The number (which is finite or infinite) defining loss at time n is put in correspondence with each trajectory h at time $n \geq 1$ for each parameter value θ. The functions $q_n^\theta(\cdot)$, $n \geq 1$ below are defined with respect to it.

(6) $q_n^\theta(h) := q_n^\theta$, a *cost* function which defines the *loss* at time n, is a measurable function of the pairs $\theta \in \Theta$ and $h \in \mathcal{H}_n$ taking values in $(-\infty, +\infty]$.

Similarly to the approach often taken in mathematical statistics, it is assumed that the choice of control can be random.

The (statistician's) *strategy* is a sequence of transition functions $\pi := \{\pi_{n+1}(\cdot|h), \ n \geq 0\}$, where $\pi_{n+1}(\cdot|h)$ is a transition function from $\mathcal{H}_n$ into $\mathcal{A}_{n+1}$ and it is assumed that for each fixed $h \in \mathcal{H}_n$ the measure $\pi_{n+1}(\cdot|h)$ is concentrated on the set of actions admissible for h. The set of all possible strategies is designated by Π.

For any $n \geq 1$ and a fixed $\theta \in \Theta$, the strategy of nature p^θ and the statistician's strategy π define a measure on $\mathcal{H}_n$. If we designate by E_θ^π the expectation of this measure, then for any function $q(h_n)$, $h_n \in \mathcal{H}_n$, bounded from below

$$E_\theta^\pi q(h_n) :=$$

$$\int_{\mathcal{X}_0} p_0^\theta(dx_0) \int_{\widetilde{\mathcal{H}}(x_0)} \pi_1(da_1|x_0) \int_{\mathcal{H}(x_0a_1)} p_1^\theta(dx_1|x_0a_1)\ldots \quad (2.33)$$

$$\int_{\widetilde{\mathcal{H}}(h_{n-1})} \pi_n(da_n|h_{n-1}) \int_{\mathcal{H}(h_{n-1}a_n)} p_n^\theta(dx_n|h_{n-1}a_n)q_n(h_{n-1}a_nx_n),$$

where $h_n := h_{n-1}a_nx_n$ and $\widetilde{\mathcal{H}}(h_{r-1})$ (correspondingly $\mathcal{H}(h_{r-1}a_r)$) for $1 \leq r \leq n-1$ is a section at the point h_{r-1} ($\tilde{h} := h_{r-1}a_r$) of the set $\widetilde{\mathcal{H}}_r$ ($\mathcal{H}_r$).

By the Ionescu–Tulcea theorem (see Neveu 1969) a fixed $\theta \in \Theta$, a strategy of nature p and a statistician's strategy π, define by (2.33) a measure P_θ^π on the infinite product space $\mathcal{X}_0 \times \mathcal{A}_1 \times \ldots$, which, as is easily seen, is concentrated on $\mathcal{H}_\infty$ (see, for example, Neveu, *op. cit.*, Theorem 3.4.1 with remarks on p. 158). Expectation with respect to this measure is also denoted by E_θ^π. The cost functions q_n^θ may also be considered as measurable functions defined on $\mathcal{H}_\infty$. The *cost of the*

strategy π for fixed θ at time $\nu \leq \infty$ is denoted by

$$w_\nu^\pi(\theta) := \sum_{n=1}^\nu E_\theta^\pi q_n^\theta, \quad (2.34)$$

in cases where the corresponding expectations and sums are defined.

Since we want to consider the functional analogous to (2.34) in the case where an initial distribution $\mu \in \mathcal{P}(\Theta)$ of the unknown parameter θ is given, and the functions $q_n^\theta(h)$ are measurable with respect to the pair θ, h, it is convenient to consider the measures P_θ^π on the product space $\Theta \times \mathcal{H}_\infty$. The measure μ and the strategy π define a measure P_μ^π on $\Theta \times \mathcal{H}_\infty$ in the following way. If we designate by E_μ^π the expectation with repect to this measure, then for any bounded measurable $f(\theta, h)$

$$E_\mu^\pi f(\theta, h) := \int_\Theta (E_\theta^\pi f(\theta, h))\mu(d\theta).$$

Then P_θ^π may be considered as P_μ^π with the measure μ concentrated at the point θ. Accordingly the cost of the strategy π up to time $\nu \leq \infty$ with initial distribution $\mu \in \mathcal{P}(\Theta)$ is denoted by

$$w_\nu^\pi(\mu) := \sum_{n=1}^\nu E_\mu^\pi q_n^\theta. \quad (2.35)$$

For the expressions on the right-hand side of (2.34) and (2.35) to make sense for all $\theta \in \Theta$ and $\pi \in \Pi$, and also for the equality

$$w_\nu^\pi(\mu) = \int_\Theta w_\nu^\pi(\theta)\mu(d\theta), \quad \nu \leq \infty \quad (2.36)$$

to hold for all $\mu \in \mathcal{P}(\Theta)$, suitable assumptions on q_n^θ are necessary. In the sequel we assume that the following condition holds:

(**A0**)   For each n the functions $q_n^\theta(h)$ are bounded from below.

The condition (A0) obviously guarantees the existence of $W_\nu^\pi(\mu)$ and equation (2.36) for $\nu < \infty$. If $\nu = \infty$, then for convergence of the series in (2.35) and for the equality (2.36) to hold, which in this case is defined as

$$\lim_{\nu \to \infty} w_\nu^\pi(\mu) = \int_\Theta \lim_{\nu \to \infty} w_\nu^\pi(\theta)\mu(d\theta), \quad (2.37)$$

additional requirements are needed.

In §2.3 we will give a condition (A2) which will be sufficient for the existence of $w_\infty^\pi(\mu)$ and for (2.37) to hold.

Remark 2.3 Since in the future we will be interested only in the functional $w_\nu^\pi(\mu)$, without loss of generality we may consider that for $n \geq 1$ the functions $q_n^\theta(h)$ depend not on $h \in \mathcal{H}_n$, but only on $\tilde{h} \in \widetilde{\mathcal{H}}_n$. It suffices in this case to replace $q_n^\theta(h)$ by

$$\overline{q}_n^\theta(\tilde{h}) := \int q_n^\theta(\tilde{h}x)p_n^\theta(dx|\tilde{h}), \qquad (2.38)$$

where the integral is taken over the section of the set $\mathcal{H}_n$ at the point $\tilde{h}$. The functional (2.35) does not change with the replacement of $q_n^\theta(h)$ by $\overline{q}_n^\theta(\tilde{h})$, which follows from (2.33). ∎

$$* \quad * \quad *$$

The aim of the statistician is optimization of $w_\nu^\pi(\mu)$ in one sense or another, but before discussion of the different criteria of optimization and related questions, we will dwell briefly on some other formulations of the general control scheme with incomplete information. Afterwards we will show that the problem given in §2.1 may be presented in the framework of the general scheme described in this section.

The description given above is conducted in terms of (sample) *path* space, i.e. in terms of trajectories and distributions on the space of trajectories. However, similarly to other problems in probability, the presentation may be given alternatively in terms of a *basic* space $(\Omega, \mathcal{F})$ on which random values are given. The space $\Theta \times \mathcal{H}_\infty$ may be taken as such a space Ω where the coordinate measurable functions (random variables) θ and x_n, a_{n+1}, $n = 0, 1, \ldots$ are given. The elements of the sets $\mathcal{H}_n$ $(\widetilde{\mathcal{H}}_{n+1})$ define the σ-algebras $\mathcal{F}_n$ $(\widetilde{\mathcal{F}}_{n+1})$. In other words, these σ-algebras are defined by the variables x_s for $0 \leq s \leq n$ and a_s for $1 \leq s \leq n$ $(1 \leq s \leq n+1)$. The σ-algebra $\mathcal{F}$ is defined as $\mathcal{F} := \mathcal{F}_\infty^\theta := \mathcal{F}_\infty \vee \sigma(\theta)$, where $\mathcal{F}_\infty := \bigvee_{n=1}^\infty \mathcal{F}_n$.

The strategy of nature $p^\theta := \{p_n^\theta(\cdot|h), \ n = 0, 1, \ldots\}$ with $\tilde{h}$ and θ replaced by the corresponding random variables defines for each value θ an initial distribution on $\mathcal{X}_0$ and a sequence of transition probabilities from $\mathcal{F}_{n+1}^\theta$ into $\mathcal{X}_{n+1}$ considered as measurable random measures. Similarly, the statistician's strategy $\pi := \{\pi_{n+1}(\cdot|h), \ n = 0, 1, \ldots\}$ defines a sequence of $\mathcal{F}_n$-measurable random measures on $\mathcal{A}_{n+1}$. The strategy of nature and the statistician's strategy define for each value θ a measure on $\mathcal{F}_\infty$, and if in addition a measure $\mu \in \mathcal{P}(\Theta)$ is given, a probability measure on $\mathcal{F}$ may be defined.

Of course, Ω need not *necessarily* coincide with path space, in which case it must be considered as an abstract measure space which must be sufficiently "rich" and on which are given the corresponding random variables taking values in Θ, $\mathcal{X}_n$, $\mathcal{A}_{n+1}$, $n = 0, 1, 2, \ldots$, and satisfying the appropriate constraints $\mathcal{H}_n$ and $\mathcal{H}_{n+1}$. In this case strategies of nature and the statistician's strategies are defined simply as sequences of random measures, measurable with respect to the corresponding σ-algebras.

At the first look, such an approach to the definition of strategies seems wider; however, this is not so by the theorem according to which a random function f, measurable with respect to the σ-algebra generated by a random variable g, may be represented as a measurable function with respect to g.

So, in discrete time, the approach in which strategies are functions of points in a path (sample) space of observations and controls, and the approach in which observations, controls and strategies are random variables given on some abstract space Ω, are equivalent. In continuous time it is convenient to conduct the presentation exclusively on the basis of the second approach. For this reason, for comparison of the two cases, we have presented the basic discrete time scheme with the same abstract approach.

We show now how the basic scheme of §2.1 is formulated in the form of a general control problem with incomplete information. As the states of system at time n it is sufficient to take values of the increments of the process X at time n and to consider that $\mathcal{X}_0 := \{e_0^m\}$ and all $\mathcal{X}_n$ for $n \geq 1$ coincide with $\widehat{S}_0^m$, all $\mathcal{A}_n$ for $n \geq 1$ coincide with $\widehat{S}^m$, Θ coincides with $\widehat{S}^N$ for $n \geq 1$, x_n may take $(m+1)$ possible values corresponding to a jump at time n on one of the m devices or to the absence of a jump, a_n may take m possible values pointing to the index of the device which is observed at time n, and θ may take N values defining the index of the hypothesis about the device parameters.

Further, $\mathcal{H}_\infty$ is a set of infinite sequences in which the first element is e_0^m, even coordinates contain elements from $\widehat{S}^m$ and if the $2s^{\text{th}}$ coordinate, $s \geq 1$ contains the element e_j^m, i.e. at time s the jth device is used, then the $(2s+1)^{\text{st}}$ coordinate contains either e_j^m (success) or e_0^m (failure). The measurable space $\mathcal{H}_n$ $(\widetilde{\mathcal{H}}_{n+1})$ is a set of finite sequences

of length $2n + 1$ ($2n + 2$) possessing the properties mentioned in the definition of $\mathcal{H}_\infty$.

If $\theta := e_i^N$, $\widetilde{h} := x_0 a_1 \ldots x_n a_{n+1} \in \widetilde{H}_{n+1}$ and $a_{n+1} := e_j^m$, then

$$p_{n+1}^\theta(e_j^m | \widetilde{h}) = \lambda_i^j, \qquad p_{n+1}^\theta(e_0^m | \widetilde{h}) = 1 - \lambda_i^j. \qquad (2.39)$$

In order to maximize the *cumulative number of successes* up to time ν, it is necessary to take

$$q_n^\theta(h) := \sum_{j=1}^m x_n^j \quad \text{if } h = x_0 a_1 \ldots a_n x_n, \quad n \geq 1, \qquad (2.40)$$

in the sum with respect to n in (2.34).

In light of Remark 2.3 and (2.39), (2.40) may be replaced by the function

$$q_n^\theta(h) := \lambda_i^j \quad \text{if } \theta := e_i^N, \quad a_n := e_j^m. \qquad (2.41)$$

To minimize loss up to time ν, $\nu \leq \infty$, we must put

$$q_n^\theta(h) := (\lambda_i - \lambda_i^j) \quad \text{if } \theta := e_i^N, \quad a_n := e_j^m, \qquad (2.42)$$

where $\lambda_i := \max_{1 \leq j \leq m} \lambda_i^j$.

Any statistician's strategy π may be considered as a sequence of vector-valued functions $\pi_{n+1}(a_1 x_1 \ldots a_n x_n)$ defined on $\mathcal{H}_n$ and taking values in S^m. Replacing x_s and a_s in the function π_{n+1} by the values of process $\Delta X(s, \omega)$ and $a(s, \omega)$ from §2.1, $1 \leq s \leq n$, we obtain an action rule in the sense of §2.1, and vice versa. Since for each n the σ-algebra $\mathcal{F}_n$ generated by the values of $\Delta X(s, \omega)$, $a(s, \omega)$, $1 \leq s \leq n$, is finitely generated, then it is obvious that any action rule β can be represented in the form of sequences of vector-valued functions

$$\beta(n+1) := \pi_{n+1}(a(1, \omega), \Delta X(1, \omega), \ldots, a(n, \omega), \Delta X(n, \omega)),$$

$$n \geq 0, \qquad (2.43)$$

defining a strategy of the statistician.

The measure defined by the action rule β on $\mathcal{F}_\infty^\theta$ defines a measure on $\Theta \times \mathcal{H}_\infty$, which may be seen to coincide with the measure arising from the strategy π obtained by formula (2.43). Corresponding values of the criterion functionals will also coincide.

Remark 2.4 Statistical decision theory is a special case of the general scheme formulated above (see, for example, Wald 1967). Indeed, it suffices to consider the control spaces $\mathcal{A}_n$ as divided into two parts. One part corresponds to the decision to *stop* the observations, and its elements define a final decision, while the second part corresponds to the decision to *continue* the observation, and its elements define the number and character of the following observations up to the time of taking the next decision. It is assumed here that all $\mathcal{X}_n$ for $n \geq 1$ contain some absorbing state $\bar{x}$. Then for $h := x_0 a_1 \ldots x_{n-1} a_n \bar{x}$, where a_n is the final decision, $a_n \neq \bar{a}$, where $\bar{a}$ is the unique control in state $\bar{x}$, the function $q_n^\theta(h)$ defines the cost connected with taking the decision to *stop* the observations. For $h := x_0 a_1 \ldots \bar{x} \bar{a} \bar{x}$, the function $q_n^\theta(h)$ is assumed to be equal to 0, and for $h := x_0 a_1 \ldots x_{n-1} a_n x_n$, where a_n is a decision to *continue*, the function $q_n^\theta(h)$ defines the cost to continue further observation. ∎

$$* \qquad * \qquad *$$

Now we consider the optimality criterion. Let Γ be a subspace of the strategy space Π, and Δ be a subspace of $\mathcal{P}(\Theta)$. The strategy $\pi^* \in \Gamma$ is called Δ-*optimal with respect to* Γ on the interval $[0, \nu)$, $\nu \leq \infty$, written $\pi^* \in \Gamma_\nu^*(\Delta)$, if

$$\sup_{\mu \in \Delta} w_\nu^{\pi^*}(\mu) = \inf_{\pi \in \Gamma} \sup_{\mu \in \Delta} w_\nu^\pi(\mu). \qquad (2.44)$$

The corresponding optimal value of the functional in (2.44) we denote by $w_\nu(\Delta, \Gamma)$, and if $\Gamma := \Pi$ we write simply $w_\nu(\Delta)$. Further, in discussing optimality criteria we consider that ν is *fixed* (finite or infinite) and sometimes we will omit the index ν in the notation $w_\nu^\pi(\mu)$, w_ν^π and $\Gamma_\nu^*(\Delta)$.

If Δ consists of single measure $\mu \in \mathcal{P}(\Theta)$, then we speak of the *Bayesian formulation*, and μ is called the *a priori* distribution. In this case $\pi^* \in \Pi^*(\mu)$ if, and only if,

$$w^{\pi^*}(\mu) = \inf_{\pi \in \Pi} w^\pi(\mu). \qquad (2.45)$$

Such a strategy π^* is called *Bayesian* with respect to the *a priori* distribution μ, and the optimal value of the functional is denoted by $w(\mu)$, or simply w, if it is obvious about which μ we speak.

Further, if Δ coincides with $\mathcal{P}(\Theta)$, then $\pi^* \in \Pi^*(\mathcal{P}(\Theta))$ if, and only if, π is a *minimax strategy*, i.e.

$$\sup_{\theta \in \Theta} w^{\pi^*}(\mu) = \inf_{\pi \in \Pi} \sup_{\theta \in \Theta} w^\pi(\theta). \tag{2.46}$$

Here we have used the equality

$$\sup_{\theta \in \Theta} w^\pi(\theta) = \sup_{\mu \in \mathcal{P}(\Theta)} w^\pi_\mu, \tag{2.47}$$

which holds for all $\pi \in \Pi$.

If, for some ρ_0, $0 \le \rho_0 \le 1$, and some $\mu_0 \in \mathcal{P}(\Theta)$, we let $\Delta := \{\mu : \mu = \rho_0 \mu_0 + (1 - \rho_0)v, \ v \in \mathcal{P}(\Theta)\}$, then we may consider the optimality criterion of *Hodges and Lehman* (see Raiffa & Schlaifer 1961), for which

$$\sup_{\mu \in \Delta} w^\pi(\mu) = \rho_0 w^\pi(\mu_0) + (1 - \rho_0) \sup_{\theta \in \Theta} w^\pi(\theta).$$

Finally, if Θ is represented as the sum of mutually exclusive sets $\Theta = \bigcup_i \Theta_i$ and the set Δ is defined as $\Delta := \{\mu \in \mathcal{P}(\Theta), \ \mu(\Theta_i) = p_i\}$ for some $p_i \ge 0$ such that $\sum p_i = 1$, then we obtain *Menges* optimality criterion (see Raiffa & Schlaifer, *op. cit.*):

$$\sup_{\mu \in \Delta} w^\pi(\mu) = \sum_i p_i \sup_{\theta \in \Theta_i} w^\pi(\theta).$$

When studying the different optimality criteria the following questions are usually interesting:

1. What are the conditions for the *existence* of the minimax and Bayesian strategy?
2. When does the value function for a finite time horizon *converge* to the value function for $\nu = \infty$ and what is the behaviour of the finite optimal strategy in the limit?
3. If optimal strategies exist, how do we *find* them and what *properties* do they have?

With the exception of §7.4, where the minimax formulation is considered, we will study problems in the Bayesian formulation in this book. In this formulation two basic approaches exist to answer the above questions.

The first approach, useful in answering questions 1 and 2, is the following. A topology is introduced on the strategy space Π in such a way that Π becomes compact in this topology and $w^\pi(\mu)$ is a lower semicontinuous function with respect to π for each fixed μ. We discuss this approach in §2.3.

The second approach is applicable *only* in the Bayesian formulation and may be used to answer questions 1–3. It is connected to the derivation of the Bellman equation for the optimal value of the criterion functional, the *value function*, and is presented in §2.4.

As usual, the discussion of both approaches begins with the general case, and then the corresponding proofs for the basic scheme of §2.1 are given. Note that questions 1–3 for the case of the basic scheme coincide with the questions formulated in §2.1 about existence and properties of optimal action rules and about the coincidence of $F_\infty(\xi)$ and $F(\xi)$.

2.3 Existence of an optimal action rule, coincidence of $F_\infty(\xi)$ and $F(\xi)$ and convexity of $F_\nu(\xi)$

First we give simple (previously known) examples showing that, generally speaking, optimal strategies may *not* exist, and finite horizon optimal values do *not* necessarily converge to infinite horizon optimal values.

In these examples, Θ consists of one point, all $\mathcal{X}_n$, $\mathcal{A}_{n+1}$, $n \ge 0$, coincide with some $\mathcal{X}$ and $\mathcal{A}$, the strategy of nature depends only on the last control and is deterministic, and the initial distribution concentrates on one point x^0, so that $p_0(x_0 = x^0) = 1$. Cost at time n depends only on the control at time n, so that $q_n(h) := q(a_n)$, and optimality is interpreted as *minimization* of the functional w^π_ν, $\nu \le \infty$.

In the directed graph (*transition digraph*) of Figure 4 the nodes correspond to states, the directed arcs to controls, and the numbers on the arcs show the value of the function q for the corresponding controls.
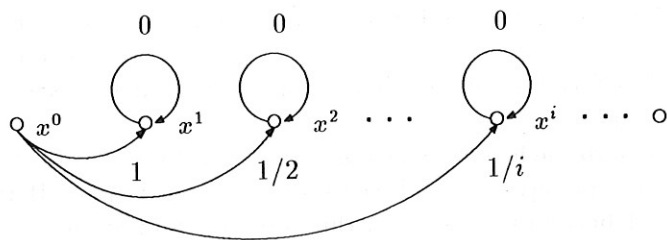
Figure 4
The transition digraph corresponding to Example 2.1

Example 2.1 *(See Figure 4.) The space $\mathcal{X}$ consists of the points $x^0, x^1, \ldots, x^n, \ldots$. The system may be transferred in one step from the initial state x^0 to any state x^i, $i \geq 1$. The cost for the transfer to state i is $1/i$, the states x^i, $i \geq 1$, are absorbing and the cost to remain in each is 0.*

In this example, in which the initial distribution is concentrated on the point x^0, $w_\nu = w_\infty = 0$ holds. Indeed, if strategy π prescribes a transfer to x^k at the first step, then further movement is determined and for any $\nu \geq 1$, $w_\nu^\pi = w_\infty^\pi = 1/k$ holds, which means $\inf_\pi w_\nu^\pi = \inf_\pi w_\infty^\pi = 0$. However, for any $\nu < \infty$ an optimal strategy does *not* exist since this infimum cannot be attained. This is connected with the fact that control space is not compact in some sense. ∎

It is easy to give examples in which, by the same reasoning, the equality $w_\infty = \lim_{\nu \to \infty} w_u$ does not hold, in spite of the fact that optimal strategies exist.

Figure 5
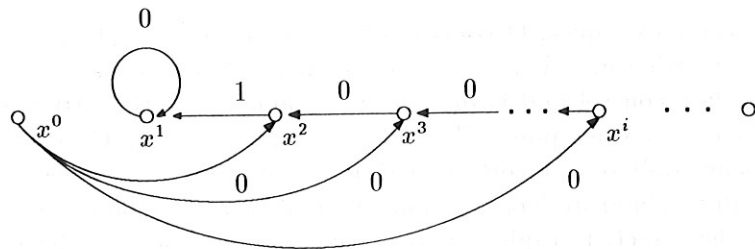The transition digraph corresponding to Example 2.2

Example 2.2 *(See Figure 5.) The space $\mathcal{X}$ consists of the points $x^0, x^1, \ldots$. The system may be transferred in one step from the initial state x^0 to any state x^i, $i \geq 1$, and the cost for this transfer equals 0. A deterministic transition occurs from state x^i, $i \geq 2$, to state x^{i-1}, with cost equal to 0 for $i \geq 3$ and to $+1$ for $i = 2$. The state x^1 is absorbing with cost 0.*

In this example, the strategy prescribing a transfer from the initial point x^0 to x^i with $i \geq \nu + 1$ at the first step is obviously optimal for the problem with horizon ν, so that $w_\nu = 0$. At the same time, for any strategy π, $w_\infty^\pi = w_\infty = 1$ holds, so that convergence of w_ν to w_∞ is not possible. ∎

Although with natural assumptions on the model, compactness of the control space guarantees the existence of an optimal strategy, this is not sufficient for convergence of w_ν to w_∞, as is shown by the following example.
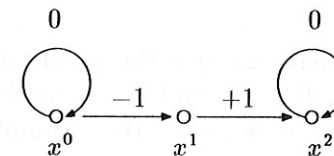

Figure 6
The transition digraph corresponding to Example 2.3

Example 2.3 *(See Figure 6.) The space $\mathcal{X}$ consists of three points x^0, x^1, x^2. It is possible to stay at x^0 (in this case the cost equals 0) or to transfer to x^1 (when the cost equals -1). A deterministic transition occurs from x^1 to x^2 with cost $+1$ and x^2 is an absorbing state with cost 0.*

In this example, the strategy from initial point x^0 which prescribes staying $\nu - 1$ times at x^0 and then transferring to x^1 is optimal in the problem with ν steps and $w_\nu = -1$. At the same time, for any strategy π, $w_\infty^\pi = 0$ holds, so that $w_\nu \not\to w_\infty$. ∎

Convergence of w_ν to w_∞ does not hold here because of two effects. First, for the optimal strategy for $\nu < \infty$ the distant moment of time make a substantial negative contribution to the value function. Second, any *extension* of the optimal strategy for finite horizon ν also makes a substantial positive contribution which does not tend to 0 as $\nu \to \infty$. Thus sufficient conditions for the equality $w_\infty = \lim_{\nu \to \infty} w_\nu$ to hold in the general case are connected with the elimination of these effects.

Now we return to the general case and *assume* that $w_\infty^\pi(\theta)$ exists for each π and θ, that for each π

$$w_\nu^\pi(\theta) \to w_\infty^\pi(\theta) \quad \text{for } \nu \to \infty \quad \text{uniformly with respect to } \theta, \quad (2.48)$$

and that $\int w_\infty^\pi(\theta)\mu(d\theta)$ exists for any $\mu \in \Delta$.

Note that from (2.48) we may in particular derive that (2.36) holds for $\nu = \infty$.

We prove now that under assumption (2.48) the inequality

$$\limsup_{\nu \to \infty} w_\nu(\Delta, \Gamma) \le w_\infty(\Delta, \Gamma) \quad (2.49)$$

always holds.

For reasons of simplicity we give the proof for the case in which Θ consists of one point, $\Gamma := \Pi$, and $w_\infty < \infty$. For arbitrary $\varepsilon > 0$, we choose π' so that $w_\infty^{\pi'} < w_\infty + \varepsilon$. By assumption $w_\nu^{\pi'} \to w_\infty^{\pi'}$ for $\nu \to \infty$, which means that

$$\limsup_{\nu \to \infty} w_\nu = \limsup_{\nu \to \infty} \inf_\pi w_\nu^\pi \le \limsup_{\nu \to \infty} w_\nu^{\pi'} < w_\infty + \varepsilon,$$

which it was required to prove.

Next we give a condition (A1) which eliminates the second effect present in Example 2.3 and guarantees under assumption (2.48) the convergence of $w_\nu(\Delta, \Gamma)$ to $w_\infty(\Delta, \Gamma)$.

For $\theta \in \Theta$, $\pi \in \Pi$ and $0 \le s \le \nu < \infty$ let

$$R_{s\nu}^\pi(\theta) := \sum_{n=s+1}^\nu E_\theta^\pi q_n^\theta(h), \quad (2.50)$$

so that $R_{0\nu}^\pi(\theta) = w_\nu^\pi(\theta)$. Expectations in (2.50) are defined for functions $q_n^\theta(h)$ bounded from below and taking values in $(-\infty, +\infty]$. Consider:

$$\textbf{(A1)} \qquad \limsup_{s \to \infty} \; \sup_{\pi \in \Pi, \theta \in \Theta, \nu > s} R_{s\nu}^\pi(\theta) \le 0.$$

Remark 2.5 Condition (A1) obviously holds in cases where the cost is nonpositive, and also in cases where there exist constants c_n, $n = 0, 1, 2, \ldots$, such that $\sum c_n < \infty$ and $E_\theta^\pi q_n^\theta < c_n$ for any π and θ (which corresponds to a model *bounded from below* in the terminology of Dynkin & Yushkevitch 1976). ∎

Denote by $\bar{z}_s := \sup_{\pi \in \Pi, \theta \in \Theta, \nu \ge s} R_{s\nu}^\pi(\theta)$. Then, obviously

$$R_{0\nu}^\pi(\theta) := w_\nu^\pi(\theta) \le w_s^\pi(\theta) + \bar{z}_s \quad \text{if } s \le \nu.$$

From this and from (A1) it is easy to see that $\lim_{u \to \infty} w_\nu^\pi(\theta)$ exists and it may be equal to $-\infty$.

We show that from (A1) and (2.48) it follows that

$$\lim_{\nu \to \infty} w_\nu(\Delta, \Gamma) = w_\infty(\Delta, \Gamma). \quad (2.51)$$

The proof is again given for the case in which Θ consists of a single point, $\Gamma := \Pi$ and $w_\infty < \infty$. Then, for each $1 \le \nu \le \infty$,

$$w_\infty = \inf_\pi (w_\nu^\pi + R_{\nu\infty}^\pi) \le \inf_\pi w_\nu^\pi + \sup_\pi R_{\nu\infty}^\pi = w_\nu + \sup_\pi R_{\nu\infty}^\pi.$$

From this

$$w_\infty \le \liminf_{\nu \to \infty} w_u + \limsup_{\nu \to \infty} \sup_\pi R_{\nu\infty}^\pi.$$

From (A1) we obtain $w_\infty \le \liminf_{\nu \to \infty} w_\nu$, which together with (2.49) gives (2.51).

However, in many cases (2.51) holds when Condition (A1) does not. This is the case for the loss minimization problem in §2.1. At the end of this section it will be proved that (2.51) holds for this problem. We show now that for this problem Condition (A1) does *not* hold, with the exception of one trivial case where all elements of the rows of the matrix corresponding to hypotheses with positive *a priori* probability are the same. Indeed, consider the action rule $\bar{\beta}$ which consists of using the jth device at each step, where j is a column of the hypothesis matrix not all of whose elements are maximal. For this action rule, by formula (2.10),

$$R_{s\nu}^{\bar{\beta}}(\xi) = (\nu - s) \sum_{i=1}^N \xi_i(\lambda_i - \lambda_i^j), \quad (2.52)$$

and, if $\xi_i > 0$ for $i = 1, \ldots, N$, then the sum on the right-hand side of (2.52) is positive, which means that $w_\infty^{\bar{\beta}}(\xi) = \infty$, i.e. (A1) does not hold.

Now, take an arbitrary model where (2.51) holds and $w_\infty \leq q < 1$; add to each state the new control a^* by which a deterministic transition into the new absorbing state x^* is performed. Let the cost for control a^* equal $+1$, and the cost in state x^* equal 0. Then it is obvious that optimal strategies for $\nu \leq \infty$ do not change and (2.51) holds. At the same time Condition (A1) does not hold.

We now formulate Condition (A2) which eliminates the first effect present in Example 2.3 and, together with compactness of the strategy space, gives existence and convergence of optimal strategies and (2.51):

$$(\mathbf{A2}) \qquad \liminf_{s \to \infty} \inf_{\pi \in \Pi, \theta \in \Theta, \nu > s} R_{s\nu}^\pi(\theta) \geq 0.$$

Remark 2.6. Notice that Condition (A2) clearly holds when the costs are nonnegative, and also in cases where there exist constants c_n, $n = 0, 1, \ldots$, such that $\sum c_n < \infty$, and $E_\theta^\pi q_n^\theta > -c_n$ for any π and θ (which corresponds to the model *bounded from above* in the terminology of Dynkin & Yushkevitch 1976). ■

Condition (A2) obviously does not imply Condition (A1) (and conversely) and, similarly to (A1), is *not necessary* for the convergence of (2.51).

Set

$$z_s := \inf_{\pi \in \Pi, \theta \in \Theta, \nu \geq s} R_{s\nu}^\pi(\theta), \qquad (2.53)$$

then

$$R_{0\nu}^\pi(\theta) = w_\nu^\pi(\theta) \geq w_s^\pi(\theta) + z_s, \qquad s \leq \nu. \qquad (2.54)$$

Now, let the model be such that we may introduce a topology on Π possessing the following properties. In this topology Π is compact and for each $\mu \in \Lambda \subset P(\Theta)$ and for each finite ν the function $w_\nu^\pi(\mu)$ is lower semicontinuous[2] with respect to the argument π. Since the operation of taking the supremum preserves lower semicontinuity and

---

[2] A function $f(x)$ is called *lower semicontinuous* if all sets $\{x : f(x) \leq c\}$ are closed. A function is lower semicontinous if and only if it is the (pointwise) limit of a non-decreasing sequence of continuous functions.

any lower semicontinuous function attains its minimum on a compact set, it follows from our assumptions that for any closed set $\Gamma \subset \Pi$ and for any $\nu < \infty$ there exists a Δ-optimal strategy with respect to Γ. Now if, in addition, Condition (A2) holds, then there exists also a Δ-optimal strategy with respect to Γ for $\nu = \infty$ and $\lim_{\nu \to \infty} w_\nu(\Delta, \Gamma) = w_\infty(\Delta, \Gamma)$. Here, if π^* is the limit point of a sequence $\{\pi_\nu\}$, where $\pi_\nu \in \Gamma_\nu^*(\Delta)$, then $\pi^* \in \Gamma_\infty^*(\Delta)$ (see the notation following Remark 2.4). These statements follow easily from (2.54) and the following simple lemma taking $\Pi := \mathcal{X}$ and $\Theta := \mathcal{Y}$.

Lemma 2.2 *Let a sequence of functions* $\{f_n(x, y), n = 0, 1, \ldots, f_0 := 0\}$ *be given on the product of some set* $\mathcal{X}$ *and a measurable space* $\mathcal{Y}$, *taking values in* $(-\infty, +\infty]$ *and satisfying for all* $\nu \geq s$ *the inequalities* $f_\nu(x, y) \geq f_s(x, y) + \varepsilon_s$, *where* $\liminf \varepsilon_s \geq 0$ *as* $s \to \infty$. *Then:*

(a) there exists a limit function $f_\infty(x, y) = \lim_{\nu \to \infty} f_\nu(x, y)$

(b) for any measure $\mu \in \mathcal{P}(\mathcal{Y})$

$$\lim_{\nu \to \infty} \int f_\nu(x, y) \mu(dy) = \int f_\infty(x, y) \mu(dy).$$

If, in addition, $\mathcal{X}$ *is a compact topological space and the integral* $\int f_\nu(x, y) \mu(dy)$ *is lower semicontinuous with respect to* x *for each* ν, $0 \leq \nu < \infty$, *and each* $\mu \in \Delta \subset \mathcal{P}(\mathcal{Y})$, *then for any closed set* $\Gamma \subset \mathcal{X}$

$$\inf_{x \in \Gamma} \sup_{\mu \in \Delta} \int f_\nu(x, y) \mu(dy) \to \inf_{x \in \Gamma} \sup_{\mu \in \Delta} \int f_\infty(x, y) \mu(dy) \qquad (2.55)$$

as $\nu \to \infty$ *and for each* ν *there exists a* x_ν *achieving the infimum in the left-hand side of (2.55). Moreover, any limit point of the sequence* $\{x_\nu\}$ *achieves the infimum in the right-hand side of (2.55).* ■

We omit the proof of the lemma, which is accomplished using standard methods of mathematical analysis.

Usually to guarantee the existence of the required topology on Π it is necessary to require some conditions on the model. So, in Schäl (1979) it is assumed that the $\mathcal{X}_n$ are Borel spaces, that the set of admissible controls for a given history is compact and that there exist measures $p_n(\cdot|\bar{h})$ such that for each θ, the measures $p_n^\theta(\cdot|\bar{h})$ are absolutely continuous with respect to these measures, and $p_n(\cdot|\bar{h})$, $p_n^\theta(\cdot|\bar{h})$