that (6.12) is well defined when c is strictly coherent and f is a tractable density. In countable spaces we noted that an equation analogous to (6.12) need not be well defined for all x when α was required to be only weakly coherent. This is also true in continuous spaces, and as a result, the restriction of tractability can be dropped from the hypothesis for weakly coherent c.

<u>Theorem 6.2</u>.

For $p \in [1, \infty]$ assume that the f in the specification is a class (p) density. If c is weakly coherent, it is (q) weak Bayes w.r.t. π. Equation (6.12) holds non-trivially on a set of positive μ measure.

by a wide class of densities commonly occurring in statistical inference. We state a theorem for these densities.

Theorem 6.1.

For $p \in [1, \infty]$ assume that the f in the specification is a tractable class (p) density. If c is strictly coherent then c is (q) weak Bayes w.r.t. π where q is determined by (5.1). Furthermore, equation (6.12) holds non-trivially, i.e. $\int f\pi d\lambda > 0$ for each x.

Proof:

Let $\{A_i\}$ and $\{B_i\}$ be sequences of sets such that (6.5.ii) and (6.13) hold respectively. Strict coherence requires $\mathcal{H}_p^c$ and the set of semi-positive functions in L^p to have an empty expanded intersection. Thus by Theorem 3.8 there exists a semi-positive function π in L^q such that π is orthogonal to $\mathcal{H}_p^c$ and

(6.14) $\quad < \pi, I_{B_i} > \; > 0$ for all i.

By Lemma 6.2 there exists a sequence $\{D_i\}$ such that $X = \sum_{i=1}^{\infty} D_i$ and (6.10) holds. Let $A_{ij} = A_i \cap D_j$. For a fixed i and j let A be any measurable subset of A_{ij}, and define a stake function s_A corresponding to A by

$$s_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{elsewhere.} \end{cases}$$

By Lemma 6.1 the expected payoff function corresponding to s_A is in $\mathcal{H}_p^c$. For ξ defined in the proof of Lemma 6.1, $|\xi(x, \theta)| \le 1$, and so by (6.10) we have

(6.15) $$\int_{A_{ij}} \int_{\Theta} |\xi(x, \theta)| f(x, \theta)\, \pi(\theta)\, d\lambda(\theta)\, d\mu(x) < \infty.$$

Applying Tonelli's and Fubini's Theorems (see Royden (1963, p. 233-234))

One possible way in which Peter could choose c is to be consist with Bayesian posterior probabilities. Later we will see the merits of this type of choice.

Definition 6.4.

For $p \in [1, \infty]$, c is said to be (p) weak Bayes w.r.t. π if and only if there exists a function π such that $\pi: \Theta \to [0, \infty]$, $\pi \in L^p$, and

$$(6.12) \qquad \int_{C_x} f(x, \theta)\, \pi(\theta)\, d\lambda(\theta) = \alpha(x) \int_{\Theta} f(x, \theta)\, \pi(\theta)\, d\lambda(\theta) \quad \text{a.e.}[\mu].$$

In discrete spaces we proved that strict coherence implied Peter's strategy to be consistent with Bayes' posterior probabilities for some strictly positive prior. Since the prior is positive, the posterior probabilties were well defined by equation (4.13). In continuous spaces we have been unable to show the existence of an a.e.$[\lambda]$ positive function which is orthogonal to a linear space containing no semi-positive function. If the prior is only semi-positive, then for an arbitrary density $f(x, \theta)$ the integral $\int f\pi d\lambda$ may be zero for some x, and then (6.12) will not uniquely define $\alpha(x)$. To circumvent this problem we consider a very mild restriction on the density.

Definition 6.5.

A density $f(x, \theta)$ is said to be tractable if there exists a disjoint decomposition of Θ into sets $B_i \in \mathcal{B}$ such that for each $x \in X$

$$(6.13) \qquad \{\theta: f(x, \theta) > 0\} \supset B_i \quad \text{for some } i.$$

Note that discrete type (p) p.m.f.'s of the previous chapters have the properties of class (p) and tractability. These properties appear to be useful in studying the continuous cases and are satisfied then there exists a sequence $\{D_i\}$ which depends on π such that

$$(6.9) \qquad X = \sum_{i=1}^{\infty} D_i$$

and

$$(6.10) \qquad \int_{D_i} \int_{\Theta} f(x, \theta)\, \pi(\theta)\, d\lambda(\theta)\, d\mu(x) < \infty \quad \text{for all } i.$$

Proof:

If $W(x) = \int_{\Theta} f(x, \theta)\, \pi(\theta)\, d\lambda(\theta)$, then Hölder's inequality assures us that $W(x)$ is well defined and finite for all $x \in \mathbf{X}$. Let $D_i' = \{x: i - 1 \leq W(x) < i\}$ for $i = 1, 2, \ldots$. Since μ is σ-finite, there exists a sequence $\{D_i''\}$ such that $X = \sum_{i=1}^{\infty} D_i''$ and $\mu(D_i'') < \infty$ for each i. Define $D_{ij}^0 = D_i' \cap D_j''$ for $i = 1, 2, \ldots$ and $j = 1, 2, \ldots$ and reorder $\{D_{ij}^0\}$ into a sequence $\{D_i\}$. Clearly (6.9) holds, and since for every D_i there corresponds a D_{ij}^0 such that

$$(6.11) \qquad \int_{D_{ij}^0} \int_{\Theta} f(x, \theta)\, \pi(\theta)\, d\lambda(\theta)\, d\mu(x) \leq \int_{D_{ij}^0} i\, d\mu(x) = i \cdot \mu(D_{ij}^0) < \infty,$$

then (6.10) follows. □

As in previous chapters we study a property of the Confidence Interval Model.

Definition 6.3.

For a fixed confidence interval specification, Peter's c is said to be strictly [weakly] coherent if and only if (i) $\mathcal{G}^c$ contains no semi-positive [positive] function, and (ii) the expanded intersection of $\mathcal{H}_p^c$ and the set of all semi-positive [positive] functions in L^p is empty. (A function g is said to be semi-positive if $g \geq 0$ a.e.$[\lambda]$ and $\lambda\{g > 0\} > 0$.)

if $p = \infty$, then $\mathcal{H}_p^c = \mathcal{G}^c$.

The next lemma shows that $\mathcal{H}_p^c$ is not empty, and later we will see that $\mathcal{H}_p^c$ contains a rich enough set of payoff functions to be useful in proving Theorem 6.1. Let us define the support, $J(s)$, of a stake function to be

$$(6.6) \qquad J(s) = \{x\colon s(x) \neq 0\}.$$

<u>Lemma 6.1</u>.

Let the density of a specification be of class (p) and suppose $\{A_i\}$ satisfies (6.5). Then $G_s^c(\theta)$ is a member of L^p for all stake functions s such that $J(s) \subset \sum_{i=1}^{N} A_i$ for some $N < \infty$.

<u>Proof</u>:

Let $\xi(x, \theta) = I_{C_{x\cdot}}(\theta) - \alpha(x)$, and M be equal to the ess. sup of $s(x)$. M is finite by the definition of a stake function. Then for all x such that $J(s) \subset \sum_{i=1}^{N} A_i$ we have

$$(6.7) \qquad |G_s^c(\theta)| = \left|\int_X \xi(x, \theta)\, s(x)\, f(x, \theta)\, d\mu(x)\right| \leq M \cdot \sum_{i=1}^{N} \int_{A_i} f(x, \theta)\, d\mu(x).$$

Applying Minkowski's inequality (see Royden (1963, p. 95)) gives

$$(6.8) \qquad \|G_s^c(\theta)\|_p \leq M \cdot \sum_{i=1}^{N} \|P(A_i|\theta)\|_p < \infty.$$

Thus $G_s^c(\theta)$ is in L^p, and so, $\mathcal{H}_p^c$ is not empty. □

In the proof of Theorem 6.1 we will need a result which we find convenient to state and prove at this time. For the remainder of this chapter q equals $p/(p-1)$ if $p \in (1, \infty]$ or ∞ if $p = 1$.

<u>Lemma 6.2</u>.

For some $p \in [1, \infty]$ assume that $f(x, \theta) \in L^p$ as a function of θ for each $x \in X$. If π is a semi-positive function in L^q, ranges over all ess. sup-bounded stake functions $s(x)$. Note: $\mathcal{G}^c$ depends on f and $\alpha(x)$ as well as c; however, f and α are fixed in the specification.

It is easy to verify that $\mathcal{G}^c$ is a linear manifold in L^∞, and as we focus attention on $\mathcal{G}^c$ we might suspect that different strategies (or confidence procedures) c, which Peter could use, will change the space of expected payoff functions. In general, this is true since $\mathcal{G}^c$ depends on c. Thus, how should Peter choose c? The next section investigates an aspect of the problem.

6.3 Implications of Coherence with Class (p) Densities.

In this section we will introduce a space of payoffs which depends on densities belonging to class (p) and use this space to study a property of Peter's strategy c called coherence. A density $f(x, \theta)$ is said to belong to class (p) where $p \in [1, \infty]$ if it satisfies two conditions:

(6.5) (i) $f(x, \underset{\sim}{\theta}) \in L^p$ for each $x \in X$

(ii) there exists a sequence $\{A_i\}_{i=1}^{\infty}$ such that $P(A_i|\underset{\sim}{\theta}) \in L^p$ for all i and $X = \sum_{i=1}^{\infty} A_i$.

This means that for a density f there is a decomposition of X into sets A_i which are "small enough" such that the L^p norm of $P(A_i|\underset{\sim}{\theta})$ is finite. As in Chapter IV we consider a class of payoffs $\mathcal{H}_p^c$ depending on class (p) densities.

Definition 6.2.

If the density of a confidence interval specification is of class (p), then $\mathcal{H}_p^c$ will denote the intersection of $\mathcal{G}^c$ and L^p. Note:

$$(6.2) \qquad \begin{array}{ll} (1-\alpha(x))s(x) & \text{if } \theta \in C_{x.} \\ -\alpha(x)s(x) & \text{if } \theta \notin C_{x.} \end{array} \quad .$$

From the payoff we see that Paul is risking $-\alpha(x)s(x)$ to win $(1-\alpha(x))s(x)$ if $\theta_0 \in C_{x.}$. In terms of odds, Peter offers Paul odds of $(1-\alpha(x))/\alpha(x)$ against $\theta \in C_{x.}$ (assuming $\alpha(x) > 0$). In usual betting terminology $\alpha(x)s(x)$ is the stake placed on $\theta \in C_{x.}$; however, it is usually convenient to refer to $s(x)$ as the stake.

The $\mathcal{G}$ measurability of $\alpha(x)$, $I_{C_{x.}}(\theta)$ and $s(x)$ imply the $\mathcal{G}$ measurability of $\nu(x, \theta)$ where

$$(6.3) \qquad \nu(x, \theta) = [I_{C_{x.}}(\theta) - \alpha(x)]s(x)f(x, \theta).$$

Assuming the integral to be well defined, an <u>expected payoff</u> or gain to Paul is

$$(6.4) \qquad G_s^c(\theta) = \int_X \nu(x, \theta)d\mu(x)$$

when Peter uses strategy c, Paul uses strategy s, and θ is true. Since $\alpha(x)$ and indicator functions are bounded, the integral will be well defined if $s(x)$ is ess. sup-bounded. For each fixed c and s, the expected payoff can be studied as a function in θ, and as in previous chapters, we will investigate the behavior of G_s^c as a function in θ as Paul varies s against a fixed c.

<u>Definition 6.1.</u>

Let Peter use strategy c for some fixed confidence interval specification where the expected payoff $G_s^c(\theta)$ is determined by (6.4). Then $\mathcal{G}^c$ will denote the space of <u>expected payoff functions</u> as s

For any set $C \in \mathcal{C}$ let $C_{x\cdot} = \{\theta: (x, \theta) \in C\}$ and $C_{\cdot\theta} = \{x: (x, \theta) \in C\}$ denote cross section sets of C.

The role of the master of ceremonies is similar to that of previous chapters which study discrete spaces. In the Confidence Interval Model (Model CI) the master of ceremonies determines an $\mathcal{A}$ measurable function $\alpha(x)$ where $0 \leq \alpha(x) \leq 1$ for all $x \in X$. One may interpret $\alpha(x)$ as the confidence level of a confidence interval when x is observed. The elements $(X, \mathcal{A}, \mu)$, $(\Theta, \mathcal{B}, \lambda)$, $(\mathcal{Y}, \mathcal{C}, \mu \times \lambda)$, f, and α denote a <u>confidence interval specification</u> and are considered to be <u>fixed</u> throughout the discussion. The remaining tasks of the master of ceremonies are to select $\theta_0 \in \Theta$ at his leisure and to observe $x \in X$ according to the distribution $P(X \leq x|\theta_0)$. The x is revealed to Peter and Paul while the θ_0 is retained until the wagers have been placed. Now we come to Peter's role.

Peter, being confronted with a confidence interval specification, must select a measurable set $C \in \mathcal{C}$ such that for each $x \in X$ he is willing to assert the probability that $\theta_0 \in C_{x\cdot}$ is $\alpha(x)$; i.e., "$P(\theta_0 \in C_{x\cdot}) = \alpha(x)$." He does this with the knowledge that Paul is allowed to test his probability assertions by placing a positive or negative bet on the event $\theta_0 \in C_{x\cdot}$ for the sample point x. Thus, Paul determines an $\mathcal{A}$ measurable stake function $s(x)$ which combined with $\alpha(x)$ determines the amount that Paul bets on $\theta_0 \in C_{x\cdot}$. The strategies of Peter and Paul are denoted by c and s respectively.

Once the wagers are made the master of ceremonies reveals θ_0, and Peter and Paul settle up. For an observed sample point x, the payoff to Paul is

CHAPTER VI

A Confidence Interval Model

6.1 Introduction.

The conditional behavior of confidence interval procedures in continuous spaces is the focal point of at least two papers, Buehler (1959) and Wallace (1959). Both of these introduce criteria for considering the appropriateness of confidence procedures. With a similar goal, we generalize Model I of Chapter IV to continuous sample and parameter spaces and study confidence procedures in view of coherence.

Our model for continuous spaces is fashioned after Buehler (1959) and is described in Section 6.2. In Section 6.3 and 6.4 we formulate a version of coherence for this model and investigate its relationship to Bayesian solutions. A typical result roughly states that if Peter is to avoid semi-positive expected payoff functions, then he must be π-Bayes where π is a function belonging to an appropriate L^p space. In Section 6.5 we discuss the relationship between coherence and the papers of Buehler and Wallace.

6.2 Description of a Model for Confidence Intervals.

Throughout this chapter $(X, \mathcal{A}, \mu)$ and $(\Theta, \mathcal{B}, \lambda)$ will denote measure spaces with σ-finite measure μ and λ. Suppose $\mathcal{Y}$ is the product space $X \times \Theta$, and $\mathcal{C}$ is the σ-field over $\mathcal{A} \times \mathcal{B}$. Let $\mu \times \lambda$ denote the usual product measure over $\mathcal{C}$. Let f be a $\mathcal{C}$ measurable function on $\mathcal{Y}$ which is a density with respect to μ, i.e.,

$$(6.1) \qquad \int_X f(x, \theta)d\mu(x) = 1 \quad \text{for all} \quad \theta \in \Theta$$

$$\int_A f(x, \theta)d\mu(x) = P(A|\theta) \quad \text{for all} \quad A \in \mathcal{A} .$$

(5.24) $$G_s^\alpha(\theta) = \Sigma c_{ij} G_{s_{ij}}^\alpha(\theta),$$

and since $G_s^\alpha \in \ell^p$ and π corresponds to a continuous linear functional on ℓ^p we have

(5.25) $$< G_s^\alpha, \pi > = \Sigma c_{ij} < G_{s_{ij}}^\alpha, \pi > = 0.$$

The conclusion follows from (5.25). □

Note that (5.21) depends on π, $p(x, \theta)$, and on the chosen sets in a complicated way. At the present we do not know of any simple conditions which show that Bayes' solutions based on improper priors are coherent. With this we conclude our discussion of sufficient conditions.

that $s(n) \to \infty$; this contradicts the sup-boundedness of s. Thus, we can conclude that the expanded intersection is empty. Hence α is strictly coherent.

5.4 ℓ^p Theory for Improper Priors.

In this section we present a method which uses ℓ^p theory to show that some improper Bayes' strategies are coherent. The technique involves finding a sufficient condition which implies that $G_s^{\alpha}(\theta)$ is a member of ℓ^p for all s if the p.m.f. is of type (p).

Theorem 5.3.

Assume Model I; the p.m.f. is of type (p), and α is (q) weak Bayes w.r.t. $\pi > 0$ $[\pi \geq 0]$. If

$$(5.21) \qquad \sum_X \sum_{i=1}^{N(x)} \{ \sum_{\theta \notin A_i(x)} |\alpha_i(x)p(x, \theta)|^p + \sum_{\theta \in A_i(x)} |(1-\alpha_i(x))p(x, \theta)|^p \}^{1/p} < \infty,$$

then α is strictly [weakly] coherent.

Proof:

For sup-bounded stake functions and ν as defined in the proof of Theorem 5.1, (5.21) implies

$$(5.22) \qquad \sum_X \sum_{i=1}^{N(x)} \|\nu(x, \theta, i)s_i(x)p(x, \theta)\|_p < \infty.$$

By Theorem 3.13-C of Taylor (1961, p. 100), (5.22) implies G_s^{α} is a member of ℓ^p. Since α is Bayes we have

$$(5.23) \qquad < G^{\alpha}_{s_{ij}}, \pi > = 0 \quad \text{for all } s_{ij} \in C'$$

where C' is defined by (4.4). Let c_{ij} be the elements of a stake function. Then

$$(5.16)\qquad \begin{aligned} s(4) &\leq s(3) + s(2) - s(1) \\ s(5) &\leq 2s(3) + -s(1) \\ s(6) &\leq 2s(3) + s(2) - 2s(1) \\ s(7) &\leq 3s(3) - 2s(1) \\ s(8) &\leq 3s(3) + s(2) - 3s(1), \end{aligned}$$

and in general for $i = 4, 6, 8, \ldots,$ we have

$$(5.17)\qquad s(i) \leq \left(\frac{i-2}{2}\right)\cdot(s(3) - s(1)) + s(2).$$

Since $s(3) - s(1) < 0$, (5.17) implies $s(i)$ is unbounded which contradicts the sup-bound property of s and thus $s(3) - s(1) \geq 0$. Since (5.15) holds, $s(2) - s(0)$ must be less than zero. By a similar argument $s(-i)$ for $i = -3, -5, \ldots$ will be unbounded. Since s is sup-bounded (5.13) cannot hold, and $\mathcal{G}^{\alpha}$ does not contain any semi-positive vector.

Although Lemma 5.1 implies an empty expanded intersection we give an alternate argument. Without loss of generality we can choose p_0 in Definition 3.1 to be zero except for $p_0(1) = 1$. Suppose s is such that $\|p_0 + p - G_s^{\alpha}\|_1 \leq \epsilon$ for some choice of p. Let $\epsilon = 1/4$ and define $t(i) = s(i+1) - s(i-1)$. Then we have

$$(5.18)\qquad 8G_s^{\alpha}(1) = t(1) - t(2) \geq 6.$$

Assume w.l.g. that $-t(2) \geq 3$. This implies

$$(5.19)\qquad \epsilon \leq \sum_{i=2}^{n} |p(i) - G_s^{\alpha}(i)| \quad \text{for all } n \geq 2$$

which gives

$$(5.20)\qquad 2 = 8\epsilon \geq \left|\sum_{i=2}^{n-1} 8p(i) - t(2) + t(n)\right| + |8p(n) - t(n) + t(n+1)|.$$

Since $-t(2) \geq 3$, this means that $t(n) \leq -1$ for all n which shows

Examples 5.1 and 5.2 show that for a specification in which the master of ceremonies uses his option and fixes a particular collection of sets $A_i(x)$, some priors in ℓ^q have posterior probabilities which are strictly coherent and others do not. In order to show that the conditions (5.6), (5.9), or (5.10) are not necessary for coherence, we present Example 5.3. A point of interest is that the uniform prior over Θ is used.

Example 5.3.

Let $\Theta = X = \{0, \pm 1, \pm 2, \text{etc.}\}$, $N(x) = 1$, $A(x) = \{x, x+1\}$, $\mathcal{G} = \{A(x)\}$, and

$$p(x, \theta) = \begin{cases} 1/4 & x = \theta-1,\ \theta,\ \theta+1,\ \theta+2 \\ 0 & \text{elsewhere.} \end{cases}$$

Suppose $\pi(\theta)$ is uniform over Θ, then $\alpha(x) = 1/2$. Note that none of (5.6), (5.9), and (5.10) hold. To show that there exists no sup-bounded s such that G_s^{α} is semi-positive, we assume without loss of generality (w.l.g.) that

$$(5.13) \qquad G_s^{\alpha}(1) > 0$$

and

$$(5.14) \qquad G_s^{\alpha}(\theta) \geq 0 \quad \text{for all} \quad \theta \in \Theta.$$

After computing the expected payoff (5.13) implies

$$(5.15) \qquad -s(0) + s(1) + s(2) - s(3) > 0.$$

We may assume w.l.g. that $s(3) - s(1) < 0$ (for not both $s(3) - s(1)$ and $s(0) - s(2)$ could be greater than zero). By (5.14) we have the following string of inequalities choose the sets so that one of the conditions holds. On the other hand, if the master of ceremonies has used his option and fixed the sets, then not all priors in ℓ^q will give posterior probabilities such that one of the conditions will be satisfied. If this is so, then other means are needed to show coherence.

Consider once again Example 5.1. We see

$$\sum_X 1\cdot\{\tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{2}\cdot\tfrac{1}{2}\} = \infty \tag{5.12}$$

and so (5.9), (5.10), and hence, (5.6) fail to hold. Possibly this is connected with the non-coherence of α. Since the α which is (∞) weak Bayes w.r.t. $\pi \equiv 1$ is not weakly coherent, one might question if there is any $\pi \in \ell^\infty$ such that a Bayesian solution is strictly or weakly coherent. Example 5.2 shows that there is at least one.

Example 5.2.

Consider the specification of Example 5.1. Let $\pi(\theta) = \pi(-\theta)$ where

Values of π

θ	0	1	2	3	4	5	6	7	8	etc.
$\pi(\theta)$	1	1	1	1/9	1/16	1	1	1/49	1/64	

Then for $\theta \geq 2$ we have

$$\frac{\pi(\theta-1)\cdot\pi(\theta+1)}{\pi(\theta-1)+\pi(\theta+1)} \leq (\theta-1)^{-2},$$

and so

$$\sum_X \frac{\pi(x-1)\pi(x+1)}{\pi(x-1)+\pi(x+1)} < 2 + 2\sum_{i=1}^{\infty} (i)^{-2}.$$

Thus (5.6) holds, and by Theorem 5.2, α which is (∞) weak Bayes w.r.t. the above π is strictly coherent.

$$(5.9) \qquad 2\sum_{X}\sum_{\Theta} p(x,\theta)\pi(\theta)\{\sum_{i=1}^{N(x)} \alpha_i(x)(1-\alpha_i(x))\} < \infty, \text{ and}$$

(ii) a sufficient condition for (5.9) is

$$(5.10) \qquad \sum_{X} \|p(x,\cdot)\|_p\{\sum_{i=1}^{N(x)} \alpha_i(x)(1-\alpha_i(x))\} < \infty.$$

Proof:

(i) By using (4.13) and interchanging the order of summation, (5.8) is equal to the left hand side of (5.9).

(ii) Since $p(x, \underset{\sim}{\theta}) \in \ell^p$ for all x and $\pi \in \ell^q$, by Hölder's inequality we have

$$(5.11) \qquad \sum_{\Theta} p(x,\theta)\pi(\theta) \le \|p(x,\cdot)\|_p \cdot \|\pi\|_q \,.$$

Thus, (5.10) implies (5.9), and combining this with Lemma 5.2 gives the conclusion. □

Theorem 5.2.

Assume Model I; p.m.f. is type (p), and α is (q) weak Bayes w.r.t. $\pi > 0$ $[\pi \ge 0]$. If (5.6), (5.9), or (5.10) holds, then α is strictly [weakly] coherent.

Proof:

By Lemma 5.2 or 5.3, we see (5.7) holds. Thus, for any sup-bounded s, the left hand side of (5.3) is finite. Now apply Tonelli's Theorem to interchange the order of summation in (5.4) and proceed as in the proof of Theorem 5.1. □

Conditions (5.6), (5.9), and (5.10) often depend in a complex manner on the improper prior π and the sets $A_i(x)$. If the master of ceremonies has not used his option, then Peter has the flexibility to choose the sets, and for an arbitrary prior in ℓ^q, perhaps he can coherent Bayesian α in either Model I or Model II. We will, however, develop some sufficient conditions concentrating on Model I. A convenient method of proving a Bayesian α to be strictly [weakly] coherent is to form the inner product of G_s^α with π, interchange the order of summation, and show the inner summation to be identically zero. Lemma 5.2 gives a sufficient condition which allows the use of Tonelli's Theorem for interchanging the order of summation.

Lemma 5.2.

Assume $p(x, \theta)$ is type (p) where $p \in [1, \infty]$, and α is (q) weak Bayes w.r.t. π. Let $W = \{x: \Sigma_\Theta p(x, \theta)\pi(\theta) > 0\}$. If

$$(5.6) \qquad \sum_W \sum_{i=1}^{N(x)} \Big(\sum_{\theta\in A_i(x)} p(x, \theta)\pi(\theta)\Big)\Big(\sum_{\theta\notin A_i(x)} p(x, \theta)\pi(\theta)\Big)\Big(\sum_{\theta\in\Theta} p(x,\theta)\pi(\theta)\Big)^{-1} < \infty,$$

then

$$(5.7) \qquad \sum_\Theta \sum_X \sum_{i=1}^{N(x)} \nu(x, \theta, i)p(x, \theta)\pi(\theta) < \infty.$$

Proof:

Since $p(x, \theta)$ and π are non-negative, by Tonelli's Theorem and the definition of α, (5.7) equals

$$(5.8) \qquad \sum_X \sum_{i=1}^{N(x)} \Big\{ \sum_{\theta\in A_i(x)} (1-\alpha_i(x))p(x, \theta)\pi(\theta) + \sum_{\theta\notin A_i(x)} \alpha_i(x)p(x,\theta)\pi(\theta)\Big\}.$$

The value of (5.8) is not changed if the sum over X is restricted to the sum over W, and by the definition of α, (5.8) equals twice the left hand side of (5.6) which is finite by assumption. □

Lemma 5.3.

Assume $p(x, \theta)$ is type (p) where $p \in [1, \infty]$, and α is (q) weak Bayes w.r.t. π. Then (i) (5.6) holds if and only if

$$s_1(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{elsewhere} \end{cases} \qquad s_2(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{elsewhere} . \end{cases}$$

Then we have

$$G_s^{\alpha}(0) = (1-\tfrac{1}{2})\cdot 1\cdot p(1, 0) + (1-\tfrac{1}{2})\cdot 1\cdot p(-1, 0) = \tfrac{1}{2}$$

$$G_s^{\alpha}(-1) = G_s^{\alpha}(1) = (1-\tfrac{1}{2})\cdot 1\cdot p(2, 1) = \tfrac{1}{4},$$

and for any $\theta > 1$,

$$G_s^{\alpha}(\theta) = (1-\tfrac{1}{2})\cdot 1\cdot p(\theta+1, \theta) - \tfrac{1}{2}\cdot 1\cdot p(\theta-1, \theta) = 0.$$

Similarly $G_s^{\alpha}(\theta)$ equals zero for $\theta < -1$. Hence, although α is weak Bayes it is not strictly coherent. To make matters worse consider the stakes

$$s_1(x) = \begin{cases} 2 - 1/|x| & \text{if } x < 0 \\ 0 & \text{elsewhere} \end{cases} \qquad s_2(x) = \begin{cases} 2 - 1/|x| & \text{if } x > 0 \\ 0 & \text{elsewhere} . \end{cases}$$

For $\theta \geq 2$ we have

$$\begin{aligned} G_s^{\alpha}(\theta) &= (1-\tfrac{1}{2})(2-(\theta+1)^{-1})p(\theta+1,\theta) - \tfrac{1}{2}(2-(\theta-1)^{-1})p(\theta-1, \theta) \\ &= (2\theta^2-2)^{-1} > 0, \end{aligned}$$

and by symmetry $G_s^{\alpha}(\theta) = G_s^{\alpha}(-\theta)$. Easy calculations give

$$G_s^{\alpha}(-1) = G_s^{\alpha}(1) = 3/8 \quad \text{and} \quad G_s^{\alpha}(0) = 1/2 .$$

Thus, α is not even weakly coherent. A similar result was achieved by Rubin (1971, p. 340) who proposed a randomized strategy in which Paul chooses a probability of betting (a unit stake) for each x. This example shows that one can avoid the randomized strategy by suitably chosen stakes.

At the present it is not known what conditions are necessary and sufficient to insure that $\pi \in \ell^P$ will produce a strictly [weakly] prior π used to compute the Bayesian posterior probabilities. That is, Peter can post probabilities which agree with posterior probabilities for any $\pi > 0$, and α will be strictly coherent. In the next section we present some sufficient conditions for a Bayes' solution based on an improper $\pi \in \ell^p$ to be coherent. In these cases Peter will have to exercise caution in the choice of the sets, and not all priors will have associated posterior probabilities which are coherent.

5.3 Priors in ℓ^p where $p \in (1, \infty]$.

Now we consider the case where the p.m.f. of a specification is of type (p); $p \in (1, \infty]$. From the preceding section we know that if α is proper Bayes, then α is coherent. If we can show a similar result holds for improper priors $\pi \in \ell^q$, then the class of strategies which are coherent is enlarged. We note that for type (p) p.m.f.'s, $< p(x, \theta), \pi(\theta) >$ is finite for any $\pi \in \ell^q$; this follows from Hölder's inequality. Thus, when $< p(x, \theta), \pi(\theta) > \; > 0$, α is well defined by (4.13). Before presenting some sufficient conditions which imply weak Bayes' strategies are coherent, we give an example for which $p(x, \theta)$ is type (1); $\pi > 0$ is in ℓ^∞, and α is (∞) weak Bayes w.r.t. π but not strictly or even weakly coherent.

Example 5.1. (Similar examples appear in Buehler (1971, p. 337), Fraser (1971, p. 49), and Lindley (1971, p. 50).)

Let $\Theta = X = \{0, \pm 1, \pm 2, \text{etc.}\}$; $N(x) = 2$; $A_1(x) = \{x+1\}$; $A_2(x) = \tilde{A}_1(x)$, and

$$p(x, \theta) = \begin{cases} \frac{1}{2} & x = \theta \pm 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Suppose $\pi(\theta)$ is the uniform prior over Θ; then $\alpha_1(x) = \alpha_2(x) = 1/2$. Let Paul use strategy s such that

Proof for Model I:

Define $\nu(x, \theta, i) = (I_{A_i}(\theta) - \alpha_i(x))$. For any sup-bounded stake function s we have

$$(5.3) \qquad \sum_{\Theta} \sum_{X} \sum_{i=1}^{N(x)} |\nu(x, \theta, i)| \cdot |s_i(x)| p(x, \theta)\pi(\theta)$$

$$\leq M \cdot N \sum_{\Theta} \sum_{X} p(x, \theta)\pi(\theta) = M \cdot N$$

where $N = \sup_{X} N(x)$ and $M = \sup_{X \text{ and } i=1 \text{ to } N(x)} |s_i(x)|$. By the assumptions of Model I, M and N are finite. Thus, by Tonelli's theorem (Royden (1963, p. 234)), we can interchange the order of summation in

$$(5.4) \qquad \sum_{\Theta} \sum_{X} \sum_{i=1}^{N(x)} \nu(x, \theta, i) s_i(x) p(x, \theta)\pi(\theta)$$

to give

$$(5.5) \qquad \sum_{X} \sum_{i=1}^{N(x)} s_i(x) \sum_{\Theta} \nu(x, \theta, i) p(x, \theta)\pi(\theta).$$

Since α is proper Bayes w.r.t. π by (4.13), the summation over Θ is zero for all x and i. Thus we have $< G_s^{\alpha}, \pi > = 0$ for all sup-bounded s. Assume $\pi > 0$; then for all s such that $G_s^{\alpha}(\theta_0) > 0$ for some $\theta_0 \in \Theta$, there exists a θ' such that $G_s^{\alpha}(\theta') < 0$. Hence $\mathcal{G}^{\alpha}$ contains no semi-positive vector. Apply Lemma 5.1 to see that the expanded intersection of $\mathcal{H}_p^{\alpha}$ and P is empty. Similar comments hold for $\pi \geq 0$.

Proof for Model II:

By applying (4.18) we see that the left hand side of an equation analogous to (5.3) is finite, and the proof follows as above. □

Note that in these theorems, since π is in ℓ^1, strict [weak] coherence does not depend on the choice of sets $\{A_i(x)\}$ or on the

$$(5.1) \qquad q = \begin{cases} p/(p-1) & \text{if } p \in (1, \infty) \\ 1 & \text{if } p = \infty \\ \infty & \text{if } p = 1. \end{cases}$$

Lemma 5.1.

Assume the p.m.f. is type (p) and α is either (q) weak Bayes w.r.t. $\pi > 0$ [$\pi \geq 0$] for $p \in [1, \infty)$ or BA Bayes w.r.t. $\gamma > 0$ [$\gamma \geq 0$] for $p = \infty$. Then the expanded intersection of $\mathcal{H}_p^\alpha$ and the set of semi-positive [positive] vectors in ℓ^p is empty.

Proof:

Let $p \in [1, \infty)$. Since α is Bayes we have $< G_{s_{ij}}^\alpha, \pi > = 0$ for all $s_{ij} \in C'$ where C' is defined by (4.4). For $G_s^\alpha(\theta) \in \ell^p$ we have

$$(5.2a) \qquad G_s^\alpha(\theta) = \Sigma\, c_{ij} G_{s_{ij}}^\alpha(\theta)$$

for suitable choice of c_{ij}. By Hölder's inequality π corresponds to a continuous linear functional on $\mathcal{H}_p^\alpha$ giving

$$(5.2b) \qquad < G_s^\alpha, \pi > = \Sigma\, c_{ij} < G_{s_{ij}}^\alpha, \pi > = 0,$$

and the conclusion follows from Theorem 3.5. The proof for $p = \infty$ is similar. □

Now we use this lemma to show proper Bayes' strategies are coherent. From the proof of Theorem 5.1 we will see that the coherence of a proper Bayes' α is independent of the sets on which the bets are placed. Thus, it is immaterial whether Peter chooses the sets or the master of ceremonies uses his option. This not the case with weak Bayes' solutions, which are discussed in later sections.

Theorem 5.1.

Assume Model I or II. If α is proper Bayes w.r.t. $\pi > 0$ [$\pi \geq 0$], then α is strictly [weakly] coherent.

CHAPTER V

Coherent Proper and Improper Bayes' Strategies
for Countable Spaces

5.1 Introduction.

Now we turn our attention to the converse problem of Chapter IV; that is, in Model I or II, will a strategy that is consistent with either proper, BA, or improper Bayes' solution be coherent. The treatment of proper Bayes' solutions is very tractable since we can interchange the order of summation by applying Tonelli's Theorem to prove that proper Bayes' strategies are coherent.

For improper Bayes' strategies, the general answer is obscured since the conditions for Tonelli's Theorem are not always satisfied. We explore this case by presenting a mixture of theorems and examples. The theorems give some sufficient conditions for interchanging the order of summation --either by Tonelli's Theorem or ℓ^p theory. The examples show that: not all improper Bayes solutions are coherent, the sufficient conditions given are not necessary, and other techniques besides interchanging the order of summation can be used to demonstrate coherence.

The coherence of BA weak Bayes' strategies will not be covered and is an open question. Section 5.2 deals with proper Bayes' strategies for Model I and II. Improper Bayes' strategies are studied in Sections 5.3 and 5.4.

5.2 Proper Bayes' Strategies.

Before we present Theorem 5.1 we give a lemma which will be used not only with proper Bayes' strategies but also with improper ones. Throughout this chapter we assume that q is defined by countably sup-bounded, i.e., $\Sigma|\alpha_i(x')s_i(x')| < 1$. By the disjointness of the sets $A_i(x')$, for all $\theta \in \Theta$ we have

$$(4.20) \qquad \sum_{i=1}^{\infty} [I_{A_i(x')}(\theta) - \alpha_i(x')]s_i(x') = 1 - \sum_{i=1}^{\infty} \alpha_i(x') > 0.$$

Since $p(x', \theta) > 0$ for some $\theta \in \Theta$, $G_s^{\alpha}(\theta)$ is semi-positive. [For weak coherence, this process at all x implies $G_s^{\alpha}(\theta)$ is positive.] This contradicts the assumption of strict [weak] coherence and thus (4.19) cannot hold, so that the probability assertions must be countably additive. □

As in Model I we note that if α is strictly coherent then the probability assertions are consistent with posterior probabilities which are well defined. Also we point out that proper priors correspond to a bounded finitely additive set function on Θ, but not all bounded finitely additive set functions correspond to proper priors. It was first conjectured that coherence for type (∞) p.m.f.'s in Model II would imply that α was consistent with a proper Bayes solution. We have been unable to show this although it may be true. This completes our discussion of the implications of coherence in these models.

Proof:

(i) follows as in Theorem 4.1. (ii) is a consequence of α being (q) weak Bayes. □

The above result is not surprising since the countable additivity follows directly from (q) weak Bayes. With a type (∞) p.m.f. the countable additivity does not follow from the Bayesian property since strict or weak coherence only implies BA weak Bayes. A technique found in the proof of Theorem 6 of Heath and Sudderth (1972) will be used to show countable additivity. If the p.m.f. of X was concentrated at a single point x_0 for all θ, then Theorem 4.4 would be similar to a part of Theorem 6 of Heath and Sudderth.

Theorem 4.4.

Assume that in Model II the p.m.f. is of type (∞), and that Peter's strategy is α. If α is strictly [weakly] coherent, then (i) there exists $\gamma \in \Gamma$ such that $\gamma(\theta) > 0$ [$\gamma(\theta) \geq 0$] for all θ, and α is BA weak Bayes w.r.t. γ, and (ii) the probability assertions $P(\theta \in A|x) = \alpha(x, A)$ agree with a countably additive probability measure on $\mathcal{Q}$ for each $x \in X$.

Proof:

As in Theorem 4.2 strict [weak] coherence implies α is BA weak Bayes w.r.t. $\gamma > 0$ [$\gamma \geq 0$] where $\gamma \in \Gamma$. Assume that for some x' Peter's probability assertions are not countably additive. This means there exists a disjoint sequence $\{A_i(x')\}_{i=1}^{\infty}$ such that

$$\Theta = \sum_{i=1}^{\infty} A_i(x') \quad \text{while} \quad \sum_{i=1}^{\infty} \alpha_i(x') < 1. \tag{4.19}$$

The stake function $s \equiv 0$ except for $s_i(x') = 1$; $i = 1,2,\ldots$ is

$T(x)$ is Paul's total stake when x is observed, and if s is countably sup-bounded then Paul's total stake at x is uniformly bounded. Since for each x the sets are required to be disjoint, we have

$$\sum_{i=1}^{\infty} |[I_{A_i(x)}(\theta) - \alpha_i(x)]s_i(x)| \leq T(x) + \sup_{i=1,2,\ldots} |s_i(x)| \tag{4.18}$$

which is uniformly bounded in x. Thus, the expected payoff (4.2) is well defined for countably sup-bounded s. We define two spaces of expected payoff functions as in Section 4.3.

<u>Definition 4.13</u>.

Assuming a fixed specification, $\mathcal{G}^{\alpha*}$ will denote the space of <u>expected payoff functions</u>, $G_s^{\alpha}(\theta)$, as s ranges over all countably sup-bounded stake functions. $\mathcal{H}_p^{\alpha*}$ will denote the intersection of $\mathcal{G}^{\alpha*}$ and ℓ^p where the p.m.f. of the specification is of type (p).

Clearly $\mathcal{G}^{\alpha*}$ is a linear space in ℓ^{∞}, and $\mathcal{H}_p^{\alpha*}$ is a linear space which contains payoff functions $G_s^{\alpha}(\theta)$ for $s \in \mathcal{C}$ (see Lemma 4.1). Let us replace $\mathcal{H}_p^{\alpha}$ and $\mathcal{G}^{\alpha}$ with $\mathcal{H}_p^{\alpha*}$ and $\mathcal{G}^{\alpha*}$ in the definition of coherence (Def. 4.9) and consider a countably additive version of Theorem 4.1.

<u>Theorem 4.3</u>.

Assume that in Model II the p.m.f. is of type (p) where $p \in [1, \infty)$, and that Peter's strategy is α. If α is strictly [weakly] coherent then (i) there exists $\pi > 0$ $[\pi \geq 0]$ such that α is (q) weak Bayes w.r.t. π where $q = p/(p-1)$ if $p \in (1, \infty)$; $q = \infty$ if $p = 1$, and (ii) the probability assertions $P(\theta \in A|x) = \alpha(x, A)$ agree with a countably additive probability measure on $\mathcal{A}$ for each fixed $x \in X$.

finitely additive probability on $\mathcal{A}(x)$. It is of interest to note that in Model I it was not necessary to include the assumption that the sets $\{A_i(x)\}$ are closed under complementation to achieve the conclusion of Theorem 4.1 and 4.2. However, to compare Peter's probability assertions with σ-additive probabilities on the subsets of Θ, we introduce these modifications.

In Model II, $\mathcal{A}$ will denote the collection of <u>all</u> subsets of Θ, and a <u>specification</u> will denote the elements X, Θ, p: $X \times \Theta$, and $\mathcal{A}$. For each $x \in X$ and $A \in \mathcal{A}$, Peter is required to state his probability of $\theta \in A$. That is, Peter asserts "$P(\theta \in A|x) = \alpha(x, A)$" <u>for all</u> $A \in \mathcal{A}$. For each x, Paul will choose a <u>countable disjoint</u> sequence $\{A_i(x)\}_{i=1}^{\infty}$ from $\mathcal{A}$ and place stakes $s_i(x)$ on these sets. In summary, the significant changes in Model II are that, for each observed sample point x, Peter must assert the probabilities of <u>all</u> subsets of the parameter space from which Paul is allowed to choose a <u>countable disjoint</u> sequence for the purpose of betting.

These modifications, however, immediately complicate the existence of the expected payoff to Paul, and in fact, (4.2) might not be well defined even for sup-bounded stake functions. To avoid this difficulty we restrict Paul to countably sup-bounded strategies s. To simplify the notation let $\alpha_i(x) = \alpha(x, A_i(x))$.

<u>Definition 4.12</u>.

Paul's strategy s is said to be <u>countably sup-bounded</u> if s is sup-bounded and

$$(4.17) \qquad \sup_{x \in X} T(x) < \infty \quad \text{where} \quad T(x) = \sum_{i=1}^{\infty} |\alpha_i(x) s_i(x)|.$$

Theorem 4.2.

Assume that in Model I the p.m.f. of a specification is type (∞), and that Peter's strategy is α. If α is strictly [weakly] coherent, then there exists $\gamma \in \Gamma$ such that $\gamma(\theta) > 0$ $[\gamma(\theta) \geq 0]$ for all $\theta \in \Theta$, and α is BA weak Bayes w.r.t. γ.

Proof:

The proof is similar to Theorem 4.1 where we appeal to Theorem 3.6 instead of Theorem 3.5. □

We point out that $\alpha_i(x)$ is well defined by equation (4.16) whenever $\int p(x, \theta)d\gamma(\theta) > 0$. Even though Theorem 4.2 does not show γ to be a countably additive set function, it does imply $\gamma(\theta) > 0$ for all θ if α is strictly coherent. Since $p(x, \theta_0) > 0$ for some $\theta_0 \in \Theta$, we have $\int p(x, \theta)d\gamma(\theta) \geq p(x, \theta_0)\gamma(\theta_0) > 0$. Thus $\alpha_i(x)$ is well defined by (4.16) whenever α is strictly coherent. On the other hand if α is weakly coherent, then $\alpha_i(x)$ need not be well defined by (4.16) for all x and i; of course this is related to the weak coherence of α.

4.4 Coherence and Its Implications in a Modified Game, Model II.

We propose to modify the Peter-Paul game introduced in Section 4.2 and study σ-additive probability assertions. The new game will be called Model II. In the previous game Peter was required to post only a finite number of probabilities on the sets $\{A_i(x)\}$. Suppose we add the assumption that for each x, the sets $\{A_i(x)\}_{i=1}^{N(x)}$ are a field, say $\mathcal{G}(x)$. Since strict [weak] coherence implies Bayes' solutions, even in the case of type (∞) p.m.f.'s, we see that for each x strict [weak] coherence implies that Peter's probabilities agree with a

$$(4.14) \qquad \sum_{\theta\epsilon\Theta} [I_{A_i(x_j)}(\theta) - \alpha_i(x_j)]p(x_j, \theta)\pi(\theta) = 0$$

which is equivalent to (4.13). Since (4.14) holds for all $s_{ij} \epsilon C'$, α is (q) weak Bayes w.r.t. $\pi > 0$.

Proof for Weakly Coherent is Similar. $\square$

We draw attention to the fact that the p in the type (p) assumption is assumed finite in Theorem 4.1. If we generalized the statement of Theorem 4.1 to include $p = \infty$, then this would mean that π used in the Bayesian solution would be a member of ℓ^1. Up to now we have been unable to obtain such a result with Model I introduced in Section 4.2. The difficulty stems from the fact that the space congruent to the conjugate space of ℓ^∞ is not ℓ^1; on the contrary, it is the space of bounded finitely additive set functions. For this reason we modify the definition of (p) weak Bayes to BA weak Bayes.

Definition 4.11.

Let Γ be the class of all bounded finitely additive set functions on all possible subsets of Θ. Then Peter's α is said to be BA weak Bayes w.r.t. γ if and only if there exists $\gamma \epsilon \Gamma$ such that

$$(4.15) \qquad \gamma(\theta) \geq 0 \quad \text{for all} \quad \theta \epsilon \Theta,$$

and

$$(4.16) \qquad \int_{\theta\epsilon A_i(x)} p(x, \theta)d\gamma(\theta) = \alpha_i(x) \int_{\theta\epsilon\Theta} p(x, \theta)d\gamma(\theta)$$

holds for all $x \epsilon X$ and $i = 1,\ldots, N(x)$.

With this modified definition we state a result for type (∞) p.m.f.'s.

[positive] vector and (ii) the expanded intersection of $\mathcal{H}_p^{\alpha}$ and the set of all semi-positive [positive] vectors in ℓ^p is empty.

Suppose we consider type (p) p.m.f.'s where $p < \infty$. In Theorem 4.1 it will be proven that if α is strictly coherent, then α must be consistent with a Bayesian solution. Let us formally define this and present Theorem 4.1.

Definition 4.10.

For $p \in [1, \infty]$ Peter's α is said to be (p) weak Bayes w.r.t. π if and only if there exists a function π such that $\pi\colon \Theta \to [0, \infty)$, $\Sigma\pi(\theta)^p < \infty$ (written $\pi \in \ell^p$), and

$$(4.13) \qquad \sum_{\theta \in A_i(x)} p(x, \theta)\pi(\theta) = \alpha_i(x) \sum_{\theta \in \Theta} p(x, \theta)\pi(\theta)$$

holds for all $x \in X$ and $i = 1$ to $N(x)$. When $p = 1$, "(p) weak" may be replaced by "proper."

Theorem 4.1.

Assume that in Model I the p.m.f. of a specification is type (p) where $p \in [1, \infty)$, and that Peter's strategy is α. If α is strictly [weakly] coherent, then there exists $\pi > 0$ [$\pi \geq 0$] such that α is (q) weak Bayes w.r.t. π where $q = p/(p-1)$ if $p \in (1, \infty)$; $q = \infty$ if $p = 1$.

Proof for Strictly Coherent.

By definition strict coherence implies the expanded intersection of $\mathcal{H}_p^{\alpha}$ and the semi-positive vectors in ℓ^p is empty. Hence, by Theorem 3.5 there exists a π such that $\pi > 0$, $\pi \in \ell^q$, and $< \pi, G_s^{\alpha} > = 0$ for all $G_s^{\alpha} \in \mathcal{H}_p^{\alpha}$. Payoffs for stake functions in C' are in $\mathcal{H}_p^{\alpha}$, and so for each $s_{ij} \in C'$, we have set of semi-positive vectors have a non-empty expanded intersection (see Definition 3.1). This is true even though the intersection of $\mathcal{G}^{\alpha}$ and the set of semi-positive vectors is empty.

To see that the expanded intersection is non-empty let P be equal to the set of semi-positive vectors in ℓ^1. Let $\varepsilon > 0$ be arbitrary but fixed. Then there exists a positive integer J such that $K_1 J^{-2} < \varepsilon$. Let Paul specify a stake function s which is identically zero except for $s(J) = 2J^2$. The expected payoff is

$$(4.12) \qquad G_s^{\alpha}(\theta) = \begin{cases} -K_1 J^{-2} & \text{if } \theta = 1 \\ K_2 & \text{if } \theta = 2 \\ J^2 & \text{if } \theta = N + 2 \\ 0 & \text{elsewhere.} \end{cases}$$

For $p_0 = (0, K_2, 0, 0, \ldots)$ and p with all elements zero except for J^2 in the $(J+2)$th position, we have $\|p_0 + p - G_s^{\alpha}\|_1 < \varepsilon$ which implies the expanded intersection of P and $\mathcal{H}_1^{\alpha}$ is non-empty.

From an inference view-point this may be undesirable since even though there is no semi-positive payoff function, there is a sequence of payoff functions which comes arbitrarily close (in ℓ^1 norm) to semi-positive vectors. Furthermore, if we let $G_s^{\alpha}(+)$ and $G_s^{\alpha}(-)$ be the positive and negative parts of G_s^{α}, then $\|G_s^{\alpha}(+)\|_1$ can be made arbitrarily large while $\|G_s^{\alpha}(-)\|_1$ goes to zero. To avoid this difficulty we suggest the following definition of coherence.

Definition 4.9.

For a type (p) p.m.f., Peter's α is said to be strictly [weakly] coherent if and only if (i) $\mathcal{G}^{\alpha}$ contains no semi-positive

Example 4.2.

Let $X = \Theta = \{1, 2, 3, \ldots\}$, $A = \{1\}^c$, $\mathcal{a} = \{A\}$, $N(x) \equiv 1$, and

$$p(x, \theta) = \begin{cases} K_1 x^{-4} & \text{for } \theta = 1 \\ K_2 x^{-2} & \text{for } \theta = 2 \\ 1 & \text{for } \theta = 3, 4, \ldots \text{ if } x = \theta - 2 \\ 0 & \text{elsewhere} \end{cases}$$

where K_1 and K_2 are equal to $(\Sigma i^{-4})^{-1}$ and $(\Sigma i^{-2})^{-1}$ respectively. Suppose Peter's strategy is $\alpha(x) \equiv \frac{1}{2}$. Then for a stake function s_n which is identically zero except for $s(n) = 2$, the expected payoff is

$$G^{\alpha}_{s_n}(\theta) = \begin{cases} -K_1 n^{-4} & \text{if } \theta = 1 \\ K_2 n^{-2} & \text{if } \theta = 2 \\ 1 & \text{if } \theta = n + 2 \\ 0 & \text{elsewhere.} \end{cases} \tag{4.10}$$

Clearly G^{α}_{s} is not semi-positive for a stake function s which contains <u>any</u> negative stake, and for any non-trivial non-negative stake function s, $G^{\alpha}_{s}(1) < 0$. Thus $\mathcal{G}^{\alpha}$ contains no semi-positive vector and α is strictly coherent by the above definition.

Suppose there is a positive vector orthogonal to $\mathcal{G}^{\alpha}$. This implies π is orthogonal to $G^{\alpha}_{s_n}$ which means

$$-K_1\pi_1 + K_2\pi_2\, n^2 + n^4\pi_n = 0. \tag{4.11}$$

Since $\pi_n > 0$ we have $K_1\pi_1 > n^2 K_2\pi_2$ for all n which is impossible because $K_2\pi_2 > 0$. Hence, there is no positive orthogonal vector.

The failure of this version of coherence to imply that α is Bayes w.r.t. $\pi > 0$ hinges on the fact that $\mathcal{H}^{\alpha}_1$ and the

These elements determine the specification. Clearly $p(x, \theta) \in \ell^1$ for all $x \in X$, and $\mathcal{H}_p^\alpha \subset \ell^1$. From (4.2) we have

$$G_s^\alpha(\theta) = \tfrac{1}{2}\{[I_A(\theta) - \alpha_1(\theta)]s_1(\theta) + [I_A(\theta) - \alpha_1(\theta+1)]s_1(\theta+1)\}. \tag{4.8}$$

Let s' denote a stake function which is zero for odd x and one for even x. Clearly s' is sup-bounded and

$$G_{s'}^\alpha(\theta) = \begin{cases} \tfrac{1}{2}[1 - \alpha_1(\theta+1)] & \text{if } \theta \in A \\ -\tfrac{1}{2}\,\alpha_1(\theta) & \text{if } \theta \notin A. \end{cases} \tag{4.9}$$

Thus, $\max\{|G_{s'}^\alpha(\theta)|, |G_{s'}^\alpha(\theta+1)|\} \geq 1/4$ for $\theta \in A$, and for finite p, $G_{s'}^\alpha(\theta) \notin \ell^p$ which implies $\mathcal{G}^\alpha \not\subset \ell^p$.

Now, let us attempt to find a suitable definition of coherence for countably infinite spaces. Our goal is an extended version which can be used in considering necessary and sufficient conditions for a Bayesian solution. A natural extension of the previous definition of coherence is to replace the space of expected payoffs $\mathcal{G}^\lambda$ with $\mathcal{G}^\alpha$. The extended definition would be: for a fixed specification, α is said to be strictly coherent if and only if $\mathcal{G}^\alpha$ contains no semi-positive vector. An analogous result based on this definition is: if α is strictly coherent, then α is Bayes w.r.t. π where $\pi > 0$. The proof of this type of a result relies on the existence of a <u>strictly</u> positive vector π which is orthogonal to the space of payoff functions $\mathcal{G}^\alpha$. This conjecture, however, is false as we see from the next example. Here we have a strategy α for which $\mathcal{G}^\alpha$ contains no semi-positive vector, but there is no positive orthogonal vector.

Lemma 4.1.

$G^\alpha_s(\theta) \in \mathcal{H}^\alpha_p$ for all $s \in \mathcal{C}$. Thus, $\mathcal{H}^\alpha_p$ is not empty.

Proof:

For any $s_{ij} \in \mathcal{C}'$ the expected payoff is

$$(4.5) \qquad G^\alpha_{s_{ij}}(\theta) = \begin{cases} [1 - \alpha_i(x_j)]p(j, \theta) & \text{if } \theta \in A_i(x_j) \\ -\alpha_i(x_j)p(j, \theta) & \text{if } \theta \notin A_i(x_j), \end{cases}$$

and so

$$(4.6) \qquad \|G^\alpha_{s_{ij}}(\theta)\|^p_p = |1 - \alpha_i(x_j)|^p \sum_{\theta\in A_i(x_j)} p^p(j,\theta) + \alpha^p_i(x_j) \sum_{\theta\notin A_i(x_j)} p^p(j,\theta)$$

$$\leq \|p(j, \theta)\|^p_p .$$

Thus, for any $s \in \mathcal{C}$

$$(4.7) \qquad \|G^\alpha_s(\theta)\|^p_p \leq \Sigma|s_i(x)|^p \, \|p(x, \theta)\|^p_p$$

where the sum is taken over a finite number of $x \in X$, and for each x over a finite number of i. All terms in the sum are finite by the definition of $\mathcal{C}$ and type (p). The conclusion follows by considering the definition of $\mathcal{H}^\alpha_p$. □

Many p.m.f.'s are of type (p) where the p is finite, and hence, $\mathcal{H}^\alpha_p$ is a non-empty linear manifold contained in ℓ^p. The significance of the p will be seen later, but in passing, we point out that even though $\mathcal{H}^\alpha_p \subset \ell^p$ for finite p, $\mathcal{G}^\alpha$ need not be contained in ℓ^p space for any finite p, as the next example shows.

Example 4.1.

Let $X = \Theta = \{0, \pm 1, \pm 2, \text{etc.}\}$, $A = \{\pm 1, \pm 3, \text{etc.}\}$, $\mathcal{Q} = \{A\}$, $N(x) \equiv 1$, and

$$p(x, \theta) = \begin{cases} 1/2 & x = \theta \\ 1/2 & x = \theta + 1 \\ 0 & \text{elsewhere} . \end{cases}$$

Definition 4.5.

Assuming a fixed specification, $\mathcal{G}^\alpha$ will denote the space of payoff functions $G_s^\alpha(\underset{\sim}{\theta})$ as s ranges over all sup-bounded stake functions.

It is easy to verify that $\mathcal{G}^\alpha$ is a linear manifold contained in ℓ^∞. Now, let us consider a second space of expected payoff functions which is defined in terms of a property of the p.m.f. in the specification.

Definition 4.6.

A function on $X \times \Theta$ which is a p.m.f. on X for every θ is said to be of type (p) if $p(x, \underset{\sim}{\theta}) \in \ell^p$ for all $x \in X$. That is, $\Sigma_\Theta |p(x, \theta)|^p < \infty$ for all $x \in X$.

Definition 4.7.

If the p.m.f. of the specification is of type (p), then let $\mathcal{H}_p^\alpha$ denote the intersection of $\mathcal{G}^\alpha$ and ℓ^p. Note if $p = \infty$, then $\mathcal{H}_p^\alpha$ is equal to $\mathcal{G}^\alpha$.

Clearly $\mathcal{H}_p^\alpha$ is a linear manifold. We will show that $\mathcal{H}_p^\alpha$ is non-empty after defining two classes of stake functions.

Definition 4.8.

Let $\mathcal{C}$ be the class of sup-bounded stake functions such that $s_i(x) = 0$ except for a finite number of x and i. Let $\mathcal{C}'$ be a class of stake functions where

$$(4.4) \qquad \mathcal{C}' = \left\{ s_{ij} : s_i(x_k) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{elsewhere} \end{cases} \text{ for } \begin{array}{l} j = 1,2,\ldots \\ i = 1 \text{ to } N(x_j) \end{array} \right\} .$$

That is for any $s_{ij} \in \mathcal{C}'$, the values of s_{ij} are zero except for a single x and i.

and for each x, the payoff is the sum of payoffs in (4.1) over i. We assume $s_i(x)$ is finite so that this sum is well defined. As in Chapter II we consider an expected payoff to Paul over the sample space.

Definition 4.3.

The expected payoff to Paul as a function of θ, α, s will be denoted by $G_s^{\alpha}(\theta)$ and is computed by

$$G_s^{\alpha}(\theta) = \sum_X \sum_{i=1}^{N(x)} [I_{A_i(x)}(\theta) - \alpha_i(x)]\, s_i(x)\, p(x, \theta). \tag{4.2}$$

The question of $G_s^{\alpha}(\theta)$ being well defined arises immediately. It is easily seen that unbounded stakes may leave (4.2) undefined regardless of the choice of α. For this reason it is convenient to restrict Paul to sup-bounded stakes.

Definition 4.4.

Pauls stakes s are said to be sup-bounded if and only if

$$\sup_{x \in X} \{ \max_{i=1,\dots,N(x)} |s_i(x)| \} = M < \infty. \tag{4.3}$$

This completes the description of the Peter-Paul game called Model I.

4.3 Coherence and Its Implications in Model I.

In Section 4.2 an expected payoff to Paul was introduced, and its existence for $\theta \in \Theta$ was discussed for sup-bounded stake functions. In this section we examine properties of $G_s^{\alpha}(\theta)$ as a vector in θ. To accommodate the study we consider two spaces of payoff functions and notice that one of these is contained in ℓ^p for suitable p. A portion of the ℓ^p theory developed in Chapter III is applied to this space yielding two main results. Let us define one of the spaces.

of odds. To be exact, Peter is confronted with a specification and an observed sample point x. If the master of ceremonies has not used his option, he must choose $N(x)$ sets from $\mathcal{Q}$ and assert the probabilities of these sets. That is, for $i = 1,2,\ldots, N(x)$ he asserts "the probability that $\theta \in A_i(x)$ is equal to $\alpha_i(x)$" and must be willing to accept bets from Paul where the payoff is determined by the stake, $\alpha_i(x)$, and the truth or falsity of $\theta \in A_i(x)$. Stated in terms of odds, Peter offers Paul odds of $(1-\alpha_i(x))/\alpha_i(x)$ to 1 against the occurrence of $\theta \in A_i(x)$. If the master of ceremonies has used his option, then Peter must state only the probabilities.

Definition 4.2.

Peter's strategy is denoted by α whether or not the master of ceremonies used his option. Thus, if he has not, α denotes the rule specifying the sets and the probabilities; if he has, α denotes only the probabilities.

With the knowledge of the specification and α, Paul determines a strategy, denoted by s, which governs the betting. That is, Paul bets with Peter on the event $\theta \in A_i(x)$ risking $s_i(x)\alpha_i(x)$ to win $(1-\alpha_i(x))s_i(x)$ if $\theta \in A_i(x)$. Notice, $s_i(x)$ is not the stake on $A_i(x)$ in usual betting terminology although we may refer to $s_i(x)$ as such. The value of $s_i(x)$ can be positive, negative, or zero; this allows Paul to bet on or against $\theta \in A_i(x)$. A stake $s_i(x) = 0$ could be interpreted as "no bet on $A_i(x)$." After the stakes are fixed, the master of ceremonies reveals θ, and they settle up as follows. For each $i = 1$ to $N(x)$, the payoff to Paul is

$$(4.1) \qquad \begin{array}{ll} (1-\alpha_i(x))s_i(x) & \text{if } \theta \in A_i(x) \\ -\alpha_i(x)s_i(x) & \text{if } \theta \notin A_i(x), \end{array}$$

The role of the master of ceremonies is four-fold. First, he specifies the collection of subsets of Θ from which Peter can choose when assigning probabilities; we denote this collection by $\mathcal{A}$. For each $x \in X$, Peter is required to choose $N(x)$ sets from $\mathcal{A}$, and for each set, state his probability. Thus, the second task of the master of ceremonies is to specify $N(x)$. In Model I we assume that $N(x)$ is an integer, and $N(x) \leq N$ for all $x \in X$ where $N < \infty$.

The third task of the master of ceremonies is optional. We leave to his discretion the power to select the $N(x)$ sets on which Peter posts probabilities. For instance, suppose $N(x) \equiv 1$; then the master of ceremonies could ask Peter to post probabilities on A for all x, where $A \subset \Theta$. This means that $A_1(x) \equiv A$ and $\mathcal{A} = \{A\}$. When the master of ceremonies is using his option, it will be mentioned specifically; otherwise, we assume Peter has the responsibility to choose the $N(x)$ sets from $\mathcal{A}$.

Definition 4.1.

The items X, Θ, $p: X \times \Theta$, $\mathcal{A}$, $\{N(x)\}$, and the option are called a specification.

The fourth task of the master of ceremonies is to select $\theta \in \Theta$ at his leisure, and then to choose x according to the p.m.f. based on θ. As in previous games the x is revealed to Peter and Paul, and after the stakes are placed, θ is disclosed so that Peter and Paul may settle up.

Peter's role in the game is similar to that of Chapter II with the major difference being that he is posting probabilities instead member of ℓ^p. If $p > 1$ then the prior may be improper.

In Chapter II we introduced a Peter-Paul game for finite spaces which was phrased in terms of odds, and we noticed that infinite odds complicated the results stated in terms of coherence. This chapter deals with probability assertions; however, one can translate them into odds and apply the method of Section 2.5 to handle infinite odds. As noted previously, the main disadvantage in using odds is that the space of expected payoffs is a half space when infinite odds occur rather than a linear subspace.

Section 4.2 describes one of the Peter-Paul games of this chapter called Model I, and Section 4.3 investigates the implications of coherence in this model. Theorems 4.1 and 4.2 imply that if Peter is to be coherent then he must give probability assertions which agree with a finitely additive probability for each x; furthermore, they must be consistent with a Bayes' solution for some prior π. In the interest of investigating σ-additive probability measures, Section 4.4 introduces modifications of Model I to give Model II. It is shown that in this model coherence implies that Peter's probability assertions must agree with a σ-additive probability for each observed sample point x.

4.2 Model I for Countably Infinite Spaces.

The goal of this section is to describe a Peter-Paul game called Model I in which the sample and parameter spaces are countably infinite. Let $\{x_i\}_{i=1}^{\infty}$ and $\{\theta_j\}_{j=1}^{\infty}$ be enumerations of X and Θ respectively, and consider a fixed function $p(x, \theta)$ which is a probability mass function (p.m.f.) on X for each θ (i.e., $\Sigma_X\, p(x, \theta) = 1$ for all θ). As in Chapter II, three persons participate in the game: the master of ceremonies, Peter, and Paul.

CHAPTER IV

Countably Infinite Parameter and Sample Spaces

4.1 Introduction.

In Chapter II a Peter-Paul game for finite parameter and sample spaces was discussed in terms of odds. In the present chapter a switch is made from odds to probability assertions, and concepts are generalized to include infinite parameter and sample spaces. The spirit of games based on probability assertions is stated by Buehler (1959):

> Many, though not all, problems of inference lead to assertions of the type "The probability that A is true is equal to α," or, "$P(A) = \alpha$." One may ask whether the person making this assertion should be willing to bet that A is true, risking an amount α to win $1 - \alpha$, and should be equally willing to bet that A is false, risking $1 - \alpha$ to win α, against an opponent who has exactly the same information as he and who is allowed to choose either side of the wager. The affirmative answer will not be defended here, but its consequences will be examined.

Continuing the theme, Cornfield (1969) defines a property called coherence and applies it to a similar situation where the inference is a statement about the probability of a subset of a parameter space given an observed sample point. The sample and parameter spaces of Cornfield (1969) are finite, and a typical result says that the inference is coherent if and only if it agrees with Bayes' rule for some prior π. Since the spaces are finite, the prior can be normed so that it is a proper prior. The introduction of countably infinite spaces, however, suggests that the definition of coherence could be altered and allows the possibility of using improper priors. In this chapter we will modify the definition of coherence for countably infinite spaces, and show that it implies a Bayes' solution where the prior is a

Theorem 3.10.

Let M be a linear manifold and P be the set of all semi-positive functions in L^{∞}. Let $\{B_i\}_{i=1}^{\infty}$ be an arbitrary fixed sequence of $\mathcal{B}$ measurable sets of $\mathcal{Y}$. If the expanded intersection of M and P is empty, then there exists a bounded finitely additive set function γ such that $\gamma(B_i) > 0$ for all i, and (3.80) and (3.82) hold.

With these theorems established we proceed to study Peter-Paul games for countable and continuous spaces.

In the cases where $\gamma(B_i) > 0$ or $< g, \chi_{B_i} > > 0$ for a sequence of B_i, it is easy to show that the expanded intersection through the elements χ_{B_i} is empty. Thus, many of the previous results could be stated in the form of necessary and sufficient conditions for separation. However, our interest lies mainly in the sufficiency.

Proof:

As in the proof of Theorem 3.4, repeatedly apply Lemma 3.3 instead of Theorem 3.3.

For $p = \infty$ we have the following analogous results. They are stated separately since continuous linear functionals on L^∞ are represented by bounded finitely additive set functions.

Lemma 3.4.

If M is a linear manifold and N is a non-empty convex cone in L^∞ such that the expanded intersection of M and N is empty, then for any $n_0 \in N$ there exists a bounded finitely additive set function γ defined on $\mathfrak{B}$ such that

$$\int m(y)\,\gamma(dy) = 0 \qquad \text{for all } m \in M \tag{3.80}$$

$$\int n_0(y)\,\gamma(dy) > 0 \tag{3.81}$$

$$\int n(y)\,\gamma(dy) \geq 0 \qquad \text{for all } n \in N. \tag{3.82}$$

Proof:

Apply Theorem 1 of Dunford and Schwartz (1958, p. 258) to Theorem 3.3. □

Theorem 3.9.

Let M be a linear mainfold and P be the set of all positive functions in L^∞. If the expanded intersection of M and P is empty, then there exists a bounded finitely additive set function γ such that $\gamma(B) \geq 0$ for all $B \in \mathfrak{B}$, and (3.80) and (3.82) hold.

Proof:

Apply Lemma 3.4 and proceed as in the proof of Theorem 3.7. □

Proof:

By Lemma 3.3, there exists $g \in L^q$ such that $\mu[g \neq 0] > 0$, (3.74) holds, and

(3.75) $\quad < g, p > \geq 0$ for all $p \in P$.

Let $B = \{y: g(y) < 0\}$, and let $\mathcal{Y} = \sum_{i=1}^{\infty} Y_i$ where $\mu(Y_i) < \infty$. Let $C_i = B \cap Y_i$ and $\chi_{C_i}(y)$ be the indicator function of C_i. Then χ_{C_i} is in the closure of P, and thus, there is a sequence $\{p_{ji}\}_{j=1}^{\infty}$ where $p_{ji} \in P$ such that $p_{ji} \to \chi_{C_i}(y)$. By the continuity of the continuous linear functional defined by g, we must have

(3.76) $\quad < g, \chi_{C_i} > \geq 0.$

Thus $\mu(C_i) = 0$ and since this holds for all i, we have $\mu(B) = 0$. This implies that g is semi-positive. $\square$

Even though we are not able to show the existence of a strictly positive orthogonal function for continuous spaces, we can show that $\int_{B_i} g du > 0$ for a countable sequence of sets $B_i \in \mathcal{B}$. This is the spirit of the next theorem.

Theorem 3.8.

Let M be a linear maniford and P be the set of all semi-positive functions in L^p where $p \in [1, \infty)$. Let $\{B_i\}_{i=1}^{\infty}$ be an arbitrary fixed sequence of $\mathcal{B}$ measurable sets of $\mathcal{Y}$. If the expanded intersection of M and P is empty, then there exists $g \in L^q$ such that

(3.77) $\quad < g, m > = 0 \quad$ for all $m \in M$

(3.78) $\quad < g, n > \geq 0 \quad$ for all $n \in P$

(3.79) $\quad < g, \chi_{B_i} > \;>\; 0 \quad$ for all B_i.

(3.72) $\quad < f, n_0 > \; > 0$

(3.73) $\quad < f, n > \; \geq 0$ for all $n \in N$.

Proof:

Theorem 3.3 shows the existence of a continuous linear functional with the desired properties. Apply Theorem 1, p. 286, if $p \in (1, \infty)$, or Theorem 5, p. 289, if $p = 1$, of Dunford and Schwartz (1958) to get the conclusion. □

The results for the continuous spaces, L^p, differ from the discrete spaces, ℓ^p. In the discrete spaces we showed that if a linear manifold has a non-empty expanded intersection with the semi-positive vectors, then there exists a positive vector orthogonal to the linear manifold. In L^p we are only able to show that if a linear manifold has non-empty expanded intersection with the positive functions, then there is a semi-positive function which is orthogonal to the linear manifold. Currently I believe that the result mentioned for ℓ^p does not hold for L^p. The proof for ℓ^p uses the property that there exists a countable basis for the convex cone which was the set of semi-positive vectors. This property does not generally hold for arbitrary L^p spaces. Our results for L^p spaces are stated in the following theorems; in the first $p \in [1, \infty)$ and in the second $p = \infty$.

Theorem 3.7.

Let M be a linear manifold in L^p and P be the set of all positive functions in L^p where $p \in [1, \infty)$. If the expanded intersection of M and P is empty, then there exists $g \in L^q$ such that g is semi-positive and

(3.74) $\quad < g, m > \; = 0$ for all $m \in M$.

3.4 Separation Theorems for L^p Space.

Let $(\mathcal{Y}, \mathcal{B}, \mu)$ be a measure space with a σ-finite measure μ. Let f be a measurable function on $\mathcal{Y}$, and define the norms

$$(3.69) \qquad \|f\|_p = \{\int |f|^p d\mu\}^{1/p} \quad \text{for} \quad p \in [1, \infty)$$

$$\|f\|_\infty = \text{ess. sup}|f(y)| = \inf\{M{:}\mu\{y{:}f(y) > M\} = 0\}.$$

Let L^p be the space of all measurable f on $\mathcal{Y}$ such that $\|f\|_p < \infty$.

Definition 3.4.

A function f in L^p is said to be semi-positive if $f \geq 0$ a.e. $[\mu]$ and $\mu\{y{:}f(y) > 0\} > 0$. A function f in L^p is said to be positive if $f > 0$ a.e. $[\mu]$.

Throughout this section $< g, f >$ will denote the inner product between $g \in L^p$ and $f \in L^q$, where q is defined by (3.64). The inner product is computed by

$$(3.70) \qquad < g, f > = \int gfd\mu ,$$

and by Hölder's inequality it is well defined. Now, we consider the representation of the continuous linear functionals of Theorem 3.3 for L^p spaces where $p \in [1, \infty)$.

Lemma 3.3.

If M is a linear manifold in L^p, $p \in [1, \infty)$, and N is a non-empty convex cone in L^p such that the expanded intersection of M and N is empty, then for any $n_0 \in N$ there exists a $f \in L^q$ such that

$$(3.71) \qquad < f, m > = 0 \quad \text{for all} \quad m \in M$$