Chapter 4 - General Theory of Optimality

4.1 Introduction

In this chapter the only requirements made of the loss function \( L(x^T, z^T, u^{t-1}) \) are that it be measurable, non-negative, and possibly infinite. The principal idea exploited here is that of the minimal conditional loss functional \( \hat{\beta}_t(z^t, [u]) \). For \([u]\) fixed, \( \hat{\beta}_t \) is a measurable function of \( z^t \), and for \( z^t \) fixed it depends on the truncated law \( u^{t-1} \) - hence the term "functional." In the usual treatment of the dynamic programming problem (Blackwell [2], Hinderer [8], etc) the minimal conditional loss function depends only on the value of the truncated law \( u^{t-1}(z^{t-1}) \) at the observation point \( z^t \) and not on the entire law \( u^{t-1} \). While the complexity of the functional \( \hat{\beta}_t(z^t, [u]) \) over the function \( \hat{\beta}_t(z^t, u^{t-1}) \) may be a disadvantage, it appears to be outweighed by the "natural" properties of \( \hat{\beta}_t \). Section 4.4, in which a comparison of the two criteria is made is not required for the development of the argument and is included primarily to show the relation of the present approach to that of the literature of dynamic programming.

The principal advantage of the functional over the function \( \beta_t(z^t, u^{t-1}) \) is that the functional \( \hat{\beta}_t(z^t, [u]) \) exists and is Borel measurable in \( z^t \) for arbitrary observation space \( Z_t \), control space \( u^t \), and class \( \mathcal{U} \) of admissible control laws. The minimal loss function \( \hat{\beta}_t(z^t, u^{t-1}) \) is at best universally measurable, and even this requires the assumption of additional (though natural) properties of the spaces \( Z_t \) and \( u^{t-1} \). More serious is the requirement that the class of admissible laws \( \mathcal{U} \) contain essentially all measurable laws. (See Strauch [16] Theorem 7.1 and Hinderer [8] Theorem 1.3.2). In section 4.4 assumptions are given under which the two criteria are the same in the sense that

\[
\hat{\beta}_t(z^t, [u]) = \hat{\beta}_t(z^t, u^{t-1}(z^{t-1})) \quad \text{a.s.} \quad \hat{\beta}_t(u^{t-1})
\]

for all \([u]\), and an example is given to show that this equality does not always hold.
In section 4.2 properties of the functionals \( \mathcal{F}_t(z^t, [u]) \) are defined - "non-negative", "compatible", "martingale", "sub-martingale", and "closed". The conditional loss functional \( \hat{\mathcal{F}}_t \) and the step-wise conditional loss functional \( \tilde{\mathcal{F}}_t \) are defined making use of the idea of the P-ess inf of a class of non-negative, measurable functions. The definition and required properties of the P-ess inf are given in the appendix section A.2. Also defined in section 4.2 are the finite and countable \( \varepsilon \)-lattice properties of a control system. In the remainder of the section, martingale properties and relationships between the conditional loss functionals \( \hat{\mathcal{F}}_t \) and \( \tilde{\mathcal{F}}_t \) are developed. Theorems 4.2.4 and 4.2.5 are the analogues for countably additive measures of Theorem 1 and Corollary 2 of section 2.14 of Dubins and Savage [5]. The relationship between excessive functions and super-martingales is discussed in section 2.12 of [5].

Section 4.3 is devoted to necessary and sufficient conditions for optimality at a control law. The motivation for the definition of optimal \( (\varepsilon \text{-optimal}) \) for \( [u] \) at \( t \) (Definitions 4.3.1 and 4.3.2) is as follows: control law \( [u] \) has for whatever reason been used and data has been collected up through time \( t \). At this point the question is asked how best to proceed from \( t \) onward. The property optimal \( (\varepsilon \text{-optimal}) \) for \( [u] \) at \( t \) is the same as \( p \)-optimal \( ((p, \varepsilon) \text{-optimal}) \) of Blackwell [2] where \( t = 1 \) and \( \mathcal{P}_{[u]}^t = p \). This property is given in Theorem 4.3.1 and could as well be used as the defining property. The defining properties given in Definition 4.3.1 and 4.3.2 were selected because they are more constructive and more convenient in applications. The properties given in Theorem 4.3.2 are the same as \( \bar{p} \)-optimal and \( \bar{p} \)-\( \varepsilon \)-optimal of Hinderer [8] again with \( t = 1 \) and \( \mathcal{P}_{[u]}^t = p \).

Under the assumption that the system has the countable \( \varepsilon \)-lattice property, it is shown in Theorem 4.3.4 that \( \varepsilon \)-optimal controls always exist. For systems which do not enjoy this property the weaker property (4.3.13) of Theorem 4.3.2 \( (\bar{p} \varepsilon \text{-optimum}) \) is probably a more basic and a better defining property for \( \varepsilon \)-optimality. In fact all the results of this chapter are of practical interest only for systems with the countable \( \varepsilon \)-lattice property.