A STUDY OF DYNAMIC PROGRAMMING AND GAMBLING SYSTEMS

By

N. D. PRABHAKAR

A thesis submitted to the Indian Statistical Institute in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Calcutta
July 1972

Acknowledgements

My biggest debt of gratitude is to Dr. Ashok Maitra under whose guidance this thesis was prepared. His seminars on dynamic programming, analytic sets etc., provided the initial inspiration for me to work in these areas and over the years his active help, valuable suggestions and kind encouragement have been generously given to me. I wish to express my heartfelt thanks to him.

I am grateful to Dr. C. R. Rao, Director of the Research and Training School of the Indian Statistical Institute for the research facilities offered at the Institute.

I take the opportunity to thank all my colleagues with whom I have had useful discussions.

Finally, it is a pleasure to thank Mr. Gour Mohon Das and Mr. Apurba Guha for the untiring patience and meticulous care with which they did the stenciling and cyclostyling of the material.

N. D. PRABHAKAR

C o n t e n t s

INTRODUCTION

The body of methods known as Dynamic Programming (d.p.) was developed by Bellman and was successfully applied by him to solve practical problems in diverse fields (see [2] and the papers cited there.). These methods revolved around an intuitively appealing principle which Bellman called the principle of optimatity. A general formulation of d.p. problem was given by Blackwell [3, 5] which is narrower in scope than Bellman's but includes many important applications [12] and offers a proper framework for asking the many interesting and mathematically sophisticated questions that arise in d.p. . Strauch, Brown, Maitra, Furukawa, Dantas, Hinderer etc., have investigated some of these questions in detail. Using slightly different terminology from [3], a few other authors like Derman, Veinott, Ross and Fisher have also worked on similar problems.

A general theory of gambling has been developed by Dubins and Savage in their book 'How to gamble if you must' ([9]). Some problems which arise from their work have been investigated by Strauch, Sudderth, Freedman, Ornstein etc. It should be mentioned that Dubins and Savage employ finitely additive probability theory and this, in itself, has led to much interest in the study of finitely additive measures.

In Chapter I of this thesis we use slight modificati of the gambling terminology to discuss simultaneously both d.p. and gambling problems over discrete time. After intr ducing the necessary definitions (section 1) we state the general problem precisely (section 2), give examples (section 3) and prove certain general results regarding the optimal reward functions and optimality equations (sections 4 and 5). In section 6 we collect some facts from measure theory which are used in the next three sections where we study measurable gambling systems. In many of the sections we frequently make digressions into related questions of interest. Consequently these sections contain results whi we do not use later.

In chapter II we consider problems over continuous time. As a fruitful theory of gambling over continuous ti is not available yet, we adopt the d.p. terminology of [3 and discuss only d.p. over continuous time. In sections to 3 we discuss an extension of the discounted d.p. probl in which actions can be taken at discrete, randomly chosen points of time and states vary continuously over time. Our methods are essentially those of [3]. In section 3 ~~and 5~~ we allow both states and actions to vary continuously over time assuming that the states $x(t)$ (in R^m) satisfy a stochastic differential equation specified by known diffusion coefficients. The resulting problem is essentially a problem of stochastic optimal control. We prove results mainly relating to measurability.

In both chapters and especially in chapter II there are quite a few questions that remain unanswered and we hope these and the large number of interesting open problems listed in the book of Dubins and Savage would attract larger contributions to these areas.

CHAPTER I

0. The Problem:

Our general problem can be described loosely as follows: Suppose we are in a gambling house wherein the rules regarding when we can play and what games we can play depend on a system with a state space F. Periodically, say once a day, we observe the state of the system - for instance, the amount of money with us on that day. Suppose the observed states over the first n days are $f_1, \ldots, f_n$, then the rules of the house specify exactly one of the three alternatives - (i) we must quit the house (ii) we must stay in the house and play (iii) we have the choice to decide whether to play or to quit. In case we quit on the n^{th} day, either by choice or by compulsion, we receive an amount $u(f_1, \ldots, f_n)$. If, on the other hand, we stay, we must choose a probability measure γ on F from a given set of probability distributions on F. The outcome of such a 'play' is that the system moves to a new state f_{n+1} distributed according to γ. It may happen that the system moves successively through the states $f_1, f_2, \ldots$ without our quitting the house on any day; in such a case we receive an amount $u(f_1, f_2, \ldots)$. A specification of what we should do in each conceivable situation is called a policy. The problem then is to find a policy which will maximize our expected reward. Sometimes there may be further constraints on the policies that we can use, in which cases we must try to maximize our expected reward by using only those policies that satisfy the given constraints. In the next section we introduce the definitions and notations which will enable us to state the problem more precisely. Note that the above problem is nonstationary over time while the gambler's problem in the formulation of Dubins and Savage [9], is stationary. But as pointed out in Chapter 12 [9], such a nonstationary problem may be regarded as a stationary one by modifying the state space F. We find it more convenient to work with the nonstationary problem as such, making suitable modifications in the definitions of terms like 'policy available at f'. Also note that we have allowed the possibility that the gambler plays indefinitely. This enables us to treat many infinite stage problems, like dynamic programming, in our framework more naturally.

1. Definitions and Notations:

As most of the terms we use are merely modifications of corresponding terms used in [9], we suppress their interpretations. Let F be a nonempty set. A partial history over F is a finite string of elements from F and a history over F is a unilateral sequence of elements from F. Let $W(F)$ denote the space of all partial histories and $H(F)$ that of all histories over F. So $W(F) = \bigcup_{n=0}^{\infty} F^n$ and $H(F) = F^N$. The empty partial history which is the only element of F^0 is denoted by $\emptyset$. If $x = (f_1, \ldots, f_m)$ and $x' = (f_1', \ldots, f_n')$ are partial histories, we write xx' or (x, x') for the catenated partial history $(f_1, \ldots, f_m, f_1', \ldots, f_n')$. If $h = (f_1'', f_2'', \ldots)$ is a history then we write xh or (x, h) for the history $(f_1, \ldots, f_m, f_1'', f_2'', \ldots)$; also we let h/n denote the partial history $(f_1'', \ldots, f_n'')$ for any $n \geq 0$ and sometimes we write h/∞ for h. For any function g on $W(F)$, g_n will denote its restriction to F^n. If g is any function on $W(F)$ or on $H(F)$ and if $x \in W(F)$, then we denote by $g[x]$ the function with same domain defined by $g[x](\cdot) = g(x, \cdot)$.

A gamble on F is a finitely additive, nonnegative and normalized set function defined on the set $S(F)$ of all subsets of F. Let $G(F)$ denote the set of all gambles on F. A strategy over F is a function on $W(F)$ into $G(F)$. If $\underline{\underline{F}}$ is a σ-algebra of subsets of F, we denote by $W(\underline{\underline{F}})$ and $H(\underline{\underline{F}})$ the σ-algebras $\sum_{n=0}^{\infty} \underline{\underline{F}}^n$ and $\underline{\underline{F}}^N$ respectively. A strategy σ over F is said to be $\underline{\underline{F}}$ measurable if (a) for all x in $W(F)$,

$\sigma(x, \cdot)$ is countably additive on $\underline{\underline{F}}$ and (b) for all A in $\underline{\underline{F}}$, $\sigma(\cdot, A)$ is $W(\underline{\underline{F}})$ measurable - i.e. $\sigma_n(\cdot, A)$ is a $\underline{\underline{F}}^n$ measurable function for each $n \geq 0$. Such a σ induces a countably additive probability measure $p_\sigma = p_\sigma(\underline{\underline{F}})$ on $H(\underline{\underline{F}})$.

Let σ be a strategy over F and g be a function on $H(F)$ into the extended real line $\bar{R} = [-\infty, \infty]$. g is said to have σ structure at most 0 if, for some σ-algebra $\underline{\underline{F}}$ of subsets of F, σ is $\underline{\underline{F}}$ measurable and g is a $H(\underline{\underline{F}})$ measurable function whose integral under $p_\sigma(\underline{\underline{F}})$ exists. We observe that if g is also $H(\underline{\underline{F}}')$ measurable and σ is $\underline{\underline{F}}'$ measurable, then $\int g \, d p_\sigma(\underline{\underline{F}}) = \int g \, d p_\sigma(\underline{\underline{F}}')$; hence we can let $\int g \, d\sigma$ denote $\int g \, d p_\sigma$. Proceeding by transfinite induction, for each ordinal $\alpha > 0$, we say that g has σ structure at most α if, we can find a subset N of F with $\sigma_0(N) = 0$ such that for each $f \notin N$, $g[f]$ has $\sigma[f]$ structure at most $\beta(f)$ for some $\beta(f) < \alpha$ and moreover $\int_{F-N} (\int g[f] \, d\sigma[f]) \, \sigma_0(df)$ exists. We then define $\int g \, d\sigma$ to be $\int_{F-N} (\int g[f] d\sigma[f]) \, \sigma_0(df)$. It is a routine matter to check by induction on α that $\int g \, d\sigma$ is well-defined. g is said to be σ structured if it has σ structure at most α for some ordinal α; the smallest such α is called the σ structure of g. The structure of an inductively integrable g in the sense of chapter 2 [9], will be referred to as the Eudoxus structure of g. If g is inductively integrable then it is σ structured for any strategy σ and its σ structure is at most the Eudoxus structure of g.

A stopping function (s.f.) over F is a function on W(F) into the set $\{0, 1\}$. Using such s.f. a over F means that if we are in a gambling house with state space F, we decide after observing a partial history x that we shall quit the house if $a(x) = 0$ and stay in it if $a(x) = 1$. A stop rule over F is a function t on H(F) into the set $\bar{N} = \{0, 1, 2, \ldots \infty\}$ such that, for every h, h' in H(F), $t(h) = t(h')$ whenever $h|t(h) = h'|t(h)$. Every s.f. a determines a stop rule t_a defined by: $t_a(h) = \infty$ if $a(h/n) = 1$ for all $n \geq 0$, otherwise $t_a(h)$ is the smallest $m \geq 0$ such that $a(h/m) = 0$. Conversely, for any stop rule t we can find at least one s.f. a such that $t = t_a$. A s.f. a is finite if t_a is finite.

The set of all strategies over F and that of all stopping functions over F are denoted by $\Sigma(F)$ and $\Lambda(F)$ respectively. A policy over F is any element of $\Sigma(F) \times \Lambda(F)$. Suppose u is a function on $W(F) \cup H(F)$ into $\bar{R}$, a is a s.f. over F and $x \in W(F)$. We denote by $R_u(a, x)$ the function on H(F) into $\bar{R}$ defined by $R_u(a, x)(h) = u(x, (h/t_{a[x]}(h)))$.

Observe that if $a(x)$ is 0 then $R_u(a, x) \equiv u(x)$ and if $a(x)$ is 1 then $R_u(a, x)[f] = R_u(a, xf)$ for all $f \in F$. If $\pi = (\sigma, a)$ is a policy, we let $I_u(\pi)(x)$ denote the integral of $R_u(a, x)$ with respect to $\sigma[x]$ whenever the latter is defined - i.e., if $R_u(a, x)$ is $\sigma[x]$ structured. It is then clear that $I_u(\pi)(x) = u(x)$ if $a(x) = 0$ and $I_u(\pi)(x) = \int_{F-N} I_u(\pi)(xf)\sigma(x, df)$ if $a(x) = 1$, where $\sigma(x)(N) = 0$.

2. Gambling System:

By a gambling system (g.s.) we mean a 4-tuple $\underline{G} = (F, \Gamma, T, u)$ where (i) F is a nonempty set; (ii) Γ is a correspondence which associates with each x in $W(F)$ a nonempty set $\Gamma(x)$ of gambles over F; (iii) T is a function on $W(F)$ into the set $\{0, 1, 2\}$ and (iv) u is a function on $W(F) \cup H(F)$ into $\bar{R}$. F is called the fortune space or state space, Γ the gambling constraint, T the stopping constraint and u the reward function for $\underline{G}$. F, Γ and u play the same role as in the gambler's problem of [9]. The role of T is this: suppose we have experienced a partial history x, then we must quit the house with a reward $u(x)$ if $T(x)$ is 0; if $T(x)$ is 1, then we must play and choose a γ from $\Gamma(x)$; if $T(x)$ is 2, then we can do either of these. The gambler's problem of [9] corresponds roughly to the case $T \equiv 2$.

Let $\underline{G} = (F, \Gamma, T, u)$ be a g.s., $x \in W(F)$ and $\pi = (\sigma, a)$ be a policy over F. π is said to be available in $\underline{G}$ at x if the following conditions are satisfied: (a) $I_u(\pi)(x)$ is defined; (b) for each $n \geq 0$, there is a subset A_n of F^n with $\sigma[x](A_n) = 1$ such that for each $y \in A_n$, $a(xy) = T(xy)$ if $T(xy)$ is 0 or 1 and $\sigma(xy) \in \Gamma(xy)$ if $a(xy) = 1$. The condition (b) says that from x and using π, in every situation that can actually arise until the time of quitting the house, π prescribes only those choices which are allowed by Γ and T. π is said to be strictly available in $\underline{G}$ at x if it is available in $\underline{G}$ at x and if condition (b) is satisfied with $A_n = F^n$ for all $n \geq 0$.

We shall denote by $A(\underline{G}, x)$ the set of all policies available in $\underline{G}$ at x and by $A_s(\underline{G}, x)$ the set of all policies strictly available in $\underline{G}$ at x. A permissible class of policies for $\underline{G}$ is a function D on $W(F)$ such that $D(x) \subset A(\underline{G}, x)$ for all x in $W(F)$. Suppose $D(x)$ is nonempty. Let $U_D^u(x) = \operatorname{Sup} \{I_u(\pi)(x) : \pi \in D(x)\}$. $U_D^u(x)$ represents our maximum expected reward if we have experienced x and are constrained to use policies from $D(x)$. Any π in $D(x)$ satisfying $I_u(\pi)(x) = U_D^u(x)$ is said to be optimal in D at x. Let $D^1(x)$ denote the projection of $D(x)$ to $\Sigma(F)$, i.e., $D^1(x) = \{\sigma \in \Sigma(F) : (\sigma, a) \in D(x)$ for some $a \in \Lambda(F)\}$. We write

$\bar{I}^u_D(\sigma, a)(x)$ for the quantity $\sup\{I_u(\sigma, a')(x): (\sigma, a') \in D(x)$, for each $n \geq 0$, $a'(xy) \geq a(xy)$ for almost all $y \in F^n$ with respect to $\sigma[x]\}$. Let $\bar{J}^u_D(\sigma)(x) = \inf\{\bar{I}^u_D(\sigma, a)(x): (\sigma, a) \in D(x)$ for each $\sigma \in D^1(x)\}$. Define $\bar{V}^u_D(\sigma)(x) = \sup\{\bar{J}^u_D(\sigma)(x) : \sigma \in D^1(x)\}$. Then $\bar{V}^u_D(x)$ represents our maximum expected reward if we have experienced x and using policies from $D(x)$ we are forced to play for as long a time as possible The quantity obtained by changing the supremum in the definition of $\bar{I}_D(\sigma, a)(x)$ into infimum is denoted as $\underline{I}^u_D(\sigma, a)(x)$. Let $\underline{J}^u_D(\sigma)(x) = \sup\{\underline{I}^u_D(\sigma, a)(x): (\sigma, a) \in D(x)\}$ and $\underline{V}^u_D(\sigma)(x) = \sup\{\underline{J}^u_D(\sigma)(x)\ \sigma \in D^1(x)\}$. A strategy σ is $\bar{V}$ optimal in D at x if $\bar{J}^u_D(\sigma)(x) = \bar{V}^u_D(x)$; similarly for $\underline{V}$ optimality. Whenever the context is clear, we shall drop the subscripts and superscripts in I_u, $\bar{I}^u_D$, $\bar{J}^u_D$, U^u_D, $\bar{V}^u_D$, $\underline{V}^u_D$ etc.

The general problem, which we shall refer to as the gambling problem (g. p.), may now be stated thus: given a g.s. $\underline{G} = (F, \Gamma, T, u)$ and a permissible class D of policies for $\underline{G}$, find $U(x)$, $\bar{V}(x)$, $\underline{V}(x)$, optimal policies and optimal strategies in D at x, whenever $D(x)$ is nonempty. It can be observed that Γ and T are only of minor importance for the specification of the problem; what are needed are the fortune space F, the reward function u, the class D of policies and the fact that $I_u(\pi)(x)$ is defined for all π in $D(x)$. However in most of the situations, the class D arises from a Γ and a T in some natural manner and this is the reason for introducing them.

3. Examples:

We now consider a few sequential problems that have been studied in the literature and discuss how they may be regarded as special cases of the gambling problem.

I. Dynamic Programming (d.p.) - (A). We begin with the d.p. problems investigated by Blackwell [3,5], Strauch [25] and others. Using the notations in [25], we suppose that a d.p. problem specified by S, A, q, r, β is given. This may be regarded as a g.p. in one of many ways. For instance, let $F = S \cup A$ assuming without loss of generality that S and A are disjoint. Let $T \equiv 1$ and the reward function u be such that $u(s_1, a_1, s_2, a_2, \ldots) = \sum_{n=1}^{\infty} \beta^{n-1} r(s_n, a_n, s_{n+1})$ if $s_n \in S$ $a_n \in A$ $n \geq 1$. Define Γ as follows:

$$\Gamma(x) = \{\gamma \in G(F): \gamma(A) = 1\} \text{ if } x = (y,s),\ y \in W(F), s \in S$$

$$= \{\bar{q}(\cdot|s,a)\} \text{ if } x = (y,s,a),\ y \in W(F),\ s \in S,\ a \in A$$

$$= G(F) \text{ otherwise.}$$

In the above definition, $\bar{q}(\cdot|s,a)$ denotes any extension of

$q(\cdot|s,a)$ to a gamble on F. Let $\underset{\sim}{G} = (F, \Gamma, T, u)$ and D be any permissible class for $\underset{\sim}{G}$ such that $D(s)$ is the set of all Borel measurable policies strictly available in $\underset{\sim}{G}$ at s, for $s \in S$. Then it is clear that $U_D(s) = \bar{V}_D(s) = \underline{V}_D(s) = v^*(s)$ for all $s \in S$.

(B) Derman [8], Ross [23] and others consider the problem of minimising the average expected loss. It is convenient to use the notations of [25] to discuss this set up as well. Let the average income from a plan π be defined as $I_*(\pi)(s) = \liminf_{n \to \infty} \frac{1}{r} I_n(\pi)(s)$. We wish to maximise this income over all measurable plans; denote this maximum by $w_*(s)$.

For reformulating this problem, we may take F and Γ as in (A) above and let $T \equiv 2$. Define u so that $u(s_1, a_1, \ldots, s_n, a_n, s_{n+1}) = \frac{1}{n} \sum_{j=1}^{n} r(s_j, a_j, s_{j+1})$ if $s_j \in s$ $a_j \in A$. For each $n \geq 1$, let a_n be the s.f. over F such that $a_n(x)$ is 1 or 0 according as the length of $x \leq 2n$ or $> 2n$. Let D be any permissible class for $\underset{\sim}{G} = (F, \Gamma, T, u)$ such that $D(s) = \{(\sigma, a_n) \in A_s(\underset{\sim}{G}, s) : \sigma$ is Borel measurable and $n \geq 1\}$. Then it is easy to verify that $\underline{V}_D(s) = w_*(s)$ for $s \in S$.

Several other formulations of discrete time d.p. problems like those of Łos [15], Bellman [2], White [28], Hinderer [12] can also be treated as gambling problems in analogous way.

II. Gambling Theory:

The gambler's problem in the sense of Dubins and Savage [9] is clearly a g.p. in our sense. (In fact, but for some technical annoyances, the converse is also true.) To spell out the details, given a gambler's problem specified by F', Γ', u', let $F = F'$, $\Gamma(x) = \Gamma'(\lambda(x))$ and $u(x) = u'(\lambda(x))$ for $x \neq \emptyset$ where $\lambda(x)$ denotes the last coordinate of x. Let $\Gamma(\emptyset)$, $u(\emptyset)$ be arbitrary and $T \equiv 2$. Define $D(x)$ to consist of all policies (σ, a) strictly available in this g.s. $\underline{G} = (F, \Gamma, T, u)$ at x such that $a[x]$ is finite. Then the functions $U_D(f)$ and $\bar{V}_D(f)$ are respectively the utility U and strategic utility V of the house Γ'.

III. Statistical Decision Theory:

A wide class of sequential decision problems may be looked upon as gambling problems. For instance, consider the problem of choosing a decision d from a decision space K about the true parameter value θ in Ω on the basis of sequentially observed random variables $x_1, x_2, \ldots$ which have a joint distribution P_θ. If we take n observations before making the decision d, we assume that our loss is $L(\theta, d) + nc$ where L is a known function on $\Omega \times K$ and c is a known constant. We wish to minimise our expected loss when θ has a known prior distribution ξ. A straightforward reformulation of this situation is obtained as follows: Let $q(\theta, f_1, \ldots, f_n$ denote the conditional distribution of x_{n+1} given $x_1 = f_1, \ldots, x_n = f_n$ and let $q(\theta)$ denote the marginal distribution of x_1, under P_θ. Define $F = \Omega \cup K \cup R$ assuming that these three sets are disjoint; here R denotes $(-\infty, \infty)$. Let Γ be a gambling constraint such that $\Gamma(\emptyset) = \{\bar{\xi}\}$ and for any x in $W(R)$, $\Gamma(\theta\, x) = \{\delta(d) : d \in K\} \cup \{\bar{q}(\theta, x)\}$ where $\bar{\xi}$ and $\bar{q}\,(\theta, x)$ are any extensions, of ξ and $q(\theta, x)$ respectively, to gambles over F. Set $T(\theta x d) = 0$ if $\theta \in \Omega$, $x \in W(R)$, $d \in K$ and $T = 1$ otherwise. Also let $u(\theta x d) = -L(\theta, d) - nc$ where n is the length of x and $u = -\infty$ otherwise. With this g.s. $\underline{G} = (F, \Gamma, T, u)$ and $D = A_s(\underline{G})$, our problem is one of finding $U_D(\emptyset)$ and the optimal policies in D at $\emptyset$.

It suffices to mention at this stage that several other problems such as those of optimal stopping, sampling with and without recall, discrete search and discrete optimal control (see [7]) can be similarly rephrased as gambling problems.

However it should be noted that if we do not adopt the approach of assuming prior distributions for unknown parameters, some of these problems involve ideas from two-person games and consequently will not fit into our optimization setup. Moreover optimization problems like d.p. have their counterparts in two-person games. So it seems worthwhile to enlarge our notion of g.p. to enable us to study minimax problems as well. Towards this end, we define a two-person (zero-sum) gambling system as a 5-tuple $\underset{\sim}{G} = (F, \Gamma, T, u, X)$ where $\underset{\sim}{G}' = (F, \Gamma, T, u)$ is a g.s. and X is a nonempty proper subset of $W(F)$. X is thought of as the set of positions at which player 1 must play and $Y = W(F) - X$ as the set of positions at which player 2 must play. A strategy for 1(2) is a function on $X(Y)$ into $\underset{\sim}{G}(F)$; thus strategies over F can be identified with pairs (ξ, η) consisting of strategies ξ for 1 and η for 2. Similarly s.fs. and policies for 1 and 2 are defined. A permissible class of policies for $\underset{\sim}{G}$ is defined as one which is permissible for $\underset{\sim}{G}'$. The value for player 1 is defined as $\bar{W}_D(x) = \operatorname{Inf} \operatorname{Sup} I_u(\pi_1, \pi_2)(x)$ where the infimum is taken over policies π_1 for 1 and supremum over policies π_2 for 2 such that $(\pi_1, \pi_2) \in D(x)$. The quantity obtained by interchanging inf and sup is denoted as $\underline{W}_D(x)$. $\underset{\sim}{G}$ is said to be determined at x if $\bar{W}_D(x) = \underline{W}_D(x)$.

Note that perfect information games of Gale and Stewart [11], stochastic games [18], game variants of optimal stopping [10] etc., may be regarded as two-person gambling problems.

4. Optimality Equations:

In this section we study some simple properties of the optimal reward functions U_D, $\bar{V}_D$, $\underline{V}_D$, $\bar{W}_D$ and $\underline{W}_D$. We first consider the optimality equations which play an important role in the theory of d.p.

<u>Definitions</u>: Let $\underline{G} = (F, \Gamma, T, u)$ be a g.s. and v be a function on $W(F)$ into $\bar{R}$ such that $\int v(xf)\,\gamma(df)$ exists whenever $T(x) \neq 0$ and $\gamma \in \Gamma(x)$. v is said to be conservative for $\underline{G}$ in case, for all x,

$$\begin{aligned} v(x) &= u(x) \quad \text{if} \quad T(x) = 0 \\ &\leq \sup_{\gamma \in \Gamma(x)} \int v(xf)\,\gamma(df) \quad \text{if} \quad T(x) = 1 \\ &\leq \max\left[u(x), \sup_{\gamma \in \Gamma(x)} \int u(xf)\,\gamma(df)\right] \text{ if } T(x) = 2. \end{aligned}$$

v is said to be excessive for $\underline{G}$ if the above relations hold with $\leq$ replaced by $\geq$. If v is both excessive and conservative for $\underline{G}$ then v is said to satisfy the U-optimality equation for $\underline{G}$.

Suppose D is a permissible class of policies for $\underline{G}$, D is said to be proper if $D(x)$ is nonempty for all x. D is said to admit continuation if, for each x in $W(F)$ and (σ, a) in $D(x)$ such that $a(x) = 1$, we have $(\sigma, a) \in D(xf)$ for almost all (a, a.) f under $\sigma(x)$. We say that D admits extension in $\underline{G}$ if the following two conditions hold: (i) whenever $T(x) = 2$ there is a $(\sigma, a) \in D(x)$ with $a(x) = 0$; (ii) suppose $T(x) \neq 0$ and $\gamma \in \Gamma(x)$. Then for any set $\{\pi_f : f \in F\}$ of policies such that $\pi_f \in D(xf)$ for all f, we can find $\pi = (\sigma, a)$ in $D(x)$ satisfying $a(x) = 1$, $\sigma(x) = \gamma$ and $\pi[xf] = \pi_f[xf]$ for a.a.f under γ. D is said to be U-regular for $\underline{G}$ if it admits both continuation and extension in $\underline{G}$

With these definitions we have the following simple result.

<u>Theorem 1:</u> Let $\underline{G} = (F, \Gamma, T, u)$ be a g.s. and D be a proper permissible class for $\underline{G}$ such that $\int U_D(xf)\, \gamma(df)$ exists for all $\gamma \in \Gamma(x)$ whenever $T(x) \neq 0$. Then,

(a) U_D is conservative for $\underline{G}$ if D admits continuation.

(b) U_D is excessive for $\underline{G}$ if D admits extension in $\underline{G}$.

(c) U_D satisfies the U-optimality equation for $\underline{G}$ if D is U-regular for $\underline{G}$.

<u>Proof</u>: Clearly (c) follows from (a) and (b). Also it is readily seen that $U_D(x) = u(x)$ if $T(x) = 0$; it is therefore enough to consider the x's for which $T(x) \neq 0$. Suppose D admits continuation and $T(x) = 1$. Then for any $\pi = (\sigma, a)$ in $D(x)$, $\sigma(x) \in \Gamma(x)$, $a(x) = 1$ and $\pi \in D(xf)$ for a.a.f under $\sigma(x)$. Hence $I(\pi)(xf)$ is defined and does not exceed $U_D(xf)$ for all $f \notin N$ where $\sigma(x)(N) = 0$. So

$$I(\pi)(x) = \int_{F-N} I(\pi)(xf)\sigma(x, df)$$

$$\leq \int U_D(xf)\, \sigma(x, df) \leq \sup_{\gamma \in \Gamma(x)} \int U_D(xf)\, \gamma(df).$$

Hence, taking supremum over π in $D(x)$, we have

$U_D(x) \leq \sup_{\gamma \in \Gamma(x)} \int U_D(xf)\gamma(df)$. The case $T(x) = 2$ is similar.

This proves (a).

Now suppose D admits extension in $\underset{\sim}{G}$ and $T(x) = 1$. Fix $\epsilon > 0$, $k > 0$, $\gamma \in \Gamma(x)$ and let $A = \{f \in F: U_D(xf) < \infty\}$. For each $f \in F$ choose a policy $\pi_f \in D(xf)$ such that $I(\pi_f)(xf) \geq U_D(xf) - \epsilon$ if $f \in A$ and $I(\pi_f)(xf) \geq k$ if $f \notin A$. Since D admits extension in $\underset{\sim}{G}$ there is a $\pi = (\sigma, a) \in D(x)$ such that $a(x) = 1$, $\sigma(x) = \gamma$ and $\pi[xf] = \pi_f[xf]$ for all $f \notin N$ where $\gamma(N) = 0$. So,

$$U_D(x) \geq I(\pi)(x) = \int_{F-N} I(\pi_f)(xf)\ \gamma(df)$$

$$\geq \int_A U_D(xf)\ \gamma(df) - \epsilon + k \cdot \gamma(F-A).$$

Taking supremum over k, $U_D(x) \geq \int U_D(xf)\ \gamma(df) - \epsilon$. As this is true for all $\epsilon > 0$ and $\gamma \in \Gamma(x)$, we have

$U_D(x) \geq \sup_{\gamma \in \Gamma(x)} \int U_D(xf)\ \gamma(df)$. The case $T(x) = 2$ is again similar. This proves (b) and completes the proof of the theorem.

Remarks: (1) Note that the gambling problems in all the examples of section 3 have or can be assumed to have proper permissible classes which admit continuation. In example II they also admit extension. The permissible classes in example I may be taken to admit extension if the action space A is countable. However, if A is uncountable, the permissible classes do not admit extension.

(2) Simple examples with deterministic gambling systems can be constructed to show that the assumptions of the theorem regarding D can not in general be weakened. For instance, to show that U_D need not be conservative for $\underline{G}$ if D does not admit continuation, we proceed as follows. Let F be a nonempty set and Γ be a gambling constraint which is deterministic in the sense that, for every x in $W(F)$, each γ in $\Gamma(x)$ is a one-point gamble $\delta(f)$ for some $f \in F$. Let $T \equiv 1$. Then for any reward function u, the g.s. $\underline{G} = (F, \Gamma, T, u)$ is such that $A(\underline{G}, x)$ does not depend on u. Also any permissible class D for $\underline{G}$ can be identified with a subset of $H(F)$ so that $U_D^u(x) = \sup\{u(x,h): h \in D(x)\}$. Now if D does not admit continuation, we can find $x_o \in W(F)$ and $h_o = (f_o, h_1)$ in $D(x_o)$ such that $h_1 \notin D(x_o f_o)$. So by defining $u(x_o, h_o) = 1$ and $u = o$ otherwise, we have $U_D^u(x_o) = 1$ and $\int U_D^u(x_o f)\,\gamma(df) = 0$ for all $\gamma \in \Gamma(x_o)$. U_D is therefore not conservative for $\underline{G}$. In a similar way the condition that D admit extension in $\underline{G}$ can be shown to be necessary, in general, for U_D to be excessive for $\underline{G}$.

(3) Suppose D is a proper permissible class for a g.s. $\underline{G} = (F, \Gamma, T, u)$ and let $\pi = (\sigma, a)$ be an optimal policy in

D at an x in W(F). Set $A = \{f \in F: U_D(xf) < \infty\}$. π' is said to satisfy the U-optimality criterion for D at x if, either $a(x) = 0$ or for each $\epsilon > 0$ and $k > 0$ we can find $N \subset F$ with $\sigma(x)(N) = 0$ such that $I(\pi)(xf) \geq U_D(xf)$ for $f \in A - N$ and $I(\pi)(xf) \geq k$ for $f \notin A \cup N$. D is said to satisfy the optimality criterion if, every optimal policy in D at any x satisfies the U-optimality criterion at x. It is easy to observe that if U_D satisfies the U-optimality equation and if D admits either extension or continuation in $\underline{G}$, then D satisfies the optimality criterion. This fact in deterministic or measurable gambling problems yields the principle of optimality in its familiar form.

Now we pass on to consider the equations satisfied by $\bar{V}_D$ and $\underline{V}_D$ in many gambling problems. For simplicity, we consider only those reward functions which are bounded, nonnegative or nonpositive.

<u>Definitions:</u> Let $\underline{G} = (F, \Gamma, T, u)$ be a g.s. and v be a function on W(F) into $\bar{R}$ which is either bounded or of constant sign. v is said to satisfy the V-optimality equation for $\underline{G}$ if,

$$\begin{aligned} v(x) &= u(x) && \text{if } T(x) = 0 \\ &= \sup_{\gamma \in \Gamma(x)} \int v(xf)\, \gamma(df) && \text{if } T(x) \neq 0. \end{aligned}$$

Let D be a permissible class for $\underline{G}$. D is said to admit substitution by stopping functions, if, for any x in $W(F)$ and policies (σ, a) and (σ', a') in $D(x)$ such that $a(x) = a'(x) = 1$, $\sigma'(x) = \sigma(x)$, $\sigma[xf] = \sigma'[xf]$ for a.a.f under $\sigma(x)$, we have $(\sigma, a') \in D(x)$. D is said to be V-regular for $\underline{G}$ if D admits continuation, extension in $\underline{G}$ and substitution by s.fs. Let $\sigma \in D^1(x)$. σ is said to stop in D at x if $a(x) = 0$ for every s.f. a with $(\sigma, a) \in D(x)$. σ is said to persistently continue in D at x if, whenever $(\sigma, a) \in D(x)$, there is a s.f. a' such that $(\sigma, a') \in D(x)$, $a'(x) = 1$ and $a'[xf] \geq a[xf]$ for a.a.f under $\sigma(x)$. Observe that if $T(x) = 0$ then every σ in $D^1(x)$ stops at x and if $T(x) = 1$ then every σ in $D^1(x)$ persistently continues at x.

<u>Lemma 2:</u> Let $\underline{G} = (F, \Gamma, T, u)$ be a g.s. such that the reward function u is either bounded or of constant sign. Let D be a proper permissible class for $\underline{G}$ which is V-regular. Then, for any x in $W(F)$ and $\sigma \in D^1(x)$,

$$\begin{aligned} \bar{J}_D(\sigma)(x) &= u(x) && \text{if } \sigma \text{ stops in } D \text{ at } x \\ &= \int \bar{J}_D(\sigma)(xf)\, \sigma(x, df) && \text{if } \sigma \text{ persistently continues in } D \text{ at } x \\ &= \operatorname{Min}\, [u(x), \int \bar{J}_D(\sigma)(xf)\, \sigma(x, df)] && \text{otherwise.} \end{aligned}$$

<u>Proof:</u> Clearly if σ stops at x then $\bar{J}_D(\sigma)(x) = u(x)$. To

complete the proof in the other two cases it is enough to show that $\operatorname{Inf}\{\bar{I}(\sigma,a)(x): (\sigma,a)\in D(x),\ a(x) = 1\} = \int \bar{J}_D(\sigma)(xf)\sigma(x,df)$. Fix a s.f. a such that $(\sigma,a)\in D(x)$ and $a(x) = 1$. Let $\epsilon > 0$ and $k > 0$. Since D admits continuation, $(\sigma,a)\in D(xf)$ for a.a.f under $\sigma(x)$ and so $\sigma\in D^1(xf)$ for all $f\notin N$, $\sigma(x)(N) = 0$. Let $A = \{f\notin N: \bar{J}_D(\sigma)(xf) < \infty\}$. For each $f\notin N$, choose a_f such that $(\sigma,a_f)\in D(xf)$, $a_f(xfy) \geq a(xfy)$ for a.a.$y\in F^n$ under $\sigma[xf]$, $n \geq 0$ satisfying

$$\begin{aligned} I_u(\sigma, a_f)(xf) &\geq \bar{J}_D(\sigma)(xf) - \epsilon && \text{if } f\in A \\ &\geq k && \text{if } f\notin N\cup A . \end{aligned}$$

For $f\in N$ choose any $(\sigma_f, a_f)\in D(xf)$. Since D admits extension in $\underline{G}$, there is a $(\sigma', a')\in D(x)$ satisfying $\sigma'(x) = \sigma(x)$, $a'(x) = 1$, $\sigma'[xf] = \sigma[xf]$ and $a'[xf] = a_f[xf]$ for a.a.f under $\sigma(x)$. As D admits substitution by s.fs., $(\sigma,a')\in D(x)$. Also, $\bar{I}(\sigma,a)(x) \geq I(\sigma,a')(x) \geq \int_A \bar{J}_D(\sigma)(xf)\sigma(x,df) - \epsilon + k\sigma(x)(F-A)$. Since ϵ and k are arbitrary, $\bar{I}(\sigma,a)(x) \geq \int \bar{J}_D(\sigma)(xf)\ \sigma(x,df)$ and consequently, $\bar{J}_D(\sigma)(x) \geq \int \bar{J}_D(\sigma)(xf)\ \sigma(x,df)$. The reverse inequality is obtained similarly. This completes the proof.

Since $\bar{J}^u_D = -\underline{J}^v_D$ where $v = -u$, we get the following expression for $\underline{J}_D$ under the assumptions of the lemma:

$$\begin{aligned} \underline{J}_D(\sigma)(x) &= u(x) && \text{if } \sigma \text{ stops in D at } x \\ &= \int \underline{J}_D(\sigma)(xf)\sigma(x,df) && \text{if } \sigma \text{ persistently continues in D at } x \\ &= \operatorname{Max}[u(x), \int \underline{J}_D(\sigma)(xf)\sigma(x,df)] && \text{otherwise.} \end{aligned}$$

Using lemma 2 and the above remark we can prove the following theorem. We omit the easy proof.

<u>Theorem 3</u>: Let $\underline{G} = (F, \Gamma, T, u)$ be a g.s. with a reward function which is either bounded or of constant sign. Suppose D is a V-regular proper permissible class of policies for $\underline{G}$ such that whenever $T(x) = 2$ every strategy in $D^1(x)$ persistently continues in D at x. Then $\bar{V}_D$ and $\underline{V}_D$ satisfy the V-optimality equation for $\underline{G}$.

<u>Remarks:</u> (1) Observe that in all the examples of section 3 the permissible classes admit substitution by s.fs. and also satisfy the condition that every σ in $D^1(x)$ persistently continues in D at x if $T(x) = 2$.

(2) As remarked earlier, the condition that D admit continuation and extension in $\underline{G}$ can be shown to be necessary in general for U_D to satisfy the U-optimality equation. In fact this can be done by considering examples in which $T \equiv 1$; but in that case $U_D = \bar{V}_D = \underline{V}_D$ and U-optimality equations and V-optimality equations are the same. Hence U-regularity of D is necessary for $\bar{V}_D$ and $\underline{V}_D$ to satisfy the V-optimality equations. We now give an example of a deterministic g.s. with a U-regular proper permissible class D which does not admit substitution by s.fs. and such that $\bar{V}_D$ does not satisfy the V-optimality equation.

The condition that a permissible class admit substitution by s.f s. is a very natural restriction, for it says that whether or not a policy π can be used after experiencing a partial history x should depend only on what π prescribes for partial histories of the form (x, x'). Consequently our example is quite artificial. Let $F = \{0, 1\}$, $\Gamma(x) = \{\partial(0), \partial(1)\}$ for all x, $T \equiv 2$. $u(x)$ is 1 if x is of even length and is 0 otherwise. We now define D: $D(x) = \{(\sigma, a) \in A_s(\underline{G}, x): a(xy) = 0$ for $y \neq \emptyset$ iff $\sigma(\emptyset) = \partial(0)\}$. It is easy to check that D admits continuation, extension in $\underline{G}$ and that every strategy in $D^1(x)$ persistently continues at x, for all $x \in W(F)$. However $\bar{V}_D(x) = 1$ if x is of odd length and $\bar{V}_D(x) = 0$ otherwise. Thus $\bar{V}_D$ does not satisfy the V-optimality equation.

Now we shall briefly mention, without proof, an analogous result regarding the functional equations satisfied by $\bar{W}_D$ and $\underline{W}_D$ in two-person gambling systems. Suppose $\underline{G} = (F, \Gamma, T, u, X)$ is a two-person g.s. and D is a permissible class. For simplicity assume u is bounded. Let D admit continuation and extension in $\underline{G}$. Suppose further that D satisfies the following two conditions: (i) Let $x \in X$. If $\pi_1 = (\xi_1, a_1)$ is a policy for player 1 and π_2^f a policy for player 2 such that (π_1, π_2^f) is in $D(xf)$ for a.a.f. under $\xi_1(x)$ and $a_1(x) = 1$ then there is a policy π_2 for 2 such that $(\pi_1, \pi_2) \in D(x)$

and for a.a.f. under $\xi_1(x)$, $\pi_2(xfy) = \pi_2^f(xfy)$ if $xfy \in Y$.

ii) Let $x \in Y$. Suppose $(\pi_1^f, \pi_2) \in D(xf)$ for a.a.f under $\xi_2(x)$ where $\pi_2 = (\xi_2, a_2)$ is a policy for 2 with $a_2(x) = 1$ then there is a policy π_1 for 1 such that $(\pi_1, \pi_2) \in D(x)$ and $\pi_1(xfy) = \pi_1^f(xfy)$ if $xfy \in X$ for a.a.f under $\xi_2(x)$.

With these conditions on D we can prove that

$$
\begin{aligned}
\bar{W}_D(x) &= u(x) && \text{if } T(x) = 0 \\
&= \sup_{\gamma \in \Gamma(x)} \int \bar{W}_D(xf)\, \gamma(df) && \text{if } T(x) = 1,\ x \in X \\
&= \inf_{\gamma \in \Gamma(x)} \int \bar{W}_D(xf)\, \gamma(df) && \text{if } T(x) = 1,\ x \in Y \\
&= \max\left[u(x), \sup_{\gamma \in \Gamma(x)} \int \bar{W}_D(xf)\, \gamma(df)\right] && \text{if } T(x) = 2,\ x \in X \\
&= \min\left[u(x), \inf_{\gamma \in \Gamma(x)} \int \bar{W}_D(xf)\, \gamma(df)\right] && \text{if } T(x) = 2,\ x \in Y.
\end{aligned}
$$

We have similar expressions for $\underline{W}_D$ with X, Y interchanged, sup and inf interchanged, max and min interchanged and $\bar{W}_D$ replaced by $\underline{W}_D$ throughout.

We conclude this section with an application of the above functional equations for $\bar{W}_D$ and $\underline{W}_D$.

Theorem 4: Let $\underline{G} = (F, \Gamma, T, u, X)$ be a two-person g.s. such that u when restricted to $H(F)$ is bounded and continuous (from the product of discrete topologies on F to the usual topology on R). Then for all $x \in W(F)$, $\overline{W}^u_D(x) = \underline{W}^u_D(x)$, where $D = A_s(\underline{G})$.

Proof: Since every bounded continuous function on $H(F)$ into R can be uniformly approximated by an inductively integrable function [9, pp. 19], we may assume that u itself is inductively integrable on $H(F)$. We prove our required result by induction on the Eudous structure of the function u restricted to $H(F)$. Assume that for all x and for all ordinals $\beta < \alpha$ we have shown that $u(x, \cdot)$ is a function of Eudoxus structure $< \beta$ implies $\overline{W}_D(x) = \underline{W}_D(x)$. Let $u(x, \cdot)$ be a function of Eudoxus structure $< \alpha$. Then for each $f \in F$ the function $u(xf, \cdot)$ is of structure $< \beta_f$ for some $\beta_f < \alpha$. So $\overline{W}_D(xf) = \underline{W}_D(xf)$. Now $A_s(\underline{G})$ is a proper permissible class which admits continuation, extension in $\underline{G}$ and also the two conditions (i) and (ii). Consequently $\overline{W}_D$ and $\underline{W}_D$ satisfy the functional equations given above. This fact together with the equality $\overline{W}_D(xf) = \underline{W}_D(xf)$ shows that $\overline{W}_D(x) = \underline{W}_D(x)$. This completes the induction step and our proof.

Remark: If we let $T \equiv 1$ and Γ be deterministic, then the two-person g.s. $\underline{G} = (F, \Gamma, T, u, X)$ corresponds to the games of perfect information studied by Gale and Stewart [11]. Then the above theorem says that if the pay-off function Φ is continuous, then the game is determined. This shows in particular that if $\Phi = I_E$ where E is a clopen subset of $H(F)$ then the game is determined. In fact much more is known, viz., if $\Phi = I_E$ where E is a $G_{\partial\sigma\partial}$ then the game is determined.

5. Mixtures of Strategies:

The typical problem investigated in this section is the following: given a g.s. $\underline{G} = (F, \Gamma, T, u)$, let $\Gamma^*(x)$ denote the convex hull of $\Gamma(x)$ for each $x \in W(F)$ and let $\underline{G}^* = (F, \Gamma^*, T, u)$. Consider $D(x) = \{(\sigma, a) \in A(\underline{G}, x): a[x]$ is finite$\}$ and the corresponding permissible class D^* for $\underline{G}^*$. Then is $U_D = U_{D^*}$?

Here we look upon $G(F)$ as a subset of $R^{S(F)}$. Thus a subset K of $G(F)$ is convex means that for any $\xi, \eta \in K$ and $0 \leq \alpha \leq 1$, $\alpha\xi + (1-\alpha)\eta \in K$. A net $\{\gamma_\alpha\}$ is said to converge to γ if $\gamma_\alpha(A) \longrightarrow \gamma(A)$ for all $A \subset F$. $G(F)$ then becomes a compact convex set. For any $K \subset G(F)$ let K_o denote the smallest closed convex set containing K. We shall need the following lemma.

<u>Lemma 5:</u> Let g be an inductively integrable function on $H(F)$ into R. Suppose $\Gamma(x) \subset G(F)$ for all $x \in W(F)$ and σ^* is a strategy over F such that $\sigma^*(x) \in (\Gamma(x))_o$ for all x.

Then for any $\epsilon > 0$ there exist strategies σ and σ' satisfying:

(a) $\sigma(x) \in \Gamma(x)$, $\sigma'(x) \in \Gamma(x)$ for all $x \in W(F)$

(b) $\int g \, d\sigma' - \epsilon \leq \int g d\sigma^* \leq \int g d\sigma + \epsilon$.

<u>Proof:</u> By induction on the Eudoxus structure of g.

<u>Theorem 6:</u> Let $\underline{G} = (F, \Gamma, T, u)$ be a g.s. such that u is either bounded or of constant sign. Let $\underline{G}_o = (F, \Gamma_o, T, u)$ where $\Gamma_o(x) = (\Gamma(x))_o$ for all x in $W(F)$. Suppose $D(x) = \{(\sigma, a) \in A_s(\underline{G}, x): a[x]$ is finite$\}$ and $D_o(x) = \{(\sigma, a) \in A(\underline{G}_o, x): a[x]$ is finite$\}$ for all x in $W(F$ Then $U_D(x) = U_{D_o}(x)$ whenever $D(x)$ is nonempty.

<u>Proof:</u> Clearly $U_D(x) \leq U_{D_o}(x)$. Let $(\sigma^*, a) \in D_o(x)$ and $\epsilon > 0$. We may assume, without loss of generality, that $\sigma^*(y) \in \Gamma_o(y)$ for all y. Now $R_u(a, x)$ is an inductively inte ble function if u is bounded, since $a[x]$ is finite. So by lemma 5, there is a σ such that

$\int R_u(a, x) d\sigma [x] \geq \int R_u(a, x) d\sigma^* [x] - \epsilon$ and $(\sigma, a) \in D(x)$.

So $U_D(x) \geq I_u(\sigma, a)(x) \geq I_u(\sigma^*, a) - \epsilon$ and ϵ, σ^*, a being arbitrary, $U_D(x) \geq U_{D_o}(x)$. If u is a non-negative unbounde function we work with the functions $R_u(a, x) \wedge n$ and use th fact that $I_u(\sigma^*, a)(x) = \sup_n \int [R_u(a, x) \wedge n] d\sigma^*[x]$. For $u \leq$ (

we make use of the first inequality in (ii) of 5..
This completes the proof of the theorem .

In applying the above theorem we would like to know what K_o is, for any given K. The following simple result answers it to a certain extent.

Theorem 6: Let K be any non-empty subset of G(F) and $\gamma \in G(F)$. Then $\gamma \in K_o$ iff there is a set X and a function g on X into K such that for some $\mu \in G(X)$,

$$\gamma(A) = \int_X g(x)(A)d\mu(x), \quad A \subset F.$$

Proof: Suppose $\gamma \in K_o$. If $\gamma = \sum_{i=1}^{k} a_i \gamma_i$ where $a_i \geq 0$ $\gamma_i \in K$ and $\sum_1^k a_i = 1$ then we take X to be K, g to be the identity function and $\mu = \sum_{i=1}^{k} a_i \partial(\gamma_i)$. Otherwise we can find a net $\{\gamma_\alpha\}$ in K_o such that $\gamma_\alpha \longrightarrow \gamma$ and each γ_α is a finite convex combination of elements from K. Hence by above we have a $\mu_\alpha \in G(K)$ such that $\gamma_\alpha(A) = \int_K \xi(A)\mu_\alpha(d\xi)$ for all α and $A \subset F$. As G(K) is compact, some subnet of $\{\mu_\alpha\}$ converges to a μ in G(K). It is then easy to check that $\gamma(A) = \int_K \xi(A)\mu(d\xi)$ for all $A \subset F$.

Conversely, suppose X is a set, $g : X \longrightarrow K$, $\mu \in G(X)$ and $\gamma(A) = \int g(x)(A)\mu(dx)$ for all $A \subset F$. Clearly $\gamma \in G(F)$. If $\mu = \sum_{i=1}^{k} a_i \partial(x_i)$ where $a_i \geq 0$ $x_i \in X$ and $\sum_1^k a_i = 1$, then

$\gamma = \sum_{i=1}^{k} a_i\, g(x_i)$. Consequently $\gamma \in K_o$. If μ does not concentrate on a finite set then we can get a net $\{\mu_\alpha\}$ in $G(X)$ converging to μ such that each μ_α concentrates on a finite set. If γ_α is the gamble on F corresponding to μ_α, i.e. if $\gamma_\alpha(A) = \int g(x)(A)\mu_\alpha(dx)$ for all $A \subset F$, then $\gamma_\alpha \in K_o$. It is easy to see that $\{\gamma_\alpha\}$ converges to γ. So, $\gamma \in K_o$.

Remark: In applying the above theorem 6, we must ensure that there is at least one policy (σ, a) available in $\underline{G}$ at x such that $a[x]$ is finite. This needs that the s.f. $T_*[x]$ is finite where T_* is defined by : $T_*(y) = 1$ if $T(y) \neq 1$ and $T_*(y) = 0$ otherwise. In gambling problems of example II (section 3), $T \equiv 2$ and so $T_* \equiv 0$. In d.p. problems in example I(A), $T_* \equiv 1$ and consequently the above theorem does not apply. But in positive d.p. problems, the optimal income v^* does not change if we let $T \equiv 2$. Thus in some extensions of positive d.p. problems in which measurability restrictions are not imposed, for instance if S and A are countable, we can use the theorem. Noticing that if $K(E) = \{\delta(f) : f \in E\}$ where $E \subset F$, then $(K(E))_o = \{\gamma \in G(F) : \gamma(E) = 1\}$ we can conclude from theorem 6 that in such situations, the actions on various days can be chosen deterministically but depending on the past history.

We do not know if, under the conditions of Theorem 6, $\bar{V}_D(x) = \bar{V}_{D_o}(x)$. We give below a weaker result in this direction. Given $\underline{G}$, $\underline{G}_o$ as in theorem 6 and $\sigma \in A^1(\underline{G}_o, x)$, σ is said to be eventually in Γ at x if, for each $h \in H(F)$ there is an $n_o \geq 1$ such that $\sigma(x,(h|n)) \in \Gamma(x,(h|n))$ for all $n \geq n_o$.

<u>Theorem 7:</u> Let $\underline{G}$ and $\underline{G}_o$ be as in theorem 6 such that the reward function u is bounded and let $D = \underline{A}(\underline{G})$. Suppose $D^*(x) = \{(\sigma,a) \in \underline{A}(\underline{G}_o, x): \sigma$ is eventually in Γ at $x\}$. Then $U_D(x) = U_{D^*}(x)$, $\bar{V}_D(x) = V_{D^*}(x)$ and $\underline{V}_D(x) = \underline{V}_{D^*}(x)$.

<u>Proof</u>: Let $(\sigma^*,a) \in D^*(x)$. If $a(x) = 0$ then $I(\sigma^*,a)(x) = u(x) \leq U_D(x)$. If $a(x) = 1$ then let $t(h)$ be the first n such that $\sigma^*(x, (h|m)) \in \Gamma(x,(h/m))$ for all $m \geq n$ $h \in H(F)$. Then t is a finite stop rule over F and consequently $g(h) = I_u(\sigma^*, a)(x, h/t(h))$ is an inductively integrable function on $H(F)$. The equation $I_u(\sigma^*,a)(x) = \int g(h)\sigma^*[x](dh)$ can be verified by induction on the Eudoxus structure of g. Now by using lemma 5, we can get a strategy σ such that $\sigma(y) \in \Gamma(y)$ for all y and $\int g(h)\sigma[x](dh) \geq I_u(\sigma^*,a)(x) - \epsilon$ for any pre-assigned $\epsilon > 0$. Define σ' as follows:

$\sigma'(x,(h|m)) = \sigma^*(x,(h|m))$ for $m \geq t(h)$ and $h \in H(F)$

$\sigma'(y) = \sigma(y)$ otherwise. Then $(\sigma', a) \in D(x)$ and $I_u(\sigma',a)(x) \geq I_u(\sigma^*, a)(x) - \epsilon$. This shows that $U_D(x) \geq U_{D^*}(x)$, but the reverse inequality is obvious. So $U_D(x) = U_{D*}(x)$.

The equalities regarding $\overline{V}$ and $\underline{V}$ can be proved along similar lines, using the equation
$\overline{J}_{D*}(\sigma^*)(x) = \int \overline{J}_{D*}(\sigma^*)(x,h|t(h))\sigma^*[x](dh)$ for strategies σ^* which persistently continue in D^* at x.

Remark: Theorem 7 can be used in discounted d.p. problems because, in such problems, any two strategies available at x and agreeing for next N days differ in their expected rewards by at most $\beta^N \|r\|$. Hence the above theorem would imply that, we can restrict ourselves to using actions deterministically. Of course here we are not considering the question whether we can also choose them measurably.

6. Some facts from measure theory:

We first fix some notations and terminology. A Borel space is a pair $(X, \underline{\underline{A}})$ where X is a set and $\underline{\underline{A}}$ is a σ-algebra of subsets of X. $\underline{\underline{A}}$ is then called a Borel structure for X. A function on X into Y is said to be $\underline{\underline{A}}, \underline{\underline{B}}$ measurable if $\underline{\underline{A}}$ is a Borel structure for X, $\underline{\underline{B}}$ is a Borel structure for Y and $f^{-1}(\underline{\underline{B}}) \subset \underline{\underline{A}}$. For any class $\underline{\underline{L}}$ of subsets of X, $\sigma(\underline{\underline{L}})$ denotes the smallest Borel structure for X containing $\underline{\underline{L}}$; $\underline{\underline{L}}$ is called a generator for $\sigma(\underline{\underline{L}})$. For each $\alpha < \omega_1$, the first uncountable ordinal, we define $\underline{\underline{L}}_\alpha$ as follows: $\underline{\underline{L}}_o = \underline{\underline{L}}$, $\underline{\underline{L}}_\alpha$ is the class of countable unions and complements of sets in $\bigcup_{\beta<\alpha} \underline{\underline{L}}_\beta$. Then $\sigma(\underline{\underline{L}}) = \bigcup_{\alpha<\omega_1} \underline{\underline{L}}_\alpha$. If $\underline{\underline{K}}$ and $\underline{\underline{L}}$ are classes of subsets of X and Y respectively, then a function f on X into Y is said to be of class α relative to ($\underline{\underline{K}}$, $\underline{\underline{L}}$) if $f^{-1}(\underline{\underline{L}}) \subset \underline{\underline{K}}_\alpha$. A collection F of functions is said to be of class α relative to ($\underline{\underline{K}}$, $\underline{\underline{L}}$) if, for each countable subclass $\underline{\underline{L}}'$ of $\underline{\underline{L}}$ we can find a countable subclass $\underline{\underline{K}}'$ of $\underline{\underline{K}}$ such that each f in F is of class α relative to ($\underline{\underline{K}}'$, $\underline{\underline{L}}'$). A collection F of $\underline{\underline{A}}$, $\underline{\underline{B}}$ measurable functions is said to be of bounded class if, for some $\alpha < \omega_1$ and generators $\underline{\underline{K}}$ and $\underline{\underline{L}}$ for $\underline{\underline{A}}$ and $\underline{\underline{B}}$, F is of class α relative to ($\underline{\underline{K}}$, $\underline{\underline{L}}$). The smallest cardinality of a generator for $\underline{\underline{A}}$ is denoted by w ($\underline{\underline{A}}$).

We denote by $\underline{\underline{R}}$ the Borel structure for R generated by the intervals. For any Borel space (X, $\underline{\underline{A}}$) let P($\underline{\underline{A}}$) be the set of all countably additive probability measures on $\underline{\underline{A}}$ and $\underline{\underline{P}}(\underline{\underline{A}})$ be the smallest Borel structure for P($\underline{\underline{A}}$) making the functions $p \longrightarrow p(A)$ measurable, $A \in \underline{\underline{A}}$. If (Y, $\underline{\underline{B}}$) is another Borel space, we denote by Q($\underline{\underline{B}}|\underline{\underline{A}}$) the set of all functions $q : X \times B \longrightarrow R$ such that $q(x, \cdot) \in P(\underline{\underline{B}})$ for all $x \in X$ and $q(\cdot, B)$ is $\underline{\underline{A}}$, $\underline{\underline{R}}$ measurable for all $B \in \underline{\underline{B}}$. In other words,

$Q(\underline{\underline{B}}|\underline{\underline{A}})$ is the set of all $\underline{\underline{A}}$, $\underline{\underline{P}}(\underline{\underline{B}})$ measurable functions. $\underline{\underline{A}} \times \underline{\underline{B}}$ or $\underline{\underline{A}}\ \underline{\underline{B}}$ will denote the Borel structure for $X \times Y$ generated by $\{A \times B: A \in \underline{\underline{A}},\ B \in \underline{\underline{B}}\}$. There is a natural mapping

$\pi : P(\underline{\underline{A}}) \times Q(\underline{\underline{B}}|\underline{\underline{A}}) \longrightarrow P(\underline{\underline{A}}\ \underline{\underline{B}})$ defined by

$\pi(p,q)(E) = \int\int I_E(x,y)q(x,dy)p(dx)$, $E \in \underline{\underline{A}}\ \underline{\underline{B}}$. $\pi(p,q)$ is also denoted as pq and the exact range of π is denoted by $P_*(\underline{\underline{A}}\ \underline{\underline{B}})$.

Let $(X, \underline{\underline{A}})$ and $(Z, \underline{\underline{C}})$ be two Borel spaces. Suppose Y is any set and φ is a function on $X \times Y$ into Z. A Borel structure $\underline{\underline{B}}$ for Y is said to be admissible relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$ if φ is $\underline{\underline{A}} \times \underline{\underline{B}}$, $\underline{\underline{C}}$ measurable. Y is said to be admissible relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$ if there is an admissible Borel structure for Y. The following theorem of Aumann [1] and B. V. Rao [22] gives necessary and sufficient conditions for Y to be admissible.

<u>Theorem 8:</u> Let $(X, \underline{\underline{A}})$ and $(Z, \underline{\underline{C}})$ be Borel spaces. Let φ be a function on $X \times Y$ into Z. Denote by F the set of all functions $\varphi(\cdot, y)$, $y \in Y$ and let $\pi : X \times F \longrightarrow Z$ be defined by $\pi(x, f) = f(x)$. Then the following statements are equivalent:

(1) Y is admissible relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$.

(2) $S(Y)$ is an admissible structure for Y relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$.

(3) There is an admissible structure $\underline{\underline{B}}$ for Y relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$ such that $w(\underline{\underline{B}}) \leq w(\underline{\underline{C}}) \cdot \aleph_0$

(4) Each $f \in F$ is $\underline{\underline{A}}, \underline{\underline{C}}$ measurable and F is admissible relative to $(\pi, \underline{\underline{A}}, \underline{\underline{C}})$.

(5) Each $f \in F$ is $\underline{\underline{A}}, \underline{\underline{C}}$ measurable and F is of bounded class.

<u>Proof:</u> Clearly (1) $\longrightarrow$ (2) and (3) $\longrightarrow$ (1)

To show (2) $\longrightarrow$ (3): Let $\underline{\underline{L}}$ be a generator for $\underline{\underline{C}}$, of cardinality $w(\underline{\underline{C}})$. Since $\varphi^{-1}(\underline{\underline{C}}) \subset \underline{\underline{A}} \times S(Y)$, for each $C \in \underline{\underline{L}}$ we can get a countable class $\underline{\underline{K}}(C) \subset S(Y)$ such that $\varphi^{-1}(C) \in \underline{\underline{A}} \times \sigma(\underline{\underline{K}}(C))$. Let $\underline{\underline{K}}$ denote the union of $\underline{\underline{K}}(C)$ as C ranges over $\underline{\underline{L}}$. Clearly $\varphi^{-1}(\underline{\underline{L}}) \subset \underline{\underline{A}} \times \sigma(\underline{\underline{K}})$. As $\varphi^{-1}(\underline{\underline{L}})$ generates $\varphi^{-1}(\underline{\underline{C}})$ it is clear that $\underline{\underline{B}} = \sigma(\underline{\underline{K}})$ is admissible relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$. Moreover $w(\underline{\underline{B}}) \leq$ cardinality of $\underline{\underline{K}} \leq w(\underline{\underline{C}}) \cdot \aleph_0$

To show (1) $\longrightarrow$ (4): Let $\underline{\underline{B}}$ be an admissible structure for Y relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$. Each $f \in F$ is then a section $\varphi(\cdot, y)$ of the measurable function φ and is therefore $\underline{\underline{A}}, \underline{\underline{C}}$ measurable. Define a function g on Y into F by $g(y) = \varphi(\cdot, y)$ and let $\underline{\underline{F}}$ be the largest Borel structure for F making g $\underline{\underline{B}}, \underline{\underline{F}}$ measurable. Then it is easy to verify that $\underline{\underline{F}}$ is admissible relative to $(\pi, \underline{\underline{A}}, \underline{\underline{C}})$.

To show (4) $\longrightarrow$ (5): Let $\mathbf{F}$ be an admissible structure for F relative to $(\pi, \mathbf{A}, \mathbf{C})$. We wish to find an $\alpha < \omega_1$ such that F is of class α relative to $(\mathbf{A}, \mathbf{C})$. Let $\mathbf{L}$ be a countable subclass of $\mathbf{C}$. $\pi^{-1}(\mathbf{L})$ being countable, we can get a countable subclass $\mathbf{K}$ of $\mathbf{A}$ such that $\pi^{-1}(\mathbf{L}) \subset \sigma(\mathbf{K}) \times \mathbf{F}$. Let $\mathbf{D} = \quad A \times B: A \in \mathbf{K}, B \in \mathbf{F} \quad$. Then for some $\alpha < \omega_1$, $\pi^{-1}(\mathbf{L}) \subset \mathbf{D}_\alpha$. It is then easy to verify that for each $L \in \mathbf{L}$ and $f \in F$, the f-section of $\pi^{-1}(L)$ is in $\mathbf{K}_\alpha$. But the f-section of $\pi^{-1}(L)$ is $f^{-1}(L)$. So $f^{-1}(\mathbf{L}) \subset \mathbf{K}_\alpha$ for all $f \in F$.

To show (5) $\longrightarrow$ (4): Let F be of class α relative to $\mathbf{A}, \mathbf{M}$ where $\mathbf{M}$ generates $\mathbf{C}$. Fix any countable class $\mathbf{L} \subset \mathbf{M}$. Then there is a countable class $\mathbf{K} \subset \mathbf{A}$ such that each $f \in F$ is of class α relative to $(\mathbf{K}, \mathbf{L})$. For each $x \in X$, Let $\bar{x}$ denote the $\sigma(\mathbf{K})$-atom containing x and for any $A \subset X$ let $\bar{A} = \{\bar{x} : x \in A\}$. Then $\{\bar{A} : A \in \sigma(\mathbf{K})\}$ forms a Borel structure for $\bar{X}$. We can define a metric on $\bar{X}$ such that the sets $\bar{A}$ as A ranges over the field generated by $\mathbf{K}$ form a clopen base for $\bar{X}$. Similarly for $(Z, \sigma(\mathbf{L}))$. Let d be a bounded metric for $\bar{Z}$ and define $\rho(f_1, f_2) = \sup_{x \in X} d\,(\overline{f_1(x)}, \overline{f_2(x)})$ for $f_1, f_2 \in F$. Then ρ is a metric on F. Consider the map g on $\bar{X} \times F$ into $\bar{Z}$ defined as $g(\bar{x}, f) = \overline{f(x)}$. This is well-defined since each $f \in F$ is $\sigma(\underline{\underline{K}}), \sigma(\underline{\underline{L}})$ measurable and consequently must be constant on each $\bar{x}$. It is easy to verify that g is continuous in f for fixed $\bar{x} \in \bar{X}$ and is of Borel class α (in the sense of Kuratowski [13]) for each fixed $f \in F$. So by a theorem of Kuratowski [13, pp. 378], g is of Borel class $\alpha + 1$ and hence measurable. In other words, there is a Borel structure $\underline{\underline{F}}(\underline{\underline{L}})$ for F such that for all $L \in \underline{\underline{L}}$, $g^{-1}(\bar{L})$ is a measurable subset of $\bar{X} \times F$. It is then clear that $\pi^{-1}(L) \in \underline{\underline{A}} \times \underline{\underline{F}}(\underline{\underline{L}})$ for all $L \in \underline{\underline{L}}$. Let $\underline{\underline{F}}$ be the smallest Borel structure for F containing $\underline{\underline{F}}(\underline{\underline{L}})$ for all countable $\underline{\underline{L}} \subset \underline{\underline{M}}$. Then π is $\underline{\underline{A}} \times \underline{\underline{F}}, \underline{\underline{C}}$ measurable. So $\underline{\underline{F}}$ is admissible relative to $(\pi, \underline{\underline{A}}, \underline{\underline{C}})$.

To show (4) ⟶ (1): If $\underline{\underline{F}}$ is an admissible Borel structure for F, let $\underline{\underline{B}}$ be the smallest Borel structure for Y making the map $y \longrightarrow \varphi(\cdot, y)$ $\underline{\underline{B}}, \underline{\underline{F}}$ measurable. Then it is easy to see that $\underline{\underline{B}}$ is admissible relative to $(\varphi, \underline{\underline{A}}, \underline{\underline{C}})$. This proves the theorem.

As a consequence of the above results we derive the following

<u>Theorem 9:</u> Let $(X, \underline{\underline{A}})$ and $(Y, \underline{\underline{B}})$ be Borel spaces and Q be a subset of $Q(\underline{\underline{B}}|\underline{\underline{A}})$. Let $\pi: P(\underline{\underline{A}}) \times Q \longrightarrow P(\underline{\underline{A}}\ \underline{\underline{B}})$ be the be the function $(p, q) \longrightarrow pq$. For any set F of measurable

functions from X into R denote by ξ the map $(x, f) \longrightarrow f(x)$. Let L be any generator for B. Then the following statements are equivalent:

(1) Q is admissible relative to $(\pi, P(A), P(A \cdot B))$.

(2) The set $E_B = \{q(\cdot, B) : q \in Q\}$ is admissible relative to (ξ, A, R) for each $B \in L$.

(3) E_B is of bounded class relative to (A, R) for all $B \in L$.

<u>Proof:</u> To show (1) $\longrightarrow$ (2): Let Q be an admissible Borel structure for Q relative to $(\pi, P(A), P(A\ B))$. Define $\partial : X \longrightarrow P(A)$ by $\partial(x)(A) = I_A(x)$ for all A in A. Then ∂ is $A, P(A)$ measurable. Consequently the function $\eta(x, q) = (\partial(x), q)$ is $A \times Q, P(A) \times Q$ measurable. For any $B \in B$, the map $\lambda_B : P(A\ B) \longrightarrow R$ defined by $\lambda_B(\mu) = \mu(X \times B)$ is measurable. Hence the maps $g_B = \lambda_B \circ \pi \circ \eta$ on $X \times Q$ into R are $A \times Q, R$ measurable. Thus Q is admissible relative to (g_B, A, R) and so $\{g_B(\cdot, q): q \in Q\}$ is admissible relative to (ξ, A, R) for each $B \in B$. But $g_B(x, q) = q(x, B)$ for all $q \in Q$. This shows that E_B is admissible relative to (ξ, A, R) for all $B \in B$.

(2) $\longleftrightarrow$ (3) follows directly from the equivalance of (4) and (5) in theorem 8.

we make use of the first inequality in (ii) of 5..
This completes the proof of the theorem .

In applying the above theorem we would like to know what K_o is, for any given K. The following simple result answers it to a certain extent.

Theorem 6: Let K be any non-empty subset of G(F) and $\gamma \in G(F)$. Then $\gamma \in K_o$ iff there is a set X and a function g on X into K such that for some $\mu \in G(X)$,

$$\gamma(A) = \int_X g(x)(A)\,d\mu(x), \quad A \subset F.$$

Proof: Suppose $\gamma \in K_o$. If $\gamma = \sum_{i=1}^{k} a_i \gamma_i$ where $a_i \geq 0$ $\gamma_i \in K$ and $\sum_1^k a_i = 1$ then we take X to be K, g to be the identity function and $\mu = \sum_{i=1}^{k} a_i \partial(\gamma_i)$. Otherwise we can find a net $\{\gamma_\alpha\}$ in K_o such that $\gamma_\alpha \longrightarrow \gamma$ and each γ_α is a finite convex combination of elements from K. Hence by above we have a $\mu_\alpha \in G(K)$ such that $\gamma_\alpha(A) = \int_K \xi(A)\mu_\alpha(d\xi)$ for all α and $A \subset F$. As G(K) is compact, some subnet of $\{\mu_\alpha\}$ converges to a μ in G(K). It is then easy to check that $\gamma(A) = \int_K \xi(A)\mu(d\xi)$ for all $A \subset F$.

Conversely, suppose X is a set, $g : X \longrightarrow K$, $\mu \in G(X)$ and $\gamma(A) = \int g(x)(A)\mu(dx)$ for all $A \subset F$. Clearly $\gamma \in G(F)$. If $\mu = \sum_{i=1}^{k} a_i \partial(x_i)$ where $a_i \geq 0$ $x_i \in X$ and $\sum_1^k a_i = 1$, then

Proof: For each $B \in \underline{\underline{L}}$ and $n \geq 1$, we can find a compact subset $K_n(B)$ of X such that $p(K_n(B)) \geq 1 - \frac{1}{n}$ and $q(\cdot, B)$ is continuous when restricted to $K_n(B)$. Let $K(B) = \bigcup_{n=1}^{\infty} K_n(B)$ Let $K(B) = \bigcup_{n=1}^{\infty} K_n(B)$ and $K = \bigcap_{B \in \underline{\underline{L}}} K(B)$. Define $q': X \times \underline{\underline{B}} \longrightarrow R$ by $q'(x, B) = q(x, B)$ if $x \in K$ and $q'(x,B) = \mu(B)$ if $x \notin K$ where μ is a fixed probability measure on $\underline{\underline{B}}$. Since $p(K) = 1$ it is clear that $pq = pq'$. Now if g denotes $q'(\cdot, B_o)$ for some $B_o \in \underline{\underline{L}}$, then we shall check that g is of Borel class 3. For any closed set C of R, $g^{-1}(C) \cap K_n(B_o)$ is closed in $K_n(B_o)$ and hence closed in X for all $n \geq 1$. So $g^{-1}(C) \cap K(B_o)$ is an F_σ in X. K being a $F_{\sigma\delta}$ in X, $g^{-1}(C) \cap K(B_o) \cap K = g^{-1}(C) \cap K$ is an $F_{\sigma\delta}$ in X. But $g^{-1}(C)$ is either $g^{-1}(C) \cap K$ or $[g^{-1}(C) \cap K] \cup (X - K)$ according as $\mu(B_o) \notin C$ or $\mu(B_o) \in C$. In either case $g^{-1}(C)$ is a $F_{\sigma\delta\sigma}$ set in X. This proves g is of Borel class 3 and completes the proof of the theorem.

Remark: If $(Y, \underline{\underline{B}})$ is also a standard analytic space, then we can find an $\alpha < \omega_1$ such that any $q' \in Q(\underline{\underline{B}}|\underline{\underline{A}})$ satisfying condition (ii) above is of Borel class α as a map from X into $P(\underline{\underline{B}})$.

The next fact we need is a result regarding mixtures of probability measures.

<u>Theorem 11:</u> Let $(X, \underline{\underline{A}})$ be a countably generated Borel space and let $\mu \in P(\underline{\underline{A}})$. Suppose Y is a separable metric space, $\underline{\underline{B}}$ its Borel σ-algebra and K is a closed convex subset of $P(\underline{\underline{B}})$ in its weak topology. Let $g: X \longrightarrow K$ be $\underline{\underline{A}}, P(\underline{\underline{B}})$ measurable and defined $\phi_\mu(E) = \int g(x)(E)\mu(dx)$ for $E \in \underline{\underline{B}}$. Then $\phi_\mu \in K$.

<u>Proof</u>: Clearly $\phi_\mu \in P(\underline{\underline{B}})$. Without loss of generality we can assume that $\underline{\underline{A}}$ contains all singleton sets of X. Let $\underline{\underline{L}}$ be a countable field generating $\underline{\underline{A}}$. We can give a metric on X such that $\underline{\underline{A}}$ is its Borel σ-algebra and $\underline{\underline{L}}$ forms a clopen base for X. Now if μ is concentrated on a finite set, say

$$\mu = \sum_{i=1}^{k} a_i \partial(x_i), \quad \text{then} \quad \phi_\mu = \sum_{i=1}^{k} a_i\, g(x_i) \quad \text{and } K \text{ being}$$

convex, $\phi_\mu \in K$. Suppose g is a continuous function and μ is any probability on $\underline{\underline{A}}$. Choose a sequence $\{\mu_n\}$ in $P(\underline{\underline{A}})$ such that each μ_n is concentrated on a finite set and $\mu_n \longrightarrow \mu$ in the weak topology of $P(\underline{\underline{A}})$. By above $\phi_n = \phi_{\mu_n} \in K$ for all n. We shall show that $\phi_n \longrightarrow \phi_\mu$. Let f be a bounded continuous function on Y. It is then easy to see that $h(x) = \int f(y) g(x)(dy)$ defines a bounded continuous function on X, as g is continuous. Consequently

$\int h(x)\mu_n(dx) \longrightarrow \int h(x)\mu(dx)$. I.e. $\int f(y)\, \phi_n(dy) \longrightarrow \int f(y)\, \phi_\mu(dy)$

This shows that $\phi_n \longrightarrow \phi_\mu$ and K being closed, $\phi_\mu \in K$.

Thus we have proved the theorem for functions g of Borel class 0 and any $\mu \in P(\underline{\underline{A}})$

We proceed by induction and assume that the theorem has been proved for all functions g' of Borel class $< \alpha$ and for any $\mu' \in P(\underline{\underline{A}})$. Let $\mu \in P(\underline{\underline{A}})$ and g be of Borel class α. Then $g = \lim_{n \to \infty} g_n$ where g_n is of class $< \alpha$ (Note that X is of dimension 0). Let $\emptyset_n$ be the element of K corresponding to g_n and μ. It is enough to show that $\emptyset_n \to \emptyset_\mu$. Let f be a bounded continuous function on Y. Since $g_n(x) \longrightarrow g(x)$ we have $\int f(y) g_n(x)(dy) \longrightarrow \int f(y) g(x)(dy)$ for all $x \in X$ and by bounded convergence theorem,

$\iint f(y) g_n(x)(dy)\mu(dx) \longrightarrow \iint f(y)g(x)(dy)\mu(dx)$. This proves that $\emptyset_n \longrightarrow \emptyset_\mu$ and completes the proof of the theorem.

The next two results were proved by Sudderth [27] for standard Borel spaces.

<u>Theorem 12:</u> Let $(X, \underline{\underline{A}})$ and $(Y, \underline{\underline{B}})$ be Borel spaces. Define a function φ on $P_*(\underline{\underline{A}}\ \underline{\underline{B}})$ into $P(\underline{\underline{P}}(\underline{\underline{B}}))$ as follows: if

$\mu = pq$ then $\varphi(\mu)(E) = p(\{x \in X : q(x) \in E\})$ for all $E \in \underline{\underline{P}}(\underline{\underline{B}})$.

Then φ is well-defined and is measurable (relative to $\underline{\underline{P}}(\underline{\underline{A}}\ \underline{\underline{B}}) \cap P_*(\underline{\underline{A}}\ \underline{\underline{B}})$ and $\underline{\underline{P}}(\underline{\underline{P}}(\underline{\underline{B}}))$).

Proof: Let $\mu = pq$ and $\mu = p'q'$. Then $p = p'$; we must show that $p(\{x: q(x) \in E\}) = p(\{x: q'(x) \in E\})$ for all $E \in \underline{\underline{P}}(\underline{\underline{B}})$. $\underline{\underline{P}}(\underline{\underline{B}})$ is generated by sets of the form $\{\xi \in P(\underline{\underline{B}}) : \xi(B) > c\}$, $B \in \underline{\underline{B}}$ $0 \leq c < 1$. So given $E \in \underline{\underline{P}}(\underline{\underline{B}})$ we can find $B_n \in \underline{\underline{B}}$ and c_n such that E is in the σ-algebra generated by $\{\xi : \xi(B_n) > c_n\}$, $n \geq 1$. Since $q(x, \cdot)$ and $q'(x, \cdot)$ are both conditional probabilities, there is an $N \in \underline{\underline{A}}$ such that $q(x, B_n) = q'(x, B_n)$ if $x \notin N$ for all $n \geq 1$ and $p(N) = 0$. Consequently $p(\{x : q(x) \in E\}) = p(\{x: q'(x) \in E\})$. This shows that φ is well-defined.

To prove φ is measurable, it is enough to show that $\mu \longrightarrow \varphi(\mu)(E)$ is measurable for all $E \in \underline{\underline{P}}(\underline{\underline{B}})$ and it is enough to consider E of the form $\{\xi : \xi(B) > c\}$ where $B \in \underline{\underline{B}}$, $0 \leq c < 1$. Since the indicator function of the set $(c, 1]$ is a limit of continuous functions on $[0, 1]$ and since continuous functions on $[0, 1]$ can be approximated uniformly by polynomials, it is enough for us to show that $\mu \longrightarrow \int g(q(x,B))p(dx)$ is measurable for any polynomial g on $[0, 1]$.

Let $\underline{\underline{C}}$ be a countably generated sub σ-algebra of $\underline{\underline{A}}$ and $\underline{\underline{C}}_n$ be finite sub-algebras of $\underline{\underline{C}}$ such that $\bigcup_{n=1}^{\infty} \underline{\underline{C}}_n$ generates $\underline{\underline{C}}$. Let $q_n(x, B)$ denote a fixed version of $E_\mu(y \in B | \underline{\underline{C}}_n \times Y)(x)$.

Then the maps $\mu \longrightarrow \int g(q_n(x,B))p(dx)$ are measurable and $q_n(x,B) \longrightarrow E_\mu(y\in B|\underline{\underline{C}} \times Y)(x)$ a.s. (p). Consequently the map $\mu \longrightarrow \int g(E_\mu(y\in B|\underline{\underline{C}} \times Y)(x)p(dx)$ is measurable.

Now the family $\{E_\mu(y\in B|\underline{\underline{C}} \times Y): \underline{\underline{C}} \subset \underline{\underline{A}}$ and $\underline{\underline{C}}$ is countably generated$\}$ is a p- martingale which is uniformly bounded. Hence by the martingale convergence theorem [19, pp.86], there is a $\underline{\underline{A}}$ measurable random variable z such that

(i) $E_\mu(z|\underline{\underline{C}} \times Y) = E_\mu(y\in B|\underline{\underline{C}} \times Y)$ for all countably generated $\underline{\underline{C}} \subset \underline{\underline{A}}$ and (ii) $E_\mu(y\in B|\underline{\underline{C}} \times Y) \longrightarrow z$ in $L^1(p)$ norm.

From (i) we conclude that $z = q(\cdot, B)$ a.s (p). From (ii) and the fact that $E_\mu(y\in B|\underline{\underline{C}} \times Y)$ is uniformly bounded by 1 we have $\int | g(E_\mu(y\in B|\underline{\underline{C}} \times Y) - g(q(x,B))|p(dx) \longrightarrow 0$. Hence $\mu \longrightarrow \int g(q(x, B))p(dx)$ is measurable. This completes the proof.

<u>Corollary 13</u>: Let $(X, \underline{\underline{A}})$ and $(Y, \underline{\underline{B}})$ be Borel spaces and let $C\in \underline{\underline{A}} \times \underline{\underline{P}}(\underline{\underline{B}})$. Define $g: P_*(\underline{\underline{A}}\ \underline{\underline{B}}) \longrightarrow R$ by

$g(\mu) = p(\{x\in X: (x, q(x))\in C\})$ if $\mu = pq$. Then g is well-defined and is measurable.

<u>Proof</u>: The fact that g is well-defined can be proved in the same way as we proved ϕ is well-defined in theorem 12. If C is a rectangle, then theorem 12 can be applied to show that g

is measurable. The proof for arbitrary C then follows, as the class of C for which g is measurable is closed under complements and countable disjoint unions.

Now we turn our attention to some questions regarding stopping rules. Let $(X, \underline{\underline{A}})$ be a Borel space and $\mathfrak{A} = \{\underline{\underline{A}}_n : n \geq 0\}$ be a sequence of increasing sub σ-algebras of $\underline{\underline{A}}$. A stopping rule for $\mathfrak{A}$ is a function t on X into the set $\bar{N} = \{0, 1, 2, \ldots \infty\}$ such that $[t \leq n] \in \underline{\underline{A}}_n$ for all $n \geq 0$. Let $T(\mathfrak{A})$ denote the set of all stopping rules. Suppose $\{Y_n : n \in \bar{N}\}$ is a family of uniformly bounded or non-negative real-valued random variables defined on $(X, \underline{\underline{A}})$. Let $v(p,t) = \int Y_{t(x)}(x)p(dx)$ for each $p \in P(\underline{\underline{A}})$, $t \in T(\mathfrak{A})$. Then for each $t \in T(\mathfrak{A})$, $v(\cdot, t)$ is a measurable function. We are interested in knowing if $\sup_{t \in T_o} v(p,t)$ is a measurable function of p for a fixed $T_o \subset T(\mathfrak{A})$. For this purpose, we introduce certain topologies on $T(\mathfrak{A})$. Given any subset M of $P(\underline{\underline{A}})$, define a topology for $T(\mathfrak{A})$ generated by the class consisting of $\{t \in T(\mathfrak{A}) : p[t \neq t_o] < \epsilon\}$, $p \in M$, $\epsilon > 0$, $t_o \in T(\mathfrak{A})$. We shall call this (completely regular) topology the M-topology for $T(\mathfrak{A})$. Suppose T_o is a subset of $T(\mathfrak{A})$ whose induced M-topology is separable, i.e. admits a countable dense subset. Then for any countable dense $T_1 \subset T_o$,

$\sup_{t \in T_1} v(p,t) = \sup_{t \in T_o} v(p,t)$ and consequently $\sup_{t \in T_1} v(p,t)$ is a measurable function on M relative to $\underline{\underline{P}}(\underline{\underline{A}}) \cap M, \underline{\underline{R}}$. Strauch [24] has shown that $T(\mathfrak{A})$ is separable in its $P(\underline{\underline{A}})$-topology if each $\underline{\underline{A}}_n$ in $\mathfrak{A}$ is countably generated. In fact Strauch assumes stop rules to be finite valued, but it is easy to see that his arguments go through even for stop rules in our sense. Since the M-topology contains M'-topology if $M \supset M'$, this shows that $T(\mathfrak{A})$ is separable in its M-topology for each M whenever $\mathfrak{A}$ consists of countably generated σ-algebras. In such a case if the M-topology for $T(\mathfrak{A})$ is pseudometrizable then every $T_o \subset T(\mathfrak{A})$ will be separable in its M-topolog[y]. One situation in which the M-topology for $T(\mathfrak{A})$ is pseudometrizable is when M is dominated by a σ-finite measurable μ. This is because if M is dominated by μ we can get a countable subset $M_o = \{p_n : n \geq 1\}$ of M which is equivalent to M; then the M-topology for $T(\mathfrak{A})$ is the same as the $\{p_o\}$-topology for $T(\mathfrak{A})$ where $p_o = \sum_{n=1}^{\infty} \frac{1}{2^n} p_n$. But the $\{p_o\}$-topology can be pseudometrized by means of the pseudometric $d(t, t') = p_o([t \neq t'])$.

7. Measurable Gambling Problems:

Definitions: A gambling system $\underline{G} = (F, \Gamma, T, u)$ is said to be $\underline{\underline{F}}$ measurable if the following conditions are satisfied: (i) $\underline{\underline{F}}$ is a Borel structure for F; (ii) for each x in W(F) and γ in $\Gamma(x)$, γ is countably additive on $\underline{\underline{F}}$ so that, by identifying gambles which agree on $\underline{\underline{F}}$, we can regard $\Gamma(x)$ also as a subset of $P(\underline{\underline{F}})$; (iii) The set

$$C = \{(x, \gamma): x \in W(F), \gamma \in \Gamma(x)\} \in W(\underline{\underline{F}}) \times \underline{\underline{P}}(\underline{\underline{F}});$$

(iv) $[T = i] \in W(\underline{\underline{F}})$ for $i = 0,1,2$; (v) u is $W(\underline{\underline{F}}) + H(\underline{\underline{F}})$, $\underline{\underline{\bar{R}}}$ measurable.

Let $\underline{G}$ be a $\underline{\underline{F}}$ measurable g.s. We denote by $A_m(\underline{G}, x)$ the set of all $\pi \in A(\underline{G}, x)$ which are $\underline{\underline{F}}$ measurable. Let $A_{sm}(\underline{G}, x) = A_s(\underline{G}, x) \cap A_m(\underline{G}, x)$. These sets may be empty. In most cases we shall assume that they are nonempty by requiring that $\underline{G}$ satisfy the conditions: (a) There is a function ξ on W(F) into $P(\underline{\underline{F}})$ such that $\xi(x) \in \Gamma(x)$ for all x and ξ is $W(\underline{\underline{F}})$, $\underline{\underline{P}}(\underline{\underline{F}})$ measurable; (b) u is either bounded or is of constant sign. Any $\underline{\underline{F}}$ measurable g.s. satisfying (a) and (b) is said to be a standard $\underline{\underline{F}}$ measurable g.s.

Let $\underline{G}$ be a standard $\underline{\underline{F}}$ measurable g.s. . Suppose Σ denotes the space of all $\underline{\underline{F}}$ measurable strategies σ over F such that $\sigma(x) \in \Gamma(x)$ for all x and let Λ denote the space of all $\underline{\underline{F}}$ measurable s.fs a over F such that $a(x) = T(x)$

whenever $T(x) \neq 2$. Then $\Sigma X \Lambda \subset \Lambda_{sm}(\underline{G}, x)$ for all x. Let $D(x) = \Sigma X \Lambda$ and $D'(x) = \Lambda_m(\underline{G}, x)$ for each x. Then it is easily verified that $U_D = U_{D'}$, $\overline{V}_D = \overline{V}_{D'}$ and $\underline{V}_D = \underline{V}_{D'}$.

In what follows let $\underline{G} = (F, \Gamma, T, u)$ be a standard $\underline{\underline{F}}$ measurable g.s. Denote by $M(x)$ the set $\{\mu \in P(H(\underline{\underline{F}})): \mu = p_{\sigma[x]}$ for some $\sigma \in \Sigma\}$ for all $x \in W(F)$. Observe that each μ in $M(x)$ can be written as $\mu = \mu_1 \mu_2 \mu_3 \ldots$ where $\mu_1 \in P(\underline{\underline{F}})$ and $\mu_n \in Q(\underline{\underline{F}} \mid \underline{\underline{F}}^{n-1})$ $n \geq 2$. The space $P_*(\underline{\underline{A}}\ \underline{\underline{B}})$ will be denoted by $P_*(\underline{\underline{F}}^n)$ when $\underline{\underline{A}}$ is $\underline{\underline{F}}^{n-1}$ and $\underline{\underline{B}}$ is $\underline{\underline{F}}$. If there is no possibility of confusion we shall write W, $\underline{\underline{W}}$, H, $\underline{\underline{H}}$ for $W(F)$, $W(\underline{\underline{F}})$, $H(F)$ and $H(\underline{\underline{F}})$ respectively. For each μ in $P(\underline{\underline{H}})$ and $n \geq 1$ let μ^n denote its marginal on $\underline{\underline{F}}^n$, i.e. $\mu^n(A) = \mu(A \times F \times F \times \ldots)$ for $A \in \underline{\underline{F}}^n$. The map $\varphi_n(\mu) = \mu^n$ is then $\underline{\underline{P}}(\underline{\underline{H}})$, $\underline{\underline{P}}(\underline{\underline{F}}^n)$ measurable. Let $P_*(\underline{\underline{H}})$ denote the set of all μ in $P(\underline{\underline{H}})$ such that $\mu^n \in P_*(\underline{\underline{F}}^n)$ for all $n \geq 2$. Let $\underline{\underline{P}}_*(\underline{\underline{H}})$ denote the Borel structure $\underline{\underline{P}}(\underline{\underline{H}}) \cap P_*(\underline{\underline{H}})$. As we observed earlier, $M(x) \subset P_*(\underline{\underline{H}})$ for all x in W. We now examine the question of measurability of these sets.

<u>Theorem 14</u>: The set $M = \{(x, \mu): x \in W, \mu \in M(x)\}$ is in $\underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$.

Proof: Observe that the mapping $c: W \times W \longrightarrow W$ defined by $c(x, y) = x.y$ is $\underline{\underline{W}} \times \underline{\underline{W}}$, $\underline{\underline{W}}$ measurable. Since the set $C = \{(x, \gamma): x \in W, \gamma \in \Gamma(x)\} \in \underline{\underline{W}} \times \underline{\underline{P}}(\underline{\underline{F}})$, this shows that $C_n = \{(x, y, \gamma): x \in W, y \in F^n, \gamma \in \Gamma(x)\} \in \underline{\underline{W}} \times \underline{\underline{F}}^n \times \underline{\underline{P}}(\underline{\underline{F}})$ for all $n \geq 1$. Let $M_o = \{(x, \mu): x \in W, \mu \in P_*(\underline{\underline{H}}), (x, \varphi_1(\mu)) \in C\}$ and for $n \geq 1$, let $M_n = \{(x, \mu): x \in W, \mu \in P_*(\underline{\underline{H}}),$ $\mu^n(y \in F^n : (x, y, \mu_{n+1}(y)) \in C_n) = 1\}$. Clearly $M_o \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$ and by corollary 13, $M_n \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$ for all n. It is easy to check that $M = \bigcap_{n=o}^{\infty} M_n$ and so $M \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$. This completes the proof.

Our next lemma is designed to cover various special cases of interest.

Lemma 15: Let $\mathcal{J}(x) = \{t_{a[x]} : a \in \Lambda\}$ for each x in W and let u be either bounded or non-negative. Then (a) $\mathcal{J}(x)$ is separable in its $P(\underline{\underline{H}})$-topology if $\underline{\underline{F}}$ is countably generated;

(b) If $\mathcal{J}_o(x)$ is a dense subset of $\mathcal{J}(x)$ in its $M(x)$-topology, then
$$\sup_{t \in \mathcal{J}(x)} \int u(x,h|t(h))\mu(dh) = \sup_{t \in \mathcal{J}_o(x)} \int u(x,h|t(h))\mu(dh)$$
for all $\mu \in M(x)$;

(c) If Λ_0 is a countable subset of Λ such that $\mathcal{J}_0(x) = \{t_{a[x]}: a \in \Lambda_0\}$ is a dense subset of $\mathcal{J}(x)$ in its $M(x)$-topology, then the map $\varphi : (x,\mu) \to \sup_{t \in \mathcal{J}(x)} \int u(x,h|t(h))\mu(dh)$ is $\underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$, $\overline{\underline{\underline{R}}}$ measurable.

(d) If $\mathcal{J}_0$ is a countable set which is a dense subset of $\mathcal{J}(x)$ in its $M(x)$-topology for all x, then the map φ is $\underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$, $\overline{\underline{\underline{R}}}$ measurable.

<u>Proof:</u> (a) Consider $X = H$, $\underline{\underline{\Lambda}} = \underline{\underline{H}}$ and $\underline{\underline{\Lambda}}_n = \underline{\underline{F}}^n \times F \times F \times \ldots (n \geq$ Then $\mathcal{O}\!\mathcal{l} = \{\underline{\underline{\Lambda}}_n : n \geq 0\}$ consists of countably generated σ-algebras. Consequently, by the extension of Strauch's result mentioned in the last section, the set $T(\mathcal{O}\!\mathcal{l})$ of all (measurable stop rules is separable in its $P(\underline{\underline{H}})$-topology. If C is any dense set in $T(\mathcal{O}\!\mathcal{l})$ and $t_1, t_2 \in T(\mathcal{O}\!\mathcal{l})$ then it is easy to see that $C' = \{t' : t' = \max(t_1, \min(t, t_2))$ for some $t \in C\}$ is dense in $\mathcal{J} = \{t' : t_1 \leq t' \leq t_2,\ t' \in T(\mathcal{O}\!\mathcal{l})\}$. Given any x in W, set $t_1 = t_{T_*[x]}$ and $t_2 = t_{T^*[x]}$, where T^* is defined by: $T^*(y) = 0$ if $T(y) = 0$ and $= 1$ otherwise. Then $t_1, t_2 \in T(\mathcal{O}\!\mathcal{l})$ and corresponding $\mathcal{J}$ is $\mathcal{J}(x)$. Hence $\mathcal{J}(x)$ is separable in its $P(\underline{\underline{H}})$-topology.

(b) is straightforward.

(c) For each $a \in \Lambda$, the map $\varphi_a(x,\mu) = \int u(x,h|t_{a[x]}(h))\mu(dh)$

is a $\underline{\underline{W}} \times \underline{\underline{P}}(\underline{\underline{H}})$ measurable function on $W \times P(\underline{\underline{H}})$. Consequently $\sup_{a \in \Lambda_o} \varphi_a(x, \mu)$ is also a measurable map on $W \times P(\underline{\underline{H}})$ since Λ_o is countable. But from (b) it follows that

$\varphi(x,\mu) = \sup_{a \in \Lambda_o} \varphi_a(x,\mu)$ on M and by theorem 14 $M \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$.

Therefore φ is $\underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$ measurable.

(d) is proved in exactly the same way as (c). This completes the proof of the lemma.

Remarks: If $(F, \underline{\underline{F}})$ is standard analytic then $P_*(\underline{\underline{H}}) = P(\underline{\underline{H}})$. In d.p. problems $T \equiv 1$ and so Λ is a singleton. In most other cases $T \equiv 2$ and $\underline{\underline{F}}$ is countably generated so that $\mathcal{T}(x) = T(\mathfrak{M})$ is separable and (d) is applicable. Various other possibilities are also covered by the lemma. For any $\Lambda \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$, let $pr(\Lambda)$ denote the projection of Λ onto W. Let $\underline{\underline{W}}^*$ be the Borel structure for W generated by the class $\{pr(\Lambda): \Lambda \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})\}$. If $(F, \underline{\underline{F}})$ is standard analytic, $\underline{\underline{W}}^*$ is such that $\underline{\underline{W}}^* \cap F^n$ is the σ-algebra on F^n generated by all analytic subsets of F^n for $n \geq 1$.

Theorem 16: Suppose $\varphi(x, \mu) = \sup_{a \in \Lambda} \int u(x,h|t_{a[x]}(h))\mu(dh)$ is $\underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}}), \underline{\underline{\bar{R}}}$ measurable on M, then U_D is a $\underline{\underline{W}}^*, \underline{\underline{\bar{R}}}$

measurable function on W.

Proof: $U_D(x) = \sup \{ I_u(\sigma, a)(x) : \sigma \in \Sigma, a \in \Lambda \}$

$$= \sup \{ \varphi \ (x, \mu) : \mu \in M(x) \}$$

So $\{ x : U_D(x) > c \} = pr (\{ (x, \mu) \in M : \varphi \ (x, \mu) > c \})$ for any $c \in R$ and hence is in $\underline{\underline{W}}^*$. This proves that U_D is $\underline{\underline{W}}^*$ measurable.

If $(F, \underline{\underline{F}})$ is standard analytic, U_D is $\underline{\underline{W}}^*$ measurable shows that U_D restricted to each F^n is measurable with respect to the completion of any probability on $\underline{\underline{F}}^n$. The situation regarding $\overline{V}_D$ is even simpler. For each σ in $D^1(x)$ it is easy to see that $\overline{J}_D(\sigma)(x) = I(\sigma, T^*)(x)$. Consequently T^* being measurable, using theorem 14, we can show that $\overline{V}_D(x)$ is a $\underline{\underline{W}}^* \ \underline{\underline{\overline{R}}}$ measurable function on W for any reward function u - bounded, positive or negative. Similarly for $\underline{V}_D$.

We now show that the situation in two-person gambling systems regarding $\overline{W}_D$ and $\underline{W}_D$ is very different. Suppose we are given a two-person g.s. $\underline{G} = (F, \Gamma, T, u, X)$ which is standard $\underline{\underline{F}}$ measurable in the sense that $\underline{G}' = (F, \Gamma, T, u)$ is standard $\underline{\underline{F}}$ measurable and $X \in W(\underline{\underline{F}})$. Assume further that $T \equiv 1$ which is the simplest situation as Λ then becomes a singleton. The question we ask is whether $\overline{W}_D$ and $\underline{W}_D$ are necessarily $\underline{\underline{W}}^*$ measurable where, as before, $D = \Sigma \times \Lambda$. We show below that even if we assume Γ to be deterministic and $(F, \underline{\underline{F}})$ to be a standard Borel space, the measurability of $\overline{W}_D$ and $\underline{W}_D$ can not be settled so easily.

Consider the following auxiliary problem: given a bounded Borel measurable function g on the unit cube $I \times I \times I$ where $I = [0, 1]$, let $f(x) = \sup_{y \in I} \inf_{z \in I} g(x,y,z)$, $x \in I$; we know that f need not be Borel measurable on I. But is it at least Lebesgue measurable? The relevance of this to the problem in the last paragraph is quite obvious. We now settle this latter question.

<u>Theorem 17:</u> The following two statements are equivalent:

i) There is a PCA set $E \subset I$ which is not Lebesgue measurable.

ii) There is a bounded Borel measurable function g on $I \times I \times I$ such that $f(x) = \sup_{y \in I} \inf_{z \in I} g(x,y,z)$ is not a Lebesgue measurable function of x.

<u>Proof:</u> To show (i) -> (ii): Let E be a non Lebesgue measurable PCA subset of I. So there is a coanalytic subset D of $I \times I$ such that E is the projection of D to the first coordinate axis. Now $I \times I - D$ being analytic,

we can find a Borel subset B of $I \times I \times I$ such that D is the projection of B to the first two coordinate spaces. In other words, for any $x \in I$, x is in E iff for some $y \in I$, there is no $z \in I$ such that $(x,y,z) \in B$. Define g to be the indicator function of B. Then f is the indicator function of E and is non Lebesgue measurable.

To show (ii) $\longrightarrow$ (i): It is enough to show that for any $c \in R$, the set $[f > c]$ is a $P\ C\ A$ subset of I. Let

$h(x,y) = \inf_{z \in I} g(x,y,z)$. Then $[f > c] = \text{Proj}_I\,[h > c]$

where Proj_I denotes the projection to the first coordinate axis.

$[h > c] = I \times I - \bigcap_{n=1}^{\infty} [h < c + \frac{1}{n}]$ and

$[h < c + \frac{1}{n}] = \text{Proj}_{I \times I}\,[g < c + \frac{1}{n}]$ where $\text{Proj}_{I \times I}$ stands for the projection to the first two coordinate axes. Since g is Borel measurable, $[g < c + \frac{1}{n}]$ is Borel ; so $[h < c + \frac{1}{n}]$ is analytic for each n. Hence $[h > c]$ is coanalytic and its projection $[f > c]$ is a $P\ C\ A$ set. This proves the theorem.

<u>Remarks:</u> (1) It is known that the statement (i) of the theorem can be proved using the axiom of constructibility and is therefore consistent with the usual axioms of set theory.

This shows that we can not hope to prove that $\underline{W}_D$ is $\underline{\underline{W}}^*$ measurable even in very special cases. For, let $F = I$, $\underline{\underline{F}}$ = Borel σ-algebra on I, $T \equiv 1$ and $\Gamma(x) = \{\partial(f): f \in I\}$ for all $x \in W$. Define $X = I \times I$ and $u(f_1, f_2, f_3, \ldots) = g(f_1, f_2, f_3)$ where g is a Borel measurable function on $I \times I \times I$. It is then easy to check that $\underline{W}_D(f) = \underline{W}_D(f) = \sup_{f' \in I} \inf_{f'' \in I} g(f, f', f'')$ on I. Consequently $\underline{W}_D$ on I need not be Lebesgue measurable and so $\underline{W}_D$ need not be $\underline{\underline{W}}^*$ measurable. Similar remarks apply to $\overline{W}_D$.

(2) In the deterministic situation considered in remark (1), whatever measurable X and u are given, we can find a suitable Borel measurable function v on $W(I) \times H(I) \times H(I)$ such that $\underline{W}_D(x) = \sup_{h \in H} \inf_{h' \in H} v(x,h,h')$. This can be shown as follows: let $h = (f_1, f_2, \ldots)$, $h' = (f_1', f_2', \ldots)$ and $x \in W(I)$. Define $v(x,h,h') = u(x,h'')$ where $h'' = (f_1'', f_2'', \ldots)$ and f_n'' is f_n if $xf_1'' \ldots f_{n-1}'' \in X$ and f_n'' is f_n' otherwise. The maps $(x,h,h') \longrightarrow f_n''$ are measurable as can be verified by induction on n. Consequently v is Borel measurable and it is easy to see that $\underline{W}_D(x) = \sup_h \inf_{h'} v(x,h,h')$. This proves that the set $[\underline{W}_D > c]$ is a P C A set for all $c \in R$. We shall see later that the same result holds for nondeterministic problems also provided $(F, \underline{\underline{F}})$ is standard analytic.

8. Measurability of the set M^o:

In this section we consider some measurability problems when the permissible class D_o is smaller than D (we are still using the notations of section 7). It is clear that the techniques of Lemma 15 and Theorem 16 are applicable for any such D_o; thus for instance if $M' = \{(x,\mu): \mu = p_{\sigma[x]}$ for some σ in $D_o^1(x)\}$ is in $\underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$ and if $T \equiv 1$ then U_{D_o} and $\overline{V}_{D_o}$ are $\underline{\underline{W}}^*$ measurable. Therefore it is of interest to see under what conditions M' is measurable. We shall discuss this in certain situations which are frequently encountered in applications. Suppose $\underline{G} = (F, \Gamma, T, u)$ is a standard $\underline{\underline{F}}$ measurable g.s. with $T \equiv 1$ and suppose $\Gamma(x)$ does not depend on x fully - for instance, it might depend on the last coordinate of x or on first and last coordinates of x etc. These cases can be considered together by assuming that there is a sub σ-algebra $\underline{\underline{W}}^o$ of $\underline{\underline{W}}$ such that the set $C = \{(x, \gamma):$ $x \in W, \gamma \in \Gamma(x)\}$ belongs to $\underline{\underline{W}}^o \times \underline{\underline{P}}(\underline{\underline{F}})$. Let Σ^o denote the set of all strategies σ in Σ which are $\underline{\underline{W}}^o$, $\underline{\underline{P}}(\underline{\underline{F}})$ measurable. Assuming Σ^o to be nonempty consider the set $M^o = \{(x, \mu):$ $\mu = p_{\sigma[x]}$ for some σ in $\Sigma^o\}$. Clearly M^o corresponds to the permissible class D_o defined by $D_o(x) = \{(\sigma, a):$ $\sigma \in \Sigma^o, \quad a \equiv 1\}$. Hence if M^o is measurable, i.e. if

$M^o \in \mathcal{W} \times \mathcal{P}_*(\mathcal{H})$ then we can conclude that U_D is $\mathcal{W}^*$ measurable.

However the question of measurability of M^o is far too general to admit an easy solution. In this connection we indicate a class of problems many of which may be regarded as special cases of the measurability problem of M^o. Let $\Omega = H(R)$, $\mathcal{A} = H(\mathcal{R})$ and x_n denote the coordinate random variables on Ω, $n \geq 1$. Consider the five sets of probability measures μ on $\mathcal{A}$ which make $\{x_n, n \geq 1\}$ respectively-(i) independent (ii) identically distributed (iii) stationary (iv) martingale and (v) Markov chain. Are these sets measurable subsets of $P(\mathcal{A})$? It is fairly easy to show that the sets (i) - (iv) are all measurable. The fact that $\mathcal{R}$ is countably generated is repeatedly made use of in the above proofs. We shall now show that the set in case (v) is an analytic subset of $P(\mathcal{A})$. Though this result is not of immediate use to us the technique used in its proof can be applied in many important situations.

<u>Theorem 18</u>: Let $\Omega = H(R)$, $\mathcal{A} = H(\mathcal{R})$ and x_n denote the coordinate random variables on Ω $(n \geq 1)$. Let $K = \{\mu \in P(\mathcal{A}) : \{x_n, n \geq 1\}$ is a Markov chain under $\mu\}$. Then K is an analytic subset of $P(\mathcal{A})$.

<u>Proof:</u> As is well-known, $(R, \mathcal{R})$ and $(P(\mathcal{R}), \mathcal{P}(\mathcal{R}))$ are both uncountable standard Borel spaces and are consequently Borel isomorphic. Let $g : R \longrightarrow P(\mathcal{R})$ be a Borel isomorphism, i.e. g is

one-to-one, onto and g, g^{-1} are measurable. Let g and g^{-1} be of Borel class β where $\beta < \omega_1$.

We have shown in theorem 10 that each probability measure λ on $\underline{\underline{R}} \times \underline{\underline{R}}$ can be written as $\lambda = pq$ where $p \in P(\underline{\underline{R}})$, $q \in Q(\underline{\underline{R}}|\underline{\underline{R}})$ and $q(\cdot, A)$ is of Borel class 3 as a function on R onto R for every A in a countable generator for $\underline{\underline{R}}$. As we remarked then, this fact implies the existence of an $\alpha < \omega_1$ such that each $\lambda \in P(\underline{\underline{R}} \times \underline{\underline{R}})$ is of the form pq where $p \in P(\underline{\underline{R}})$ and $q : R \longrightarrow P(\underline{\underline{R}})$ is of Borel class α. Now choose and fix a Borel measurable function U on $R \times R$ into R such that every Borel measurable function f on R into R of Borel class $\alpha + \beta$ is a section of U, i.e. for some $a \in R$ $f = U(a, \cdot)$. Such a universal function exists [13, pp. 369,393]. Define $V : R \times R \longrightarrow P(\underline{\underline{R}})$ by $V(a, x) = g(U(a, x))$. Then V is Borel measurable. Further if $q : R \longrightarrow P(\underline{\underline{R}})$ is of Borel class α, let $f : R \longrightarrow R$ be defined by $f(x) = g^{-1}(q(x))$. Then f is of Borel class $\alpha + \beta$ and so $f = U(a, \cdot)$ for some $a \in R$. Therefore $q = V(a, \cdot)$. Consequently any $\lambda \in P(\underline{\underline{R}} \times \underline{\underline{R}})$ can be written as $\lambda = p.V(a, \cdot)$ for some $p \in P(\underline{\underline{R}})$ and $a \in R$.

Consider the map $\ell : P(\underline{\underline{R}}) \times \Omega \longrightarrow P(\underline{\underline{A}})$ defined by $\ell(p, a_1, a_2, \ldots) = p.V(a_1, \cdot)\, V(a_2, \cdot)\, V(a_3, \cdot) \ldots$. ℓ is Borel measurable and it is easy to verify that $K = \text{Proj}_I \{(\mu, p, a_1, a_2, \ldots) : \mu = \ell(p, a_1, a_2, \ldots)\}$ where

Proj_I denotes the projection to the $P(\underline{\underline{A}})$ axis. Hence K is analytic. This completes the proof.

The above proof can be modified to show that many other interesting classes of measures on $P(\underline{\underline{A}})$ are also analytic. Similarly in d.p. problems the sets of measures corresponding to Markov policies, semi-markov policies, stationary policies etc., can be shown to be analytic by analogous arguments.

Turning our attention back to general M^o, we shall now establish the measurability of M^o in a relatively simple situation.

Theorem 19: Suppose $\underline{\underline{F}}$ is countably generated, $F^n \in \underline{\underline{W}}^o$ for all $n \geq 0$ and there are countably many $\underline{\underline{W}}^o$-atoms $E_m (m \geq 1)$ such that $W = \bigcup_{m=1}^{\infty} E_m$. Then $M^o \in \underline{\underline{W}} \times \underline{\underline{P}}_* (\underline{\underline{H}})$.

Proof: Define $\varphi_n : W \times P(\underline{\underline{H}}) \longrightarrow P(\underline{\underline{W}} \times \underline{\underline{F}}^n)$ by $\varphi_n(x, \mu) = \partial(x) \times \mu^n$. Clearly φ_n is measurable for all $n \geq 1$. Let $c : W \times W \longrightarrow W$ be defined by $c(x,y) = x.y$. c is $\underline{\underline{W}} \times \underline{\underline{W}}$, $\underline{\underline{W}}$ measurable. Fix $n \geq 0$, $k \geq 1$ and a $\underline{\underline{W}}^o$ atom $E \subset F^{n+k}$ and let $A_o = c^{-1}(E)$. Define $L_{n,k}(E) = \{(x, \mu) : x \in F^n$, $\mu \in P(\underline{\underline{H}})$ and for any $B \in \underline{\underline{F}}$, $A \in c^{-1}(E) \cap \underline{\underline{F}}^{n+k}$,

$\varphi_{k+1}(x,\mu)(A \times B)\varphi_k(x,\mu)(A_o) = \varphi_{k+1}(x,\mu)(A_o \times B)\ \varphi_k(x, \mu)(A)\}$.

Since $\underline{\underline{F}}$ is countably generated, each $\underline{\underline{F}}^{n+k}$ is countably generated; consequently in the definition of $L_{nk}(E)$ it is enough to vary Λ and B over countable fields generating $\bar{c}^{1}(E) \cap \underline{\underline{F}}^{n+k}$ and $\underline{\underline{F}}$ respectively. This fact and the measurability of the maps φ_k and φ_{k+1} show that $L_{nk}(E) \in \underline{\underline{F}}^n \times \underline{\underline{P}}(\underline{\underline{H}})$. Let L_{nk} denote the intersection of $L_{nk}(E)$ over all $\underline{\underline{W}}^o$ atoms E contained in F^{n+k}. Note that each $\underline{\underline{W}}^o$-atom is contained in some F^m.

Set $L_n = \bigcap_{k=1}^{\infty} L_{nk}$ and $L = \bigcup_{n=o}^{\infty} L_n$. Each L_{nk} is measurable as there are only countably many $\underline{\underline{W}}^o$-atoms and so $L \in \underline{\underline{W}} \times \underline{\underline{P}}(\underline{\underline{H}})$. We shall now show that $M^o = L \cap M$ where $M = \{(x, \mu) : \mu = p_{\sigma[x]}$ for some $\sigma \in \Sigma\}$. Since $M \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$, this will show that $M^o \in \underline{\underline{W}} \times \underline{\underline{P}}_*(\underline{\underline{H}})$.

(a) Let $(x, \mu) \in M^o$. So $\mu = p_{\sigma[x]}$ for some $\sigma \in \Sigma^o$. Clearly $\sigma \in \Sigma$ and hence $(x, \mu) \in M$. Let $x \in F^n$, $k \geq 1$ and E be a $\underline{\underline{W}}^o$-atom with $E \subset F^{n+k}$. We must show that $(x, \mu) \in L_{nk}(E)$. Let $\Lambda \in \bar{c}^{1}(E) \cap \underline{\underline{F}}^{n+k}$ and $C = \{y \in F^k : (x,y) \in \Lambda\}$.

Observe that $\varphi_{k+1}(x,\mu)(\Lambda \times B) = \mu^{k+1}(C \times B) = \int_C \mu_{k+1}(y,B)\mu^k(dy)$

$$= \int_C \sigma(xy, B)\, \mu^k(dy).$$

But if $y \in C$ then $(x, y) \in \Lambda$ and $c(x,y) = x.y \in E$. As σ is $\underline{\underline{W}}^o$, $P(\underline{\underline{F}})$ measurable, σ is constant on E. Hence

$\sigma(xy, B) = \xi(B)$ for some $\xi \in P(\underline{\underline{F}})$. So $\varphi_{k+1}(x,\mu)(\Lambda \times B) =$ $= \xi(B) . \mu^k(C)$. Similarly $\varphi_{k+1}(x,\mu)(\Lambda_o \times B) = \xi(B) . \mu^k(C_o)$ where $C_o = \{y \in F^k : (x,y) \in \Lambda_o\}$. Since $C \subset C_o$ if $\mu^k(C_o) = 0$ the required equality in $L_{\mu^k}(E)$ is trivial to check.

Otherwise $\xi(B) = \dfrac{\varphi_{k+1}(x,\mu)(\Lambda_o \times B)}{\mu^k(C_o)}$. Therefore

$\varphi_{k+1}(x,\mu)(\Lambda \times B)\ \mu^k(C_o) = \varphi_{k+1}(x,\mu)(\Lambda_o \times B)\mu^k(C)$. However $\mu^k(C_o) = \varphi_k(x,\mu)(\Lambda_o)$ and $\mu^k(C) = \varphi_k(x,\mu)(A)$. We thus have $(x,\mu) \in L_{nk}(E)$. So $M^o \subset L \cap M$.

(b) Let $(x,\mu) \in L \cap M$. $\mu = p_{\sigma[x]}$ for some $\sigma \in \Sigma$. Since we have assumed that Σ^o is nonempty, choose a $\sigma^* \in \Sigma^o$. Suppose $x \in F^n$. Define a strategy σ' as follows:

$$\sigma'(xy)(B) = \varphi_{k+1}(x,\mu)(A_o \times B)/\mu^k(C_o) \quad \text{if } y \in C_o \text{ and } \mu^k(C_o) > 0$$

$$= \sigma^*(xy)(B) \quad \text{if } y \in C_o \text{ and } \mu^k(C_o) = 0$$

$\sigma'(x') = \sigma(x)$ if x and x' belong to same $\underline{\underline{W}}^o$ atom and $\sigma'(z) = \sigma^*(z)$ for other $z \in W$. In this definition A_o and C_o have the same meaning as in last paragraph. Then σ' is $\underline{\underline{W}}^o$, $\underline{\underline{P}}(\underline{\underline{F}})$ measurable. It is enough to show that $p_{\sigma[x]} = p_{\sigma'[x]}$. If $\eta = p_{\sigma'[x]}$ then $\eta^1 = \sigma'(x) = \sigma(x) = \mu^1$. Assume that $\eta^k = \mu^k$; we shall prove that $\eta^{k+1} = \mu^{k+1}$. For any $C \in \underline{\underline{F}}^k$, $B \in \underline{\underline{F}}$,

$$\eta^{k+1}((C \cap C_o) \times B) = \int_{C \cap C_o} \sigma'(xy)(B)\, \eta^k(dy)$$

$$= \varphi_{k+1}(x,\mu)(A_o \times B) \cdot \eta^k(C \cap C_o)/\mu^k(C_o) \qquad \text{if } \mu^k(C_o) > 0$$

$$= \varphi_{k+1}(x,\mu)(A \times B) \quad \text{where } A = \{(x,y) : y \in C \cap C_o\}, \text{ since}$$

$(x,\mu) \in L$ and $\eta^k(C \cap C_o) = \mu^k(C \cap C_o)$

$$= \mu^{k+1}((C \cap C_o) \times B).$$

If $\mu^k(C_o) = 0$ then $\eta^k(C_o) = 0$ and so $\eta^{k+1}((C \cap C_o) \times B) = 0 = \mu^{k+1}((C \cap C_o) \times B)$. Since this is true for each C_o corresponding to an atom $E \subset F^{n+k}$ and since there are only countably many atoms, $\eta^{k+1}(C \times B) = \mu^{k+1}(C \times B)$ for each

$C \in \underline{\underline{F}}^k$, $B \in \underline{\underline{E}}$ This proves that $\eta^{k+1} = \mu^{k+1}$. Thus $\eta = \mu$ and so $(x, \mu) \in M^o$. This completes our proof.

Remark: Suppose I is a countable set and $\underline{\underline{I}} = S(I)$. Let ρ_k be any equivalence relation on I^k for each $k \geq 1$. For any probability measure η on $H(\underline{\underline{I}})$ let $\eta_k(x_1, \ldots, x_k)$ denote the conditional distribution of x_{k+1} given $x_1, \ldots, x_k$. Consider the set Q of all probability measures η on $H(\underline{\underline{I}})$ such that for all k and $x_1, \ldots, x_k$, $\eta_k(x_1, \ldots, x_k) = \eta_k(x_1', \ldots, x_k')$ whenever $(x_1', \ldots, x_k') \rho_k (x_1, \ldots, x_k)$. The above theorem shows that Q is a measurable subset of $P(H(\underline{\underline{I}}))$ By defining ρ_k such that $(x_1, \ldots, x_k) \rho_k (x_1', \ldots, x_k')$ iff $x_k = x_k'$ we see that the corresponding Q is the set of probability measures making the coordinate random variables $\{x_n, \ n \geq 1\}$ a Markov chain with discrete state space I.

We conclude this section with a result regarding two-person measurable gambling systems. Let $\underline{G} = (F, \Gamma, T, u, X)$ be a standard $\underline{\underline{F}}$ measurable two-person g.s such that $T \equiv 1$ and $(F, \underline{\underline{F}})$ is a standard Borel space. Assume moreover that for some nonempty proper subset N_o of natural numbers $X = \bigcup_{n \in N_o} F^n$. Let $D = \Sigma \times \Lambda$ in usual notations. Then we have the following

<u>Theorem 20:</u> For each $c \in R$, the set $\{x: \underline{W}_D(x) > c\}$ is a P C A set in $W(F)$.

<u>Proof:</u> If F is countable there is nothing to show. So let F be uncountable and because of Borel isomorphism we may assume without loss of generality that F is the unit interval $I = [0, 1]$ and $\underline{\underline{F}} = \underline{\underline{I}}$ the Borel σ-algebra on I. As in the proof of theorem 18, we can get an $\alpha_n < \omega_1$ such that each $\lambda \in P(\underline{\underline{I}}^{n+1})$ can be expressed as $\lambda = pq$ where $p \in P(\underline{\underline{I}}^n)$ and $q : I^n \longrightarrow P(\underline{\underline{I}})$ is a Borel measurable function of Borel class α_n, for each $n \geq 1$. Let g be a fixed Borel isomorphism of I onto $P(\underline{\underline{I}})$ such that both g and g^{-1} are of Borel class say β ($\beta < \omega_1$). Let $\alpha = \sup_{n \geq 1} \alpha_n$ and choose a Borel measurable function U_n on $I \times I^n$ into I such that each Borel measurable f on I^n into I of Borel class $\alpha + \beta$ is some section $U_n(a, \cdot)$ of U_n. Such U_n exists for each $n \geq 1$. Define $V_n: I \times I^n \longrightarrow P(\underline{\underline{I}})$ by $V_n(a, x_1, \ldots, x_n) =$ $g(U_n(a, x_1, \ldots, x_n))$. Then V_n is Borel measurable and each $q : I^n \longrightarrow P(\underline{\underline{I}})$ of Borel class α is some section $V_n(a, \cdot)$ of V_n. Consequently each $\lambda \in P(\underline{\underline{I}}^{n+1})$ can be expressed as $\lambda = p\, V_n(a, \cdot)$ for some $p \in P(\underline{\underline{I}}^n)$ and $a \in I$. Hence for each $\mu \in P(H(\underline{\underline{I}}))$ we can find $p \in P(\underline{\underline{I}})$ and a sequence $a_1, a_2, \ldots$

from I such that $\mu = p\, V_1(a_1, \cdot)\ V_2(a_2, \cdot)\ V_3(a_3, \cdot) \ldots$.

Let $\chi : P(\underline{\underline{I}}) \times H(I) \longrightarrow P(H(\underline{\underline{I}}))$ be defined by

$\chi(p, a_1, a_2, \ldots) = p\, V_1(a_1, \cdot)\ V_2(a_2, \cdot) \ldots$ and

$\varphi : W(I) \times P(\underline{\underline{I}}) \times H(I) \longrightarrow \bar{R}$ by $\varphi(x, p, \omega) = \int u(x, h)\, \chi(p, \omega) dh$

Then χ and φ are Borel measurable. We have already shown that the set $M = \{(x, \mu) : \mu = p_{\sigma[x]}$ for some $\sigma \in \Sigma\}$ is Borel measurable. Let $c \in R$ be given. Define

$\varphi_c(x, p, \omega) = \varphi(x, p, \omega)$ if $(x, \xi(p, \omega)) \in M$ and

$= -\left|\frac{c}{2}\right|$ otherwise. Then φ_c is also Borel measurable.

Now we shall show that the sets $[\underline{W}_D > c] \cap X$ and $[\underline{W}_D > c] \cap Y$ are both $P\,\check{C}\,A$ sets, Y being $W(I) - X$. It can be verified that $[\underline{W}_D > c] \cap X =$

$\{x \in X : \sup \inf \varphi_c(x, p, \omega) > c\}$ where the supremum is taken over $p \in P(\underline{\underline{I}})$ and sequences $\{a_n : n \in N_o\}$ from I and the infimum is over sequences $\{a_n : n \notin N_o\}$ from I. As φ_c is a Borel measurable function of x, ~~p,~~ ω it is clear that the set $[\underline{W}_D > c] \cap X$ is a $P\,\check{C}\,A$ subset of $W(I)$. Similarly the set $[\underline{W}_D > c] \cap Y = \{x \in Y : \sup \inf \varphi_c(x, p, \omega) > c\}$ where the supremum is taken over sequences $\{a_n : n \in N_o\}$ from I and the infimum over $p \in P(\underline{\underline{I}})$ and sequences $\{a_n : n \notin N_o\}$ from I.

So $[W_{-D} > c] \cap Y$ is also a P C A set. Consequently $[\underline{W}_D > c]$ is a P C A subset of W(I). This completes the proof of the theorem.

Remarks: (1) The sets $[\overline{W}_D > c]$ can be seen to be C P C A subsets. If we replace D by any permissible class such that the corresponding set M is Borel measurable then the same result can be proved for such a class also. Thus for instance in stochastic games [18] the natural permissible class is such that the corresponding M is measurable. Hence the upper and lower value functions of a stochastic game are measurable with respect to the σ-algebra generated by P C A sets.

(2) How far the assumptions $T \equiv 1$ and $X = \bigcup_{n \in N_0} E^n$ are needed is not known.

(3) It is known that under the axiom of existence of a measurable cardinal, every P C A set in I is Lebesgue measurable. So, in such a case, the upper and lower values of a stochastic game are Lebesgue measurable.

9. Optimal reward in measurable problems

Strauch [25] has shown that in the d.p. problem, any measurable policy can be replaced by a random semi-markov policy with the same total expected return. We shall first prove similar results for measurable gambling problems. It should be noted that such results provide another approach for deciding if U_D is measurable in problems of section 7.

The result of Strauch, stated above, proceeds by taking suitable conditional distributions. As these conditional distributions need not exist in the setup considered so far, we have to impose more restrictions on the class of measurable gambling systems. We first consider a situation very similar to the d.p. problem: Let $\underline{G} = (F, \Gamma, T, u)$ be a standard $\underline{\underline{F}}$ measurable g.s. where $(F, \underline{\underline{F}})$ is a standard analytic space and $T \equiv 1$. For each x in W we assume that $\Gamma(x)$ is a closed convex subset of $P(\underline{\underline{F}})$ in its weak topology. For each $n \geq 1$ we are given a subset $C_n = \{i_1(n), \ldots, i_{k_n}(n)\}$ of $\{1, 2, \ldots, n\}$ such that $\Gamma(f_1, \ldots, f_n)$ and $u(f_1, \ldots, f_n)$ depend only on the coordinates f_j, $j \in C_n$. i.e. if $x = (f_1, \ldots, f_n)$ and $x' = (f_1', \ldots, f_n')$ are such that $f_{i_j}(n) = f'_{i_j}(n)$ for $1 \leq j \leq k_n$ then $\Gamma(x) = \Gamma(x')$ and $u(x) = u(x')$. Further we assume that $C_{n+1} \subset C_n \cup \{n+1\}$ for $n \geq 1$ and that $u(h) = \sum_{n=1}^{\infty} u(h/n)$ for every $h \in H$. Then we have the following

<u>Theorem 21:</u> Let $\sigma \in \Sigma$, i.e. σ is an $\underline{\underline{F}}$ measurable strategy such that $\sigma(x) \in \Gamma(x)$ for all x. Then there is a strategy

σ^* in Σ such that (i) $I_u(\sigma^*, T)(\emptyset) = I_u(\sigma, T)(\emptyset)$ and (ii) for each $n \geq 1$, $\sigma^*(f_1,\ldots,f_n)$ depends only on $f_{i_1},\ldots, f_{i_k}$ if $C_n = \{i_1,\ldots, i_k\}$.

Proof: Let $\sigma^*(\emptyset) = \sigma(\emptyset)$ and for each $n \geq 1$, let $\sigma^*(f_1,\ldots,f_n)$ be the conditional distribution under p_σ of f_{n+1} given $f_{i_1},\ldots, f_{i_k}$ (for typographic convenience we write $i_1,\ldots,i_k$ instead of $i_1(n),\ldots, i_{k_n}(n)$. Then σ^* clearly satisfies (ii).

We shall now show that $\sigma^*(x) \in \Gamma(x)$ for all x. If g is any function on F^n such that $g(f_1,\ldots, f_n)$ depends only on $f_{i_1},\ldots, f_{i_k}$ then we write $g_n(f_{i_1},\ldots, f_{i_k})$ for $g(f_1,\ldots, f_n)$. Let $\mu_n(f_{i_1},\ldots, f_{i_k})$ be the conditional distribution under p_σ of the variables $f_1,\ldots, f_n$ given $f_{i_1},\ldots, f_{i_k}$. It is then easy to check that for all $A \in \underline{\underline{F}}$,

$$\sigma^*_n(f_{i_1},\ldots, f_{i_k})(A) = \int_{F^n} \sigma(f_1,\ldots,f_n)(A)\,d\mu_n(\cdot\,|f_{i_1},\ldots,f_{i_k}).$$

Since $\sigma(f_1,\ldots,f_n) \in \Gamma_n(f_{i_1},\ldots,f_{i_k})$ and $\Gamma_n(f_{i_1},\ldots,f_{i_k})$ is closed and convex, theorem 11 applies to show that $\sigma^*_n(f_{i_1},\ldots,f_{i_k}) \in \Gamma_n(f_{i_1},\ldots, f_{i_k})$. This is true for all $n \geq 1$ and $(f_1,\ldots,f_n)$. Also $\sigma^*(\emptyset) = \sigma(\emptyset) \in \Gamma(\emptyset)$. Hence $\sigma^* \in \Sigma$.

As $u(f_1, f_2, \ldots) = \sum_{n=1}^{\infty} u_n(f_{i_1}, \ldots, f_{i_k})$ it is enough to show that $\int u_n(f_{i_1}, \ldots, f_{i_k}) dp_\sigma = \int u_n(f_{i_1}, \ldots, f_{i_k}) dp_{\sigma*}$ for all $n \geq 1$. We shall show that for any $n \geq 1$ and any $\underline{\underline{F}}^n$ measurable function $g_n(f_{i}, \ldots, f_{i_k})$ on F^n which is either bounded or of constant sign, $\int g_n(f_{i_1}, \ldots, f_{i_k}) dp_\sigma = \int g_n(f_{i_1}, \ldots, f_{i_k}) d$

The marginal distribution of f_1 under p_σ as well as under $p_{\sigma*}$ being same, the above assertion is true for $n = 1$. Assuming the assertion to be valid for n consider g_{n+1} on F^{n+1}. If $C_{n+1} = C_n$ there is nothing to show. Otherwise $C_{n+1} = C_n \cup \{n+1\}$ and so g_{n+1} is a function of $f_{i_1}, \ldots, f_{i_k}, f_{n+1}$. Let E_σ and $E_{\sigma*}$ denote expectations under p_σ and $p_{\sigma*}$ respectively.

$$E_\sigma[g_{n+1}(f_{i_1}, \ldots, f_{i_k}, f_{n+1})] = E_\sigma\{ E_\sigma[g_{n+1}(f_{i_1}, \ldots, f_{i_k}, f_{n+1}) / f_{i_1}, \ldots, f_{i_k}] \}$$

$= E_\sigma[g'_n(f_{i_1}, \ldots, f_{i_k})]$ where

$$g'_n(f_{i_1}, \ldots, f_{i_k}) = E_\sigma[g_{n+1}(f_{i_1}, \ldots, f_{i_k}, f_{n+1}) / f_{i_1}, \ldots, f_{i_k}]$$

$$= \int_F g_{n+1}(f_{i_1}, \ldots, f_{i_k}, f_{n+1}) \sigma^*_n(f_{i_1}, \ldots, f_{i_k}, d$$

by the property of conditional distribution.

By induction hypothesis, $E_\sigma[g_n'(f_{i_1}, \ldots, f_{i_k})] = E_{\sigma*}[g_n'(f_{i_1}, \ldots, f_{i_k})]$

$= \int g_{n+1}(f_{i_1}, \ldots, f_{i_k}, f_{n+1})\sigma_n^*(f_{i_1}, \ldots, f_{i_k})(df_{n+1})dp_{\sigma*}$

$= E_{\sigma*}[g_{n+1}(f_{i_1}, \ldots, f_{i_k}, f_{n+1})]$. This proves the assertion for for $n+1$ and completes the proof of the theorem.

A rather obvious modification of the above proof yields the following result.

<u>Theorem 22:</u> Let $\sigma \in \Sigma$ and $m \geq 1$. Then there is a $\sigma^* \in \Sigma$ such that (i) $I_u(\sigma^*, T)(y) = I_u(\sigma, T)(y)$ for all $y \in F^m$ and (ii) for each $n \geq m$, $\sigma^*(f_1, \ldots, f_n)$ depends only on $f_1, \ldots, f_m, f_{i_1}, \ldots, f_{i_k}$ where $C_n = \{i_1, \ldots, i_k\}$.

<u>Remark:</u> (1) In the d.p. problem an application of Theorem 22 gives the result of Strauch referred to at the beginning of this section. To see this we note first that the reward function $r(s,a,s')$ can, without loss of generality be assumed to depend only on s,a by replacing it by $r_1(s,a) = \int r(s,a,s')\, q(ds'|s,a)$. This does not alter the expected reward from any plan. Now define $C_{2n-1} = \{2n-1\}$ and

$C_{2n} = \{2n-1, 2n\}$ for all $n \geq 1$. Then the assumptions of theorem 22 are satisfied and the conclusion with $m=1$ shows that for any measurable plan π there is a random semi-markov plan π' such that $I(\pi')(s) = I(\pi)(s)$ for all $s \in S$. Similarly, given a probability measure p on S, we can let $\Gamma(\emptyset) = \{p\}$ and apply theorem 21 to conclude that there is a random Markov plan π'' such that $pI(\pi) = pI(\pi'')$.

(2) An example due to Blackwell [25, Example 4.1] shows that in the d.p. situation, given a measurable plan π there need not exist any Markov plan π' such that $I(\pi)(s) = I(\pi')(s)$ for all $s \in S$. The same example therefore shows that in Theorem 22 condition (ii) can not be strengthened to say that for each $n \geq m$, $\sigma^*(f_1, \ldots, f_n)$ depends only on $f_{i_1}, \ldots, f_{i_k}$.

(3) Observe that theorems of 21 and 22 go through whenever the conditional distributions σ^*_n and μ_n exist for all n. Thus for instance, if each gamble in $\Gamma(x)$ is concentrated on a countable set and is countably additive $(x \in W)$ then the above theorems would apply with $\underline{\underline{F}} = S(F)$. Ornstein has considered this situation in [21]; theorem 22 applied to his problem would imply (using his notations) that if $V(x)$ is closed and convex for all $x \in X$, then given any strategy s we can find a semi-markov strategy t such that $F_s(x) = F_t(x)$ for all $x \in X$.

(4) We should mention that the condition $T \equiv 1$ can be modified or relaxed to some extent in these theorems. Clearly it is enough if, for each n, either $T_n \equiv 0$ or $T_n \equiv 1$; the proofs also remain valid if we use only those measurable s.f. a such that t_a is finite and $[t_a = n]$ depends only on $f_{i_1}, \ldots, f_{i_k}$ for each n.

We now consider briefly a setup which is more general than that discussed in theorems 21 and 22. Suppose $\underline{G} = (F, \Gamma, T, u)$ is a standard $\underline{\underline{F}}$ measurable g.s.; $(F, \underline{\underline{F}})$ is a standard analytic space, $T \equiv 1$; $\Gamma(x)$ is a closed convex subset of $P(\underline{\underline{F}})$ for each x; for each $n \geq 1$, $\underline{\underline{F}}_n$ is a countably generated sub-σalgebra of $\underline{\underline{F}}^n$ such that Γ_n and u_n are $\underline{\underline{F}}_n$ measurable - i.e. u_n is $\underline{\underline{F}}_n, \bar{\underline{\underline{R}}}$ measurable and the set $\{(x, \gamma): x \in F^n, \gamma \in \Gamma(x)\} \in \underline{\underline{F}}_n \times \underline{\underline{P}}(\underline{\underline{F}})$ and contains the graph of a $\underline{\underline{F}}_n, \underline{\underline{P}}(\underline{\underline{F}})$ measurable function. Further we assume $\underline{\underline{F}}_{n+1} \subset \underline{\underline{F}}_n \times \underline{\underline{F}}$ for all n and that $u(h) = \sum_{n=1}^{\infty} u(h/n)$.

Let $\sigma \in \Sigma$ and $m \geq 0$. Let $\underline{\underline{F}}_{n,m}$ be the Borel structure for F^n generated by $\underline{\underline{F}}^m \times F^{n-m} \cup \underline{\underline{F}}_n$, for all $n \geq m$. Suppose for each $n > m$ we can find a conditional joint distribution under p_σ of $f_1, \ldots, f_n$ given $\underline{\underline{F}}_{n,m}$ which, moreover, is everywhere proper (in the sense of [6]). Then we can carry out obvious modifications in the proofs of theorems 21 and 22 to get a strategy $\sigma^* \in \Sigma$ satisfying (i) $I(\sigma^*, T)(y) = I(\sigma, T)(y)$ for all $y \in F^m$ and (ii) σ^*_n is $\underline{\underline{F}}_{n,m}$, $\underline{\underline{P}}(\underline{\underline{F}})$ measurable for $n \geq m$.

However an everywhere proper conditional distribution given $\underline{\underline{F}}_{n,m}$ need not exist. In fact, suppose $(X, \underline{\underline{A}})$ is a standard Borel space and $\underline{\underline{B}}$ is a countably generated sub σ-algebra of $\underline{\underline{A}}$, then Blackwell and Ryll-Nardzewski [6] have shown that a $q \in Q(\underline{\underline{B}}|\underline{\underline{A}})$ satisfying $q(x, B) = I_B(x)$ for all $x \in X$, $B \in \underline{\underline{B}}$ exists iff there is a $\underline{\underline{B}}$, $\underline{\underline{A}}$ measurable function g on X into X such that $g(x)$ and x belong to the same $\underline{\underline{B}}$-atom for all x. Thus given $p \in P(\underline{\underline{A}})$, in order that we can find a $q \in Q(\underline{\underline{B}}|\underline{\underline{A}})$ satisfying $q(x, B) = I_B(x)$ and $\int_B q(x, A)p(dx) = p(A \cap B)$ for all $x \in X$, $B \in \underline{\underline{B}}$, $A \in \underline{\underline{A}}$, it is necessary that a measurable selection choosing one point from each $\underline{\underline{B}}$-atom exists. We do not know if this condition is also sufficient.

We note that the assumption regarding the conditional distribution given $\underline{\underline{F}}_{n,m}$ being everywhere proper is needed only to ensu that $\sigma^*(x) \in \Gamma(x)$ for all x. Consequently the assumption is unnecessary if Γ is a constant on each F^n. Moreover, as we know that almost everywhere proper distributions always exist, σ^* can be shown to be 'essentially available in Γ' (see [26]) - i.e. $(\sigma^*, T) \in A_m(G, x)$ for all $x \in F^m$ and satisfies (i) and (ii).

We now turn our attention to a different problem. Blackwell has shown (Theorem 1 of [3]) that in a discounted d.p. problem, for any $p \in P(S)$ and $\epsilon > 0$ there is a (p, ϵ) optimal plan. We prove below the corresponding result for a measurable g.s.

<u>Theorem 23:</u> Let $\underline{G} = (F, \Gamma, T, u)$ be a standard $\underline{\underline{F}}$ measurable g.s. and $D = \Sigma \times \Lambda$. Suppose $n \geq 1$, $p \in P(\underline{\underline{F}}^n)$, $\epsilon > 0$ and $K > 0$ are given. Then there is a $\pi^* \in D$ such that $p(\{x \in F^n : I(\pi^*)(x) \geq I(\pi)(x) - \epsilon$ or $I(\pi^*)(x) \geq K\}) = 1$ for any $\pi \in D$.

<u>Proof:</u> Let $v(x)$ be the essential supremum of the family $\{I(\pi)(x) : \pi \in D\}$ relative to p. Hence there is a sequence $\pi^1, \pi^2, \ldots$ in D such that $v(x) = \sup_m I(\pi^m)(x)$ a.s.(p). Let $w = \sup_m I(\pi^m)$ and X_m be the set all x in F^n such that m is the smallest integer k satisfying $I(\pi^k)(x) \geq w(x) - \epsilon$ or $I(\pi^k)(x) \geq K$ according as $w(x) < \infty$ or $w(x) = \infty$. Then X_m's form a measurable partition of F^n. Define π^* such

that $\pi^* = \pi^1$ on F^k for $k < n$ and $\pi^*(xy) = \pi^m(xy)$ if $x \in X_m$, $y \in W$. Then $\pi^* \in D$ and for a.a. x, $I(\pi^*)(x) \geq v(x) - \epsilon$ or $I(\pi^*)(x) \geq K$ according as $v(x) < \infty$ or $v(x) = \infty$.

Now for any $\pi \in D$, $I(\pi) \leq v$ a.s. (p) by the property of essential supremum. Consequently it is easy to check that $p(\{I(\pi^*) \geq I(\pi) - \epsilon \text{ or } I(\pi^*) \geq K\}) = 1$.

<u>Remark</u>: If U_D is bounded above on F^n, as in the case of discounted d.p., then choosing $K > \sup_{x \in F^n} U_D(x)$ we get a (p, ϵ) optimal policy - i.e. a $\pi^* \in D$ such that $p(\{I(\pi^*) \geq I(\pi) - \epsilon\}) = 1$ for any $\pi \in D$. However it is easy to give examples in which $U_D \equiv \infty$ but $I(\pi)$ is bounded for all π in D, so that there need not be a (p, ϵ) optimal policy in general.

A straightforward modification of the proof of theorem 23 yields the following result.

<u>Theorem 24</u>: Let $\underline{G} = (F, \Gamma, T, u)$ be a standard $\underline{\underline{F}}$ measurable g.s. and D be a permissible class such that $D(x) = \Sigma \times \Lambda(x)$ for all $x \in W$. Suppose $\bar{J}_D(\sigma)(x)$ is $\underline{\underline{F}}^n$ measurable on F^n for all σ in Σ. Then for any given $p \in P(\underline{\underline{F}}^n)$, $\epsilon > 0$ and $K > 0$ there is a $\sigma^* \in \Sigma$ such that

$$p\left(\{x \in F^n : \bar{J}_D(\sigma^*)(x) \geq \bar{J}_D(\sigma)(x) - \epsilon \text{ or } \bar{J}_D(\sigma^*)(x) \geq K\}\right) = 1$$

for every σ in Σ.

The above theorem raises the question as to when $\bar{J}_D(\sigma)$ is measurable for all σ in Σ. If $D = \Sigma \times \Lambda$ then $\bar{J}_D(\sigma) = I(\sigma, T^*)$ as noted earlier and so $\bar{J}_D(\sigma)$ is measurable. The more interesting permissible class is $D(x) = \Sigma \times \Lambda_0(x)$ where $\Lambda_0(x)$ consists of all s.fs a in Λ such that $a[x]$ is finite, $x \in W$. For each x and history h let $N(x,h)$ be the set of all positive integers $n \geq 0$ such that $T(x,h|n) \neq 1$ and $t_{T_*[x]}(h) \leq n \leq t_{T^*[x]}(h)$. Define $g(x, h) =$

$$\inf_{n \in N(x,h)} \ \sup_{\substack{m \in N(x,h) \\ m \geq n}} u(x,h|m)$$

for $x \in W$, $h \in H$. The measurability of the map g is easy to check: The maps $(x, h) \longrightarrow t_{T_*[x]}(h)$ and $(x,h) \longrightarrow t_{T^*[x]}(h)$ are measurable. Define $v_{nm}(x,h) = u(x, h|m)$ if $m \geq n$ and $m \in N(x,h)$ and $= -\infty$ otherwise. Then v_{nm} is measurable and so $\sup_m v_{nm}$ is measurable. Let $w_n(x,h) = \sup_m v_{nm}(x,h)$ if $n \in N(x,h)$ and $= \infty$ otherwise. Then w_n is measurable and so $\inf_n w_n$ is also measurable. However $\inf_n w_n(x,h) = g(x,h)$.

Sudderth [26, Thms. 3.1, 4.2] has shown that

$\bar{J}_D(\sigma)(x) = \int g(x,h)p_{\sigma[x]}(dh)$ if u is bounded. He works with the case $T \equiv 2$ so that $g(x,h) = \overline{\mathrm{Lim}}_{m \to \infty} u(x, h|m)$ but it is easy to modify u so as to apply his theorem for a general T. (Of course we are assuming in this discussion that $T_*[x]$ is finite for all x; otherwise $\Lambda_0(x)$ is empty.) Since g is measurable relative to $\underline{\underline{W}} \times \underline{\underline{H}}$, $\bar{J}_D(\sigma)$ is $\underline{\underline{W}}$ measurable for all σ in Σ.

Suppose $D(x) = \Sigma \times \Lambda(x)$ is a permissible class such that $\Lambda(x) \subset \Lambda$ for all x. In order to decide whether $\bar{J}_D(\sigma)$ is measurable for all σ, the following lemma is useful in certain situations.

<u>Lemma 25:</u> Let $(X, \underline{\underline{A}})$ be a Borel space and $\mathfrak{A} = \{\underline{\underline{A}}_n : n \geq 0\}$ be a sequence of increasing sub σ-algebras of $\underline{\underline{A}}$. Let $\{Y_n : 0 \leq n \leq \infty\}$ be a sequence of uniformly bounded real random variables on X. Suppose $\mathfrak{J}$ is a set of stopping rules with respect to $\mathfrak{A}$ such that $\max(t, t') = t \vee t' \in \mathfrak{J}$ for all $t, t' \in \mathfrak{J}$, $M \subset P(\underline{\underline{A}})$ and C is a dense subset of in its M-topology. Then

$$\inf_{t \in \mathfrak{J}} \sup_{\substack{t' \in \mathfrak{J} \\ t' \geq t}} \int Y_{t'} \, d\mu = \operatorname{Inf}_{t \in C} \sup_{t' \in C} \int Y_{t \vee t'} \, d\mu$$

~~for~~ each $\mu \in M$.

Proof: For each $t \in \mathcal{J}$, $\sup_{\substack{t' \in \mathcal{J} \\ t' \geq t}} \int Y_{t'}\, d\mu = \sup_{t' \in \mathcal{J}} \int Y_{t \vee t'}\, d\mu$

because $\mathcal{J}$ is closed under maxima. Let $\overline{Y}_n(x) = Y_{t(x) \vee n}(x)$

for $0 \leq n \leq \infty$. As C is dense in $\mathcal{J}$ it is easy to see that

$\sup_{t' \in \mathcal{J}} \int \overline{Y}_{t'}\, d\mu = \sup_{t' \in C} \int \overline{Y}_{t'}\, d\mu$ for $\mu \in M$. Hence

$\sup_{t' \in \mathcal{J}} \int Y_{t \vee t'}\, d\mu = \sup_{t' \in C} \int Y_{t \vee t'}\, d\mu$ and t being arbitrary

this shows that $\inf_{t \in \mathcal{J}} \sup_{t' \in} \int Y_{t \vee t'}\, d\mu \leq \inf_{t \in C} \sup_{t' \in C} \int Y_{t \vee t'}\, d\mu$

for all $\mu \in M$. To show the reverse inequality let $\mu \in M$, $t \in \mathcal{J}$, $\epsilon > 0$ and $K \geq \sup_{0 \leq n \leq \infty} |Y_n|$. Choose $t_o \in C$ such that

$\mu([t \neq t_o]) < \epsilon$. Then $\mu([t \vee t' \neq t_o \vee t']) < \epsilon$ for all t'

and consequently $\int Y_{t \vee t'}\, d\mu \leq \int Y_{t_o \vee t'}\, d\mu + \epsilon K$. So

$\sup_{t' \in C} \int Y_{t_o \vee t'}\, d\mu \leq \sup_{t' \in} \int Y_{t \vee t'}\, d\mu + \epsilon K$. As t, ϵ are arbitrary the required inequality follows. This completes the proof of the lemma.

Theorem 26: Let $\underline{G} = (F, \Gamma, T, u)$ be a standard $\underline{\underline{F}}$ measurable g.s. such that u is bounded. Let $D(x) = \Sigma \times \Lambda(x)$, $\Lambda(x) \subset \Lambda$ and $\mathcal{J}(x) = \{t_{a[x]} : a \in \Lambda(x)\}$ for all x in W. Suppose there is a countable set C of measurable stopping rules such that C is a dense subset of $\mathcal{J}(x)$ in its

$\underline{\underline{P}}(\underline{\underline{H}})$-topology for all x. Then the map $x \longrightarrow \bar{J}_D(\sigma)(x)$ is $\underline{\underline{W}}, \underline{\underline{R}}$ measurable for any $\sigma \in \Sigma$.

<u>Proof</u>: - For any fixed σ in Σ, the map $x \longrightarrow p_{\sigma[x]}$ is $\underline{\underline{W}}, \underline{\underline{P}}(\underline{\underline{H}})$ measurable and for any fixed measurable stopping rule t, the function $(x, h) \longrightarrow \mathring{u}(x, h|t(h))$ is $\underline{\underline{W}} \times \underline{\underline{H}}, \underline{\underline{R}}$ measurable. So the map $x \longrightarrow \int u(x, h|t(h))\, p_{\sigma[x]}(dh)$ is $\underline{\underline{W}}, \underline{\underline{R}}$ measurable. By Lemma 25 it is clear that

$$\bar{J}_D(\sigma)(x) = \inf_{t \in C} \sup_{t' \in C} \int u(x, h|t(h) \vee t'(h))\, p_{\sigma[x]}(dh)$$

for all $x \in W$. Since C is countable this shows that $\bar{J}_D(\sigma)$ is a measurable function of x.

We shall now study conditions under which the optimal reward functions satisfy the optimality equations in a measurable gambling problem. We consider a standard $\underline{\underline{F}}$ measurable g.s.$\underline{G} = (F, \Gamma, T, u)$ and confine our attention to two permissible classes associated with it: (i) $D_0(x) = \Sigma \times \Lambda$ for all $x \in W$; (ii) $D_1(x) = \Sigma \times \Lambda(x)$ where $\Lambda(x) = \{a \in \Lambda : a[x]$ is finite$\}$ for all $x \in W$.

Note that both D_0 and D_1 admit continuation. Consequently U_{D_0} and U_{D_1} are conservative for $\underline{G}$. D_0 and D_1

also admit substitution by s.fs and if $\underline{\underline{F}} = S(F)$ then they admit extension in $\underline{G}$ as well. Thus if $\underline{\underline{F}} = S(F)$ then U_D satisfies U-optimality equation for $\underline{G}$ and $\bar{V}_D$ satisfies V-optimality equations for $\underline{G}$ where D is D_0 or D_1. This shows, for instance, that in d.p. problems in which the state and action spaces are countable and also in problems con-sidered in Ornstein [21] and Blackwell [4], the optimal reward functions satisfy the optimality equations. In the next two theorems we consider situations as in general d.p. problems or in measurable gambling problems of Strauch [24] and Sudderth [26, 27].

<u>Theorem 27:</u> Suppose $\underline{G} = (F, \Gamma, T, u)$ is a standard $\underline{\underline{F}}$ measurable g.s. such that $T \equiv 2$, u is bounded and $(F, \underline{\underline{F}})$ is a standard analytic space. Then

$$\bar{V}_{D_1}(x) = \sup_{\gamma \in \Gamma(x)} \int \bar{V}_{D_1}(xf)\, \gamma(df) \quad \text{for all } x \text{ in } W.$$

<u>Proof:</u> As already noted $\bar{V}_{D_1}(x) \leq \sup_{\gamma \in \Gamma(x)} \int \bar{V}_{D_1}(xf)\, \gamma(df)$ for all x. To prove the reverse inequality, let $x \in W$, $\gamma \in \Gamma(x)$ and $\epsilon > 0$. Since $(F, \underline{\underline{F}})$ is a standard analytic space $P_*(\underline{\underline{H}}) = P(\underline{\underline{H}})$ and so the set $M = \{(f, \mu): f \in F, \mu = p_{\sigma[xf]}$ for some $\sigma \in \Sigma\}$ is a Borel subset of $F \times P(\underline{\underline{H}})$.

By Sudderth's theorem [26] $\overline{J}_{D_1}(\sigma)(y) = \int g(y;\, h) p_{\sigma[y]}(dh)$ for all $\sigma \in \Sigma$, $y \in W$ where $g(y;\, h) = \overline{\lim}_{n \to \infty} u(y,\, h|n)$. Hence $\overline{V}_{D_1}(xf) = \sup_{\mu:(f,\mu)\in M} \int g(xf;\, h)\, \mu(dh)$ is a universally measurable function of f. Consequently we can choose a Borel set $N \subset F$ and a Borel measurable function v such that $\gamma(N) = 0$ and $v(f) = \overline{V}_{D_1}(xf)$ for all $f \notin N$. Let $L = \{(f,\mu) \in M : f \notin N$ and $\int g(xf;\, h)\mu(dh) > v(f) - \epsilon$ or $f \in N\}$. Then L is a Borel subset of $F \times P(\underline{\underline{H}})$ and every f-section of L is nonempty. Hence by a theorem of von Neumann [16, Thm. 6.3], we can find a Borel set $N' \subset F$ and a Borel measurable function φ on F into $P(\underline{\underline{H}})$ such that $\gamma(N') = 0$ and $(f,\, \varphi(f)) \in L$ for all $f \notin N'$. Since M admits a Borel selection, we can assume without loss of generality that $(f,\, \varphi(f)) \in M$ for all $f \in N'$. So for each $f \in F$ there is a $\sigma_f \in \Sigma$ such that $\varphi(f) = p_{\sigma_f[xf]}$. We can find a $\sigma \in \Sigma$ such that $\sigma(x) = \gamma$ and $\sigma[xf] = \sigma_f[xf]$ for all f. Then

$$\overline{J}_{D_1}(\sigma)(x) = \int g(x;\, h) p_{\sigma[x]}(dh) = \iint g(xf;\, h)\varphi(f)(dh)\, \gamma(df)$$
$$> \int [v(f) - \epsilon]\, \gamma(df) = \int \overline{V}_{D_1}(xf)\, \gamma(df) - \epsilon.$$

So $\overline{V}_{D_1}(x) \geq \int \overline{V}_{D_1}(xf)\, \gamma(df)$. This completes our proof.

<u>Theorem 28</u>: Suppose $\underline{G} = (F, \Gamma, T, u)$ is a standard $\underline{\underline{F}}$ measurable g.s. such that $(F, \underline{\underline{F}})$ is a standard analytic space, $T \equiv 2$ and u is bounded. Then U_{D_0} and U_{D_1} satisfy the U-optimality equation for $\underline{G}$.

<u>Proof</u>: As already shown, U_{D_0} and U_{D_1} are conservative for $\underline{G}$. Therefore it is enough to show that $U_D(x) \geq \int U_D(xf)\,\gamma(df)$ for each $\gamma \in \Gamma(x)$ and $D = D_0$ or D_1. Fix $\epsilon > 0$ and let $\mathcal{J}(xf)$ be $\{t_{a[xf]} : a \in \Lambda\}$ or $\{t_{a[xf]} : a \in \Lambda(xf)\}$ according as D is D_0 or D is D_1. Observe that there is a set C of measurable stopping rules such that C is a dense subset of $\mathcal{J}(xf)$ in its $P(\underline{\underline{H}})$-topology. So the map $w(f,\mu) = \sup_{t \in \mathcal{J}(xf)} \int u(xf,h|t(h))\mu(dh)$ is a measurable function on $F \times P(\underline{\underline{H}})$ since the supremum can be taken over the countable set C. If M denotes the Borel set $\{(f,\mu) \in F \times P(\underline{\underline{H}}) : \mu = p_{\sigma[xf]}$ for some $\sigma \in \Sigma\}$ then

$U_D(xf) = \sup_{\mu:(f,\mu) \in M} w(f,\mu)$ and hence $U_D(xf)$ is a universally measurable function of f. Choose a Borel measurable function v on F and a Borel subset N of F such that $\gamma(N) = 0$ and $v(f) = U_D(xf)$ for $f \notin N$. Let $L = \{(f,\mu) \in M: f \notin N$ and $w(f,\mu) > v(f) - \epsilon$ or $f \notin N\}$. L is a Borel subset of

$F \times P(\underline{\underline{H}})$ whose every f-section is nonempty. Hence, by von Neumann's theorem [16] we can get a Borel function φ on F into $P(\underline{\underline{H}})$ and a γ-null set N' such that $(f, \varphi(f)) \in M$ for all f and $(f, \varphi(f)) \in L$ for all $f \notin N'$. Let C be ordered as $t_1, t_2, \ldots$. Let F_n denote the set of all $f \in F$ such that n is the smallest natural number satisfying $\int u(xf, h|t_n(h)) \varphi(f)(dh) > w(f, \varphi(f)) - \epsilon$. For each n and f, choose a s.f. a_n^f such that $t_n = t_{a_n^f[xf]}$. Define a s.f a by $a(x) = 1$, $a(xfy) = a_n^f(xfy)$ if $f \in F_n$ $y \in W$ and $a = 0$ otherwise. Also choose $\sigma \in \Sigma$ such that $p_{\sigma[xf]} = \varphi(f)$ for all f. Then $(\sigma, a) \in D$ and it is easy to check that $I(\sigma, a)(x) > \int [v(f) - \epsilon] \gamma(df)$. As $\epsilon > 0$ is arbitrary this proves that $U_D(x) \geq \int U_D(xf) \gamma(df)$ and completes the proof of the theorem.

Remarks: (1) The assumptions that $T \equiv 2$ can be relaxed to some extent. We used only the fact that C is countable dense in $\mathcal{J}(xf)$ for all f. Similarly regarding the class of s.fs permissible in D, all we need is that this admit a 'measurable' continuation. The assumption of boundedness of u may also be dropped. The proof is also valid for instance, when $T \equiv 1$.

(2) In d.p. problems by using the above theorem we have the equation $U_{D_0}(s) = \sup_{a \in A} U_{D_0}(s,a)$

$$= \sup_{a \in A} \int U_{D_0}(s, a, s')\, q(s,a,ds').$$

In the total expected reward criterion, it is clear from the stationarity in the problem, that $U_{D_0}(s,a,s') = r(s,a,s') + \beta U_{D_0}(s')$. Consequently $U_{D_0}(s) = \sup_{a \in A} \int [r(s,a,s') + \beta U_{D_0}(s')]\, q(s,a,ds')$.

In the case of average expected reward, $U_{D_0}(s,a,s') = U_{D_0}(s')$ so that the optimality equation takes the form

$$U_{D_0}(s) = \sup_{a \in A} \int U_{D_0}(s')\, q(s,a,ds').$$

(3) In measurable gambling problems considered by Strauch [24] and Sudderth [26, 27], the above theorem implies that $U_{D_1}(f) \geq \sup_{\gamma \in \Gamma(f)} \int U_{D_1}(ff')\gamma(df')$. But again $U_{D_1}(ff')=U_{D_1}(f')$ because of stationarity. Hence $U_{D_1}(f) \geq \sup_{\gamma \in \Gamma(f)} \int U_{D_1}(f')\gamma(df')$ also $U_{D_1}(f) \geq u(f)$ for all f. These two facts show, by a result of Dubins and Savage [9, pp. 28] that U_{D_1} exceeds the utility U of the house. Since the reverse inequality always holds $U = U_{D_1}$, which shows that in a measurable problem it is

enough to use measurable policies. This result was proved by Strauch [24].

Optimality equations play a crucial role in the study of many other sequential problems, see e.g. [7].

We now turn our attention to the optimal reward $\bar{V}_{D_1}$ and optimal strategies. Dubins and Savage have given neat characterizations of optimal strategies by studying strategies which they call thrifty and equalizing [9, Chapter 3]. We examine, very briefly, how far their results go through in a measurable setting. We let $\underline{G} = (F, \Gamma, T, u)$ be a standard $\underline{\underline{F}}$ measurable g.s in which u is bounded and $T \equiv 2$. Let $D(x) = \{(\sigma, a): \sigma \in \Sigma, a \in \Lambda$ and $a[x]$ is finite$\}$ for all $x \in W$. Let v denote $\bar{V}_D$.

<u>Theorem 29</u>: $\bar{J}^v_D(\sigma)(x) \geq \bar{J}^u_D(\sigma)(x)$ for all $\sigma \in \Sigma$, $x \in W$. If $(F, \underline{\underline{F}})$ is a standard analytic space then v is the smallest function w on W such that (a) $w(x) \geq \int w(xf)\, \gamma(df)$ for all $\gamma \in \Gamma(x)$ and (b) $\bar{J}^w_D(\sigma)(x) \geq \bar{J}^u_D(\sigma)(x)$ for all $\sigma \in \Sigma$, $x \in W$.

<u>Proof:</u> By Sudderth's theorem [26, 4.1], $\bar{J}^u_D(\sigma)(x) =$

$\int g(x; h)\, p_{\sigma[x]}(dh)$ where $g(x; h) = \limsup_{n \to \infty} u(x, h|n)$.

Hence $\bar{J}^u_D(\sigma)(x) = \int \bar{J}^u_D(\sigma)(xf)\ \sigma(x, df)$.

From this equality it is easy to see that

$\bar{J}^u_D(\sigma)(x) = \int \bar{J}^u_D(\sigma)(x, h|t(h))\ \sigma[x](dh)$ for any finite stop rule t, using an induction argument on the Eudoxus structure of t. Consequently, for any s.f. a such that $a[x]$ is finite we have

$$\bar{J}^u_D(\sigma)(x) = \int \bar{J}^u_D(\sigma)(x, h|t_{a[x]}(h))\ \sigma[x]\ (dh)$$

$$\leq \int v(x, h|t_{a[x]}(h))\ \sigma[x](dh) = I_v(\sigma, a)(x)$$

So $\bar{J}^u_D(\sigma)(x) \leq \bar{J}^v_D(\sigma)(x)$ for each $\sigma \in \Sigma$, $x \in W$.

Suppose $(F, \underline{\underline{F}})$ is a standard analytic space. Then v satisfies (a) on account of Theorem 27 and v satisfies (b) as shown above. Let w be any function on W satisfying (a) and (b). Since w satisfies (a) a result of Dubins and Savage [9, pp. 28] shows that $w(x) \geq I_w(\sigma, a)(x)$ for every $\sigma \in \Sigma$ and s.f. a such that $a[x]$ is finite. Consequently $w(x) \geq \bar{J}^w_D(\sigma)(x) \geq \bar{J}^u_D(\sigma)(x)$ by (b). Taking supremum over all $\sigma \in \Sigma$ we have $w(x) \geq v(x)$. This proves the theorem.

Theorem 30. Let $(F, \underline{\underline{F}})$ be a standard analytic space, $\sigma \in \Sigma$ and $x \in W$. Define, for each $n \geq 1$, $w_n(y) = 1$ if $u(y) \geq v(y) - \frac{1}{n}$ and $= 0$ otherwise $(y \in W)$. Also let

$$A = \{h \in H: v(x, h|n) = \int v(x, h|n, f)\, \sigma(x, h|n)(df) \text{ for all } n \geq 0\}.$$

Then i) $\bar{J}_D^u(\sigma)(x) = \bar{J}_D^v(\sigma)(x)$ iff $\bar{J}_D^{w_n}(\sigma)(x) = 1$ for all $n \geq 1$,

and ii) $\bar{J}_D^v(\sigma)(x) = v(x)$ iff $p_{\sigma[x]}(A) = 1$.

Proof: (i) follows directly from Theorem 7.2 of Chapter 3 in Dubins and Savage [9]. We merely have to note that, in view of Sudderth's result [Theorem 2, 26] and the universal measurability of v, $\bar{J}_D^v(\sigma)(x) = v(\sigma[x])$, $\bar{J}_D^{w_n}(\sigma)(x) = w_n(\sigma[x])$ and $\bar{J}_D^u(\sigma)(x) = u(\sigma[x])$ in the notations of [9].

(ii) A is universally measurable since v is so. Moreover $v(y) \geq \int v(yf)\, \gamma(df)$ for all $\gamma \in \Gamma(y)$ and $y \in W$. Consequently the required result follows from Theorems 6.1 and 6.2 of Chapter 3 in Dubins and Savage [9].

Remarks: (1) Dubins and Savage call strategies satisfying $\bar{J}_D^u(\sigma)(x) = \bar{J}_D^v(\sigma)(x)$ equalizing and those for which

$\bar{J}_D^v(\sigma)(x) = v(x)$ thrifty. Blackwell [4] has given highly insightful interpretations of these two properties. He also observed that in discounted and negative d.p. problems, with the criterion of total expected reward, every strategy is equalizing. In the discounted this fact is clear from (i) of Theorem 30. To prove it in the negative case, one needs an extension of the theorem to unbounded u which is easy but we shall not enter into it.

(2) It is worth noting that theorems 27-30 remain valid in the situation considered in Ornstein [21] - i.e. where each γ in Γ is a discrete probability and $\underline{\underline{F}} = S(F)$.

Sudderth [27] has proved that in leavable gambling houses which are also measurable there exist (p, ϵ) optimal stationary strategies. We briefly indicate, without proof, an extension of this result: suppose $\underline{G} = (F, \Gamma, T, u)$ is a standard $\underline{\underline{F}}$ measurable g.s. such that $(F, \underline{\underline{F}})$ is a standard analytic space, $T \equiv 2$, u is bounded and $D(x) = \Sigma \times \Lambda(x)$ where $\Lambda(x) = \{a \in \Lambda : a[x] \text{ is finite}\}$. Assume, as in theorem 21, that for each $n \geq 1$, a subset $C_n = \{i_1, \ldots, i_k\}$ of $\{1, 2, \ldots, n\}$ is given such that

$u(f_1, \ldots, f_n)$ depends only on $f_{i_1}, \ldots, f_{i_k}$ and $C_{n+1} \subset C_n \cup \{n+1\}$. Suppose there is a Borel measurable $\varphi : W \to P(F)$ satisfying (i) $\varphi(x) \in \Gamma(x)$ for all x and (ii) for any $\sigma \in \Sigma$, $x \in W$ the strategy σ' defined by $\sigma'(y) = \sigma(y)$ if $y \neq x$ and $\sigma'(x) = \varphi(x)$ is such that $\bar{J}_D(\sigma)(x) = u(x)$. Under these conditions we have the following theorem. A proof can be given along the lines of theorem 3 [27].

<u>Theorem 31:</u> Given any $m \geq 1$, $p \in P(F^m)$, $\epsilon > 0$ there exists a $\sigma^* \in \Sigma$ satisfying

(a) $p(\{x \in F^m : \bar{J}_D(\sigma^*)(x) \geq U_D(x) - \epsilon\}) = 1$ and

(b) for every $n \geq 1$, $\sigma^*(f_1, \ldots, f_n)$ depends only on $f_{i_1}, \ldots, f_{i_k}$.

By setting $\varphi(x) = \partial(x)$ $m = 1$ and $C_n = \{n\}$ we get Sudderth's theorem.

We shall conclude this section and the chapter with a very brief reference to continuous gambling systems. Maitra [17] has proved the existence of stationary optimal plans in discounted d.p. problems when the action space A is compact, the reward function $r(s, a)$ is upper semicontinuous and the transition function q is continuous. He uses a selection theorem due to Dubins and Savage [9, pp. 38]. The selection theorems of Kuratowski and Ryll-Nardzewski [14] may also be used in this connection and they yield slightly stronger results. Certain extensions to more general continuous gambling problems are also possible.

REFERENCES

[1] Aumann, R. J. (1961): Borel structures for function spaces.
Illinois J. Math. 5, 614-630.

[2] Bellman, R. (1957): Dynamic programming.
Princeton Univ Press.

[3] Blackwell, D. (1965): Discounted dynamic programming.
Ann. Math. Statist. 36, 226-235.

[4] ... ... (1970): On stationary policies.
Jour. Roy. Stat. Soc. (Ser. A) 133, 33-37

[5] ... ... (1965): Positive dynamic programming.
Proc. V Berkeley Symp. 1, 415-418.

[6] Blackwell, D. and Ryll-Nardzewski, C. (1963): Nonexistence of everywhere proper conditional distributions.
Ann. Math. Statist. 34, 223-225.

[7] De Groot, M. H. (1970): Optimal statistical decisions
McGraw-Hill.

[8] Derman, C. (1966): Denumerable state Markovian decision processes.
Ann. Math. Statist. 37, 1545-1554.

[9] Dubins, L.E. and Savage, L.J. (1965): How to gamble if you must
McGraw-Hill

[10] Dynkin, E. B. (1969): A game theoretic version of an optimal stopping problem (Russian) Dokl. Akad Nauk SSSR 185, 16-19. (English Translation: Soviet Math. Dokl. 10, 270-274).

[11] Gale, D. and Stewart, F.M. (1953): Infinite games with perfect information. Contributions to the theory of games II, 245-266.

[12] Hinderer, K. (1970): Foundations of non-stationary dynamic programming with discrete time parameter. Lecture notes in operations research and mathematical systems, Springer-Verlag.

[13] Kuratowski, K. (1966): Topology, vol. I. Academic Press.

[14] Kuratowski, K. and C. Ryll-Nardzewski (1965): A general theorem on selectors. Bull. Polish Acad. Sci. XIII, 397-403.

[15] Łos, J. (1965): Horizon in dynamic programs. Proc. V Berkeley Symp. 1, 479-490.

[16] Mackey, G.W. (1957): Borel structure in groups and their duals. Trans. Amer. Math. Soc. 85, 134-165.

[17] Maitra, A. (1968): Discounted dynamic programming on compact metric spaces. Sankhyā (Ser. A) 30, 211-216.

[18] Maitra, A. and Parthasarathy, T. (1970): On stochastic games.
Jour. Optimization Theory Appl. 5, 289-300

[19] Meyer, P. A. (1966): Probability and potentials
Blaisdell Publ.

[20] Neveu, J. (1965): Mathematical foundations of the calculus of probability
Holden-Day.

[21] Ornstein, D. (1969): On the existence of stationary optimal strategies.
Proc. Amer. Math. Soc. 20, 563-569.

[22] Rao, B. V. (1971): Borel structures for function spaces.
Colloquium Math. XXIII, 33-38.

[23] Ross, S. M. (1968): Non-discounted denumerable Markovian decision models.
Ann. Math. Statist. 39, 412-423.

[24] Strauch, R. E. (1967): Measurable gambling houses.
Trans. Amer. Math. Soc. 126, 64-72.

[25] ... ... (1966): Negative dynamic programming.
Ann. Math. Statist. 37, 871-890.

[26] Sudderth, W. D. (1971): On measurable gambling problems.
Ann. Math. Statist. 42, 260-269.

[27] ... ... (1969): On the existence of good stationary strategies.
Trans. Amer. Math. Soc. 135, 399-414.

[28] White, D. J. (1969): Dynamic programming
Math. Economic Texts No.1.,
Oliver and Boyd.

CHAPTER II

0. Introduction:

In this chapter we study some gambling problems over continuous time. We do not attempt to formulate these problems in as general terms as we did in discrete time, but content ourselves with discussions of certain special classes of them. For this reason we find the terminology of dynamic programming (d.p.) more suited for our purpose than the gambling terminology.

We first consider a continuous-time d.p. problem which admits of a treatment by discrete-time methods. Suppose we have a system with a state space S. We are allowed to observe the states of the system and choose appropriate actions at discrete, randomly chosen, points of time. We start in a state $x_o \in S$; we choose an action a_o from a given set A of actions available to us and also decide an instant $t_o > 0$ at which we shall next observe the system. The system moves according to a transition mechanism $q(t, x_o, a_o, \cdot)$ and at each instant t upto time t_o we receive a reward $r(x_t, a_o)$ if x_t is the state at time t. If the state at time t_o is x_1 then we choose $a_1 \in A$ and $t_1 > 0$ depending on x_o, a_o, t_o and x_1 and observe the system next at time $t_o + t_1$. For each t, $t_o < t \leq t_o + t_1$ we get a reward $r(x_t, a_1)$ where x_t is distributed according to $q(t - t_o, x_1, a_1, \cdot)$ and so on. We wish to maximise our total expected reward over the infinite future. In order that the expectations exist we shall discount our reward at time t by the factor e^{-t}. The resulting problem is a very natural extension of the discounted d.p. problem of Blackwell [1] from the discrete time to continuous time. Our methods are also straightforward extensions of those used by Blackwell [1] and Strauch [6]. A fairly complete treatment of the finite case (i.e. S and A finite) using the average expected cost criterion, over continuous time, was carried out by Jewell [5] Howard [4] and Chitgopekar [2].

The above problem is specified by 4 objects : S, A, q and r. We assume S and A to be Borel subsets of complete separable metric spaces. Let r be a bounded Borel measurable real-valued function on $S \times A$ and q be a function on $[0, \infty) \times S \times A \times \underline{\underline{S}}$ into $[0, 1]$ satisfying :

i) $q(t, x, a, \cdot)$ is a probability measure on the class $\underline{\underline{S}}$ of all Borel subsets of S ;

ii) $q(\cdot, \cdot, \cdot, E)$ is Borel measurable;

iii) $q(t + s, x, a, E) = \int q(t, x, a, dy)\, q(s, y, a, E)$ and

iv) $q(0, x, a, E) = I_E(x)$ for all $t \geq 0$, $s \geq 0$, $x \in S$, $a \in A$, $E \in \underline{\underline{S}}$.

We shall use the notations of [1,6] without explicit mention. Let $B = (0, \infty) \times A$ and regard q as a transition function on S given $S \times B$ - i.e. $q \in Q(S|B)$. A plan π is defined to be a sequence $(\pi_0, \pi_1, \pi_2, \ldots)$ such that $\pi_n \in Q(B|X_n)$ where $X_n = SB \ldots SB\, S$ ($2n + 1$ factors), $n \geq 0$. Each π defines a $e_\pi \in Q(\Omega|S)$ where $\Omega = BSBS \ldots$, namely $e_\pi = \pi_0 q\, \pi_1 q \ldots$. Denote the coordinate functions on $S\Omega$ by $x_0, (a_0, t_0), x_1, (a_1, t_1), \ldots$. We shall call π to be complete if $\sum_{i=0}^{\infty} t_i = \infty$ a.s. with respect to $e_\pi(\cdot|x_0)$ for each $x_0 \in S$. π is said to be separated if for some $\partial > 0$ it is true that $t_i \geq \partial$ for all $i \geq 0$ a.s. under $e_\pi(\cdot|x_0)$ for each $x_0 \in S$. π is said to be semi-Markov if for each $n \geq 0$, $\pi_n(\cdot|x_0, a_0, t_0, \ldots, x_{n-1}, a_{n-1}, t_{n-1},$ depends only on $x_0, \sum_{i=0}^{n-1} t_i, x_n$. It is said to be Markov if $\pi_n(\cdot|x_0, a_0, t_0, \ldots, x_{n-1}, a_{n-1}, t_{n-1}, x_n)$ depends only on $\sum_{i=0}^{n-1} t_i, x_n$ for $n \geq 0$. A Markov π is non-random if for some $\partial > 0$ and for each $n \geq 0$, there is a

$$f_n : [0, \infty] \times S \longrightarrow A \times [a, \infty)$$ such that

$\pi_n(\cdot | x_0, \ldots, x_n) = f_n(\sum_0^{n-1} t_i, x_n)$ and it is stationary if $f_n = f$ for all $n \geq 0$.

1. Optimal rewards:

Let $\pi = (\pi_0, \pi_1, \ldots)$ be any plan. We shall now define the total expected reward from π. Define

$\rho : [0, \infty) \times S \times \Omega \longrightarrow R$ by

$\rho(t, x_0, \omega) = e^{-t} \int r(y, a_n) q(t - \sum_0^{n-1} t_i, x_n, a_n, dy)$ if $\sum_{i=0}^{n-1} t_i \leq t < \sum_{i=0}^{n} t_i$, and $= 0$ otherwise, where

$\omega = (a_0, t_0, x_1, a_1, t_1, \ldots)$ and $t_{-1} = 0$. Clearly ρ is Borel measurable and we let

$$I(\pi)(x) = \int_0^\infty \int_\Omega \rho(t, x, \omega) e_\pi(\cdot | x) dt.$$

$I(\pi)(x)$ is the total expected reward from π starting at the initial state x. Let $v_0^* = \sup I(\pi)$ where the supremum is taken over all plans π. Similarly let v_1^* and v_2^* denote the suprema of $I(\pi)$ as π ranges over all complete plans and separated plans respectively. Plainly $v_0^* \geq v_1^* \geq v_2^*$.

It is easy to see that v_0^* and v_1^* need not, in general, be the same. For example: let $S = [0, \infty)$, A be a singleton; let $q(t, x, a, \cdot) = \partial(x + t)$ for $x \geq 0$, $t \geq 0$ and $r(x, a) = 1$ if $0 \leq x \leq 1$ and $= -1$ otherwise. Then $v_0^*(0) = 1 - \frac{1}{e}$ and $v_1^*(0) = 1 - \frac{2}{e}$.

Now we shall show that $v_1^* = v_2^*$ in any problem. Let (S, A, q, r) be a given problem, π any complete plan, $x_0 \in S$ and $\epsilon > 0$. We have to find a separated plan π^* such that $I(\pi^*)(x_0) \geq I(\pi)(x_0) - \epsilon$. Let μ denote the probability $e_\pi(\cdot|x_0)$ and M the supremum norm of r. Given any $\eta > 0$ choose $N > 0$ such that $M e^{-N} < \eta$ and also choose $k \geq 0$ such that $\mu([\sum_{i=0}^{k} t_i \geq N]) \geq 1 - \eta$. Then it is easy to see that, for any plan π' which agrees with π for first $k+1$ coordinates (i.e. $\pi_i' = \pi_i$ $0 \leq i \leq k$) we have $|I(\pi)(x_0) - I(\pi')(x_0)| \leq (2M+1)\eta$. Also for any $\eta' > 0$ and $k \geq 0$ we can find a $\partial > 0$ such that

$\mu([t_i \geq \partial, 0 \leq i \leq k]) \geq 1 - \eta'$. Fix η and η' suitably small, depending on ϵ, and define π^* as follows:

$$\pi_i^*(h_i, C) = \pi_i(h_i, C \cap (A \times [\partial, \infty))) + \int_A I_C(a, \partial)\,\pi_i(h_i, da)$$

where C is a Borel subset of B, $h_i \in X_i$, $i \geq 0$ and ∂, k

are chosen corresponding to η', η. Thus the distributions of t_i are truncated below at ∂. It needs a straightforward computation to verify that $I(\pi^*)(x_o) \geq I(\pi)(x_o) - \epsilon$ for properly chosen η, η'. We leave the details to the reader.

Even though $v_1^* = v_2^*$ in any problem there may be no separated plans which do uniformly better than a given complete plan. We give below an example in which an optimal complete plan π and a $p \in P(S)$ exist such that for any $\epsilon < 1$ and separated plan π', $p([I(\pi') \geq I(\pi) - \epsilon]) < 1$.

Let $S = [0, \infty)$, A : the set of natural numbers and $q(t, x, a, \cdot) = \partial(x + t)$ for all $x \in S$, $t \geq 0$, $a \in A$. The reward function r is such that $r(x, a) = 1$ if

$1 + \frac{1}{2} + \ldots + \frac{1}{a-1} \leq x < 1 + \frac{1}{2} + \ldots + \frac{1}{a}$ and $= 0$ otherwise.

Define π as follows: $\pi_n(\cdot | x_o, a_o, t_o, \ldots, x_{n-1}, a_{n-1}, t_{n-1}, x_n)$ $= \partial((a_n, t_n))$ where $1 + \frac{1}{2} + \ldots + \frac{1}{a_n - 1} \leq x_o + \sum_0^{n-1} t_i < 1 + \frac{1}{2} + \ldots + \frac{1}{a_n}$ and $t_n = 1 + \frac{1}{2} + \ldots + \frac{1}{a_n} - x_o - \sum_0^{n-1} t_i$, $(n \geq 0)$. Then π is a complete plan and $I(\pi) \equiv 1$ so that π is optimal. Let p be any probability measure on the Borel subsets of S which is absolutely continuous with respect to the Lebesgue measure λ. Let π' be any separated plan and $\epsilon < 1$. Let ∂^* be the smallest strictly positive number such that

$$e_\pi[(\{t_i \geq \partial^* \text{ for all } i \geq 0\}) \mid x_o] = 1 \text{ for all } x_o \in S.$$

Choose $a^* \in A$ such that $\frac{1}{a^*} < \frac{1}{2}\partial^*$. Then for any $a > a^*$ and $x \in [1 + \frac{1}{2} + \ldots + \frac{1}{a-1}, 1 + \frac{1}{2} + \ldots + \frac{1}{a})$ we have

$$I(\pi')(x) \leq \int_0^{\partial^* - \frac{1}{a}} e^{-t}\, dt + \int_{\partial^*}^{\infty} e^{-t}\, dt = 1 - e^{-\partial^*}(e^{-\frac{1}{a}} - 1).$$

Consequently if $a < \frac{1}{\text{Log}\,(1 + \epsilon\, e^{\partial^*})}$, $I(\pi')(x) < 1 - \epsilon$ and so $\lambda\,[\,I(\pi')(x) \geq I(\pi)(x) - \epsilon] \neq 0$.

After these preliminary observations we turn our attention to the measurability properties of these 3 reward functions. It is easy to see that these functions need not be Borel measurable. For example: let $S = I \cup I \times I$, $A = I$ where I denotes the interval $[0, 1]$; let $q(t, x, a, \cdot) = \partial((x, a))$ if $x \in I$ and $q(t, (x, y), a, \cdot) = \partial((x, y))$ if $(x, y) \in I \times I$ for $t \geq 0$, $a \in A$. Fix a Borel subset C of $I \times I$ whose projection D to the first coordinate space is non Borel and define the reward function r by $r(x, a) = 0$ if $x \in I$ and $r((x, y), a) = I_C(x, y))$ if $(x, y) \in I \times I$ for $a \in A$. Plainly, $v_o^* = v_1^* = v_2^*$ and v_o^* on I is the indicator function I_D of D. Consequently the optimal reward is not Borel measurable.

However the reward functions are universally measurable. To show this, it is enough for us to note the following points: (i) The set $\Sigma_o = \{(x, \mu): \mu = e_\pi(\cdot|x)$ for some plan $\pi\}$ is a Borel subset of $S \times P(\Omega)$ where $P(\Omega)$ is the set of all probability measures on the Borel sets of Ω. Strauch's theorem [6, Lemma 7.2] can be used directly to prove this fact, by working with the action space B.

(ii) The set $E = \{\mu \in P(\Omega): \mu([\sum_{i=o}^{\infty} t_i = \infty]) = 1\}$ is a Borel subset of $P(\Omega)$ and for each $m \geq 1$ the set $E_m = \{\mu \in P(\Omega): \mu([t_i \geq \frac{1}{m}$ for all $i \geq 0]) = 1\}$ is also Borel.

(iii) Define $u : [0, \infty) \times S \times A \times [0, \infty) \longrightarrow R$ as follows:

$$u(s, x, a, t) = e^{-s} \int_o^t e^{-v} \int r(y, a)q(v, x, a, dy)\, dv.$$ Then

u is Borel measurable and so the function $g: S \times \Omega \longrightarrow R$ defined by $g(x, \omega) = \sum_{n=o}^{\infty} u(\sum_{i=o}^{n-1} t_i, x_n, a_n, t_n)$ where $x_o = x$ and $\omega = (a_o, t_o, x_1, a_1, t_1, \ldots)$ is also Borel measurable. Observe that $I(\pi)(x) = \int g(x, \omega)\, e_\pi(d\omega|x)$.

(iv) Let $\Sigma_1 = \Sigma_o \cap \{(x, \mu): x \in S, \mu \in E\}$ and $\Sigma_2 = \Sigma_o \cap \{(x, \mu): x \in S, \mu \in E_m$ for some $m \geq 1\}$. It is easy to see that $v_i^*(x) = \sup_{\mu:(x,\mu) \in \Sigma_i} \int g(x, \omega)\mu(d\omega)$ $i = 0, 1, 2,$ Consequently v_i^*'s are universally measurable.

Imitating the proof of theorem 8.1 in Strauch [6] we can also show the existence of (p, ϵ) optimal plans for v_o^* and v_1^* criteria, i.e. given $p \in P(S)$ and $\epsilon > 0$ there exists a plan π and a complete plan π' such that

$$p([I(\pi) \geq v_o^* - \epsilon]) = 1 \text{ and } p([I(\pi') \geq v_1^* - \epsilon]) = 1.$$

However, as the example on page 97 shows, a (p, ϵ) optimal separated plan need not exist.

2. <u>Stationary separated plans:</u>

As in the discrete-time situation, we shall show that semi-Markov plansare enough in continuous-time problems also. Our proof is an obvious modification of Strauch's proof of the corresponding result in discrete time.

<u>Theorem 1:</u> Given a plan π there exists a plan π^* which is semi-Markov such that $I(\pi^*) = I(\pi)$. If π is complete, π^* can be chosen to be complete and if π is separated,

π^* can be chosen to be separated.

<u>Proof:</u> Let π^*_n be the conditional distribution under e_π of (a_n, t_n) given $x_o, \sum_{i=o}^{n-1} t_i, x_n$, for each $n \geq 0$. We shall show that for any Borel measurable function w defined on $S \times [0, \infty) \times S \times A \times [0, \infty) \times S$ and any $n \geq 0$,

$$e_\pi w(x_o, \sum_{i=o}^{n-1} t_i, x_n, a_n, t_n, x_{n+1}) =$$

$$e_{\pi^*} w(x_o, \sum_{i=o}^{n-1} t_i, x_n, a_n, t_n, x_{n+1}).$$

The assertion is clearly true for $n = 0$ since $\pi^*_o = \pi_o$. Assume that we have proved the assertion for all $m < n$. Using the letter E to denote conditional expectation under e_π

$$e_\pi w(x_o, \sum_0^{n-1} t_i, x_n, a_n, t_n, x_{n+1}) =$$

$$= E[w(x_o, \sum_0^{n-1} t_i, x_n, a_n, t_n, x_{n+1}) \mid x_o]$$

$$= E\left\{E[w(x_o, \sum_0^{n-1} t_i, x_n, a_n, t_n, x_{n+1}) \mid x_o, \sum_0^{n-1} t_i, x_n] \mid x_o\right\}$$

$$= E \; w'(x_o, \sum_0^{n-1} t_i, x_n) \mid x_o \quad \text{where}$$

$$w'(x_o, \sum_0^{n-1} t_i, x_n) = E[w(x_o, \sum_0^{n-1} t_i, x_n, a_n, t_n, x_{n+1}) \mid x_o, \sum_0^{n-1} t_i, x_n]$$

$$= \pi^*_n \, q \, w$$ by property of conditional distribution. Also $w'(x_o, \sum_0^{n-1} t_i, x_n)$ can be considered as

a function of $(x_0, \sum_0^{n-2} t_i, x_{n-1}, a_{n-1}, t_{n-1}, x_n)$ and then using the induction hypothesis $E\{w'(x_0, \sum_0^{n-1} t_i, x_n) | x_0\} =$

$e_{\pi^*} w'(x_0, \sum_0^{n-1} t_i, x_n) = e_{\pi^*} \pi_n^* qw$

$= e_{\pi^*} w(x_0, \sum_0^{n-1} t_i, x_n, a_n, t_n, x_{n+1})$. Thus the assertion has been proved for all $n \geq 0$.

Now, let $u(s, x, a, t) = \bar{e}^{s} \int_0^t \bar{e}^{v} \int r(y,a) q(v,x,a,dy) dv.$

Then $I(\pi)(x) = \sum_{n=0}^{\infty} \int u(\sum_0^{n-1} t_i, x_n, a_n, t_n) e_\pi(\cdot | x)$

$$= \sum_{n=0}^{\infty} \int u(\sum_0^{n-1} t_i, x_n, a_n, t_n)(e_{\pi^*}(\cdot | x) = I(\pi^*)(x).$$

From the definition π^* is semi-Markov. Note that the distribution of $\sum_{i=0}^{\infty} t_i$ is same under both e_π and e_{π^*}. So if π is complete, π^* is also complete. It is clear that if $t_n \geq \partial$ a.s. under e_π, the same is true under e_{π^*}. Consequently π^* is separated if π is. This completes the proof.

<u>Theorem 1:</u> Given a separated plan π, $p \in P(S)$ and $\epsilon > 0$ there exists a separated Markov plan π^* such that $p([I(\pi^*) \geq I(\pi) - \epsilon]) = 1.$

Proof: Let $\partial > 0$ be such that $t_n \geq \partial$ a.s. under $e_\pi(\cdot|x)$ for all $n \geq 0$ and $x \in S$. If π' is a plan which agrees with π for k coordinates i.e. $\pi'_i = \pi_i$ $0 \leq i \leq k$, then $|I(\pi)(x) - I(\pi')(x)| \leq e^{-k\partial}$. M where M is the supremum of $|r|$. Consequently we can assume that π is already Markov from some point on, say $n > N$. So let $\pi = (\pi_0, \ldots, \pi_N, f_{N+1}, f_{N+2}, \ldots)$ where $f_j : [0,\infty) \times S \longrightarrow [0, \infty) \times \Lambda$ are Borel functions. We shall now show that for any $\eta > 0$ there is a Borel measurable function $f_N : [0, \infty) \times S \longrightarrow [0, \infty) \times \Lambda$ such that $p([I(\pi') \geq I(\pi) - \eta]) = 1$ where $\pi' = (\pi_0, \ldots, \pi_{N-1}, f_N, f_{N+1}, \ldots)$. Using this fact N times with $\eta = \frac{\epsilon}{N}$ produces the necessary Markov π^*.

We write $I(\pi) = \pi_0 q\, \pi_1 q \ldots \pi_{N-1} q\, (u_1 + e^{-\sum_0^{N-1} t_i} \pi_N q\, u_2)$

where $u_1(x_0, a_0, t_0, \ldots, x_{N-1}, a_{N-1}, t_{N-1}, x_N) =$

$$= \sum_{n=0}^{N-1} u\Big(\sum_{i=0}^{n-1} t_i, x_n, a_n, t_n\Big),$$

$$u_2\Big(\sum_0^{N-1} t_i, x_N, a_N, t_N, x_{N+1}\Big) = \int_0^{t_N} e^{-s} \int r(y, a_N) q(s, x_N, a_N, dy)\, ds +$$

$+ e^{-t_N} I(\bar{\pi})(x_{N+1})$ and $\bar{\pi} = (f_{N+1}, f_{N+2}, \ldots)$ after time $\sum_0^{N-1} t_i$. It is enough to find f_N such that

$$(*) \ldots \quad p\left\{\pi_0 q \ldots \pi_{N-1}\, q\, f_N\, w \geq \pi_0 q \ldots \pi_N q\, \pi_N\, w - \eta\right\} = 1$$

where $w = e^{-\sum_0^{N-1} t_i} \int u_2(x_N, a_N, t_N, x_{N+1}) q(t_N, x_N, a_N, dx_{N+1})$.

Consider the probability $m = p\pi_0\, q \ldots \pi_N$ on $SB \ldots SB$ (2N+2 factors). For any f_N,

$z_1 = \pi_0 q \ldots \pi_N q\, f_N\, w\, (x_0)$ is a version of the conditional expectation

$$E(w(\sum_0^{N-1} t_i, x_N, f_N(\sum_0^{N-1} t_i, x_N))|x_0)$$

and $z_2 = \pi_0 q \ldots \pi_{N-1}\, q\, \pi_N\, w\, (x_0)$ is a version of

$E(w(\sum_0^{N-1} t_i, x_N, a_N, t_N)|x_0)$. If we choose f_N so that

$w(\sum_0^{N-1} t_i, x_N, f_N(\sum_0^{N-1} t_i, x_N)) \geq w(\sum_0^{N-1} t_i, x_N, a_N, t_N) - \eta$

with probability 1 then $z_1 \geq z_2$ with probability 1 and consequently f_N satisfies the required condition (*). To show that such an f_N exists we use the Lemma of Blackwell [1, sec.2] by setting $X = [0, \infty) \times S$, $Y = A \times [\partial, \infty)$, $u = w$, $\epsilon = \eta$ and q = the conditional joint distribution of (a_N, t_N) given $(\sum_0^{N-1} t_i, x_N)$ under m. The degenerate f_N that we obtain from the lemma also satisfies $f_N \geq \partial$. This proves our result.

In the case of complete plans, we can get slightly weaker results using similar methods but we shall not enter into them here. The example in Blackwell [1, Example 3] of a plan which can not be ϵ-dominated by a Markov plan serves also as an example in continuous-time to show that, in general, we can not restrict ourselves to policies π where π_n depends only on $(\sum_0^{n-1} t_i, x_n)$.

Let f be any Borel measurable function on $[0, \infty) \times S$ into $A \times [\partial, \infty)$ where $\partial > 0$. Let $\underline{\underline{M}}$ denote the space of all bounded measurable functions on $[0, \infty) \times S$. We shall define an operator $L(f)$ on $\underline{\underline{M}}$ into $\underline{\underline{M}}$.

$$(L(f)v)(s,x) = \bar{e}^{s} \int_0^{g(s,x)} \bar{e}^{t} \int r(y,h(s,x))q(t,x,h(s,x), dy)dt + \bar{e}^{s-g(s,x)} \int v(s+g(s,x), y)q(g(s,x), x, h(s,x), dy)$$

for any $v \in \underline{\underline{M}}$ where $f = (h,g)$, $h: [0, \infty) \times S \to A$ and $g : [0, \infty) \times S \to [\partial, \infty)$. The following properties of $L(f)$ are immediate: (i) $L(f)v_1 \geq L(f)v_2$ if $v_1 \geq v_2$;

(ii) $L(f)(v+c) \leq L(f)v + \bar{e}^{\partial}.c$ for any constant $c \geq 0$;

(iii) for any Markov plan $\pi = (f_o, f_1, f_2, \dots)$,

$(L(f)v)(o, x) = I(f, \pi)(x)$ where $v(t,y) = I(\pi)(y)$, $t \geq 0$ and (f, π) is the plan $(f, f_o, f_1, f_2, \dots)$.

For any separated Markov (non-random) plan $\pi = (f_0, f_1, f_2, \ldots\ )$ we shall say that a Borel function f on $[0, \infty) \times S$ into $A \times [0, \infty)$ is π-generated if the set $[0, \infty) \times S$ can be partitioned into Borel sets $E_0, E_1, E_2 \ldots$ such that $f = f_n$ on E_n $(n \geq 0)$. A separated Markov $\pi' = (f_0', f_1', \ldots\ )$ is said to be π-generated if each f_n' is π-generated. With each separated Markov π associate the operator $T(\pi) : \underline{\underline{M}} \longrightarrow \underline{\underline{M}}$ defined by $T(\pi)v = \sup_{n \geq 0} L(f_n)v$. Important properties of this operator are contained in the following

<u>Theorem 2:</u> (a) $T(\pi)\, v_1 \geq T(\pi)\, v_2$ if $v_1, v_2 \in \underline{\underline{M}}$, $v_1 \geq v_2$

(b) There is a $\partial > 0$ such that $T(\pi)(v + c) \leq T(\pi)v + \bar{e}^{\partial}.c$ for any $v \in \underline{\underline{M}}$ and constant $c \geq 0$.

(c) For any π-generated f, $L(f)v \leq T(\pi)v$ for all $v \in \underline{\underline{M}}$.

(d) For any $v \in \underline{\underline{M}}$ and $\epsilon > 0$ there is a π-generated f such that $L(f)v \geq T(\pi)v - \epsilon$.

(e) There is a $u^* \in \underline{\underline{M}}$ such that $T(\pi)u^* = u^*$.

(f) For any π-generated Markov plan π', $I(\pi')(x) \leq u^*(0,x)$ and for every $\epsilon > 0$ there is a π-generated f such that $I(f^{\infty}) \geq I(\pi) - \epsilon$.

Proof: (a) is immediate.

(b) Since π is separated there is a $\partial > 0$ such that $g_n \geq \partial$ for each $n \geq 0$ where $\pi = (f_0, f_1, \ldots)$ and $f_n = (h_n, g_n)$. Then it is easy to see that $T(\pi)(v + c) \leq T(\pi)v + \bar{e}^{\partial} . c$ for $c \geq 0$.

(c) Let $f = f_n$ on E_n where E_n's form a measurable partition of $[0, \infty) \times S$. Hence $(L(f)v)(s, x) = (L(f_n)v)(s,x)$ if $(s,x) \in E_n$. So $(L(f)v)(s,x) \leq (T(\pi)v)(s,x)$ for all (s,x).

(d) Let E_n denote the set of all (s,x) in $[0,\infty) \times S$ such that n is the smallest integer $m \geq 0$ such that $(L(f_m)v)(s,x) \geq (T(\pi)v)(s,x) - \epsilon$. Then E_n's form a measurable partition of $[0, \infty) \times S$. Define $f = f_n$ on E_n. Then f is π-generated with $(L(f)v)(s,x) = (L(f_n)v)(s,x)$ on E_n and so $L(f)v \geq T(\pi)v - \epsilon$.

(e) $\underline{\underline{M}}$ is a Banach space with the supremum norm

$$\| v \| = \sup_{s, x} |v(s,x)|.$$

For any $v_1, v_2 \in \underline{\underline{M}}$ $v_1 \leq v_2 + \| v_1 - v_2 \|$ and so by (a)

$$T(\pi)v_1 \leq T(\pi)(v_2 + \| v_1 - v_2 \|) \leq T(\pi)v_2 + \bar{e}^{\partial} . \| v_1 - v_2 \|$$

using (b) for the second inequality. Interchanging the roles of v_1, v_2, we see that $\| T(\pi)v_1 - T(\pi)v_2 \| \leq \bar{e}^{\partial} \| v_1 - v_2 \|$ and as $\partial > 0$, this shows that $T(\pi)$ is a contraction mapping on $\underline{M}$.

Hence by Banach's fixed point theorem, there is a unique $u^* \in \underline{\underline{M}}$ such that $T(\pi)u^* = u^*$ and, in fact, for any $v \in \underline{\underline{M}}$, $T^n(\pi)v \longrightarrow u^*$.

(f) Let $\pi' = (f'_0, f'_1, \ldots)$ and for each $n \geq 0$ let $u_n(s,x)$ denote the expected income from the policy $(f'_{n+1}, f'_{n+2}, \ldots)$ starting from x at time s. Then

$I(\pi')(x) = (L(f'_0) \ldots L(f'_n)u_n)(0,x)$. Since π' is π-generated each $L(f'_i)$ is a contraction with modulus $e^{-\partial}$. So

$\| L(f'_0) \ldots L(f'_n)u_n - L(f'_0) \ldots L(f'_n)u^* \| \leq e^{-n\partial} \| u_n - u^* \|$

$\leq e^{-n\partial}.(\| r \| + \| u^* \|)$. Thus $(L(f'_0)\ldots L(f'_n)u^*)(0,x) \to I(\pi')(x)$ as $n \to \infty$. Since each f'_i is π-generated, $L(f'_0) \ldots L(f'_n)u^* \leq T^n(\pi)\ u^*$ and u^* being the fixed point of $T(\pi)$, $T(\pi)u^* = u^*$. So $I(\pi')(x) \leq u^*(0,x)$.

Given $\epsilon > 0$ choose a π-generated f such that $L(f)u^* \geq T(\pi)u^* - \epsilon = u^* - \epsilon$. Hence $L^n(f)u^* \geq u^* - \epsilon$ for each $n \geq 1$. As $L(f) = T(f^\infty)$ and $T^n(f^\infty)u^*$ converges to the unique fixed point w^* of $T(f^\infty)$, we have $w^* \geq u^* - \epsilon$. But $w^*(0,x) = I(f^\infty)(x)$ and $u^*(0,x) \geq I(\pi)(x)$. So $I(f^\infty) \geq I(\pi) - \epsilon$. This completes the proof of the theorem.

As we mentioned at the very beginning our treatment imitates that of Blackwell [1] verbatim But there are several interesting side-problems which pertain to the fact that the states vary continuously over time and consequently do not arise in discrete-time. We do not enter into them here. We are not also discussing the important special cases of countable or finite state and action spaces. As we remarked earlier, when the state and action spaces are finite and the criterion is average expected cost, the problems have been analysed quite extensively by Chitgopekar [2] and others.

3. Continuously Varying Actions:

We now consider problems in which both states and actions vary continuously over time. A plan in such a case specifies an action at each instant of time depending on all previous states and actions. Since the existence of a probability measure e_π corresponding to a plan π is then difficult to establish even for relatively simple classes of plans, we shall make several assumptions on the problem. The resulting problem becomes essentially a stochastic control problem. However our assumptions are weaker than those usually made in control theory and our approach is also different.

We assume that the state space S is the m-dimensional Euclidean space R^m and the action space A is a Borel subset of a complete separable metric space. Corresponding to the transition function q in discrete-time, we assume that certain diffusion coefficients $a_{ij}(x)$, $1 \leq i, j \leq m$ and $b_j(x,u)$, $1 \leq j \leq m$ are given for $x \in S$, $u \in A$ such that the following conditions are satisfied:

i) The maps $a_{ij}: S \to R$ are bounded continuous and the maps $b_j : S \times A \to R$ are bounded **Borel measurable.**

ii) The matrix $a(x) = ((a_{ij}(x)))$ is symmetric and there is a $K < \infty$ such that $0 \leq \langle \theta, a(x)\theta \rangle \leq K|\theta|^2$ for all $\theta \in R^m$ $x \in S$ where $\langle \cdot, \cdot \rangle$ denotes the inner product and $|\cdot|$ the norm in R^m. We are also given a bounded Borel measurable reward function r on $S \times A$. Thus the problem is specified by m, A, a, b, r.

We first consider Markov plans. A Markov plan is a Borel measurable function u on $[0, \infty) \times S$ into A. For each such u define $b_u : [0, \infty) \times S \to R$ by $b_u(t,x) = b(x, u(t, x))$. We wish to find a probability measure P_x for each $x \in S$ such that, roughly speaking, the evolution of the states $x(t)$, $t \geq 0$ under P_x starts from $x(0) = x$ and at each instant t has local covariance matrix $\frac{1}{2}\, a(x(t))$ and local drift vector

$b_u(t, x(t))$. We take the underlying measurable space on which P_x should be defined to be (Ω, Σ): Ω is the space of all continuous functions on $[0, \infty)$ into S vanishing outside bounded intervals and Σ is the class of Borel subsets of Ω when Ω is equipped with the topology of uniform convergence on compacta. Let $x(t,\omega) = \omega(t)$, $t \geq 0$, $\omega \in \Omega$ denote the coordinate variables. (Ω, Σ) is a standard Borel space and Σ is the smallest Borel structure for Ω making $x(t)$, $t \geq 0$ measurable. The condition on P_x that we stated above can now be formulated as a stochastic differential equation of the form:

$dx = b_u(t, x(t))dt + \sigma(x(t))d\beta$ where σ is the positive square-root of the matrix a and β is a Brownian motion on m-dimensions. We find it more convenient to state it differently following [7].

<u>Definition:</u> Let u be a Markov plan and $x \in S$. We shall say that a probability measure P on (Ω, Σ) corresponds to (x,u) if, for each $\theta \in R^m$, the family $\{Y_\theta(t) : t \geq 0\}$ is a P-martingale where $Y_\theta(t)$ is defined as follows:

$$Y_\theta(t) = \exp[< \theta, x(t) - x(0) > - < \theta, \int_0^t b_u(s, x(s))ds > - \frac{1}{2}\int_0^t < \theta, a(x(s))\theta > ds]$$

and further, $P[x(0) = x] = 1$.

The martingale condition above is equivalent to the stochastic differential equation of the last paragraph in case a is uniformly elliptic, i.e. if there exists $K' > 0$ such that $\langle \theta, a(x)\theta \rangle \geq K'|\theta|^2$ for all $x \in S$, $\theta \in R^m$. This is a result of Stroock and Varadhan [7, Corollary 3.2].

If there is a unique P corresponding to (x,u) we may define the expected reward, by using u and starting from x, as the integral $\int_0^\infty e^{-t} \int r(x(t), u(t, x(t)))dP\, dt$. We must therefore check that there exists one and only one probability measure corresponding to any $x \in S$ and any Markov plan u. This fact follows directly from the main result of Stroock and Varadhan [7, Theorem 6.2]. The unique probability P corresponding to (x,u) is denoted by P_x^u and let

$$I(u)(x) = \int_0^\infty e^{-t} \int r(x(t), u(t, x(t)))\, dP_x^u\, dt.$$

We shall show that for any fixed u, $I(u)(x)$ is a measurable function of x. For this, it is enough to prove that P_x^u is a measurable map of x. It suffices to show that $P_x^u [x(t_1) \in B_1, \ldots, x(t_n) \in B_n]$ is a Borel measurable function of x for every choice of Borel subsets $B_1, \ldots, B_n$ of S and $0 \leq t_1 < \ldots < t_n$. Since $\{x(t): t \geq 0\}$ is a Markov process under P_x^u [7] we shall be done if we prove that

$P(s,y,t,B)$ is a measurable function of y for each $0 \leq s < t$ a and Borel $B \subset S$, where $P(s,y,t,B)$ denotes the transition probability: $P_x^u [x(t) \in B | x(s)](y)$. In fact Stroock and Varadhan [8, Theorem 7.1] have shown that $P(s,y,t,B)$ is a continuous function of (s,y). Thus $I(u)(x)$ is a Borel measurable function of x. Let $v^*(x) = \sup I(u)(x)$ where supremum is taken over all Markov plans u.

Let u be any fixed Markov plan. Partition $[0, \infty) \times S$ into cubes of the type $C_{k,n_1,\ldots,n_m} = [k, k+1) \times [n_1, n_1+1) \times \ldots \times [n_m, n_m+1)$ where $n_1, \ldots, n_m, k$ are integers and $k \geq 0$. On each of these cubes we can find a $v_{k,n_1,\ldots,n_m}$ which is a Borel function into A of Borel class 3 such that $u = v_{k,n_1,\ldots,n_m}$ a.e. relative to Lebesgue measure. This fact can be proved along the same lines as Theorem 10 of chapter I. Define v on $[0, \infty) \times S$ into A by setting $v = v_{k,n_1,\ldots,n_m}$ on $C_{k,n_1,\ldots,n_m}$. Then v is of Borel class 4 and $u = v$ a.e.

We shall first show $P_x^u = P_x^v$ for all $x \in S$. Even though this can be verified directly we find it easier to derive it from the more general results of Stroock and Varadhan [8, Corollary 9.3] by setting: $s = 0$ $b^n = b_v$ and $b = b_u$ so that $P^n_{o,x} = P_x^v$ for all n $P_{o,x} = P_x^u$ and $P^n_{o,x}$ converges to $P_{o,x}$ implies $P_x^u = P_x^v$.

Next we prove that $I(u) = I(v)$. Observe that if $r(x, u)$ did not depend on $u \in A$ then we have already established this equality since $P_x^u = P_x^v$. Since $u = v$ a.e. relative to Lebesgue measure there is a Lebesgue null set N such that $u(s,x) = v(s,x)$ if $(s,x) \notin N$. If $P(o, x, t, \cdot)$ denotes the transition probability under P_x^u, then it is shown in Section 8 of Stroock and Varadhan [8] that $P(o, x, t, \cdot)$ has a density for almost all t. I.e. there is a one-dimensional Lebesgue null set N_0 such that for all $t \notin N_0$, $P(o,x,t, B) = \int_B p(o,x,t,y)\lambda(dy)$, $B \subset S$, for a suitable Borel function $p(o,x,t,\cdot)$; λ denotes the m-dimensional Lebesgue measure.

Now $I(u)(x) = \int_0^\infty e^{-t} \int r(x(t), u(t,x(t)))\, dP_x^u\, dt$

and $\int r(x(t)), u(t,x(t)))dP_x^u = \int_S r(y, u(t,y))\, P(o,x,t,dy)$

$= \int r(y, u(t,y))\, p(o,x,t,y)\lambda(dy)$ for all $t \notin N_0$.

Hence $I(u)(x) = \int_{N_0^c} e^{-t} \int r(y, u(t,y))\, p(o,x,t,y)\lambda(dy)dt$

$$= \int_{N_0^c \times S} e^{-t}\, r(y, u(t,y))\, p(o,x,t,y)\lambda(dy)dt$$

$$= \int_{(N_0^c \times S) \cap N^c} e^{-t}\, r(y, u(t,y))\, p(o,x,t,y) \lambda(dy) dt$$

$$= \int_{(N_0^c \times S) \cap N^c} e^{-t}\, r(y, v(t,y))\, p(o,x,t,y) \lambda(dy) dt$$

$$= \int_0^\infty e^{-t} \int r(x(t), v(t,x(t)))\, dP_x^v\, dt = I(v)(x).$$

Thus we have proved the following:

Theorem 3: Given a Markov plan u there exists a Markov plan v such that (i) v is of Borel class 4 and (ii) $I(u)(x) = I(v)(x)$ for all $x \in S$.

Using the above theorem we shall show that the optimal reward over Markov plans is a universally measurable function.

Theorem 4: $v^*(x) = \sup_u I(u)(x)$ is a universally measurable function on S.

Proof: From theorem 3 it follows that $v^*(x) = \sup_{u \in C} I(u)(x)$ where C denotes the class of all Borel functions on $[0, \infty) \times S$ into A which are of Borel class 4. Choose and fix a Borel measurable function U defined on $I \times [0, \infty) \times S$ into A, where $I = [0, 1]$, such that each $u \in C$ is a section $U(z, \cdot)$ for some $z \in I$ and conversely each $U(z, \cdot) \in C$.

Then $v^*(x) = \sup_{z \in I} I(U(z, \cdot))(x)$.

Consequently it is enough for us to show that $I(U(z, \cdot))(x)$ is a measurable function of z and x.

Let P_x^z denote the measure $P_x^{U(z, \cdot)}$ on Σ. We shall first show that the map $(x, z) \longrightarrow P_x^z$ is Borel measurable. Let $t \geq 0$ and B be a set in Σ which is in the σ-algebra Σ_t generated by $\{x(s), 0 \leq s \leq t\}$. It suffices to show that $P_x^z(B)$ is measurable in x and z. For this, we make use of a result of Stroock and Varadhan [7, Theorem 6.2,] in which the density of P_x^z with respect to Q_x on Σ_t is given, where Q_x is the measure on Σ which corresponds to the diffusion coefficients $[a, 0]$. According to the expression given therein, $P_x^z(B) = \int_B f(z, \omega)\, Q_x(d\,\omega)$ where

$$f(z, \omega) = \exp\Big[\int_0^t \langle b_z(s, \omega), a^{-1}(x(s, \omega))\, dx(s)\rangle -$$

$$- \frac{1}{2}\int_0^t \langle b_z(s, \omega), a^{-1}(x(s, \omega))\, b_z(s, \omega)\rangle\, ds\Big]$$

and $b_z(s, \omega) = b(x(s, \omega), U(z, s, x(s, \omega)))$.

Now $b_z(s, \omega)$ is a measurable function of s, z and ω and so $f(z, \omega)$ is measurable in z and ω Moreover $x \longrightarrow Q_x$ is continuous and this implies that $P_x^z(B)$ is measurable in x and z.

Now $I(U(z, \cdot))(x) = \int g(z, \omega) P_x^z (d\omega)$ where $g(z, \omega) = \int_0^\infty e^{-t} r(x(t, \omega), U(z, t, x(t, \omega)) dt$. As g is a measurable function of z and ω and P_x^z is measurable in x and z we may conclude that $I(U(z, \cdot))(x)$ is a measurable function of x and z. This completes the proof of the theorem.

The following result is of interest in many cases.

<u>Theorem 5.</u> Suppose b and r are continuous functions on $S \times A$. Let $\{u_n : n \geq 0\}$ be a sequence of Markov plans such that $u_n \longrightarrow u_o$ in measure (i.e. Lebesgue measure on $[0, \infty) \times S$). Then $I(u_n)(x) \longrightarrow I(u_o)(x)$ for all x.

<u>Proof</u>: Since b is continuous, $b_{u_n} \longrightarrow b_{u_o}$ in measure as $n \to \infty$. Then, using the result of Stroock and Varadhan [8, Corollary 9.3], $P_x^{u_n} \longrightarrow P_x^{u_o}$ in the sense of variation on each Σ_t.

Let $x \in S$ and $\epsilon > 0$ be given. Choose T such that $e^{-T} \cdot \| r \| < \frac{\epsilon}{4}$. Choose $N \geq 1$ such that for any $n \geq N$ the total variation $\| \cdot \|_{\Sigma_T}$ of $(P_x^{u_n} - P_x^{u_o})$ on Σ_T is at most $\frac{\epsilon}{2 \| r \|}$. Then for any $n \geq N$,

$$|I(u_n)(x) - I(u_o)(x)| \leq \left| \int_o^T \int \bar{e}^t r_n(t, x(t)) d P_x^{u_n} dt - \right.$$

$$\left. - \int_o^T \bar{e}^t r_o(t, x(t)) dP_x^{u_o} dt \right| + \frac{\varepsilon}{2}$$

(where $r_m = r_{u_m}$)

$$\leq \left| \int_o^T \int \bar{e}^t r_n d P_x^{u_n} dt - \int_o^T \int \bar{e}^t r_n dP_x^{u_o} dt \right| +$$

$$\int_o^T \int \bar{e}^t |r_n - r_o| dP_x^{u_o} dt + \frac{\varepsilon}{2}$$

$$\leq \| r \| \cdot \| P_x^{u_n} - P_x^{u_o} \|_{\Sigma_T} . + \int_o^T \int \bar{e}^t |r_n - r_o| dP_x^{u_o} dt + \frac{\varepsilon}{2}$$

$$\leq \frac{\varepsilon}{2} + \int_o^T \int \bar{e}^t |r_n - r_o| d P_x^{u_o} dt + \frac{\varepsilon}{2}.$$

Hence it is enough to show that $\int_o^T \int \bar{e}^t |r_n - r_o| d P_x^{u_o} dt$ converges to 0 as $n \to \infty$. This is easy since $r_n(y, v) \to r_o(y, v)$ converges in measure and the transition probability $P^{u_o}(0, x, t, \cdot)$ corresponding to $P_x^{u_o}$ has a density for almost all t. This completes the proof of the theorem.

As a digression we shall prove the following result which corresponds to Lemma 7.2 of Strauch [6].

<u>Theorem 6:</u> Let $M = \{(x, \mu) : x \in S, \mu = P_x^u$ for some Markov

plan $u \}$. Then M is an analytic subset of $S \times P(\Omega)$ where $P(\Omega)$ denotes the set of all probability measures on (Ω, Σ).

Proof: By the proof of theorem 3 it is clear that $M = \{(x, \mu): x \in S, \ \mu = P_x^u$ for some $u \in C\}$, where C denotes the class Markov plans of Borel class 4. As in the proof of theorem 4 choose a Borel measurable function U on $I \times [0, \infty) \times S$ into A such that $C = \{U(z, \cdot): z \in I\}$. Then clearly $M = \text{Proj} \{(x, \mu, z): x \in S, \ \mu = P_x^{U(z,\cdot)}, \ z \in I\}$ where Proj denotes projection to the first two coordinate axes.

Now the condition $\mu = P_x^{U(z, \cdot)}$ is equivalent to saying that $\{Y_\theta^z(t), \ t \geq 0\}$ is a μ- martingale for all $\theta \in R^m$ and $\mu[x(0) = x] = 1$. Here $Y_\theta^z(t)$ refers to the defining martingale for the plan $U(z, \cdot)$. Observe that $Y_\theta^z(t)$ is continuous in t and θ. Hence it is enough to require that $\{Y_\theta^z(t) : t \in Q_o\}$ is a μ-martingale for all $\theta \in Q_1$ where Q_o and Q_1 are countable dense subsets of $[0, \infty)$ and R^m respectively.

Consequently $M = \text{Proj} \{(x, \mu, z): \{Y_\theta^z(t) : t \in Q_o\}$ is a μ-martingale for all $\theta \in Q_1$ and $\mu[x(0) = x] = 1\}$.

Now the set within brackets on the right hand side is easily seen to be a Borel subset of $S \times P(\Omega) \times I$, using the measurability of $Y_\theta^z(t)$ in θ, z and t. This shows that M is analytic and completes the proof of the theorem.

As we mentioned earlier the problem we have been considering is a problem of optimal control. It should therefore be mentioned that in control theory, the usual assumptions made on a, b, r are much more stringent than the ones imposed here. We refer the reader in this connection to the excellent review article of Fleming [3] and the papers cited there. The results we have proved here cover very little ground and much deeper studies are needed to bring the theory to the level obtaining in discrete time d.p. theory. For instance here we work only with Markov plans. More general plans, i.e. those which depend on the entire past can, of course, be defined. Existence and uniqueness of measures corresponding to such plans can be proved, using the results of Stroock and Varadhan [7, 8]. But in the absence of further conditions on a, b, r like the existence of two continuous derivatives, we are unable to show that the supremum over general plans is the same as that over Markov plans. The existence of (p, ϵ) optimal Markov plans etc., are even more remote.

REFERENCES

[1] Blackwell, D. (1965): Discounted dynamic programming. Ann. Math. Statist. 36, 226-235.

[2] Chitgopekar, S. S. (1967): Continuous time Markovian sequential control processes. Florida Statistics Univ. Report M 127.

[3] Fleming, W. H. (1969): Optimal continuous - parameter stochastic control. SIAM Review 11, 470-509.

[4] Howard, R. (1963): Semi-Markovian control systems. Tech. Report 3, Operations Research Centre, M.I.T.

[5] Jewell, W. (1963): Markov-renewal programming I., Formulation, finite return models. Op. Res. 11, 938-948.

[6] Strauch, R. E. (1966): Negative dynamic programming. Ann. Math. Statist. 37, 871-890.

[7] Stroock, D. W. and Varadhan, S.R.S. (1969): Diffusion processes with continuous coefficients I, XXII, 345-400.

[8] ... ... (1969): Diffusion processes with continuous coefficients II, XXII, 479-530.