Experimental Inquiries: A Mathematical Approach

Guy C. Corynen
Lawrence Livermore Laboratory
University of California

The need for a mathematical theory of simulation has developed particularly in recent years because the problems which the mathematical and social sciences are attempting to attack have become increasingly complex and have acquired an increasingly interdisciplinary character.

In this paper, an initial effort has been made to endow the practice of simulation with more rigorous contents. A definition of simulation is developed in terms of two central concepts: a *model* and an *experimental inquiry*.

The former concept was formalized by the author in earlier papers. The latter is developed here as a fourtuple.

$$I = \langle \Psi , Q, M, D \rangle ,$$

where Ψ is a set of *experimental objects*, Q is a set of *questions*, M is a sampling plan or *experimental design*, and D is a *decision strategy*.

The essential purpose of this paper is to summarize efforts which have been directed towards the development of a framework to rigorously unify the important methodological principles underlying the practice of simulation, and initial results demonstrate the usefulness of the formulation. As an application to the modeling of utility structures, it is shown that, if we associate continuous utility functions u and u' with the modeled and the model, respectively, a weak relation between u' and u guarantees that the assignment of utilities is consistent.

Finally, the developed simulation framework is used to determine the probability that simulation predictions are ε - correct for simulation experiments of arbitrary length n.

Introduction

In this paper we address the experimental issues associated with models. In particular, we intend to develop a framework within which simulation problems can be discussed. This framework is an extension of the modeling structure developed in Refs. 1 and 2 and involves one new central concept: an *experimental inquiry*.

An *experimental inquiry* I is a 4-tuple $\langle \Psi, Q, M, D \rangle$, where Ψ is a set of *experimental objects*, Q is a set of *questions*, M is a *sampling plan* or *experimental design*, and D is a *decision strategy*. We shall see that simulation is a particular type of inquiry called a *simulation inquiry*, SI.

The report is organized as follows. First we summarize our view of "model." Then we discuss in detail what is meant by an inquiry. This is done in Chapters 2 through 5, where experimental objects Ψ, questions Q, experimental designs M, and decision strategies D are discussed respectively. Then simulation inquiries are addressed in Chapter 6, where a precise definition of *simulation* is developed which distinguishes simulation activities from modeling activities. This last section also presents some theorems and an example.

Because much of the work in this paper is definitional, a guiding objective was to present a reasonably general treatment. For instance, the sampling plan M is quite general and allows the discussion of stopping rules and other sequential designs. The decision strategy D is equally flexible, and we show specifically that standard statistical techniques such as sequential hypothesis testing, estimation, and confidence sets are incorporated in a natural way.

Since this paper depends upon some of the concepts developed earlier, [1] it will be convenient to summarize these at this time.

The concept of "model" is developed from the following ingredients:

1. Two sets of objects S and S′, the modeled *(prototypes)* and the models *(model candidates)*.
2. Two *performance spaces* V and V′ for S and S′, whose elements are values of two performance criteria chosen by the modeler.
3. Two *performance criteria* or evaluators Φ and Φ'. With these criteria the modeler isolates the prototype and model features of interest and partially characterizes the process of imitation. For most of the remaining work the objects S and S′ will be mathematical objects (as opposed to physical objects, for instance), in which case these criteria are maps Φ: S → V and Φ': S′ → V′.
4. An *interpretation relation* $H_V \subset V' \times V$. This is the object which permits the modeler to indicate things about the prototype on the basis of properties or observations from the model. For the remainder of this work, it shall be a function H_V: V′ → V.
5. Two *topologies* T_V and T_V' on V and V′, which are intended to characterize experimental or observational errors and the inaccuracies that may or may not be tolerated by the modeler.

These topologies are also intended to capture the process of approximation from our point of view and to allow the discussion of the relevant convergence and continuity issues.

Let us state our immediate objective: Given two objects $s \in S$ and $s' \in S'$, define the conditions under which s' is a model of s.

First we require the following definition:

Definition 1.1: A *modeling criterion* for two sets of objects S and S' is a 4-tuple:

$$C = \langle \Phi, \Phi', T_V, T_{V'} \rangle,$$

where

Φ: $S \to V$ is the *prototype performance function* into the *prototype performance space* V.

Φ': $S' \to V'$ is the model *performance* function into the model *performance space* V',

T_V and $T_{V'}$ are two *performance topologies* for V and V', respectively.

Figure 1 illustrates the interrelationship between these various terms and the interpretation H_V: $V' \to V$. The purpose of the pair $\langle C, H_V \rangle$ is to summarize the modeler's view, and we shall occasionally refer to $\langle C, H_V \rangle$ itself as "the modeler."

One more concept must be defined before we can proceed with the major definition.

Definition 1.2: Given a topological space $\langle V, T_V \rangle$, then $v_2 \in V$ is T_V-*close* to† $v_1 \in V$ iff $\nexists O(v_1) \in T_V \ni v_2 \notin O(v_1)$. We write this as v_2 $CL(T_V)v_1$.

Note that v_2 $CL(T_V)v_1 \neq v_1$ $CL(T_V)v_2$ in general. When this symmetry holds, however, we call v_1 and v_2 T_V-*equivalent* or T_V-*indistinguishable*, and we write v_2 $I(T_V)v_1$. In other words, $v_1 \in V$ is T_V-equivalent to $v_2 \in V$ iff T_V does not *separate* v_1 and v_2, i.e., iff $(\nexists O(v_1) \in T_V)$ $(\nexists O(v_2) \in T_V)$ $[(v_1 \notin O(v_2)) \wedge (v_2 \notin O(v_1))]$.

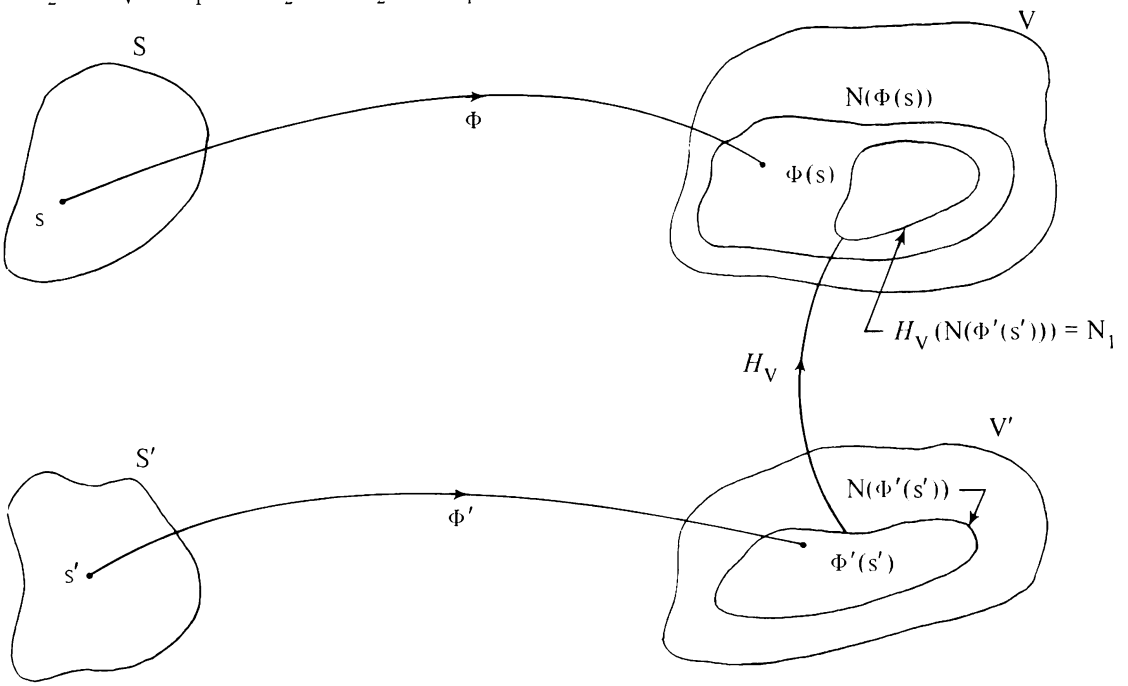

Fig. 1. The relationships between the terms of Definition 1.3.

† This definition differs from the original definitions in Refs. 1 and 2 in order to conform more closely with accepted terminology. O(v) is any open set containing v.

Let us now briefly describe the substance of our definition. We are given two objects s and s' and the modeler's viewpoint $\langle C, H_V \rangle$, where $C = \langle \Phi, \Phi', T_V, T_{V'} \rangle$. Then s' is a model of s relative to $\langle C, H_V \rangle$, and we write s' m(C, H_V)s if H_V interprets the actual performance $\Phi'_{act}(s')$ of s' as closely as T_V demands, i.e., if $H_V(\Phi'_{act}(s'))$ is *close* to $\Phi(s)$. Given that $\Phi'_{act}(s')$ is known no better than $T_{V'}$ allows — i.e., all the modeler knows is that his measurement or calculation $\Phi'_{meas}(s')$ is $T_{V'}$-*close* to $\Phi'_{act}(s')$ — we say that s' m(C, H_V)s if $H_V(v')CL(T_V)\Phi(s)$ for every $v' \in V' \ni v' CL(T_{V'})\Phi(s')$. We now state this formally. Let $\mathcal{N}_T(p)$ be the T-neighborhood system of some point p, and $O_T(p)$ a T-open set containing p. (When the topology T is clear from the context, we do not use it explicitly and write $\mathcal{N}(p)$ or $O(p)$.)

Definition 1.3: Given two sets S and S', a criterion $C = \langle \Phi, \Phi', T_V, T_{V'} \rangle$, an interpretation H_V, and two objects $s \in S$ and $s' \in S'$. Then s' *is a model of* s *relative to* C *and under* H_V if, and only if, the inverse image under H_V of every neighborhood $N(\Phi(s))$ contains a neighborhood $N(\Phi'(s'))$ of $\Phi'(s')$. A succinct statement of Definition 1.3 is: s' m(C, H_V)s iff $\forall N(\Phi(s)) \in \mathcal{N}_{T_V}(\Phi(s)) \; \exists N(\Phi'(s')) \in \mathcal{N}_{T_V}(\Phi'(s')) \ni N(\Phi'(s')) \subset H_V^{-1}(N(\Phi(s)))$.

The fundamental idea behind Definition 1.3 is that we want the important characteristics $\Phi(s)$ and $\Phi'(s')$ to agree "within T_V" under a suitable transformation H_V. The ideal situation is where $H_V \circ \Phi'(s') = \Phi(s)$. Given that certain errors and approximations † are considered, we only care to (or are limited to) "end up" in every neighborhood of $\Phi(s)$; however, exact agreement (the equality above) is not required. Furthermore, we cannot in general make explicit use of $\Phi'(s')$ or $\Phi(s)$, since they are usually not known. All we can use are points which are at best indistinguishable from $\Phi'(s')$ and $\Phi(s)$.

In most practical situations, the modeler often cares to get only ϵ-close to the true value $\Phi(s)$ of $s \in S$, where $\epsilon > 0$ is some fixed number. The following definition is intended to encompass these important cases.

Definition 1.4: Let T_V and $T_{V'}$ be generated by metrics d_V and $d_{V'}$, respectively, $\epsilon > 0$ be a fixed number, and *ceteris paribus*. Then s' is an ϵ-model of s relative to‡ C under $H_{V'}$, s' m(C, H_V, ϵ)s, iff $\exists \delta > 0 \ni N(\Phi'(s'), \delta) \subset H_V^{-1}(N(\Phi(s), \epsilon))$.

In reference to simulation models, we shall employ the notion of an indexed set.

Definition 1.5: An *indexed set (indexed class)*, represented by $S_X = \{s_\chi \in S: \; x \in X\}$ is a map $\psi: X \to S$ from a given set X (the *index set*) into another set S (the set of *indexed elements*). Since S_X is thus a map, the expression $S_X(x) = s_\chi$ makes sense. We shall also need the concept of a structure and a substructure.††

Definition 1.6: A *function structure*‡‡ is an object $FS = \{A, B; F\}$ with $F = \langle \{f_t: \; t \in T\}, \alpha \rangle$, where $\{f_t: \; t \in T\}$ is a family of $\alpha(t)$-ary functions from A to B, where $\alpha: \; T \to N_0^+$ (the nonnegative integers) is the *arity function* of F. The function α determines the "arity" of elements from F, and

---

† We have only considered deterministic errors and approximations thus far. Future work will extend the present framework to include various types of statistical errors and approximations.

‡ When metrics d_V and $d_{V'}$ generate T_V and $T_{V'}$, we shall also write $C = \langle \Phi, \Phi', d_V, d_{V'} \rangle$.

†† Here our interpretation of the terms "structure" and "substructure" is in basic agreement with the meanings used in set theory and mathematical logic,[3] and in general topology.[4]

‡‡ The term "function structure" is also used for relations or a combination of functions and relations, in which case it is simply called a relational structure or just a *structure*, respectively.

$\forall t \in T$, f_t: $\prod_{u=1}^{\alpha(t)} A_u \to B$, $A_u = A$. We thus consider only functions with at most countable arity, and when α is one-one it can be omitted, and we shall also write $F = \{f_t : t \in T\}$ in that case, where each f_t is a t-ary function from A to B.

We shall also need the concept of a *substructure*

Definition 1.7: A function structure $(FS)' = \{A', B', F'\}$ where $F' = \langle \{f'_{t'} : t' \in T'\}, \alpha' \rangle$, is a *substructure* of a function structure $FS = \{A, B, \langle \{f_t : t \in T\}, \alpha \rangle\}$ iff

1. $A' \subset A$ and $B' \subset B$,
2. $T' \subset T$ and α agrees with α' on T',
3. $\forall t' \in T'$, $a_1, a_2, \ldots, a_{\alpha'(t')} \in A'$,
   $f'_{t'}(a_1, a_2, \ldots, a_{\alpha'(t')}) = f_{t'}(a_1, a_2, \ldots, a_{\alpha'(t')})$.

2. Experimental Objects

In this section we define the class of objects which will be the subject of experimentation as defined in this chapter. These *experimental objects* — also referred to as *experimental units* in statistics[5] — are treated here as mathematical entities, but they actually represent "physically real" objects such as computers or their programs, chemical plants, economies, etc. This distinction is here irrelevant, however, as we pointed out in Ref. 1 when models were discussed.

Whereas in Chapter 1 the sets of prototypes S and model candidates S' represent arbitrary objects, we shall assume for the remainder of this work that experimental objects have a particular structure, as the following definition indicates.

Definition 2.1: An *experimental object* (Fig. 2) is an indexed class $S_X = \{s_x : x \in X\}$ whose elements are *relational structures* $s_X = \{\mathcal{U}, \mathcal{Y}, \{r_X\}\}$ (see Ref. 1). The sets $\mathcal{U}$ and $\mathcal{Y}$ are called the *input* and *output* sets, respectively, the relation $r_X \subset \mathcal{U} \times \mathcal{Y}$ the input-output relation of s_X, and X is a set underlying a topological space $\mathcal{X} = (X, T_X)$, the *parameter space*.

In many cases, for instance when S_X consists of dynamical systems for which the concept of state is defined, the elements s_x will be function structures and r_x will be a function $\lambda_x : \mathcal{U} \to \mathcal{Y}$. In these cases each input therefore produces only one output and r_x is a function λ_x, and this is what we shall assume for the remainder of this work.

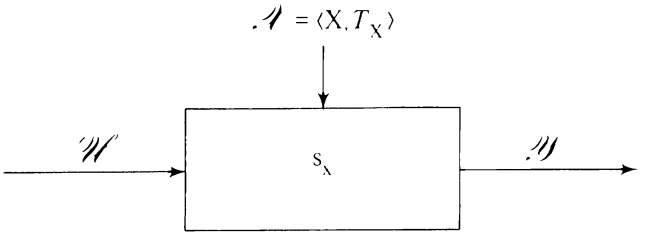

Fig. 2. A general experimental object.

In the context of statistical experiments, particular interpretations are attached to the sets $\mathscr{U}$ and $\mathscr{Y}$ and the space $\mathscr{A}$, as follows.

The set X represents structural variables, environmental conditions, internal parameters, policy levels (see example below), starting states, etc., and is quite arbitrary in the sense that it may contain scalars, vectors, functions, matrices, and so on. The particular structure of $\mathscr{A}$ as a topological space was discussed in Ref. 1.

The output space $\mathscr{Y}$ is the Borel space $\Pi_{\beta_T} = \langle \prod_{t\in T} R_t, \beta_T \rangle$, T some index set,† and is called the *observation space* (see Ref. 2 for the definition of β_T as a cross product $\prod_{t\in T} \beta_t$). This space consists of the observable *outcomes* $\prod_{t\in T} R_t$ – the set of all real functions on T – and the observable events β_T.[6]

The input space $\mathscr{U}$ is called the *driving space* and represents the conditions under which the object is "driven" with random numbers, sequences, or trajectories *(trace-driven)*, in accordance with some probability law P. This is particularly the case in simulation experiments where $\mathscr{U}$ is a set of numbers called *seeds* which are used to initiate a sequence of random numbers via some generating scheme such as the Lehmer scheme, for instance. Accordingly, we represent the input (driving) space $\mathscr{U}$ as a probability space $P = \langle \Omega, \mathscr{A}, \mathrm{P} \rangle$, where Ω is a set of random numbers, seeds, or traces. If the model structure itself is random, this randomness is often isolated as an environmental input and P would then also characterize this randomness. Note, however, that this space is fixed relative to S_X in the sense that each $s_x \in S_X$ is subjected to the same input conditions $\mathscr{U}$ or P. We thus obtain the structure of Fig. 3, where $\forall\, x \in X$, $s_x = \{\Omega, \mathscr{Y}; \{\lambda_x\}\}$, a function structure.

In a given experimental situation, a statistical experimental object S_X can be viewed as producing *outcomes* in $R_T = \prod_{t\in T} R_t$, called *samples*, in response to stimuli such as seeds, random numbers, and traces selected from Ω in accordance with the probability law P. If P represents all internal and exogenous randomness of the object, and we take some $x \in X$, then the response of s_x to any stimulus $\omega \in \Omega$ is a sample‡ (realization, time series) $\lambda_x(\omega) = \mathscr{Y}_{T,x}(\omega) = \langle y_t : t \in T \rangle \in R_T$ of a stochastic process $\mathscr{Y}_{T,x}$ with index set T, which we call the *output process* of s_x. The output of S_X is thus a family†† $\mathscr{Y}_{T,X} = \{\mathscr{Y}_{T,x} : x \in X\}$ of stochastic processes.

Similarly, if a fixed $\omega \in \Omega$ is considered, $\mathscr{Y}_{T,X}(\omega) = \{\mathscr{Y}_{T,x}(\omega) : x \in X\}$ is a family of realizations or time series, one for each $\mathscr{Y}_{T,x} \in \mathscr{Y}_{T,X}$. If each $\mathscr{Y}_{T,x}$ is ergodic, this family contains all the statistical information of the collection $\mathscr{Y}_{T,X}$ (see Ref. 2).

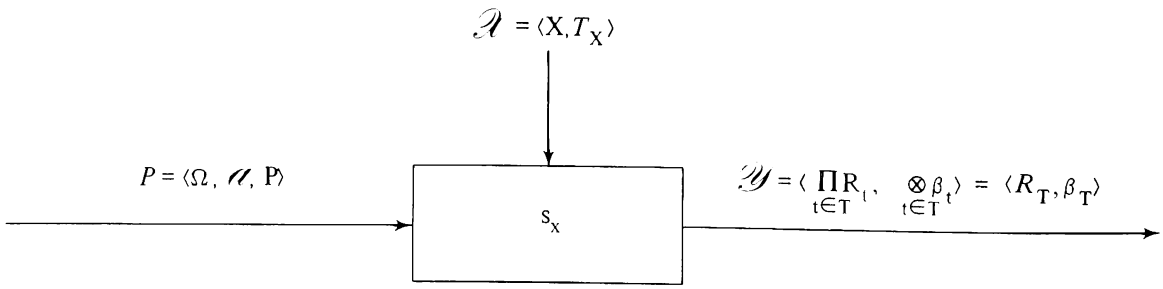

Fig. 3. A statistical experimental object.

†T will usually denote the time scale $[0, \infty)$.

‡Also denoted by $y_x(\omega, \cdot)$ to reflect the index $x \in X$ and the dependence of the sample upon ω, which is fixed in this case, and $t(\omega)$, which is variable (consider $y_x(\cdot, t)$ similarly).

††Also written as $\mathscr{Y}_X$ when the index set T is evident.

In conclusion, the set Ψ of experimental objects in an inquiry $I = \langle \Psi, Q, M, D \rangle$ constitutes a family of indexed sets $S_X = \{s_x \in S: x \in X\}$, where S is some fixed set of *indexed elements* (Definition 1.5) and X is a set underlying the topological parameter space $\mathcal{T} = \langle X, T_X \rangle$. The essential way in which these indexed sets differ is in their indexing scheme (see Definition 1.5, where S_X is defined as a map $\psi: X \to S$), and their elements s_x are function structures $\{\Omega, \mathcal{Y}; \{\lambda_x\}\}$ (Definition 1.6).

Remark 2.1: We have treated the general case where each *encounter* $(x, \omega, y_x(\omega, \cdot))$ with the experimental object produces a time series $y_x(\omega, \cdot)$. In many cases, the output of the system S_X is simply a family of random variables $Z_X = \{Z_x: x \in X\}$. The observation space is then simply $\Pi_\beta = \langle R, \beta \rangle$, and each encounter $(x, \omega, y_x(\omega))$ then produces a sample $y_x(\omega)$ of the random variable Z_x.

Remark 2.2: Sometimes the boundaries of the objects considered for experimentation are not easily defined. If these objects have random internal structures, or are exposed to environmental conditions which cannot be isolated in the form of an input space *P*, then this space is often not included, as in Fig. 4. In this case, we still assume that each s_x produces an output process defined on a fixed, but not necessarily specified, probability space $P = \langle \Omega, \mathcal{A}, P \rangle$.

Example 2.1: A Chemical Production Problem. Consider the production of chemicals by a process which is sensitive to random environmental conditions such as temperature. Let the factor space $\mathcal{X}$ represent the amount of catalyst introduced in the process, and let $\mathcal{Y}_X$ be the yield of the process. Then this physical situation is described by the experimental object of Fig. 4. Next consider the simulation of the above problem where the environmental conditions are modeled with a random-number generator *P*. Then Fig. 3 illustrates the experimental object ("model").

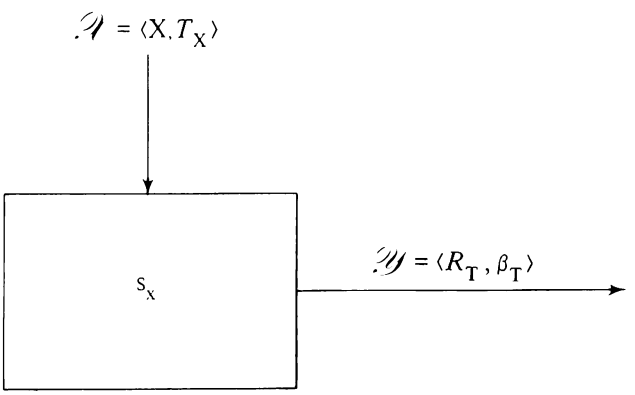

Fig. 4. An autonomous experimental object.

3. The Question Set Q

Let us now define what we shall mean by a "question" (Section 3.1) and when questions are well posed in the present context (Section 3.2), and then present some examples (Section 3.3).

3.1. DEFINING A QUESTION

Consistent with the linguistic definition[†] of Harrah,[7] we define a question as follows.

Definition 3.1: A *question* is a mapping $q: S \to Z$, where S is the *domain of the question (about which* the question is asked), and Z is the *response set*.

An element $z \in Z$ is called a *response to* q and a pair $(s_1, z_1) \in S \times Z$ is called an *answer to* q, which is a *correct answer to* q if $q(s_1) = z_1$.

Example 3.1: Consider a set S of random variables and let it be asked, "what is the expected value of $X \in S$?" Then $q: S \to R^1$ is the expected value operator $E(\cdot)$, and (X,a) is the correct answer to q if $E(X) = a$.

3.2. QUESTIONS IN EXPERIMENTAL INQUIRIES

In the context of experimental inquiries, a more specific interpretation is attached to questions, as the following definition shows (refer to Ref. 1, where experimental objects and the modeling of indexed classes are discussed).

Definition 3.2: Consider a set Ψ of experimental objects S_X whose elements are from some set S, and a parameter space $\mathcal{X} = \langle X, T_X \rangle$. Then an experimental question (Fig. 5) is a function $q: \Psi_q \to Z$, the *response set*, where $\Psi_q = \{S_X/_{X_q} : S_X \in \Psi, X_q \subset X\}$ and S_X is treated as a map from X to S. We shall denote the elements of Ψ_q by S_q. Thus a question usually does not "ask about" all parameter levels X but rather about a chosen subset $X_q \subset X$.

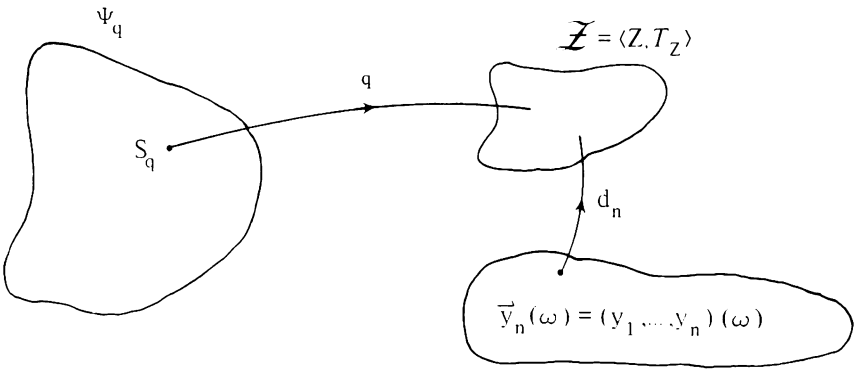

Fig. 5. Asking general questions.

[†] Where a question is a statement of a particular form: the disjunction of all its possible answers.

Example 3.2: Consider the chemical production problem of Example 2.1. Let the analyst be interested in the sensitivity of yield to changes in factor level x about a given operating level $x_0 \in X$. If T_X is generated by a metric d on X, he may wish to vary x over a finite subset X_0 of $N(x_0, \epsilon)$, an ϵ-neighborhood of x_0. He would then phrase his experimental question q about the class of objects $\Psi_q = \{S_q: S_q = S_X/_{X_q}, S_X \in \Psi, X_q = X_0\}$.

Our approach to experimentation is similar to that used in response-surface studies[5,8] where, in response to a question q: $\Psi_q \to Z$, experiments are usually conducted in stages, as follows.

First, "local" questions $q_0: S \to Z_0$ are asked about individual elements $S_q(x)$ of $S_q \in \Psi_q$, $x \in X_q$. Thus a *local* question is one where the parameter level x is held fixed. Then a *collective* question $q_1: \Psi_q \to Z_1$ is asked about Ψ_q that "collects" the responses $q_0(S_q(x))$, $x \in X_q$, into a response $q_1(S_q) = \{(x, q_0 \circ S_q(x)): x \in X_q\}$ for $S_q \in \Psi_q$. Finally, a *global* question $q_2: Z_1 \to Z$ is asked about the collective responses $q_1(S_q)$. Clearly, when such a decomposition of q into component questions q_0, q_1, q_2 is possible, the combined result of asking the component questions must be the same as asking q directly. Hence $q = q_2 \circ q_1$, where $q_1: \Psi_q \to \prod_{x \in X_q} (Z_0)_x = Z_1$ is defined by $(\forall S_q \in \Psi_q)(q_1(S_q) = \{(x, q_0 \circ S_q(x)): x \in X_q \subset X)\}$, and $q_0: S \to Z_0$ is a (local) question about the elements S.

Example 3.3: A Global Question about Indexed Classes. Consider Example 3.1. Let Ψ consist of indexed sets S_X of random variables Y_x, $x \in X$. Thus $(\forall x \in X)(S_X(x) = Y_x)$. Let q be, "What is the value $x^* \in X_q$ of the parameter at which the expected value $E(Y_x)$ is a maximum?" Then $q_0(s_x) = q_0 \circ S_q(x) = E(Y_x)$, $q_1(S_q) = \{(x, E(Y_x): x \in X_q\}$, and q_2 is sup $\{E(Y_x): x \in X_q\}$.

Remark 3.1: Example 3.1 may be discussed as a parameter-estimation problem (to simplify the discussion, a controlled parameter x is not included), where S is an indexed set of distributions $F_\Theta = \{F_\theta: \theta \in \Theta\}$ and q: $F_\Theta \to \theta$ is the parameter operator, in this case the expected value operator $E(\cdot)$, i.e., $q(F_\theta) = \theta$.

3.3. RELATIONSHIP BETWEEN QUESTIONS AND THE PERFORMANCES OF CHAPTER 1

We shall make no *mathematical* distinction between questions and performances. Thus, exactly as in Chapter 1 (see also Ref. 1), q may be considered as a function from Ψ_q to a topological space $\mathcal{Z} = \langle Z, T_Z \rangle$, where T_Z represents the various inaccuracies associated with q, as in the case of performances.

In this topological situation, an answer $(S_q, z) \in (\Psi_q \times Z)$ to q is correct (T_Z-correct) if z is T_Z-close to $q(S_q)$, i.e., $z\ CL(T_Z)q(S_q)$ (see Definition 1.2). If the decomposition q into $q_2 \circ q_1$ is used, q_1 may be considered as a function from Ψ_q to the space $\mathcal{Z}_1 = \langle Z_1, T_{Z_1} \rangle$, where $Z_1 = \prod_{x \in X_q} (Z_0)_x$, and T_{Z_1} is the Tychonoff product $\prod_{x \in X_q} (T_X \times T_{Z_0})_x$, and q_0 is a map from[†] S to a topological space $\mathcal{Z}_0 = \langle Z_0, T_{Z_0} \rangle$. Next, the "global" question (performance) q_2 is a mapping from $\mathcal{Z}_1$ to the space $\mathcal{Z} = \langle Z, T_Z \rangle$. Regardless of the accuracy of q_1, the inaccuracies embodied in T_Z are partially determined by T_{Z_1}, and we assume consistency in T_Z in the sense that $q_2^{-1}(T_Z) \subset T_{Z_1}$, i.e., q_2 is continuous.

[†] Each $S_q \in \Psi_q$ is of the form $\{s_x \in S: x \in X_q\}$.

To conclude, in an experimental inquiry $I = \langle \Psi, Q, M, D \rangle$, the set of questions Q *about* Ψ consists of questions q about Ψ_q. Thus $Q = \{q_j : \Psi_{q_j} \to \langle Z, T_Z \rangle, j \in J$ some index set , where $(\forall j \in J)$ $(\Psi_{q_j} = \{S_{q_j} : S_{q_j} = S_X/_{X_j}, S_X \in \Psi, X_j \subset X\})$.

4. The Experimental Design *M*

In a statistical inquiry, the essential purpose of an experimental object is to produce samples of the output processes. The purpose of an *experimental design* or *sampling plan M* is to state precisely how this sampling is to be accomplished by specifying how many samples (not necessarily independent) are to be collected for each controlled parameter setting, and in what order. There are two essential sampling methods: *fixed sampling* and *sequential sampling*. With the former method, the number of samples to be taken at each parameter setting is specified in advance of experimentation. With the latter, every time a new observation is made, the sampling plan specifies whether sampling is to be continued or stopped with that observation. If *M* is part of a *sequential decision procedure* SEQ = $\langle M, D \rangle$, where *D* is a decision strategy (see Chapter 5), *M* would then also specify at which observation times a decision is to be made. Clearly, fixed sampling is a special case of sequential sampling.

In the present context of experimental inquiries, sequential sampling in turn can be executed in two distinct ways: *ordered sampling*, where the parameter settings are stated in advance, and *mixed sampling*, where the next parameter setting depends upon observations already taken. In the mixed case, therefore, the design may require that additional samples be taken for another parameter value at which observations were already taken earlier in the experiment.

In this chapter, we shall first present the general (mixed) case. Then an important special case will be discussed, and we conclude with an example.

4.1. DEFINING A GENERAL SAMPLING PLAN *M*

Consider an experimental object, such as in Fig. 3, whose output for a parameter level sequence $\langle x_1, x_2, x_3, \ldots \rangle$ is a process $(Y_1(x_1), Y_2(x_2), \ldots)$, let $\Sigma_X = \sigma(T_X)$, the Borel† sets on X generated by T_X, and let $F(Y_1, Y_2, \ldots) = \sigma(\bigcup Y_i^{-1}(\beta))$, the σ-algebra generated by $\bigcup Y_i^{-1}(\beta)$), then we define M as follows.

Definition 4.1: A *sampling plan* or *experimental design* $M(\Psi)$ for a class of experimental objects Ψ is a measure-valued function μ on the probability space $\langle \Omega, \ , P \rangle$ with the following properties:

(1) $\forall \omega \in \Omega, \mu(\omega) \in M(X, \Sigma_X)$, the measures on $\langle X, \Sigma_X \rangle$.

(2) $\{\omega : \mu(\omega) = \upsilon \in M(X, \Sigma_X)\} \in F(Y_1(x_1), Y_2(x_2), \ldots, Y_k(x_k), \ldots, Y_{\upsilon(X)}(x_{\upsilon(X)}))$, where‡

$\#\{Y_k(x_k) : x_k \in A\} = \upsilon(A), A \in \Sigma_X$.

---

†These are *generalized* Borel sets[9,11] and are not necessarily the same as the standard Borel sets β on the real line R^1.

‡#(A) is the cardinality of a set A.

The interpretation of μ as a generalized stopping variable[12] is evident. Indeed, let X be a singleton $\{x_0\}$ and let $M(X, \Sigma_X) = M_N(X, \Sigma_X)$, the interger-valued measures on $\langle X, \Sigma_X \rangle$. Thus each $\upsilon_N \in M_N(X, \Sigma_X)$ satisfies the condition $(\forall A \in \Sigma_X)(\upsilon_N(A) \in N)$, where N are the natural numbers. Since $X = \{x_0\}$ in this case, we only need to consider the event $\{x_0\}$, and each υ_N is specified by its value at $\{x_0\}$ and equals an integer n_{υ_N} there.

Expressing this special case in our framework, $M(\Psi)$ is a random variable μ on $\langle \Omega, \mathcal{A}, P \rangle$ with the following properties:

(a) $\forall \omega \in \Omega, \mu(\omega) \in N$

(b) $\{\omega: \mu(\omega) = m \in N\} \in F(Y_1(x_0), Y_2(x_0), \ldots, Y_m(x_0))$, where $\#\{Y_k(x_0): x_0 = x_0\} = m$, obviously.

Clearly, μ is a stopping variable (time).[12] Our generalization is thus an extension of this stopping concept to the case where, instead of having a single process $\mathcal{Y}_{x_0} = (Y_1(x_0), \ldots, Y_k(x_0), \ldots)$, we have a family of indexed processes $\mathcal{Y}_X = \{\mathcal{Y}_x: x \in X\}$. In the general case, therefore, for any $A \in \Sigma_X$ and any $\omega \in \Omega$, $\mu(\omega)(A)$ specifies how many samples are to be taken at a parameter level in A, and $\mu(\omega)(X)$ is the total number of samples taken for that particular $\omega \in \Omega$.[13]

An important special case in practice is where the parameter levels constitute an at most countable partition $PAR(X) = \{x_i: i \in I\}$ of X, the number of samples taken at some level $x_k \in X$ are independent of the number taken at any other level $x_j \neq x_k$, and the experiment is executed by sampling sequentially over all the values of $\{x_i: i \in I\} = X$ in accordance with a particular order $x_1, x_2, x_3, \ldots$. In this case, the output process is $Y_1(x_1), Y_2(x_1), \ldots, Y_{k_1}(x_1)$, $Y_1(x_2), \ldots, Y_{k_1}(x_1)$, and condition (2) of Definition 4.1 may be rewritten as

(2') $(\forall x_j \in X)[\{\omega: \mu(\omega)(x_j) = k_j\} \in F(Y_1(x_j), Y_2(x_j), \ldots, Y_{k_j}(x_j))]$.

Now that we have stated what is meant by a design $M(\Psi)$ for a class Ψ of objects, let us bring this in the context of an experimental inquiry I. Considering a set of questions Q for the inquiry I, we now define a sampling plan or experimental design M for the inquiry.

Definition 4.2: A *sampling plan* or *experimental design* M for an inquiry I with question set Q is a set of designs $M = \{M(\Psi_q): q \in Q\}$, where each $M(\Psi_q)$ is a design for Ψ_q (i.e., to answer q).

Expressing this definition in terms of generalized stopping variables, $M = \{\mu_q: \Omega \to M(X_q, \Sigma_{X_q})\}$.

4.2. EXAMPLE

Although the merits of sequential sampling techniques are most apparent in a decision theoretic setting where losses and risks are considered, let us cite a simple example of a scheme which illustrates some aspects of sequential sampling and which is used in simulation practice where the modeler intends to sample until he believes that system transients have decayed below a certain level.

Consider the system of Fig. 3, where the analyst is faced with a class Ψ of experimental objects S_X whose elements s_x generate an output process $\mathcal{Y}_x$, and the assumptions underlying part (2') of Definition 4.2 are satisfied. If the analyst is interested in the expected value of $\mathcal{Y}_x$ and believes that these processes are stationary in the mean beyond a certain finite but unknown "transient time" τ, then he might consider the following sampling scheme.

For each $x_j \in X$, stop sampling at $\bar{Y}_{k_j}(x_j)$ if $\text{Max}\{|\bar{Y}_{k_j}(x_j) - \bar{Y}_{k_j - n_j}(x_j)|: n = 1,2,\ldots,n_j < k\} < \epsilon > 0$, where $\bar{Y}_m(x_j) = \frac{1}{m}\left(\sum_{i=1}^{m} Y_i(x_j)\right)$, and n_j is some fixed number for each x_j.

In other words, the sampling process at level x_j stops at k_j when the last n_j estimates $\bar{Y}_m(x_j)$ for the mean $E(\mathcal{Y}_{x_j})$ do not fluctuate about $\bar{Y}_{k_j}(x_j)$ by more than $\epsilon > 0$. Of course, if he has some estimate $n_\tau = \hat{\tau}$ for τ, the analyst may want to reject all samples up to n_τ by subtracting $\bar{Y}_n(x_j)$ from all $\bar{Y}_m(x_j)$, $m > n_\tau$.

5. The Decision Strategy *D*

Now that the experimenter has stated his questions Q about Ψ and he has an experimental sampling plan *M*, he must consider how he will answer or estimate the answers to Q by executing experiments† upon elements of Ψ in accordance with the design *M*. These experiments are usually finite and rarely provide the correct answer (see Definition 3.1). He must therefore "estimate" or "decide" what the correct answers are from the (finite) body of collected data.

In this section we show that this process of "guessing" the correct answer can be satisfactorily discussed within the framework of statistical decision theory.[14-16] Specifically, we first discuss what is meant by a *decision strategy* (Section 5.1); then we show in Sections 5-2 through 5-4 that the techniques of sequential hypothesis testing, parameter estimation, and confidence sets can be expressed in this framework very naturally. Finally, we define what is meant by a decision strategy *D for an inquiry* I (Section 5.5), and we present an example in Section 5.6.

5.1. DECISION STRATEGIES

A *sequential statistical decision problem*[17] is a 4-tuple $\langle \mathcal{X}, \mathbf{A}, L, \mathcal{Y} \rangle$ where $\mathcal{X}$ and **A** are topological spaces $\langle X, T_X \rangle$ and $\langle A, T_A \rangle$ called the *parameter space* (states of nature) and the *action space*, respectively.[18] The function $L: X \times A \to R^1$ is called the *loss function*, and $\mathcal{Y}$ is a stochastic process on a probability space $P = \langle \Omega, \mathcal{A}, P \rangle$, called the *observation process*.

Coupled with such a decision problem is a *sequential decision strategy* $D = \{d_n: R^n \to A, n = 1,2,\ldots\}$, where each d_n is a measurable mapping called a *decision rule*, and a family of representation probability spaces ‡ $\{\langle R_N, \beta_N, P_{\mathcal{L},x}\rangle: x \in X\}$ which contains the space induced by $\mathcal{Y}$ and *P*. When X is a set of "possible states of nature," it is the purpose of the decision strategy to determine or estimate which $x \in X$ is the true "state of nature," as follows.

The statistician observes a sample $(y_1, y_2, \ldots, y_n) \in R^n$ from the random variables $(Y_1, Y_2, \ldots, Y_n)$ of $\mathcal{Y}$, then he makes the decision $d_n(y_1, \ldots, y_n) \in A$ and incurs a *loss* $L(x, d_n(y_1, y_2, \ldots, y_n))$ where $x \in X$ is the (usually unknown) state of nature. Since $(Y_1, Y_2, \ldots, Y_n)$ are random variables, $d_n(Y_1, Y_2, \ldots Y_n)$ is also a random variable. Consequently, the experimenter's objective is to select or compute d_n such that the expected value of the loss $E_x[L(x, d_n(Y_1, \ldots, Y_n))]$, called the *risk function* $R(x, d_n)$ of d_n, is "minimized" in some sense, obtaining a function $R^*(x)$ of x.

In the context of experimental inquiries as we view them, this decision-theoretic framework applies in a slightly modified form. The parameter set X includes two types of parameters: state

†See Section 6.1 for a precise definition of experiments.
‡N are the natural numbers (also see Section 4.1).

of nature Θ and controlled parameters X_c (the familiar index set). This is expressed by setting X equal to the cross-product $X_c \times \Theta$. A typical decision problem is then simply to estimate the true state of nature for experimental objects, for a set of parameter levels $X_q \subset X_c$, in response to a question q and consistent with an experimental design $M(\Psi_q)$ (see Chapter 4).

In this report, we do not address ourselves to the general decision-theoretic issues related to the selection of decision rules and loss functions. Rather, we shall limit the discussion of inquiries to the case where sequential decision strategies have been determined and will leave the issues associated with the selection of such strategies in the background. In addition, we shall assume that the decision strategy D consists of fixed† and nonrandomized‡ (simple) decision rules.

Let us now show that the important sequential statistical techniques can be expressed in this framework quite naturally. In this discussion the controlled parameter set X_c is not included because it is not important to the validity of the arguments and would merely complicate the discussion. We thus assume that $X = \Theta$.

5.2. TESTS OF HYPOTHESES

Consider a family of spaces $P_\Theta = \{\langle \Omega, \mathcal{A}, P_\theta \rangle: \theta \in \Theta\}$ and a random variable $Y: \Omega \to R^N$. Take $P_{\theta_0} \in P_\Theta$, where θ_0 is unknown, and set up the hypothesis (*null* hypothesis) H_0 that $\theta_0 \in A \subset \Theta$ against another hypothesis (the *alternative*) H_1 that $\theta_0 \in A^{-1} = \Theta - A$. Then a *simple test of* H_0 is a 2-tuple $T = \langle T_n, \beta_n \rangle$, n fixed, where $T_n: R^n \to R^1$ and $B_n \subset R^1$, and H_0 is rejected on the *(critical)* region $C = \{\omega: T_n \circ Y(\omega) \in B_n\} \subset \Omega$ or its representation region $C_n = \{\vec{y} = (y_1, \ldots y_n) \in R^n: T_n(\vec{y}) \in B_n\} \subset R^n$.

Under the same conditions as above, a *sequential* test of H_0 is a sequence $T = \langle (T_n, B_n): n = 1,2,\ldots \rangle$ where, $\forall n = 1,2,\ldots, T_n: R^n \to R^1$, $B_n \subset R^1$ and H_0 is rejected on the regions $\{\vec{y} = (y_1, \ldots, y_n) \in R^n: T_n(\vec{y}) \in B_n, n = 1,2,\ldots\}$ and H_0 is accepted on $\{\vec{y} = (y_1, \ldots, y_n) \in R^n: T_n(\vec{y}) \in B'_n = B_n^{-1}, n = 1,2,\ldots\}$.

As a decision problem, T defines a strategy $D = \{d_n: n = 1,2,\ldots\}$ as follows: $\forall n \in N$, $d_n(y_1, \ldots, y_n) = a_1 \in A$ iff $T_n(y_1, \ldots, y_n) \in B_n$, and $\forall n \in N$, $d_n(y_1, \ldots y_n) = a_0 \in A$ iff $T_n(y_1, \ldots, y_n) \in B'_n$.

If an *indifference region* is defined, i.e., $B_n \cup B'_n \neq R^1$, then the sampling rule (Chapter 4) would call for sampling to continue until a sample falls in some B_m or B'_m, $m \in N$.

5.3. ESTIMATION PROBLEMS

A *sequential estimation strategy (statistic)* is a family of statistics $E = \{T_n: R^n \to \langle \Theta, T_\Theta \rangle, n = 1,2,3,\ldots\}$, where $\langle \Theta, T_\Theta \rangle$ is a topological *parameter space* and each T_n is a function of n observation random variables. Clearly, by setting $T_n = d_n$, $\Theta = A$ (the action space), we obtain the familiar decision-theoretic setting.

5.4. CONFIDENCE SETS AND INTERVALS

Let $P_\Theta = \{\langle \Omega, \mathcal{A}, P_\theta \rangle: \theta \in \Theta\}$ be a family of probability spaces, each specified by a particular value of $\theta \in \Theta$. Let $Y^{(n)}$ be a random variable from Ω to R^n, and T_n a function from R^n to 2^Θ. Then $T_n(y_1, \ldots, y_n) = B \subset \Theta$ is a $(1 - \alpha)$-*confidence set estimator for* $\theta \in \Theta$ iff $(\forall \theta \in \Theta)$ $(P_\theta \{\omega: \theta \in T_n \circ Y^{(n)}(\omega) = B(\omega) \subset \Theta\} = 1 - \alpha)$. Then a confidence set test CT for θ is a statistic $CT = T_n: R^n \to 2^\Theta$ such that $P_{Y^{(n)}\Theta} \{(y, \ldots, y_n): \theta \in T_n(y_1, \ldots, y_n)\} = k$, where $P_{Y^{(n)}\Theta}$ is the

†The rules are not selected at random.
‡A randomized rule produces probability measures on the action space A.

probability induced upon R^n by $Y^{(n)}$ and P_θ. Finally, a sequential confidence set test SCT for θ is a family of confidence tests $SCT = \{I_n : R^n \to 2^{(\Theta)}, n = 1,2,3,...\}$. The interpretation of SCT as a sequential decision structure is now clear.

Remark 5.1: Usually, when $\Theta = R$ for instance, estimators I_n are induced by simpler vector functions $\begin{bmatrix} T_{n,1} \\ T_{n,2} \end{bmatrix} : R^n \to \Theta \times \Theta$ in an obvious way by defining $I_n(y_1,...,y_n) = [T_{n,1}(y_1,y_2,...,y_n), T_{n,2}(y_1,...,y_n)] \subset \Theta = R$. In this more familiar case, the confidence sets are called *confidence intervals*.

5.5. THE SEQUENTIAL DECISION STRATEGY D_1

In this section, we discuss how a decision strategy D is expressed as a decision strategy D_1 for the inquiry I.

Recall that a decision rule is a function from the sample space R^n to the action space $\langle A, T_A \rangle$, and that a question q is a mapping from $^{\dagger}\Psi_{X_q}$ to $\langle Z, T_Z \rangle$. These decision functions d_n are usually considered as "estimators"[16] of the parameter (state of nature) $\theta \in \Theta$, in which case the sets A and Θ are assumed to be the same.

Similarly, in an inquiry as formulated above, the purpose of actions is to inform the statistician about the correct answers to q. These actions must therefore by interpreted in terms of possible responses $z \in Z$ to q. Accordingly we assume that $A = Z$. We shall also not distinguish between the topologies T_A and T_Z (see Chapter 1). Thus $\langle A, T_A \rangle = \langle Z, T_Z \rangle$.

Note that, since a sampling plan M is also specified, there is no need to make decisions when M requires that sampling continue. Therefore, decision rules d_n need not really be defined on all of R^n, $n = 1,2,...$. However, it is usually simpler and more convenient to make the decision strategy D independent of the sampling plan M.[17,19] In fact, this choice is often mandatory because the experimenter may not know any of the experimental outcomes in advance. We can now define a decision strategy D.

Definition 5.1: Given a set of experimental objects Ψ, a set of questions $Q = \{q_j : \Psi_q \to \langle Z, T_Z \rangle, j \in J\}$.

Then a *decision strategy to answer* $q \in Q$ is a set of decision rules $D_q = \{d_{q,n} : R^n \to \langle Z, T_Z \rangle\}$. A *decision strategy for the inquiry* I is then a *decision strategy to answer* Q, and is a family of strategies $D_Q = \{D_q : q \in Q\}$.

Remark 5.2: Corresponding to the decomposition of a question q into $q_2 \cdot q_1$ (Section 3.2), a strategy can be structured similarly. The strategy D_q then has two components: a "local" component $D_{q,1}$ and a "global" component $D_{q,2}$.

The local component is defined by $D_{q,1} = \{d_{n,x} : R^n \to X \times Z_0, x \in X_q, n = 1,2,...\}$, where each $d_{n,x}$ is a decision rule associated with parameter level $x \in X_q$ and is defined by $d_{n,x}(y_1,...,y_n) = (x,z_0) \in X \times Z_0$. The rules $d_{n,x}$ estimate the local performance[‡] of experimental objects S_q at some fixed parameter setting $x \in X_q$.

[†] A set of indexed classes with parameter set X_q.
[‡] Recall that every question is also a performance function (Section 3.3).

The global component is employed at the (random) conclusion of the experiment and is then simply a fixed decision rule $D_{q,2} = \{d_{q,2}: \prod_{x \in X_q} (X \times Z_0)_x \to Z\}$, which operates on the results of the decisions $d_{n,x}$. Corresponding to the special case of Section 4.1, when decisions $d_{n,x}$ are the same for each level x, then $D_{q,1}$ is determined by the strategy $D = \{d_n: R^n \to Z_0, n = 1,2,...\}$, and $d_{n,x}$ are defined by $d_{n,x}(y_1,...,y_n) = (x, d_n(y_1,...,y_n))$.

5.6. EXAMPLE: A GENERAL REGRESSION EXPERIMENT

Consider a family Ψ of experimental objects such as in Fig. 4, where each $s_x \in S_X$ generates a random variable Y_x. For any $S_X \in \Psi$, let us be interested in the response surface $RS = \{(x, E(Y_x)): x \in X\}$. In this case, the question set Q_1 consists of a single question $\{q\}$ which can be decomposed into (see Section 3.2) $q = q_2 \circ q_1$, where q_2 is the identity mapping.

In order to develop a decision strategy $D_{\{q\}}$ to answer q, let us assume that RS is behaved as follows:

$E(Y_x) = \sum_{i=1}^{M} \theta_i f_i(x)$, where f_i, i = 1,...,M, is a given set of functions (usually continuous and linearly independent[13]), called the *regression functions* (on x), and θ_i are unknown. Then a decision-theoretic approach is to find "best" estimates $\hat{\theta}_i$, i = 1,2,...,M, i.e., those that minimize some risk function $R = E(L(E(Y_x), \sum_{i=1}^{M} \hat{\theta}_i f_i(x)))$.

Consider now the special case of *linear regression*, where $f_1(x) = 1$, $f_2(x) = x$, and $f_i(x) = 0$ for $i > 2$, and where the Y_x are normally distributed and have a common variance σ^2. Let us estimate the answer to q by minimizing the mean of the *mean-squared-error* loss function $L(E(Y_x), \hat{E}(Y_x)) = [(E(Y_x) - \theta_2 \bar{x}) - (\hat{\theta}_1 + \hat{\theta}_2(x - \bar{x}))]^2$, where $\bar{x}$ is the arithmetic mean of all the parameter levels in X and was introduced to simplify later expressions.

Next let a (non-sequential) experiment (see Chapter 6) $\langle \langle (x_j, y_{x_{j,i}}): i \in I_j \rangle: j \in J \rangle$ be executed, where $\langle x_j, y_{x_{j,i}}: i \in I_j \rangle$ is a local experiment of length[†] $\#(I_j)$ at parameter level x_j, $y_{x_{j,i}}$ is the *ith* observation at x_j, and $\#(J)$ is the number of parameter levels used. Let $\{I_j: j \in J\}$ be fixed before experimentation. This determines a fixed design *M*, and the experiment is of length[‡] $\prod_{j \in J} \#(I_j) = n$.

Now consider the decision strategy *D* (independently from the design *M*). The best estimates for the above problem[14] are the following:

$$\hat{\theta}_1 = \sum_{j \in J} \sum_{i \in I_j} y_{x_{j,i}}, \text{ and } \hat{\theta}_2 = \frac{s_{x,y}}{s_x^2}, \text{ where } {}^{\ddagger} \; s_{x,y} = \frac{\sum_{j \in J} \sum_{i \in I_j} y_{x_{j,i}} (x_j - \bar{x})}{\prod_{j \in J} \#(I_j)},$$

$$s_x^2 = \frac{\sum_{j \in J} x_j (x_j - \bar{x})}{\sum_{j \in J} \#(I_j) \cdot x_j (x_j - \bar{x})}, \qquad \bar{x} = \frac{\sum_{j \in J} \#(I_j) \cdot x_j}{\prod_{j \in J} \#(I_j)}.$$

---

† $\#(\alpha)$ is the cardinality of a set α

‡ $\prod_{j \in J} \#(I_j)$ is the arithmetic product of the $\#(I_j)$.

Hence $\widehat{E(Y_x) - \theta_2 x} = \theta_1 + \theta_2(x - \bar{x}) = \hat{\theta}_1 + \hat{\theta}_2(x - \bar{x})$. This least-squares estimate is unbiased and has the smallest variance of all estimators consisting of linear combinations of the observations.[14] Finally, the decision functions (rules) d_n produce, in response to question q_1, actions $\sum_{i=1}^{M} \hat{\theta}_1 I_1(x) = \hat{\theta}_1 + \hat{\theta}_2 x \in Z_1$, i.e., $d_n((y_1, \ldots, y_n)_x) = d_{n,1}(y_1, \ldots, y_n) + d_{n,2}(y_1, \ldots, y_n)x$, where $d_{n,1} = \hat{\theta}_{1,n}$ and $d_{n,2} = \hat{\theta}_{2,n}$, and $\hat{\theta}_{1,n}$ and $\hat{\theta}_{2,n}$ are defined as above with $\prod_{j \in I} \neq (I_j) = n$.

6. Simulation Experiments

The purpose of this final chapter is essentially threefold. First, we shall discuss general experiments in the context of inquiries (Section 6.1). In Section 6.2, we shall bring to bear much of the axiomatic modeling framework developed in this paper and address one of our early objectives: the definition of simulation. Then, we relate our modeling theory of earlier chapters to simulation experiments by incorporating it into the more general framework of *simulation inquiries*.

Although much of this work is definitional and is intended to constitute a framework for further research, two theorems and one example are developed in the last section (Section 6.3) which illustrate some of the important modeling issues in simulation.

Expressed in terms of the modeling framework of Chapter 1, the first theorem shows that, under reasonable conditions, continuous functions on the performance spaces $V \supset \Phi(S)$ and $V' \supset \Phi'(S')$ preserve modeling relations. The second deals with the "information" conveyed by a sample taken from an object known or assumed to be a model and develops an expression for the probability that decisions made with the model be correct.

6.1. EXPERIMENTS IN INQUIRIES

Given an experimental object, an *experimental encounter* is a pair (x, y_x), where y_x is the *outcome (sample, observation)* of the encounter at parameter level x.

An *experiment* is a sequence of encounters $E = \langle (x, y_x)_i : i \in I, x \in X \rangle$, I some index set. When the parameter set X is ordered as $\{x_1, x_2, x_3, \ldots\}$, and sampling is performed in accordance with that order, E can be expressed as $E = \langle \langle (x_j, y_{x_{j,i}}) : i \in I_j \rangle : j \in J \rangle$, a sequence of "local" experiments at fixed levels $x_1, x_2, \ldots, x_J$, where $\neq(I_j)$ is the number of encounters at level x_j (recall our discussion of experimental designs).

Sometimes, in parameter estimation problems for instance, the parameter space $\mathcal{A}$ is $\otimes = \langle \Theta, T_\Theta \rangle$, a space of noncontrolled parameters.‡ Then our experiment is a set of observations $E = \langle y_i : i \in I \rangle$.

When the object is driven (Fig. 3) with inputs selected independently[5] from a set Ω underlying a probability space‡‡ $P = \langle \Omega, \mathcal{H}, P \rangle$ in accordance with a probability law P (when the object is driven with random numbers, for instance), then a *controlled encounter* is a triple $(x, \omega, y_x(\omega))$ where $y_x(\omega)$ is the outcome at level x due to input ω. Then a *controlled experiment* is a sequence of controlled encounters $E_P = \{(x, \omega, y_x(\omega))_i : i \in I\}$. The following definitions place these concepts in the perspective of an inquiry $I = \langle \Psi, Q, M, D \rangle$.

‡ The mixed case $\mathcal{A} = \mathcal{A}_c \times \Theta$ is also possible.

‡‡ Which also represents all the exogenous and endogenous random effects that affect the object.

Definition 6.1: An experiment $E = \langle (x,y_x)_i : i \in I, x \in X_q \subset X \rangle$ is *an experiment* E_q *to answer* $q \in Q$ iff $(\exists \omega \in \Omega)(\forall A \in \Sigma_{X_q})(\#\{(x,y_x) \in E : x \in A\} = \mu_q(\omega)(A))$. (See Definition 4.1)

The definition simply states that the experiment E_q must be executed consistently with the design $M(\Psi_q) \in M$ for $q \in Q$.

An *experiment for an inquiry* I is then defined as a union of experiments $E_q = \bigcup_{q \in Q} E_q$, and similarly for controlled experiments.

6.2. SIMULATION INQUIRIES

The groundwork has now been laid for describing what we mean by a simulation inquiry and for defining when an object simulates another.

A typical simulation problem can be informally described as follows. Consider an experimental object (prototype) S^*_X which is unknown except for the fact that it belongs to a prototype class Ψ, and a set of questions Q about Ψ. Then consider a model candidate $S^*_{X'}$ which is known to belong to Ψ', and a corresponding set Q' of questions about Ψ'. (See Chapter 1 for an analogous treatment of performances.) The simulation problem is then typically to obtain answers to Q from answers to Q'. In terms of inquiries, we approach this situation as follows.

First we require an inquiry $I' = \langle \Psi', Q', M', D' \rangle$ to answer the questions Q', and an interpretation $H_Z : Z' \to Z$ (see Definition 3.2 and Section 3.3) with which estimates to Q' can be interpreted as answers to Q. None of these answers may be correct, of course. Finally, in order that we obtain a *simulation* inquiry, we require that $S^*_{X'}$ be a model of S^*_X *relative to these questions*, as follows.

We assume two sets Ψ and Ψ' of experimental objects S_X and $S_{X'}$ with parameter spaces $\mathcal{X} = \langle X, T_X \rangle$ and $\mathcal{X}' = \langle X', T_{X'} \rangle$, respectively. Recall that a question set Q is a set $Q = \{q_j : \Psi_{q_j} \to \langle Z, T_Z \rangle, j \in J \text{ some index set}\}$, where $\Psi_{q_j} = \{S_{q_j} = S_X/_{X_j} : S_X \in \Psi, X_j \subset X\}$, and similarly for Q'.

Definition 6.2: Consider a 5-tuple $SI = \langle S^*_X, S_{X'}, Q, I', H_Z \rangle$, where

1. $S^*_X \in \Psi$ and $S^*_{X'} \in \Psi'$.
2. Q is a set of questions about Ψ.
3. I' is an inquiry $\langle \Psi', Q', M', D' \rangle$ (with questions Q' about Ψ').
4. $H_Z : Z' \to Z$ is an interpretation map.

Then SI is a *simulation inquiry* if $(\forall q \in Q)(S^*_{X'}\, m(C_q, H_Z) S^*_X)$, where $C_q = \langle q, q', T_Z, T_{Z'} \rangle$ is a modeling criterion (see Definition 1.1), and $q' \in Q'$ is the question corresponding to $q \in Q$.

The relationship $(\forall q \in Q)(S^*_{X'}\, m(C_q, H_Z) S^*_X)$, which is required to obtain a *simulation* inquiry, is precisely what we shall identify with a *simulation model*.

Definition 6.3: Consider two sets of experimental objects Ψ and Ψ', sets of questions Q and Q', and an interpretation H_Z, as in Definition 6.2, and two elements $S^*_X \in \Psi$ and $S^*_{X'} \in \Psi'$. Then $S^*_{X'}$ *simulates* S^*_X *relative to the criterion* $C_s = \langle Q, Q', T_Z, T_{Z'} \rangle$ *and interpretation* H_Z if $(\forall q \in Q)(S^*_X\, m(C_q, H_Z) S^*_X)$, and (C_s, H_Z), $S^*_{X'}$, and S^*_X are from a simulation inquiry SI. When these conditions are satisfied, $S^*_{X'}$ is called a *simulation model* of S^*_X for the *inquiry* SI, and we write $S^*_{X'}\, sim(C_s, H_Z) S^*_X$ or $S^*_{X'}\, sim(SI) S_X$. In any simulation inquiry SI, therefore, $S^*_{X'}\, sim(SI) S_X$.

The essential distinction between modeling and simulation is that simulation can take place *only* in an active experimental environment where it is part of a particular inquiry. Thus a *modeling*

relationship becomes a simulation relationship once the modeling criterion $C = \langle \Phi, \Phi', T_V, T_{V'} \rangle$ can be properly interpreted as part of a simulation inquiry, as in Definition 6.2. This represents precisely the position that *simulation is an experimental activity* whereas modeling is not, and disagrees with views held in specific areas such as automata theory, as we illustrated in Ref. 1, where automata simulation was treated purely as a modeling problem.

As a final comment, a simulation inquiry requires that $S^*_{X'}$ *simulate* S^*_X. This fact is of course rarely *known* in advance, since then there would often be no purpose in simulating at all. However, it is almost always assumed in practice, and it is the modeler's responsibility to formulate and and develop the object $S^*_{X'}$ with enough care, and possibly with some verification and validation,[5-8] in order that this assumption be reasonable.

6.3. THE PRESERVATION OF MODELING RELATIONS WITH CONTINUOUS FUNCTIONS AND THE PROBABILITY THAT MODEL DECISIONS ARE CORRECT

Recall that, when questions q can be decomposed into $q_2 \circ q_1$, as in Section 3.3, the global component q_2 of q was assumed continuous. The following theorem shows that, if $S_{X'}$ is a simulation model of S_X for the questions q_1 and[‡] q'_1 alone (and an interpretation H_{Z_1}), then when any continuous q_2 is composed with q_1 as a new question q (similarly for q'), a natural condition on H_Z: $Z' \to Z$ assures that $S_{X'}$ is also a model of S_X relative to q, q', and H_Z (see Fig. 6).

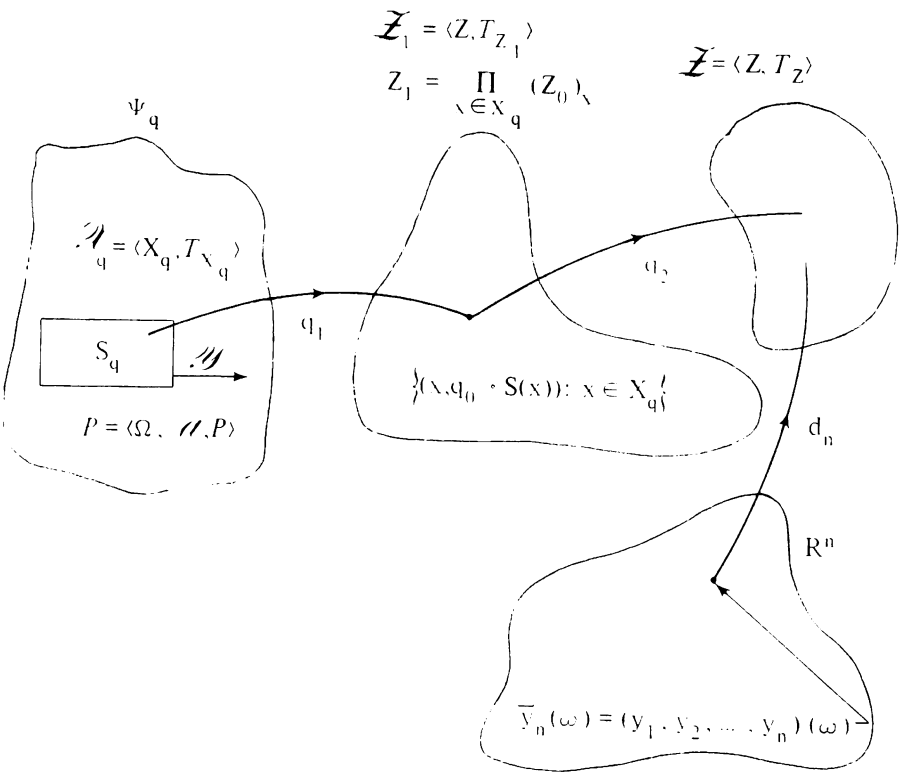

Fig. 6. Asking response-surface questions about experimental objects.

[‡] q'_1 is the model question corresponding to q_1.

Theorem 6.1

Consider two objects $S_X \in \Psi$ and $S_X' \in \Psi'$, two questions $q_1 : \Psi \to \mathscr{Y} = \langle V, T_V \rangle$ and $q_1' : \Psi' \to \mathscr{Y}' = \langle V', T_V' \rangle$, and an interpretation $H_V : V' \to V$. Then take two continuous questions $q_2 : V \to \mathscr{Z} = \langle Z, T_Z \rangle$ and $q'_2 : V' \to \mathscr{Z}' = \langle Z', T_Z' \rangle$ and another interpretation $H_Z : Z' \to Z$, and let $q = q_2 \circ q_1$ and $q' = q'_2 \circ q'_1$. Then, if $C_1 = \langle q_1, q'_1, T_V, T_V' \rangle$ and $C_2 = \langle q_2, q'_2, T_Z, T_Z' \rangle$, $(S_X', m(C_1, H_V) S_X) \wedge (q'_1(S_X') \; m(C_2, H_Z) H_V \circ q'_1(S_X')) \Rightarrow S_X' \; m(C_q, H_Z) S_X$, where $C_q = \langle q, q', T_Z, T_Z' \rangle$.

Proof: To simplify notation, denote S_X by s and S_X' by s'. $s' \; m(C_1, H_V) s \Rightarrow (\forall N(q(s)))$ $[H_V(q'_1(s')) \in q_2^{-1}(N(q(s)))]$, since q_2^{-1} maps each neighborhood N(q(s)) of q(s) to a neighborhood $N(q_1(s))$ of $q_1(s)$. But $H_V(q'_1(s')) \in q_2^{-1}(N(q(s))) \Rightarrow q_2 \circ H_V \circ q'_1(s') \in q_2 \circ q_2^{-1}(N(q(s)))$ $\subset N(q(s))$. Therefore, $q_2 \circ H_V \circ q'_1(s') \in N(q(s))$. Hence every neighborhood of q(s) is a neighborhood of $q_2 \circ H_V \circ q'_1(s')$, and $q'(s') \; m(C_2, H_Z) H_V \circ q_1'(s') \Rightarrow (\forall N(q_2 \circ H_V \circ q'_1(s')))$ $(N(q'(s')) \subset H_Z^{-1}(N(q_2 \circ H_V \circ q'_1(s')))) \Rightarrow (\forall N(q(s)))(\exists N(q'(s')) \subset H_Z^{-1}(N(q(s)))) \Rightarrow$ $s' \; m(C_q, H_Z) s$, as desired.

The following general conclusion can be drawn from Theorem 6.1. If $s' \; m(C_1, H_V) s$ and q_2, q'_2, and H_Z are chosen such that $(\forall v' \in V')$ $(v' \; m(C_2, H_Z) H_V(v'))$, then s' is a model of s for the overall (composed) questions $q = q_2 \circ q_1$ and $q' = q'_2 \circ q'_1$, whatever $s' \in S'$ and $s \in S$. Hence continuous functions q_2 and q_2' which are related in the above manner preserve modeling relations. This choice is reasonable because the purpose of H_V itself is to relate models to their prototypes. When q_2 and q'_2 are homomorphic† under (H_V, H_Z), for instance, this condition is automatically satisfied.

Example 6.1 The modeling of problems in utility theory is often accomplished in two parts (Figure 6.A): the simulation of the dynamic and stochastic structure generating the consequences or alternatives (the "technological" aspects of the problem), and the modeling of the utility structure.

If $U^o : V^o \to R$ is a utility function for basic consequences V^o, and the alternatives V are the lotteries (probability distributions) $PM(V^o)$ on V^o, it is reasonable[17] to define the utility function $U : PM(V^o) \to R$ by $(\forall P \in PM(V^o))(U(P) = E_p(U^o))$, where $E_p(f)$ is the expected value of f under P. When U^o is continuous, U is continuous in the weak topology[11,12], where convergence is equivalent to convergence in distribution (measure). If s' is an alternatives model of s, and q'(s') is a utility model of $H_V \circ q'(s')$, s' is a utility model of s (for $U^o q$ and $U' \circ q'$) and utilities obtained from the model can be used to correctly infer utilities of the modeled.

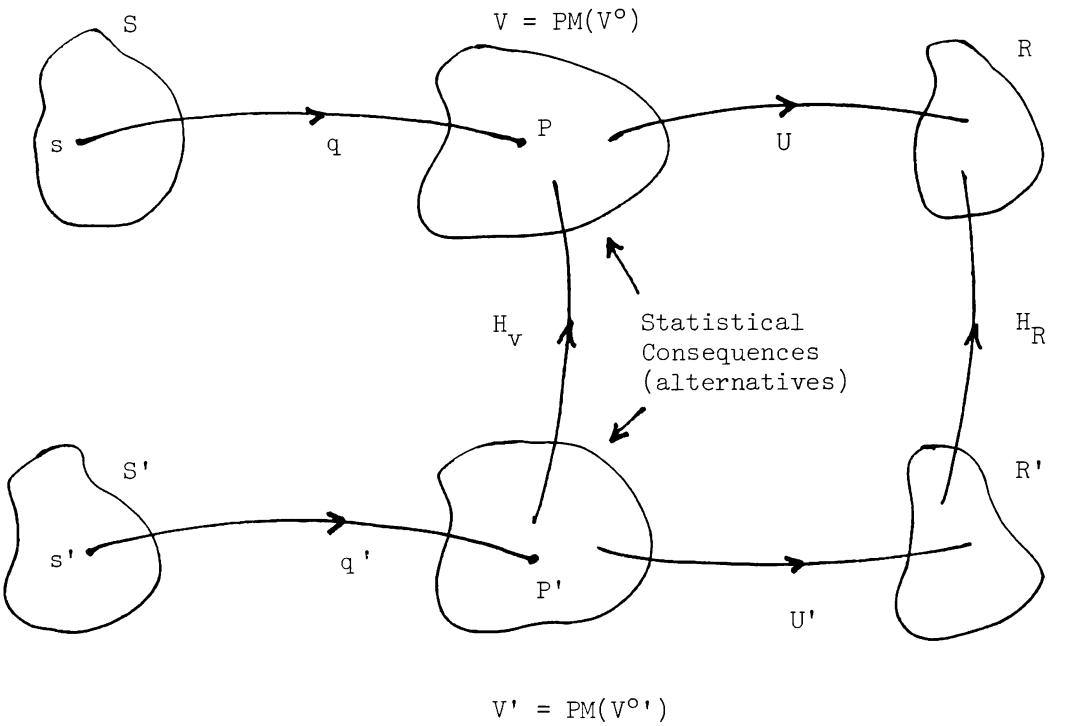

Figure 6.A. The Modeling of Problems in Utility Theory.

Let us now turn to another issue.

We have stated earlier that experimental objects are often known or assumed to be simulation models of some prototype. This assumption in itself is usually insufficient to answer any questions about the prototype, however, and that is of course why a simulation inquiry is contemplated. But what information does such a model convey about its prototype in the course of such an inquiry? When does the simulator stop sampling? What decision rules d'_n should be used?

One way to consider these "questions" is to ask another: what is the probability P_{q,d'_n} that any given decision $z'_n = d'_n(y'_1, \ldots, y'_n)$ is correct for a question q, i.e., provides the *correct* answer to q (after suitable interpretation)? The following theorem attacks precisely this question.

Theorem 6.2

Let $q(S^*_X) = z_q \in Z$, the correct answer to q about S^*_X, and $q(S^*_{X'}) = z'_{q'} \in Z'$. Let[‡] $S^*_{X'}$ $\mathrm{sim}(C_q, H_Z)S^*_X$ and[‡‡] $H_Z^{-1}(\mathcal{N}(z_q)) \cap \mathcal{N}(z'_{q'}) = \mathcal{N}(z'_{q'}) \cap H_Z^{-1}(\mathcal{N}_Z)$, and let there exist a neighborhood $N_0(z'_{q'}) \in \mathcal{N}(z'_{q'})$ which contains an at most countable number of $N(z'_{q'}) \in H_Z^{-1}(\mathcal{N}(z_q))$ with $P(z'_n \in N(z'_{q'})) > 0$. Then the probability P_{q,d'_n} that $z'_n = d'_n(y'_1, y'_2, \ldots, y'_n)$ is correct (*T*-correct) for q is $P_{q,d'_n} = \inf\{P(z'_n \in \mathcal{N}(z'_{q'})): N(z'_{q'}) \in H_Z^{-1}(\mathcal{N}_Z)\}$.

[†] $(\forall v' \in V')\, H_Z \circ q'_2(v') = q_2 \circ H_V(v')$.

[‡] $\mathcal{N}(z)$ are the neighborhoods of $z \in Z$ from T_Z.

[‡‡] This will always be the case when $S^*_{X'}$ $\mathrm{sim}(C_q, H_Z)S^*_X$, except in pathological situations which we shall not discuss here.

Proof: $P_{q,d'_n} = P[(\forall N(z_q) \in \mathcal{N}(z_q))(H_Z(z'_n) \in N(z_q))]$. Since $S^*_{X'}$ is a model of S^*_X, $P_{q,d'_n} = P[(\forall N(z_q) \in \mathcal{N}(z_q))(z'_n \in H_Z^{-1}(N(z_q)))] = P[(\forall N(z'_{q'}) \in H_Z^{-1}(\mathcal{N}(z_q)))(z'_n \in N(z'_{q'}))]$. This probability exists, since we assume $\exists N_0(z'_{q'}) \in \mathcal{N}(z'_{q'})$, which contains an at most countable number of $N(z'_{q'}) \in H_Z^{-1}(\mathcal{N}(z_q))$, with $P(z'_n \in N(z'_{q'})) > 0$. Thus $P_{q,d'_n} = P(z'_n \in \bigcap \{N(z'_{q'}) \in H_Z^{-1}(\mathcal{N}(z_q))\}) = \inf\{P(z'_n \in N(z'_{q'})): N(z'_{q'}) \in H_Z^{-1}(\mathcal{N}(z_q))\}$.

Since $H_Z^{-1}(\mathcal{N}(z_q)) \cap \mathcal{N}(z'_{q'}) = \mathcal{N}(z'_{q'}) \cap H_Z^{-1}(\mathcal{N}_Z)$, where $\mathcal{N}_Z$ are all the neighborhoods from T_Z, we obtain the desired result, and $P_{q,d_n} = \inf\{P(z'_n \in N(z'_{q'})): N(z'_{q'}) \in H_Z^{-1}(\mathcal{N}_Z)\}$.

This theorem indicates that the fact that $S^*_{X'}$ simulates S^*_X is very informative (for q) and allows the estimation of the statistical "quality" of d'_n exclusively in terms of neighborhoods of the model, and regardless of the true value of z_q. Clearly, these probabilities or their estimates can be used as stopping rules, as the following example suggests. Note, however, that the probability law of d'_n and the values of $z'_{q'}$ are usually not known in advance of experimentation; otherwise there would be no purpose in experimenting, since $H_Z(z'_{q'})$ is then *known* to be a correct answer to q. In general, therefore, P_{q,d'_n} can only be estimated as the experiment proceeds.

Example 6.2 Consider Example 3.3 and let the modeler be interested only in a fixed value $x_0 \in X$ of the parameter. Thus Ψ can be considered as a family of random variables Y_{x_0} (similarly for Ψ'), and q is the question, "What is the expected value $E(Y_{x_0}) = z_q$ of $Y_{x_0} \in \Psi$?" This corresponds to question q_0 in Example 3.3.

Let $\mathcal{Z}$ and $\mathcal{Z}'$ be metric spaces $\langle Z, T_d \rangle$ and $\langle Z', T_{d'} \rangle$. Let $Y'_{x'_0}$ be a model of Y_{x_0} in the sense that $H_Z(E(Y'_{x'_0})) = E(Y_{x_0})$ (Fig. 7). Let the simulator be interested in evaluating the probability

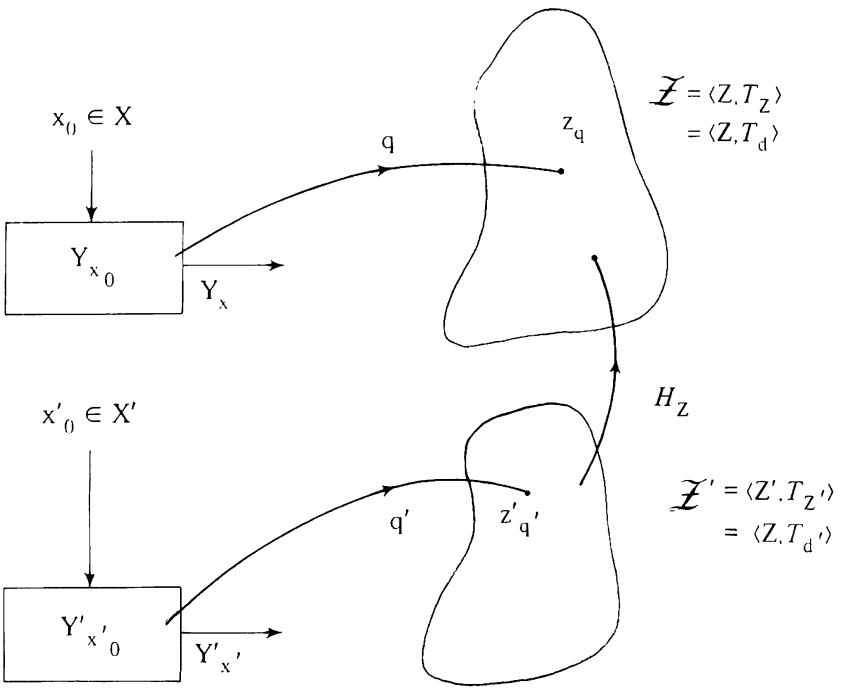

Fig. 7. The objects of Example 6.1.

$P_{q,d_n,\epsilon} = P(H_V(d'_n) \in N(E(Y_{x_0}), \epsilon)) = P(|H_V(d'_n) - E(Y_{x_0})| < \epsilon)$ that the decision $d'_n = \overline{Y}'_n = \frac{1}{n}\sum_{i=1}^{n} Y'_i$ be ϵ-correct for $q = E(\cdot)$, where $(Y'_1, Y'_2, \ldots, Y'_n)$ is a sample from Y_{x_0}, d'_n is unbiased for $E(Y'_{x'_0}) = z'_{q'}$.

Further assume that $\overline{Y}'_n$ has distribution † $NORM(E(Y'_{x'_0}), \frac{\sigma^2}{n})$ where σ^2 is the variance of $Y'_{x'_0}$ (Fig. 8). (This assumption may be based on the fact that the central limit theorem may be invoked, and the quality of this assumption depends on the degree of independence and the size of the sample.) Finally, assume that H_V^{-1} maps $N(E(Y_{x_0}), \epsilon)$ to a δ-neighborhood $N(E(Y'_{x'_0}), \delta)$, regardless of the actual value of $E(Y_{x_0})$ and $E(Y'_{x'_0})$ (as $\epsilon \to 0$ this assumption is arbitrarily accurate for standard metrics on R^1). Then $P(d'_n$ is ϵ-correct for $q) = P_{q,d'_n,\epsilon} = P(d'_n \in H_V^{-1}(N(z_q, \epsilon))) = P(d'_n \in N(z'_{q'}, \delta)) = P(d'_n \in [z'_{q'} - \delta, z'_{q'} + \delta]) =$

$$P((d'_n \le z'_{q'} + \delta) \wedge (d'_n \ge z'_{q'} - \delta)) = P(-\delta \le d'_n - z'_n \le \delta) = P\left(\frac{-\delta}{\sigma/\sqrt{n}} \le \frac{d'_n - z'_{q'}}{\sigma/\sqrt{n}} \le \frac{\delta}{\sigma/\sqrt{n}}\right) = \Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(\frac{-\delta\sqrt{n}}{\sigma}\right),$$

where Φ is the cumulative distribution function of NORM(0, 1).

So in this case, the analyst can compute $P_{q,d'_n,\epsilon}$ exactly, without statistical estimation. If this probability is used to determine a sampling plan, by specifying that sampling continue until $P_{q,d'_n,\epsilon} = \alpha$, for instance, a fixed sampling rule is obtained ("stop sampling at $n(\alpha)$"), and the decision strategy D equals $\{d'_{n(\alpha)}\}$.

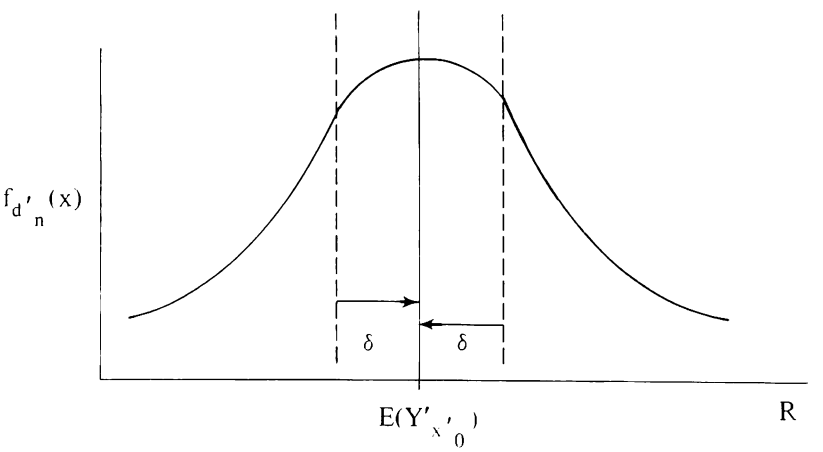

Fig. 8. The density function of the rule d'_n.

† NORM(a, b) is the normal distribution with mean a and variance b.

References

1. G.C. Corynen, *A Mathematical Theory of Modeling*, Lawrence Livermore Laboratory, Rept. UCRL-51726 (1975).
2. G.C. Corynen, *The Modeling of Stochastic Objects*, Lawrence Livermore Laboratory, Rept. UCRL-51760 (1975).
3. J.R. Shoenfield, *Mathematical Logic* (Addison-Wesley, Reading, Mass., 1967).
4. J. Dugundji, *Topology* (Allyn and Bacon, Boston, 1966).
5. G.A. Mihram, *Simulation: Statistical Foundations and Methodology* (Academic Press, New York, 1972).
6. A. Renyi, *Foundations of Probability* (Holden-Day, San Francisco, 1970).
7. D. Harrah, *Communication: A Logical Model* (M.I.T. Press, Cambridge, Mass., 1963).
8. T.H. Naylor, *Computer Simulation Experiments with Models of Economic Systems* (John Wiley & Sons, New York, 1971).
9. J. Neveu, *Mathematical Foundations of the Calculus of Probability* (Holden-Day, San Francisco, 1965).
10. J.W. Gibbs, *Elementary Principles in Statistical Mechanisms* (Yale University Press, New Haven, 1902).
11. H.L. Royden, *Real Analysis* (Macmillan, New York, 1968).
12. L. Breiman, *Probability* (Addison-Wesley, Reading, Mass., 1968).
13. C.L. Atwood, "Optimal and Efficient Designs of Experiments," *Ann. Math. Stat. 40* (5), 1570 (1969).
14. B.W. Lindgren, *Statistical Theory* (Macmillan, New York, 1968).
15. T.S. Ferguson, *Mathematical Statistics* (Academic Press, New York, 1967).
16. A. Wald, *Statistical Decision Functions* (Chelsea, New York, 1971).
17. M.H. DeGroot, *Optimal Statistical Decisions* (McGraw-Hill, New York, 1971).
18. J.S. Rustagi, *Optimizing Methods in Statistics* (Academic Press, New York, 1971).
19. S.S. Wilks, *Mathematical Statistics* (John Wiley & Sons, New York, 1962).