A DECISION THEORETIC APPROACH TO MODEL SELECTION FOR STRUCTURAL RELIABILITY

by

MIRCEA GRIGORIU

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
at the
Massachusetts Institute of Technology
January, 1976

Signature redacted

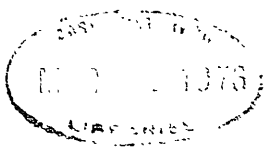
Signature of Author . . . . . .
Department of Civil Engineering
January 28, 1976

Signature redacted

Certified by. . . . . .
Thesis Supervisor

Signature redacted

Accepted by . . . . . .
Chairman, Departmental Committee of Graduate
Students of the Department of Civil Engineering.



ABSTRACT

A DECISION THEORETIC APPROACH
TO MODEL SELECTION FOR
STRUCTURAL RELIABILITY.

by

MIRCEA GRIGORIU

Submitted to the Department of Civil Engineering on
January 28, 1976 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Consider a random sequence $\underline{X}(t)$, which is observed during a finite time interval; the aim of the study is to find "the best" model of $\underline{X}(t)$ (for instance, "the best" distribution of the yearly wind extremes) for structural safety applications.

Typically in reliability analysis, model selection is done in an inference context, the main purpose being the determination of that model from some small convenient set which "best" fits the data (the notion of "best" might be based on a maximum likelihood criterior or on a chi-square test, etc.).

In contrast, in this work, the model selection problem is solved in a Bayesian decision framework. The essential ingredient of the procedure is the so-called utility function which assigns a value to each "action-behavior of nature" pair (for example, the action might represent the selection of a model, say of the code strategy, or of a design level

for a (deterministic) engineering system for which $\underline{X}(t)$ is an input). A penalty for model complexity is also included in the decision formulation.

Details are given for the case of independent sequences (with normal, gamma and Poisson distributions), normal (and log-normal) autoregressive sequences and renewal sequences of gamma interarrival times with random magnitudes.

As a case study, the design of a deterministic two-state (failure/survival) system, loaded by the white univariate sequence X(t), extreme type I distributed with parameter values corresponding closely to the yearly wind speed extremes recorded in Boston for a period of 42 years, is extensively analysed.

Thesis Supervisor: Daniele Veneziano

Title: Asst. Prof. of Civil Engineering

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to Professor Daniele Veneziano, his thesis supervisor, and to Professor C. Allin Cornell, his research project supervisor, for continued advice and constant encouragement during the course of this work.

The author wishes also to thank Professor I.Rodriguez-Iturbe who gave him the opportunity to work on a Water Resource project, Professor E. H. Vanmarcke, Professor C. F. Daganzo and Dr. T. G. Harmon for their valuable suggestions.

The work for this study was supported by National Science Foundation (RANN) Grant No. GI44060, "Structural Loads Analysis and Specification". The support of these sources is gratefully acknowledged.

TABLE OF CONTENTS

	Page
Title page	1
Abstract	2
Acknowledgements	4
Table of Contents	5
List of Figures	9
List of Symbols	13
CHAPTER I. INTRODUCTION	14
I.1. STATE OF UNCERTAINTY. METAMODEL	18
I.2. HYPOTHESIS TESTING. LITERATURE REVIEW	21
I.3. RESULTS AND ORGANIZATION	30
FIGURES	36
CHAPTER II. ESSENTIALS AND GENERALITIES	37
II.1. STATISTICAL DECISION PROBLEM	39
II.2. GENERALIZED BAYESIAN DECISION THEORY	46
II.3. CONCLUSIONS	63
FIGURES	67
CHAPTER III. INDEPENDENT MODELS	71
III.1. INTRODUCTION	71
III.2. DECISION STRATEGY IN INDEPENDENT SEQUENCES	73
III.2.1. Introduction	73

Page

III.2.2. Prior to Posterior Analysis; Predicion Analysis 77

III.2.2.1. Gaussian Sequence 77

III.2.2.2. Gamma Sequence 83

III.2.2.3. Poisson Sequence 85

III.2.2.4. Posterior Probabilities of Models 86

III.2.3. Loss Function 90

III.3. APPLICATIONS 92

III.3.1. Example 1: Optimal design of a structure loaded by an extreme type I white stationary sequence, with parameter values corresponding closely to Boston yearly wind speed extremes 93

III.3.2. Example 2: On prediction of extreme wind speeds from parent distribution. Optimal design of a structure for wind speeds of long return periods. 108

III.4 CONCLUSIONS 114

FIGURES 116

CHAPTER IV. AUTOREGRESSIVE MODELS 127

IV.1. INTRODUCTION 127

IV.2. MARKOV SEQUENCES 130

IV.2.1. Markov Sequences in Regression Form 130

IV.2.2. Error between two Markov Sequences 132

Page

IV.2.3. Sensitivity Equations in Prediction 135

IV.3. DECISION STRATEGY FOR MARKOV SEQUENCES 140

IV.3.1. Introduction 140

IV.3.2. Prior to Posterior Analysis; Prediction Analysis 141

IV.3.2.1. Posterior Distribution of the Unknown Parameters 142

IV.3.2.2. Simple Predictive Distribution 143

IV.3.2.3. Simultaneous Predictive Distribution 146

IV.3.2.4. Simultaneous Distribution. Probabilistic Solution 148

IV.3.2.5. Simultaneous Distribution. Statistical Solution 160

IV.3.2.6. Posterior Probability of Models 167

IV.3.3. Loss Functions. Decision Rules 171

IV.3.3.1. Optimal Modeling of Random Sequences with Application to the Design of a Structure and to the Operation of a Reservoir. 171

IV.3.3.2. Optimal Operation of an (Existing) System S. The case where S is a reservoir. 176

IV.3.3.3. Optimal Design of a System 178

Example 1: Design of a structural system against wind 178

Example 2: Design of a reservoir size 180

IV.4. CONCLUSIONS 183

Page

FIGURES 186

CHAPTER V. RENEWAL SEQUENCES WITH RANDOM MAGNITUDES 208

V.1. DECISION STRATEGY FOR RENEWAL SEQUENCES WITH RANDOM MAGNITUDES. THE CASE WHEN p IS KNOWN 214

V.1.1. Decision strategy for Renewal Sequences; p=0 215

V.1.1.1. Probabilistic Solution 215

V.1.1.2. Statistical Solution 223

V.1.2. Decision strategy for Renewal Sequences; p≠0 but known 233

V.2. DECISION STRATEGY FOR RENEWAL SEQUENCES WITH RANDOM MAGNITUDES. THE CASE WHEN p IS NOT KNOWN 236

V.3. CONCLUSIONS 241

FIGURES 243

CHAPTER VI. CONCLUSIONS AND SUGGESTIONS 248

REFERENCES 252

BIOGRAPHY 262

LIST OF FIGURES

		Page
CHAPTER I		
Figure I.1.	Reference design problem	36
Figure I.2.	Operation of a reservoir	36
CHAPTER II		
Figure II.1.	Metamodels of the input variable $\underline{X}(t)$	67
Figure II.2.	Hypothetical randomized and nonrandomized risk sets of A, when A and are finite.	68
Figure II.3.	Hypothetical randomized and nonrandomized risk sets of A and A_c, when A and are finite.	68
Figure II.4.	Typical expected cost functions $E[d/\eta]$, Eq. (II.19).	69
Figure II.5.	Hypothetical nonrandomized risk set of A_c when A and are finite. The case when the penalty for complexity is considered.	69
Figure II.6.	Hypothetical expected failure probability of S designed in accordance with a level III code.	70
Figure II.7.	Searching for the optimal metamodel	70
CHAPTER III		
Figure III.1.	Hypothetical realization of the independent sequence X(t)	116

Page

Figure III.2. Typical cost functions in Structural Reliability 116

Figure III.3. Expected cost E[d/η] in Eq. (II.19) for T=10, τ =1, c_F=2 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ 117

Figure III.4. Expected cost E[d/η] in Eq. (II.19) for T=40, τ=1, c_F=2 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ 118

Figure III.5. Expected cost E[d/η] in Eq. (II.19) for T=10, τ=1, c_F=20 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ 119

Figure III.6. Expected cost E[d/η] in Eq. (II.19) for T=40, τ=1, c_F=20 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ 120

Figure III.7. Expected cost E[d/η] in Eq. (II.19) for T=40, τ=50, c_F=2 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ 121

Figure III.8. Expected cost E[d/η] in Eq. (II.19) for T=40, τ=50, c_F=20 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ 122

Figure III.9. Randomized and nonrandomized risk sets corresponding to various code strategies. 123

Figure III.10. Empirical CDF 124

Figure III.11. Predicted and observed crossing rate 125

Figure III.12. Wind speed return period 126

CHAPTER IV

Figure IV.1. Calculation of the distribution of the extreme, $\tilde{y}_{m,\tau}$, and of the sum, $\tilde{y}_{s,\tau}$, variables 186

Figure IV.2. Time evolution of the marginal distribution 187

Page

of the state X(t) of a GMS. 187

Figure IV.3. Calculation of the mean rate of upcrossings for a GMS. 187

Figure IV.4. Mean rate of upcrossings of the zero mean unit variance first order GMS 188

Figure IV.5. Hypothetical conjugate prior and posterior PDF on the correlation coefficients of a GMS 189

Figure IV.6. Normalized expected error in Eq. (IV.104) versus RO1. 190

Figure IV.7. One step prediction error in Eq. (IV.109) versus RO1 199

Figure IV.8. Optimal operation of a system, S, excited by a partially known input, X(t). 206

Figure IV.9. Hypothetical penalty associated with the water release y. 206

Figure IV.10. Hypothetical realization of the wind speed process. 207

CHAPTER V

Figure V.1. Renewal point sequence, $X_0(t)$ 243

Figure V.2. Renewal wave sequence, $X_1(t)$. The case when $P(X_1=0) = p$ is zero 243

Figure V.3. Renewal wave sequence. The case when $P(X_1=0) = p$ is not zero 244

Page

Figure V.4. Observation vectors corresponding to the processes $X_q(t)$. The case when $p = 0$ 245

Figure V.5. Observation vectors corresponding to the process $X_1(t)$. The case when $p \neq 0$ and $X_1(t)$ is observable or unobservable 246

Figure V.6. Borges processes 247

LIST OF SYMBOLS

$\underline{X}(t)$	- input variable
E	- experiment
S	- system excited by $\underline{X}(t)$
τ	- life time of S
$f(\cdot)$	- optimal code strategy
$c(\cdot)$	- initial cost function
c_F	- failure cost
c_f	- complexity cost of $f(\cdot)$
P_f	- target failure probability
A	- space of pure strategies
A^*	- space of randomized strategies
A_c	- space of composite strategies
δ	- randomized strategy
η	- composite strategy
	- complete metamodel
*	- operational metamodel
(opt)	- optimal metamodel
EX1	- Extreme type 1 population
EX2	- Extreme type II population
G	- Gamma population
N	- Normal population
P	- Poisson population
GMS	- Gauss Markov sequence

CHAPTER I

INTRODUCTION

Consider the deterministic two-states (failure/survival) system S in Fig. I.1 loaded by the random process X(t). The system is designed to resist loads of intensity $x \leq d$.

The final goal of the present study is to optimize various engineering decisions; for instance, with reference to Fig.(I.1), the aim is to find the optimal (in some sense) design level $d^{(opt)}$.

For the majority of constructions the evaluations of $d^{(opt)}$ is done in two phases: first, the "code committee" establishes the specifications to be used in the design of any individual system and second, the "designer" calculates $d^{(opt)}$ in accordance with the code format. So-called Level III and IV code formats (Table (I.1)) to be defined more carefully below require that the code committee specify explicitly the probability distribution f(x) (in the simple case that X(t) = X is simply a random variable).

What has typically been done has been to choose f(x) in Table (I.1) as being the best probability distribution function (PDF) of X in the sense of statistical inference (for example, to maximize the likelihood).

Table (I.1)

Code Format	Code Specifications
Level III	criterion: P_f = target failure probability f(x) = code strategy
Level IV	criterion: Formal utility optimization f(x) = code strategy

Note: For a detailed version of this table see Sec. II.3 (Table (II.38)).

In the approach taken in this study f(x) is generally chosen from the set of all probability density functions. The proposed selection procedure is of decision theoretic type (see Sec. II.2). The goodness of any particular choice of f(x) is measured by the expected cost of the code, E[f], which equals the weighted sum

$$E[f] = \sum_S F(S)\, E_S[f] \qquad (I.2)$$

where, $E_S[f]$ is the cost of an optimal system S designed under the code strategy f(x) and F(S) denotes the relative frequency of systems of type S to which the code will be applied.

The detailed steps of the evaluation of the optimal action f(x) in the sense of minimizing (I.2) are presented in Sec. II.2.

For given code strategy, f(x), the design value of the system S in Fig. I.1 at level III, $d^{(opt)} = d_{III}$, is the P_f - upper fractile of f(x), i.e.

$$\int_{d_{III}}^{+\infty} f(x)dx = P_f \qquad (I.3)$$

while at level IV, $d^{(opt)} = d_{IV}$ is the result of the minimization of the expected cost of the design $x = d$, $E[d]$, being for example:

$$E_S[d] = C_S(d) + C_{F,S}[1 - F(d)] \qquad (I.4)$$

In Eq. (I.4), $C_S(d)$ and $C_{F,S}$ denote the initial cost of designing S at $x = d$ and the cost of failure of S, respectively, and $F(x)$ is the cumulative distribution function (CDF) of X.

The generalization of the Table (I.1) and the equations (I.2), (I.3), and (I.4) to the case of stochastic systems (say, with random resistance) or of a system with several modes of failure or of a vector-valued input load is conceptually (although not always numerically) straightforward (see [41], [61], [83]).

Three classes of engineering problems corresponding to the case when S is deterministic and the input variable, X(t), is univariate sequence are considered explicitly:

i) Optimal modeling of random sequences: The aim here is to find the model of X(t) which "best" fits the data (i.e. which minimizes a certain type of error). See detailed illustration in Sec. IV.3.3 for the case when X(t) is discrete time Gauss Markov process; see also the discussions in Chapter II and Chapter V.

II) Optimal operation of an existing system: Assume now that the system S is deterministic and that S has been physically realized; for instance, S might be an existing reservoir or dam. The design level d of S is here the capacity of the reservoir, X(t) denotes the amount of water

entering the reservoir and Y(t) is the water release per unit of time. The problem here is to find the output Y(t) which is optimum in some sense (detailed discussion in Sec. IV.3.3); see Fig.(1.2).

iii) Optimal design of a system: This problem constitutes one of the central goals of Structural Reliability. The aim here is to find the "best" (in utility sense) design level of S (see Fig. I.1) which in fact represents the optimal trade off, say, between safety and initial investment.

The optimal design of S is considered in detail in Chapter III, Chapter IV, and Chapter V.

The problems i), ii), and iii) represent several selected tasks in applied reliability which, however, cover a broad field of applications. Interestingly enough all these problems can be solved with the general decision framework developed in Chapter II. Moreover, it may be seen that most of the engineering problems involving optimization of various objective functions can be handled with the general development in this study (see Sec. II.2).

Sec. I.1 considers various design cases. The practical importance of statistical design is emphasized. Also the basic steps of the methodology developed in this study to evaluate f(x) in (I.1) are presented. Sec. I.2 contains a brief review of the literature on hypothesis testing; Sec. I.3 describes the essential results of this study and gives the organization of the whole thesis.

I.1 STATE OF UNCERTAINTY. METAMODEL

Consider again the design problem in Fig.(I.1)and let E denote an informative experiment (for instance E might consist of T measurements of the input sequence X(t); when $T < \infty$, X(t) is partially known and when $T \to \infty$, X(t) tends to be perfectly known).

Assume first that the probabilistic law of X(t) is perfectly known; then the design of S in Fig. (I.1) is called probabilistic design. In this case the inductive component of uncertainty (the uncertainty on the law of X(t)) is zero. The deductive component of uncertainty (uncertainty in the value of the state X(t)) is, instead, nonzero if X(t) is a random process.

When the probabilistic sequence X(t) with only partially known law is the input of S in Fig.(I.1), then the design of S is referred to as a statistical design.

Typically (see comments of Table (I.1)) the statistical design problem is solved in an inferential context: a family of models of X(t), $\mathcal{H}^{(I)} = \{H\}$, is postulated and then the "best" inferential model (say, the most probable model) is used in the design of S (Fig. (I.1)) as if it were the true model (see the way in which f(x) is used in Eqs. (I.3) and (I.4)).

This approach may lead to confusion especially when there are two or more models which have approximately equal probability of being true. For example, the sequence of yearly wind extremes is modeled in the U.S. code by extreme type II and in the Canadian code by extreme type I distributions.

In contrast, this study uses Bayesian Decision Theory to infer the best utilitarian (i.e. the best in the utility sense defined by Eq. (I.2)) characterization of X(t) (see Table (I.1)) corresponding, say, to the purpose of "best" designing S (Eq. (I.4)). The steps of the proposed procedure are the following:

First step: Let $\mathcal{H} = \{H\}$ be the complete (i.e. exhaustive) family of models corresponding to X(t); thus, the metamodel $\mathcal{H}$ of X(t) in Fig. (I.1) contains all the models of univariate sequences so that in general $\mathcal{H}$ is an infinite set. In the forgoing discussion any family (finite or infinite) of models corresponding to the input variable is also called metamodel or reference set of X(t).

For obvious computational reasons $\mathcal{H}$ is generally replaced by a finite reference set $\mathcal{H}^*, \mathcal{H}^* \subseteq \mathcal{H}$, containing computationally simple models. (For example, if X is a random variable, $\mathcal{H}^*$ might be the set of all univariate distributions available in the literature.)

Since $\mathcal{H}^*$ still may contain extremely many models, it too must be reduced. In this work the subset $\mathcal{H}^{(opt)}$ of $\mathcal{H}^*$, also called the optimal metamodel, is found; $\mathcal{H}^{(opt)}$ is the optimal strategy in the decision framework considered at metamodel level (Sec. II.2). This decision analysis recognizes explicitly the cost of complexity. Thus, the optimal reference set $\mathcal{H}^{(opt)}$ is a formal trade-off between accuracy and simplicity. The evaluation of $\mathcal{H}^{(opt)}$ is the responsibility of the code committee.

Second step: The optimal code strategy f(x) (see Table (I.1)) is selected by the code committee to minimize the expected cost in Eq. (I.2), where E[f] represents now the expected cost of f(x) conditional on $\mathcal{H}^{(opt)}$.

Third step: The design of a particular structural system, say the sys-

tem S in Fig. I.1 is obtained from Eq. (I.3) or Eq. (I.4), by the designer. The optimal code strategy, f(x), evaluated in this study for level III and level IV code formats can also be used to implement the Paloheimo's (level II) code proposal [61], which may be viewed as an approximate solution of the level III code format. The specifications of Paloheimo's code consist in a set of safety indeces, $\{\beta_i\}$, corresponding to each component, X_i, of the input variable $\underline{X}$. The evaluation of $\{\beta_i\}$ is done as follows: a target failure probability, P_f, is assumed and then the distributions $f_i(x_i)$ of X_i are inferred. The index of safety β_i satisfies:

$$\int_{m_i+\beta_i\sigma_i}^{+\infty} f_i(x_i)dx_i = P_f \qquad (I.5)$$

where m_i, and σ_i are the mean and the standard deviation of X_i, respectively.

We propose to replace the inferential safety factors $\{\beta_i\}$ by the utilitarian safety indeces obtained from Eq. (I.5) where $f_i(x_i)$ denotes the optimal code strategy for X_i.

Section I.2 starts with a simple example aiming at a better understanding of the difference between the inferential and the utilitarian strategies. Then, a critical review of various results in hypothesis testing which are of interest in the model selection problem is reported. Also, a deductive scheme (see computational details in Sec. II.2) with feedback is proposed to evaluate the optimal metamodel, $\mathcal{H}^{(opt)}$, of the input variable.

I.2 HYPOTHESIS TESTING. LITERATURE REVIEW

The choice between hypotheses can be done in an inferential or utilitarian manner. To stress the differences between these two methodologies, consider the following example [79]. Suppose we are to guess whether the queen is over 40 or not. The following table

Table (I.6)

		Actual Situations	
		Queen is over 40 (>40)	Queen is under 40 (≤40)
Gambler's Actions	Queen is over 40 (> 40)	\$1,000	- \$1,000,000
	Queen is under 40 (≤ 40)	\$0	\$1,000

gives the utility (utility = - cost or loss) of the gambler for all possible situations; thus, for instance when the gambler chooses the action (>40) but the actual state is (≤40) his tongue will be cut off (since, probably, the gambler values his tongue quite high, he assigns a very large negative utility to this event). Assume also that from the available evidence (informative experiment E) the probability of the states of nature (>40) and (≤40) are, respectively, P(>40) = 9/10 and P(≤40) = 1/10.

Then the expected cost of the actions that the gambler may consider are:

$$\text{-action (>40)} : \Sigma(\text{Prob})(\text{Utility}) = (\tfrac{9}{10})(1{,}000) + \tfrac{1}{10}(-1{,}000{,}000) =$$

$$- 99{,}100$$

$$\text{-action } (\leq 40) : \Sigma(\text{Prob})(\text{Utility}) = (\tfrac{9}{10})(0) + \tfrac{1}{10}(1{,}000) = +100 \qquad (I.7)$$

Therefore the best utilitarian action is (≤40) since it maximizes the expected utility (minimizes loss). By comparison the best inferential action in the sense of the most probable state of nature, is (>40). Whether inferential strategy depends only on the evidence, E, the utilitarian strategy tried to achieve the best compromise between purpose and evidence. (For instance if the gambler's tongue were already cut off, then in (I.6) - $1,000,000 should be replaced by $0, so that in this situation the action (>40) is the best utilitarian strategy.)

In this simple example, the space of the states of nature and the space of actions are identical, finite and with perfectly known elements (states of nature and actions). The same problem can be formulated in the context of hypothesis testing (say, H_1 = (>40) and H_2 = (≤40), where H_i denotes the i^{th} hypothesis). In general, the expected cost of the action H_j (see Zellner [92], p. 295, and Eq. (I.7)) is:

$$E[H_j] = \sum_{i=1}^{m} L(H_i, H_j)\, p(H_i) \qquad (I.8)$$

where, m is the number of hypotheses; in this case m = 2 · $L(H_i, H_j)$ is the loss (-utility) when action H_j is taken and the actual state of nature is H_i (for instance, in Table (I.6), $L(H_2, H_1)$ = +1,000,000 dollars) and $p(H_i)$ is the probability that the hypothesis H_i is true conditional on $\mathcal{H} = \{H_i\}_{i=1}^{m}$ being true and on E; for instance in (I.7) $p(H_1)$ = 9/10 and $p(H_2)$ = 1/10. (In

this case since $H_1 \cup H_2$ exhausts all the possiblities, the probabilities $p(H_i)$, i=1,2, are conditional only on E)

Then, the best (optimal) utilitarian action, $H_{opt}^{(u)} = H^{(opt)}$, satisfies

$$E[H_{opt}^{(u)}] = \min_{1 \le i \le m} \{E[H_i]\} \tag{I.9}$$

Interestingly enough, the best (optimal) inferential action, $H_{opt}^{(i)}$, or more precisely the most probable hypothesis, can be found from Eq. (I.9). In fact taking in Eq. (I.8), the loss function

$$L(H_i,H_j) = \begin{cases} 1 & , \quad i \neq j \\ 0 & , \quad i = j \end{cases} \tag{I.10}$$

and minimizing as in Eq. (I.9), it is found that, $H_{opt}^{(i)}$, has to satisfy:

$$p(H_{opt}^{(i)}) = \max_{1 \le i \le m} \{p(H_i)\} \tag{I.11}$$

that is the most probable model is the optimal utilitarian strategy for a particular binary utility. Moreover, if $L(H_i,H_i) << L(H_i,H_j)$ for any $j \neq i$ and $L(H_i,H_j) \cong L(H_k,H_\ell)$ for any $i \neq j$, $k \neq \ell$ the rule (I.11) still holds approximately.

Detailed presentation of the hypothesis testing problem in the context of the Decision Theory are found in Zellner [92], Ferguson [27], DeGroot [21], Raiffa and Schlaifer [66], etc. Although the literature dealing with the problem of hypothesis testing of decision type is very rich the application of this body of knowledge to model selection is almost nonexistent.

The problem of finding the optimal model, for a random sequence $\underline{X}(t)$

is generally solved inferentially (for instance, if $\mathcal{H} = \{H\}$ denotes various models of $\underline{X}(t)$ as plausible candidates, then the optimal model may be chosen as the most probable model; see Eq. (I.11)). Smoollwood [80] formulated the model selection problem (conditional on the reference set $\mathcal{H}$) as a (Bayesian) Decision problem; the weak points of his approach are the subjective choice of $\mathcal{H}$ and the fact that no penalty is associated with the model complexity.

The existing statistical literature dealing with the problem of hypothesis testing with application to model selection may be divided into three categories: modeling independent sequences, model for (Gauss) Markov Sequences, and evaluation of the best polynomial regression. First, consider the case when $\underline{X}(t)$ is independent sequence. Inferential best models of $\underline{X}(t)$ (say, the maximum likelihood model) can be found in any classical statistics book (see for instance Wilks [91]). Guttman [37] solved the Goodness-of-fit problem for $\underline{X}(t)$ by combining a Baysian with a classical sampling argument; in fact he used a "chi-square like" statistic, measuring the discrepancy between the observed frequencies and those predicted by the Bayesian distributions of the models considered in test.

Second, assume that $\underline{X}(t)$ is a discrete/continuous time Gauss Markov process generated by a linear dynamic system [30,76] and that the state $\underline{X}(t)$ is observed through a (linear) noisy mechanism. The problem of signal detection is extensively treated [75, 49, 50, 51].

Regarding the problem, of more direct interest here, i.e. finding "the best" representation of $\underline{X}(t)$, three routes are taken. Consider the discrete time case [77]. Assuming various Gauss Markov models $\mathcal{H} = \{H\}$, each with known parameters, of $\underline{X}(t)$, the posterior probability, p''_H of

$H \varepsilon \mathcal{H}$, is readily obtained from 1) the output of the Kalman-Bucy filter of $H \varepsilon \mathcal{H}$ (see [76]), 2) the informative experiment E (consisting of noisy observation of $\underline{X}(t)$) and 3) the prior probability, p'_H, of $H \varepsilon \mathcal{H}$. The essential disadvantage of this procedure is the limited capability of comparing models due to the computational cost of operating many Kalman-Bucy filters.

A different approach is given by Identification Theory. In this case a Markovian law is assumed for $\underline{X}(t)$ which depends on a set of free parameters, $\underline{p}$ (see Schweppe [76], Chapter 14, [3, 13, 39]). Some difficulties may arise since not all systems are globally identifiable; also, in general, identification leads to nonlinear filtering problems. (A system is said to be globally identifiable if for any $\underline{p}_1$ and $\underline{p}_2 \neq \underline{p}_1$ the outputs are different.)

Finally, under fairly loose conditions on the covariance function of the Gauss Markov process $\underline{X}(t)$ it is shown [44,89] that a linear system (also called realization) generating the process $\underline{X}(t)$ exists. Clearly, if the mean and covariance function of $\underline{X}(t)$ were perfectly known, then the minimum realization model of $\underline{X}(t)$ would be the best model (where minimum refers to the order of the Markov sequences modeling $\underline{X}(t)$). Although very attractive, the use of the Minimum Realization Theory is not advisable when little is known about the covariance function of $\underline{X}(t)$.

Third, consider the problem of finding the degree and the parameters of a polynomial regression. The inferential literature on this topic is quite broad; in Chapter 6 , Draper and Smith [23] consider the problem of selecting the "best" regression equation, Guttman [37] uses a Bayesian goodness-of-fit procedure to infer the best degree of a polynomial regression, Zellner [92] dedicates Chapters III, IV, and VIII to univariate/

multivariate regression models, Raiffa and Schlaifer [66] also derived posterior results corresponding to the regression equations (Chapter 13). Halpern [38,39] gives a Bayesian Decision formulation to the problem of finding the degree of a polynomial regression, $y = P(x)$, in which he allows for complexity. Clearly, Halpern's problem is a special hypothesis testing task, where the hypotheses have increasing degrees of generality in the sense that hypothesis H_δ (H_δ denotes polynomial regression of degree δ) coincides with $H_{\delta+1}$ when the coefficient of $x^{\delta+1}$ in $H_{\delta+1}$ is zero; because of this, the determination of the references set of regressions, $\mathcal{H}$, is not as important as it appears to be in a general model selection problem.

It follows that typically the decision regarding the optimal characterization of X(t) is done <u>deductively</u> (in a nonparametric sense), meaning that a set of models $\mathcal{H} = \{H\}$ of $\underline{X}(t)$ is postulated and then following Eq. (I.11)/Eq. (I.9) "the optimum" inferential/utilitarian model of $\underline{X}(t)$ is evaluated. The quality of the chosen metamodel $\mathcal{H}$ is never questioned.

So, Smallwood [80], speaking about the formulation of the metamodel says:

> "... This step of the process requires a great deal of creativity and insight on the part of the modular and is the part of the operation that we know the least about. ..."

Halpern [39] in his optimal polynomial regression describes the choice of the metamodel as follows:

> "... rather than assume that the model is known, append a priori probability, p'_δ, to the hypothesis that $\Delta = \delta$ for δ from 0 to δ_0. ..."

(where $\Delta = \delta$ denotes the hypothesis H_δ). If in Halpern's approach $\delta_o \to \infty$, then the optimal polynomial regression results; since for computational reasons δ_o is taken finite only a suboptimal regression polynom, $H^{(s\,opt)}$, can be evaluated. Because of the character of hypothesis in this problem ($..., \subseteq H_\delta \subseteq H_{\delta+1} \subseteq ...$) and of the capability of polynomial (with $\delta < \infty$) to represent various functions it seems that $H^{(s\,opt)}$ is a good approximation of $H^{(opt)}$ when δ_o is taken large enough (see details in the original paper [39]).

Schweppe [76], p. 425, considering the identification problem states:

> "... A basic principle of system identification is that you cannot identify a completely unknown system. In general, it is necessary to put some restrictions on the class of systems in which lies the system to be identified. Consider models of the following structures:"

and he assumes a law of $\underline{X}(t)$ (i.e. a linear dynamic system driven by white noise) with partially known parameters.

Zellner [92], p. 295, considering the "two-state-two-action" decision problem introduces the procedure by noting:

> "Further, we recognize that, by assumption, there are two possible states of the world - H_0 true or H_1 true",

and this list can be even substantially extended.

At this point it appears legitimate to ask ourselves how well one can choose the metamodel, $\mathcal{H}$, of $\underline{X}(t)$ or the related question (see Sec. II.2 where the largest metamodel of $\underline{X}(t)$ is considered) about the quality of man's prior probability assignment. The psychological literature related to the latter problem is very rich. Several conclusions in connec-

tion with the ability of humans as processors of information are presented next (see also the survey papers [40], [46]).

Hogarth [45] considers that human performance when facing complex problems is very low. In his view, man has limited information processing capacity, and consequently:

- man's perception of information is not comprehensive but selective,
- man makes much use of heuristic and cognitive simplification mechanism, and
- man processes information sequentially.

Hogarth [45] concluded (p. 273) that:

> "man is a selective, stepwise information processing system with limited capacity and he is ill-equipped for accessing subjective probability distributions. Furthermore, man frequently just ignores uncertainty."

In our context, Hogarth's result implies that the man may choose an "acceptable" (inferential) metamodel, $\mathcal{H}$ of $\underline{X}(t)$, only by chance with a probability which is very small if one considers the large number of possible metamodels of $\underline{X}(t)$.

However, the radicalism of Hogarth relative to the limitation of human beings is rejected (see [45]) by both Winkler and Edward. Besides other very interesting observations, Edward says:

> "psychologists in general seem to have rather strange ground rules about which aspects of human performance depend on human capacities and which do not. Roughly speaking the rules seem to prohibit external aids, to prohibit mental arithmetic...".

In fact, similar ideas of considering the man-machine couple were

the base of the Probabilistic Information-Processing System (PIP), [24], where, say, the man is responsible for the initial (inductive) step defining the set of hypotheses $\mathcal{H} = \{H\}$ and the machine performs the (deductive) step of evaluating the expected utility of various actions. The tendency to "conservatism"(people are unable to extract from data nearly as much certainty as is justified by the data in the light of Bayesian model) and to "radicalism" (opposite to conservatism when man has to decide between very many possibilities) of the human being was one of the justifications of PIP. (See also [53]).

The "limited" capability of human beings to perform complex tasks is a matter of definition (see for example Hograth and Edward conclusions in [45]) in the sense that the external aids may be accepted or not when considering the capacity of man to solve complex problems. In our view men are tool users so that the human performance depends on the capability of the machine; also the training is a significant factor in solving various simple/complex problems.

To evaluate the optimal (utilitarian) metamodel, $\mathcal{H}^{(opt)}$, of $\underline{X}(t)$, a feedback PIP-like procedure is proposed. The reason to evaluate $\mathcal{H}^{(opt)}$ instead of working with the complete set of models is only computational (see also Section (I.1)). For instance man may start from a pivotal metamodel $\mathcal{H}_1$ of $\underline{X}(t)$ and then conditional on $\mathcal{H}^*$ (where $\mathcal{H}^* \subseteq \mathcal{H}$ is finite and contains all the models of $\underline{X}(t)$ which are analytically simple) the expected cost (at code level) of choosing $\mathcal{H}_1$, $E[\mathcal{H}_1/\mathcal{H}^*]$ is evaluated by the machine (deductive step). Then, a covering $\mathcal{H}_2$ of $\mathcal{H}_1$, i.e. $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \mathcal{H}^*$, is considered and $E[\mathcal{H}_2/\mathcal{H}^*]$ is computed. The enlargement $\mathcal{H}_2$ of $\mathcal{H}_1$ is significant for the code optimization problem if $E[\mathcal{H}_2/\mathcal{H}^*] < E[\mathcal{H}_1/\mathcal{H}^*]$. (This is

the feedback phase of the analysis.) Sequential application of these two steps leads to the evaluation of the optimal metamodel $\mathcal{H}^{(opt)}$ of $\underline{X}(t)$. Thus, instead of the classical deductive procedure, one might use a deductive scheme with feedback to calculate the reference set of the input variable.

The evaluation of $\mathcal{H}^{(opt)}$ constitutes the first step towards the computation of the optimal code strategy, f(x) (see Table (I.1)). f(x) minimizes the expected cost in Eq. (I.2) where E[f] is the expected cost of the code corresponding to action f(x) conditional on $\mathcal{H}^{(opt)}$ (details in Sec. II.2). The fact that $\mathcal{H}^{(opt)}$ is preferred to $\mathcal{H}$ does not imply that we are against working with the complete reference set, $\mathcal{H}$, but that in the context of our methodology we have to limit ourselves to only finite and simple-to-operate metamodels. $\mathcal{H}^{(opt)}$ is such a metamodel which is the "best" balance between accuracy and simplicity corresponding to the purpose of code optimization.

The next section presents the organization and the main results of this study.

I.3 RESULTS AND ORGANIZATION

All the results in the thesis rest on Bayes' theorem, although parallel results can be obtained in the frequentist context. The Bayesian formulation is preferred since:

- it is no logical difficulty to interpret in engineering problems any Bayesian result, after the prior-posterior use of the Bayes' theorem is accepted,
- in general the Bayesian distribution are (computationally) suitable in most of the practical situations,
- the Bayes' solution uses besides E the relevant subjective knowledge related to the decision problem at hand.

In contrast, any classical result, such as tolerance regions, R, has to be interpreted in a frequency sense (so, the classical interpretation of R of coverage p is that on average p% of the future realizations of the input variable fall in R). This interpretation is not desirable in various engineering problems. Also, the difficulty of evaluating the pivotal statistics as well as the fact that the frequentist approach cannot use other information but sample (or more generally, E) decreases our interest in the classical solution. One of the essential differences between the Bayesian and the frequentist methodologies is related to the space where the uncertainty is quantified. For instance, let $X(t) = X$ (Fig. (I.1)) be the input random variable whose PDF, $f_X(x/\underline{\theta})$, depends on the unknown parameter $\underline{\theta}$. Then, the Bayesian procedure measures the uncertainty in the model of X by the posterior distribution, $f(\underline{\theta}/E)$. For example, if $f(\underline{\theta}/E)$ has small variance, then a posteriori we are quite confident in the value of $\underline{\theta}$, i.e. in the model of X. In other words a (parametric) Bayesian statistician locates his uncertainty in the space of parameters. This fact is considered wrong by frequentist since the nature has a unique state $\underline{\theta} = \underline{\theta}_N$ and therefore it is meaningless to allow $\underline{\theta}$ to have values other than $\underline{\theta}_N$. In contrast a (parametric) frequentist statis-

tician locates the uncertainty in the space of outcomes of X.

The fact that the parametrical Bayesian statistical confidence regions coincide with parametrical frequentist statistical regions when non-informative priors on the unknown parameters $\underline{\theta}$ of the distribution of X are used (see for instance Zellner [92], p. 43) is a virtue for Bayesians. In fact assuming that Bayesians are right, a frequentist looks (in the context of statistical confidence regions) like an ignorant Bayesian. On the contrary, frequentist statisticians regard this situation as another incorrect procedure which at some limits give the correct answer [16].

It is not the purpose of this thesis to analyze the Bayesian versus the frequentist methodology. As pointed out at the beginning of this section since the Bayesian approach leads to straightforward interpretation of the results, is attractive numerically, is able to use, besides E, subjective knowledge which may be essential in some engineering tasks and fits very well the Decision Theory framework, the Bayesian approach was considered "the best" methodology towards the goal of optimizing engineering decisions.

The organization and the main results of the thesis are reported next: Chapter II, Section II.1 reviews the basic results of classical Decision Theory (DT). An extension of DT for model/metamodel selection, namely, Generalized Bayesian Decision Theory (GBDT) is developed in Section II.2. The essential results in Section II.2 are the algorithm to find the optimal metamodel, $\mathcal{H}^{(opt)}$, of X(t) corresponding to a precise goal (for instance the aim of the design problem in Fig. I.1 is to find the design level $d^{(opt)}$ minimizing the expected cost in Eq. (I.4)), and the evaluation (at code level) of the optimal strategy f(x) (see Table

(I.1)); f(x) together with $\mathcal{H}^{(opt)}$ represent the main results in the thesis towards the goal of code optimization. Also (see Sec. (I.1)), the code strategy f(x) may have significant practical implications in various engineering problems; for instance, Sec. I.1 suggests the use of f(x) in the implementation of Paloheimo's level II code format. All the other chapters develop mostly applications of the methodology in Chapter II for various kinds of random sequences X(t).

Thus Chapter III, Section III.2 analyzes univariate independent sequences having Gaussian, Gamma and Poisson distributions. The essential output is the one-sided (predictive) confidence regions. Also, the use of the noninformative prior on the unknown parameters of various models $H \in \mathcal{H}$ ($\mathcal{H}$ is the metamodel of $\underline{X}(t)$) is discussed in a special paragraph.

Sec. III.3 considers the design of the structural system S in Fig. (I.1) against wind. Results are presented for both, level III and level IV code formats. The following conclusions are drawn:

- In Boston, in accordance with the evidence consisting of 42 years of data (1933-1974), the sequence of yearly wind extreme, $X_y(t)$, is better modeled by the extreme type I distribution than the extreme type II distribution, in the maximum likelihood sense. This result does not necessarily imply that $X_y(t)$ has an extreme type I distribution. Because of the sensitivity of the inferential scheme to the experiment E (see Tables (III.48) and (III.49)) the utility approach illustrated in Sec. III.3.1 seems adequate.
- Considering the problem of designing S for long return periods of the wind when only few years of records are available, in

[19] and [32] is proposed to use, say, the mean hourly wind speed sequence $X_h(t)$. Since $X_h(t)$ is a correlated sequence, a careful evaluation of the Rice's crossing result [19,32] is first developed towards designing S. It appears, however, (Sec. III.3.2) that (at least for the data in [32]) the correlation of $X_h(t)$ is much less significant than the marginal PDF of $X_h(t)$ for the calculation of the design level of S corresponding to long return period wind speeds.

Chapter IV considers Gaussian Markov Sequences (GMS) generated by linear dynamic systems driven by white noise. Sec. IV.1 considers the design problem in Fig. (I.1) for the case when X(t) is GMS. Sec. IV.2 presents several useful results from the theory of Markov Sequences generated by linear systems. Sec. IV.3 develops the ingredients required by the Bayesian Decision framework for Gaussian Markov processes. Specifically, Sec. IV.3.1 summarizes the elements needed in analysis and Sec. IV.3.2 presents the prior to posterior analysis. It is found that if the conjugate prior distribution for the unknown parameters of the GMS is used, no steady-state Bayesian distribution of X(t) can be obtained. Sec. IV.3.3 develops detailed solutions of the problems (i), (ii) and (iii) considered at the beginning of Chap. I.

Sensitivity plots giving the penalty associated with the decision of using a GMS of order i instead of a GMS of order j provide a direct solution of task (i). The problem (ii) is a direct consequence of the posterior results derived in the body of Chap. IV. Unfortunately for important problem (iii) only an approximate solution is found. A peripheral result related to (iii) is that the mean rate of upcrossings of level

X=d of any (finite) order GMS X(t) is well approximated by the mean rate of upcrossings of the independent sequence having the same marginal distribution as X(t) provided d is large (say, $d \geq m_{X(t)} + 2.5\sigma_{X(t)}$, where $m_{X(t)}$ and $\sigma_{X(t)}$ denote the instantaneously mean and standard deviation of X(t), respectively. This observation is used also for the evaluation of the mean rate of upcrossings and return periods of wind starting from mean hourly data and assuming independence.

Chap. V analyzes the design of S in Fig. (I.1) loaded by renewal sequences of random magnitudes. Specific results are obtained for the case when the interarrival times are gamma distributed with integer shape parameter.

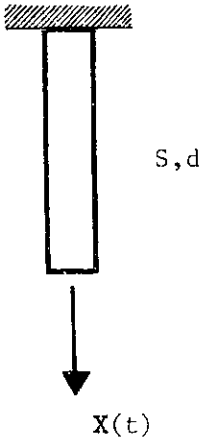


Fig. I.1. Reference design problem.
S is a deterministic two-state (failure/survival) system designed at x=d. X(t) is the input variable.

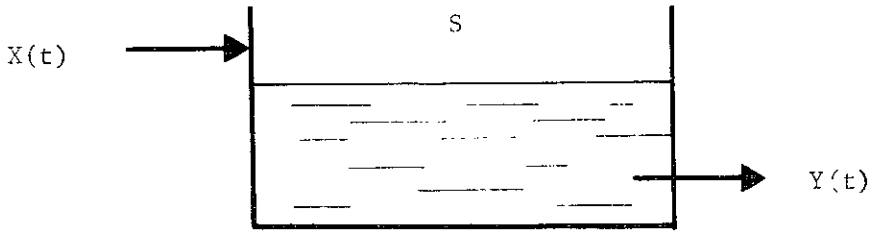


Fig. I.2. Operation of a reservoir.
S is a reservoir of size d. X(t) and Y(t) are the amount of water entering S and the water release at time t, respectively.

CHAPTER II

ESSENTIALS AND GENERALITIES

This chapter formalizes some of the concepts introduced in Chapter I. The final goal of the analysis (see also Chap. I) is to evaluate the optimal action that the engineer has to take in a specific context. For instance, in the problem in Fig. (I.1) the purpose of the study may be to find the optimal (in a utility sense) design level, $d^{(opt)}$, of S when the input is a partially known random processes X(t). The emphasis in this work is on the problems associated with only partially known distributions or models of X(t). In the context of Codified Structural Reliability the final goal of optimal design the systems $\{S_i\}$ controlled by a code C, is achieved in three steps. First, the optimal metamodel, $\mathcal{H}^{(opt)}$, is found; the evaluation of $\mathcal{H}^{(opt)}$ (where $\mathcal{H}^{(opt)}$ contains only a few analytically simple models, but is still sufficiently accurate for the purpose of the analysis) is needed since the complete metamodel, $\mathcal{H}$, of X(t) ($\mathcal{H}$ contains all the possible models of X(t)) may be an infinite set containing models which are not simple computationally. $\mathcal{H}^{(opt)}$ represents the optimal action in the decision framework in Section II.2 satisfying $\mathcal{H}^{(opt)} \subseteq \mathcal{H}^*$, where $\mathcal{H}^*$ is called the operational metamodel; $\mathcal{H}^*$ (defined in (II.13)) con-

tains only models of $\underline{X}(t)$ which can be processed through the decision framework with a reasonable computation cost. The relation between these three metamodels is: $\mathcal{H}^{(opt)} \subseteq \mathcal{H}^{*} \subseteq \mathcal{H}$ (see also Fig. II.1). Second, at the code level the optimal strategy, f(x), (see Table (I.1)) is found as a function of the code format. Thus, the code specifications are as follows: at level I a set of characteristic values (actual code format), at level II a set $\{\beta_i\}$ of safety factors associated with the components, $X_i(t)$, of the input variable $\underline{X}(t)$ in [56], (see also [15], [71], [73]), at level III the "distribution" $f(\underline{x})$ of $\underline{X}(t)$ and the target failure probability P_f, and at level IV the "distribution" of $\underline{X}(t)$, [48]. (At level III and level IV "distribution" of $\underline{X}(t)$ stands for the optimal code strategy which may have the form of a probability distribution function; see Table (I.1). We shall see however that $f(\underline{x})$ is not necessarily the PDF of a model H of $\underline{X}(t)$.) Third, $f(\underline{x})$, is used by the designer, in accordance with code rules to design individual systems (see illustration in Eq. (I.3) and Eq. (I.4) for level III and level IV design, respectively).

The organization of the chapter is as follows. Sec. II.1 presents the basic methodology of (Bayesian) Decision Theory (DT) and several variations of it, which will be used extensively in this study. Sec. II.2 presents a generalized version of DT. First, conditional on a finite metamodel, the optimal (level III and IV) strategy, f(x), is evaluated and then, two procedures to select $\mathcal{H}^{(opt)}$ are developed. The application of the results in this section to code optimization is discussed briefly. Also, the sense of suboptimality of the output of the analysis is explained. Sec. II.3 considers some of the limitations of the procedure in Sec. II.2 Also, the content of the chapters III, IV, and V is briefly reported.

II.1 STATISTICAL DECISION PROBLEM

This section contains some basic elements of Decision Theory which are used throughout this study.

Let $\underline{X}(t)$ be the (partially known) multivariate process relevant to the design problem at hand. For example, in the simple case of Fig. (I.1) $\underline{X}(t)$ is a scalar load process.

Denote by $\mathcal{H} = \{H\}$ the complete metamodel of $\underline{X}(t)$, by $A = \{a\}$ the space of actions (strategies) available to the statistician, and by $L(a,H)$ the expected loss if action "a" is chosen and H is the true law of $\underline{X}(t)$.

Then the expected loss of action "a" conditional on $\mathcal{H}$ (assumed here finite) and on the experiment E is:

$$E[a] = \sum_{H \in \mathcal{H}} L(a,H)\, p_E(H) \qquad \text{(II.1)}$$

where, $p_E(H)$ is the (Bayesian) posterior probability of H, $H \in \mathcal{H}$. Eq. (II.1) is a generalization of Eq. (I.8) for the case when A is different from $\mathcal{H}$.

The optimal action, a_{opt}, satisfies:

$$E[a_{opt}] = \inf_{a \in A} \{E[a]\} \qquad \text{(II.2)}$$

For instance, the metamodel in the example of Sec. I.2 is

$$\mathcal{H} = \{H_1 = (\text{queen's age} > 40),\ H_2 = (\text{queen's age} \leq 40)\}$$

and the action space has the form:

$$A = \{a_1 = (\text{queen's age} > 40),\ a_2 = (\text{queen's age} \leq 40)\}$$

In this case the experiment E leading to the posterior probabilities

$p_E(H_1) = 9/10$ and $p_E(H_2) = 1/10$ might consist of discussions with persons having knowledge about the queen's age. This example is quite particular since $\mathcal{H} = A$ and $\mathcal{H}$ is finite.

The ingredients of a general Bayesian decision problem are:

1. - the Bayesian description of the metamodel $\mathcal{H}$, i.e. a (prior) probability distribution on the models in $\mathcal{H}$. In addition, one must give the prior density, $f'_H(\underline{\theta}_H)$, of the unknown parameters, $\underline{\theta}_H$, and the probabilistic characterization (conditional on $\underline{\theta}_H$) of each model H in $\mathcal{H}$.

2. - the experiment, E, which in this study consists of a set of perfect/noisy observations of $\underline{X}(t)$ (where "noisy" observations may mean that $\underline{X}(t)$ is measured with errors) over a finite length of time. (In general the measurements are not continuous and one can measure either instantaneously values of $\underline{X}(t)$ or various functions of $\underline{X}(t)$.

3. - the loss function, L(a,H), which depends on the deterministic loss c(a,$\underline{y}$) of action "a" when the state of nature is $\underline{X}(t) = \underline{y}$ and on the simple/simultaneous predictive distribution (see Zellner [92], p. 29, 187 and Chap. III, IV, and V of this thesis) of the state of H, $H \in \mathcal{H}$, conditional on E.

The expected loss L(a,H) is:

$$L(a,H) = E_{\underline{Y}}[c(a,\underline{y})/H] \qquad \text{(II.3)}$$

where, $E_{\underline{y}}[\cdot/H]$ is the a posteriori expectation of c(a,$\underline{y}$) over all possible outcomes $\underline{y}$ of the state of H.

The Bayesian (posterior) expected cost of $a \in A$, conditional on H, is

$$E[a/\mathcal{H}] = \sum_{H \in \mathcal{H}} L(a,H)\, p''_H \qquad \text{(II.4)}$$

where, p''_H is the (posterior) probability of $H \in \mathcal{H}$,

and the optimal action a_{opt}, (Bayesian strategy) is obtained through minimization of $E[a/\mathcal{H}]$ in Eq. (II.4).

$$E[a_{opt}/\mathcal{H}] = \inf_{a \in A} \{E[a/\mathcal{H}]\} \qquad \text{(II.5)}$$

Thus the optimal action, a_{opt}, is obtained from the decision step (Eq. (II.5)). To evaluate $E[a/\mathcal{H}]$ in Eq. (II.4), first the inferential step consisting of:

- computation of the posterior probability, p''_H, of $H \in \mathcal{H}$, and
- the computation of (see Eq. (II.3)) the simple/simultaneous predictive distribution of $\underline{X}(t)$ conditional on $H \in \mathcal{H}$ being true

has to be completed first.

A different way of selecting the "best" strategy might be based on the minimax rule. In this case, the best strategy or the minimiax strategy, $\tilde{a}_{opt}$, satisfies:

$$\sup_{H \in \mathcal{H}} \{L(\tilde{a}_{opt}, H)\} = \inf_{a \in A} \sup_{H \in \mathcal{H}} \{L(a,H)\} \qquad \text{(II.6)}$$

The minimax strategy represents the basis for decision in Game Theory. In fact (see Ferguson [27], Chap. I) any Decision Task can be viewed as a two person game between the statistician (player 1) and nature (player 2).

The optimal action, $\tilde{a}_{opt}$, in Eq. (II.6) minimizes the maximum loss of player 1; in Eq. (II.6) it is implicitly assumed that both players know the rules of the game and that each player is always trying to minimize his maximum (expected) loss, $\sup_{H \in \mathcal{H}} \{L(a,H)\}$.

However, certain conceptual differences between Game Theory (GT) and Decision Theory (DT) are essential in the choice of the order in A; some of these differences are presented next:

1. - In GT both players are aiming to minimize their maximum (expected) loss; this is not true in the "game" between the statistician and nature since nature follows its own laws and clearly is not going to modify its behaviour in function of various actions that the statistician may choose.

2. - Also, in DT the notion of loss is meaningful only for <u>one</u> player.

3. - In GT the rules of the game are known; in DT the rules of the game for the statistician are only partially known. In fact the behaviour of player 2 (nature) is infered from the experiment E.

One might conclude that the essential element in DT is the ability of the statistician to estimate the law of nature which together with the loss L(a,H) defines the game for the statistician.

Since the Bayesian model $H_B = \sum_{H \in \mathcal{H}} p''_H H$ is (inferentially) more accurate than any other combination of the model in $\mathcal{H}$ and since the purpose is to evaluate best the rules of the game, the expected cost, $E_{\underline{y}}[c(a,\underline{y})/H_B]$, of $c(a,\underline{y})$ with respect to the distribution of H_B should be evaluated; it is $E_{\underline{y}}[c(a,\underline{y})/H_B] = E[a/\mathcal{H}]$ in Eq. (II.4). Considering the fact that nature behaves independently of the statistician the minimax rule (II.6) does not seem appropriate for the game statistician - nature. Consequently, only the Bayesian strategy (Eq. (II.5)) is considered in this study.

Assume for the moment that A and $\mathcal{H}$ are finite; let $\{H_1, H_2, \ldots, H_m\}$ be the elements of $\mathcal{H}$ and $\{a_1, a_2, \ldots, a_n\}$ the elements of A (the hypothesis that A is finite is released later). Our goal now is to evaluate the modifications of the optimal action, a_{opt}, corresponding to A when various enlargements of the action space are considered. Consequently, two addi-

tional spaces of actions, A^* (the space of randomized strategies) and A_c (the space of composite strategies) will be defined. For the both new action spaces the optimal strategy is evaluated here without accounting for simplicity (of the actions).

The dependence of the optimal strategy on the action space is easily obtained if the geometric interpretation of the Bayes rule (Eq. (II.5)) is considered.

The expected cost of action $a_i \in A$ (see Eq. (II.4)) is:

$$E[a_i/\mathcal{H}] = \sum_{k=1}^{m} L(a_i,H_k)\ p''_k \qquad \text{(II.7)}$$

where, $L(a_i,H_k)$ and p''_k have the same meaning as in Eq. (II.4)

Consider first the space of randomized actions (strategies), $A^* = \{\delta : \delta = \text{probability on } A\}$; the randomized strategy $\delta \in A^*$ chooses action a_i with probability $\delta(a_i)$, i=1,2,...,n. The expected loss, $R(\delta,H_j)$, under the randomized strategy δ and given that H_j is the actual state of nature is generally (see [27], p. 23, 35) linear in $\{L(a_i,H_j)\}_i$, i.e.

$$R(\delta,H_j) = \sum_{i=1}^{n} L(a_i,H_j)\ \delta(a_i) \qquad \text{(II.8)}$$

$$(j = 1,2,...,m)$$

and the expected loss of strategy δ for a given $\mathcal{H}$ has the form (see Eq. (II.7))

$$E[\delta/\mathcal{H}] = \sum_{j=1}^{m} R(\delta,H_j)\ p''_j \qquad \text{(II.9)}$$

Let $\delta = \delta_i$ be the randomized strategy which weights with one a_i and with zero $a_j \neq a_i$. Then, from Eq. (II.8) it follows that $R(\delta_i,H_j) = L(a_i,H_j)$

and therefore (see Eq. (II.9) and Eq. (II.7)) $E[\delta_i/\mathcal{H}] = E[a_i/\mathcal{H}]$, meaning that the actions δ_i and a_i are equivalent in the expected cost sense.

Denote by $\underline{\xi} = (\xi_1,\ldots,\xi_m)$ the point in R^m corresponding to δ, where $\xi_j = R(\delta,H_j)$, $j=1,\ldots,m$. The set of points $\{\underline{\xi}\}$ corresponding to all $\delta\varepsilon A^*$, V, is called the randomized risk set and the set of points $\underline{\xi}\varepsilon R^m$ corresponding only to pure actions $\{a_i\}$, V_o, is called the nonrandomized risk set. For example, in the case when $\mathcal{H} = \{H_1,H_2\}$ and $A = \{a_1,a_2,\ldots a_5\}$ (i.e. n=5 and m=2), $V_o = \{\underline{\xi}^1,\underline{\xi}^2,\ldots,\underline{\xi}^5\}$ where the coordinates of $\underline{\xi}^k$, $k=1,\ldots,5$, are $(L(a_k,H_1), L(a_k,H_2))$ and V is the triangular region Δ_{2-3-5} (see Fig. (II.2)). Any point $\underline{\xi}\varepsilon\Delta_{2-3-5}$ corresponds to a certain randomized strategy $\delta\varepsilon A^*$ and has the coordintes $(R(\delta,H_1), R(\delta,H_2))$.

If Eq. (II.8) is satisfied the following statements are true:

(a) The randomized risk set, V, is the smallest convex covering (the hull) of the nonrandomized risk set V_o, ([27], p. 75). (Let A and B be two sets; then B is a covering of A if $B \supseteq A$. The hull of A, $\mathcal{H}(A)$ is: $\mathcal{H}(A) = \bigcap_{\substack{B\supseteq A \\ B=\text{convex}}} B$, where B is convex if for any two points $b_1,b_2\varepsilon B$, the segment $\overline{b_1b_2}$ also belongs to B.)

(b) Since (see [27], p. 37, and Eq. (II.5))

$$E[a_{opt}/\mathcal{H}] = \min_{1\le i\le n} \{E[a_i/\mathcal{H}]\} = \inf_{\underline{\xi}\varepsilon V} \{\sum_{k=1}^{m} p_k''\xi_k\} \qquad \text{(II.10)}$$

the Bayes point (i.e. the point in $\underline{\xi}$-space corresponding to the best strategy, Fig. (II.2)) is obtained by translating towards the origin the hyperplane with normal $(p_1'',\ldots,p_m'')$ until one or more tangency points are found; the two situations correspond to the cases of unique and multiple Bayes solutions, respectively.

(c) Unless the Bayesian solution is not unique ([27], p. 43) the Bayesian strategy is a pure strategy. From Eq. (II.10) and Fig. (II.2) it results that the optimal action, a_{opt}, (it is assumed here that the Bayesian solution is unique) corresponding to A coincides with the optimal action, δ_{opt}, in A^*, in the sense $E[a_{opt}/\mathcal{H}] = E[\delta_{opt}/\mathcal{H}]$. Also, the points in the space $\{\xi\}$ defined by a_{opt} and δ_{opt} coincide.

Consider now the space of composite actions $A_c = \{a_c\}$ where a_c is formally expressed as the weighted sum $a_c = \sum_{i=1}^{n} \eta(a_i)a_i$, where $\sum_{i=1}^{n} \eta(a_i)=1.0$ and $\eta(a_i) \geq 0$ for all i (a_c will also be referred to as the composite action η). For example, if $A \equiv \mathcal{H}$ and $\mathcal{H}$ is the metamodel of the random variable X, then the composite action $H_c = \sum_{i=1}^{m} \eta(H_i)H_i$ is characterized by the distribution

$$f_{\eta}(x) = \sum_{i=1}^{n} \eta(H_i) f_{H_i}(x) \qquad \text{(II.11)}$$

where, $f_{H_i}(x)$ is the distribution of H_i.

In fact the pair $\{A_c,\mathcal{H}\}$ defines a new decision problem with the same states of nature, $\mathcal{H}$, as the original problem, $\{A,\mathcal{H}\}$, while the space of actions, A_c, is a continuous space of (pure) strategies, a_c; A_c is probably the obvious extension of A to the space of its linear combinations. One may consider more elaborated coverings of A; for instance, when $A = \mathcal{H}$, instead of A_c (see Eq. (II.11)) the space $\tilde{A}$ of all density functions ($\tilde{A} \supseteq A_c \supseteq A$) can be taken.

It will be seen that in most reliability problems the expected cost, $R(\eta,H_j)$, when action η(or H_c) is chosen and H_j is the true law of, say, input variable X is in fact a nonlinear function of $\{L(H_i,H_j)\}_i$ (see illustration in Fig.(II.3) for the case when $\mathcal{H} = A = \{H_1,H_2\}$). Consequently

the optimal strategy is not necessarily a pure strategy in A. For instance, if the (nonrandomized) risk set corresponding to the space of composite actions, A_c, has the form in Fig. (II.3), the optimal strategy $H_c^{(opt)} = \sum_{i=1}^{2} \eta_{(H_i)}^{(opt)} H_i$, corresponds to the Bayesian point $\underline{\xi}^{(opt)}$. The explicit form of the optimal composite strategy, $H_c^{(opt)}$, is:

$$f(x) = \sum_{i=1}^{2} \eta_{(H_i)}^{(opt)} f_{H_i}(x) \qquad \text{(II.12)}$$

The last considerations (see Eq. (II.12) and Fig. (II.3)) are used and emphasized in Sec. II.2 where the optimal code strategy f(x) (see also Table (I.1)) is evaluated.

II.2 GENERALIZED BAYESIAN DECISION THEORY

As in Sec. II.1 denoted by $\underline{X}(t)$ the input variable of a certain engineering task. The final goal of finding, say, the optimal design level, $d^{(opt)}$, of S (Fig. II.1)) is approached in two phases. First, conditional on the reference set of models, $\mathcal{H} = \{H\}$, of X(t) the optimal (level III and level IV) code strategy, $f(\underline{x})$, is evaluated following basically the methodology in Sec. II.1 (see the case when the composite strategies, H_c, are considered, Eq. (II.11)).

Second, the optimal reference set for $\underline{X}(t)$, $\mathcal{H}^{(opt)}$, is found by solving a second decision problem at the level of metamodel. Clearly, in actual applications these two steps follow each other in reverse order, i.e. first $\mathcal{H}^{(opt)}$ is evaluated and then the optimal strategy f(x) is computed conditional on $\mathcal{H}^{(opt)}$.

The need for an analytical approach to select $\mathcal{H}^{(opt)}$ as opposed to making a subjective choice may be justified by the following example. Assume that due to the lack of experience or lack of available information one chooses from the complete metamodel $\mathcal{H}$ of the random variable X the operational metamodel $\mathcal{H}'$ containing only models which have very low (posterior) probability of being true. For example, let X denote the yearly wind extremes and, say, $\mathcal{H}'$ = {Weibull distribution}; then the optimal design, d_S', of S in Fig. (I.1) conditional on $\mathcal{H}'$ will be much smaller than the present design level d_S'', where d_S'' is the optimal design level of S conditional on $\mathcal{H}''$ = {extreme type I distribution}. Consequently, the design d_S' may be quite unconservative, if, say, $\mathcal{H}''$ would be the actual law of nature.

Our main disagreement with the usual practice is not related to the subjective choice of the operational metamodel of X but to the fact that only <u>one</u> metamodel is selected and then deductively the optimal strategy is evaluated; the sensitivity of the optimal action to the metamodel is never considered. Consequently, the following qualitative rule has to guide the choise of the operational metamodel, $\mathcal{H}^*$.

> "The operational metamodel $\mathcal{H}^*$ must contain <u>all</u> types of models of $\underline{X}(t)$ which are sufficiently simple analytically; the prior weight, p_H, of various models H in $\mathcal{H}^*$ is related to the subjective interpretation of the physical mechanism of $\underline{X}(t)$." (II.13)

For example, the prior probabilities $\{p_H'\}$ on $\mathcal{H}^*$ may be the result of a group of individuals (say, code committee) reaching consensus ([22]); also Winkler [90] considers procedures of assigning prior probabilities.

Although very important, the rule (II.13) may lead to significant nu-

merical difficulties when for instance, (II.13) requires many models in $\mathcal{H}^*$ which are not too simple numerically. Thus, a trade off between accuracy (rule (II.13)) and the possibility to handle the problem in a practical sense has to be made. This requirement is expressed by so-called simplicity argument.

> "The metamodel, $\mathcal{H}^*$, should contain only a few simple-to-operate models." (II.14)

where (in the Bayesian context of this study) to operate a model, H, means to use the simple/simultaneous predictive distribution of H and the posterior probability of H being true (conditional on $\mathcal{H}^*$) in the various computations required by the decision framework.

We are not going to discuss here the philosophical aspects of simplicity; only a few conclusions which will clarify the meaning of the rule (II.14) are considered next.

Many scientists have expressed the conviction that the basic laws of nature are simple, from which comes the natural philosophical justification for simpler representations of various states of nature.

The question of simplicity was first related to the curve fitting problem. Assume that measurements $(x_i,y_i)i=1,2,\ldots,n$ have been made, the problem is to find the functional relation $y=f(x)$, between x and y. If no assumption on the form of $f(x)$ is done, any function $g(x)$ satisfying $y_i = g(x_i)$, $i=1,\ldots,n$, is a valid choice; consequently the function $f(x)$ <u>cannot</u> be identified if $n<\infty$. Thus, the choice of a particular model of $f(x)$ is done generally in two steps: first a functional form $g(x)$ (typi-

cally analytically simple) is postulated and then the parameters of g(x) are inferred from the measurements.

Popper [29] identifies simplicity with falsifiability. He proposed two criteria to compare the falsifiability of two hypotheses: (i) the inclusion criterion (if hypothesis H_1 implies hypothesis H_2 and H_2 does not imply H_1, then H_1 is more falsifiable than H_2) and (ii) the dimension criterion (H_1 is more falsifiable than H_2 if less points are needed to reject H_1 than H_2). Popper's conclusion is that science should search for the simplest, i.e. the most falsifiable hypothesis. Popper's position is not without difficulties (see [29], p. 247).

On the other hand Bunge [29] has shown that it is meaningless to speak about <u>overall simplicity</u> because, generally,the degrees of complexity in various respects are not additive. In Bunge's view, the following aspects of simplicity might be considered (instead of overall simplicity) when modeling a certain state/law of nature (see details at p. 280 in the original paper [29]): syntactical simplicity (economy of forms), semantical simplicity (economy of presuppositions), epistomological simplicity (economy of transcendental terms. Epistemological requirements are: consistency with observations, explanatory power, predictive power, etc.; therefore epistemological simplicity is not desirable in a theory.) and pragmatical simplicity (economy of work).

For instance, the Theory of Elasticity (TE) as a model of the elastic deformable mediums is syntactically simple (primitive concepts in TE are Newton's Laws and the notion of an elastic medium) but it is not pragmatically simple,say, compared with the Strength of Material (in fact only few problems are given closed form analytical solutions). Consequently,

TE is neither overall simple nor overall complex. Since generally in engineering problems what counts is the simplicity of the algorithm (the engineers are basically theory-users), it is the tendency in engineering to identify the simplicity with pragmatical simplicity.

The decision framework developed in this study accounts only for pragmatical simplicity.

Consider now the practical problem of best designing the system S in Fig. (I.1).

The optimal code committee strategy, f(x), (see Table (I.1)) and the designer's best action (see Eq. (I.3) and Eq. (I.4) are evaluated next. First, at code level the optimal strategy f(x) is calculated conditional on the complete metamodel $\mathcal{H}$ considered for the moment finite; various action spaces and their optimal strategies are summarized in Table (II.38). Once f(x) is obtained the design of any individual system follows the routine described in Chap. I (see Eq. (I.3), Eq. (I.4)).

Second, the assumption that $\mathcal{H}$ is both known and finite is released and a decision framework providing $\mathcal{H}^{(opt)}$ as the best (conditional on $\mathcal{H}^{*}$) compromize between accuracy and simplicity is developed.

Let, $\mathcal{H} = \{H_i\}_{i=1}^{m}$, m finite, be the complete metamodel of the random variable X = X(t) in Fig. (I.1). The optimal design of S, $d^{(opt)}$, is obtained in two phases: code phase and designer phase.

First is calculated the optimal level IV design, d_{IV}, of S.

(a) code phase: with output the optimal strategy f(x) corresponding to, say, the input variable X of S in Fig. (I.1).

Denote by $f_\eta(x/E)$, or shortly η, (see Eq. (II.11)), the set of all the strategies that code committee may consider, where

$$f_{\eta}(x/E) = \sum_{i=1}^{m} \eta(H_i)\, f_{H_i}(x/E) \qquad \text{(II.15)}$$

with, $\sum_{i=1}^{m} \eta(H_i) = 1.0$, $\eta(H_i) \geq 0$ for all i, and $f_{H_i}(x/E)$ is the distribution of X conditional on H_i and the informative experiment E.

In Eq. (II.15), $f_{H_i}(x/E)$ may be the simple/simultaneous Bayesian distribution of H_i, or the PDF in H_i where the unknown parameters, $\underline{\theta}_i$, are replaced by their point estimates $\underline{\hat{\theta}}_i$, etc. Also, strategies $\eta = \eta_j$, where $\eta_j(H_i)$ is zero if $i \neq j$ and is one when $i = j$, are accepted in Eq. (II.15) clearly $f_{\eta_j}(x/E) = f_{H_j}(x/E)$ i.e. $\{\eta_j\}$ constitutes the space of pure strategies. See a summary of all possible (level III and level IV) code strategies in Table (II.38).

Table (II.16)

		State of Nature				
		H_1	. . .	H_i	. . .	H_m
Code Committee Strategies	η	$R(\eta;H_1)$	. . .	$R(\eta;H_i)$	. . .	$R(\eta;H_m)$

Let $R(\eta,H_i)$ (see Table (II.16)) be the expected cost of the code if action η is chosen and H_i is the true law of X. Then,

$$R(\eta,H_i) = c_{\eta} + E[d_{\eta}^{(opt)}/H_i] \qquad \text{(II.17)}$$

where, c_{η} is the computational cost of, say, evaluating the optimal design of S corresponding to density $f_{\eta}(x/E)$; since minimization in Eq. (II.19) is generally done numerically, the

penalty c_η may be related to the cost to implement and to use subroutines calculating the CDF $F_\eta(x/E)$. $E[d/H_i]$ has the same form as in Eq. (I.4); also (see Eq. (I.2)), $E[d/H_i]$ may be interpreted as the code expected cost corresponding to the design x=d when the law of X is H_i

$$E[d/H_i] = c(d) + c_F[1 - F_{H_i}(d/E)] \tag{II.18}$$

The (code) expected cost of the design of level X=d corresponding to $\eta(F_\eta(x/E)$ is the cumulative distribution function (CDF) of $f_\eta(x/E))$ is (see also Eq. (II.18)):

$$E[d/\eta] = c(d) + c_F[1 - F_\eta(d/E)] \tag{II.19}$$

and consequently the optimal design level, $d_\eta^{(opt)}$, of η satisfies

$$E[d_\eta^{(opt)}/\eta] = \inf_d \{E[d/\eta]\} \tag{II.20}$$

Figure (II.4) shows hypothetical relationships of the type (II.18) and (II.19) for the case when m=2; $d_i^{(opt)}$ denotes the best design conditional on H_i.

It is easy to see that $E[d_\eta^{(opt)}/H_i]$ is not linear in $E[d_j^{(opt)}/H_i]$ j=1,2, i=1,2; in the context of Sec. II.1 it means that the expected cost of the composite strategy η is nonlinear in the expected cost of the pure strategies. As a consequence (see also Fig. (II.3)) the best strategy is not necessarily a pure strategy. The graphical evaluation of the Bayesian strategy is shown in Fig. (II.5); the optimal strategy, $\eta^{(opt)}$, corresponding to the Bayesian point $\underline{\xi}^{(opt)}$, minimizes the expected cost

$$E[\eta] = \sum_{i=1}^{m} R(\eta, H_i) p''_{H_i} \qquad \text{(II.21)}$$

where, p''_{H_i} is the posterior probability that model H_i is true.

The discontinuity of the risk set $\overline{V}_0 = \{\underline{\xi}^1 \cup \underline{\xi}^2 \cup \tilde{V}_0\}$ (see Fig. (II.5))is implied by the inequality $c_\eta \geq c_{\eta_i}$, saying that in general it is more "expensive" to operate a composite strategy, η, than any of the pure strategies, η_i (where $f_{\eta_i}(x/E) = f_{H_i}(x/E)$) entering η.

Thus, when $\mathcal{H} = \{H\}_{i=1}^{m}$ is known and finite, from the minimization versus η in Eq. (II.21) "the optimal" level IV code strategy, $\eta_{IV}^{(opt)}$, is found. If in Eq. (II.15) only pure strategies, η_i, are considered, then $\eta_{IV}^{(opt)}$ is itself a pure strategy; in contrast, when η in Eq. (II.15) is not restricted to the class of pure strategies, in general, $\eta_{IV}^{(opt)}$ is a linear combination of pure strategies (see various possibilities in Table (II.38)).

However, the (utilitarian) level IV strategy

$$f_{IV}(x) = \sum_{i=1}^{m} \eta_{IV}^{(opt)}(H_i) f_{H_i}(x/E) \qquad \text{(II.22)}$$

is not necessarily the best since the actions of the code committee were restrained to the space $A_c = \{f_\eta(x/E)\}$, where $f_\eta(x/E)$ is defined in Eq. (II.15). Let $\tilde{A}$ be the space of all (univariate) densities; then, following the routine to evaluate $f_{IV}(x)$ in Eq. (II.22), the optimal strategy $\tilde{f}(x)$ of $\tilde{A}$ is calculated. Since $\tilde{A} \supseteq A_c$ the strategy $\tilde{f}(x)$ is better than $f_{IV}(x)$ in the sense of the expected cost in Eq. (II.21); consequently, $f_{IV}(x)$ is only a <u>suboptimal</u> strategy. However, if it is true that the cost of operating any action in $\tilde{A} \setminus A_c$ is infinity, then $f_{IV}(x) = \tilde{f}(x)$

is the optimal action of the code committee.

The restriction to the action space A_c is recommended here only for a computational reason; to calculate $f_{IV}(x)$ one minimizes $E[\eta]$ in Eq. (II.21) with respect to parameters $\{\eta(H_i)\}_{i=1}^{m}$. In contrast the evaluation of $\tilde{f}(x)$ requires the minimization in Eq. (II.21) with respect to the elements of the infinite set $\tilde{A}$. Also, it is expected that $f_{IV}(x)$ is generally (say, for typical ratios $c(d)/c_F$) a good approximation of $\tilde{f}(x)$.

The final step of the analysis calculates the design level of a given particular system S.

(b) <u>designer phase</u>: which at level IV and for the cost structure in Eq. (I.4) (see also Eq. (II.18)) consists in minimizing the expected cost $E_S[d/\eta_{IV}^{(opt)}]$ of designing S at x=d in accordance with the strategy $f_{IV}(x)$. From Eq. (II.19) it is

$$E_S[d/\eta_{IV}^{(opt)}] = c_S(d) + c_{F,S}[1- F_{\eta_{IV}^{(opt)}}(d/E)] \qquad (II.23)$$

where $c_S(d)$ is the initial cost of designing S at x=d and $c_{F,S}$ is the cost of failure.

The optimal design level of S, d_{IV}, satisfies $E_S[d_{IV}/\eta_{IV}^{(opt)}] = \inf_d\{E_S[d/\eta_{IV}^{(opt)}]\}$. Conditional on $\mathcal{H}$, two sources of suboptimality of d_{IV} can be recognized: the first source is the suboptimality of $f_{IV}(x)$; the second source is due to the difference between the costs of the code (see Eq. (I.2), Eq. (II.19) and Eq. (II.23)), and of the specific system, S.

If the first source of suboptimality cannot be removed, the second one can if for instance a level V code format is accepted; in this context the code gives the designer only the metamodel and full optimization as in (a) is done for each individual system, S.

A similar algorithm is developed for level III code format. Consider for the moment only pure strategies η_i, Eq. (II.15), and evaluate $\{d_i\ (P_f)\}_{i=1}^{m}$, where $d_i(P_f)$ is the P_f - upper fractile of strategy $f_\eta(x/E) = f_{\eta_i}(x/E)$ and P_f represents the target failure probability; P_f is either chosen directly or calculated through optimization of the code parameters. (The case when composite strategies are considered is analogous.) Then, the expected cost of each pair: action η_i, state of nature, H_j, can be calculated (see Eq. (II.17)):

$$R(\eta_i,H_j) = c_{\eta_i} + E[d_i(P_f)/H_j] \tag{II.24}$$

Since at level III the design of S corresponding to strategy η consists of the evaluation of the $(1-P_f)$ - fractile of $f_\eta(x/E)$, pure strategies with tabulated CDF or with CDF in analytical closed-form will receive a much lower penalty than, say, any composite strategy.

In the case when only pure strategies, H_i, are considered, the risk set, V_0, contains a finite number of points. For example, if $A = \mathcal{H} = \{H_1,H_2\}$ as in Fig. (II.5), the level III risk set has two points: $\{\underline{\xi}^1, \underline{\xi}^2\}$ and the optimal strategy is easily found, either graphically or analytically.

As in the case of level IV format the output of the code is the (sub)optimal strategy, $\eta_{III}^{(opt)} = \eta_{i_o}$, where η_{i_o} minimizes the expected cost (see also Eq. (II.21)).

$$E[\eta_i] = \sum_{i=1}^{m} R(\eta_i,H_j)P''_{H_i} \tag{II.25}$$

The designer of the system S, designs S at (see also Eq. (I.3)) $d_{III} = d_{H_{i_o}}(P_f)$; the expected failure probability of S designed at

$x = d_{III}$ is:

$$E[P_f/d_{III}] = \sum_{i=1}^{m} [1 - F_{H_i}(d_{III}/E)]p''_{H_i} \qquad (II.26)$$

Figure(II.6) shows the estimation of the design d_{III} for the case when $f_{III}(x) = f_{H_1}(x/E)$ and m=2; $f_B(x/E) = \sum_{i=1}^{2} p''_{H_i} f_{H_i}(x/E)$ is the density of the Bayesian composite strategy, obtained from Eq. (II.15) for $\eta(H_i) = p''_{H_i}$. From the hypothetical situation in Fig. (II.6) it appears that $E[P_f/d_{III}]$ may be much larger than the target failure probability. However for typical cost functions, the difference between $E[P_f/d_{III}]$ and P_f is expected not to be significant.

Comparing the outputs of level III and level IV without accounting for the expenditure of evaluating the optimal design of S, one finds that $E[\eta_{III}^{(opt)}] \leq E[\eta_{IV}^{(opt)}]$. However, the use of $\eta_{III}^{(opt)}$ at designer level, requires only the evaluation of $(1-P_f)$ fractile of $f_{III}(x) = \sum_{i=1}^{m} \eta_{III}^{(opt)}(H_i) f_{H_i}(x/E)$ versus full optimization (see Eq. (II.23)) when $\eta_{IV}^{(opt)}$ is considered.

The evaluation of the optimal code strategies $\eta_{III}^{(opt)}$ and $\eta_{IV}^{(opt)}$, was done conditional on the metamodel, $\mathcal{H} = \{H_i\}_{i=1}^{m}$ with $m < \infty$, of the input variable X(t). Generally the complete metamodel $\mathcal{H}$ is an infinite set so that the operational metamodel, $\mathcal{H}^*$, is considered instead; since $\mathcal{H}^*$ may contain many models with various degrees of analytical difficulty the subset $\mathcal{H}^{(opt)}$ of $\mathcal{H}^*$ satisfying a certain optimality condition is generally evaluated.

Two schemes to calculate $\mathcal{H}^{(opt)}$ are presented next:

1. Sensitivity Analysis Assume two metamodels $\mathcal{H}_j \subseteq \mathcal{H}$, j=1,2 corresponding to the input variable X(t) and denote by $\eta(\mathcal{H}_j)$ the optimal

strategy (for instance, $\eta(\mathcal{H}_j) = \eta_{IV}^{(opt)}$ in Eq. (II.21) under $\mathcal{H}_j$) conditional on $\mathcal{H}_j$.

Consider now that $\mathcal{H}_2 \supseteq \mathcal{H}_1$, i.e. $\mathcal{H}_2$ is an enlargement or covering of $\mathcal{H}_1$, and evaluate the difference

$$\Delta_{12} = E[\eta(\mathcal{H}_1)/\mathcal{H}_2] - E[\eta(\mathcal{H}_2)/\mathcal{H}_2] \geq 0 \qquad \text{(II.27)}$$

where, $E[\eta/\mathcal{H}_j]$ is the expected cost of the strategy η conditional on $\mathcal{H}_j$ (see Eq. (II.21) and Eq. (II.25)).

The difference Δ_{12} is a measure of the relative importance of the models $H \in \mathcal{H}_2 \setminus \mathcal{H}_1$ in the decision problem at hand.

Assume for instance that $\mathcal{H}_2$ is obtained from $\mathcal{H}_1$ by adding a new model H. Then, H is significant for the optimal action if the error in Eq. (II.27) is large; in this case $\mathcal{H}_2$ should replace $\mathcal{H}_1$.

Thus, a search procedure leading to the final decision (metamodel $\mathcal{H}^{(opt)}$) might develop along the following steps:

- consider a metamodel $\mathcal{H}_1$; $\mathcal{H}_1$ is chosen only on subjective basis,
- enlarge $\mathcal{H}_1$ in various "directions" ("direction" here stands for type of model) and evaluate the error in (II.27) and retain (reject) the model $H \in \mathcal{H}_2 \setminus \mathcal{H}_1$ if it produces a large (small) cost variation Δ_{12} (Eq. (II.27))
- after all the "directions", say, contained in $\mathcal{H}^* \setminus \mathcal{H}_1$ have been checked the resulting reference set $\mathcal{H}^{(opt)}$ represents the optimal metamodel in the sense of Eq. (II.27) (and conditional on $\mathcal{H}^*$); consequently $\mathcal{H}^{(opt)}$ is linked to the specific decision problem considered by the statistician. Then, con-

ditional on $\mathcal{H}^{(opt)}$, the strategies $\eta_{III}^{(opt)}$ and $\eta_{IV}^{(opt)}$ are readily obtained (see Eq. (II.21), (II.29)).

Assume, for instance, that X(t) is a scalar independent sequence of yearly wind extremes and that, based on his experience, the statistician chooses $\mathcal{H}_1$ = {EX1}, where EX1 denotes the extreme distribution of type I. A first step might be to check the sensitivity of the optimal action when $\mathcal{H}_1$ is enlarged to $\mathcal{H}_2$ = {EX1, EX2}, where EX2 is the extreme type II distribution. Other directions of checking might be (see Fig. (II.7)) those of symmetrical models (e.g.: normal distribution), extreme (minima) models (e.g.: Weibull distribution), etc.; although these models seem not to conform to the data, they still can be important in a utility sense quantified in Eq. (II.27). Detailed analysis of this example is presented in Chap. III.

This approach is useful in decision problems where only a few models of X(t) seem to be significant. Plots of the expected cost of designing S at x=d versus d conditional on $\mathcal{H}_j$ (see Eq. (II.23)) may be useful to evaluate the relative importance of various models in the context of a specific decision problem.

The suboptimality of $\mathcal{H}^{(opt)}$ evaluated in this way comes from the following facts:

- not all models have been checked but only most of the significant types of models in $\mathcal{H}^*$ were considered.
- the algorithm evaluating $\mathcal{H}^{(opt)}$ is sequential; thus, the possibility exists to introduce in $\mathcal{H}^{(opt)}$ less significant models and to reject more significant models for a given decision problem.

- the strategy $\eta(\mathcal{H}_j)$ in Eq. (II.27) is only suboptimal (in the minimum expected cost sense; see Eq. (II.21)).

So, in this approximation one starts with a pivotal model/metamodel $\mathcal{H}_1 \subseteq \mathcal{H}^*$; then $\mathcal{H}_1$ is enlarged until the optimal strategy becomes insensitive to any other new model accepted in the analysis. The metamodel satisfying this property and which contains the minimum number of models, analytically simple is the (sub)optimal metamodel $\mathcal{H}^{(opt)}$.

The next approximation takes an opposite route, in the sense that here $\mathcal{H}^{(opt)}$ is the result of simultaneous comparison of various models $\mathcal{H}_i$ included in $\mathcal{H}^*$. In fact $\mathcal{H}^{(opt)}$ is the best action in the decision framework developed at metamodel level. It is pointed out that $\mathcal{H}^*$ may be very large since the evaluation of $\mathcal{H}^{(opt)}$ is performed only once (for given purpose) and since the decision framework giving $\mathcal{H}^{(opt)}$ does not bring numerical difficulties (see the development below).

2. Decision framework in $\mathcal{H}^*$ Consider first two metamodels $\mathcal{H}_1, \mathcal{H}_2$ ($\mathcal{H}_1$ and $\mathcal{H}_2$ may or may not have elements in common) corresponding to the load X(t) in Fig. (I.1); the problem is to evaluate which of the metamodels, $\mathcal{H}_1$ and $\mathcal{H}_2$, is the best in utility sense (the utility is linked, say, to the cost functions of the structural system S in Fig. (I.1)). For this purpose, the following statistical game (see (II.28)) is defined:

		"Nature"	
		$\mathcal{H}_1$	$\mathcal{H}_2$
Actions	$\mathcal{H}_1$	$\bar{L}(\mathcal{H}_1, \mathcal{H}_1)$	$\bar{L}(\mathcal{H}_1, \mathcal{H}_2)$
	$\mathcal{H}_2$	$\bar{L}(\mathcal{H}_2, \mathcal{H}_1)$	$\bar{L}(\mathcal{H}_2, \mathcal{H}_2)$

(II.28)

where, $\bar{L}(\mathcal{H}_i,\mathcal{H}_j) = c(\mathcal{H}_i) + E[d^{(i)}_{opt}/\mathcal{H}_j]$ (II.29)

denotes the expected loss for the statistician when he chooses action $\mathcal{H}_i$ instead of $\mathcal{H}_j$,

$$E[d^{(i)}_{opt}/\mathcal{H}_j] = \inf_d\{E[d/\mathcal{H}_j]\} \quad \text{(II.30)}$$

For example (see also Eq. (II.23) for the case when $\eta(H) = p''_H,\ H\varepsilon\mathcal{H}_j$)

$$E[d/\mathcal{H}_j] = c_S(d) + c_{F,S}\sum_{H\varepsilon\mathcal{H}_j}[1-F_H(d/E)]p''_H \quad \text{(II.31)}$$

is the expected cost of the design x=d corresponding to the Bayesian strategy $f_{B,j}(x) = \sum_{H\varepsilon\mathcal{H}_j} p''_H f_H(x/E)$ of the metamodel $\mathcal{H}_j$. The cost of operationg the metamodel $\mathcal{H}_i$ may have the form:

$$c(\mathcal{H}_i) = \sum_{H\varepsilon\mathcal{H}_i} c(H) \quad \text{(II.32)}$$

where c(H) is the computational cost of using the predictive distribution $f_H(x/E)$, and the posterior probability, p''_H, of H, in this decision framework.

From Eq. (II.28) the expected cost, $E[\mathcal{H}_i]$, of action $\mathcal{H}_i$ is readily obtained

$$E[\mathcal{H}_i] = \sum_j \bar{L}(\mathcal{H}_i,\mathcal{H}_j)p''_{\mathcal{H}_j} \quad \text{(II.33)}$$

$$\text{where, } p''_{\mathcal{H}_j} \alpha \sum_{H\varepsilon\mathcal{H}_j} p''_H \quad \text{(II.34)}$$

and it is said that $\mathcal{H}_1$ is better than $\mathcal{H}_2$ if

$$E[\mathcal{H}_1] \leq E[\mathcal{H}_2] \quad \text{(II.35)}$$

The extension of the decision problem (II.28) to the more general

case of evaluating the optimal metamodel $\mathcal{H}^{(opt)}$ is straightforward. Let $\mathcal{H}$ and $\mathcal{H}^*$ be the complete metamodel and the operational metamodel, respectively, of the input variable X(t). The expected cost of action $\mathcal{H}_i$ (see (II.28) and Eq. (II.23)) has the form:

$$E[\mathcal{H}_i] = \bar{L}(\mathcal{H}_i, \mathcal{H}^*) p''_{\mathcal{H}^*} + \bar{L}(\mathcal{H}_i, \hat{\mathcal{H}}) p''_{\hat{\mathcal{H}}} \qquad \text{(II.36)}$$

$$\text{where, } \hat{\mathcal{H}} = \mathcal{H}^* \setminus \mathcal{H}$$

From Eq. (II.29), Eq. (II.30), and Eq. (II.31) it follows that $E[\mathcal{H}_i]$ in Eq. (II.36) is the (code) expected cost corresponding to the Bayesian strategy $f_{B,i}(x) = \sum_{H \in \mathcal{H}_i} p''_H f_H(x/E)$ of $\mathcal{H}_i$ (see Eq. (II.17) and Eq. (II.21)). Consequently, $\mathcal{H}^{(opt)}$ is defined theoretically by the Bayesian strategy of type $f_{B,i}(x)$ which minimizes (II.36). Since only metamodels in $\mathcal{H}^*$ are practically considered in the analysis, $\mathcal{H}^{(opt)}$ is only suboptimal. However, if <u>either</u> $\bar{L}(\mathcal{H}_i, \hat{\mathcal{H}})$ does not depend on $\mathcal{H}_i \subseteq \mathcal{H}^*$ and the penalty of using a metamodel not in $\mathcal{H}^*$ is infinity <u>or</u> $p''_{\hat{\mathcal{H}}}$ is zero, Eq. (II.36) becomes:

$$E[\mathcal{H}_i] \alpha \bar{L}(\mathcal{H}_i, \mathcal{H}^*) \qquad \text{(II.37)}$$

$$\text{with, } \mathcal{H}_i \subseteq \mathcal{H}^*$$

and the consideration only of metamodels in $\mathcal{H}^*$ provides the optimal metamodel $\mathcal{H}^{(opt)}$.

In summary the decision framework at metamodel level developed here can be used to compare two metamodels (see Eq. (II.35)) and to evaluate the (sub)optimal metamodel, $\mathcal{H}^{(opt)}$ (see Eq. (II.37)).

The essential advantage of the metamodel-decision framework consists in the capability of the methodology to compare simultaneously very many models using as input information only the characterization of each

individual model. In this way the (sub)optimal set of models, $\mathcal{H}^{(opt)}$, corresponding to a specific engineering problem is readily identified and then used by code in the evaluation of the (sub)optimal strategy, $\eta^{(opt)}$, to be transferred to designer.

It may be observed that the calculation of $\mathcal{H}^{(opt)}$ is not always required. For example, when the code committee considers composite actions the evaluation of the optimal code strategy directly from $\mathcal{H}^*$ becomes soon numerically prohibitive when the size of $\mathcal{H}^*$ increases. However this is not necessarily so if only pure actions are considered. Consequently $\mathcal{H}^{(opt)}$ is recommended to be evaluated when $\mathcal{H}^*$ is a large set and composite actions are accepted by the code.

The decision framework in Eq. (II.37) may be extended to the general problem of code optimization. Following [67] and [68] the code is viewed as a mapping of known functional form with several free parameters, $\underline{p}$, defined on the space of structural topology, material, etc. with values in the space of structural member dimensions. The optimization is done in two steps: first, conditional on the metamodel, $\mathcal{H}_i$, of the design variable X(t) the best value of the code parameters, $\underline{p}(\mathcal{H}_i)$, is computed and second, by comparing the optimal cost of code corresponding to various $\mathcal{H}_i$, the best metamodel, $\mathcal{H}^{(opt)}$ and the best code parameters, $\underline{p}(\mathcal{H}^{(opt)})$, result.

The next section analyses the advantages and the limitations of the application of the generalized Bayesian Decision Theory developed in this section to systems design.

II.3 CONCLUSIONS

The evaluation of the optimal design, $d^{(opt)}$, of the system S in Fig. (I.1)was achieved in two phases. First,the optimal metamodel,$\mathcal{H}^{(opt)}$, and the optimal code strategy, f(x), (see various choices of f(x) in Table (II.38)) were calculated at code level. Second, $d^{(opt)}$ was computed in accordance with the code format (see Eq. (I.3) and Eq. (I.4)).

The sources of suboptimality of $d^{(opt)}$ are the following:

- suboptimality of the reference set $\mathcal{H}^{(opt)}$,
- suboptimality of the code strategy, $\eta^{(opt)}$; conditional on $\mathcal{H}^{(opt)}$ and A_c, $\eta^{(opt)}$ is the optimal (at code level) in the expected cost sense, say, of Eq. (II.21), and
- assume that $\eta^{(opt)}$ is optimal for the code C. Then, since the cost functions of S are different generally from the cost functions of C, $\eta^{(opt)}$ is not optimal for S (where S is any structural system controlled by C).

The analysis in Sec. II.2 has some limitations in that it does not allow one to evaluate the unconditional expected cost associated with the optimal strategy $\eta^{(opt)}$, or the difference between the suboptimal design level of S recommended by $\eta^{(opt)}$ (see for instance Eq. (II.22)) and the actual optimal level.

However, the essential feature of the methodology in Sec. II.2 lies in its capability of analyzing and comparing a very broad class of models when $\mathcal{H}^{(opt)}$ is evaluated.

The next chapters deal mostly with the solution of the Decision Theory problem for various types of models; specifically, Chap. III dis-

Table (II.38a)

Code Format	Code Specifications		
	$\tilde{A}$ = space of all densities	**A** = space of pure strategies $f_{\eta_i}(x/E) = f_{H_i}(x/E)$ (see Eq (II.15) for $\eta=\eta_i$)	
		$f_{H_i}(x/E)$=Bayesian PDF	$f_{H_i}(x/E)=f_{H_i}(x/\hat{\underline{\theta}}_{H_i})$ Estimate of PDF
Level III (Eq(II.25))	Not considered explicitly in this study	criterion: P_f $f_{III}(x)=f_{H'_{III}}(x/E)$	criterion: P_f $f_{III}(x)=f_{H''_{III}}(x/\hat{\underline{\theta}}_{H''_{III}})$
Level IV (Eq(II.21))		criterion: F.U.O. $f_{IV}(x)=f_{H'_{IV}}(x/E)$	criterion: F.U.O. $f_{IV}(x)=f_{H''_{IV}}(x/\hat{\underline{\theta}}_{H''_{IV}})$

The following are the notations (see also the next page):

(i) The Bayesian PDF of H is:

$$f_H(x/E) = \int f_H(x/\underline{\theta}_H) f_H(\underline{\theta}_H/E)\, d\underline{\theta}_H$$

where $f_H(\underline{\theta}_H/E)$ is the posterior PDF of the unknown parameters $\underline{\theta}_H$ of H.

(ii) The estimate of PDF of H is:

$$f_H(x/\hat{\underline{\theta}}_H) = f_H(x/\underline{\theta}=\hat{\underline{\theta}}_H)$$

(iii) P_f = code target failure probability

Table (II.38b)

Code Format	Code Specifications	
	A_c = space of composite strategies $f_\eta(x/E) = \sum_{i=1}^{m} \eta(H_i) f_{H_i}(x/E)$ (see Eq.(II.15))	
	$f_{H_i}(x/E)$=Bayesian PDF	$f_{H_i}(x/E)=f_{H_i}(x/\hat{\underline{\theta}}_{H_i})$ Estimate of PDF
Level III (Eq.(II.25))	criterion: P_f $f_{III}(x)= \sum_{i=1}^{m} \eta'_{III}(H_i) f_{H_i}(x/E)$	$f_{III}(x)=\sum_{i=1}^{m} \eta''_{III}(H_i) f_{H_i}(x/\hat{\underline{\theta}}_{H_i})$
Level IV (Eq.(II.21))	criterion: F.U.O. $f_{IV}(x)= \sum_{i=1}^{m} \eta'_{IV}(H_i) f_{H_i}(x/E)$	criterion: F.U.O. $f_{IV}(x)= \sum_{i=1}^{m} \eta''_{IV}(H_i) f_{H_i}(x/\hat{\underline{\theta}}_{H_i})$

(iv) F.U.O. = Formal utility optimization

(v) η'_j, η''_j, H'_j, H''_j denote the optimal strategy $\eta_j^{(opt)}$, j=III, IV, in each specific case.

cusses the case of independent sequences with normal, gamma and Poisson distributions, Chap. IV treats Gauss (log normal) Markov Sequences and Chap. V aims to find the optimal strategy in renewal sequences of random magnitudes.

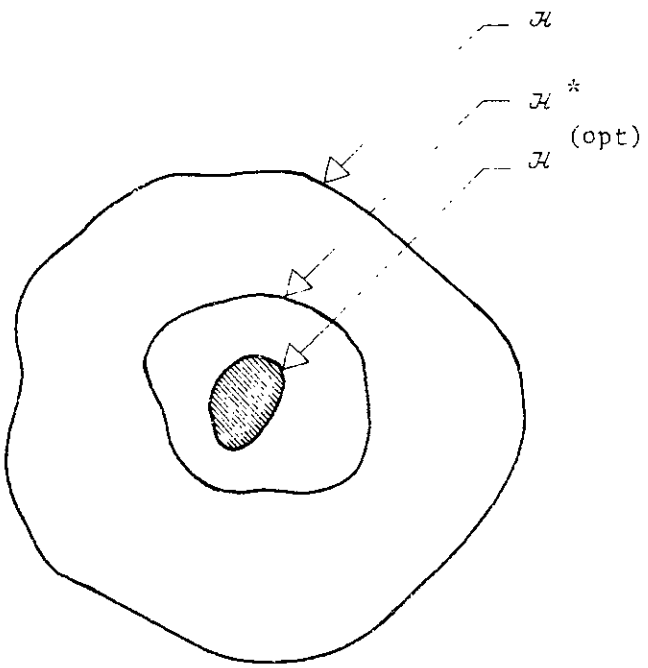


Fig. II.1. Metamodels of the input variable $\underline{X}(t)$

$\mathcal{H}$ = the complete metamodel of $\underline{X}(t)$

$\mathcal{H}^*$ = the operational metamodel of $\underline{X}(t)$

$\mathcal{H}^{(opt)}$ = the optimal metamodel of $X(t)$

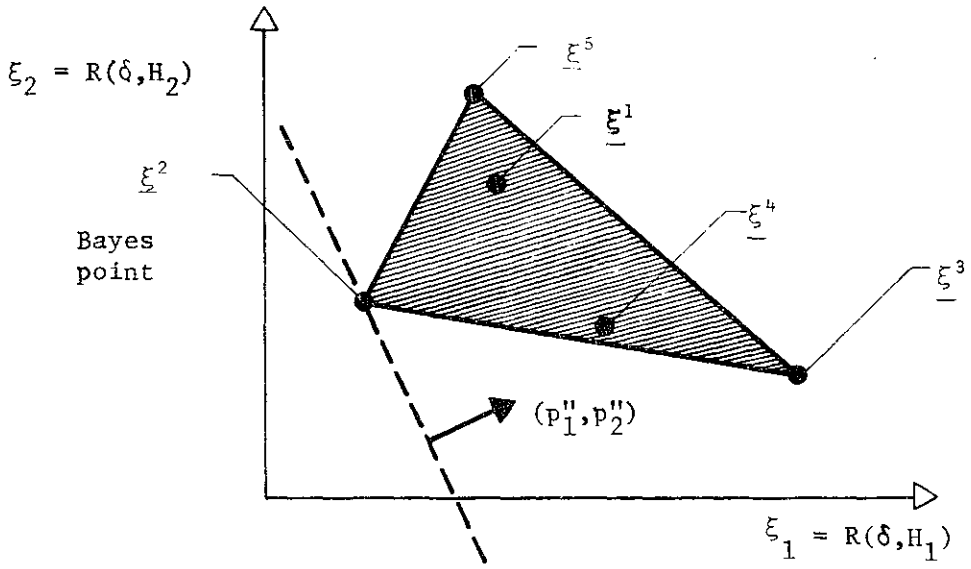


Fig. II.2. Hypothetical randomized and nonrandomized risk sets of A, when A and are finite. Illustration for $A = \{a_i\}_{1\le i\le 5}, \mathcal{H} = \{H_1,H_2\}$.

$\{\underline{\xi}^i\}_{1\le i\le 5}$ = the nonrandomized risk set of A

The triangle Δ_{2-3-5} = the randomized risk set of A

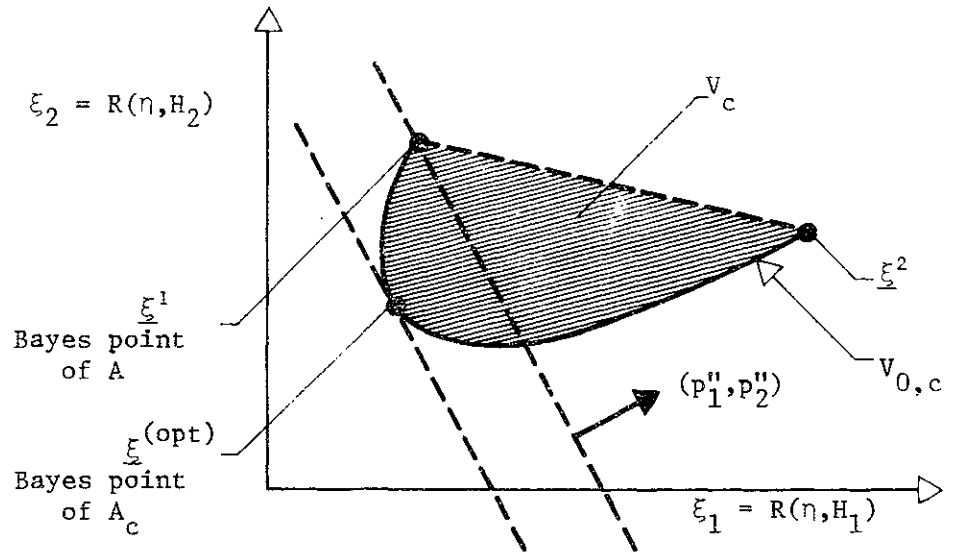


Fig. II.3. Hypothetical randomized and nonrandomized risk sets of A and A_c when A and are finite. Illustration for $A = \{a_1,a_2\}, \mathcal{H} = \{H_1,H_2\}$.

$\{\underline{\xi}^1,\underline{\xi}^2\}$ = the nonrandomized risk set of A

The segment $\underline{\xi}^1\underline{\xi}^2$ = the randomized risk set of A

The curve $V_{0,c}$ = the nonrandomized risk set of A_c

The region V_c = the randomized risk set of A_c

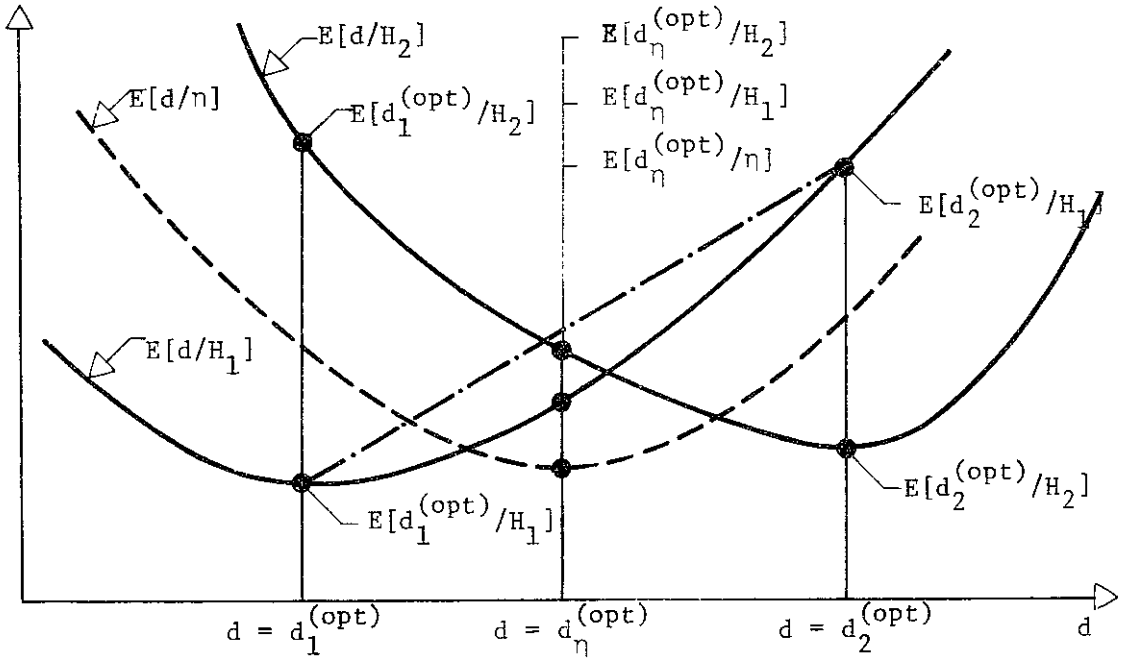


Fig. II.4. Typical expected cost functions, $E[d/\eta]$, in Eq. (II.19)
η = code strategy
$d_\eta^{(opt)}$ = the optimal design level under η (Eq.(II.20)).

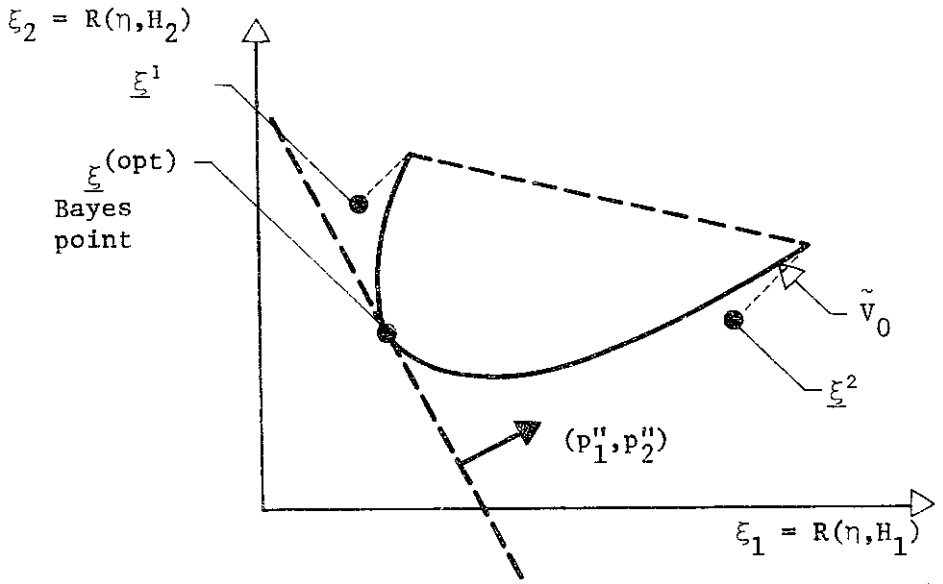


Fig. II.5. Hypothetical nonrandomized risk set of A_c when A and are finite: Illustration for the case when $A = \{a_1,a_2\}$, $\mathcal{H} = \{H_1,H_2\}$. and when the penalty for complexity is considered.
$\{\underline{\xi}^1 \cup \tilde{V}_0 \cup \underline{\xi}^2\}$= the nonrandomized risk set of A_c (compare with Fig. (II.3)).

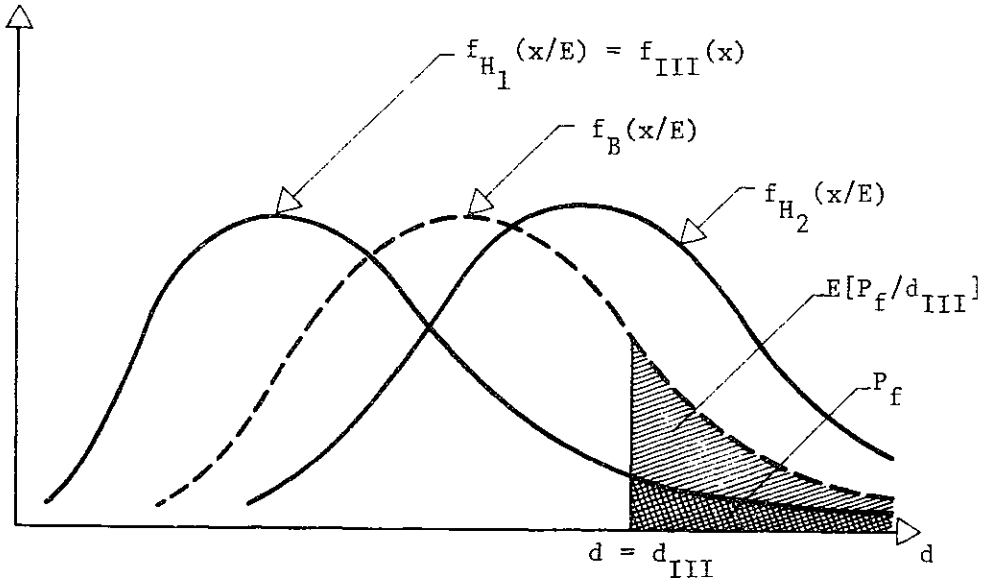


Fig. II.6. Hypothetical expected failure probability of S designed in accordance with a level III code.

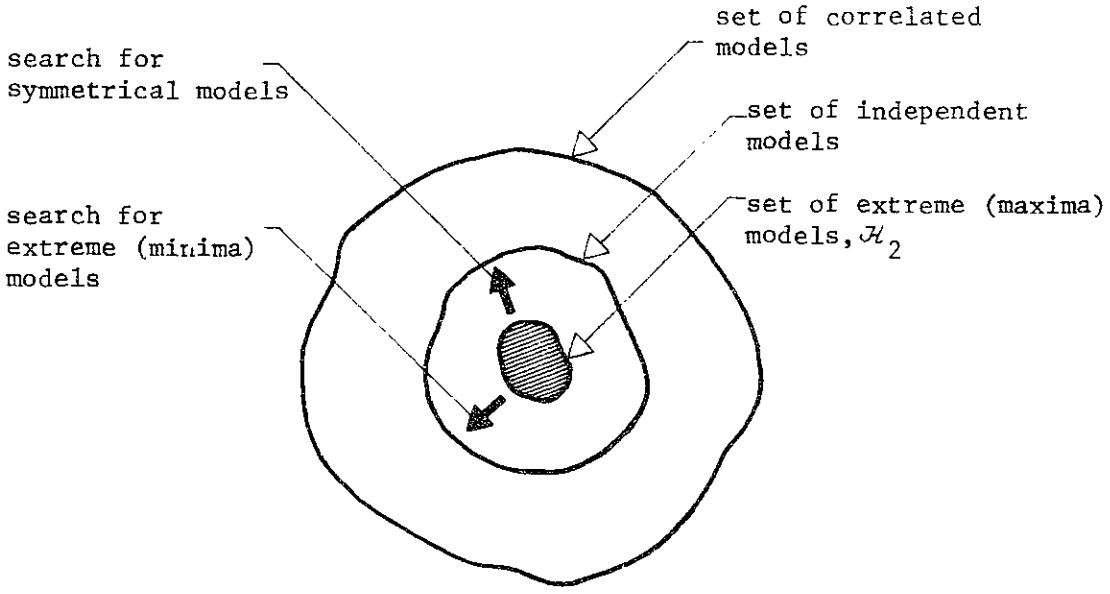


Fig. II.7. Searching for the optimal metamodel.

CHAPTER III

INDEPENDENT MODELS

III.1 INTRODUCTION

Let S be a deterministic two states (failure/survival) system and X(t) be the input variable of S (see Fig. (I.1)). This chapter's main goal is to evaluate the optimal (in the sense defined by Eq. (I.3) or Eq. (I.4)) design level, $d^{(opt)}$, of S when the input X(t) is a partially known stationary independent (or white) sequence (see Fig. (III.1)).

The complete probabilistic characterization of X(t) in Fig. (III.1) is given by the time step of the sequence, Δ, and by the marginal distribution of X(t); since Δ is considered perfectly known (for example if X(t) is the yearly wind extremes process, then Δ = 1 year), the time step of X(t) is taken always as one without loss of generality. Thus, the uncertainty in X(t) is related only to the degree of knowledge about the marginal distribution of the input variable.

Following Sec. II.2 the optimal design of S is reached in two phases. First, at code level, the operational metamodel, $\mathcal{M}^*$, of X(t) is postulated and then, if necessary (see Table (II.38)) the optimal metamodel, $\mathcal{M}^{(opt)}$,

is calculated.

The ultimate goal of this phase is to find the optimal (pure or composite) code strategy f(x). Second, at designer level, f(x) is processed following the rules of the code format (see Table (I.1) for level III and level IV code) to evaluate the optimal design level of any individual system controlled by the code.

The organization of this chapter is the following: Sec. III.2 produces the ingredients required by the decision framework in Sec. II.2. Thus, Sec. III.2.1 presents the general scheme of the prior to posterior analysis in stationary independent sequences, Sec. III.2.2 calculates predictive and posterior probabilities of Gaussian, Gamma, and Poisson stationary white sequences; also Sec. III.2.2.4 considers various problems which may appear when a noninformative prior is used for the unknown parameters of the models H in $\mathcal{H}^*$. Sec. III.2.3 presents briefly several cost functions which may be considered in the decision framework.

Sec. III.3 develops two applications related to the design of structural systems of the same type as S in Fig. (I.1) against wind. Sec. III.3.1 considers the controversial problem of what should be the distribution of the yearly wind speed extremes, $F_{year}(x)$; as is known $F_{year}(x)$ is assumed now to be of extreme type I in Canada and of extreme type II in U.S.A.

It is shown that a full inferential approach to model selection (using, say, maximum likelihood criterion) may lead to distributions which may predict very low or very large design values for S; this is the drawback of having generally data only in the body of the distribution so that almost any model in $\mathcal{H}^*$ may maximize the likelihood for a certain out-

come $\underline{Z}$. The utility approach to model selection is proposed instead; consequently, the code is not supposed to give the "true" distribution of the yearly extremes but the optimal strategy f(x) (see Table (I.1)) linked to a chosen utility and evidence $\underline{Z}$.

Sec. III.3.2 calculates the level III design, d_{III}, of the structural system in Fig. (I.1) subjected to the correlated input sequence X(t) (X(t) is, say, the mean hourly wind speed sequence) when the probability of failure P_f (see Table (I.1)) is small. It is concluded that the correlation structure of X(t) is much less significant for d_{III} than the marginal distribution of X(t) is; this conclusion seems not to be in perfect agreement with the methods in [19] and [32] which allow one to account for the correlation of X(t) when evaluating d_{III}.

Sec. III.4 discusses some of the practical implications of the results obtained in this chapter.

III.2 DECISION STRATEGY IN INDEPENDENT SEQUENCES

III.2.1 Introduction

Consider the design problem in Fig. (I.1) and assume that the input variable X(t) is a white stationary sequence. The final goal of the analysis is to evaluate the optimal (say, in the utility sense expressed by Eq. (I.4)) design level, $d^{(opt)}$, of S.

The Bayesian decision framework in Sec. II.2 is used to calculate $d^{(opt)}$; the ingredients of the methodology for the specific case of independent stationary input are the following:

1. the Bayesian description of the operational metamodel $\mathcal{H}^* = \{H\}$, where H stands for the distributional form of X(t), $f_H(x/\underline{\theta}_H)$, defined up to the unknown parameters $\underline{\theta}_H$. The prior uncertainty on the models in $\mathcal{H}^*$ is quantified by the probabilities p'_H ($\sum_{H\in\mathcal{H}^*} p'_H = 1.0$, and $p'_H \geq 0$ for all H in $\mathcal{H}^*$) of H being true and the prior belief in various possible values of the parameters $\underline{\theta}_H$ of H modeled by the prior PDF $f'_H(\underline{\theta}_H)$.

2. the experiment E which consists here of perfect measurements of X(t) during a finite period of time T; consequently the sample information is gathered in

$$\underline{Z}^t = \{Z_1 = X(1),\ Z_2 = X(2),\ \ldots\ ,\ Z_T = X(T)\} \qquad \text{(III.1)}$$

also, referred to as the observation or the measurement vector.

3. the loss (utility) function quantifies the purpose of the specific engineering problem considered in the analysis.

It was shown in Sec. II.2 that the consideration of the space of composite strategies, A_c, is advisable in most engineering problers. Consequently, both spaces of pure and composite strategies are considered in this chapter. (However, the use of composite actions is not always justified; see for example, Chap. IV.)

The inferential step of the decision approach is solved in Sec. III.2.2. The scheme of the procedure that provides posterior and predictive results is developed in [92] and is briefly presented here. First ([92], Chap. II) conditional on H being the true model of X(t), the posterior distribution, $f''_H(\underline{\theta}_H)$, of $\underline{\theta}_H$ is readily calculated from Bayes' rule (applied for the prior PDF $f'_H(\underline{\theta}_H)$ and the informative experiment $\underline{Z}$ in Eq. (III.1)).

$$f''_H(\underline{\theta}_H) = f_H(\underline{\theta}_H/\underline{Z}) \propto f'_H(\underline{\theta}_H)\, f_H(\underline{Z}/\underline{\theta}_H) \tag{III.2}$$

$$\text{where, } f_H(\underline{Z}/\underline{\theta}_H) = \prod_{t=1}^{T} f_H(Z_t/\underline{\theta}_H) \tag{III.3}$$

is the likelihood function of H at $\underline{Z}$.

The predictive distribution, $f_H(\tilde{\underline{y}}_\tau/\underline{Z})$, of H is the solution of the integration

$$f_H(\tilde{\underline{y}}_\tau/\underline{Z}) = \int f_H(\tilde{\underline{y}}_\tau/\underline{\theta}_H)\, f''(\underline{\theta}_H)\, d\underline{\theta}_H \tag{III.4}$$

$$\text{where } \tilde{\underline{y}}^t_\tau = \{\tilde{y}_1 = x(T+1),\ \tilde{y}_2 = x(T+2),\ldots,\ \tilde{y}_\tau = x(T+\tau)\} \tag{III.5}$$

is the vector of the next τ future realizations of the input variable X(t).

When $\tau=1$, $f_H(\tilde{\underline{y}}_\tau/\underline{Z}) = f_H(\tilde{y}/\underline{Z})$ is the <u>simple</u> predictive distribution of H and when $\tau>1$, $f_H(\tilde{\underline{y}}_\tau/\underline{Z})$ is the <u>simultaneous</u> predictive distribution of H; analytical solutions of $f_H(\tilde{\underline{y}}_\tau/\underline{Z})$ in Eq. (III.4) in closed form have been obtained for $\tau=1$ and distributions on $\underline{\theta}_H$ of conjugate type in [1,2, 36,37,92] for several probability laws. The evaluation of $f_H(\tilde{\underline{y}}_\tau/\underline{Z})$ where $\tau>1$ leads generally to numerical difficulties, even for small τ. When dealing with extreme values, however, an approximation is available. Veneziano ([84], p. 367) noticed that if X(t) is a normal white sequence, say, with unknown mean and variance, then the predictive probability $P(\tilde{y}_1 \le d, \ldots, \tilde{y}_\tau \le d/\underline{Z}) = F_{H,\tau}(d/\underline{Z})$ is well approximated by:

$$F_{H,\tau}(d/\underline{Z}) \cong [F_H(d/\underline{Z})]^\tau \tag{III.5}$$

where, $F_H(x/\underline{Z})$ is the simple predictive CDF ($\tau=1$ in Eq. (III.4)), provided the level x=d is sufficiently large, say, $F_H(d/\underline{Z}) \ge 0.99$. The

fact that the equality (III.5) is accurate only for high levels is not a significant restriction since in most of the reliability problems (see for example the design of S in Fig. (I.1)) the level x=d is high enough to satisfy Eq. (III.5).

The independence assumption leading to approximation (III.5) is accepted for all the models of X(t) considered in this chapter; consequently, Eq. (III.4) will be explicitly solved only for τ=1.

Clearly, when the size T of $\underline{Z}$ goes to infinity (i.e. the case of perfect knowledge) the distribution $f''_H(\underline{\theta}_H)$ becomes a unit spike located at the actual value, $\underline{\theta}_H^{(0)}$, of $\underline{\theta}_H$ and Eq. (III.5) is exactly satisfied for all d. This also happens when instead of working with $f''(\underline{\theta}_H)$ the point estimate $\hat{\underline{\theta}}_H$ of $\underline{\theta}_H$ is used as if it were the true value of $\underline{\theta}_H$. Computationally, the use of point estimates of $\underline{\theta}_H$ is very attractive; however, when T is small $\hat{\underline{\theta}}_H$ has typically large variance and therefore the "predictive" distribution $f_H(\tilde{y}/\underline{\theta}=\hat{\underline{\theta}}_H)$ may not be accurate.

Then, the posterior probability, p''_H, of H being the true model of X(t) is calculated (see [92], Chap. X)

$$p''_H = c^{-1} p'_H f_H(\underline{Z}) \tag{III.6}$$

$$\text{where, } f_H(\underline{Z}) = \int f_H(\underline{Z}/\underline{\theta}_H) f'_H(\underline{\theta}_H) d\underline{\theta}_H \tag{III.7}$$

Also, $f_H(\underline{Z})$ may be obtained from the ratio (see Eq. (III.2))

$$f_H(\underline{Z}) = \frac{f'_H(\underline{\theta}_H) f_H(\underline{Z}/\underline{\theta}_H)}{f''_H(\underline{\theta}_H)} \tag{III.7'}$$

$$\text{and } c = \sum_{H\in\mathcal{H}^*} p''_H \tag{III.8}$$

From Eq. (III.6) and Eq. (III.8) it follows that p''_H has to be inter-

preted as the probability of H conditional on $\mathcal{H}^*$; the unconditional probability of H results if the normalized constant is $c = \sum_{H \in \mathcal{H}} p''_H$, where $\mathcal{H}$ is the complete metamodel of X(t) (see Sec. II.2).

After the inferential step is completed, the optimal decision is readily evaluated (see Sec. III.2.3) for several cost functions of interest in Structural Reliability.

III.2.2 Prior to Posterior Analysis; Prediction Analysis

The inferential results required by the decision approach in Sec. II.2 are obtained in this section for various models (distributions) of the white sequence X(t). Thus, Sec. III.2.2.1, Sec. III.2.2.2, and Sec. III.2.2.3 find the solutions of Eq. (III.2) to Eq. (III.6) for Normal, Gamma, and Poisson Sequences, respectively when informative (noninformative) prior distributions of conjugate (Jeffrey's [47]) kind are assumed. Sec. III.2.2.4 considers the implications of using noninformative prior distributions on $\underline{\theta}_H$ in the calculation of p''_H in Eq. (III.6).

III.2.2.1 Gaussian Sequence

Let m and 1/h be the mean and the variance of the normal input sequence X(t); X(t) is sampled during time T (see observation vector $\underline{Z}$ in Eq. (III.1)). The marginal distribution function, $f_N(x/m,h)$, of X(t) is

$$f_N(x/m,h) = \left(\frac{h}{2\Pi}\right)^{1/2} \exp\left\{-\frac{h}{2}(x-m)^2\right\} \qquad \text{(III.9)}$$

The posterior distribution of(m,h), the simple predictive distribution of X(t) and $f_N(\underline{Z})$ in Eq. (III.7) are found next for various cases of prior knowledge on the parameters (m,h) defining completely $f_N(x/m,h)$.

(1) unknown mean and variance -

The informative conjugate prior PDF on (m,h) is a normal-gamma distribution (see [66], p.55) with parameters (m', n', v', ν'), i.e. $(m/h) \varepsilon N(m', \frac{1}{hn'})$ and $h \varepsilon G2(v', \nu')$; see also Chap. 7 in [66]. It follows that the conjugate prior distribution on (m,h) is:

$$f'_N(m,h) = \left(\frac{hn'}{2\pi}\right)^{1/2} \exp\left[-\frac{hn'}{2}(m-m')^2\right] \frac{(v'\nu'/2)^{\nu'/2}}{\Gamma(\nu'/2)} h^{\frac{\nu'}{2}-1} \exp\left[-\frac{v'\nu'h}{2}\right]$$

$$\text{in which, } -\infty < m, m' < \infty, \text{ and } h, v', n', \nu' > 0 \qquad \text{(III.10)}$$

The posterior distribution of (m,h) for given $\underline{Z}$, is still normal-gamma with parameters (m'', n'', ν'', v'') depending on the sufficient statistic $(T, \hat{Z}, v)$, where

$$n'' = n' + T$$

$$\nu'' = \nu' + T$$

$$m'' = (n'm' + T\hat{Z})/n''$$

$$v''\nu'' = v'\nu' + n'm'^2 + (T-1)v + T\hat{Z}^2 - n''m''^2$$

$$\hat{Z} = \frac{1}{T}\sum_{t=1}^{T} Z_t \quad , \quad v = \frac{1}{T-1}\sum_{t=1}^{T}(Z_t - \hat{Z})^2 \qquad \text{(III.11)}$$

The simple predictive distribution of X(t), $f_N(\tilde{y}/\underline{Z})$ follows from Eq. (III.4) when $\tau=1$ and $f''_H(\underline{\theta}_H) = f''_N(m,h)$, [1,2,36]. It is obtained that the one-step predictive variable $\tilde{y}/\underline{Z}$ is Student t with parameters $(m'', n''/[(n''+1)v''], \nu'')$ (see [92], p. 366); consequently, the normalized variable

$$t_{\nu''} = (\tilde{y}-m'')/\left[v''(1+1/n'')\right]^{1/2} \qquad \text{(III.12)}$$

is standardized student t with ν'' degrees of freedom.

Two derivates of $t_{\nu''}$ in Eq (III.12) will be frequently used to calculate both the code and the designer optimal strategies (see also Sec. II.2). First, the probability, $P_{f,N(m,h)}(d)$, that $\tilde{y}/\underline{Z}$ exeeds the level x=d is

$$P_{f,N(m,h)}(d) = P(\tilde{y} > d/\underline{Z}) = 1 - F_{t_{\nu''}}\left(\frac{d - m''}{[v''(1+1/n'')]^{1/2}}\right) \qquad \text{(III.13)}$$

where, $F_{t_{\nu''}}(x)$ is the CDF of the standardized Student t distribution with ν'' degrees of freedom.

Second, the P_f-upper fractile, $d_{N(m,h)}(P_f)$, of $(\tilde{y}/\underline{Z})$ which has the form,

$$d_{N(m,h)}(P_f) = m'' + [v''(1+1/n'')]^{1/2} t_{\nu''};(1-P_f) \qquad \text{(III.14)}$$

where, $t_{\nu''};(1-P_f)$ is the P_f - upper fractile of the standardized Student t distribution with ν'' degrees of freedom.

If apriori the parameters (m,h) are completely unknown, then (see [47,92]) the distribution in (III.10) is replaced by the Jeffrey's (or vague, or noninformative) prior PDF

$$f_{N(m,h)}(m,h) \quad \alpha \quad 1/h \qquad \text{(III.15)}$$

The results (III.12) to (III.14) hold also in this case if in Eq. (III.11) the prior parameters $n' \to 0$, $\nu' \to -1$, and $v' \to 0$.

The distribution of the observation vector (see Eq. (III.7)) for the conjugate informative prior on (m,h) in Eq. (III.10) is:

$$f_{N(m,h)}(\underline{Z}) = (2\Pi)^{-(\nu''-\nu')/2} \left(\frac{n'}{n''}\right)^{\frac{1}{2}} \frac{(v'\nu'/2)^{\nu'/2}}{(v''\nu''/2)^{\nu''/2}} \frac{\Gamma(\nu''/2)}{\Gamma(\nu'/2)} \qquad \text{(III.16)}$$

where, the posterior parameters $\{m'',n'',v'',\nu''\}$ are calculated from $\{m',n',v',\nu'\}$ and $\underline{Z}$ in Eq. (III.11)

(2) known mean, unknown scale -

The informative conjugate prior PDF on the scale parameter h is a gamma-2 distribution with parameters (v',ν'), [66], i.e.

$$f'_N(h) = \frac{(v'\nu'/2)^{\nu'/2}}{\Gamma(\nu'/2)} \; h^{\nu'/2-1} \; e^{-\frac{1}{2}v'\nu'h} \qquad \text{(III.17)}$$

where $h,v',\nu' > 0$

so that the posterior distribution of h has the same form with parameters (v'',ν'') depending on the statistic (T,w) and on the initial choice (v',ν'):

$$\nu'' = \nu' + T$$

$$v''\nu'' = v'\nu' + Tw$$

$$w = \frac{1}{T} \sum_{t=1}^{T} (Z_t - m)^2 \qquad \text{(III.18)}$$

The simple predictive distribution is student t with parameters (m, $1/\nu''$, ν''), [1,2]; consequently, the normalized variable $t_{\nu''} = (\tilde{y}-m)/\sqrt{v''}$ is student t with ν'' degrees of freedom.

Thus, the probability that $(\tilde{y}/\underline{Z})$ is larger than x=d, is

$$P_{f,N(h)}(d) = P(\tilde{y}>d/\underline{Z}) = 1 - F_{t_{\nu''}}\left(\frac{d-m}{\sqrt{v''}}\right) \qquad \text{(III.19)}$$

and the P_f - upper fractile of $(\tilde{y}/\underline{Z})$ results (see also Eq. (III.14)):

$$d_{N(h)}(P_f) = m + \sqrt{v''}\, t_{\nu'';(1-P_f)} \qquad \text{(III.20)}$$

The vague prior situation corresponds to $f'_{N(h)}(h)$ as in Eq. (III.15). In this case the results (III.19), (III.20) hold if $\nu' \to 0$ and v' is finite in Eq. (III.18).

The unconditional likelihood of $\underline{Z}$ for $f'_{N(h)}(h)$ in Eq. (III.17) is:

$$f_{N(h)}(\underline{Z}) = (2\Pi)^{-(\nu''-\nu')/2} \frac{(v'\nu'/2)^{\nu'/2}}{(v''\nu''/2)^{\nu''/2}} \frac{\Gamma(\nu''/2)}{\Gamma(\nu'/2)} \qquad \text{(III.21)}$$

where, v'', ν'' are calculated in Eq. (III.18).

(3) unknown mean, known scale -

The informative conjugate prior corresponding to m is normal with mean m' and variance $1/(hn')$, [66], i.e.

$$f'(m) = \left(\frac{hn'}{2\Pi}\right)^{1/2} \exp\left[-\frac{hn'}{2}(m-m')^2\right] \qquad \text{(III.22)}$$

where, $-\infty < m, m' < \infty$, $n' > 0$

and the posterior mean $(m/\underline{Z}) \in N(m'', 1/(hn''))$ where m'' and n'' are obtained from (m',n') and the statistic $(T, \hat{Z})$ from equations

$$n'' = n' + T$$

$$n''m'' = n'm' + T\hat{Z}$$

$$\hat{Z} = \frac{1}{T} \sum_{t=1}^{T} Z_t \qquad \text{(III.23)}$$

The normalized simple predictive variable $(\tilde{y}-m'')/\sqrt{(1+1/n'')/h}$ is standardized normal, [1]; consequently (see Eq. (III.13) and Eq. (III.14))

$$P_{N(m)}(d) = P(\tilde{y}>d/\underline{Z}) = 1 - \Phi(\frac{d - m''}{\sqrt{(1+1/n'')/h}}) \qquad \text{(III.24)}$$

where, $\Phi(\cdot)$ is the CDF of the standard normal variable, and

$$d_{N(m)}(P_f) = m'' + \sqrt{(1+1/n'')/h}\ \Phi_{(1-P_f)} \qquad \text{(III.25)}$$

in which, $\Phi_{(1-P_f)}$ is the $(1-P_f)$ - fractile of $\Phi(\cdot)$.

The posterior results (III.24) and (III.25) corresponding to prior vagueness $f'(m) \alpha$ constant hold if in Eq. (III.23) n' goes to zero.

The unconditional distribution of $\underline{Z}$ (see Eq. (III.16)) corresponding to informative prior state of knowledge is:

$$f_{N(m)}(\underline{Z}) = (2\Pi)^{-(n''-n')/2} (\frac{n'}{n''})^{\frac{1}{2}} h^{(n''-n')/2} e^{-\frac{hn''v''}{2}} \qquad \text{(III.26)}$$

where, $n''v'' = n'm'^2 - n''m''^2 + (T-1)v + T\hat{Z}^2$, and

$$v = \frac{1}{T-1} \sum_{t=1}^{T} (Z_t-\hat{Z})^2 \qquad \text{(III.27)}$$

Consequently, when the scale parameter is known and the mean of X(t) is unknown, the necessary statistic $(T,\hat{Z})$ to calculate the predictive distribution $f_{N(m)}(\tilde{y}/\underline{Z})$ is not sufficient to evaluate $f_{N(m)}(\underline{Z})$. In fact for the distribution of $\underline{Z}$ in Eq. (III.26) the sufficient statistic is $\{T, \hat{Z}, v\}$ and the posterior parameters (n'',m'',v'') are found from Eq. (III.23) and (III.27).

III.2.2.2 Gamma Sequence

Consider now that the input sequence X(t) is gamma distributed with known shape parameter k and unknown decay parameter λ; consequently the distribution of $X(t) \in G(k,\lambda)$ has the form:

$$f_G(x/K,\lambda) = \frac{e^{-\lambda x}\lambda^k x^{k-1}}{\Gamma(k)} \qquad \text{(III.28)}$$

in which, $x \geq 0$, and $k,\lambda > 0$

The conjugate prior distribution of λ is gamma with parameters (s,μ), [66], and therefore the posterior PDF of λ is also gamma of parameters (a,b) depending on (s,μ), k and the sufficient statistic $(T,\hat{Z} = \frac{1}{T}\sum_{t=1}^{T} Z_t)$. The dependence is given in the equations (see also [1,2]):

$$a = Tk + s$$

$$b = T\hat{Z} + \mu \qquad \text{(III.29)}$$

On integrating Eq. (III.4) for $\tau = 1$, the simple predictive distribution of the white stationary gamma sequence X(t) (of known shape parameter k) is obtained; $(\tilde{y}/\underline{Z})$ is inverted-beta-2 distributed with parameters (k, a,b) (see [66], p. 221 and [1])

$$f_{G(\lambda)}(\tilde{y}/\underline{Z}) = \frac{b^a}{B(a,k)} \frac{\tilde{y}^{k-1}}{(\tilde{y}+b)^{a+k}} \qquad \text{(III.30)}$$

where, $B(\cdot,\cdot)$ is the beta function.

The probability that $(\tilde{y}/\underline{Z})$ (see Eq. (III.30)) exceeds x=d is

$$P_{G(\lambda)}(d) = P(\tilde{y}>d/\underline{Z}) = I_{b/(b+d)}(a,k) \qquad \text{(III.31)}$$

where, I is the incomplete beta function as tabulated by Pearson [64],

The P_f - upper fractile of $(\tilde{y}/\underline{Z})$ has the form

$$d_{G(\lambda)}(P_f) = b[1 - B(a,k;P_f)]/B(a,k;P_f) \qquad \text{(III.32)}$$

where, $B(a,k;P_f)$ is the P_f - upper fractile of the incomplete beta distribution i.e. it satisfies (see also, (III.31))

$$I_{B(a,k;P_f)}(a,k) = P_f.$$

When the prior PDF on λ is diffuse, (i.e. the prior distribution on λ is proportional to $1/\lambda$; see [11,12,47]), the previous results still hold if in (III.29) s and μ go to zero.

The density of the observation vector $\underline{Z}$ corresponding to the informative PDF (of conjugate type) is:

$$f_{G(\lambda)}(\underline{Z}) = \left(\prod_{t=1}^{T} Z_t\right)^{k-1} \frac{\Gamma(a)}{[\Gamma(k)]^T} \frac{\mu^s}{b^a} \qquad \text{(III.33)}$$

As in Eq. (III.26) the sufficient statistic $(T,\hat{Z}, \prod_{t=1}^{T} Z_t)$ to calculate (III.33) contains strictly the statistic one needs to find the predictive distribution of X(t).

In particular, when k=1 the sequence X(t) is exponentially distributed; so, the results in this section apply also for white exponential processes.

III.2.2.3 Poisson Sequence

Consider a Poisson Sequence of unknown occurance rate λ and let X(t) be the interarrival time process; consequently X(t) is an independent stationary sequence exponentially distributed, i.e.

$$f_E(x/\lambda) = \lambda e^{-\lambda x} \qquad \text{(III.34)}$$

in which, $\lambda > 0$ and $x \geq 0$

The informative conjugate prior on λ is gamma with parameters (s,μ), [66]. The posterior distribution of λ in terms of the sufficient statistics $(T,\hat{Z} = \frac{1}{T}\sum_{t=1}^{T} Z_t)$ is still gamma of parameters (a,b), where:

$$a = T + s$$

$$b = T\hat{Z} + \mu \qquad \text{(III.35)}$$

The density of the predictive variable $\tilde{y}/\underline{Z}$ having Poisson distribution of mean $\lambda\ell$ is of negative binomial form (see [66], p. 236) with parameters $(\ell/(b+\ell),a)$

$$f_{P(\lambda)}(\tilde{y}/\underline{Z}) = \frac{\Gamma(a+\tilde{y})}{\tilde{y}!\,\Gamma(a)} \left(\frac{b}{b+\ell}\right)^a \left(\frac{\ell}{b+\ell}\right)^{\tilde{y}} \qquad \text{(III.36)}$$

The probability of $(\tilde{y}/\underline{Z})$ exceeding k, $P_{P(\lambda)}(k)$, and the $(1-P_f)$ fractile of $(\tilde{y}/\underline{Z})$, $k_{P(\lambda)}(P_f)$, [2], are:

$$P_{P(\lambda)}(k) = P(\tilde{y}>k/\underline{Z}) = \begin{cases} I_{\ell/(\ell+b)}(k+1,a) & , \quad k+1 \geq a \\ 1-I_{b/(\ell+b)}(a,k+1) & , \quad k+1 < a \end{cases} \qquad \text{(III.37)}$$

and

$$k_{P(\lambda)}(P_f) = \min \{\gamma : I_{b/b+\ell)}(a,) \geq 1 - P_f\} - 1 \qquad \text{(III.38)}$$

where, I is the incomplete beta function tablulated in [64]

The posterior results (III.36), (III.37), and (III.38) corresponding to a noninformative prior have the same form; however, in this case, the parameters (a,b) are obtained from Eq. (III.35) with $\mu,s \to 0$.

From Eq. (III.7), the unconditional likelihood function corresponding to the conjugate prior information results:

$$f_{P(\lambda)}(\underline{Z}) = \frac{\mu^s}{b^a} \frac{\Gamma(a)}{\Gamma(s)} \qquad \text{(III.37)}$$

where, a,b are calculated in Eq. (III.35)

The results in this section are used in Chapter V; the development is presented here since the inter-arrival time generated by a Poisson process is a white stationary sequence.

The next section discusses the possibility of using a noninformative prior distribution (of Jeffrey's type) on $\underline{\theta}_H$ in the evaluation of the posterior probability p''_H (see Eq. (III.6)). It appears that, generally, improper prior PDF cannot be processed to calculate posterior probabilities of the models in $\mathcal{H}^*$.

III.2.2.4 Posterior Probabilities of Models

Assume that the optimal metamodel $\mathcal{H}^{(opt)}$ of X(t) (see Sec. II.2) has been obtained and consider the case when all prior distributions $f'_H(\underline{\theta}_H)$ in Eq. (III.7) are proper (i.e. $f'_H(\underline{\theta}_H)$ is informative for $\underline{\theta}_H$ for all H in

$\mathcal{H}^{(opt)}$); then, the posterior probability, p''_H, of any H in $\mathcal{H}^{(opt)}$ is obtained from (III.6). For instance, if $\mathcal{H}^{(opt)} = \{N(m,h), G(\lambda)\}$ where N(m,h) and G(λ) denote the normal model with unknown mean and variance and the gamma model with unknown decay parameter, respectively, the posterior probabilities of the models in $\mathcal{H}^{(opt)}$ are:

$$p''_{N(m,h)} = c^{-1} p'_{N(m,h)} f_{N(m,h)}(\underline{Z})$$

$$p''_{G(\lambda)} = c^{-1} p'_{G(\lambda)} f_{G(\lambda)}(\underline{Z}) \qquad \text{(III.40)}$$

where, $f_{N(m,h)}(\underline{Z})$ and $f_{G(\lambda)}(\underline{Z})$ are calculated in Eq. (III.16) and Eq. (III.33); $p'_{N(m,h)}$ and $p'_{G(\lambda)}$ are the prior belief in N(m,h) and G(λ), respectively. The constant c is chosen such that $p''_{N(m,h)} + p''_{G(\lambda)} = 1.0$.

To assign $f'_H(\underline{\theta}_H)$ is not always an easy task. Two cases may be distinguished. First, consider the situation when in addition to the measurements $\underline{Z}$ some other information, say regional statistics $\underline{S}'_R$ (see [85,86]), are available; then, the parameters of the prior conjugate PDF $f'_H(\underline{\theta}_H)$ of H are estimated from the $\underline{S}'_R$. In this case all the results obtained in this chapter hold.

Second, assume that there is no regional information in which case either an informative or a diffuse (subjective) prior PDF on $\underline{\theta}_H$ may be assigned. When choosing an informative prior PDF on $\underline{\theta}_H$ the difficulty is that one cannot generally find consistent $f'_H(\underline{\theta}_H)$ (they are not consistent in the sense that the elements of the family $\{f'_H(\underline{\theta}_H)\}$ do not correspond to a unique $\underline{S}'_R$ when the $f'_H(\underline{\theta}_H)$ are subjectively selected; also, one can

not choose first $\underline{S}'_R$ and then calculate $f'_H(\underline{\theta}_H)$ for all H since various $\underline{S}'_R$ may favour, in the decision framework, different models) can not be generally found. If instead the prior position is of vagueness about $\underline{\theta}_H$ of all H, then in most of the cases the posterior probabilities p''_H do not exist. This latter problem was also noticed by Halpern [38] in his Bayesian decision framework to optimal polynomial regression selection. Because of the peculiarity of the metamodel $\mathcal{H}^{(opt)} = \tilde{\mathcal{H}}$ in [38] (any model H in $\tilde{\mathcal{H}}$ is a degenerate version of the highest degree polynomial regression accepted in $\mathcal{H}$; see also Sec. IV.3.2.6) a solution of p''_H corresponding to diffuse prior $f_H^{(0)}(\underline{\theta}_H)$ was proposed. Halpern's approach seems to be limited, however, to $\tilde{\mathcal{H}}$-like metamodels.

Consider now another special case when the optimal metamodel $\tilde{\tilde{\mathcal{H}}} = \mathcal{H}^{(opt)}$ contains models, H, of the same distributional type, $f(x/\underline{\theta}_H)$ but having different values of the known parameters $\underline{\theta}_{H,k}$, where $\underline{\theta}_H^t = \{\underline{\theta}_{H,k}^t, \underline{\theta}_v^t\}$. For example, $\tilde{\tilde{\mathcal{H}}}$ might be a set of gamma models $\{G(k_i,\lambda)\}_{i=1,2,\ldots,m}$, where k_i is the known shape parameter and λ is unknown. (As an aside, the reader should observe that all our probelms could have been formulated to include alternative parameter values as well as alternative distributional type in the set H. Instead, throughout this work we have chosen to include only distributional types in H and to treat parameter uncertainty through $f'_H(\underline{\theta}_H)$ distributions conditional on H. The two approaches are equivalent).

Assuming in this special case that the prior distribution, $f'(\underline{\theta}_v)$, on $\underline{\theta}_v$ is the same for all H in $\tilde{\tilde{\mathcal{H}}}$, from Eq. (III.6) and Eq. (III.7') follows:

$$p''_H = c^{-1} p'_H \frac{f_H(\underline{Z}/\underline{\theta}_{H,K},\underline{\theta}_v)}{f''(\underline{\theta}_v)} \qquad \text{(III.41)}$$

Since p''_H in Eq. (III.41) does not depend explicitly on $f'_H(\underline{\theta}_H)$, the equation (III.41) holds also for diffuse prior on $\underline{\theta}_v$. The value of the ratio $R_H = f_H(\underline{Z}/\underline{\theta}_{H,K}, \underline{\theta}_v)/f''_H(\underline{\theta}_v)$ is given below for several $\tilde{\tilde{\mathcal{H}}}$-like meta-models:

(a) $\tilde{\tilde{\mathcal{H}}}$ is a set of Gamma dstributions: $\{G(k_i, \lambda)\}$

$$R_{G(k_i,\lambda)} \alpha \frac{\Gamma(Tk_i)}{[\Gamma(k_i)]^T} \frac{(\prod_{t=1}^{T} Z_t)^{k_i-1}}{(\sum_{t=1}^{T} Z_t)^{Tk_i}} \qquad \text{(III.42)}$$

(b) $\tilde{\tilde{\mathcal{H}}}$ contains normal models of known mean and unknown scale parameter: $\{N(m_i, 1/h)\}$

$$R_{N(m_i;1/h)} \alpha \sum_{t=1}^{T} (Z_t - m)^2 \qquad \text{(III.43)}$$

(c) $\tilde{\tilde{\mathcal{H}}}$ is a set of normal distributions with unknown mean but known scale parameter $N(m,1/h_i)$

$$R_{N(m,1/h_i)} \alpha h_i^{\frac{T-1}{2}} \exp\{-\frac{h_i}{2}[\sum_{t=1}^{T} (Z_t-\hat{Z})^2]\} \qquad \text{(III.44)}$$

In general, however, it must be concluded that when $\mathcal{H}^{(opt)}$ contains models of various PDF forms (and no regional information exists) then neither consistent inrormative prior PDF's, $f'_H(\underline{\theta}_H)$, nor posterior probabilities, p''_H, corresponding to vague prior PDF's $f_H^{(0)}(\underline{\theta}_H)$ exist. Consequently, it is recommended to use always informative prior PDF's of the conjugate type (in which case all the results in this chapter hold) corresponding either to $\underline{S}'_R$ when there is regional information or to subjective know-

ledge, say, containing little information (e.g., $f'_H(\underline{\theta}_H)$ is flat over a certain range of outcomes of X(t)) where there is no regional prior knowledge. In the latter case, although $\{f'_H(\underline{\theta}_H)\}$ are not consistent in the sense defined above, they may not influence significantly the analysis in most practical situations.

These last results complete the inferential step required by the decision approach in Sec. II.2. The next section considers several cost functions linked to the reliability problem of designing S in Fig. I.1.

Sec. III.3 presents two examples which are doubly important: first, since the theoretical development in Sec. II.2 is illustrated and second, due to the practical results related to the design of structural systems against wind.

III.2.3 Loss Function

The utility function leading to the selection of the code strategy f(x) (see Sec. II.2) is defined by the penalty, c_f, associated with the complexity of f(x) (c_f is a function of the analytical form of f(x) and on the criterion (code format) to calculate the optimal design of the system S ruled by the code), the initial cost, c(d), of designing S at x=d, and $c(\tilde{y};d)$. The last component of the loss gives the penalty when the design level is x=d and the future outcome of the input is $\tilde{y}$. The functions c(d) and $c(\tilde{y};d)$ are deterministic and completely defined by the type of structure, material quality, workmanship, cost of failure, etc.

Let c(d) and $c(\tilde{y};d)$ be the cost functions of the system S; then the expected cost of the design x=d, E[d], under code strategy f(x) is (see also Chap. I)

$$E[d] = c(d) + \int_{-\infty}^{\infty} c(\tilde{y};d)\, f(\tilde{y})\, d\tilde{y} \qquad \text{(III.45)}$$

Typically in Structural Reliability, the cost functions $c(d)$ and $c(\tilde{y};d)$ are chosen (see Esteva [25]) in the form:

$$c(d) = \alpha + \beta\, d^{k} \qquad \text{(III.46)}$$

$$k=1,2,\ldots$$

and

$$c(\tilde{y};d) = \gamma\, u(\tilde{y}-d) \qquad \text{(III.47)}$$

where $u(\cdot)$ is the unit step function located in the origin.

The cost $c(\tilde{y};d)$ in Eq. (III.47) also referred to as the failure cost, corresponds to an ideal system which collapses when $\tilde{y} > d$ and suffers no damages as long as $\tilde{y} \leq d$ (see Eq. (I.4)).

Hypothetical initial and failure cost functions are shown in Fig. (III.2).

When $c(\tilde{y};d)$ has the form in Fig. (III.2(b)) the calculation of the integral in Eq. (III.45) reduces to the evaluation of the probability, $P_f(d)$, of exceeding $x=d$; consequently, $E[d]$ is easily computed. All of the examples considered in the next section assume cost functions as shown in Fig. (III.2).

More accurate functions, $c(\tilde{y};d)$, may be considered but generally the computation of the expected cost in Eq. (III.45) increases substantially. For instance, one may take a failure cost $c(\tilde{y};d)$ linear in $\tilde{y}$ for $\tilde{y} \geq d$; in this case and for $f(x)$ being a Gaussian density the expected cost of

failure $\int_{\infty}^{\infty} c(\tilde{y};d)\, f(\tilde{y})d\tilde{y}$ is tabulated in [66] at p. 356.

The cost functions in Eq. (III.46) and (III.47) with $k = 4$ and various values of the parameters α, β, and γ are used in the wind examples developed in the next section.

III.3 APPLICATIONS

The design of a structural system S (see Fig. (I.1)) subjected to the wind load, X(t), is considered here. The yearly wind speed extremes are assumed to follow an extreme type I distribution in Canada and an extreme type II distribution in U.S.A. This disagreement is not necessarily a lack of consistency but may simply be the drawback of using inferential approaches to select the "optimal" model of X(t) from the observation of the input variable during the finite time T.

In Sec. III.3.1 the input X(t) is assumed to be a white stationary sequence extreme type I distributed with parameters $\alpha = 0.16$, and $\mu = 50$ (see various extreme distributions in [5], p. 273-286), values which correspond closely to the yearly wind speed extremes recorded in Boston for a period of 42 years (1933-1974). Following Sec. II.2, and assuming that the statistician learns about X(t) only from the observation vector $\underline{Z}$ (see Eq. (III.1)), the optimal metamodel, $\mathcal{H}^{(opt)}$, of X(t), the optimal level IV code strategy, $f_{IV}(x)$ (see Table (I.1)) and the best level IV design, d_{IV}, of S are obtained. Also, a brief exemplification of the computation providing the optimal level III design, d_{III}, of the structural system S is presented.

Sec. III.3.2 assumes the more general case when X(t) is a stationary correlated sequence. The goal here is to evaluate the sensitivity of, say, the level III design, d_{III}, of S to the correlation structure of X(t). It appears that when $P(X(t) \leq d_{III})$ is close to one (i.e. design in the upper tail of X(t)), the original input may be replaced by the stationary independent sequence $X^{(i)}(t)$ having the same marginal distribution as X(t). It is also shown that the essential element for evaluating d_{III} is the marginal distribution of X(t) or $X^{(i)}(t)$. Comparison between our approach and the solutions in [19] and [32] are also presented.

III.3.1 Example 1

Consider the problem of finding the optimal code strategy, f(x), when the input variable X(t) denotes the wind yearly speed extremes and the code controls the design of systems of the type described in Fig. (I.1).

The analysis in this section assumes that X(t) is an independent stationary sequence.

Typically, the optimal code strategy is obtained inferentially. First, a set of distributions, $\mathcal{H}^* = \{H\}$, of X(t) is postulated and then, say, the maximum likelihood model $H_{ML} \in \mathcal{H}^*$ is selected for X(t). This procedure may lead to either overdesign or underdesign situations.

To support this statement the following experiment was conducted. From an extreme type I distribution with $\alpha = 0.16$ and $\mu = 50$, ten sequences of independent samples of sizes T = 10 and T = 40 were drawn (each line in the Table (III.48) and Table (III.49) corresponds to one sequence of length T = 10 and T = 40, respectively); see for sample generation [51]. Assume that the statistician learns about X(t) only from the observation

vector and let $\mathcal{H}^*$ be the set $\{EX1(\hat{\alpha},\hat{\mu}), EX2(\hat{k},\hat{v}), N(\hat{m},\hat{h})\}$, where $EX1(\alpha,\mu)$ denotes the extreme type I distribution with parameters (α,μ), $EX2(k,v)$ is the extreme type II distribution with parameters (k,v) and $N(m,h)$ is the normal distribution with mean m and scale parameter h. The notations $(\hat{\alpha},\hat{\mu})$, $(\hat{k},\hat{v})$ and $(\hat{m},\hat{h})$ stand for the maximum likelihood estimates of (α,μ), (k,v), and (m,h), respectively. The last three columns in Tables (III.48) and (III.49) give quantities which are proportional to the likelihoods $f_{EX1}(\underline{z}/\hat{\alpha},\hat{\mu})$, $f_{EX2}(\underline{z}/\hat{k},\hat{v})$ and $f_N(\underline{z}/\hat{m},\hat{h})$. (From Eq. (III.6) and assuming a noninformative prior on the models in $\mathcal{H}^*$, i.e. $p'_{EX1} = p'_{EX2} = p'_N = \frac{1}{3}$, it results that the last three columns in (III.48) and (III.49) represent also the posterior probability of the models in $\mathcal{H}^*$ corresponding to a prior distribution $f'_H(\underline{\theta}_H)$ which weights with one the maximum likelihood parameters $\hat{\underline{\theta}}_H$ and with zero $\underline{\theta}_H \neq \hat{\underline{\theta}}_H$.

The last row in the Table (III.49) corresponds to the yearly wind speed extremes sequence recorded in Boston during 42 years (1933-74). For this sample the likelihood $f_H(\underline{z}/\underline{\theta}_H)$, $H \in \mathcal{H}^*$, is maximized by the $EX1(\hat{\alpha},\hat{\mu})$ distribution. This does not necessarily imply that the sequence of yearly wind extremes in Boston has an extreme type I distribution. It may be possible for the maximum likelihood model in $\mathcal{H}^*$ to vary when, say, the observation vector increases. In fact, from Tables (III.48) and (III.49) it appears that any of the models in $\mathcal{H}^*$ may be selected provided that the inferential rule requires one to choose for X(t) the maximum likelihood model. Consequently, when $N(\hat{m},\hat{h})$ is chosen, S may be underdesigned (N has thinner tail than EX1) and when $EX2(\hat{k},\hat{v})$ is selected, S is generally overdesigned (EX2 has thicker tail than EX1). This drawback of the inferential methodology is avoided in the utility approach proposed in

this work, which tries to achieve the best compromise between the purpose (here, to best design S in the expected cost sense) and the evidence (i.e. the observation $\underline{Z}$). The details of the utility approach developed in this study are found in Sec. II.2. The numerical examples considered next assume the operational metamodel $\mathcal{H}^*$ and the space of pure strategies $A = \{EX1(\hat{\alpha},\hat{\mu}), EX2(\hat{k},\hat{v}), N(\hat{m},\hat{h}), St(m'',n'',v'')\}$, where St(m'',n'',v'') is the predictive PDF of N(m,h) corresponding to diffuse prior on (m,h) (see Sec. III.2.2). The particular form (the maximum likelihood estimates, $\hat{\underline{\theta}}_H$, are used in the PDF of $H \in \mathcal{H}^*$ as it would be the actual value of $\underline{\theta}_H$) of the operational metamodel $\mathcal{H}^*$ considered in the following examples is chosen since the predictive PDF of EX1 and EX2 distributions do not have an analytical closed-form. Consequently, (simplicity argument, see Sec. II.2) the metamodels containing the predictive PDF of the extreme distributions are highly penalized. The predictive PDF of the Gaussian model is not used in $\mathcal{H}^*$ to insure equivalent treatment of the models in $\mathcal{H}^*$. Since the maximum likelihood estimates of the parameters of the EX1 and EX2 distributions have not an explicit solution [4,14,54,70,81,82], $\hat{\alpha}$, $\hat{\mu}$, $\hat{k}$, and $\hat{v}$ were numerically evaluated.

The penalty associated with the complexity of the models in $\mathcal{H}^*$ is not introduced explicitly in the analysis.

Numerical results are given for the particular case of a code, C, which controls the design of a set of systems $\{S_k\}$ having the same cost functions as S in Fig. (I.1). It follows that $\mathcal{H}^{(opt)}$ and f(x) are also "the best" (in the sense defined in Sec. II.2) with respect to only the utility of S in Fig. (I.1). Consequently, $\mathcal{H}^{(opt)}$ and f(x) are calculated only in function of the cost functions of S (see also [67,68]).

Table (III.48)

NT=10	STATISTICS			MAX. LIKE. PARAMETERS						MAX. LIKE. (normalized)		
				EX1		EX2		N				
Run #	SAMPLE MEAN	SAMPLE VARIANCE	SAMPLE PRODUCT	$\hat{\alpha}$	$\hat{\mu}$	$\hat{k}$	$\hat{v}$	$\hat{m}$	$\hat{h}$	EX1	EX2	N
1	53.57	87.9891	52.83	0.1317	49.35	6.9210	48.51	53.57	0.01263	0.3532	0.3216	0.3253
2	52.94	26.3659	52.71	0.2098	50.63	10.271	50.43	52.94	0.04214	0.2393	0.1619	0.5988
3	51.51	71.0647	50.51	0.1571	47.71	7.721	47.61	51.51	0.01564	0.4094	0.3817	0.2089
4	54.70	52.6293	54.26	0.1618	51.43	8.421	51.06	54.70	0.02111	0.3424	0.2547	0.4029
5	52.12	41.5603	59.78	0.2389	49.22	10.621	49.17	52.12	0.02673	0.4184	0.4734	0.1081
6	53.27	49.5095	52.89	0.2223	50.60	10.621	50.73	53.27	0.02244	0.4423	0.5016	0.0561
7	51.84	67.5017	51.20	0.1311	47.61	6.621	46.89	51.81	0.01646	0.2604	0.1048	0.6349
8	56.11	112.7582	55.30	0.1508	51.83	8.321	51.36	56.11	0.00985	0.3843	0.5555	0.0602
9	55.44	92.6233	54.75	0.1682	51.61	9.771	51.24	55.44	0.011996	0.3316	0.6112	0.0572
10	50.91	62.9323	50.37	0.1617	47.34	7.821	47.06	50.91	0.01766	0.3766	0.3398	0.2836

TABLE (III.49)

NT=40	STATISTICS			MAX. LIKE. PARAMETERS						MAX. LIKE. (normalized)		
Run #	SAMPLE MEAN	SAMPLE VARIANCE	SAMPLE PRODUCT	EX1		EX2		N				
				$\hat{\alpha}$	$\hat{\mu}$	$\hat{k}$	$\hat{v}$	$\hat{m}$	$\hat{h}$	EX1	EX2	N
1	53.18	56.2901	52.67	0.156	49.80	7.87	49.15	53.18	0.01822	0.4252	0.1588	0.4160
2	53.33	65.5513	52.77	0.1534	49.68	7.37	49.29	53.33	0.1565	0.8200	0.1348	0.0451
3	55.17	111.0625	54.27	0.1317	50.42	6.87	49.9	55.17	0.009235	0.4294	0.5669	0.0036
4	53.22	20.5449	53.03	0.2480	51.18	10.62	51.19	53.22	0.04992	0.3413	0.0486	0.6101
5	55.93	59.1282	55.44	0.1768	52.47	9,62	52.20	55.93	0.01735	0.3674	0.6221	0.01046
6	54.73	59.9615	54.21	0.1556	51.24	8.27	50.59	54.73	0.071	0.5802	0.2833	0.1365
7	53.57	54.5881	53.08	0.1486	50.25	7.27	49.52	53.57	0.01879	0.2285	0.0270	0.7445
8	52.20	85.8654	51.54	0.1684	48.53	8.37	48.25	52.20	0.01194	0.3586	0.6413	0.00003
9	53.25	73.7788	52.62	0.1543	49.39	7.57	48.72	53.25	0.01567	0.5835	0.3988	0.0176
10	53.30	65.4343	52.74	0.1536	49.66	7.57	49.26	53.30	0.01567	0.7397	0.2107	0.0496
Actual data in Boston (42 yrs)	52.23	69.6707	51.62	0.1587	48.47	7.721	48.28	52.23	0.01470	0.5508	0.4336	0.0156

The form of the cost functions of S is taken from [25] (see also Sec. III.2.3); the coefficients have been adjusted to be adequate for the design against wind. The scaled version of the expected cost E[d] in Eq. (I.4) has the initial cost function

$$c(d) = 0.02\ (d/70)^4 \tag{III.50}$$

For the failure cost, c_F, two values were assumed, i.e. 2 and 20.

The expected cost of C (or S in our specific case) corresponding to level x=d and systems of life time τ years when the code strategy is f(x) has the form:

$$E[d] = c(d) + c_F\{1 - [F(d)]^\tau\} \tag{III.51}$$

where, F(x) is the CDF of f(x)

If d is large, such that F(d) is close to one and $\tau[1-F(d)] << 1$, Eq. (III.51) may be approximated by:

$$E[d] \overset{\sim}{=} c(d) + c_F\tau[1 - F(d)] \tag{III.51'}$$

The calculation of the optimal metamodel, $\mathcal{H}^{(opt)}$, and of the optimal level IV code strategy, $f_{IV}(x)$, are presented next. In order to evaluate $f_{IV}(x)$, composite strategies consisting of only linear combinations of pair of pure actions (let A_c' denote their set) have been analysed.

First the one year system S (τ = 1 year) is considered for two failure costs c_F = 2, c_F = 20 and two observation vectors $\underline{Z}$ of sizes T = 10 and T = 40 (see their statistics in the first line of Table (III.48) and of Table (III.49), respectively). The optimal metamodel, the optimal code

strategy and the optimal design of S are obtained in each case. Then, several results when τ = 50 years, T = 40 and c_F = 2 and c_F = 20 are also presented.

Following Sec. II.2, the evaluation of the optimal level IV design, d_{IV}, of S in Fig. (I.1) is separated in two phases: (1) the code phase which calculates the optimal metamodel, $\mathcal{H}^{(opt)}$, of X(t) and the optimal code strategy, $f_{IV}(x)$, and (2) the designer phase which computes d_{IV} as the minimum point of E[d] in Eq. (III.51), where c(d) and c_F are the cost functions of S, and $F(x) = F_{IV}(x)$. ($F_{IV}(x)$ is the CDF of $f_{IV}(x)$.)

To illustrate the evaluation of $\mathcal{H}^{(opt)}$ we consider as possible candidates in $\mathcal{H}^*$ only the metamodels $\mathcal{H}_1 = \{EX1(\hat{\alpha},\hat{\mu}), EX2(\hat{k},\hat{v})\}$ and $\mathcal{H}_2 = \mathcal{H}^*$. The comparison between $\mathcal{H}_1$ and $\mathcal{H}_2$ may be also viewed as the check of the sensitivity of the optimal design when $\mathcal{H}_1$ is enlarged to $\mathcal{H}_2$.

Let $d_{opt}^{(i)}$ be the optimal design level of the Bayesian strategy, $H_{B,i} = \sum_{H\in\mathcal{H}_i} p''_H H$, of $\mathcal{H}_i$, i = 1,2; $d_{opt}^{(i)}$ is obtained as the minimum point of $E[d/\mathcal{H}_i]$ evaluated from Eq. (III.51) when $F_{IV}(x)$ is replaced by $F_{B,i}(x) = \sum_{H\in\mathcal{H}_i} p''_H F_H(x/\hat{\underline{\theta}}_H)$ (p''_H is the posterior probability of H conditional on $\mathcal{H}_i$ and $F_H(x/\hat{\underline{\theta}}_H)$ is the CDF of H for $\underline{\theta}_H$ equal to its maximum likelihood estimate $\hat{\underline{\theta}}_H$).

The expected cost, $E[d/\eta]$ (see Eq. (III.19)), of designing S at x=d under the strategies EX1, EX2, N, St and $H_{B,i}$ are plotted versus d in Fig. (III.3) to Fig. (III.8) and denoted with $1,2,3,4,5_{(i)}$, respectively. The optimal design values under $\eta(d_\eta^{(opt)})$, for η = EX1, EX2, N, St, and Bayesian, $(d_{opt}^{(i)})$, strategies are also shown.

From Eq. (II.28), the expected cost, $E[\mathcal{H}_i]$, of $\mathcal{H}_i$, i=1,2, results:

$$\begin{bmatrix} E[\mathcal{H}_1] \\ E[\mathcal{H}_2] \end{bmatrix} = \begin{bmatrix} E[d_{opt}^{(1)}/\mathcal{H}_1] & E[d_{opt}^{(1)}/\mathcal{H}_2] \\ E[d_{opt}^{(2)}/\mathcal{H}_1] & E[d_{opt}^{(2)}/\mathcal{H}_2] \end{bmatrix} \begin{bmatrix} p''_{\mathcal{H}_1} \\ p''_{\mathcal{H}_2} \end{bmatrix} \qquad \text{(III.52)}$$

(matrix multiplication)

$$\text{where,} \quad p''_{\mathcal{H}_1} \propto p''_{EX1} + p''_{EX2}$$
$$p''_{\mathcal{H}_2} \propto p''_{EX1} + p''_{EX2} + p''_{N} \qquad \text{(III.53)}$$

The expected cost $E[d_{opt}^{(i)}/\mathcal{H}_j]$ in Eq. (III.52) is the ordinate of the expected cost of the Bayesian strategy $H_{B,j}$, i.e. $E[d/\mathcal{H}_j] = 0.02(\frac{d}{70})^4 + c_F[1-F_{B,j}(d)]$, (the curve 5(j) in Fig. (III.3) to Fig. (III.8)) at $d = d_{opt}^{(i)}$. For example, when $c_F = 2$, $T = 10$ and $\tau = 1$, then (see Fig. (III.3)) $d_{opt}^{(1)} = 87.00$, $E[d_{opt}^{(1)}/\mathcal{H}_1] = 0.0716448$, and $E[d_{opt}^{(1)}/\mathcal{H}_2] = 0.0639200$. The posterior probabilities in Eq. (III.53) are obtained from the first line in Table (III.48); it follows that $p''_{\mathcal{H}_1} \propto 0.3532 + 0.3216 = 0.6748$ and $p''_{\mathcal{H}_2} \propto 1.0$.

The expected cost $E[\mathcal{H}_i]$ in Eq. (III.52) and the optimal design levels, $d_{opt}^{(i)}$, i=1,2 are given in Table (III.54) for the case when $\tau = 1$ year, $T = 10; 40$ and $c_F = 2; 20$.

Table (III.54)

	$c_F = 2$				$c_F = 20$			
	$E[\mathcal{H}_1]$	$d_{opt}^{(1)}$	$E[\mathcal{H}_2]$	$d_{opt}^{(2)}$	$E[\mathcal{H}_1]$	$d_{opt}^{(1)}$	$E[\mathcal{H}_2]$	$d_{opt}^{(2)}$
T = 10	0.0670	87.00	0.0671	84.50	0.1434	105.00	0.1431	102.00
T = 40	0.0522	83.50	0.0519	81.00	0.1066	98.50	0.0999	94.50

From Table (III.54) results that since $E[\mathcal{H}_1] \simeq E[\mathcal{H}_2]$ the enlargement $\mathcal{H}_2$ of $\mathcal{H}_1$ is not significant for the purpose of optimal designing S and the cost functions assumed in the analysis. In fact, if the penalty for metamodel complexity is introduced (see Eq. (II.29) and Eq. (II.33)) then we would expect $E[\mathcal{H}_1]$ to be less than $E[\mathcal{H}_2]$ in most situations considered in (III.54) because $\mathcal{H}_2$ contains the normal model whose CDF has not an explicit analytical form while EX1 and EX2 distributions have. Therefore, (see Eq. (II.32)) $c(\mathcal{H}_1) < c(\mathcal{H}_2)$ so that it is expected that $\mathcal{H}^{(opt)} = \mathcal{H}_1$. From (III.54) it appears that $d_{opt}^{(i)}$ increases with the failure cost c_F; the fact that $d_{opt}^{(i)}$ decreases with T is the drawback of having in this example smaller sample variance at T = 40 (56.2901) than at T = 10 (87.9891) (see Tables (III.48) and (III.49)).

To complete the code phase the optimal code level IV strategy, $f_{IV}(x)$, is calculated next for $\mathcal{H}^{(opt)} = \mathcal{H}_1$ (also $f_{IV}(x)$ conditional on $\mathcal{H}_2$ is evaluated) and for the space of composite strategies A_c'. As defined, any $H_c \in A_c'$ has the form $H_c = \eta(H_\ell)H_\ell + \eta(H_k)H_k$, where $\eta(H_\ell) + \eta(H_k) = 1.0$, $\eta(H_\ell) \geq 0$, $\eta(H_k) \geq 0$, and $H_\ell, H_k \in A$. Numerical evaluation of the expected cost of the H_c was done only for discrete values of $\eta(H_\ell)$, namely $\eta(H_\ell) = 0.0\ (0.1)\ 1.0$.

From Eq. (II.20), the optimal design levels $\{d_\eta^{(opt)}\}$ of pure and composite strategies result. For instance, when the composite strategy is $\{\eta(EX2)EX2 + \eta(N)N\}$, the expected cost of designing S at x=d under η is:

$$E[d/\eta] = \eta(EX2)\ E[d/EX2] + \eta(N)\ E[d/N] \qquad \text{(III.55)}$$

The decomposition of the expected cost of the composite strategy η in the form (III.55) is very useful in practical calculations because it shows

that only the expected cost functions corresponding to the pure strategies have to be generated and stored in the computer.

From Eq. (II.17) and Eq. (II.29) (again without accounting for the cost of complexity), the expected cost, $E[\eta]$, of the strategy η corresponding to $\mathcal{H}^{(opt)} = \mathcal{H}_1$ is

$$E[\eta] = E[\eta/\mathcal{H}_1] = p''_{EX1}\, E[d_\eta^{(opt)}/EX1] + p''_{EX2}\, E[d_\eta^{(opt)}/EX2] \quad \text{(III.56)}$$

and the expected cost, $E[\eta/\mathcal{H}_2]$, of η under $\mathcal{H}_2$ has the form:

$$E[\eta/\mathcal{H}_2] = p''_{EX1}\, E[d_\eta^{(opt)}/EX1] + p''_{EX2}\, E[d_\eta^{(opt)}/EX2] + p''_{N}\, E[d_\eta^{(opt)}/N] \quad \text{(III.57)}$$

The optimal strategy $\eta_{IV}^{(opt)}$ providing $f_{IV}(x)$ (see Eq. (II.22)) minimizes $E[\eta]$ ($E[\eta/\mathcal{H}_2]$) in Eq. (III.56) (Eq. (III.57)).

In addition to the plots $E[d/\eta]$ in Eq. (II.19) for η being a pure strategy and $F[\cdot/H_{B,i}]$, i=1,2, Fig. (III.3) to Fig. (III.8) show also tables giving the optimal (without considering the penalty for nonsimplicity) pure and composite strategies. For example (see Fig. (III.3)) when $\tau = 1$, $c_F = 2$, $T = 10$ and the analysis is done conditional on $\mathcal{H}_2$, the best pure and composite strategies are, respectively, {St} and {0.2N + 0.8EX1} or {0.6St + 0.4EX1}. When the sample size is $T = 40$, (see Fig. (III.4)) EX1 is the best pure action and {0.3N + 0.7EX1} is the optimal composite strategy conditional on $\mathcal{H}_2$. When accounting for nonsimplicity the optimal strategies may be {EX1} or {0.2N + 0.8EX1} for $T = 10$ and {EX1} for $T = 40$. This might be so because the strategies {EX1} and {EX2} have their CDF's in explicit analytical form while {N} and {St} do not. One may obtain the CDF of {N} and {St} from tables depending on one parameter

and two parameters, respectively. Consequently, the penalty for complexity (see also Eq. (II.17)) corresponding to the pure strategies in A satisfies

$$c_{EX1} = c_{EX2} \leq c_N \leq c_{St} \tag{III.58}$$

Assume that the penalty for nonsimplicity of the composite strategy H_c is obtained by adding the penalties c_H of the pure actions H entering H_c. Then, for T = 10 (and this particular sample) it may occur that {0.2N + 0.8EX1} is smaller than c_{St} (see Fig.(III.3) and therefore the composite action {0.2N + 0.8EX1} may remain better than {St} even after the penalty for complexity is considered. In contrast, for T = 40 (see Fig. (III.4)), because the cost of operating {0.2N + 0.8EX1} is much larger than c_{EX1}, the pure strategy {EX1} is expected to be the code output.

Similar observations hold for all the cases presented in Fig. (III.3) to Fig. (III.8); the last two figures consider the case when T = 40, τ = 50 years and c_F = 2;20.

Fig. (III.9) shows the randomized and the nonrandomized risk sets (see Sec. II.1) where T = 40, τ = 1, c_F = 20 and $\mathcal{H}^{(opt)} = \mathcal{H}_1$. Since when T = 40 the strategies {N} and {St} practically coincide, only the pure strategies space A = { EX1$(\hat{\alpha},\hat{\mu})$, EX2$(\hat{k},\hat{v})$, N$(\hat{m},\hat{h})$ } has been considered. Also, the point corresponding to the optimal composite strategy {0.3EX2 + 0.7EX1} is represented.

Table (III.59) gives the coordinates of the pure strategies in A and of the composite action {0.3EX2 + 0.7EX1} in the space $\underline{\xi} = (\xi_1,\xi_2)$, where $\xi_1 = E[d_\eta^{(opt)}/EX1]$ and $\xi_2 = E[d_\eta^{(opt)}/EX2]$ ($d_\eta^{(opt)}$ has the same meaning as in Eq. (III.56) or Eq. (III.57)).

From Table (III.59) and Fig. (III.9) it appears that in this case

Table (III.59)

Strategy	$E[d_\eta^{(opt)}/EX1]$	$E[d_\eta^{(opt)}/EX2]$	Point in the Fig. (III.11)
EX1	0.0825046	0.1827418	$\underline{\xi}_{EX1}$
EX2	0.1043772	0.1520870	$\underline{\xi}_{EX2}$
N	0.1878176	0.4622742	$\underline{\xi}_{N}$
0.3EX2 + 0.7EX1	0.0870281	0.1624088	$\underline{\xi}_{(0.3EX2 + 0.7EX1)}$

(defined by T = 40, τ = 1, c_F = 20, $\mathcal{H}^{(opt)} = \mathcal{H}_1$, see also Fig. (III.6)) the normal strategy is not doing well although it has a high posterior probability (see Table (III.49)). This is so because of the very large penalty applied to N because $d_N^{(opt)}$ is low. Since the cost for complexity of EX1 and EX2 is small, the strategy {0.3EX2 + 0.7EX1} may remain optimal even after the penalties for nonsimplicity are considered.

Consider now the calculation of the optimal level III code strategy, $f_{III}(x)$, for the code C whose cost functions coincide with those of S in Fig. (I.1). Assume that a sample of size T = 10 is available (see first row in Table (III.48)) and let $\mathcal{H}^*$ be the operational metamodel and P_f = 0.01 the (chosen) target failure probability (see Table (I.1)).

To evaluate $f_{III}(x)$ (code phase), first the optimal metamodel has to be identified. Consider also here as possible candidates for $\mathcal{H}^{(opt)}$ the metamodels $\mathcal{H}_1$ and $\mathcal{H}_2$. Denoting by $d^{(i)}(P_f)$ the $(1-P_f)$ fractile of the Bayesian CDF $F_{B,i}(x)$ it follows: $d^{(1)}(0.01) = 89.00$ and $d^{(2)}(0.01) = 85.00$. Then assuming τ = 1 year, c_F = 2 (see Fig. (III.3)), the expected cost $E[d^{(i)}(P_f)/\mathcal{H}_j]$, i,j = 1,2, is the ordinate of the curve $5_{(j)}$ at $d=d^{(i)}(P_f)$. It follows that, $E[d^{(1)}(0.01)/\mathcal{H}_1] = 0.0720965$ and $E[d^{(1)}(0.01)/\mathcal{H}_2] =$ $= 0.0695782$. From Eq. (III.52) where $E[d_{opt}^{(i)}/\mathcal{H}_j]$ is replaced by

$E[d^{(i)}(P_f)/\mathcal{H}_j]$, the expected cost of the metamodel $\mathcal{H}_1$ and $\mathcal{H}_2$ are found: $E[\mathcal{H}_1] = 0.0706$, $E[\mathcal{H}_2] = 0.0669$. Since $E[\mathcal{H}_2] < E[\mathcal{H}_1]$, $\mathcal{H}_2$ is said to be better than $\mathcal{H}_1$ if no penalty for complexity is assigned. If the penalty for nonsimplicity is considered the order between $\mathcal{H}_1$ and $\mathcal{H}_2$ may change.

Table (III.60) gives the $(1-P_f)$ fractiles of the distributions in A (only pure strategies are considered here), $d_H(P_f)$, and the expected cost of the pure strategies (see also Eq. (II.25)) conditional on $\mathcal{H}_i$. For example the expected cost of the strategy EX1 under $\mathcal{H}_j$ is equal to the ordinate of the curve $5_{(j)}$ (see Fig. (III.3)) at $d = d_{EX1}(0.01)$, where $d_{EX1}(0.01) = 84.30$ is the P_f - upper fractile of EX1.

Table (III.60)

		Expected cost of the strategy	
Strategy	$d_H(P_f)$	conditional on $\mathcal{H}_1$	conditional on $\mathcal{H}_2$
EX1	84.30	0.0731	0.0632
EX2	94.20	0.0780	0.0738
N	74.40	0.1117	0.0876
St	78.70	0.0854	0.0696

From (III.60), it appears that the optimal level III code strategy is $f_{III}(x) = f_{EX1}(x/\hat{\alpha},\hat{\mu})$, where $\hat{\alpha} = 0.1317$, $\hat{\mu} = 49.3524$.

The designer phase consists of the evaluation of the $(1-P_f)$ fractile of $f_{III}(x)$. Thus, the optimal level III design of S for $P_f = 0.01$ is $d_{III} = d_{EX1}(0.01) = 84.30$.

From this example the following conclusion may be drawn:

(1) The optimal code strategy f(x) depends directly on the observation

vector $\underline{Z}$, the cost functions of S and the penalties for complexity (see Fig. (III.3) to Fig. (III.12).

(2) In general there is a composite strategy which is optimal, i.e. it is better (in the sense of expected cost in Eq. (III.56)) than any other pure or composite action.

(3) From Fig. (III.3) to Fig. (III.8) it appears that to find $\eta_{IV}^{(opt)}$ is equivalent to the identification of the linear combination, η, of pure strategies which minimizes the distance $\delta(\eta) = |d_{\eta}^{(opt)} - d_{opt}|$ (where $d_{\eta}^{(opt)}$ and d_{opt} are, respectively, the optimal design levels of η and of the Bayesian strategy of $\mathcal{H}^{(opt)}$). Consequently a composite strategy η may become the best only if it contains pure actions having there optimum designs both above and below d_{opt}. For example, in Fig. (III.7), the combination $\{\eta(N)N + \eta(EX1)EX1\}$ cannoc become the best since $d_N^{(opt)} < d_{EX1}^{(opt)} < d_{opt} = 106$; in fact always $\{EX1\}$ is better than any combination $\{\eta(N)N + \eta(EX1)EX1\}$. However, this is only a necessary condition.

This observation shows also the importance of d_{opt} or more accurately of the Bayesian model $H_B^{(opt)}$, of $\mathcal{H}^{(opt)}$ in the identification of the code strategy. Consequently, the decision framework in Sec. II.2 which allows one to compare many metamodels on the grounds of the evidence $\underline{Z}$ and the purpose (the cost functions) seems adequate.

(4) In summary the code phase solves two problems: first it evaluates the reference $H_B^{(opt)}$ (of $\mathcal{H}^{(opt)}$) which in a utility sense models best X(t) and second, the optimal strategy f(x) is calculated using Eq. (III.56).

It is helpful to contrast our Bayesian <u>parametrical</u> approach to metamodel selection leading to the optimal Bayesian model $H_B^{(opt)}$ with a Baye-

sian nonparametrical procedure to estimate the distribution of the input variable X(t) [28]. In fact one may replace in Eq. (III.56) the CDF $F_B^{(opt)}(x)$ of $H_B^{(opt)}$ by the best estimate $\hat{F}_T(x/\underline{Z})$ of the actual CDF, F(x), of X(t) obtained by Ferguson [28] in a Bayesian nonparametrical context (see details in the original paper [28]). The form of $\hat{F}_T(x/\underline{Z})$ is:

$$\hat{F}_T(x/\underline{Z}) = p_T F_o(x) + (1-p_T)\, F_T(x/\underline{Z}) \qquad \text{(III.61)}$$

where, $p_T = \frac{\alpha(R)}{\alpha(R)+T}$ and $\alpha(R)$ is a measure on the real line,

$F_T(x/\underline{Z}) = \frac{1}{T} \sum_{t=1}^{T} u(x-Z_t)$ is the empirical cumulative distribution function; u(y) is the unit step function located at y=0, and $F_o(x)$ is the prior CDF of X(t).

The measure $\alpha(R)$ may be viewed as a quantification of the faith in the prior guess $F_o(x)$. For example when $\alpha(R)$ is small compared with the sample size T, little weight is given to the prior choice $F_o(x)$. In fact when $\alpha(R) \to 0$, $\hat{F}_T(x/\underline{Z})$ is the best estimate of F(x) corresponding to the noninformative Dirichlet prior and converges to $F_T(x/\underline{Z})$.

The expected cost of the strategy η, assuming that $\hat{F}_T(x/\underline{Z})$ describes X(t) results from Eq. (III.56) or Eq. (III.57) where the CDF of the Bayesian strategy of $\mathcal{H}_i$ is replaced by $\hat{F}_T(x/\underline{Z})$:

$$E[\eta] = c(d_\eta^{(opt)}) + c_F[1 - \hat{F}_T(d_\eta^{(opt)}/\underline{Z})] \qquad \text{(III.62)}$$

In all the cases considered in Fig. (III.3) to Fig. (III.8) the optimal design, $d_\eta^{(opt)}$, is always higher than the largest observation in

$\underline{Z}$ (see Fig. (III.10); consequently, in all the cases

$$\hat{F}_T(d_\eta^{(opt)}/\underline{Z}) = p_T\, F(d_\eta^{(opt)}) + (1-p_T) \qquad \text{(III.63)}$$

Thus, the goodness of the strategy η (Eq. (III.62)) is obtained only from the cost functions and prior informations; the observation vector $\underline{Z}$ is not used at all. Clearly, a somewhat similar situation occurs also in our parametrical approach leading to $H_B^{(opt)}$ in the sense that the parameters of the models in $\mathcal{H}^{(opt)}$ are obtained from data located in the body of the distribution while the goodness of η is entirely based on the upper tail of $H_B^{(opt)}$ whose shape is defined by the type of models and their weights in $\mathcal{H}^{(opt)}$. However, $H_B^{(opt)}$ is the result of an extensive sensitivity analysis (see Sec. II.2) of the optimal design level of S to many models and also corresponds to the purpose quantified in cost functions, versus $\hat{F}_T(x/\underline{Z})$ in Eq. (III.63) which depends only on the statistician's intuition.

The next section considers the design of S in Fig. (I.1) when X(t) is a stationary correlated sequence and the system S is designed at a level x=d which is a high fractile of the marginal distribution of X(t).

III.3.2 Example 2

Consider the design against wind of the structural system S in Fig. (I.1), and let X(t) be the mean hourly wind speeds sequence. Assume that X(t) is measured for a short time period, say, five years and that the goal is to design S for wind speeds of return periods up to 100 years. The standard procedure is to retain the maximum wind speed recorded in each year and then to infer the distribution of the yearly wind extremes, $F_{year}(x)$ However, $F_{year}(x)$ corresponding to few data points (five in our

example) has generally very large uncertainty which may lead to unrealistic designs of S. Even if the shape of the distribution, $f(x/\underline{\theta})$, of X(t) is known, the estimates of $\underline{\theta}$ have generally large variances, or in a Bayesian context, the posterior probability of the unknown parameters (see Eq. (III.2)), $f''(\underline{\theta})$ does not differ much from $f'(\underline{\theta})$.

To avoid this problem Davenport [19], Gomes and Vickery [32] proposed the computation of the wind speeds corresponding to various return periods from the mean rate of upcrossings of the correlated process X(t), where X(t) is the mean hourly wind speed process.

As is well known [69] the average rate of upcrossings of level x=d, ν_{d+}, has the form (assuming that X(t) is stationary)

$$\nu_{d+} = E[\dot{X}+/X=d]\, f_x(d) \tag{III.64}$$

where, $f_x(x)$ is the marginal distribution of X(t) and $E[\dot{x}+/x=d]$ denotes the first moment of the positive tail of $\dot{x}$ distribution conditional on x=d, i.e.

$$E[\dot{X}+/X=d] = \int_0^\infty \dot{x}\, f_{\dot{X}/X}(\dot{x}/d)d\dot{x} \tag{III.65}$$

Then, the average number of upcrossings of x=d during time τ is $\nu_{d+}\tau$ and consequently the mean return period of the wind speed X(t) = d is $R_d = 1/(\nu_{d+}\tau)$. Thus, the evaluation of ν_{d+} leads directly to R_d.

The procedure to evaluate R_d in [32] is presented next; let $\underline{Z}$ be the observation vector containing all the mean hourly wind speeds recorded in the five years of data. The evaluation of ν_{d+} follows from Eq. (III.64) under the following two assumptions:

(i) - The original sequence X(t) and the derivative $\dot{X}(t)$ process

are assumed independent; consequently Eq. (III.64) becomes simply

$$\nu_{d+} = E[\dot{X}+]\, f_X(d) \qquad \text{(III.66)}$$

The average $E[\dot{X}+]$ is evaluated numerically from $\underline{Z}$; various approximations are proposed to calculate $E[\dot{X}+]$.

(ii) - The marginal distribution of X(t) is assumed to be Weibull with parameters k, and c:

$$F_X(x) = 1 - \exp[-(x/c)^k] \qquad \text{(III.67)}$$

The parameters (k,c) in Eq. (III.67) are estimated in [32] based on five years of mean hourly wind speed records (1965-1969) from Observatory Hill, Sydney. Their values are, [32]: k = 1.492 and c = 7.7(knots). The corresponding mean and standard deviation are, respectively, m = 6.869 and r = 4.963:

Conditional on the assumptions (i) and (ii) the mean rate wind up-crossings, ν_{d+}, and the return periods, R_d, are computed and plotted (see Fig. (III.11) and Fig. (III.12), respectively).

The fact that the return periods in [32] or in [19] differ by less than 5% from the Gumbel line based on eighteen yearly maximum mean hourly wind speeds when $R_d \in (5,100)$ years is viewed in [32] as a success of the crossing approach. On the contrary, we are inclined to believe that the "success" of the crossing solution in [19] and in [32] rests solely on the proper choice of the marginal distribution of X(t).

In Sec. IV.3.2.4 (see also Fig. (IV.4) it is found that the mean rate

of upcrossings, $\tilde{\nu}_{d+}$, of x=d for a Gaussian correlated sequence, $X_G(t)$, varies slowly with the one lag correlation coefficient of $X_G(t)$ if the threshold d is large, say $d \geq m_x + 2.5\sigma_x$. In this case $\tilde{\nu}_{d+}$ is well approximated by the average rate of upcrossings, $\tilde{\nu}_{d+}^{(i)}$, of the independent sequence $X_G^{(i)}(t) \in N(m_x, \sigma_x^2)$:

$$\tilde{\nu}_{d+}^{(i)} = \Phi\left(\frac{d-m_x}{\sigma_x}\right)\left[1 + \Phi\left(\frac{d-m_x}{\sigma_x}\right)\right] \qquad \text{(III.68)}$$

where, $\Phi(\cdot)$ is the cumulative distribution function of the standardized Gaussian variable.

The same idea is used here to calculate approximate values of the winds speeds for various return periods; in this case ν_{d+} in Eq. (III.66) is replaced by (see Eq. (III.68)) $\nu_{d+}^{(i)}$, where (see also Fig. (III.11)):

$$\nu_{d+}^{(i)} = \{1 - [\exp(-\frac{d}{c})^k]\} \exp[-(\frac{d}{c})^k] \qquad \text{(III.69}$$

The return period $R_d^{(i)}$ corresponding to $\nu_{d+}^{(i)}$ with k = 1.492, and c = 7.7(knots) is shown in Fig. (III.12). The maximum discrepancy between $R_d^{(i)}$ and the Gumbel line over the range 5 to 100 years is 10%. Therefore, for this data set at least, the replacement of the original sequence X(t) by the white sequence $X^{(i)}(t)$ with the same marginal distribution as X(t) generates a good approximation of R_d. Thus, the large amount of computation to find ν_{d+} in [19,32] seems not necessarily satisfactory.

The shape of the marginal density of X(t) appears to be essential towards evaluating R_d. The Table (III.70) gives the ratios between the return period of X(t) assumed Weibull, $R_{d,W}$, to the return periods of X(t) considered Gaussian ($R_{d,N}$), extreme type I, ($R_{d,EXI}$),and extreme type II

($R_{d,EX2}$), respectively, for design levels d = 10 (5) 40 knots. The parameters of all marginal distributions of X(t) considered in (III.70) are calculated from the sample mean m = 6.896 knots and the sample variance $\sigma^2 = 4.963^2$ knots given in [32] (see also Eq. (II.67))

Table (III.70)

Level,d	$R_{d,W}/R_{d,N}$	$R_{d,W}/R_{d,EX1}$	$R_{d,W}/R_{d,EX2}$
10	1.3092	0.9999	0.6761
15	1.1677	0.9207	0.4192
20	0.4986	0.9835	0.5106
25	0.0967	1.2667	0.9881
30	0.0082	1.9547	2.7052
35	0.0003	3.5629	9.7849
40	~0.0000	7.5678	44.7323

Consequently, comparing the error related to the independence assumption on X(t) with the error corresponding to various choices of the marginal PDF of X(t) (see Table (III.70)), it results, for the observation vector used in [32], that:

(i) The correlation structure of the mean hourly wind can be neglected when evaluating the level III design of S. From Fig. (IV.4) and the fact that the one lag correlation coefficient, ρ, found in [32] is $\rho = 0.8551$, one may expect that this conclusion is more general than stated here. However, numerical studies are needed. In fact the solution in [32] and the approach proposed in this

section are tested against only one observation vector.

(ii) From Table (III.70) the importance of the appropriate choice of the marginal distribution of X(t) is evident. When the correct PDF (here Weibull) is not identified, the error in the design level d is large for return periods of, say, 50 to 100 years which are of interest in Structural Reliability. Consequently, the selection of the marginal PDF of X(t) is essential for the design of structural systems. Thus, a utility approach is proposed next to best design the structural system S in Fig. (I.1).

The decision framework developed in this work may be directly used to evaluate the optimal code strategy, i.e. by marginal PDF of X(t) (in the further development X(t) is replaced by $X^{(i)}(t)$). Then, the optimal design of S is readily obtained. Assume that a level III code format is to be implemented for wind and, say, that the observations consist of measurements of the hourly average wind speed during several years.

Let P_f be the input failure probability (see Table (I.1)) and $d(P_f)$ be the P_f - upper fractile of the strategy η. By first approximating the probability $P(\max_{1\leq t\leq\tau}\{X(t)\} > d)$ by $\nu_{d+}\tau$ (this approximation is broadly accepted) and then $\nu_{d+}\tau$ by $\nu^{(i)}_{d+}\tau$, the expected cost of η(see Eq. (II.24) and Eq. (II.25)) is:

$$E[\eta] = c_\eta + c(d_\eta(P_f)) + c_F \sum_{H\in\mathcal{H}^{(opt)}} [\nu^{(i)}_{d_\eta(P_f),H}\tau]p''_H \qquad (III.70)$$

where, c_η, $c(\cdot)$, c_F, p''_H have the same meaning as in Eq. (II.24), $\mathcal{H}^{(opt)}$ is the optimal (code)metamodel, and $\nu^{(i)}_{d_\eta(P_f),H}$ is the mean rate of upcrossings of X(t) at level $X = d_\eta(P_f)$ correspond-

ing to $X^{(i)}(t)$ with the PDF given by H.

By minimizing the expected cost in Eq. (III.70) versus η, the optimal level III code strategy, $\eta_{III}^{(opt)}$, results, and therefore the best level III design, d_{III}, of S is the $(1 - P_f)$ fractile of $\eta_{III}^{(opt)}$; d_{III} also represents the $1/P_f$ - years wind speed return period (say, τ in Eq. (III.70) is (365 days) × (24 hours) = 8760 hours) conditional on the marginal PDF of X(t) being $\eta_{III}^{(opt)}$.

III.4 Conclusions

The development in this chapter follows the decision framework established in Sec. II.2 for the case when X(t) is a white stationary sequence. Specific inferential results (such as posterior PDF of $\underline{\theta}_H$, predictive PDF $f_H(\tilde{\underline{y}}_\tau/\underline{Z})$ for $\tau = 1$ and unconditional likelihood function $f_H(\underline{Z})$) have been obtained for Gaussian, Gamma and Poisson sequences. Only the simple predictive distribution, $f_H(\tilde{y}/\underline{Z})$, has been obtained since generally $f_H(\tilde{\underline{y}}_\tau/\underline{Z})$, $\tau > 1$ does not have an analytical closed-form, and because in many structural reliability problems $f_H(\tilde{\underline{y}}_\tau/\underline{Z})$ can be (approximately) obtained from $f_H(\tilde{y}/\underline{Z})$ and τ (see Eq. (III.5)).

Several cost functions corresponding to problems of type (iii) in Chap. I are suggested in Sec. II.2.3. Utility functions which are adequate to problems (i) and (ii) may be also developed.

The main output of this chapter consists of the examples presented in Sec. II.3. From Sec. III.3.1 it appears that, the dispute related

to the type of distribution to be assigned to the yearly wind speed extremes is linked to the present inferential approach to model selection. As shown in Tables (III.48) and (III.49) for various realizations of X(t) different models become inferentially the best.

The example in Sec. III.3.1 has the merit of illustrating in detail the steps of the decision framework developed in Sec. II.2 for the case when X(t) is a stationary independent sequence. Thus if in Sec. III.3.1 one uses the results in the last line of Table (III.49) corresponding to the wind yearly extremes recorded in Boston during 42 years (1933-1974) the optimal design of the system S in Fig. (I.1) loaded by wind is obtained.

Sec. III.3.2 considers a level III code format. A result is that for small P_f (see Table (I.1)) the correlation of X(t) (where it exists) is not significant for d_{III}. On the contrary d_{III} is very sensitive to the marginal PDF of X(t). Based on Chap. II and several approximations in Chap. IV an approximate procedure is proposed to handle the design of S in Fig. (I.1) subjected to stationary correlated sequences when P_f is small.

As a general conclusion of the results in Sec. III.3 it appears necessary to search among many models (i.e. to take a large $\mathcal{H}^*$) when evaluating $\mathcal{H}^{(opt)}$ since the best code strategy f(x) depends strongly on the Bayesian model $H_B^{(opt)}$ of $\mathcal{H}^{(opt)}$.

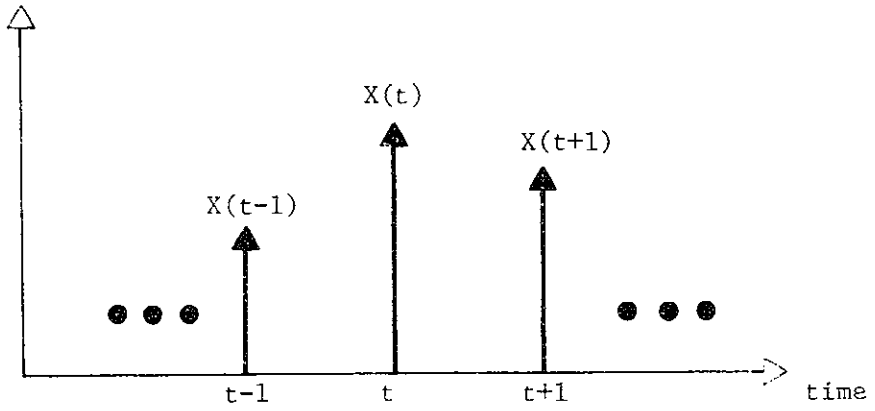


Fig. III.1. Hypothetical realization of an independent sequence.

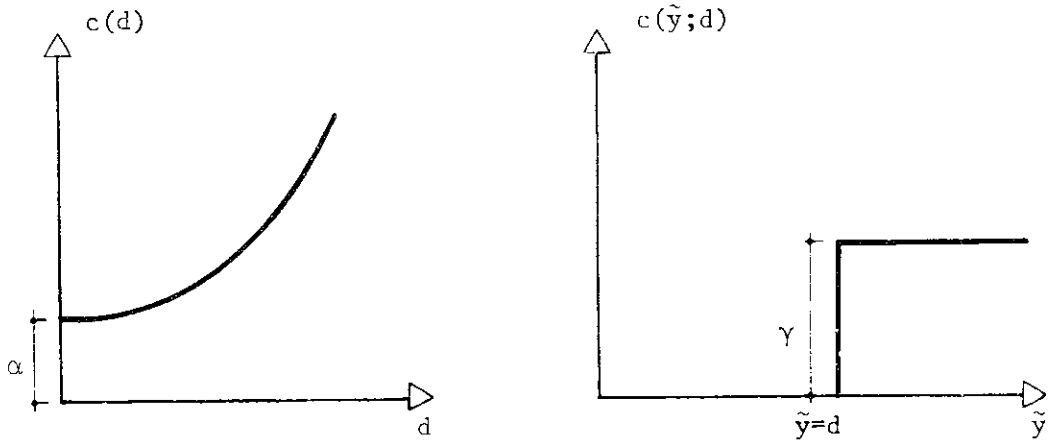
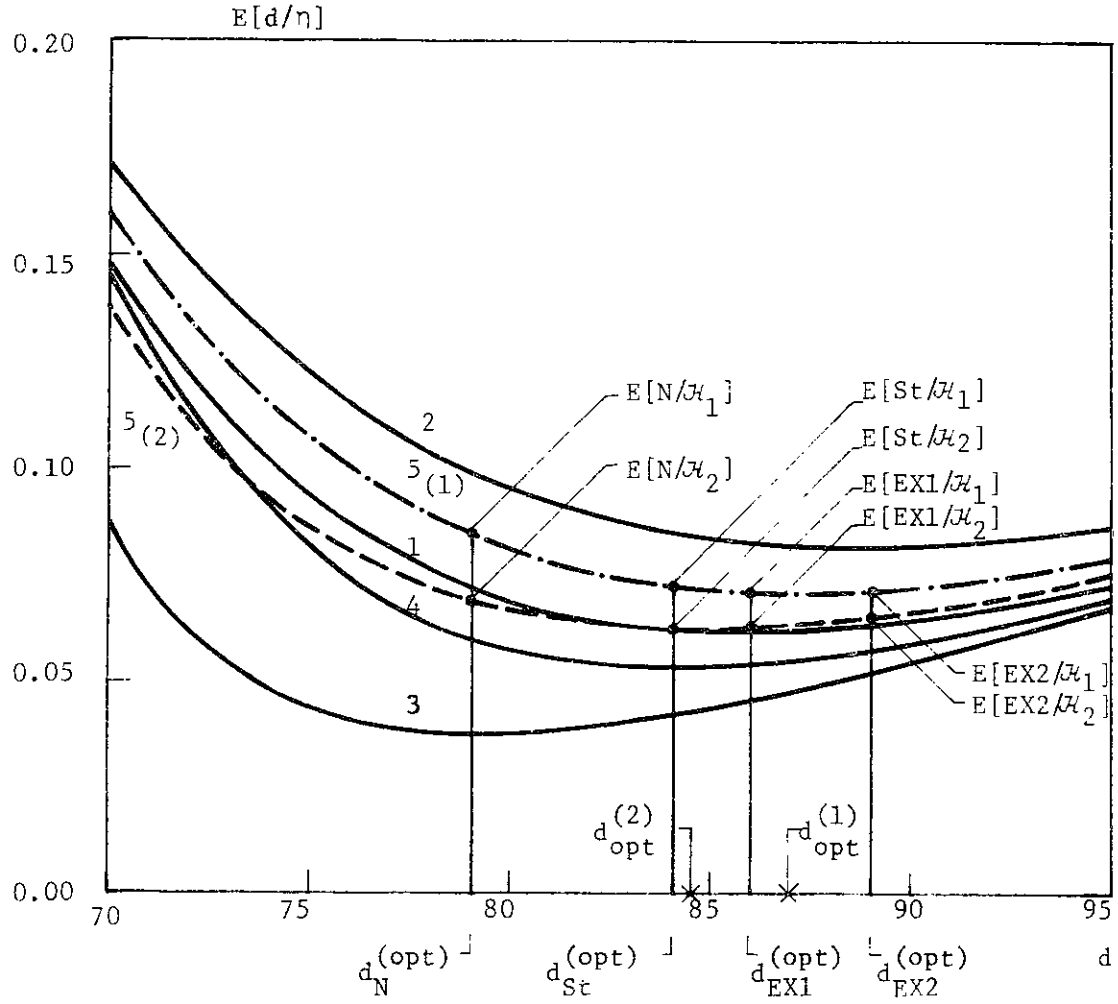
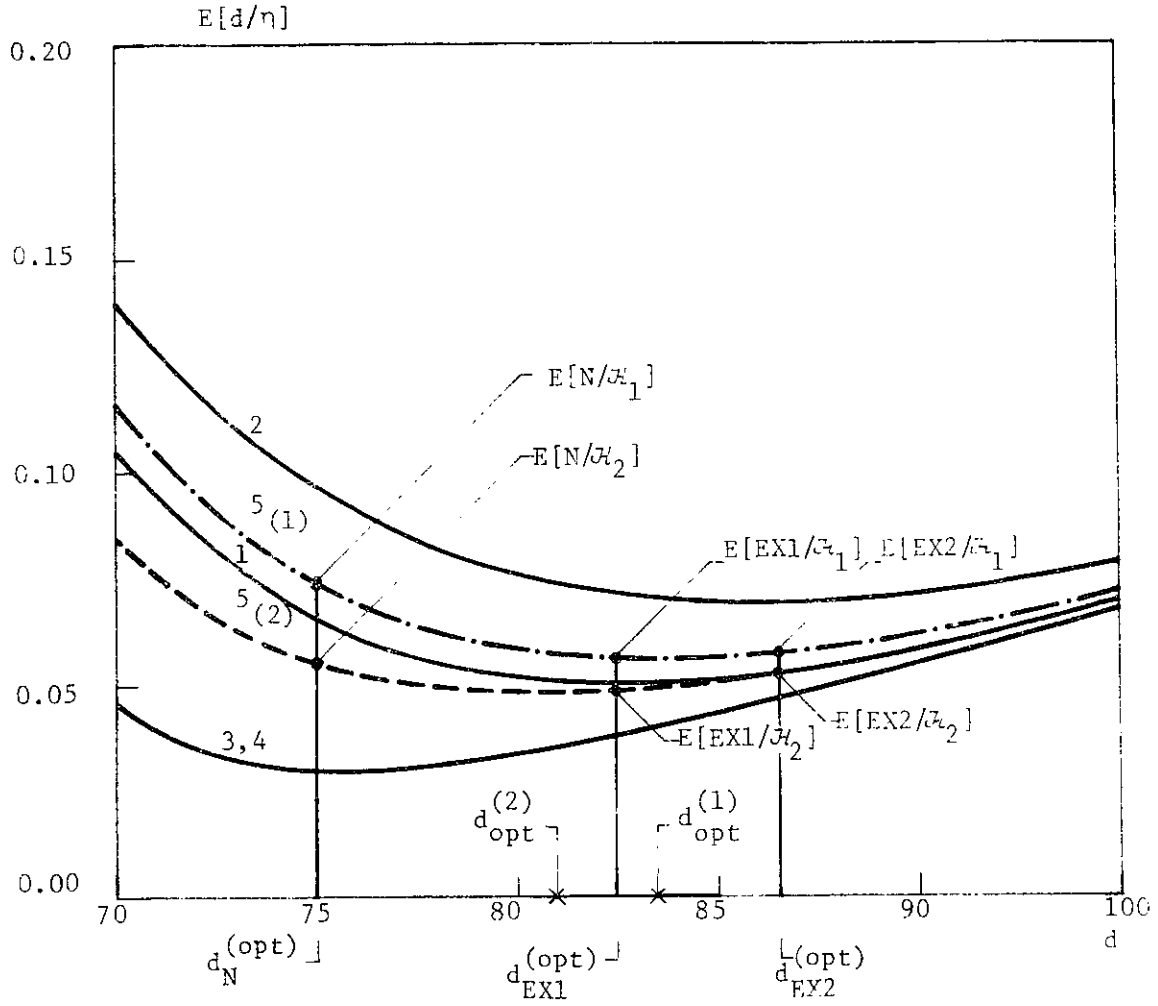


Fig. III.2. Typical cost functions in Structural Reliability.
c(d) = the initial cost of designing S at X=d
γ = the cost of failure.



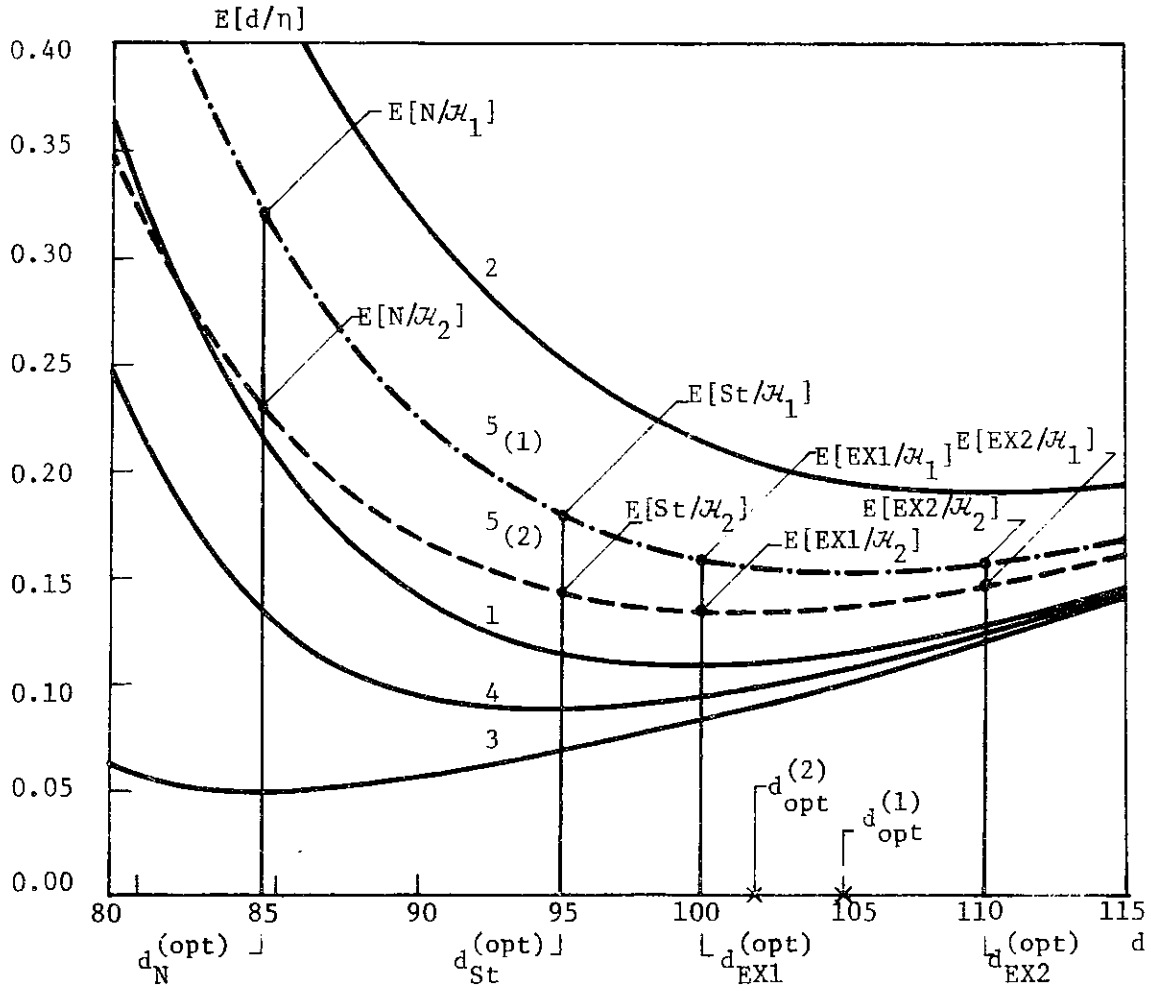
Optimal strategies under $\mathcal{H}_1$	$E[\eta/\mathcal{H}_1]$	Optimal strategies under $\mathcal{H}_2$	$E[\eta/\mathcal{H}_2]$
pure: EX1	0.0718845	pure: St	0.0632475
composite: 0.2N + 0.8EX2 0.3St + 0.7EX2 0.4EX2 + 0.6EX1	0.0716449	composite: 0.2N + 0.8EX1 0.6St + 0.4EX1	0.0631738

Fig. III.3. Expected cost $E[d/\eta]$ in Eq. (II.19) for T=10, c_I=1, c_F=2 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ (curves 1,2,3,4, $5_{(1)}$ and $5_{(2)}$, respectively). The table gives the expected cost of the optimal pure and composite strategies under $\mathcal{H}_1$ and $\mathcal{H}_2$.



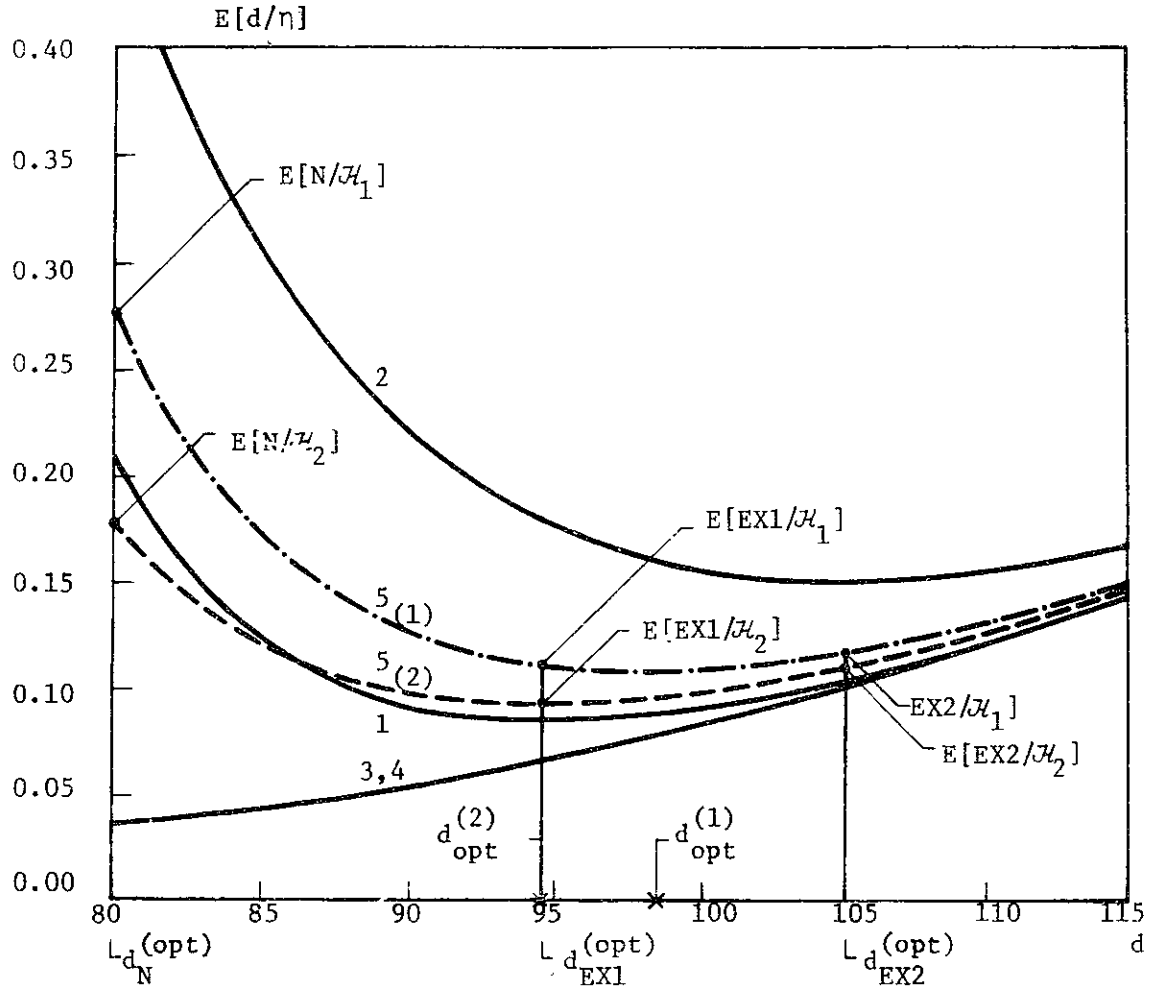
Optimal strategies under $\mathcal{H}_1$	E[η/$\mathcal{H}_1$]	Optimal strategies under $\mathcal{H}_2$	E[η/$\mathcal{H}_2$]
pure: EX1	0.0566097	pure: EX1	0.0491-49
composite: 0.3EX2 + 0.7EX1	0.0564161	composite: 0.3N + 0.7EX1	0.0486 49

Fig. III.4. Expected cost E[d/η] in Eq. (II.19) for T=45, η=1, c_F=2 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ (curves 1,2,3,4, $5_{(1)}$ and $5_{(2)}$, respectively). The table gives the expected cost of the optimal pure and composite strategies under $\mathcal{H}_1$ and $\mathcal{H}_2$.



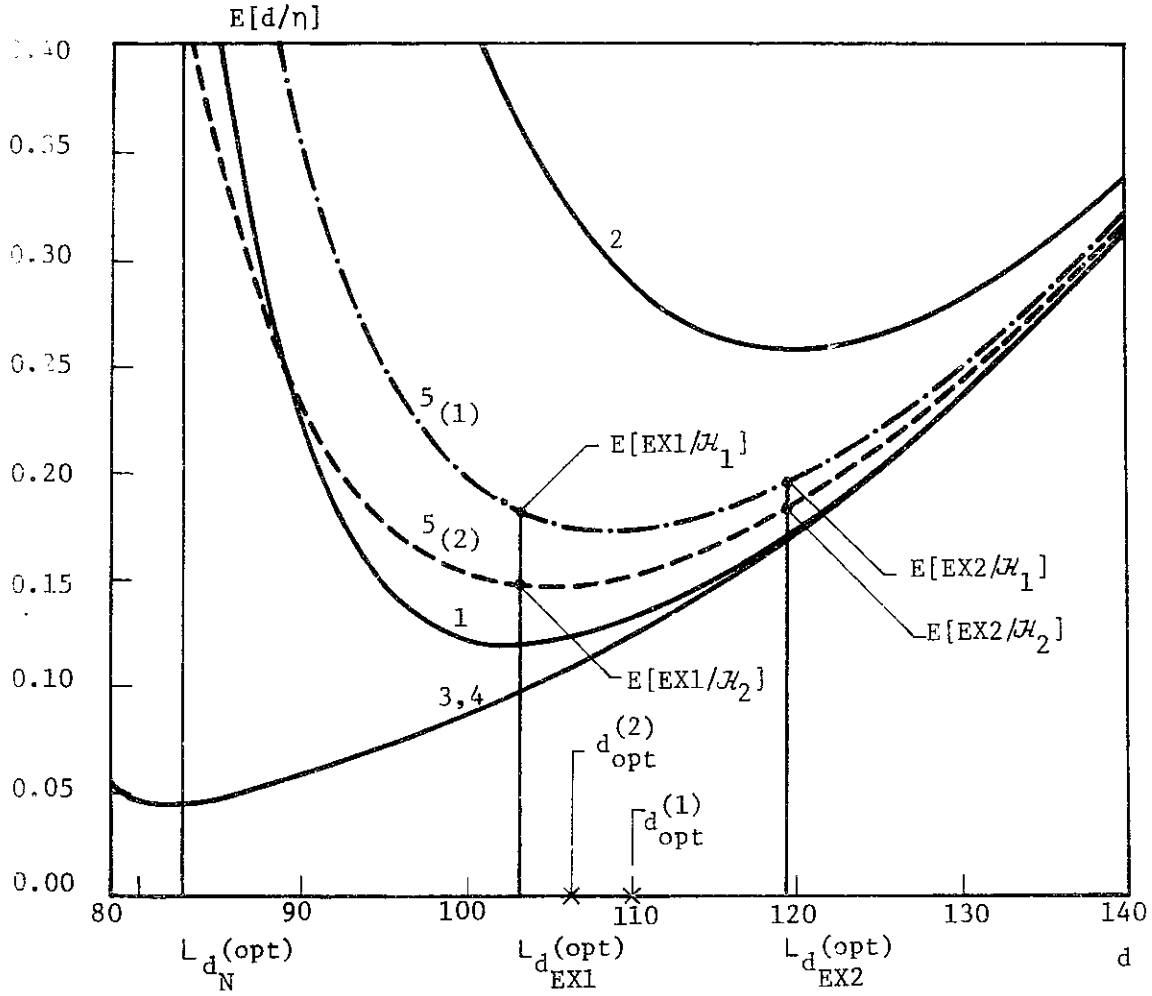
Optimal strategies under $\mathcal{H}_1$	$E[\eta/\mathcal{H}_1]$	Optimal strategies under $\mathcal{H}_2$	$E[\eta/\mathcal{H}_2]$
pure: EX2	0.1584725	pure: EX1	0.1351943
composite: 0.4St + 0.6EX2 0.4EX2 + 0.6EX1	0.1535596	composite: 0.2EX2 + 0.8EX1 0.6St + 0.4EX2	0.1345025

Fig. III.5. Expected cost $E[d/\eta]$ in Eq. (II.19) for T=10, τ=1, c_F=20 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ (curves 1,2,3,4, $5_{(1)}$ and $5_{(2)}$, respectively). The table gives the expected cost of the optimal pure and composite strategies under $\mathcal{H}_1$ and $\mathcal{H}_2$.



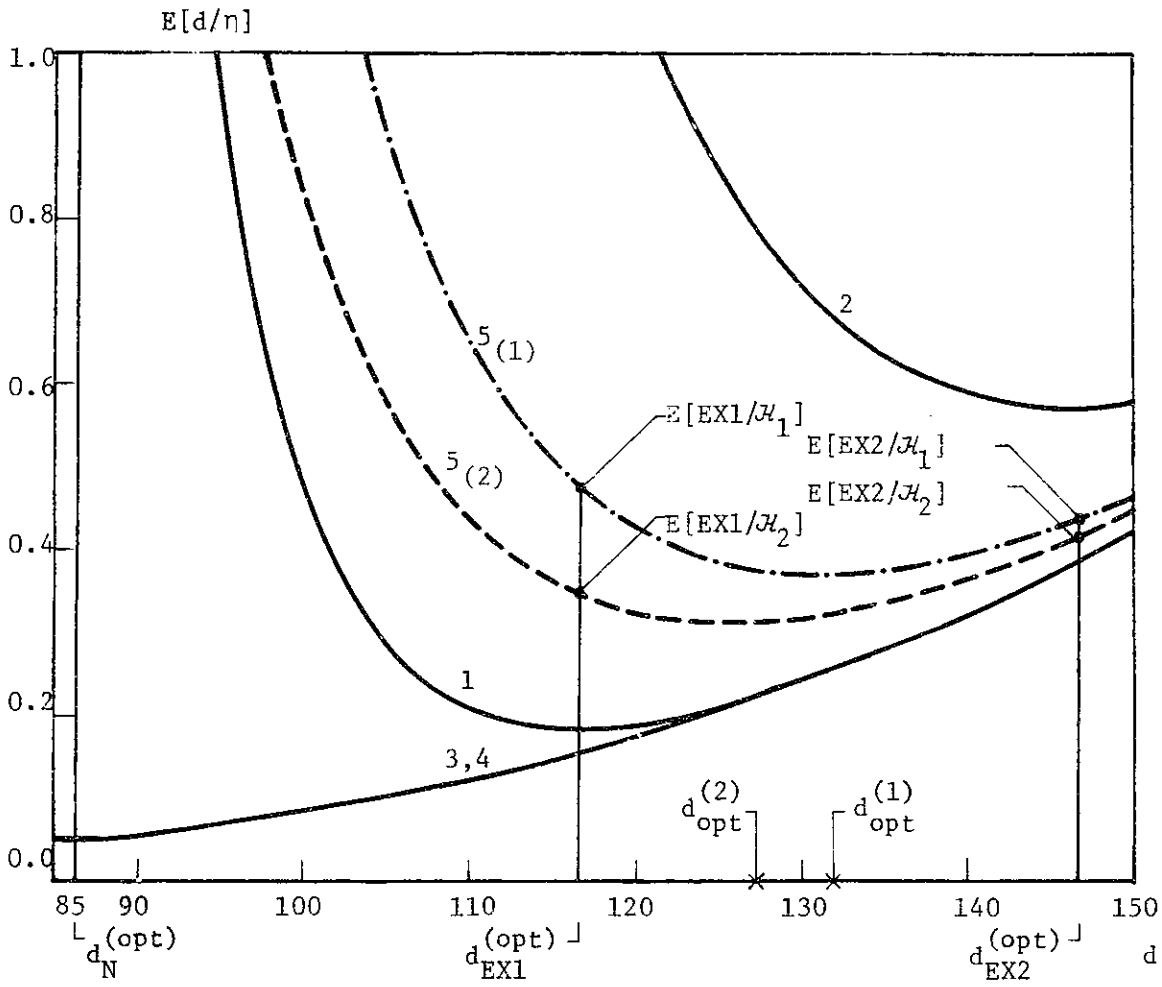
Optimal strategies under $\mathcal{H}_1$	$E[\eta/\mathcal{H}_1]$	Optimal strategies under $\mathcal{H}_2$	$E[\eta/\mathcal{H}_2]$
pure: EX1	0.1117200	pure: EX1	0.0928803
composite: 0.3EX2 + 0.7EX1	0.108565	composite: EX1	0.0928803

Fig. III.6. Expected cost $E[d/\eta]$ in Eq. (II.19) for T=40, τ=1, c_F=20 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ (curves 1,2,3,4, $5_{(1)}$ and $5_{(2)}$, respectively). The table gives the expected cost of the optimal pure and composite strategies under $\mathcal{H}_1$ and $\mathcal{H}_2$.



Optimal strategies under $\mathcal{H}_1$	$E[\eta/\mathcal{H}_1]$	Optimal strategies under $\mathcal{H}_2$	$E[\eta/\mathcal{H}_2]$
pure: EX1	0.1865978	pure: EX1	0.1479771
Composite: 0.3EX2 + 0.7EX1	0.1738968	Composite: 0.1EX2 + 0.9EX1	0.1471157

Fig. III.7. Expected cost $E[d/\eta]$ in Eq. (II.19) for T=40, τ=50, c_I=2 and η=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ (curves 1,2,3,4, $5_{(1)}$ and $5_{(2)}$, respectively). The table gives the expected cost of the optimal pure and composite strategies under $\mathcal{H}_1$ and $\mathcal{H}_2$.



Optimal strategies under $\mathcal{H}_1$	$E[\eta/\mathcal{H}_1]$	Optimal strategies under $\mathcal{H}_2$	$E[\eta/\mathcal{H}_2]$
pure: EX2	0.4335150	pure: EX1	0.3392990
composite: 0.3EX2 + 0.7EX1	0.3672119	composite: 0.8N + 0.2EX2	0.3081866

Fig. III.8. Expected cost $E[d/\eta]$ in Eq. (II.19) for T=40, τ=50, c_F=20 and r=EX1,EX2,N,St,$H_{B,1}$,$H_{B,2}$ (curves 1,2,3,4, $5_{(1)}$ and $5_{(2)}$, respectively). The table gives the expected cost of the optimal pure and composite strategies under $\mathcal{H}_1$ and $\mathcal{H}_2$.

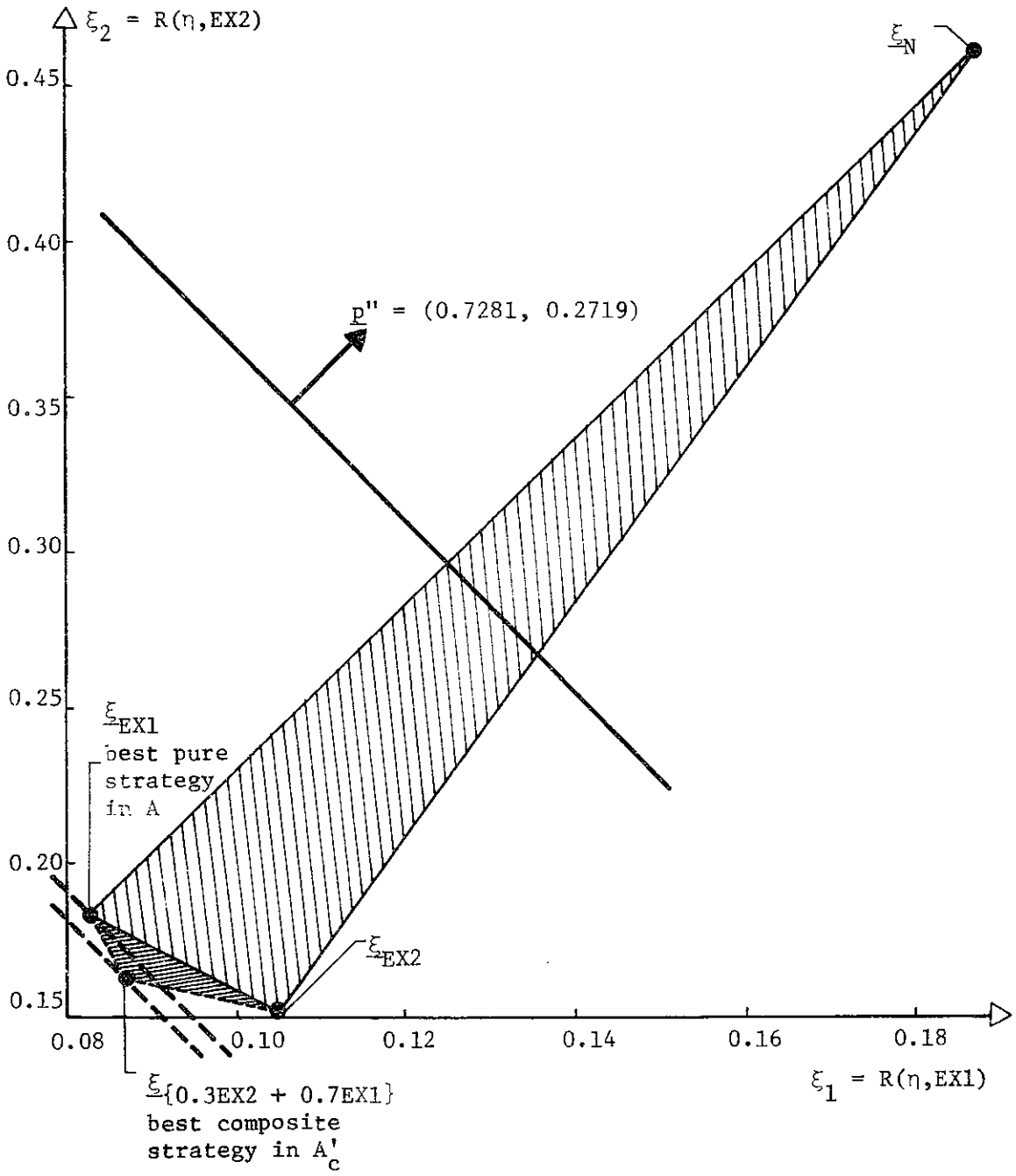


Fig. III.9. Randomized and nonrandomized risk sets corresponding to various code strategies.

$\xi_{\{0.3EX2 + 0.7EX1\}}$ = the risk point of the composite strategy {0.3EX2 + 0.7EX1}.

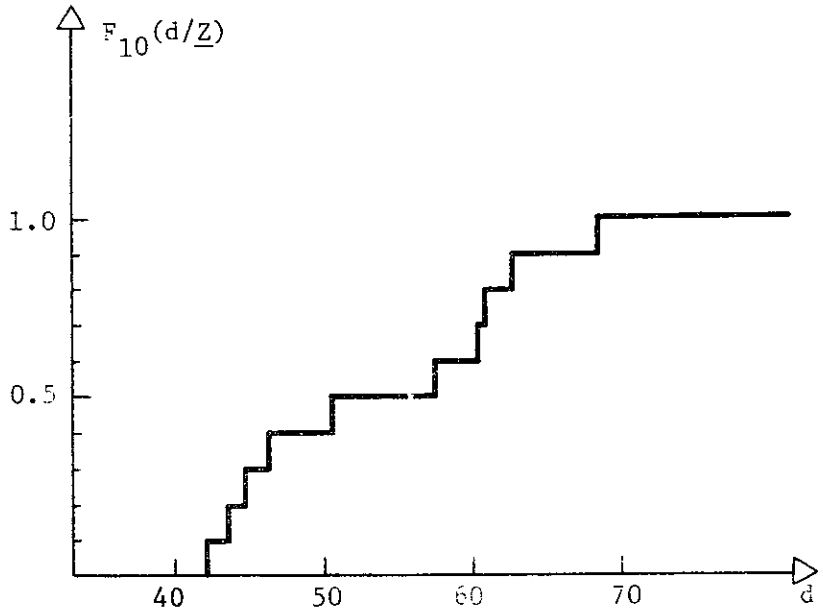


Fig. III.10. Empirical CDF. Illustration for the observation vector whose statistics are given in the first line of Table (III.48).

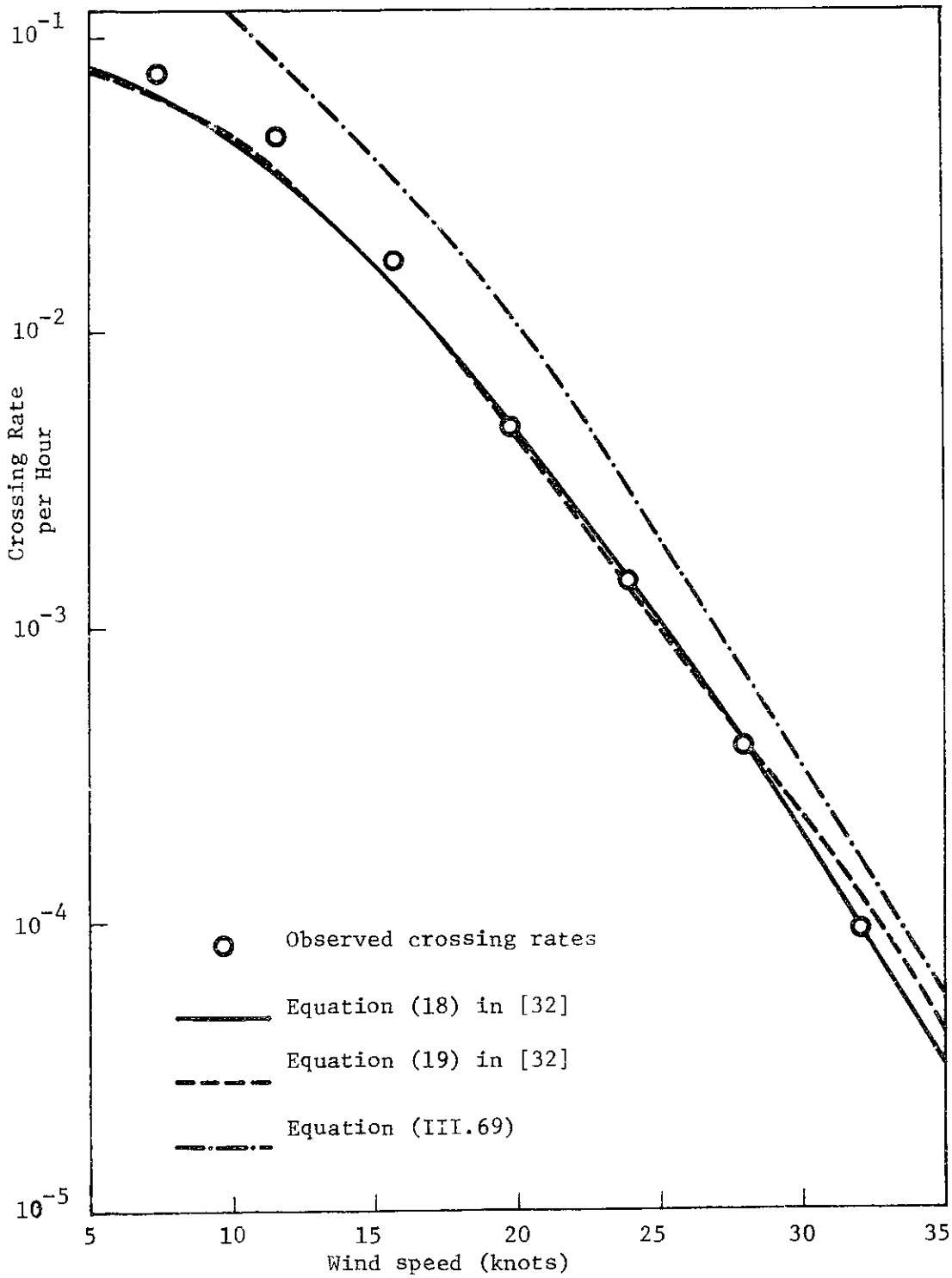


Fig. III.11. Predictive and observed crossing rate; figure base in reference [32].

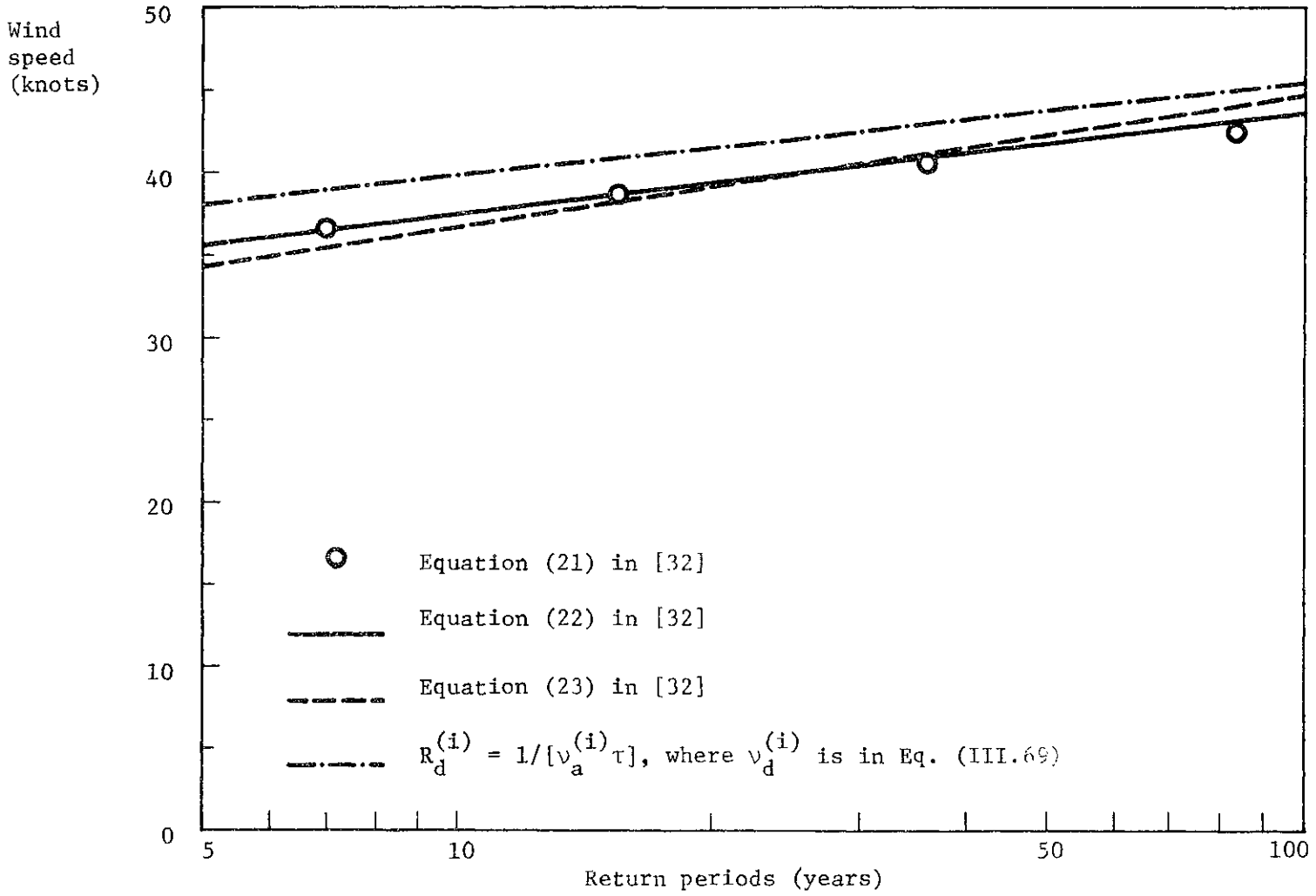


Figure III.12. Wind speed return period; figure base in reference [32].

CHAPTER IV

AUTOREGRESSIVE MODELS

IV.1 INTRODUCTION

Consider a deterministic system, S, (see Fig. (I.1) and Fig. (I.2) subjected to the (partially known) input sequence of Markovian type, X(t). The goal of this chapter is to optimize various engineering decisions; specific results are found for the problems (i), (ii) and (iii) defined at the beginning of Chap. I, when X(t) is a univariate Gauss Markov Sequence (GMS) generated by linear systems. The general form of a scalar k order, $k < \infty$, GMS (k-GMS) considered in this chapter, [13,30,76], is:

$$X(t+1) = \beta_0 + \beta_1 X(t) + \ldots + \beta_k X(t-k+1) + w(t) \qquad \text{(IV.1)}$$

where, $\{w(t)\}$ is a normal independent and identically distributed sequence with mean zero, and variance $1/h_k$; $\{w(t)\}$ is also called the driving white noise in Eq. (IV.1).

This study is restricted to Markovian Sequences of Gaussian type due only to computational facilities; however if there is a deterministic

(strictly monotonical) function, g(x), such that g(X(t)) is GMS, the methodology developed in this chapter applies also for the new state g(X(t)). For example, if g(x) = ln(x), the sequence X(t) is log normal distributed.

To achieve the final goal of, say, obtaining the optimal design of the system in Fig. (I.1) the steps in Sec. II.2 are followed; first the operational metamodel $\mathcal{H}^* = \{H_k\}_{k\in K}$ is postulated, where H_k is k - GMS in Eq. (IV.1) and $K = \{0,1,2,\ldots,k_o\}$, with k_o large enough. The peculiarity of $\mathcal{H}^*$ here in comparison with the general form of the operational metamodel consists in the fact that H_{k+1} identifies with H_k for $\beta_{k+1} \equiv 0$; because of this and the fact that X(t) is assumed Markovian only pure strategies (see Sec. III.1 and Sec. II.2) are to be considered. Thus, the optimal metamodel, $\mathcal{H}^{(opt)}$, is not necessarily to be calculated in this case since the optimal (pure) strategy, $H^{(opt)}$, is readily recovered directly from $\mathcal{H}^*$ (see also Table (II.38)).

The organization of this chapter is as follows. Sec. IV.2 develops the mathematical background to evaluate various types of errors related to the problem of optimal modeling X(t). Two specific forms of this error corresponding to the problem of the optimal size of a reservoir and to the problem of optimal operation of a reservoir are considered.

Section IV.3 solves the problems (i), (ii), and (iii) (see Chap. I) following the decision framework developed in Chap. II for the case when X(t) is GMS. The ingredients of the analysis are summarized in Sec. IV.3.1. Sec. IV.3.2 deals with the modification of the prior knowledge into posterior knowledge conditional on the operational metamodel $\mathcal{H}^*$ and the informative experiment. Special attention is given to the simul-

taneous prediction problem. Sec. IV.3.2.3 presents the problem in general terms, Sec. IV.3.2.4 deals with the (probabilistic) distribution of the extremes and sum of the states generated by GMS. The results in Sec. IV.3.2.4 are generalized in Sec. IV.3.2.5 to account for statistical uncertainty. From Sec. IV.3.2.3 it appears that the correlation and the order of the GMS, X(t), is not generally significant in reliability problems where failure occurs when X(t) exceeds high levels d, say, $d=m+2.5\sigma$ (where m and σ are the mean and the standard deviation of X(t), respectively). This observation suggests to treat these types (i.e. high threshold d) of safety problems by directly modeling X(t) as an independent sequence. An example in this sense is presented in Chap. III where the mean hourly wind data are used to infer the yearly wind extremes.

The result in Sec. IV.3.2.3 which concludes not to use correlated models in the reliability problems considered above brings significant computational simplifications.

Sec. IV.3.3 solves several engineering problems of the type (i), (ii) and (iii) considered in Chap. I. The optimal model, $H^{(opt)}$, of the sequence X(t) corresponding to the evaluation of the optimal size of a reservoir and to finding the optimal operation of a dam is the object of Sec. IV.3.3.1. Plots showing the error when the model H_ℓ is used instead of H_k for $\ell,k=1,2$ are presented. These curves give a rational base for choosing the order of the GMS of X(t) even without the effective evaluation of $H^{(opt)}$.

Sec. IV.3.3.2 calculates the optimal operation of a reservoir, and Sec. IV.3.3.3 considers the problem of optimal system design. Unfortunately (unless the problems of little practical interest when only one future

realization of X(t) is considered in design, i.e. the life of the system S is one unit of time) the problem of designing S is given only approximate solutions.

Sec. IV.4 analyses the results in this chapter considering their practical implications.

IV.2 MARKOV SEQUENCES

This section presents the mathematical background related to GMS defined in Eq. (IV.1) required in the further development of this chapter (the reader familiar with the theory of GMS may skip Sec. IV.2.2 and Sec. IV.2.3).

Thus, Sec. IV.2.1 derives the regression form of a GMS evolving over a finite period of time. This form is necessary to perform the prior to posterior analysis in Sec. IV.3.2; the errors evaluated in Sec. IV.2.2 and Sec. IV.2.3 are used to calculate the optimal model in two specific problems of type (i) (see also Sec. IV.3.3).

The results obtained in Section IV.2 do not require any distributional assumption on the white noise w(t) in Eq. (IV.1).

IV.2.1 Markov Sequences in Regression Form

Consider the scalar discrete time Markov Sequence of order k (k - MS)

$$X(t+1) = \beta_0 + \beta_1 X(t) + \ldots + \beta_k X(t-k+1) + w(t) \qquad \text{(IV.2)}$$

where, $E[w(t)] = 0$ for any t, and

$$E[w(t)w(t')] = \begin{cases} 1/h_k, & t = t' \\ 0, & t \neq t' \end{cases}$$

Then, Eq. (IV.2) written at t = k +1, k+2, ..., T is equivalent to the regression form:

$$\underline{Y}_k = \underline{X}_k \underline{\beta}_k + \underline{w} \qquad \text{(IV.3)}$$

where,

$$\underline{Y}_k = \begin{bmatrix} X(k+1) \\ X(k+2) \\ \vdots \\ X(T) \end{bmatrix}, \quad \underline{X}_k = \begin{bmatrix} 1\ X(k)\ldots X(1) \\ 1\ X(k+1)\ldots X(2) \\ \vdots \\ 1\ X(T-1)\ldots X(-k+T) \end{bmatrix} \qquad \text{(IV.4)}$$

$$\underline{\beta}_k = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \underline{w} = \begin{bmatrix} w(k) \\ w(k+1) \\ \vdots \\ w(T-1) \end{bmatrix}$$

If the state X(t) is measured at time t = 1,2,...,T, then the matrices $\underline{Y}_k$ and $\underline{X}_k$ are perfectly known; consequently Eq. (IV.3) is the standard form of a multiple regression model (Zellner [92], ρ. 65) where $\underline{\beta}_k$ is the vector of regression coefficients and $\underline{w}$ is the vector of disturbance.

The regression equations of form (IV.3) where the error term $\underline{w}$ is Gaussian with zero mean and covariance matrix $(1/h_k)\underline{I}$ ($\underline{I}$ is the identity maxtrix) is considered for instance in the references (Zellner [92], in

Chap. IV, Raiffa and Schlaifer [66] in Chap. 13).

IV.2.2 Error between two Markov Sequences

Consider the design of the reservoir size in Fig. (I.2) where the input variable is a MS. Let X(t) and $X_f(\cdot)$ be the actual and the postulated quantity of water entering the reservoir per unit of time, respectively. Then, the goodnes of the postulated model may be appreciated by the variance of the variable $D = [\sum_{t=1}^{\tau} X(t) - \sum_{t=1}^{\tau} X_f(t)]$, measuring the difference between the actual and the assumed quantity of water flowing in S during τ time units. Since the law of nature is not known, the error (i.e. the variance of D) evaluated in this section may be interpreted as the sensitivity of, say, the design of S when the input $X_f(t)$ is used instead of X(t).

Let X(t) be the state of the k - MS in Eq. (IV.2); then, always X(t) can be written in the following canonical form:

$$\underline{X}(t+1) = \underline{\xi} + \underline{\phi}\,\underline{X}(t) + \underline{G}\,w(t) \qquad \text{(IV.5)}$$

where,

$$\underline{X}(t) = \begin{bmatrix} X(t) \\ X(t-1) \\ \cdot \\ \cdot \\ \cdot \\ X(t-k+1) \end{bmatrix}, \quad \underline{\phi} = \left[\begin{array}{c|ccc} \beta_1 & \beta_2 & \cdots & \beta_k \\ \hline 0 & 1 & & \underline{0} \\ \cdot & & \ddots & \\ \cdot & & & \\ 0 & \underline{0} & & 1 \end{array}\right], \quad \underline{G} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad \underline{\xi} = \begin{bmatrix} \beta_0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \qquad \text{(IV.6)}$$

Consider then two scalar MS of states X(t) and $X_f(t)$ having the

canonical forms (see Eq. (IV.5)):

$$\underline{X}(t+1) = \underline{\xi} + \underline{\phi}\,\underline{X}(t) + \underline{G}\,w(t) \qquad \text{(IV.7)}$$

$$\underline{X}_f(t+1) = \underline{\xi}_f + \underline{\phi}_f(t) + \underline{G}_f\,w_f(t) \qquad \text{(IV.8)}$$

On augmenting the state $\underline{X}(t)$ and $\underline{X}_f(t)$, the following linear system is found

$$\overline{\underline{X}}(t+1) = \underline{\eta} + \overline{\underline{\phi}}\,\overline{\underline{X}}(t) + \overline{\underline{G}}\,\overline{\underline{w}}(t) \qquad \text{(IV.9)}$$

where,

$$\overline{\underline{X}}(t) = \begin{bmatrix} \underline{X}(t) \\ \underline{X}_f(t) \end{bmatrix}, \quad \overline{\underline{\phi}} = \begin{bmatrix} \underline{\phi} & \underline{0} \\ \underline{0} & \underline{\phi}_f \end{bmatrix}, \quad \overline{\underline{G}} = \begin{bmatrix} \underline{G} & \underline{0} \\ \underline{0} & \underline{G}_f \end{bmatrix},$$

$$\underline{\eta} = \begin{bmatrix} \underline{\xi} \\ \underline{\xi}_f \end{bmatrix}, \quad \overline{\underline{w}}(t) = \begin{bmatrix} w(t) \\ w_f(t) \end{bmatrix} \qquad \text{(IV.10)}$$

The steady-state covariance equation (see [76], p.58 or [30], p. 76) of the system (IV.9) (it is assumed here that all the eigenvalues of $\overline{\underline{\Phi}}$ lie inside the unit disc centered in the origin of the complex plane; the effect of this assumption is discussed in Sec. IV.3.3.1) satisfies the equation:

$$\underline{\Gamma} = \overline{\underline{\phi}}\,\underline{\Gamma}\,\overline{\underline{\phi}}^t + \overline{\underline{G}}\,\overline{\underline{Q}}\,\overline{\underline{G}}^t \qquad \text{(IV.11)}$$

where, $\underline{\Gamma}$ is the steady-state covariance matrix of the state $\overline{\underline{X}}(t)$ and $\overline{\underline{Q}}$ is the covariance matrix of the noise $\overline{\underline{w}}(t)$; the superscript "t" stands for matrix transposition.

Under the same conditions leading to Eq. (IV.11), the steady-state covariance function, $\underline{c}(\zeta)$, [30] is:

$$\underline{c}(\zeta) = \begin{cases} (\overline{\underline{\Phi}})^{\xi}\,\underline{\Gamma} & , \quad \xi \geq 0 \\ \underline{\Gamma}(\overline{\underline{\Phi}}^{\,t})^{\zeta} & , \quad \xi < 0 \end{cases} \tag{IV.12}$$

Since we focus only on one component in each vectorial state X(t) and $\underline{X}_f(t)$ (in fact any component of these vectorial states is the scalar state we are interested in), say, the components $X_1(t) = X(t)$ and $X_{f,1}(t) = X_f(t)$, the variance of the variable D (it is assumed that the average of X(t) and $X_f(t)$ coincide at any t) is:

$$E[(\sum_{t=1}^{\tau} X(t) - \sum_{t=1}^{\tau} X_f(t))^2] =$$

$$= \sum_{t,t'=1}^{\tau} \{[\underline{c}(t-t')]_{11} - [\underline{c}(t-t')]_{1,k+1} - [\underline{c}(t-t')]_{k+1,1} +$$

$$+ [\underline{c}(t-t')]_{k+1,k+1}\} \tag{IV.13}$$

where, k is the order of the MS generating X(t).

If the states X(t) and $X_f(t)$ are not correlated then $[\underline{c}(\zeta)]_{1,k+1} = 0$ for any ζ so that the error in Eq. (IV.13) is given by the sum of the covariance functions of X(t) and $X_f(t)$. In this case the error in (IV.13) is obtained by calculating independently the covariance functions of X(t) and $X_f(t)$. For our purpose the states $\underline{X}(t)$ and $\underline{X}_f(t)$ will be correlated; in general the noises driving the systems (IV.7) and (IV.8) will be the same at all times.

IV.2.3 Sensitivity Equations in Prediction

Consider the (existing) reservoir S in Fig. (I.2); let X(t) and $X_f(t)$ be the actual and the postulated amount of water entering S at time t, respectively. Assume that X(t) is measured a finite period of time ending at the present moment T. The optimal water release, y(T+1), depends on the future amount of water, say, X(T+1). Since the value of X(T+1) is not known, y(T+1) may be inferred from the best prediction, $\hat{X}(T+1|T)$, of X(T+1) conditional on all the measurements of X(t) ending at T. Assume that instead of $\hat{X}(T+1|T)$, the predictor $\hat{X}_f(T+1|T)$ to be defined below is considered. Then, the goodness of the postulated model is measured by the error:

$$EPRED = E[(X(T+1) - \hat{X}_f(T+1|T))^2] \qquad (IV.14)$$

where, $\hat{X}_f(T+1|T)$ is the "best" prediction of X(T+1) obtained from the model generating $X_f(t)$ and the measurements of X(t) at t = ...,T-1, T.

As in Sec. IV.2.2, since the law of X(t) is not known, (IV.14) is generally interpreted as the error (in one step prediction) when $X_f(t)$ is used instead of X(t) (see also the use of EPRED in Sec. IV.3.3.1).

To evaluate EPRED in Eq. (IV.14) the sensitivity equations of the theory of Kalman Bucy filters (KBF) are used. The main references on this topic are [17,18,30,33,43,59,60], (see also [76] for a general approach of KBF).

Clearly, one can evaluate the error in Eq. (IV.14) by considering a static estimation problem (see [76], p. 132); some numerical problems may

arise in this approach when the observation vector is very large. As will be seen this does not happen when KBF technique is considered since in this case the observation vector is processed sequentially.

Consider first the basic idea of the Kalman Bucy Filtering Theory. Let $\underline{X}(t)$ be the state of the zero mean (no loss of generality is associated with this assumption) MS generated by the linear system

$$\underline{X}(t+1) = \underline{\Phi}\,\underline{X}(t) + \underline{G}\,\underline{w}(t) \tag{IV.15}$$

where, $E[\underline{w}(t)] = \underline{0}$ for any t, and

$$E[\underline{w}(t_1)\,\underline{w}^t(t_2)] = \begin{cases} \underline{Q} & , \quad t_1 = t_2 \\ \underline{0} & , \quad t_1 \neq t_2 \end{cases}$$

observed through the linear noisy mechanism

$$\underline{Z}(t) = \underline{H}\,\underline{X}(t) + \underline{v}(t) \tag{IV.16}$$

where, $\underline{v}(t)$ is the error in the observation (for perfect observations $\underline{v}(t) = \underline{0}$ at all t); $\underline{v}(t)$ is white with mean zero and covariance $\underline{R}$.

The KBF assumes that the law of the random sequence, $\underline{X}(t)$, is perfectly known (i.e., that the matrices $\underline{\Phi}$, $\underline{G}$, and $\underline{Q}$ in Eq. (IV.15) are known) while the state $\underline{X}(t)$ is not known; consequently, the aim is to find the best estimate $\hat{\underline{X}}(t|t)$ and the best (one step) prediction $\hat{\underline{X}}(t+1|t)$ of the state $\underline{X}(t)$ and $\underline{X}(t+1)$, respectively, given the "noisy" measurements $\{\underline{Z}(1),\ldots,\underline{Z}(t)\}$.

All estimates and predictors considered in this section are linear and best in the minimum error covariance sense (see details in [76],

p. 97).

The KBF procedure relies on the following two theorems:

(1) Consider the equation [77]:

$$\underline{X}_2 = \underline{A}\,\underline{X}_1 + \underline{w} \tag{IV.17}$$

where, $\underline{X}_1, \underline{X}_2$ and $\underline{w}$ are random vectors such that $\underline{w}$ is zero mean and uncorrelated with $\underline{X}_1$.

Consider an informative experiment $\underline{Z}$ of $\underline{X}_1$ and denote by

$$\hat{\underline{X}}_1 = \underline{B}^*\underline{Z}$$

the "best" linear estimate of $\underline{X}_1$.

Then, the best estimate of $\underline{X}_2$ given $\underline{Z}$ is

$$\hat{\underline{X}}_2 = \underline{A}\,\hat{\underline{X}}_1 = \underline{A}\,\underline{B}^*\underline{Z} \tag{IV.18}$$

Consequently, denoting by $\hat{\underline{X}}(t|t)$ the best estimate of $\underline{X}(t)$ given the observation vector $\underline{Z}^t = \{\underline{Z}(1),\ldots,\underline{Z}(t)\}$, the best estimate of $\underline{X}(t+1)$ given $\underline{Z}$ (compare Eq. (IV.15) with Eq. (IV.17)), $\hat{\underline{X}}(t+1|t)$, which is also called the best one step prediction of $\underline{X}(t+1)$ is

$$\hat{\underline{X}}(t+1|t) = \underline{\Phi}\,\hat{\underline{X}}(t|t) \tag{IV.19}$$

(2) Consider the following correlated, zero mean, random variables $\underline{X}$, $\underline{Z}_1$, $\underline{Z}_2$. Then, the best estimate of $\underline{X}$ given $\underline{Z}_1$ and $\underline{Z}_2$, $\hat{\underline{X}}|_{\underline{Z}_1,\underline{Z}_2}$, is:

$$\hat{\underline{X}}|_{\underline{Z}_1,\underline{Z}_2} = \hat{\underline{X}}|_{\underline{Z}_1} + \hat{\underline{X}}|_{\underline{\nu}_2} \tag{IV.20}$$

where, $\hat{\underline{X}}|_{\underline{Z}_1}$ and $\hat{\underline{X}}|_{\underline{\nu}_2}$ are respectively, the best linear estimate of $\underline{X}$ given $\underline{Z}$ and the best linear estimate of $\underline{X}$ given $\underline{\nu}_2$, where $\underline{\nu}_2 = \underline{Z}_2 - \hat{\underline{Z}}_2|_{\underline{Z}_1}$ represents the new information for $\underline{X}$ brought by $\underline{Z}_2$ with respect to $\underline{Z}_1$ and is called the innovation vector; $\hat{\underline{Z}}_2|_{\underline{Z}_1}$ is the best estimate of $\underline{Z}_2$ given $\underline{Z}_1$.

Assume that in some way the estimate $\hat{\underline{X}}(t|t)$ was obtained; then, the estimate $\hat{\underline{X}}(t+1|t+1)$ of the state $\underline{X}(t+1)$ corresponding to observation $\{\underline{Z}(1),\ldots,\underline{Z}(t+1)\}$ (see Eq. (IV.19) and Eq. (IV.20)) has the form:

$$\hat{\underline{X}}(t+1|t+1) = \hat{\underline{X}}(t+1|t) + \hat{\underline{X}}(t+1)|_{\underline{\nu}(t+1)}$$

where, $\underline{\nu}(t+1) = \underline{Z}(t+1) - \underline{H}\,\hat{\underline{X}}(t+1|t)$

or (see [76], p. 136) since $\hat{\underline{X}}(t+1)|_{\underline{\nu}(t+1)} = \underline{K}(t+1)\underline{\nu}(t+1)$, it is:

$$\hat{\underline{X}}(t+1|t+1) = \underline{\phi}\,\hat{\underline{X}}(t|t) + \underline{K}(t+1)\underline{\nu}(t+1) \qquad \text{(IV.21)}$$

where $\underline{K}(t+1)$ is called the filter gain at step t+1.

Assume now that the Eq. (IV.15) and Eq. (IV.16) represent the evolution of the actual state and the actual observation mechanism, respectively. Consider that instead of the system (IV.15) the Markovian sequence $\underline{X}_f(t)$ is used to estimate/predict the actual state $\underline{X}(t)$, where,

$$\underline{X}_f(t+1) = \underline{\phi}_f\underline{X}_f(t) + \underline{G}_f\underline{w}_f(t) \qquad \text{(IV.22)}$$

has the observation pattern

$$\underline{Z}_f(t) = \underline{H}_f\underline{X}_f(t) + \underline{v}_f(t) \qquad \text{(IV.23)}$$

in which $\underline{w}_f(t)$ and $\underline{v}_f(t)$ are white noises of covariance matrices

$\underline{Q}_f$ and $\underline{R}_f$, respectively.

Let $\hat{\underline{X}}_f(t|t)$ be the best estimate of $\underline{X}(t)$ for given measurement $\{\underline{Z}(1),\ldots,\underline{Z}(t)\}$ in Eq. (IV.16) (in fact only $\underline{Z}(t)$ is actually measured); then, the optimal one step prediction of X(t+1) is

$$\hat{\underline{X}}_f(t+1|t) = \underline{\phi}_f\hat{\underline{X}}_f(t|t) \tag{IV.24}$$

The best estimate of $\underline{X}(t+1)$ conditional on observations $\{\underline{Z}(1),\underline{Z}(2),\ldots,\underline{Z}(t+1)\}$ corresponding to model (IV.22), (IV.23) results from Eq. (IV.21) and Eq. (IV.24)

$$\hat{\underline{X}}_f(t+1|t+1) = \underline{\phi}_f\hat{\underline{X}}_f(t|t) + \underline{K}_f(t+1)[\underline{Z}(t+1) - \underline{H}_f\underline{\phi}_f\hat{\underline{X}}_f(t|t)] \tag{IV.25}$$

or considering also Eq. (IV.16), it follows

$$\hat{\underline{X}}_f(t+1|t+1) = [\underline{I} - \underline{K}_f(t+1)\underline{H}_f\underline{\phi}_f]\hat{\underline{X}}_f(t|t) + \underline{K}_f\underline{H}\ \underline{X}(t+1) + \underline{K}_f(t+1)\underline{v}(t+1) \tag{IV.26}$$

On augmenting the states of the linear system (IV.15) and (IV.26) (see also [35]) one finds a linear system of state

$$\underline{y}(t) = \begin{bmatrix} \underline{X}(t) \\ \hat{\underline{X}}_f(t|t) \end{bmatrix} \tag{IV.27}$$

driven by the noise $\{\underline{w}^t(t)\underline{v}^t(t+1)\}$.

From Eq. (IV.15), Eq. (IV.19), and the covariance matrix of the state $\underline{y}(t)$ (see Eq. (IV.11) for the expression of the steady-state covariance matrix) the error EPRED in Eq. (IV.14) (for $X(t) = X_1(t)$ and $\hat{X}_f(t+1|t) = \hat{X}_{f,1}(t+1|t)$) is readily obtained.

IV.3 DECISION STRATEGY FOR MARKOV SEQUENCES

IV.3.1 Introduction

Let X(t) be the Gauss Markov input sequence relevant to a certain engineering problem; for instance, X(t) may be the load applied to S in Fig. (I.1).

The goal of the analysis in this section is to evaluate the optimal Gauss Markov model of X(t), i.e. the optimal code strategy (see Chap. I); also, at designer level the best design or the best operation of a specific system are found. To achieve this purpose the Bayesian Decision framework developed in Sec. II.1 and Sec. II.2 is used. The ingredients required in the analysis for the specific case of GMS are the following:

1. - the Bayesian description of the operational metamodel $\mathcal{H}^* = \{H_k\}_{k\in K}$, where H_k denotes the k-GMS in Eq. (IV.1) and $K = \{0,1,2,\ldots, k_0\}$. The prior uncertainty in $\mathcal{H}^*$ is measured by the prior probabilities, p'_k, of H_k being true (conditional on $\mathcal{H}^*$); also, the prior probability distribution of the unknown parameters in H_k, $f'_k(\underline{\beta}_k, h_k)$, has to be input.

2. - the experiment E, consists of perfect/noisy measurements of X(t) during a finite duration T. If X(t) is perfectly observed, then the measurements (observations) vector

$$\underline{Z}^t = \{Z_1 = X(1),\ldots, Z_T = X(T)\} \tag{IV.28}$$

contains all the sampling information.

3. - the loss (utility) function, which depends on the specific problem; several examples corresponding to the tasks (i), (ii) and (iii) of Chap. I are given in Sec. IV.3.3.

Since the elements of $\mathcal{H}^*$ have increasing degree of complexity (i.e. H_{k+1} becomes H_k for $\beta_{k+1} \equiv 0$) and since X(t) is assumed to be a Markovian process, it appears inconsistent to consider other than pure actions, H_k (see discussion about various action spaces in Sec. II.1). Consequently, the intermediate step of finding the optimal metamodel, $\mathcal{H}^{(opt)}$, (required only from computational reasons when composite actions are considered) is not necessary here because the optimal (pure) strategy, $H^{(opt)}$, is readily recovered from $\mathcal{H}^*$.

The next section derives the inferential elements (the posterior probability, p''_k, of H_k conditional on $\mathcal{H}^*$ and $\underline{Z}$ and the simple/simultaneous predictive distribution of H_k for given $\underline{Z}$ in Eq. (IV.28)) and Sec. IV.3.3 presents various cost functions corresponding to problems (i), (ii), and (iii) in Chap. I and evaluates the optimal code and designer decision.

IV.3.2 Prior to Posterior Analysis; Prediction Analysis

Consider the k order GMS in Eq. (IV.1), where the parameters $\underline{\beta}_k$ and h_k are not known, and assume that the measurements $\underline{Z}$ (Eq. (IV.28)) of the state X(t) are given.

The regression form of the dynamic system (IV.1) for the initial conditions $\{Z_1, Z_2, \ldots, Z_k\}$ (see also Eq. (IV.3) and Eq. (IV.28)) is:

$$\underline{Z}_k = \underline{X}_k \underline{\beta}_k + \underline{w} \qquad \text{(IV.29)}$$

where, $\underline{Z}_k{}^t = \{Z_{k+1} Z_{k+2} \cdots Z_T\}$

$\underline{\beta}_k{}^t = \{\beta_0 \ \beta_1 \cdots \beta_k\}$

$\underline{w}^t = \{w(k) \ w(k+1) \ \ldots \ w(T-1)\}$

$$\underline{X}_k = \begin{bmatrix} 1 & Z_k & Z_{k-1} & \cdots & Z_1 \\ 1 & Z_{k+1} & Z_k & \cdots & Z_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & Z_{T-1} & Z_{T-2} & \cdots & Z_{T-k} \end{bmatrix} \qquad \text{(IV.30)}$$

Consequently, the matrices $\underline{Z}_k$ and $\underline{X}_k$ in Eq. (IV.29) are perfectly known, being defined completely by the observation vector $\underline{Z}$. The regression form (IV.29) of the k-GMS in Eq. (IV.1), facilitates the evaluation of the posterior probability distribution of the parameters $(\underline{\beta}_k, h_k)$.

IV.3.2.1. Posterior distribution of the unknown parameters

Following ([92], p. 70; [66],p. 343; [38]) a conjugate informative prior probability distribution for $(\underline{\beta}_k, h_k)$, $f'_k(\underline{\beta}_k, h_k)$, is assumed. The conjugate distribution $f'_k(\underline{\beta}_k, h_k)$ in the regression equation (IV.29) is normal-gamma (in the analysis the order k of the Markov sequence is assumed known) or equivalently, the distribution of $\underline{\beta}_k$, conditional on h_k and k is multivariate normal with mean $\underline{\beta}'_k$ and variance $\underline{\Sigma}'_k/h_k$, i.e.

$$\underline{\beta}_k/h_k, k \in N(\underline{\beta}'_k, \underline{\Sigma}'_k/h_k) \qquad \text{(IV.31)}$$

and the distribution of the scale paramter h_k conditional on k is gamma-2 with parameters(v'_k, ν'_k) (see [66], p. 226, [94])

$$h_k/k \in G2(v'_k, \nu'_k) \qquad \text{(IV.32)}$$

The posterior distribution of $(\underline{\beta}_k, h_k)$ is then (see [92], p. 71; or [66], p. 343; or [38,39]) still normal-gamma, i.e.

$$\underline{\beta}_k/h_k, k, \underline{Z} \in N(\underline{\beta}''_k, \underline{\Sigma}''_k/h_k) \qquad \text{(IV.33)}$$

$$h_k / \; k, \underline{Z} \in G2(v_k'', \nu_k'') \tag{IV.34}$$

with parameters:

$$\begin{aligned}
(\underline{\Sigma}_k'')^{-1} &= (\underline{\Sigma}_k')^{-1} + \underline{X}_k{}^t \underline{X}_k \\
\underline{\beta}_k'' &= \underline{\Sigma}_k''[(\underline{\Sigma}_k')^{-1}\underline{\beta}_k' + \underline{X}_k{}^t \underline{Z}_k] \\
v_k''\nu_k'' &= v_k'\nu_k' + \underline{\beta}_k'^t(\underline{\Sigma}_k')^{-1}\underline{\beta}_k' - \underline{\beta}_k''^t(\underline{\Sigma}_k'')^{-1}\underline{\beta}_k'' + \underline{Z}_k{}^t \underline{Z}_k \\
\nu_k'' &= \nu_k' + T - k
\end{aligned} \tag{IV.35}$$

Direct application of the updating equations (IV.35) for large observation vectors $\underline{Z}$ may lead to numerical difficulties. To avoid this, first $\underline{Z}$ is partitioned, say, $\underline{Z}^t = \{\underline{Z}^t_{(1)}, \ldots, \underline{Z}^t_{(i)}, \ldots, \underline{Z}^t_{(q)}\}$, and then Eq. (IV.35) is applied sequentially for each piece $\underline{Z}_{(i)}$ of $\underline{Z}$ where the PDF of $(\underline{\beta}_k, h_k)$ before observing $\underline{Z}_{(i)}$ equals the posterior distribution of $(\underline{\beta}_k, h_k)$ after the measurements $\{\underline{Z}^t_{(1)}, \ldots, \underline{Z}^t_{(i-1)}\}$ have been processed.

From the analytical form of H_k and the posterior probability, $f_k''(\underline{\beta}_k, h_k)$ of the unknown parameters, $(\underline{\beta}_k, h_k)$, in H_k, the simple and simultaneous probability distribution of H_k are evaluated. Sec. IV.3.2.2 and Sec. IV.3.2.3 calculate the simple and the simultaneous predictive distribution of a finite order GMS, respectively. Approximate results for the simultaneous predictive distribution are evaluated in Sec. IV.3.2.4 and Sec. IV.3.2.5. Sec. IV.3.2.6 provides the posterior probability that H_k (conditional on $\mathcal{H}^*$) is true.

IV.3.2.2 Simple predictive distribution

The one step predictive distribution for a first order Gaussian autoregressive sequence (k=1) is given in ([92], p. 188). The general solu-

tion of the predictive distribution for any GMS of finite order is derived next. Assume that t=T is the present time when the last observation of the state X(t) generated by Eq. (IV.1) is made. The problem is to find the distribution of the next value of the state $\tilde{y} = X(T+1)$, based on the "Bayesian" description of H_k (see Sec. IV.3.1) and on the informative vector $\underline{Z}$ in (IV.28).

Equation (IV.1) at t=T gives

$$\tilde{y} = \underline{x}_k^t \underline{\beta}_k + w(T) \tag{IV.36}$$

$$\text{where, } \underline{x}_k^t = \{1 \; Z_T \; Z_{T-1} \; \cdots \; Z_{T-k+1}\} \tag{IV.37}$$

Then, conditional on h_k, $\tilde{y}$ is Gauss ($\underline{x}_k$ is perfectly known) with mean $\underline{x}_k^t \underline{\beta}_k''$ and variance $(1+\underline{x}_k^t \underline{\Sigma}_k'' \underline{x}_k)/h_k$; consequently the simple predictive distribution, $f_k(\tilde{y}/\underline{Z})$, is the result of the integration

$$f_k(\tilde{y}|\underline{Z}) = \int_0^\infty f_k(\tilde{y}|h_k,\underline{Z}) \; f_k(h_k|\underline{Z}) \; dh_k \tag{IV.38}$$

$$\text{where, } \tilde{y}|h_k,k,\underline{Z} \in N(\underline{x}_k^t \underline{\beta}_k'', [1+\underline{x}_k^t \underline{\Sigma}_k'' \underline{x}_k]/h_k) \text{ and}$$

$$h_k|k,\underline{Z} \in G2(v_k'', \nu_k'') \tag{IV.39}$$

On integrating (IV.38) it results that the variable

$$t_{\nu_k''} = (\tilde{y} - \underline{x}_k^t \underline{\beta}_k'')/\sqrt{v_k''(1 + \underline{x}_k^t \underline{\Sigma}_k'' \underline{x}_k)} \tag{IV.40}$$

is standardized student t with ν_k'' degrees of freedom; thus, from tables of standard student t distribution, one obtains:

- the probability that the next future value, $\tilde{y}$, of the state X(t) in Eq. (IV.1) exceeds a level d:

$$P_{f,k}(d) = P(\tilde{y}>d/\underline{Z},k) = 1 - F_{t_{\nu_k''}}\left(\frac{d - \underline{x}_k^t \underline{\beta}_k''}{\sqrt{v_k''(1+\underline{x}_k^t \underline{\Sigma}_k'' \underline{x}_k)}}\right) \tag{IV.41}$$

where, $F_{t_{\nu''_k}}(\cdot)$ denotes the cumulative distribution function of a standard student variable with ν''_k degrees of freedom, and,

- the P_f - upper fractile of the predictive distribution (IV.38)

$$d_k(P_f) = \underline{x}_k^t \underline{\beta}''_k + \sqrt{v_k(1+\underline{x}_k^t \underline{\Sigma}''_k \underline{x}_k)}\ t_{\nu''_k;(1-P_f)} \qquad \text{(IV.42)}$$

where, $t_{\nu''_k;(1-P_f)}$ represents the $(1-P_f)$-fractile of the standardized student t with ν''_k degrees of freedom.

In the particular case when ν''_k is large, say $\nu''_k \geq 30$, (see [92], p. 367), the student t variable $\tilde{y}$ in Eq. (IV.40) can be approximated by the Gaussian variable $\tilde{\tilde{y}}$, where

$$\tilde{y}/k,\underline{Z} \in N(\underline{x}_k^t \underline{\beta}'' \ , \ v''(1+\underline{x}_k^t \underline{\Sigma}''_k \underline{x}_k)) \qquad \text{(IV.43)}$$

Thus in this case (i.e. when $\nu''_k \geq 30$) equations (IV.41) and (IV.42) become respectively

$$P_{f,d}(d) \cong 1 - \Phi\left(\frac{d - \underline{x}_k^t \underline{\beta}''_k}{\sqrt{v''(1+\underline{x}_k^t \underline{\Sigma}''_k \underline{x}_k)}}\right) \qquad \text{(IV.44)}$$

where, $\Phi(\cdot)$ is the cumulative distribution function of the standard Gauss variable, and

$$d_k(P_f) \cong \underline{x}_k^t \underline{\beta}''_k + \sqrt{v''_k(1+\underline{x}_k^t \underline{\Sigma}''_k \underline{x}_k)}\ \Phi_{(1-P_f)} \qquad \text{(IV.45)}$$

where, $\Phi_{(1-P_f)}$ is the P_f - upper fractile of the

standardized normal distribution.

The results (IV.40), (IV.41), and (IV.42) or their approximations (IV.43), (IV.44), and (IV.45) represent the basic output of this section. The next section considers the more general case of evaluating the distribution of several future realizations of X(t) conditional on H_k and on $\underline{Z}$.

IV.3.2.3 Simultaneous predictive distribution

As in Section IV.3.2.2 we are given the order of the model, say k, its structure (see Eq. (IV.1)), the Bayesian description of H_k (i.e. the prior knowledge on the parameters $f'_k(\underline{\beta}_k, h_k)$) and the informative vector $\underline{Z} = \{Z_T = X(T),\ Z_{T-1} = X_{(T-1)},\ \ldots\ ,\ Z_1 = X(1)\}$. The aim of this section is to find the predictive distribution of the next τ future realizations of the state X(t) in Eq. (IV.1); therefore the following vectorial variable is considered:

$$\underline{\tilde{y}}_\tau^t = \{\tilde{y}_1 = X(T+1)\ ,\ \tilde{y}_2 = X(T+2)\ ,\ \ldots\ ,\ \tilde{y} = X(T+\tau)\} \qquad \text{(IV.46)}$$

Since,

$$f_k(\underline{\tilde{y}}_\tau|\underline{Z}) = \prod_{i=0}^{\tau-1} f_k(\tilde{y}_{\tau-i}|y_{\tau-i-1},\ \ldots\ ,\ \tilde{y}_1,\underline{Z}) \qquad \text{(IV.47)}$$

where, the variable $\tilde{y}_{\tau-i}|\tilde{y}_{\tau-i-1},\ldots,\tilde{y}_1,\underline{Z}$ has the same distribution as $\tilde{y}$ in Eq. (IV.40) for the informative vector $\{\tilde{y}_{\tau-i-1},\ldots,\tilde{y}_1,\underline{Z}\}$.

the analytical form of $f_k(\underline{\tilde{y}}_\tau|\underline{Z})$ is readily obtained from the results of Section IV.3.2.2. It can be seen that the predictive distribution of $\underline{\tilde{y}}_\tau$ is a product of τ student t distributions obtained from Eq. (IV.40) for

various observation vectors.

Consider now the following two derivates of $\tilde{y}_\tau$

$$\tilde{y}_{m,\tau} = \max_{T+1\leq t\leq T+\tau} \{X(t)\} \tag{IV.48}$$

$$\tilde{y}_{s,\tau} = \sum_{t=T+1}^{T+\tau} X(t) \tag{IV.49}$$

The predictive distribution of the extreme $\tilde{y}_{m,\tau}$ is of main interest in Structural Reliability; for example, consider the design of system S (Fig. (I.1)) of life time τ. The variable $\tilde{y}_{s,\tau}$ appears when the design of a reservoir size is considered (see Sec. IV.2.2 and Sec. IV.3.3.1).

The cumulative distribution function of the random variables (IV.48) and (IV.49) is derived by integrating Eq. (IV.47),

$$P(\tilde{y}_{\eta,\tau} \leq a) = P(\tilde{\underline{y}}_\tau \in D_\eta(a)) = \int_{D_\eta(a)} f_k(\tilde{\underline{y}}_\tau|\underline{Z})\, d\tilde{\underline{y}}_\tau \tag{IV.50}$$

where, η is m or s and

$$D_m(a) = \prod_{i=1}^{\tau} (-\infty, a) \tag{IV.51}$$

$$D_s(a) = \{\underline{y} : \sum_{i=1}^{\tau} y_i \leq a\} \tag{IV.52}$$

The domains $D_m(a)$ and $D_s(a)$ are shown in Fig. (IV.1), for $\tau = 2$.

Clearly, the computation of the multiple integral (IV.50) becomes soon prohibitive when τ increases (say $\tau>3$). Thus, the evaluation of the CDF of $\tilde{y}_{\tau,m}$ and $\tilde{y}_{s,m}$ from Eq. (IV.50) is practically inconvenient since in most of the engineering problems τ is large (for example, if the scale time in H_k is one day, then τ = 18,250 units for 50 years system S).

Consequently, approximate solutions of the probability $P(\tilde{y}_{\eta,\tau} \leq a)$

in Eq. (IV.50) are derived next. First, Sec. IV.3.2.4, the distribution of $\tilde{y}_{m,\tau}$ and $\tilde{y}_{s,\tau}$ is evaluated under the assumption that the order and the parameters of the GMS generating X(t) are perfectly known. Then, Sec. IV.3.2.5, approximate expressions of $P(\tilde{y}_{\eta,\tau} \leq a)$ for the case when only the order of the GMS X(t) is perfectly known are derived.

IV.3.2.4 Simultaneous distribution. Probabilistic Solution

In this section the order of the GMS as well as its parameters $(\underline{\beta}_k, h_k)$ are assumed perfectly known.

Consider first the variable $\tilde{y}_{m,\tau}$ in Eq. (IV.48), i.e. the maximum of X(t) during the time τ. Several approximations for the distribution of this variable are proposed next.

(a) Consider for simplicity the standard first order Gauss Markov sequence

$$X(t+1) = \rho\, X(t) + \sqrt{1-\rho^2}\, w(t) \qquad \text{(IV.53)}$$

where, $w(t) \in N(0,1)$ is the driving white sequence and ρ represents the (known) correlation coefficient at lag one

Assume that at t=0 the state equals Z. Then, the mean, m(t), and variance, $\gamma(t)$, equations of the state X(t), for $t \geq 0$ are respectively:

$$m(t) = \rho^t\, Z \qquad \text{(IV.54)}$$

$$\gamma(t+1) = \rho^2\, \gamma(t) + 1-\rho^2 \qquad \text{(IV.55)}$$

If $|\rho| < 1$, then the equations (IV.54) and (IV.55) become when t increases indefinitely,

$$m_\infty = \lim_{t\to\infty} m(t) = 0 \qquad \text{(IV.54')}$$

$$\gamma_\infty = \rho^2 \gamma_\infty + 1-\rho^2 \qquad \text{(IV.55')}$$

$$\text{or } \gamma_\infty = 1.0$$

where, $\gamma_\infty = \lim_{t\to\infty} \gamma(t)$

The evolution of the distribution of the state X(t), at t=0,1,2,.., is shown in the Fig. (IV.2).

From Eq. (IV.54) and (IV.55) result that the steady state solution (i.e. $X(t) \in N(m_\infty,\gamma_\infty)$) is approached sooner for less correlated sequences (in fact when $\rho=0$ there is no transient distribution).

Denote by $\underline{\Gamma} = \{\gamma_{ij}\}$ the covariance matrix of the vector $\underline{X}^t = \{X(1), \ldots,X(\tau)\}$ and by $\underline{\Gamma}^* = \{\gamma^*_{ij}\}$ the covariance matrix obtained from $\underline{\Gamma}$ by setting the off-diagonal elements of $\underline{\Gamma}$ equal to zero (in other words $\underline{\Gamma}^*$ is the covariance matrix of the vector $\underline{X}$ when the variables X(1), X(2), ..., X(τ) are assumed independent); consequently, $\gamma_{ij} \geq \gamma^*_{ij}$ for any $i,j = 1,2,\ldots,\tau$, provided that all $\gamma_{ij} \geq 0$. It follows that [78],

$$P(\tilde{y}_{m,\tau} \leq a) = P(\max_{1\leq t\leq\tau} \{X(t)\} \leq a) \geq \prod_{t=1}^{\tau} \Phi\left(\frac{a-\rho^t Z}{\sqrt{\gamma(t)}}\right) \qquad \text{(IV.56)}$$

where, $\Phi(\cdot)$ is the cumulative distribution function of the standard normal variable.

Clearly, when ρ is small the approximate solution (IV.56) is expected to work well. Moreover (see approximation (c), to follow, which

considers the steady-state behaviour of X(t)), (IV.56) seems to be an acceptable approximation for most values of ρ, say $|\rho| < 0.70$, if for example $a-\rho^t Z \geq 3\sqrt{\gamma(t)}$ at any $t=1,2,\ldots,\tau$ (see Fig. (IV.4)).

The approximation of $P(\tilde{y}_{m,\tau} \leq a)$ in Eq. (IV.56) depends strongly on the starting point Z when τ is small and ρ is large. But, if the future time τ is large, (IV.56) may be replaced by the steady-state approximation

$$P(\tilde{y}_{m,\tau} \leq a) \simeq [\Phi(a)]^{\tau} \qquad \text{(IV.56')}$$

The result (IV.56) corresponds to standard first order GMS (see Eq. (IV.53)). The propagation of the mean and covariance in finite order GMS (see Sec. IV.2.2) are readily obtained by state augmentation; consequently, the marginal mean and variance of the state can be calculated. Then, Eq. (IV.56) or Eq. (IV.56') provides an approximation of the cumulative distribution of the extreme $\tilde{y}_{m,\tau}$ in Eq. (IV.48). The essential advantage of approximation (a) is its computational simplicity.

(b) As a second approximation, consider the continuous analogue of the discrete time Markov sequence in Eq. (IV.53) (see details in [76], pp. 499-503).

$$\dot{X}(t) = -\beta X(t) + \sqrt{2\beta}\, w(t) \qquad \text{(IV.57)}$$

where, $\beta = -\ell n(\rho)$ and

$w(t)$ is standard Gaussian white noise.

It can be shown that the density $f(x,t;Z)$ of state X(t) in (IV.57) when $X(0) = Z$ satisfies the Fokker-Plank equation (see for instance [88],

p. 120)

$$\frac{\partial f}{\partial t} = \beta \frac{\partial}{\partial x}(xf) + (2\beta)\frac{\partial^2 f}{\partial x^2} \tag{IV.58}$$

The distribution of the first passage time of level x=a ([88], p. 20) is obtained from the solution of Eq.(IV.58) with, respectively, the initial and boundary conditions

$$\lim_{t\to 0} f(x,t;Z) = \begin{cases} 0 & , \quad x \neq Z \\ 1 & , \quad x = Z \end{cases}$$
$$f(a,t;Z) = 0, \text{ for any } t \tag{IV.59}$$

The evaluation of the distribution of first passage time can be also calculated as follows. First, from the solution of (IV.57)

$$X(t) = e^{-\beta t}Z + \int_0^t e^{-\beta(t-\tau)}\sqrt{2\beta}\, w(\tau)d\tau \tag{IV.60}$$

the distribution f(x,t;Z) satisfying the initial condition X(0) = Z is calculated; in fact due to linearity in (IV.60) and since $w(\tau)$ is Gaussian, it results $X(t) \in N(e^{-\beta t}Z,\ 1-e^{-2\beta t})$.

Then, following ([88], p. 9), the density of the first passage time variable of level x=a is

$$q(a;t) = -(2\beta)\left(\frac{\partial f(x,t;Z)}{\partial x}\right)\Big|_{x=a} \tag{IV.61}$$

Consequently, the CDF of the extremes $\tilde{y}_{m,\tau}$ results:

$$P(\tilde{y}_{m,\tau} \leq a) = \int_\tau^\infty q(a;t)dt \tag{IV.62}$$

The exact solution (IV.62) of the first passage time corresponding

to the continuous system (IV.57) is known only for Markovian processes of order one. This is a severe limitation since our purpose is to select the "optimal" order of the Markovian model of X(t). Also the accounting for uncertainty in the parameters of the model (β in Eq. (IV.57)) leads generally to difficult numerical integrations. Because of these drawbacks the solution in Eq. (IV.62) is of little interest in the further development of this chapter. Consequently, improvements of the approximations in (a) based on the crossing theory of random processes are developed next; Fig. (IV.4), Table (IV.76) and Eq. (IV.80) summarize the main results obtained in the next approximation.

(c) Consider the k order GMS in Eq. (IV.1); denoting by

$$\underline{X}(t) = \begin{bmatrix} X(t) \\ X(t-1) \\ \cdot \\ \cdot \\ \cdot \\ X(t-k+1) \end{bmatrix}, \quad \underline{\xi}_k = \begin{bmatrix} \beta_0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad \underline{\phi}_k = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_k \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \quad \underline{G}_k = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{IV.63}$$

the linear system (IV.1) takes the form:

$$\underline{X}(t+1) = \underline{\xi}_k + \underline{\phi}_k \underline{X}(t) + \underline{G}_k w(t) \tag{IV.64}$$

The propagation of the mean, $\underline{m}_k(t)$, and covariance $\underline{\Gamma}_k(t)$ of $\underline{X}(t)$ is calculated from the equations:

$$\underline{m}_k(t+1) = \underline{\xi}_k + \underline{\phi}_k \underline{m}_k(t) \tag{IV.65}$$

$$\underline{\Gamma}_k(t+1) = \underline{\phi}_k \underline{\Gamma}_k(t) \underline{\phi}_k^t + \frac{1}{h_k} \underline{G}_k \underline{G}_k^t \tag{IV.66}$$

The asymptotic (when $t \to \infty$) analogue of Eq. (IV.65) and (IV.66) exists only if all the eigenvalues of $\underline{\phi}_k$, $\lambda(\underline{\phi}_k)$, satisfy $|\lambda(\underline{\phi}_k)| < 1$. In this case the steady state mean $\underline{m}_{k,\infty}$ is:

$$\underline{m}_{k,\infty} = (\underline{I}_k - \underline{\phi}_k)^{-1} \underline{\xi}_k \qquad \text{(IV.65')}$$

and the steady state covariance, $\underline{\Gamma}_{k,\infty}$, satisfies the matrix equation:

$$\underline{\Gamma}_{k,\infty} = \underline{\phi}_k \underline{\Gamma}_{k,\infty} \underline{\phi}_k^t + \frac{1}{h_k} \underline{G}_k \underline{G}_k^t \qquad \text{(IV.66')}$$

Since the process, $\underline{X}(t)$ in Eq. (IV.64) is Gaussian, the steady-state joint distribution $f_k(X(t), X(t+1))$ is completely defined by $\underline{m}_{k,\infty}$ and $\underline{\Gamma}_{k,\infty}$ for $k > 1$. (Only steady state results are considered here; the transient response can be solved analogously.) The mean rate of upcrossings of X(t) in Eq. (IV.1) of level X=a (see Fig. (IV.3), is:

$$\nu_a^{(k)} = P(\{X(t+1) > a\} \cap \{X(t) \leq a\}) = \int_D f_k(x_1, x_2) dx_1 dx_2 \qquad \text{(IV.67)}$$

so that, the steady state mean rate of upcrossings is:

$$\nu_a^{(k)} = \int_{-\infty}^{(a-m_k)/\sigma_k} \phi(y) \left[1 - \Phi\left(\frac{a - m_{k,1/2}}{\sqrt{\gamma_{k,1/2}}}\right)\right] dy \qquad \text{(IV.68)}$$

where, $\phi(\cdot)$ and $\Phi(\cdot)$ represent, respectively, the PDF and CDF of the standard normal variable, m_k and σ_k are the steady state mean and standard deviation of the state X(t), in Eq. (IV.1), respectively (m_k and σ_k are obtained from Eq. (IV.65') and Eq. (IV.66'), respectively),

$$m_{k,1/2} = m_k + \frac{\gamma_{k,12}}{\sigma_k^2} y,$$

$$\gamma_{k,12} = E[(X(t+1) - m_k)(X(t) - m_k)], \text{ and}$$

$$\gamma_{k,1/2} = \sigma_k^2 - \gamma_{k,12}^2/\sigma_k^2$$

The (steady state) average rate of upcrossings of level X=a for the standard first order GMS (see Eq. (IV.53)) with the first lag correlation coefficient, ρ, is:

$$\nu_a^{(1)}|_\rho = \int_{-\infty}^{a} \phi(y)[1-\Phi(\frac{a-\rho y}{\sqrt{1-\rho^2}})]dy \qquad \text{(IV.69)}$$

where, $\phi(\cdot)$ and $\Phi(\cdot)$ have the same meaning as in Eq. (IV.68)

Comparing equations (IV.68) and (IV.69) it appears that the average rate of upcrossings of a k-GMS is computed exactly as the average rate of upcrossings of the standard 1-GMS in Eq. (IV.53), for the normalized threshold $a_1 = (a-m_k)/\sigma_k$ and one lay correlation coefficient $\rho = \gamma_{k,12}/\sigma_k^2$, i.e.

$$\nu_a^{(k)} = \nu_{a_1 = \frac{a-m_k}{\sigma_k}}^{(1)}|_{\rho=\gamma_{k,12}/\sigma_k^2} \qquad \text{(IV.70)}$$

Fig. (IV.4) shows plots of $\nu_a^{(1)}|_\rho$ in Eq. (IV.69) versus the threshold X=a for various values of the correlation coefficient ρ. It appears that when the level X=a is high (say $a \geq 3.0$) the mean rate of upcrossings is not sensitive to the correlation structure of the process X(t); consequently, the independent approximation in Eq. (IV.56) is expected to be sufficient-

ly accurate for most of the reliability problems where the failure event corresponds to exceedance of high levels X=a.

Consider now the modified version of (IV.56') where τ is replaced (see also [9]) by a reduced future time $\tau_a^{(k)}$ depending on the mean rate of upcrossings of level X=a for given model H_k. Following ([8], p. 194) the probability, $Q(\tau)$, that the extreme $\tilde{y}_{m,\tau}$ exceeds the level X=a when $X(0) < a$, is

$$Q(\tau) = \sum_{i=1}^{\bar{\tau}} Q_i(a,\ 0\leq t\leq\tau) \qquad \text{(IV.71)}$$

where, $\bar{\tau} = [(\tau+1)/2]$ ($[\cdot]$ denotes the largest integer function) and $Q_i(a,\ 0\leq t\leq\tau)$ is the probability of i upcrossings of level x=a during time τ.

The average number of upcrossings during time τ is:

$$N_+(a;\ 0\leq t\leq\tau) = \sum_{i=1}^{\bar{\tau}} i\ Q_i(a;\ 0\leq t\leq\tau)$$

so that, it is:

$$N_+(a;\ 0\leq t\leq\tau) - Q(\tau) = \sum_{i=2}^{\bar{\tau}} (i-1)\ Q_i(a;\ 0\leq t\leq\tau) \geq 0$$

Consequently,

$$Q(\tau) \leq N_+(a;\ 0\leq t\leq\tau) = \nu_a\tau \qquad \text{(IV.72)}$$

i.e. the average number of upcrossings during time τ, $\nu_a\tau$, provides an upper bound for the failure probability $Q(\tau)$. (The last equality in Eq. (IV.72) holds if the sequence X(t) is stationary.) Moreover when $\nu_a\tau$ is small [8] it is

$$Q(\tau) = P(\tilde{y}_{m,\tau} > a) = P(\max_{1\le t\le\tau}\{X(t)\} > a) \approx \nu_a \tau \qquad \text{(IV.73)}$$

From the approximate value of the probability of exceedence (IV.73) "the equivalent" time of independence, $\tau_a^{(k)}$, corresponding to the k-order GMS in (IV.1) and level X=a is readily evaluated. Consider the k-GMS, X(t), where $X(t) \in N(m_k,\sigma_k^2)$ in steady state and 0-GMS, $X^{(0)}(t)$, having the same marginal distribution as X(t). On equating the approximate probability of exceeding the level X=a (see Eq. (IV.73)) of both sequences X(t) and $X^{(0)}(t)$, it follows:

$$\tau_a^{(k)} = (\nu_a^{(k)}/\nu_a^{(0)})\tau \qquad \text{(IV.74)}$$

where, $\nu_a^{(0)} = \Phi(\frac{a-m_k}{\sigma_k})[1-\Phi(\frac{a-m_k}{\sigma_k})]$ is the mean upcrossings rate of the independent sequence $X^{(0)}(t) \in N(m_k,\sigma_k^2)$. Clearly, Eq. (IV.74) and Eq. (IV.75) hold only when the approximation in Eq. (IV.73) may be accepted.

Then, the distribution of the extreme $\tilde{y}_{m,\tau}$ in Eq. (IV.48) is approximated by:

$$P(\max_{1\le t\le\tau}\{X(t)\} \le a) \cong (F_X(a))^{\tau_a^{(k)}} \qquad \text{(IV.75)}$$

where, X(t) is generated by the linear system (IV.1), $F_X(x)$ is the marginal CDF of the state X(t) and $\tau_a^{(k)}$ represents the number of "independent" realizations of X(t) associated with the level X=a.

Clearly, both approximate results (IV.56) and (IV.75) coincide with

the exact solution when $\{X(t)\}$ is an independent sequence with marginal distribution $F_X(x)$. In addition, from Fig. (IV.4), Eq. (IV.74) and Table (IV.76) it appears that, unless the one lag correlation coefficient ρ in Eq. (IV.53) is large the independent approximation (IV.56') does well for high levels X=a (say, for $a \geq m+3\sigma$ where m and σ are the mean and the variance of X(t), respectively). However, if the correlation coefficient is, say, larger than 0.7 (see Table (IV.76)) the approximation (IV.75) may be preferred.

Table (IV.76)

ρ	$\nu_a^{(1)}/\nu_a^{(0)}$	
	a = 3 r.m.s.	a = 4 r.m.s.
0.0	1.0000	1.0000
0.2	0.9928	1.0000
0.4	0.9668	1.0000
0.5	0.9396	0.9961
0.6	0.8964	0.9767
0.7	0.8293	0.9349
0.8	0.7239	0.8509
0.9	0.5476	0.6757
0.95	0.4006	0.5081
0.96	0.3608	0.4603
0.97	0.3147	0.4038
0.98	0.2588	0.3342
0.99	0.1845	0.2397

Consider now the random variable $\tilde{y}_{s,\tau}$ defined by Eq. (IV.49), where the state X(t) is generated by the dynamic system (IV.1) of known parameters $(\underline{\beta}_k, h_k)$. From Equations (IV.65') and (IV.66') the steady state mean

$\underline{m}_{k,\infty}$ and the steady state covariance $\underline{\Gamma}_{k,\infty}$ of $\underline{X}(t)$ are obtained. The steady state covariance function, $\underline{c}_k(\zeta)$, of $\underline{X}(t)$ is (see [30]):

$$\underline{c}_k(\zeta) = \begin{cases} \underline{\phi}_k^{\zeta}\underline{\Gamma}_{k,\infty} & , \quad \zeta \geq 0 \\ \underline{\Gamma}_{k,\infty}(\underline{\phi}^t)^{\zeta} & , \quad \zeta < 0 \end{cases} \tag{IV.77}$$

where, $\underline{c}_k(\zeta) = E[(\underline{X}(t) - \underline{m}_{k,\infty})(\underline{X}(t') - \underline{m}_{k,\infty})^t]$

with $\zeta = t - t'$

Since the multivariate random variable $\underline{\tilde{Y}}^t = \{X(1) \ldots X(\tau)\}$ is normal and since $\tilde{y}_{s,\tau}$ is a linear transformation of $\underline{\tilde{Y}}$ it follows that $\tilde{y}_{s,\tau}$ is a univariate normal random variable. Consider for illustration the first order GMS:

$$X(t+1) = \beta_0 + \beta_1 X(t) + w(t) \tag{IV.78}$$

where, $w(t) \varepsilon N(0, 1/h_1)$

The steady state mean and covariance of the process X(t) in Eq. (IV.78) are, respectively:

$$m_1 = \beta_0/(1-\beta_1) \tag{IV.79}$$

$$c_1(\zeta) = \beta_1^{|\zeta|}/[h_1/(1-\beta_1^2)]$$

Thus, $\underline{\tilde{Y}}^t = \{X(1), \ldots, X(\tau)\}$ has the steady state mean $\underline{I}_1(\frac{\beta_0}{1-\beta_1})$ ($\underline{I}_1^t = \{1,\ldots,1\}$, $(1,\tau)$ matrix) and covariance

$$\underline{c}_{-1} = \frac{1}{h_1(1-\beta_1^2)} \begin{bmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & \beta_1^{\tau-1} \\ \beta_1 & 1 & \beta_1 & \cdots & \beta_1^{\tau-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_1^{\tau-1} & \beta_1^{\tau-2} & \cdots & \cdots & 1 \end{bmatrix}$$

Consequently, the linear combination of $\underset{\sim}{\tilde{Y}}$ (see [76], p. 520) which gives the variable $\tilde{y}_{s,\tau}$ has the distribution:

$$\tilde{y}_{s,\tau} = \sum_{t=1}^{\tau} X(t) = \underline{1}_{-1}^{t} \underset{\sim}{\tilde{Y}} \in N\left(\frac{\beta_0 \tau}{1-\beta_1}, \frac{\tau+2(\tau-1)\beta_1+2(\tau-2)\beta_1^2+\ldots+2\beta_1^{\tau-1}}{h_1(1-\beta_1^2)}\right) \tag{IV.80}$$

Thus, when the parameters $(\underline{\beta}_k, h_k)$ are assumed perfectly known the exact probabilistic characterization of the variable $\tilde{y}_{s,\tau}$ is readily obtained.

Clearly, any result containing many future realizations (i.e. large τ) leads to the problem of existence of the steady state probability distribution. If the steady state conditions are not satisfied, the variance of the steady state distribution of X(t) goes to infinity. This fact will restrict substantially the use of the results in this section to the case when the parameters of the linear system in Eq. (IV.1) are not perfectly known. The scope of the next section is to extend the output of this section to the case when the parameters $(\underline{\beta}_k, h_k)$ in (IV.1) are partially known.

IV.3.2.5 Simultaneous distribution. Statistical Solution

Since in Structural Reliability we focus on the long run (large τ) behaviour of the variables $\tilde{y}_{m,\tau}$ (Eq. (IV.48)) and $\tilde{y}_{s,\tau}$ (Eq. (IV.49)), the first step in a statistical design problem should be the solution of the steady state (Bayesian) distribution. (Recall that τ is the future time of study expressed in units of, the time step of the GMS; for instance, if τ is one month and the time step of the sequence X(t) is one hour, then τ = (30 days) × (24 hours) = 720 units.)

The evaluation of the steady state distribution leads to certain difficulties due primarily to the choice of the prior distribution on the parameters. More specifically, consider the first order GMS with (for simplicity) $\beta_0 = 0$

$$X(t+1) = \beta_1 X(t) + w(t) \qquad \text{(IV.81)}$$

The propagation of the marginal (Bayesian) distribution of the $\tilde{y}_t = X(T+t)$ for given $\underline{Z}$ in Eq. (IV.28) is obtained by integrating over $\tilde{y}_{t-1}^{t} = \{\tilde{y}_{t-1}, \tilde{y}_{t-2}, \ldots, \tilde{y}_1\}$ in Eq. (IV.47) as follows:

$$f_1(\tilde{y}_t|\underline{Z}) = \int_{\tilde{\underline{y}}_{t-1}} f_1(\tilde{y}_t, \tilde{\underline{y}}_{t-1}|\underline{Z})\, d\tilde{\underline{y}}_{t-1} \qquad \text{(IV.82)}$$

where, $f_1(\tilde{y}_t, \tilde{\underline{y}}_{t-1}|\underline{Z})$ is the simultaneous predictive distribution of the vector $\{\tilde{y}_t, \tilde{\underline{y}}_{t-1}\}$

Hypothetical prior and posterior distributions of the unknown parameter β_1 conditional on h_1 are shown in Fig. (IV.5).

Since the conjugate prior of β_1/h_1 is normal, the posterior of β_1/h_1 is still normal with mean β_1'' and variance Σ_1''/h_1 (see Eq. (IV.35)). Conse-

quently, with finite non-zero probability

$$P'' = \Phi\left(\frac{-1.0-\beta_1''}{\sqrt{\Sigma_1''/h_1}}\right) + \Phi^*\left(\frac{1.0-\beta_1''}{\sqrt{\Sigma_1''/h_1}}\right) > 0 \qquad \text{(IV.83)}$$

where, $\Phi(\cdot)$ is the CDF of the standard Gauss variable, and

$$\Phi^*(\cdot) = 1.0 - \Phi(\cdot)$$

the steady state condition, $|\beta_1| < 1$ (see Sec. IV.3.2), is violated. Thus, the computation in (IV.82) does diverge, meaning that the steady state Bayesian distribution (for the informative prior distribution of conjugate type) $f_1(\tilde{y}_\infty|\underline{Z}) = \lim_{t\to\infty} f_1(\tilde{y}_t|\underline{Z})$, has infinite variance.

To overcome this difficulty the following considerations might be made:

(a) Since the processes we deal with have generally finite variance (i.e. finite power), the prior distribution (see Fig. (IV.5)) on the parameters defining the correlation structure (β_1 in Eq. (IV.81)) is not correct. Thus, a different prior distribution such that $f_1'(\beta_1/h_1) = 0$ for $\beta_1 \notin (-1,1)$ and any $h_1 > 0$ should be assigned. This assumption leads to a posterior distribution, $f_1''(\beta_1|h_1)$, which satisfies the same condition and therefore the steady state requirement, $|\beta_1| < 1$, is fulfilled with probability one. Although attractive logically, this approach has no practical appeal since no analytical closed form solution for the predictive distribution, posterior probability of the model being true, etc. is known, if other than a conjugate prior distribution is assumed.

(b) A crude approximation might be to use an estimate $(\hat{\underline{\beta}}_k, \hat{h}_k)$ of $(\underline{\beta}_k, h_k)$ as being the true value of the parameters of H_k (see Eq. (IV.1)).

If the estimate $\hat{\underline{\beta}}_k$ satisfies the steady state condition (see details in Eq. (IV.65), (IV.66), (IV.65'), and (IV.66')) then the marginal steady state distribution of the state of H_k is readily obtained by using the results in Sec. IV.3.2.4. Since the estimate $(\hat{\underline{\beta}}_k, \hat{h}_k)$ may have a large variance the error in this approximation can be significant.

(c) Instead of using one set of values for the unknown parameters $(\underline{\beta}_k, h_k)$, several sets $(\beta_{k,i}, h_{k,j})_{i,j}$ might be considered. This corresponds to the discretization of the prior distribution $f'_k(\underline{\beta}_k, h_k)$ of H_k; consequently H_k is replaced by the set of models $H_{k,ij}, k\varepsilon K$ ($H_{k,ij}$ is obtained from the model H_k for $\underline{\beta}_k = \beta_{k,i}$ and $h_k = h_{k,j}$) with known parameters and prior believes $p'_{k,ij} = p'_k \cdot p'_{ij}$ (where p'_{ij} is the prior belief in $H_{k,ij}$ conditional on H_k). Consider for exemplification the first order GMS, H_1:

$$X(t+1) = \beta_0 + \beta_1 X(t) + w(t) \qquad \text{(IV.84)}$$

Assume for instance that ten possible values for each parameter are accepted. It follows that 10^3 (probabilistic) models should be considered in the analysis. Clearly, such an approach is numerically prohibitive, even for k = 1. Consider then the mixed approach where the parameters $\{\beta_1, \beta_2, \ldots, \beta_k\}$ in Eq. (IV.1) have fixed values and only β_0 and h_k are assumed unknown. The regression equation (IV.29) becomes

$$\underline{Z}_k = \underline{I}_1 \beta_0 + \underline{\mathcal{Z}}_k + \underline{w} \qquad \text{(IV.85)}$$

where, $\underline{Z}_k$ and $\underline{w}$ have the same meaning as in (IV.29) and

$$\underline{\mathcal{Z}}_k = \begin{bmatrix} Z_k & Z_{k-1} & \cdots & Z_1 \\ Z_{k+1} & Z_k & \cdots & Z_2 \\ & & & \\ Z_{T-1} & Z_{T=2} & \cdots & Z_{T-k} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \text{(matrix multiplication)}$$

$$\underline{I}_1^t = \{1 \ldots\ldots 1\} \qquad \text{(T-k elements)}$$

As in Eq. (IV.31),(IV.32), the conjugate prior distribution on the unknown parameters

$$\beta_0/h_k,k \in N(\beta'_{0k}, \Sigma'_k/h_k)$$

$$h_k/k \in G2(v'_k,\nu'_k) \tag{IV.86}$$

are considered. Consequently (see Eq. (IV.33) and (IV.34)) the posterior distribution of the unknown parameters has the form:

$$\beta_0/h_k,k \in N(\beta''_{ok}, \Sigma''_k/h_k)$$

$$h_k/k \in G2(v''_k,\nu''_k) \tag{IV.87}$$

where (see Eq. (IV.35)) it is

$$1/\Sigma''_k = 1/\Sigma'_k + (T-k)$$

$$\beta''_{ok} = \Sigma''_k[\beta'_{ok}/\Sigma'_k + \underline{I}_1^t\underline{\tilde{Z}}_k]$$

$$v''_k\nu''_k = v'_k\nu'_k + \beta'_{ok}{}^2/\Sigma'_k - \beta''_{ok}{}^2/\Sigma''_k + \underline{\tilde{Z}}_k^t\underline{\tilde{Z}}_k$$

$$\nu''_k = \nu'_k + T - k \tag{IV.88}$$

$$\text{with the notation } \underline{\tilde{Z}}_k = \underline{Z}_k - \underline{\hat{\beta}}_k$$

Following Eq. (IV.36), (IV.37), (IV.38) it results that the simple predictive distribution of the state is univariate student with ν''_k degrees of freedom, mean $\beta''_0 + \beta_1 Z_T + \beta_2 Z_{T-1} + \ldots + \beta_k Z_{T-k+1}$ and variance

$v''_k \nu''_k (1+\Sigma''_k)/(\nu''_k-2)$.

Referring again to the sequence (IV.84), the predictive distribution for various β_1 values, is readily obtained. Assuming that β_1 can take only the values $\{\beta_{11}, \beta_{12}, \ldots, \beta_{ij_1}\}$ with (prior) probabilities $\{p'_{11}, p'_{12}, \ldots, p'_{1j_1}\}$ ($\sum_{i=1}^{j} p'_{1i} = 1$) respectively, it follows that the model H_1 with (unknown) β_0, β_1 and h_1 and prior probability p'_1 can be approximated by the set, $\{H_{1i}\}_{i=1,\ldots,j_1}$, of models ($H_{1i}$ has unknown β_0 and h_1 but known correlation coefficient $\beta_1 = \beta_{1i}$) with prior probabilities $p'_{1,1i} = (p'_1)(p'_{1i})$, $i = 1,2,\ldots,j_1$.

Clearly, this procedure is feasible for GMS of low order, say $k = 1,2$. This latter approach requires much less computation than the methodology developed at the beginning of this paragraph. For example, assuming ten values for each parameter in Eq. (IV.84), the number of models to be operated is 10^3 and 10 in the first and in the second methodology, respectively.

(d) Assume that τ in Eq. (IV.48) is sufficiently large so that the search for the steady state solution of the state X(t) in Eq. (IV.1) is justified. Then, if the asymptotic predictive distribution $\lim_{t\to\infty} \{f_k(\tilde{y}_t|\underline{Z})\} = f_k(\tilde{y}_\infty|\underline{Z})$ exists, and is symmetric, the Gaussian assumption on $f_k(y_\infty|\underline{Z})$ may be appropriate; since the steady-state mean, $E[\tilde{y}_\infty/\underline{Z}]$, and variance, $VAR[\tilde{y}_\infty/\underline{Z}]$ of $\tilde{y}_\infty/\underline{Z}$ can be obtained recursively, then the full probabilistic characterization of $\tilde{y}_\infty/\underline{Z}$ results.

In summary, the following (statistical) approximations for the distribution of the extreme, $\tilde{y}_{m,\tau}$, in Eq. (IV.48) and of the sum, $\tilde{y}_{s,\tau}$, in Eq. (IV.49) are recommended:

(1). Approximations for $y_{m,\tau}$

$(1)_1$ - Use the method in paragraph (b) of this section, under the steady state assumption. Then, from Eq. (IV.56') the probability that $\tilde{y}_{m,\tau}$ exceeds the level X=a is

$$P(\tilde{y}_{m,\tau}>a/\underline{Z},\underline{\hat{\beta}}_k,\hat{h}_k) \cong 1-[F_{\tilde{y}_\infty|\underline{\hat{\beta}}_k,\hat{h}_k}(a)]^{\tau} \cong \tau[1-F_{\tilde{y}_\infty|\underline{\hat{\beta}}_k,\hat{h}_k}(a)] \quad (IV.89)$$

where, $F_{\tilde{y}_\infty|\underline{\hat{\beta}}_k,\hat{h}_k}(\cdot)$ is the cumulative distribution function of the asymptotic state $\tilde{y}_\infty$ conditional on $\underline{\beta}_k = \underline{\hat{\beta}}_k$ and $h_k = \hat{h}_k$.

On the contrary if $\hat{\tau}_a^{(k)} = \tau_a^{(k)}|_{\underline{\beta}_k = \hat{\beta}_k, h_k=\hat{h}_k}$ differs significantly from τ (see Eq. (IV.74)) the approximation (IV.75) is proposed instead; consequently the complementary CDF of $\tilde{y}_{m,\tau}$ is evaluated from:

$$P(\tilde{y}_{m,\tau}>a\ \underline{Z},\underline{\hat{\beta}}_k,\hat{h}_k) \cong 1-[F_{\tilde{y}_\infty|\underline{\hat{\beta}}_k,\hat{h}_k}(a)]^{\hat{\tau}_a^{(k)}} \cong \hat{\tau}_a^{(k)}[1-F_{\tilde{y}_\infty|\underline{\hat{\beta}}_k,\hat{h}_k}(a)] \quad (IV.90)$$

$(1)_2$ - Use the method in paragraph (c) when no discretization of β_0 and h_k is assumed; then analogous relations as in Eq. (IV.89) and (IV.90) are obtained. The only difference lies in the computation of $\hat{\tau}_a^{(k)}$ (see (IV.90)) which in this case is obtained by integrating $\hat{\tau}_a^{(k)}$ for given $\beta_0, \beta_1, \ldots, \beta_k, h_k$ with respect to the posterior distribution $f''_k(\beta_0,h_k)$.

Clearly,the approximations $(1)_1$ and $(1)_2$ might be crude but it seems that they are the only solutions which can be generally handled numerically.

$(1)_3$ - If the steady state solution does exist approximations of the dis-

tribution of the extreme $\tilde{y}_{m,\tau}$ similar to (IV.89) and (IV.90) are easily obtained. For instance, within the Gaussian assumption upon the steady state Bayesian distribution the extreme $\tilde{y}_{m,\tau}$ generated by system (IV.81) has the probability of exceeding a level X=a

$$P(\tilde{y}_{m,\tau} > a|\underline{Z}) \cong 1 - [\Phi(\frac{a - E[\tilde{y}_\infty/\underline{Z}]}{\sqrt{Var[\tilde{y}_\infty/\underline{Z}]}})]^{\tau} \qquad \text{(IV.91)}$$

when $\tau \cong \hat{\tau}_a^{(k)}$, and

$$P(\tilde{y}_{m,\tau} > a/\underline{Z}) \cong 1 - [\Phi(\frac{a - E[\tilde{y}_\infty/\underline{Z}]}{\sqrt{Var\ [\tilde{y}_\infty|\underline{Z}]}})]^{\hat{\tau}_a^{(k)}} \qquad \text{(IV.92)}$$

where, $\hat{\tau}_a^{(k)}$ is obtained from Eq. (IV.74) for $\underline{\beta}_k = \hat{\underline{\beta}}_k$, $h = \hat{h}_k$, $(\hat{\underline{\beta}}_k, \hat{h}_k)$ being estimates of $(\underline{\beta}_k, h_k)$

when $\hat{\tau}_a^{(k)}$ differs significantly from τ.

(2) Approximation for $\tilde{y}_{s,\tau}$ in Eq. (IV.49)

$(2)_1$ - Use estimates $(\hat{\underline{\beta}}_k, \hat{h}_k)$ for $(\underline{\beta}_k, h_k)$ (see paragraph (b)), which leads directly to the distribution of the sum $\tilde{y}_{s,\tau}$ (see illustration for the case of first order GMS in Eq. (IV.80)). In this approximation the statistical steady state result is identical to the probabilistic steady state solution with parameters $\underline{\beta}_k = \hat{\underline{\beta}}_k$, $h_k = \hat{h}_k$.

$(2)_2$ - Use discretization of $\underline{\beta}_k$, say $\{\underline{\beta}_{k,i}\}_i$ and obtain the distribution of $\tilde{y}_{s,\tau}$ for known $\underline{\beta}_k = \underline{\beta}_{k,i}$ conditional of h_k. Then, on integrating $\tilde{y}_{s,\tau}/h_k, \underline{\beta}_{k,i}$ weighted by $f''_k(h_k)$ the resulting variable $\tilde{y}_{s,\tau}/\underline{\beta}_{k,i}$ is univariate student t with ν''_k degrees of freedom for any value $\underline{\beta}_{k,i}$. If the prior belief in model H_k is p'_k, then the prior probability of H_k with

$\underline{\beta}_k = \underline{\beta}_{k,i}$, i.e. the model $H_{k,i}$, is $p'_{k,i} = p'_k \cdot p'_{i|k}$ ($p'_{i|k}$ is the prior belief in $H_{k,i}$ conditional of H_k being correct).

Consequently, one replaces the model H_k by the set of models $\{H_{k,i}\}_i$ where only the variance, $1/h_k$, of the driving noise is not assumed perfectly known. Thus, for each model $\{H_{k,i}\}_{\substack{k=1,2,\dots \\ i=1,2,\dots}}$ the distribution of the sum, $\tilde{y}_{s,\tau}$, is readily obtained.

The disadvantage of this approach consists in the large number of models $\{H_{k,i}\}_{k,i}$ which have to be processed; in fact when higher order models H_k are to be considered in the analysis this methodology is practically inconvenient. Therefore, in most of the cases either the approximation $(2)_1$ or a very crude discretization in $(2)_2$ is preferred.

IV.3.2.6 Posterior Probability of Models

So far, the analysis was done conditional on the order, k, of the GMS. Consider now the operational metamodel $\mathcal{H}^* = \{H_k\}_{k\in K}$ (see Sec. IV.3.1); the posterior probability p''_k of the model H_k, $k\in K$, conditional on the metamodel $\mathcal{H}^*$ is computed next.

Following ([92], p. 297) it is

$$p''_k \propto p'_k \, f_k(\underline{Z}_k) \tag{IV.93}$$

where, p'_k is the prior probability of model H_k and $f_k(\underline{Z}_k)$ is the likelihood of the observation vector of model H_k, i.e.

$$f_k(\underline{Z}_k) = \int f_k(\underline{Z}_k|\underline{\beta}_k, h_k) \, f'_k(\underline{\beta}_k, h_k) \, d\underline{\beta}_k \, dh_k \tag{IV.94}$$

Consider the derivation of $f_k(\underline{Z}_k)$ in Eq. (IV.94) for various informative situations:

(1) - Assume that all the parameters $(\underline{\beta}_k, h_k)$ which define H_k

are perfectly known, say $\underline{\beta}_k = \underline{\beta}_{k,o}$ and $h_k = h_{k,o}$; consequently

$$f'_k(\underline{\beta}_k, h_k) = \begin{cases} 1 \ , \ \underline{\beta}_k = \underline{\beta}_{k,o} \text{ and } h_k = h_{k,o} \\ 0 \ , \text{ otherwise} \end{cases} \tag{IV.95}$$

and the density $f_k(\underline{Z}_k)$ in Eq. (IV.94) becomes

$$f_k(\underline{Z}_k) = f_k(\underline{Z}_k | \underline{\beta}_{ko}, h_{k,o}) = \left(\frac{h_{k,o}}{2\pi}\right)^{T-k} \exp\left[-\frac{h_{k,o}}{2}(\underline{Z}_k - \underline{X}_k \underline{\beta}_{k,o})^t (\underline{Z}_k - \underline{X}_k \underline{\beta}_{k,o})\right] \tag{IV.96}$$

This is the likelihood which has to be used in the method developed in paragraph (b) of Sec. IV.3.2.5.

(2) - Consider now the most general case when all the parameters $(\underline{\beta}_k, h_k)$ are assumed not to be known. From the regression equation

$$\underline{y}_k = \underline{X}_k \underline{\beta}_k + \underline{w} \tag{IV.97}$$

where, $\underline{X}_k$, $\underline{\beta}_k$, $\underline{w}$ have the same meaning as in Eq. (IV.29) ($f'_k(\underline{\beta}_k, h_k)$ is the distribution of $(\underline{\beta}_k, h_k)$ considered in the regressive form (IV.97)) and $\underline{y}_k$ is defined by the equality (IV.97)

and Eq. (IV.94), the distribution of $\underline{y}_k$ is

$$f_k(\underline{y}_k) = \int dh_k \, f'_k(h_k) \int d\underline{\beta}_k \, f_k(\underline{y}_k | \underline{\beta}_k, h_k) \, f'_k(\underline{\beta}_k | h_k) \tag{IV.98}$$

On integrating in (IV.98) first with respect to $\underline{\beta}_k$ and then over h_k the distribution $f_k(\underline{y}_k)$ is obtained. Its value at $\underline{y}_k = \underline{Z}_k$ is (see also Halpern in Eq. (IV.35)).

$$f_k(\underline{Z}_k) = (2\pi)^{-(\nu_k''-\nu_k')/2} \frac{|\underline{\Sigma}_k''|^{\frac{1}{2}}}{|\underline{\Sigma}_k'|^{\frac{1}{2}}} \frac{(v_k'\nu_k'/2)^{\nu_k'/2}}{(v_k''\nu_k''/2)^{\nu_k''/2}} \frac{\Gamma(\nu_k''/2)}{\Gamma(\nu_k'/2)} \qquad \text{(IV.99)}$$

(3) - Various particular cases of the result (IV.99) can be easily obtained. For instance if $\beta_1,\ldots,\beta_k$ are known and only β_0 and h_k are considered unknown (see Eq. (IV.85) through (IV.88)) a straighforward computation leads to

$$f_k(\underline{Z}_k) = (2\pi)^{-(\nu_k''-\nu_k')/2} \sqrt{\frac{\Sigma_k''}{\Sigma_k'}} \frac{(v_k'\nu_k'/2)^{\nu_k'/2}}{(v_k''\nu_k''/2)^{\nu_k''/2}} \frac{\Gamma(v_k''/2)}{\Gamma(v_k'/2)} \qquad \text{(IV.100)}$$

where, Σ_k'', v_k'', ν_k'' are computed in Eq. (IV.88).

This last result corresponds to the second approximation developed in the paragraph (c) of Sec. IV.3.2.5.

Once, $f_k(\underline{Z}_k)$ is found the posterior (Bayesian) probability of the model H_k, p_k'', being the true model, conditional on the metamodel $\mathcal{H}^*$ and the informative experiment $\underline{Z}$ is readily obtained. From Eq. (IV.93) and either results (IV.96) or (IV.99) or (IV.100) it is

$$p_k'' = c^{-1} p_k' f_k(\underline{Z}_k) \qquad \text{(IV.101)}$$

$$\text{where,} \quad c = \sum_{k\in K} p_k' f_k(\underline{Z}_k) \qquad \text{(IV.102)}$$

All the results derived in this section are valid if the prior probability distribution $f_k'(\underline{\beta}_k, h_k)$, for any $k\in K$, is informative; the probabi-

lities p''_k of H_k corresponding to vague prior knowledge on $(\underline{\beta}_k, h_k)$ can not be obtained by increasing indefinitely the variance of $f'_k(\underline{\beta}_h, h_k)$ in either of the results (IV.96) or (IV.99) or (IV.100) (see also Sec. III.2). On the other hand, on integrating Eq. (IV.98) for the noninformative prior $f'_k(\underline{\beta}_k, h_k) = C_k/h_k$ (see [92], p. 41-55, and [38,39]) it results:

$$p''_k = p'_k c\, c_k\, (2\pi)^{-\nu''_k/2} \frac{|\underline{\Sigma}''_k|^{1/2}\Gamma(\nu''_k/2)}{(\nu''_k v''_k/2)^{\nu''_k/2}} \qquad \text{(IV.102)}$$

where, $\nu''_k = T-2k-1$, and

c has to satisfy $\sum_{k\in K} p''_k = 1.0$

The difficulty comes when explicit values are to be assigned to the coefficients $\{c_k\}$. Halpern ([38], pp. 61-66, pp. 125-259) recommends taking $c_k = (2\pi)^{-k/2}$. (The small difference between the exponent of (2π) in Eq. (IV.102) and in Eq. (5.23) of [38], is due to the fact that the regression equation associated with a k-GMS has the dimension T-k, where T is the length of the informative vector $\underline{Z}$).

Consider now the decision framework in equations (II.3),(II.4), and (II.5). This section calculated the posterior probability p''_k of H_k entering the evaluation of $E[a/\mathcal{H}^*]$, (see Eq. (II.4)) where a and $\mathcal{H}^*$ are the action and the operational metamodel, respectively.

Sec. IV.3.3 presents cost functions, $c(a,\underline{y})$ corresponding to various engineering problems, where $c(a,\underline{y})$ denotes the deterministic loss when action a is chosen and the next τ future realizations, $\underline{\tilde{y}}$, of X(t) in Eq. (IV.1) are $\underline{\tilde{y}}^t = \underline{y}^t = \{y_1, y_2, \ldots, y_\tau\}$. Consequently, $L(a,H_k)$ in Eq. (II.3) is generally the average of $c(a,\underline{y})$ with respect to the simple

(simultaneous) predictive distribution $f_k(\tilde{\underline{y}}/\underline{Z})$ in Eq. (IV.40) (Eq. (IV.47)). Since generally one is interested in functions of $\tilde{\underline{y}}$, the predictive distributions of $\tilde{y}_{m,\tau}$ in Eq. (IV.48) and $\tilde{y}_{s,\tau}$ in Eq. (IV.49) have been evaluated in Sec. IV.3.2.5; in fact only the cost functions $c(a,\tilde{y})$, $c(a,\tilde{y}_{m,\tau})$ and $c(a,\tilde{y}_{s,\tau})$ are considered in the next section.

IV.3.3. Loss Functions. Decision Rules

The general framework of the decision theoretical approach to model selection is outlined in Sec. II.2. This section presents several illustrations of problems of type (i), (ii), and (iii) considered in Chap. I for the case when the input variable X(t) is a GMS. Because of the special character of the operational metamodel, $\mathcal{H}^*$, (see Sec. IV.1) only pure strategies (see Sec. II.2) are considered.

IV.3.3.1 Optimal Modeling of Random Sequences

The development in this section rests on the results presented in Sec. IV.2.2 and Sec. IV.2.3 (see also the task (i) in Chap. I). Consider then the random sequence X(t) measured (with random errors) during a finite time period T (for example X(t) might represent wind daily extremes, or daily flows in a river), [53]. The problem here is to find the model which fits best, in some sense, the available information. The cost functions associated with the ℓ-GMS strategy, H_ℓ, are obtained from the errors $\tilde{L}(H_\ell,H_k)$ of choosing H_ℓ instead of H_k, $k\in K$, ($\tilde{L}(H_\ell,H_k)$ corresponds to a specific goal; see Sec. IV.2.2 and Sec. IV.2.3) and from the penalty c_ℓ associated with the analytical complexity of H_ℓ.

Example 1. Consider the operational metamodel $\mathcal{H}^* = \{H_k\}_{k\in K}$ and let $X^{(k)}(t)$ be the state of H_k. Then, (see Sec. IV.2.2) the error

$$\tilde{L}(H_\ell, H_k) = E[(\sum_{t=T+1}^{T+\tau} X^{(\ell)}(t) - \sum_{t=T+1}^{T+\tau} X^{(k)}(t))^2] \qquad \text{(IV.103)}$$

where, T is the present time

seems adequate to find the optimal (code) strategy $H^{(opt)}$ when the evaluation of a reservoir size is considered. In the forgoing discussion $\tilde{L}(H_\ell, H_k)$ is interpreted (see also Sec. IV.2.2) as the variance of the random variable $D = \sum_{t=T+1}^{T+\tau} X^{(\ell)}(t) - \sum_{t=T+1}^{T+\tau} X^{(k)}(t)$ measuring the difference between the total amount of water flowing in the reservoir during the future time τ when H_ℓ is chosen instead of H_k.

The normalization of (IV.103) with respect to the second moment of the total quantity of water which enters the reservoir S (Fig. (1.2)) under H_k, i.e.

$$\tilde{\ell}(H_\ell, H_k) = E[(\sum_{t=T+1}^{T+\tau} X^{(\ell)}(t) - \sum_{t=T+1}^{T+\tau} X^{(k)}(t))^2] / E[(\sum_{t=T+1}^{T+\tau} X^{(k)}(t))^2] \qquad \text{(IV.104)}$$

was numerically evaluated in the steady state regime [35] for GMS of order $k,\ell = 1,2$, with mean zero and variance one. Figures (IV.6a,b,c,d,e,f,g, h,i) show the error $\tilde{\ell}(H_\ell, H_k)$ versus $\beta_{1|H_k} \equiv$ RO1 $\in [0.05, 0.95]$ for $\tau = 6$ and for the following selected values of the parameters of H_k and H_ℓ: $\beta_{2|H_k} \equiv$ RO2 = - 0.4, 0.0, 0.4; $\beta_{1|H_\ell} \equiv$ ROF1 = 0.05(0.15)0.95 and $\beta_{2|H_k}$ = ROF2 = - 0.4, 0.0, 0.4. The basic idea in [35] was to evaluate confidence intervals, $\mathcal{J}$, of the actual correlation coefficient (RO1, RO2) and then from the plots in Fig. (IV.6) to infer the best strategy H_ℓ, i.e. the model which minimizes the maximum error $\tilde{\ell}(H_\ell, H_k)$ which can be achieved in $\mathcal{J}$. In addition to this minimax strategy various measures, m(RO1,RO2), on (RO1,

RO2) can be considered in which case the best H_ℓ will minimize the expectation of $\tilde{\ell}(H_\ell,H_k)$ with respect to m(RO1,RO2). In the decision framework of this study the measure m(•) equals the posterior probability distribution of the unknown parameters $(\underline{\beta}_k,h_k)$ in H_k; following [38,39] and considering the results in Sec. IV.3.2.1 the expected loss when H_ℓ is chosen instead of H_k is:

$$L(H_\ell,H_k) = \int d\underline{\beta}_k dh_k \; f_k''(\underline{\beta}_k,h_k)\; \tilde{L}(H_\ell,H_k) \qquad \text{(IV.105)}$$

where, $f_k''(\beta_k,h_k)$ is the posterior distribution of the parameters $(\underline{\beta}_k,h_k)$ defining the model H_k (see Eq. (IV.33) and Eq. (IV.34)); the result of the integration in (IV.105) is a function of the parameters $(\underline{\beta}_\ell,h_\ell)$ of action H_ℓ.

Since τ (see Eq. (IV.103)) is generally large and the steady state conditions may be violated with nonzero probability one has to, say, discretize $f_k'(\underline{\beta}_k,h_k)$ (see sec. IV.3.2.5); consequently, the posterior $f_k''(\underline{\beta}_k,h_k)$ will be a discrete distribution so that $L(H_\ell,H_k)$ in (IV.105) becomes a weighted sum of ordinates measured in Fig. (IV.6).

The expected cost of the strategy H_ℓ (see Eq. (II.21) and Eq. (II.25)) conditional on $(\underline{\beta}_\ell,h_\ell)$ is:

$$E[H_\ell|\underline{\beta}_\ell,h_\ell] = c_\ell + \sum_{k\in K} L(H_\ell,H_k)\; p_k'' \qquad \text{(IV.106)}$$

where, c_ℓ is the penalty associated with the complexity of H_ℓ (see also [39]) and

p_k'' is the posterior probability of model H_k.

The optimal code strategy is obtained in two steps: first, the "best" values, $(\underline{\beta}_\ell, h_\ell)$, of the parameters, $(\underline{\beta}_\ell, h_\ell)$, of H_ℓ are obtained from:

$$E[H_\ell | \check{\underline{\beta}}_\ell, \check{h}_\ell] = \inf_{(\underline{\beta}_\ell, h_\ell)} \{E[H_\ell | \underline{\beta}_\ell, h_\ell]\} \qquad \text{(IV.107)}$$

and second, the "best" strategy (strategy in this context means order of the GMS and parameter values), $H^{(opt)} = H_\ell^*$ with parameters $(\check{\underline{\beta}}_\ell^*, \check{h}_\ell^*)$ from Eq. (IV.107) is obtained from:

$$E[H_\ell^* | \underline{\beta}_\ell^*, h_\ell^*] = \min_{k \in K} \{E[H_\ell | \hat{\underline{\beta}}_\ell, \hat{h}_\ell]\} \qquad \text{(IV.108)}$$

The evaluation of $H^{(opt)}$ requires generally a significant amount of work; however this is not as important since $H^{(opt)}$ is evaluated once, say, at code level. Then, $H^{(opt)}$ is used as if it were the actual law of the input variable X(t) (see Fig. (I.2)) in the designer phase to calculate the size of any specific reservoir S.

Example 2. Assume now that the choice of the GMS is linked to the optimal operation of a system (see Sec. IV.2.3); the error EPRED in Eq. (IV.14) seems adequate for this purpose. Numerical evaluation of EPRED has been done for the case when X(t) and $X_f(t)$ are zero mean, unit variance (in steady state) states generated by 2 - GMS with correlation coefficients β_1 = RO1 ε [0.05, 0.95], β_2 = RO2 = - 0.4, 0.4 and $\beta_{1,f}$ = ROF1 = 0.05(0.15) 0.95, β_{2f} = ROF2 = - 0.4, 0.0, 0.4, respectively. Also, the situation in which both X(t) and $X_f(t)$ are 1-GMS is considered. Figures (IV.7) present EPRED versus RO1 for various combinations of (RO2, ROF1, ROF2) and for an intensity, R, of the observation noise of 0.1 (units of the steady state variance of X(t)).

These plots can be directly used [35], to choose the "best" model of the input sequence, which here is also referred to as the optimal code strategy. Considering the decision framework in Sec. II.2 (see also Example 1 of this section), the penalty, $\tilde{L}(H_\ell, H_k)$, when H_ℓ is chosen instead of H_k is

$$\tilde{L}(H_\ell, H_k) = E[(\hat{X}^{(\ell)}(T+1|T) - X^{(k)}(T+1))^2] \qquad \text{(IV.109)}$$

i.e. $\tilde{L}(H_\ell, H_k)$ = EPRED in Eq. (IV.14) where $X(t)$ and $X_f(t)$ are replaced by $X^{(k)}(t)$ and $X^{(\ell)}(t)$, respectively.

The posterior value, $L(H_\ell, H_k)$, of $\tilde{L}(H_\ell, H_k)$ is obtained from Eq. (IV.105). As in the first example it is recommended to replace H_k by the set $H_{k,i}$ having known correlation structure, $\underline{\beta}_{k,i}$ (see approximation in paragraph (c) of Sec. IV.3.2.5). For instance, assuming that h_k is known and that the posterior of $H_{k,i}$ is $p''_{k,i}$, the expected cost $L(H_\ell, H_k)$ becomes the subsequent weighted sum of ordinates, $\{\tilde{L}(H_\ell, H_{k,i})\}_i$, in Fig. (IV.7):

$$L(H_\ell, H_k) = \sum_i p''_{k,i} \tilde{L}(H_\ell, H_{k,i}) \qquad \text{(IV.110)}$$

Then, following Eq. (IV.106), Eq. (IV.107), and Eq. (IV.108) the optimal code strategy is evaluated.

The loss considered in the examples of this section correspond to the case when the cost functions of the systems ruled by code are not well known; consequently, utility functions corresponding to various errors (see $L(H_\ell, H_k)$ in Eq. (IV.103) and Eq. (IV.110)) have been considered.

The next two sections assume that the cost functions of all the systems designed or operated in accordance with the code rules are perfectly known. Consequently, losses expressed in dollar units are to be consi-

dered.

IV.3.3.2 Optimal Operation of an (Existing) System

Consider for illustration a reservoir S of known size q which at time T contains L(T) cubic meters (c.m.) of water (see (ii) in Chap. I). Suppose, we choose to release y = y(T+1) c.m. of water at time T+1 (time T is the present time, say in a scale of months) and let X(T+1) be the actual (unknown) quantity of water (c.m.) which flows in S at future time T+1.

Let $\mathcal{H}^* = \{H_k\}_{k\in K}$ be the operational metamodel of X(t) and $\underline{Z}$ in (IV.38) be the informative experiment.

Then, in accordance with the strategy of releasing y c.m. of water at time T+1, the following components of the loss are considered in the analysis:

- penalty $c_{y_o}(y)$ of releasing y instead of y_o, where y_o measures the demand (see Fig. (IV.9) and,
- two constant penalties corresponding to the cases when the water overflows the reservoir, D_f, and when the reservoir is emptied, D_d.

It is easy to see that the sets $\{x : x < y - L(T)\}$ and $\{x : x > y - L(T) + q\}$ give all the possible inputs, {x}, which (for chosen y and conditional on L(T) and q) lead, respectively, to dries and overflows of the reservoir. Consequently (see Eq. (II.19)), the expected cost of the release y conditional on the code strategy H_ℓ is

$$E[y|H_\ell] = c_{y_o}(y) + D_d F_{H_\ell}(y-L(T)) + D_f[1-F_{H_\ell}(y-L(T)+q)] \qquad \text{(IV.111)}$$

where, $F_{H_\ell}(\cdot)$ is the one step predictive cumulative distribution of H_ℓ for given $\underline{Z}$ (see Sec. IV.3.2.2).

and the optimal release, $y^{(\ell)}$, under H_ℓ satisfies

$$E(y^{(\ell)}/H_\ell] = \inf_{\{y:y\geq 0\}} \{E[y/H_\ell]\} \qquad \text{(IV.112)}$$

Then, the expected cost of the strategy H_ℓ, $E[H_\ell]$, (conditional on $\mathcal{H}^*$) follows from Eq. (II.21)

$$E[H_\ell] = \sum_{k\in K} R(H_\ell, H_k) p''_k \qquad \text{(IV.113)}$$

$$\text{where, } R(H_\ell, H_k) = c_\ell + E[y^{(\ell)}/H_k] \qquad \text{(IV.114)}$$

has the same meaning as $R(\eta, H_k)$ in Eq. (II.17).

The optimal code (pure) strategy, $H^{(opt)} = H_\ell{}^*$, minimizes the expected cost in (IV.113). Then the optimal release, y_s, of any specific system S is obtained from Eq. (IV.112) where H_ℓ is replaced by $H^{(opt)}$ and the cost functions in Eq. (IV.111) correspond to S.

Clearly, the decision framework considered in this example is a simplified scheme of the complex problem of optimal operating a system; several assumptions we have made are:

- the operation rule was very simple, namely to give as much water as is available in order to achieve y,
- the input sequence X(t) was assumed stationary; however if the non-stationarity of X(t) is only in the mean, m(t), and variance $\sigma^2(t)$, the development in this section applies if X(t) is replaced by the normalized sequence.

$$\tilde{X}(t) = \frac{X(t) - m(t)}{\sigma(t)}$$

- in the evaluation of the optimal strategy, y_s, associated with S

only the next future realization of the actual state X(t) is considered. Clearly, a more accurate approach would be to consider at least a few future realizations of X(t); however this extension leads to a significant increase in computation.

The next section considers the important reliability problem of optimal system design (see (iii) in Chap. I).

IV.3.3.3 Optimal Design of a System

Consider the (deterministic) system S in Fig. (I.1). The optimal design level of S, $d^{(opt)}$, subjected to the random input X(t) is analyzed next. Since the design examples we have chosen for illustration depend essentially on the y_m, (in Eq. (IV.48)) and y_s, (in Eq. (IV.49)), (see Sec. IV.3.2.5) only approximate solutions will be obtained.

Example 1. Consider the design against wind of a tall building for τ years. The problem is, first, to evaluate the optimal, say level IV, code strategy (see Table (II.38)) and second, to best design any system S, ruled by the code when the informative experiment consists in J years of, say, mean hourly wind speed readings.

Two extreme situations may be considered. If J is large (say, $J \geq 40$) then an independent model with time step of one year (see Sec. III.2) is recommended. When J is small (say, $J \leq 10$) one has to choose a model (correlated or not; see Fig. (IV.4)) with a time scale of, say, days, or hours in order to use most of the data points. To be more specific we are interested in the modeling of the yearly wind extremes from, say, mean hourly speeds. To achieve this goal it is assumed that at the monthly scale the mean hourly process is stationary; (see Fig. (IV.10)). Let $\underline{Z}_i^t = \{\{V_{i1}(t)\}_t, \{V_{i2}(t)\}_t, \ldots, \{V_{iJ}(t)\}_t\}$ be the informative vector con-

taining all the hourly measurements in month i, i=1,2,...,12, and let us consider only pure code strategies, $\{H_\ell\}$, (see Sec. II.1) i.e. GMS of the same order all over the year; however the predictive distribution of H_ℓ varies with the month.

The following steps lead to the evaluation of the optimal code strategy, $H^{(opt)} = H_\ell^*$. First, the distribution, $F_{H_\ell}^{(i)}(\cdot)$, of the extreme $\tilde{y}_{m,\zeta}^{(i)} = \max_{1\le t\le\zeta}\{V_i(t)\}$, (where ζ = 720 hours and $V_i(t)$ is the extreme wind speed in the t^{th} day of month i) in month i is inferred from $\underline{Z}_i$ for given H_ℓ, following any of the approximations in Sec. IV.3.2.5.

Then, assuming independence between the monthly extremes the cumulative distribution of the yearly extremes, $F_{H_\ell}(\cdot)$, conditional on H_ℓ is $F_{H_\ell}(y) = \prod_{i=1}^{12} F_{H_\ell}^{(i)}(y)$. From the independence assumption between yearly extremes, the predictive distribution of the maximum wind in the next τ year is readily calculated. (The independent assumption of monthly and yearly extremes is generally accepted.)

The expected cost of the code corresponding to the design level X=d and the strategy H_ℓ is the same as in Eq. (II.9):

$$E[d|H_\ell] = c(d) + c_F[1 - (F_{H_\ell}(d))^\tau] \qquad \text{(IV.115)}$$

The optimal design, $d^{(\ell)}$, under H_ℓ minimizes the expected cost in Eq. (IV.115). The expected cost of strategy H_ℓ, $E[H_\ell]$, is obtained from Eq. (II.21) by replacing η with H_ℓ; it results

$$E[H_\ell] = \sum_{k\in K} R(H_\ell, H_k)p_k'' \qquad \text{(IV.116)}$$

where, (see Eq. (IV.17))

$$R(H_\ell, H_k) = c_\ell + E[d^{(\ell)}/H_k] \qquad \text{(IV.117)}$$

The penalty c_ℓ measures the computational cost of using the strategy H_ℓ.

and consequently the optimal (τ-years) code strategy $H_\tau^{(opt)} = H_\ell^*$, satisfies

$$E[H_\ell^*] = \min_{\ell \in K} \{E[H_\ell]\} \tag{IV.118}$$

Then (see Sec. II.2), the strategy $H^{(opt)}$, i.e. the τ-years extreme cumulative distribution $[F_{H_\tau^{(opt)}}(\cdot)]$, is used to evaluate the optimal design of any specific system. For instance, the expected cost of the design X=d of, say, system S in Fig. (I.1) is obtained from Eq. (IV.115) where $\{c(d), c_F\}$ denote the cost functions of S and $F_{H_\ell}(\cdot)$ is replaced by $F_{H_\tau^{(opt)}}(\cdot)$. Then, by minimizing the expected cost, $E^{(S)}[d/H_\ell^*]$, associated with S, the optimal level IV wind design $d_{IV} = d^{(opt)}$ of S (see also Table (II.38)) results.

This methodology is expected to produce good results even for highly correlated sequences and moderate levels X=d. The major source of error in this approach consists in the evaluation of the predictive distribution, $F_{H_\ell}^{(i)}(\cdot)$, of monthly extremes. From the wind example in Chap. III it appears that (see also Fig. (IV.4)) the hourly mean sequence may be assumed independent if the level X=d is sufficiently large; consequently a similar independent model is developed in Chap. III. Consider now the design of the size of reservoir S in Fig. (I.2).

Example 2. The problem here is to find the optimal size of a reservoir when the amount of water which enters the reservoir is random and partially known in the sense that J years of, say, daily flows of water, $V_{ij}(t)$, are

available($V_{ij}(t)$ is the flow in the t^{th} day of month i, of year j), [87].

Let us first assume that the law of the flow process is perfectly known (probabilistic design).

The initial goal is to find the distribution of the sum

$$V_i = \sum_{t=1}^{30} V_i(t) \tag{IV.119}$$

where, $V_i(t)$ is the random process which gives the flow in the t^{th} day of month i.

The distribution of the variable (IV.119) is readily obtained following the procedure in Sec. IV.3.2.4 (see for illustration Eq. (IV.80). Assuming that $\{V_i\}$, i=1,...,12, are independent (Gaussian) variables, the excess or the deficit of water in month m, Q_m, is still normally distributed with mean and variance directly obtained from Eq. (IV.120)

$$Q_m = \sum_{i=1}^{m} V_i - \frac{m}{12} \sum_{i=1}^{12} V_i \tag{IV.120}$$

Consequently, the distribution of the extremes of $\{Q_m\}$ m=1,...,12 under the assumption of independence of the variable Q_m for various m, are:

$$F_{\max_{1\le m\le 12}\{Q_m\}}(\xi) \cong \prod_{m=1}^{12} F_{Q_m}(\xi) \tag{IV.121}$$

$$F_{\min_{1\le m\le 12}\{Q_m\}}(\xi) \cong \prod_{m=1}^{12} (1-F_{S_m}(\xi)) \tag{IV.122}$$

and the range per year is $R = \max_{1\le m\le 12}\{Q_m\} - \min_{1\le m\le 12}\{Q_m\} \ge 0$.

Assuming (for the simplicity of exposition) independence between the variables in Eq. (IV.121) and Eq. (IV.122) (more accurate assumptions can

be considered) the range R is obtained directly by convolution. Denote the cumulative distribution function of R by $F_R(\cdot)$; then, the loss in τ years associated with the size q is

$$E[q] = c(q) + D_f[1 - F_R^{\tau}(q)] \qquad \text{(IV.123)}$$

where, c(q) is the cost of building a reservoir of size q and D_f is the penalty associated with the failure to provide the water demand due to the fact that not all the water is stored. (Clearly, when the total amount of water is stored no penalty for lack of water is assigned. However, to fulfill this requirement $q \to \infty$ and consequently $c(q) \to \infty$.)

To account for statistical uncertainty, the development in the example 1 of this section is followed. Assume that $V_i(t)$ is stationary on a monthly scale and evaluate the predictive distribution of V_i for all i conditional on the strategy H_ℓ (see also Sec. IV.3.2.5). Then, from Eq. (IV.120) through Eq. (IV.123) and from the definition of R, the predictive cumulative distribution of the range, $F_{R|H_\ell}(\cdot)$, and the expected cost of designing at level q, $E[q/H_\ell]$, are evaluated conditional on H_ℓ. The optimal design level, $q^{(\ell)}$, under H_ℓ satisfies:

$$E[q^{(\ell)}/H_\ell] = \inf_{\{q:q\geq 0\}} \{E[q|H_\ell]\} \qquad \text{(IV.124)}$$

Finally, the expected cost, $E[H_\ell]$, of H_ℓ is calculated following Eq. (IV.116); the optimal (τ- years) code strategy, $H_\tau^{(opt)} = H_\ell^*$, minimizes $\{E[H_\ell]\}_\ell$. The designer step consists in the computation of the optimal

size of S by minimizing E[q] in (IV.123), where the cost functions correspond to S and $F_R(\cdot)$ is replaced by $F_{H_\tau(opt)}(\cdot)$.

IV.4 CONCLUSIONS

The development in this chapter applies the general decision framework established in Sec. II.2 for the case when the input variable (see for instance Fig. (I.1)), X(t), is a univariate (Gauss) Markov Sequence. The inferential phase of the analysis is considered in Sec. IV.3.1 and Sec. IV.3.2. Related to the cost functions required by the decision approach two situations (see Sec. IV.3.3) have been considered. First, (Sec. IV.3.3.2) following Halpern [38], losses corresponding to errors which seem to be significant in various engineering problems, such as the optimal design of a reservoir size or the optimal operation of a dam, are calculated. Sensitivity plots giving the (expected) error when, say, a ℓ-GMS, H_ℓ, is chosen instead of a k-GMS, H_k, for a variety of values of the parameters in H_ℓ and H_k, $\ell,k=1,2$, are shown. Second, Sec. IV.3.3.2 and Sec. IV.3.3.3, it is assumed that dollar values can be assigned to any pair {designer action, nature outcome}; then, the deterministic cost functions obtained in this way are directly used to evaluate the optimal code and designer strategies following Sec. II.2.

In most of the engineering problems the optimal designer action depends on, say, the next τ future realizations of the input variable X(t), $\tilde{\underline{y}}_\tau$; consequently, the simultaneous predictive distribution, $f_k(\tilde{\underline{y}}_\tau/\underline{Z})$, of $\tilde{\underline{y}}_\tau$ (or functions of this random variable) conditional on H_k, $k\in K$, has to

be evaluated. Because of the numerical difficulties of using $f_k(\tilde{\underline{y}}_\tau/\underline{Z})$ in various calculations required by our approach the search for approximate expressions of the simultaneous predictive distribution is justified. To be specific approximations of the predictive distribution of the maximum $\tilde{y}_{m,\tau} = \max_{T+1\leq t\leq T+\tau}\{X(t)\}$ and of the sum $\tilde{y}_{s,\tau} = \sum_{t=T+1}^{T+\tau} X(t)$ have been derived. From the approximate expression of the distribution of $\tilde{y}_{m,\tau}$ it appears that in reliability problems where the failure event occurs when X(t) exceeds at least once during τ the threshold $d = m + 3.0\sigma$ (m and σ are the mean and the standard deviation of X(t), respectively) and the one lag correlation coefficient of X(t) is, say, smaller than 0.7, the optimal design, $d^{(opt)}$, (see Fig. (I.1)) of S is not sensitive to the correlation structure of X(t). On the contrary, the form of the marginal distribution of X(t) seems to be significant for the value of $d^{(opt)}$ (see Chap. III). These observations have important practical implications, for example, assuming that the mean hourly wind speeds form an independent sequence, the distribution of the wind yearly extremes can be well estimated (see also Sec. III.3.2).

On the other hand in problems such as the optimal design of a reservoir size, $q^{(opt)}$, where the variable $\tilde{y}_{s,\tau}$ has to be considered the correlation structure of the input variable X(t) seems to be significant for $q^{(opt)}$. Since the evaluation of the optimal model of X(t) requires laborious computation the error curves in Fig. (IV.6) (the same discussion applies for the plots in Fig. (IV.7)) may be used directly to calculate the "best" Markovian model of X(t).

In summary, it seems that the correlation, $c(\tau)$, of the input sequence X(t) is not always significant for the evaluation of the optimal

designer strategy. Consider the situation when the designer action is not sensitive to $c(\tau)$ and let $\mathcal{M}^*$ be the operational metamodel containing independent and correlated models. Then, the optimal code strategy is likely to be an independent sequence, $H^{(I)}$, since $H^{(I)}$ is sufficiently accurate and is much less costly to use in the decision framework than any correlated strategy of Markovian type. If this is the case one may start directly with an operational metamodel containing only independent models. When independent models are not adequate, in most situations (unless only the one-step predictive distribution of H_k is required; see Sec. IV.3.3.2) only approximate solutions are of practical interest.

In all illustrations of problems (i), (ii), and (iii) defined in Chap. I, the evaluation of the optimal action of "best" designing/operating a system, S, is achieved in two steps. First, the optimal (pure) strategy, $H^{(opt)}$, of the code committee is calculated; then, $H^{(opt)}$ and the cost functions of any specific system, S, to be designed are processed and, say, the optimal design level, $d^{(opt)}$, of S (designer step) is found.

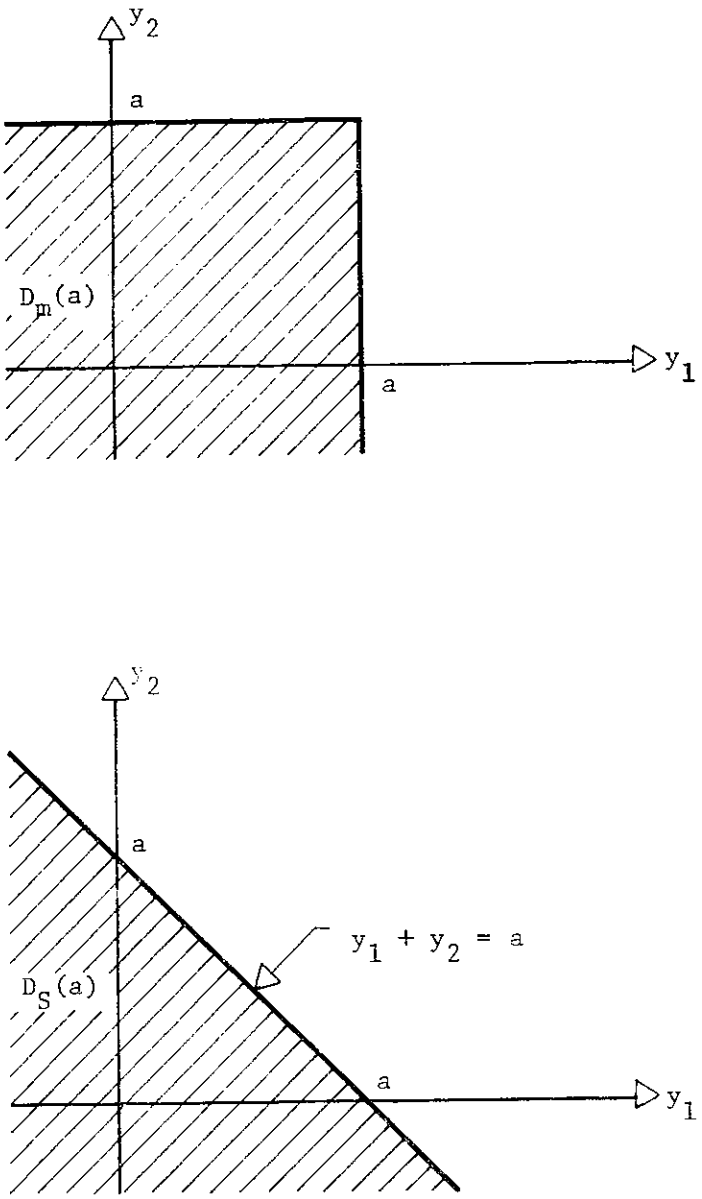


Fig. IV.1. Calculation of the distribution of the extreme, $\tilde{y}_{m,\tau}$, and of the sum, $\tilde{y}_{S,\tau}$, variables. The case when $\tau = 2$.

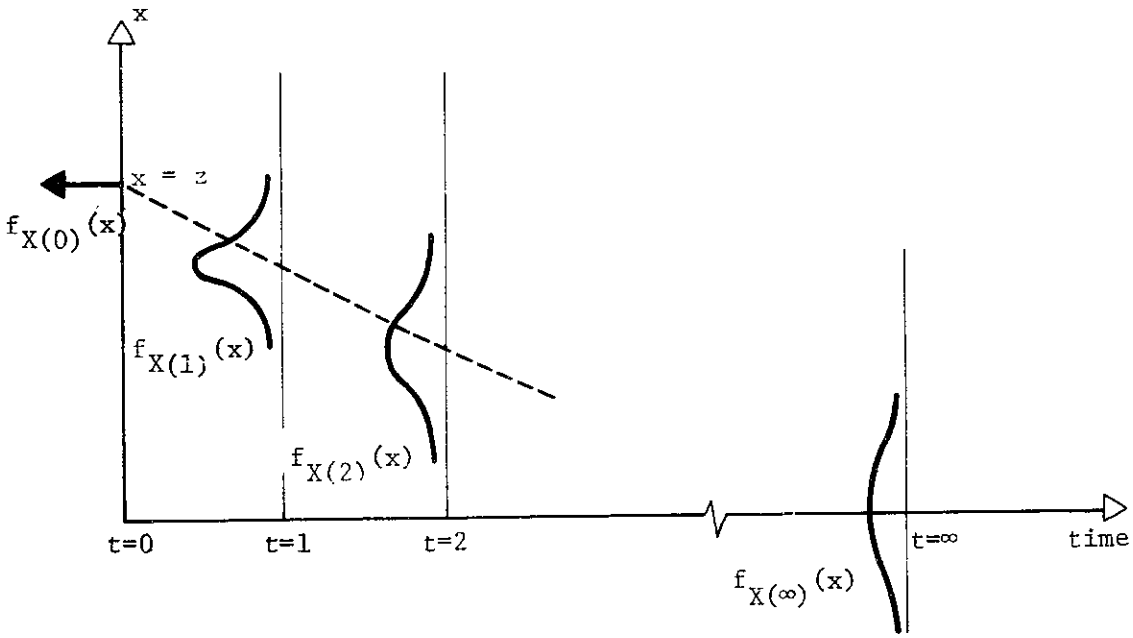


Fig. IV.2. Time evolution of the marginal distribution of the state X(t) of a GMS. The case when the asymptotic mean is zero: z is the value of the state at the present time.

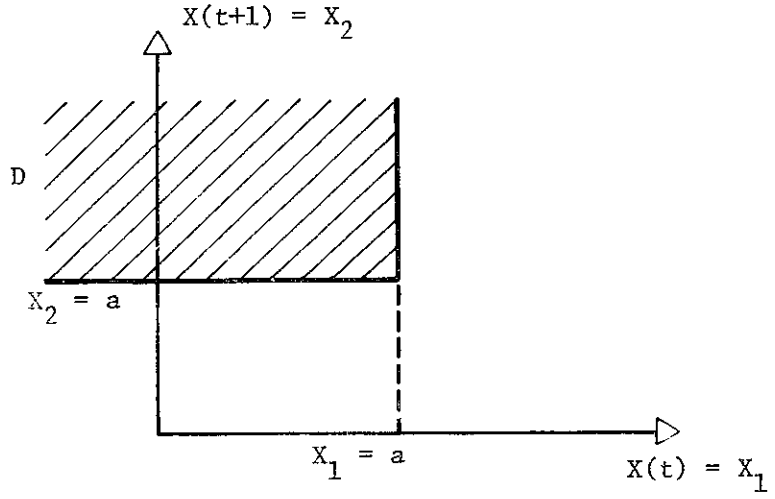


Fig. IV.3. Calculation of the mean rate of upcrossings for a GMS.

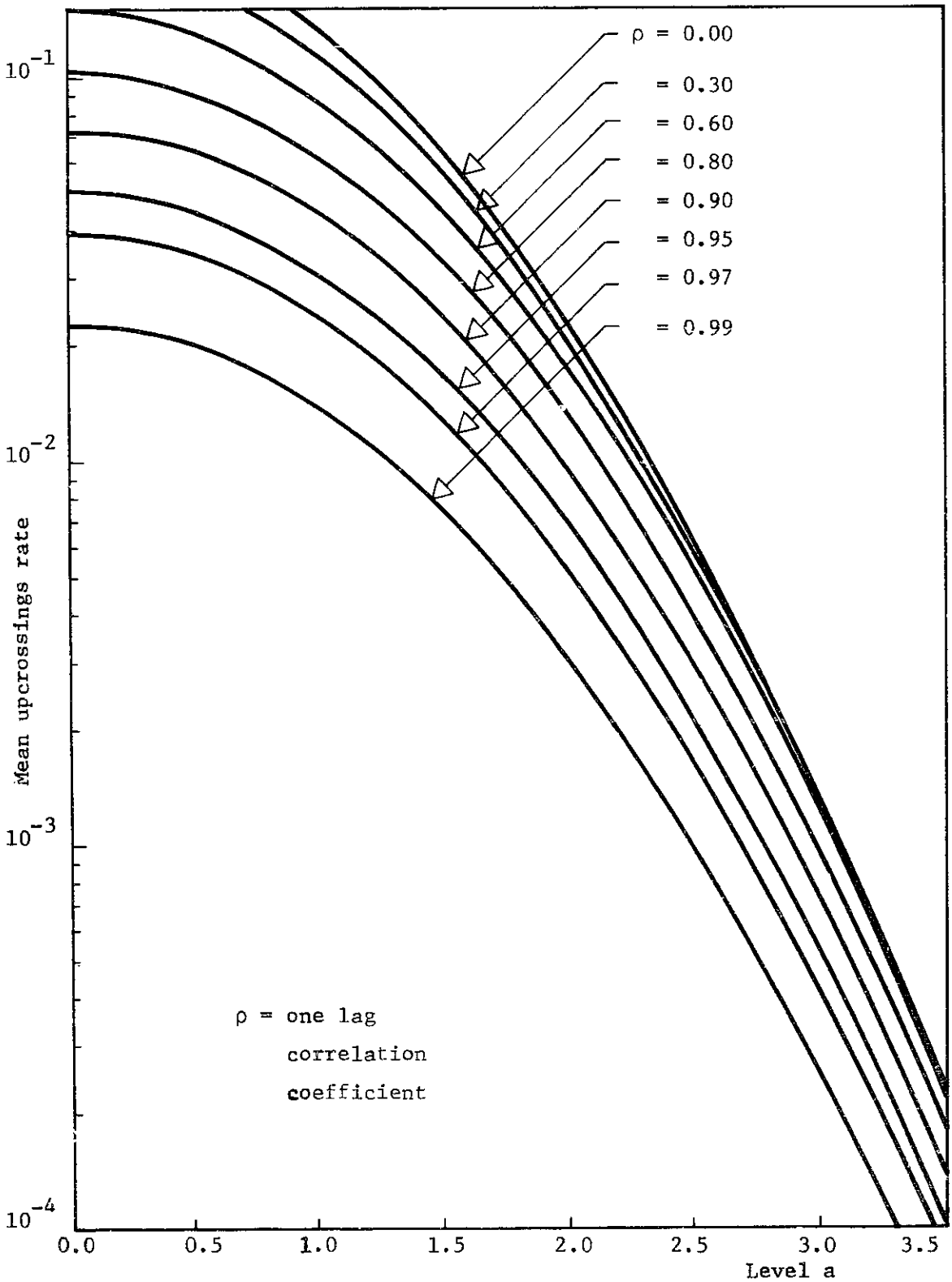


Fig. IV.4. Mean rate of upcrossings of the zero mean unit variance first order GMS.

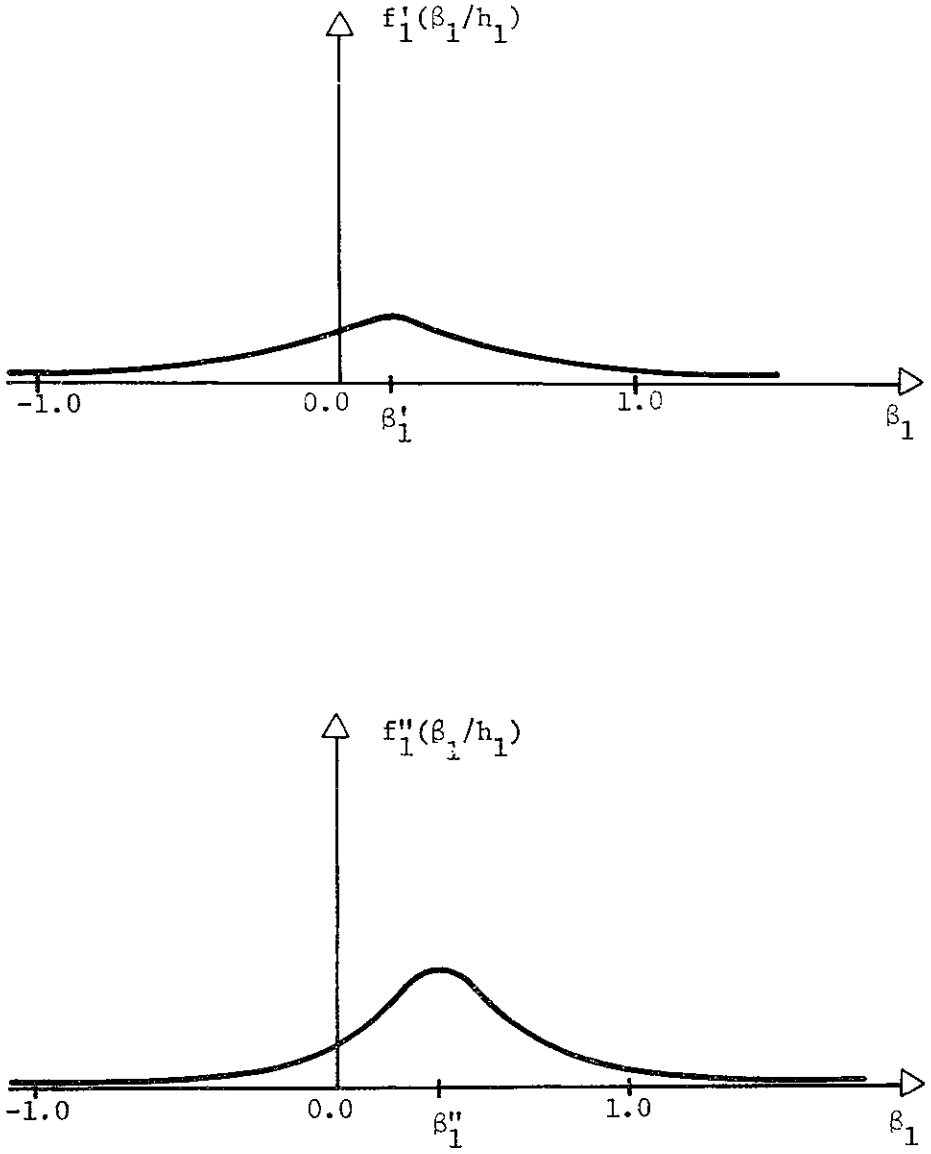


Fig. IV.5. Hypothetical conjugate prior and posterior PDF of the correlation coefficients of a GMS. Examplification for the first order GMS in Eq. (IV.81).

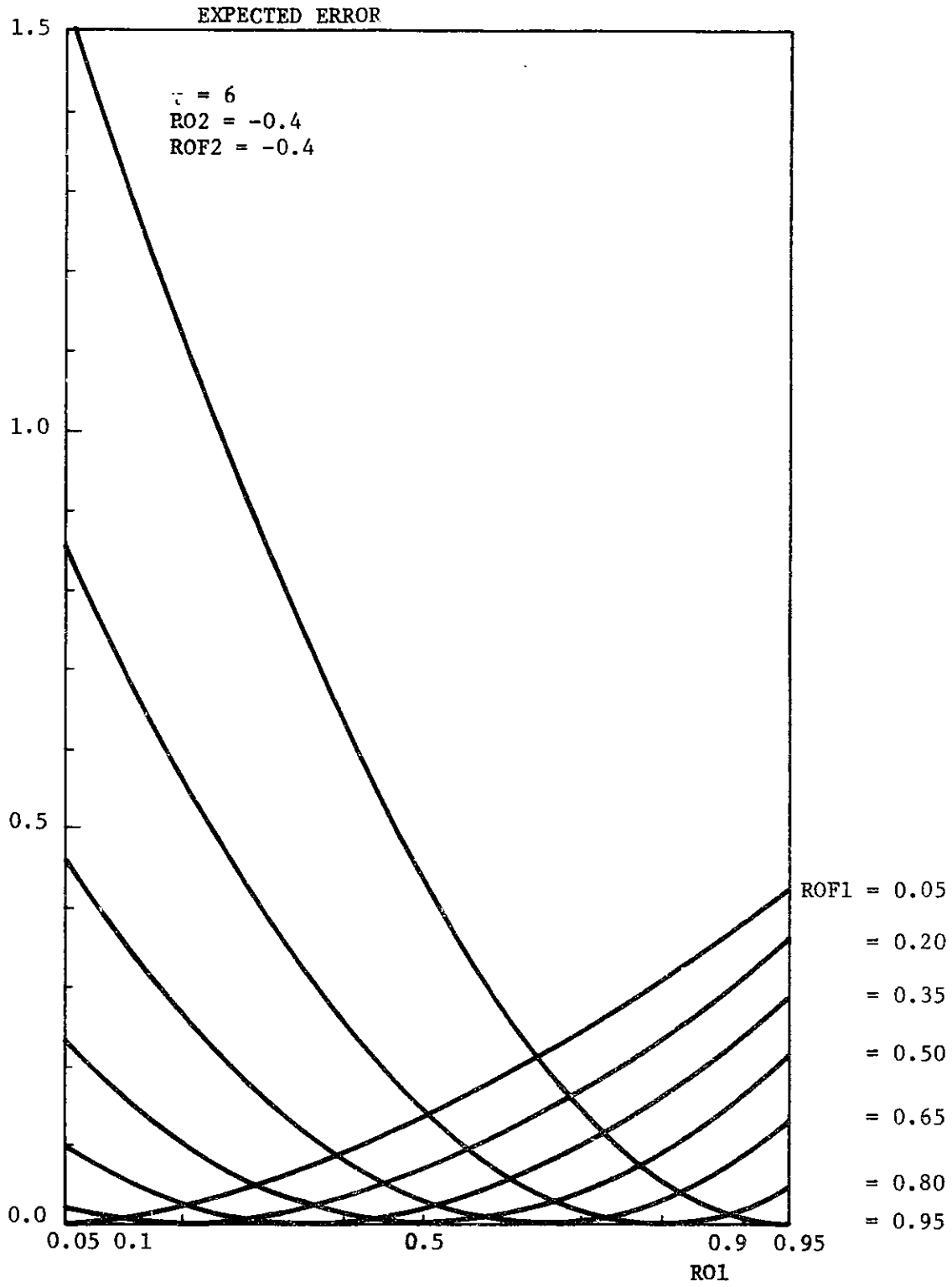


Fig. IV.6a. Normalized expected error in Eq. (IV.104) versus RO1 The case when RO2 = - 0.4, ROF2 = -0.4, τ = 6.

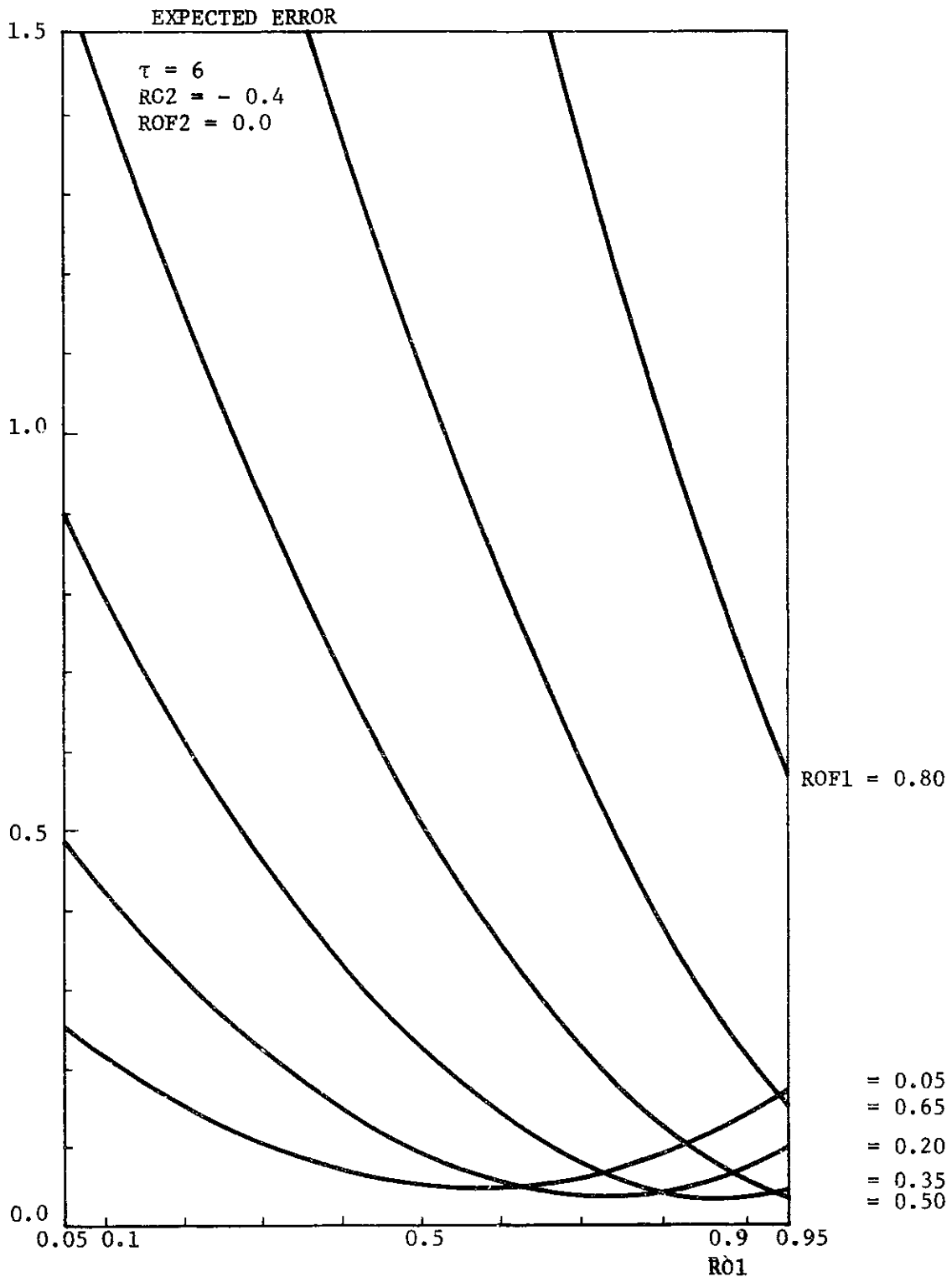


Fig. IV.6b. Normalized expected error in Eq. (IV.104) versus RO1 The case when RO2 = -0.4, ROF2 = 0.0, τ = 6.

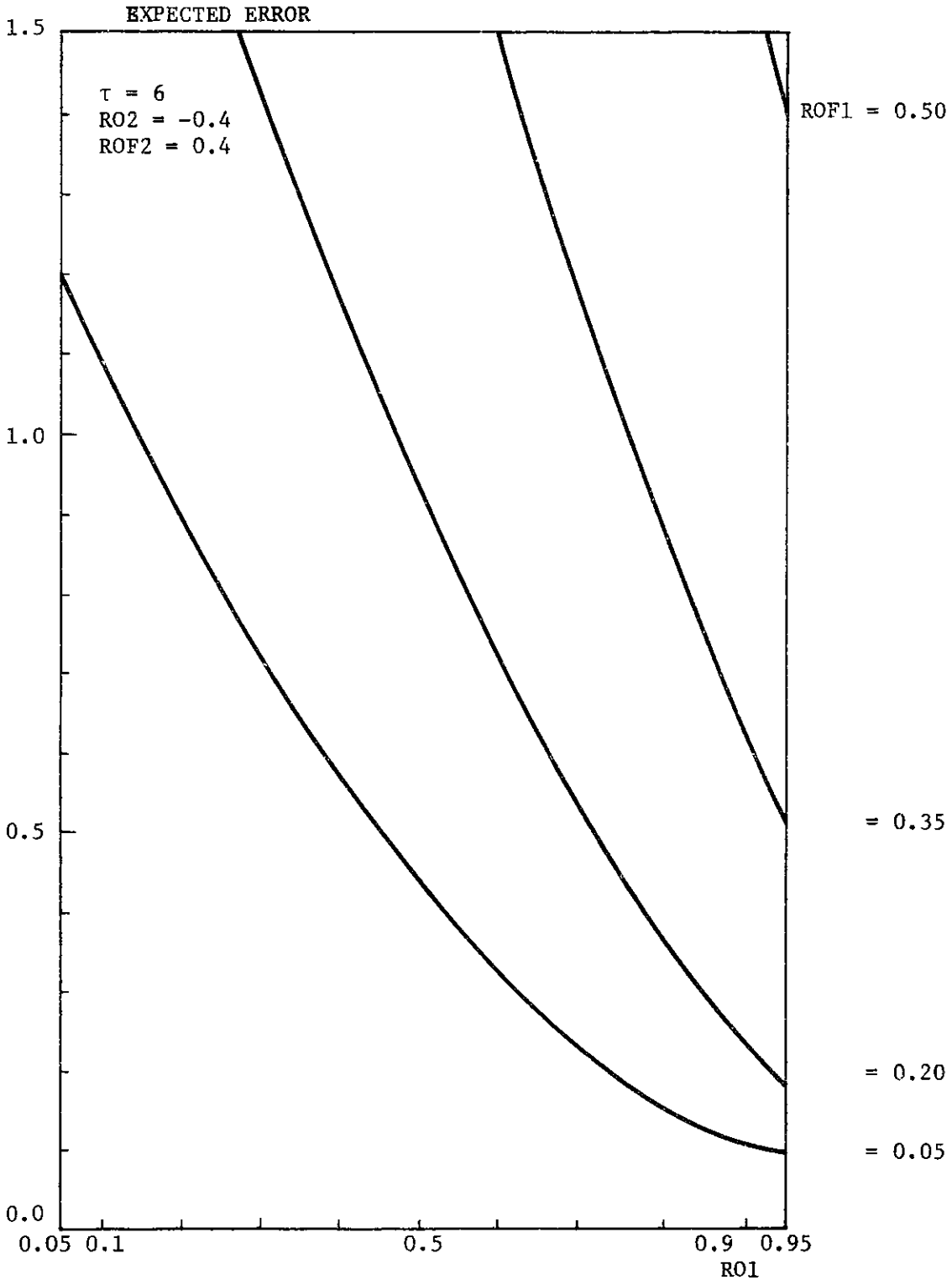


Fig. IV.6c. Normalized expected error in Eq. (IV.104) versus RO1. The case when RO2 = - 0.4, ROF2 = 0.4, τ = 6.

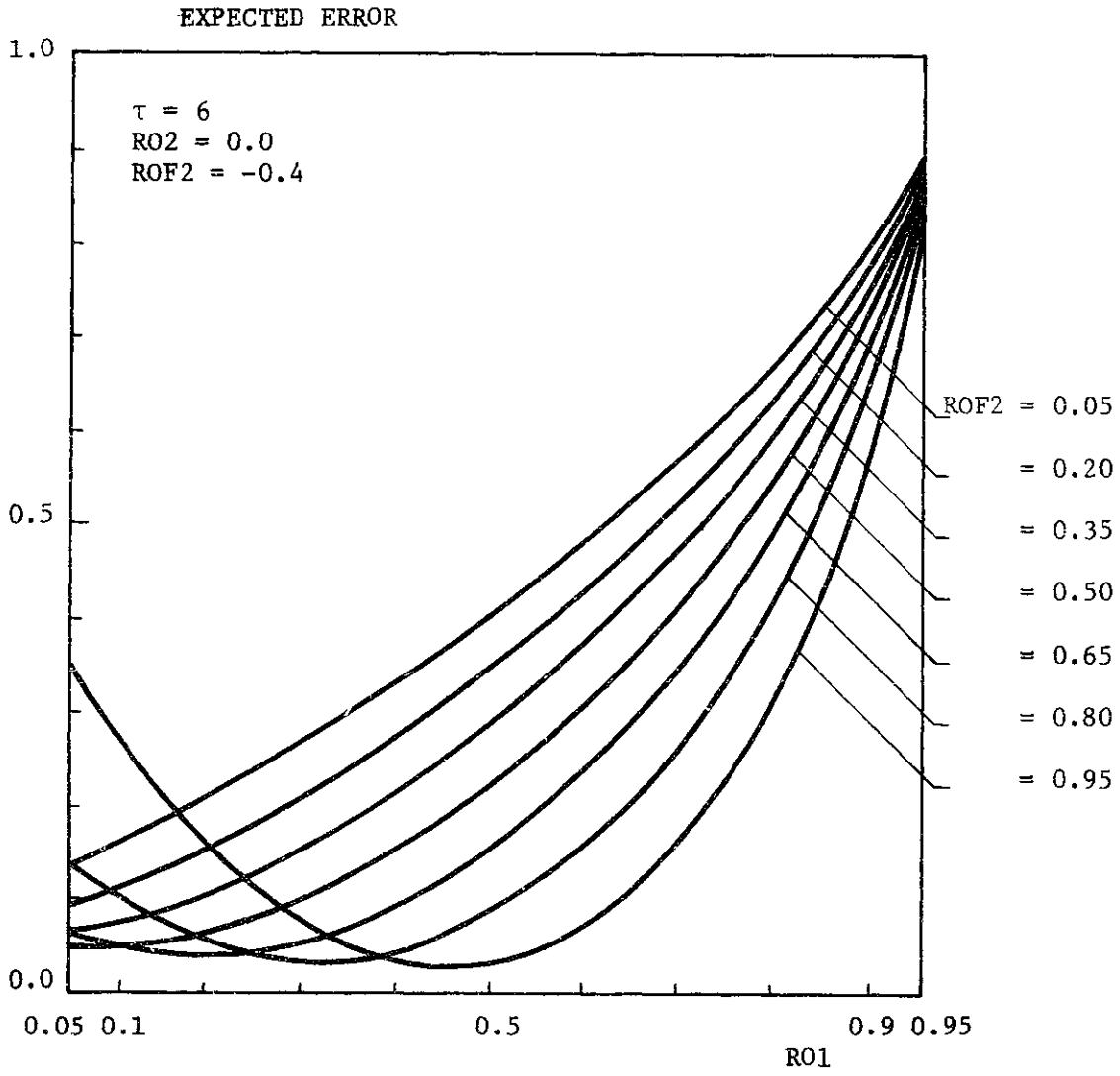


Fig. IV.6d. Normalized expected error in Eq. (IV.104) versus RO1. The case when RO2 = 0.0, ROF2 = -0.4, τ = 6.

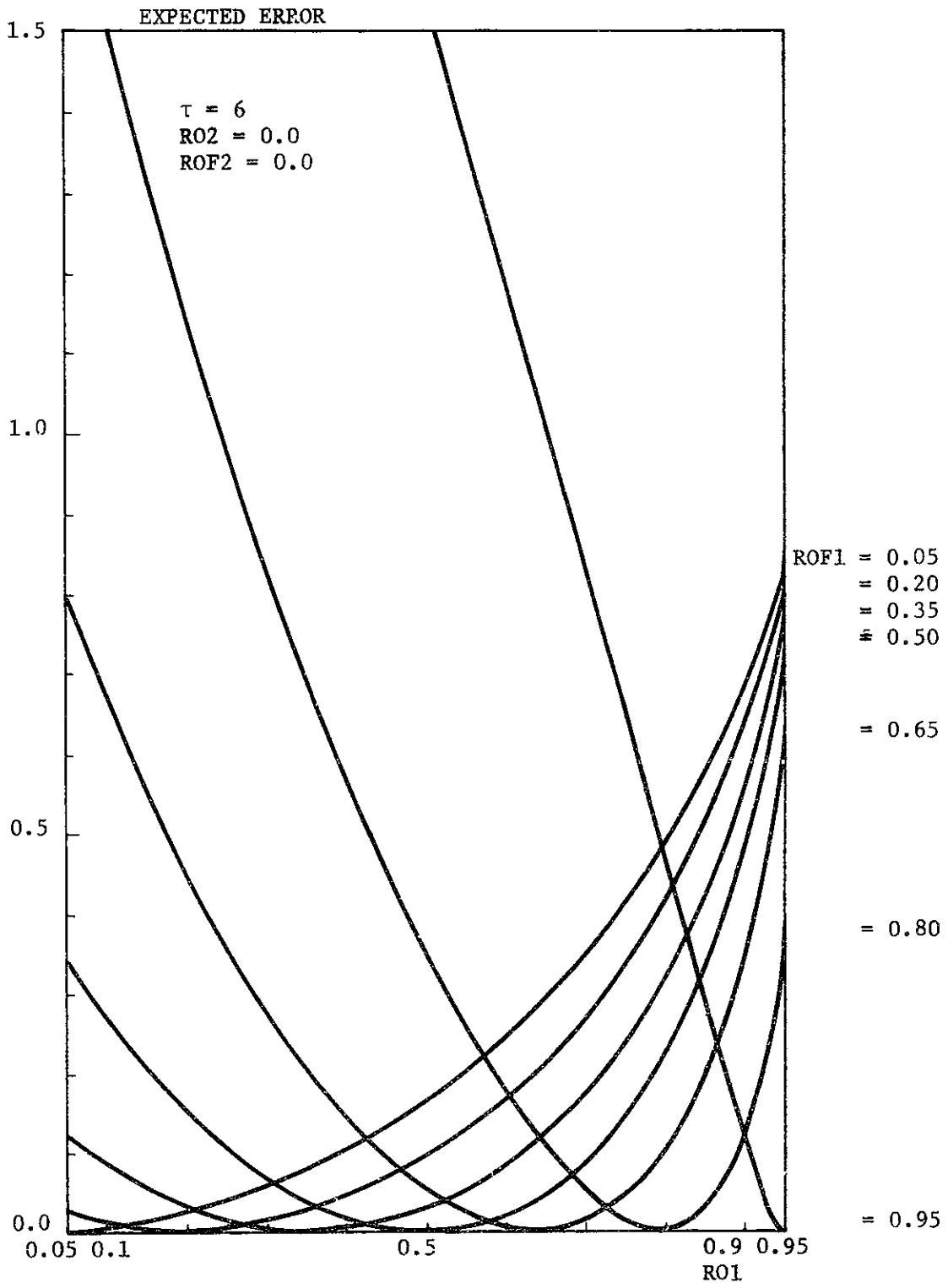


Fig. IV.6e. Normalized expected error in Eq. (IV.104) versus RO1. The case when RO2 = 0.0, ROF2 = 0.0, τ = 6.

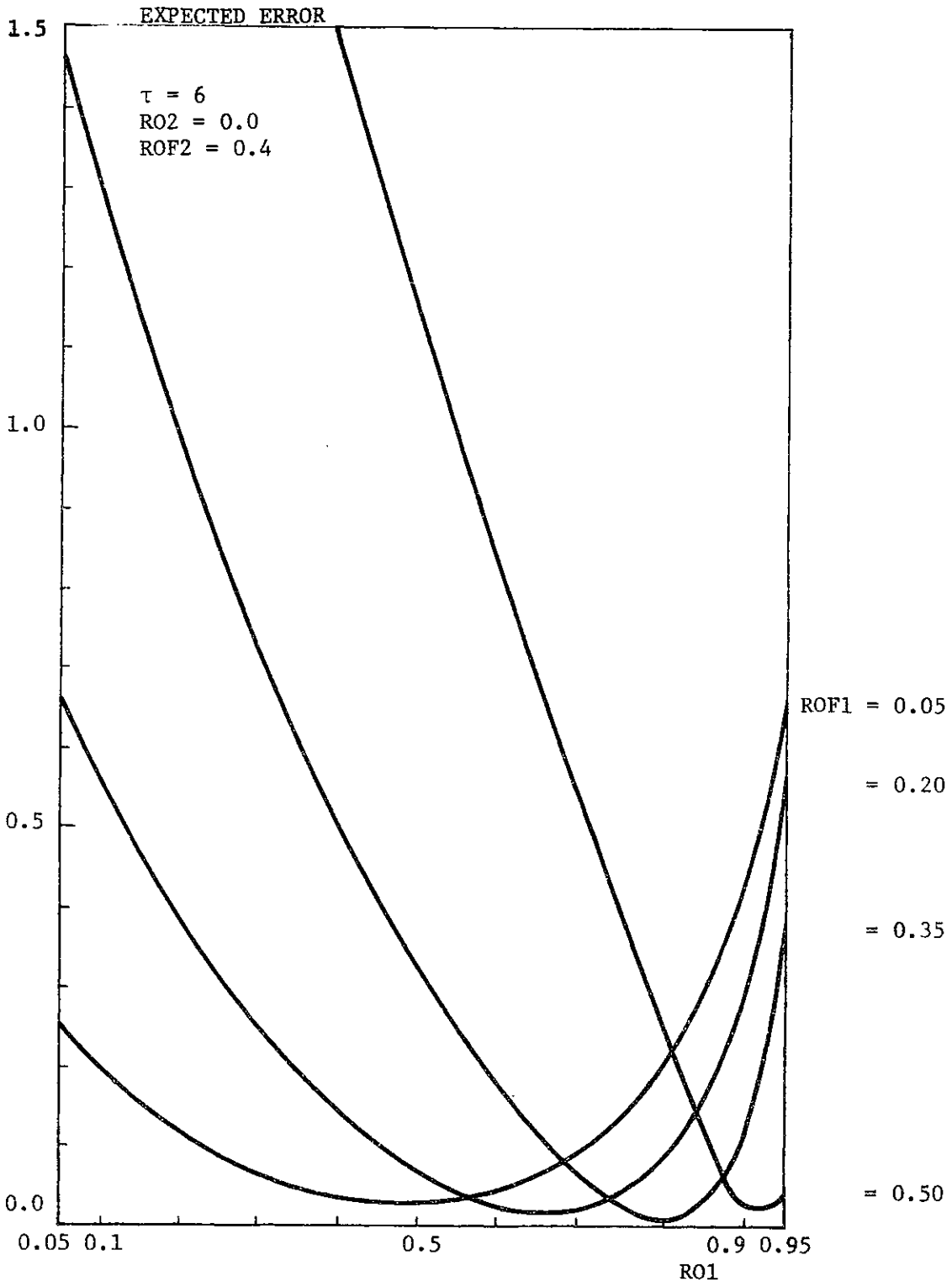


Fig. IV.6f. Normalized expected error in Eq. (IV.104) versus RO1. The case when RO2 = 0.0, ROF2 = 0.4, τ = 6.

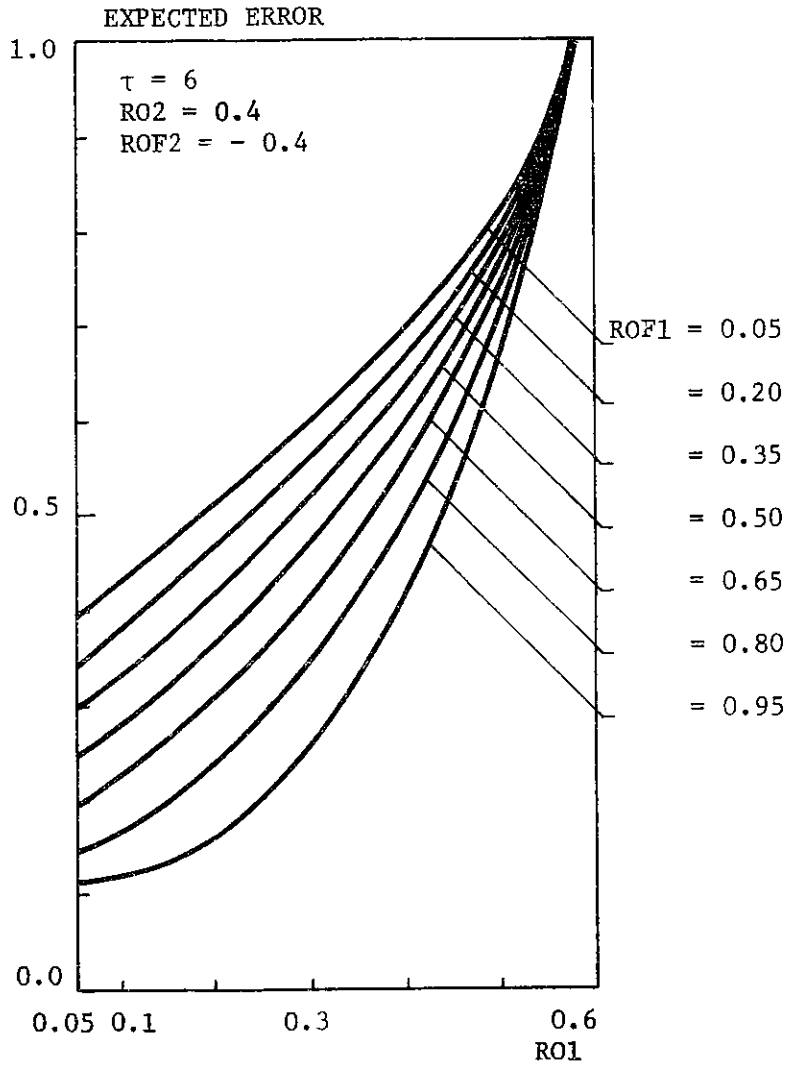


Fig. IV.6g. Normalized expected error in Eq. (IV.104) versus RO1. The case when RO2 = 0.4, ROF2 = -0.4, τ = 6.

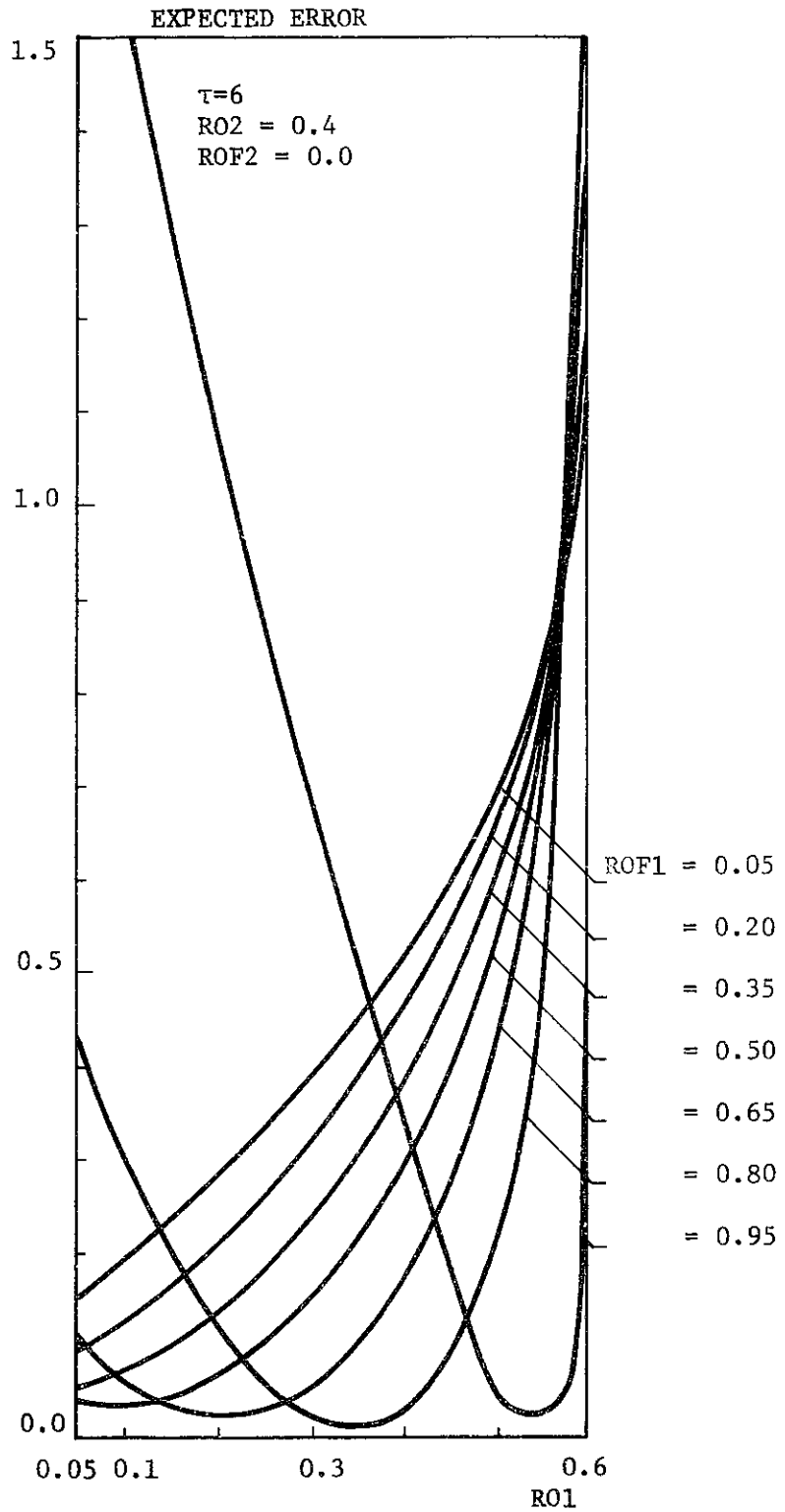


Fig. IV.6h. Normalized expected error in Eq. (IV.104) versus RO1. The case when RO2 = 0.4, ROF2 = 0.0, τ = 6.

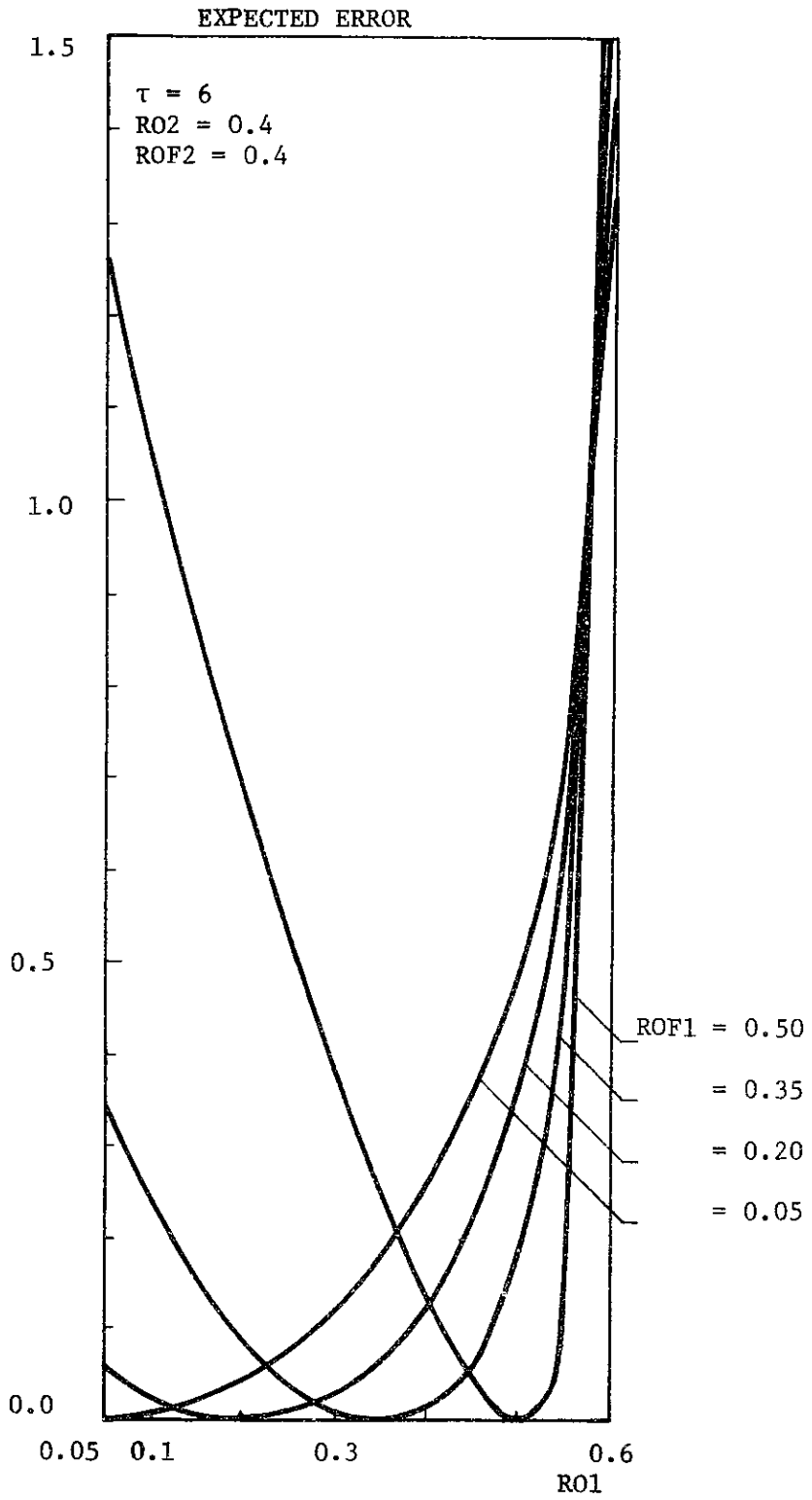


Fig. IV.6i. Normalized expected error in Eq. (IV.104) versus RO1. The case when RO2 = 0.4, ROF2 = 0.4, τ = 6.

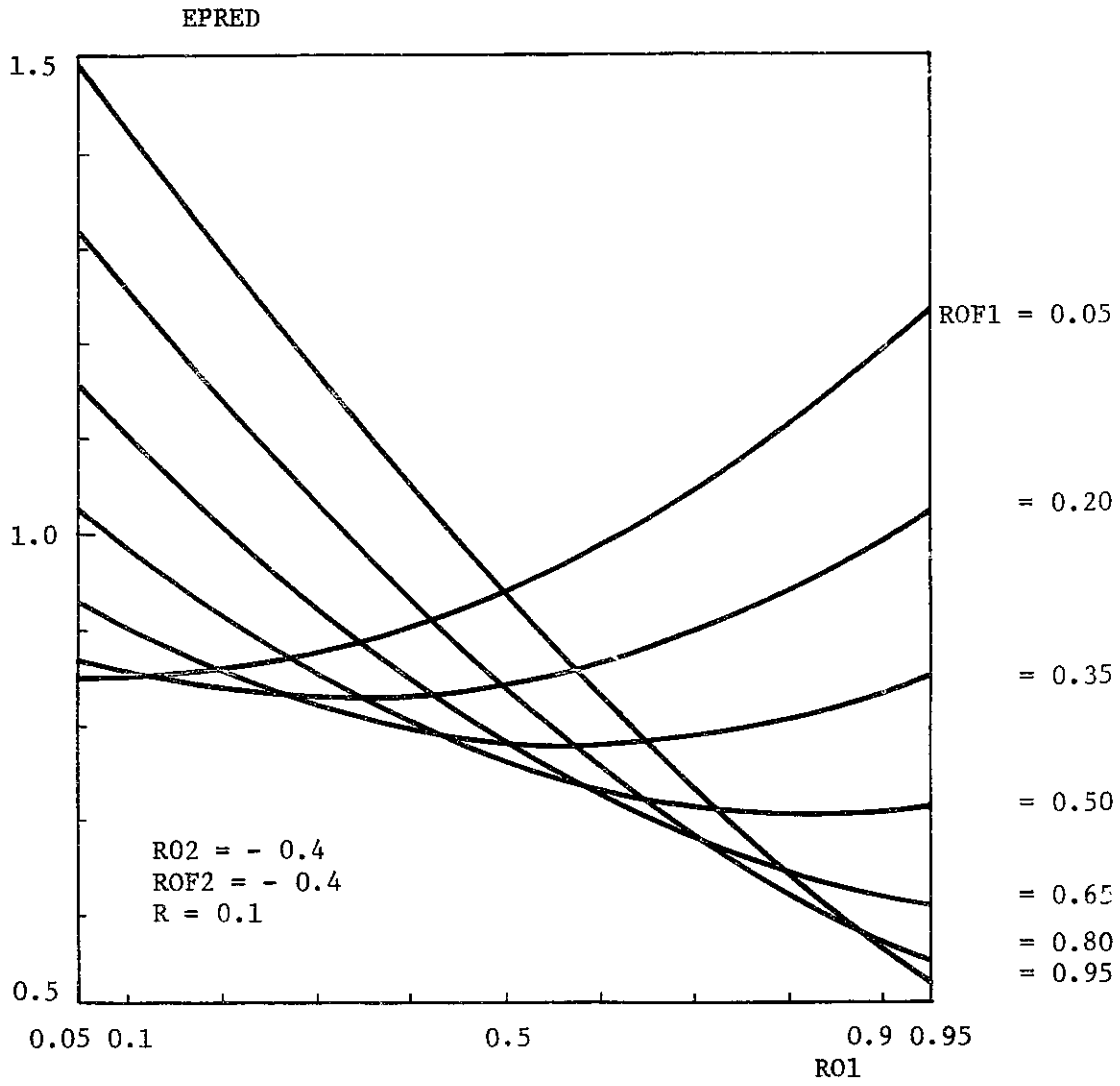


Fig. IV.7a. One step prediction error in Eq. (IV.109) versus RO1. The case when RO2 = - 0.4, ROF2 = - 0.4, R = 0.1

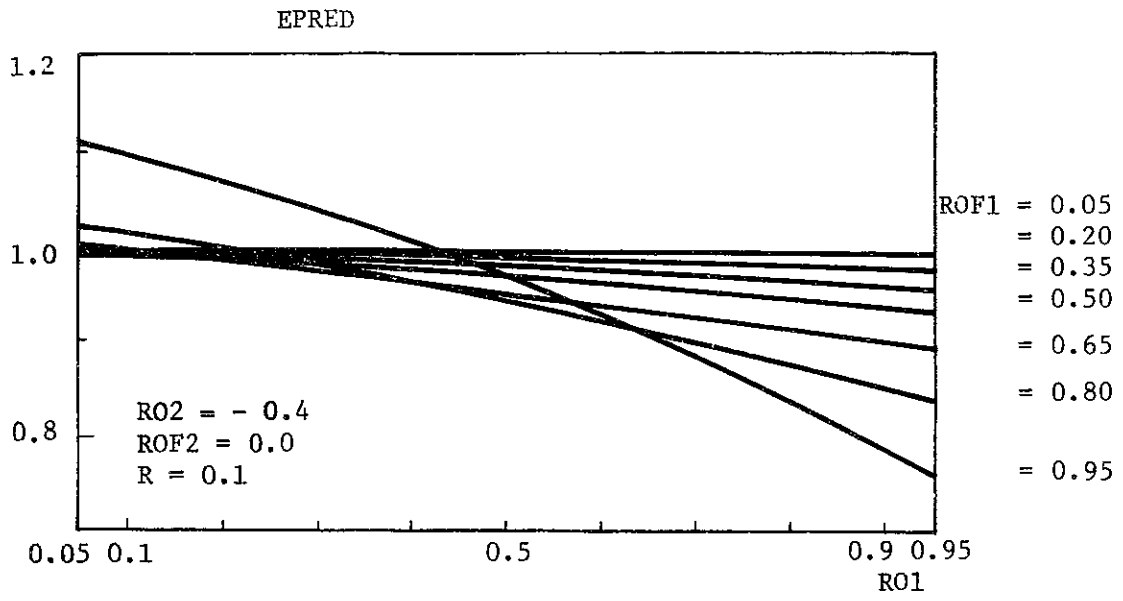


Fig. IV.7b. One step prediction error in Eq. (IV.109) versus RO1. The case when RO2 =-0.4, ROF2 = 0.0, R = 0.1.

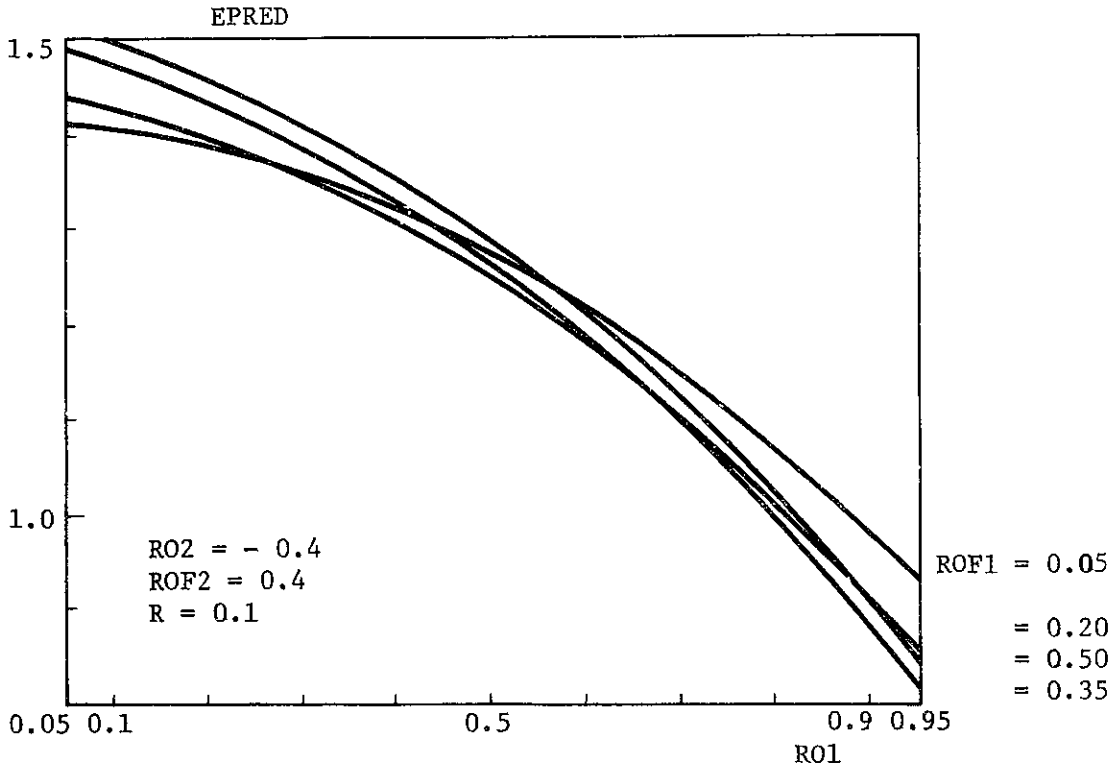


Fig. IV.7c. One step prediction error in Eq. (IV.109) versus RO1. The case when RO2 = - 0.4, ROF2 = 0.4, R = 0.1

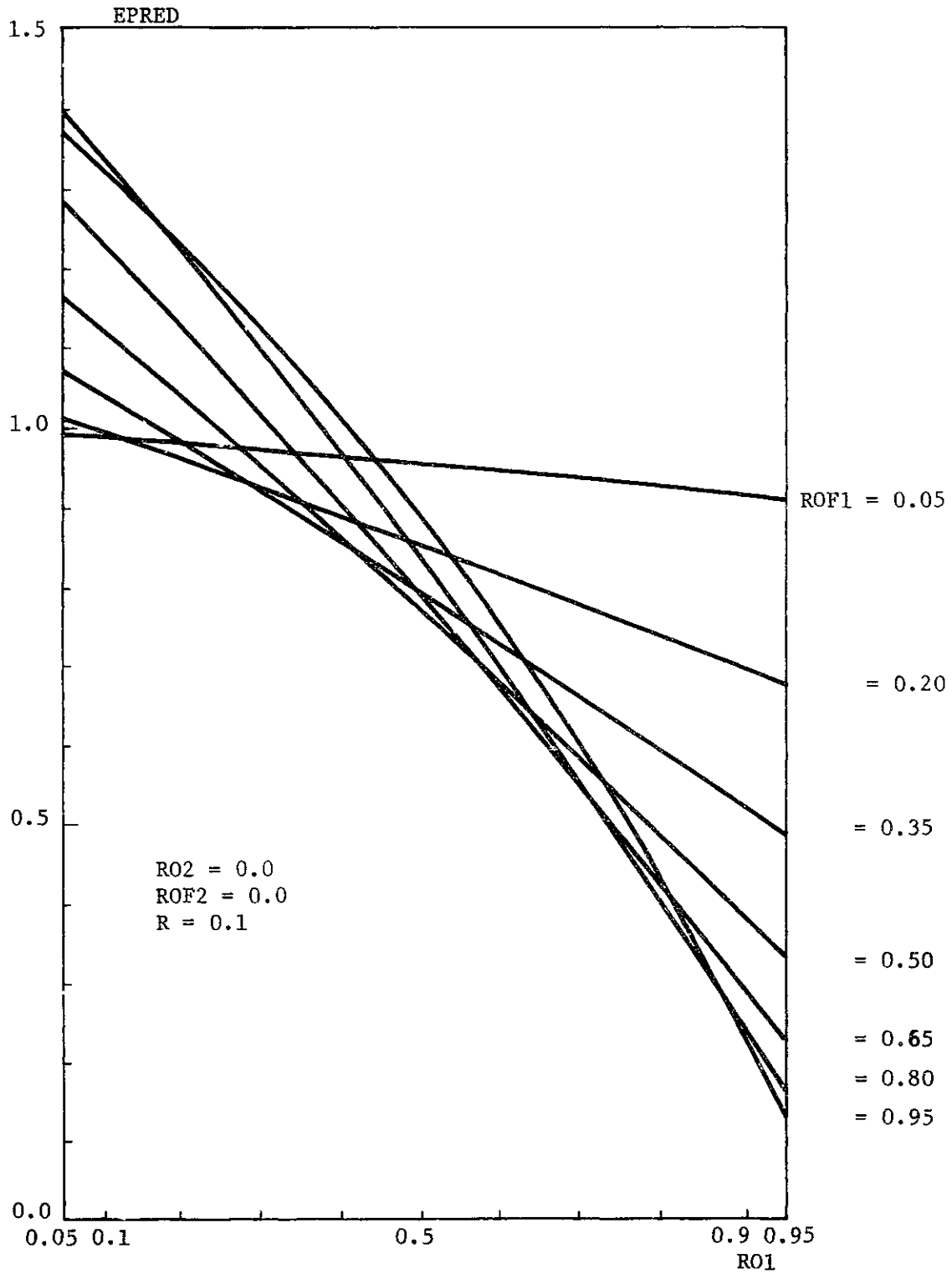


Fig. IV.7d. One step prediction error in Eq. (IV.109) versus RO1. The case when RO2 = 0.0, ROF2 = 0.0, R = 0.1.

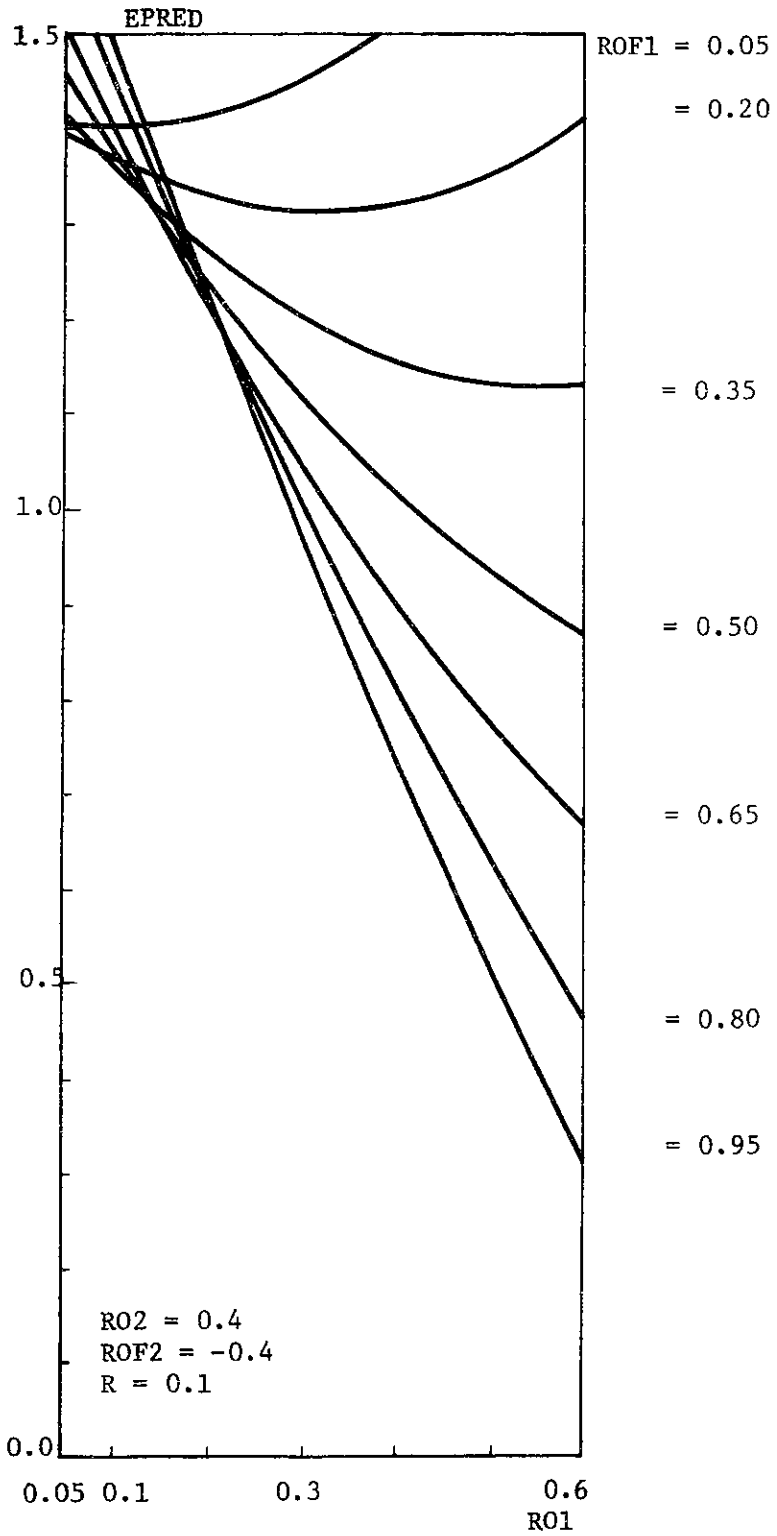


Fig. IV.7e. One step prediction error in Eq. (IV.109) versus RO1. The case when RO2 = 0.4, ROF2 = -0.4, R = 0.1.

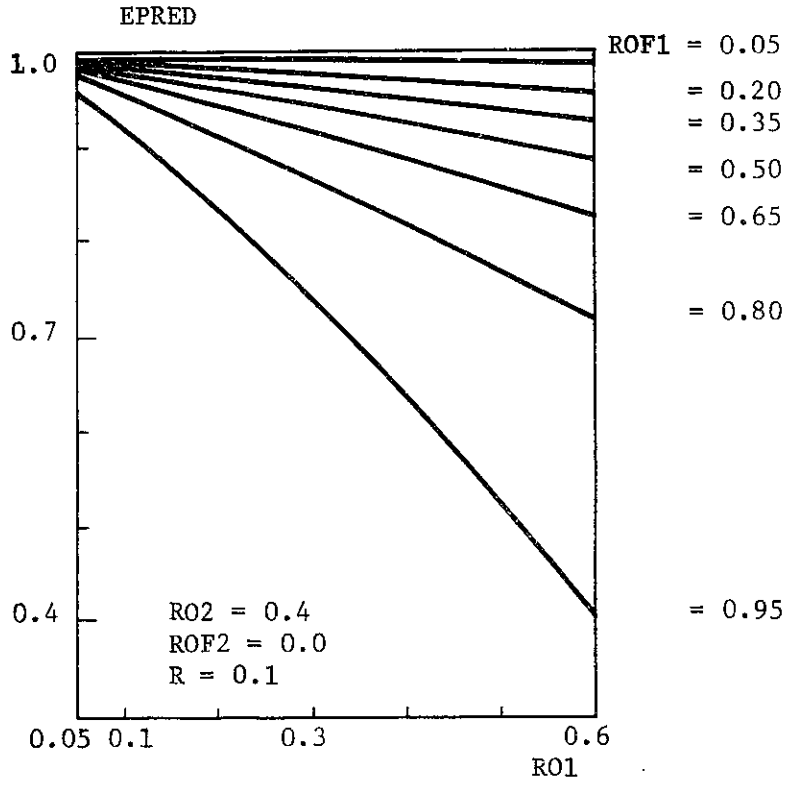


Fig. IV.7f. One step prediction error in Eq. (IV.109) versus RO1. The case when RO2 = 0.4, ROF2 = 0.0, R = 0.1.

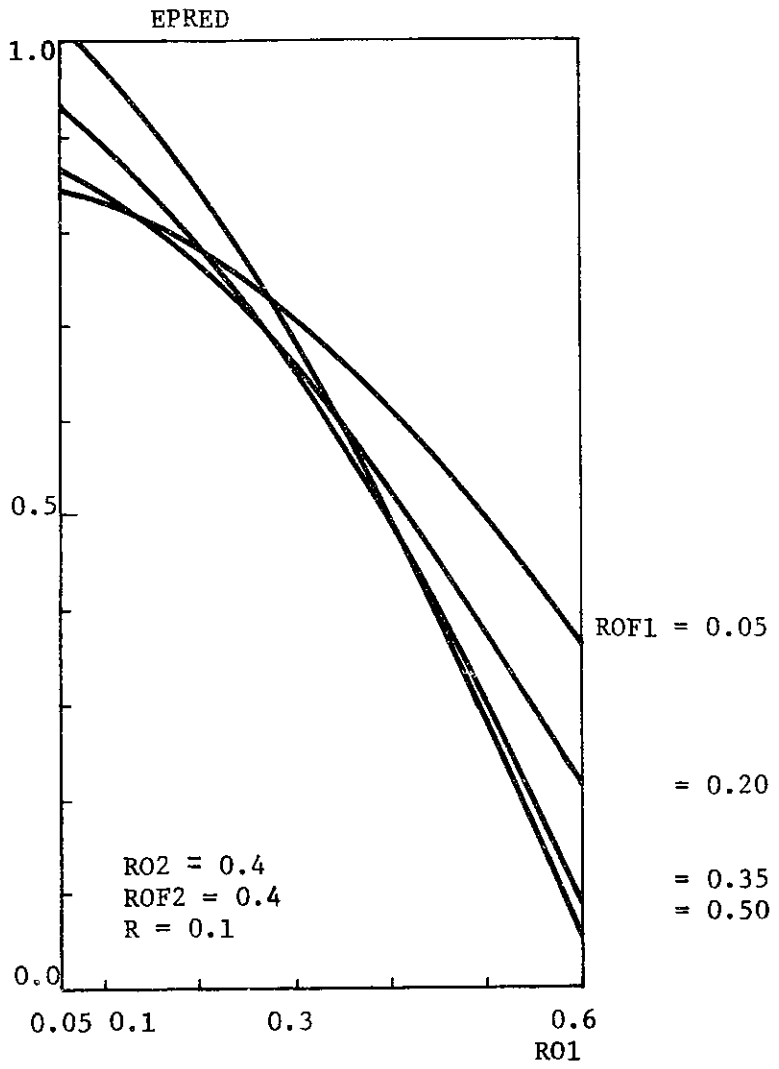


Fig. IV.7g. One step prediction error in Eq. (IV.109) versus RO1. The case when RO2 = 0.4, ROF2 = 0.4, R = 0.1.

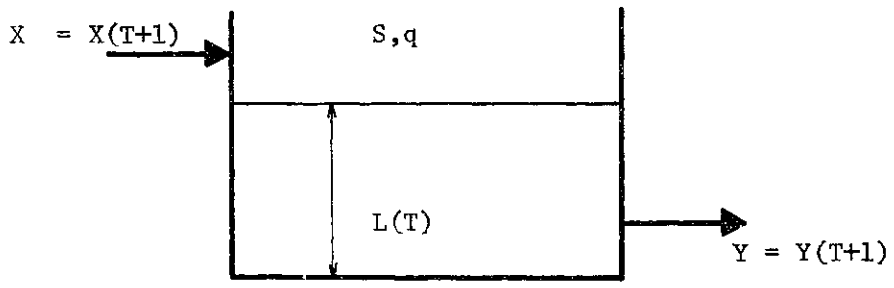


Fig. IV.8. Optimal operation of a system, S, excited by a partially known input, X(t). Examplification for the case where:
S = a reservoir of capacity q.
L(T) = the amount of water in S at present time t=T
X(T+1) = the future (unknown) quantity of water which will enter S at t=T+1
Y(T+1) = the water to be released at t = T+1.

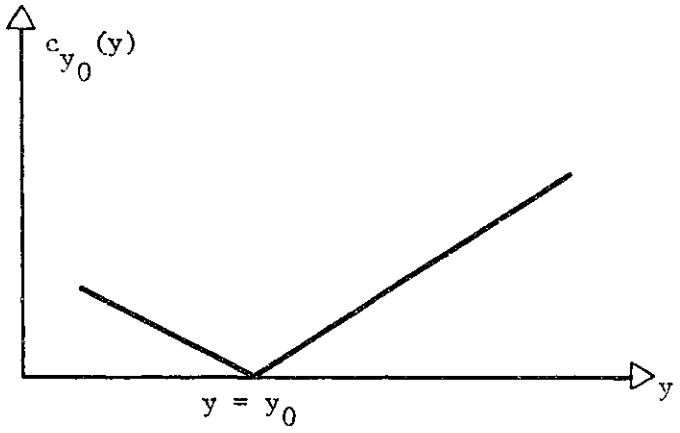


Fig. IV.9. Hypothetical penalty associated with the water release y.

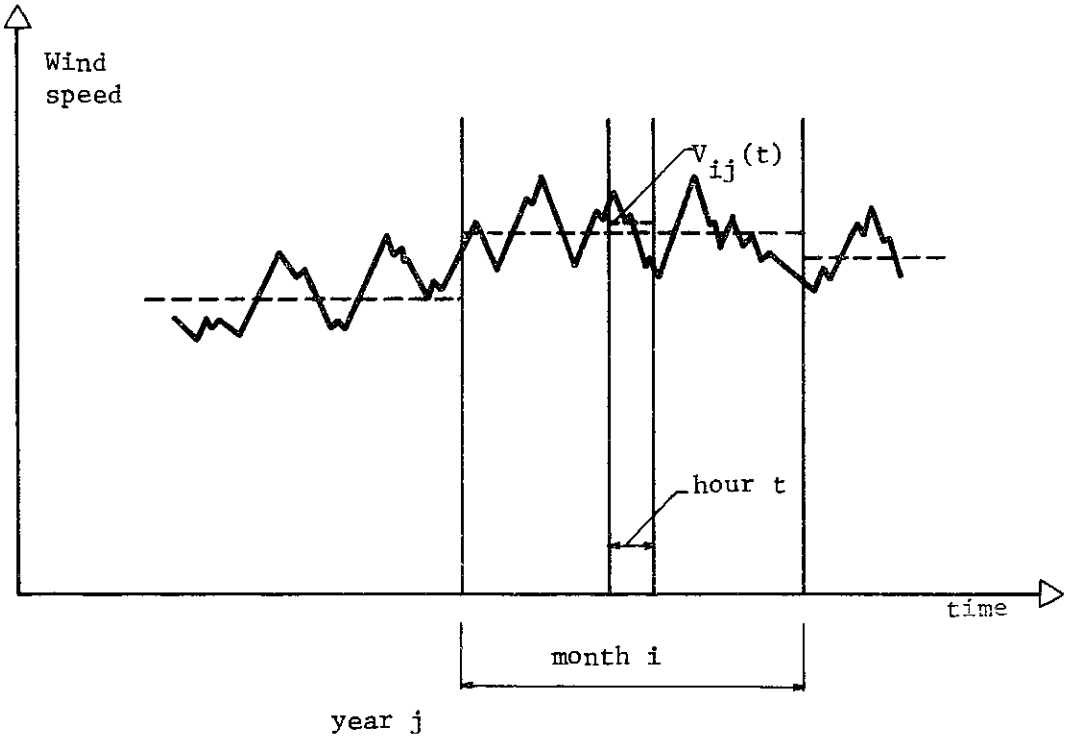


Fig. IV.10. Hypothetical realization of the wind speed process $V_{ij}(t)$ = the hourly average calculated in the t^{th} day of month i, year j.

CHAPTER V

RENEWAL SEQUENCES WITH RANDOM MAGNITUDES

Consider the design of the structural system S in Fig. (I.1) loaded by the renewal sequence $X_q(t)$, q = 0,1, to be defined below. Hypothetical realizations of $X_q(t)$ are shown in Fig. (V.1) for q = 0 and in Fig. (V.2) for q = 1.

Thus, $X_0(t)$ has renewal occurances of interarrival time τ_o and independent, identically distributed event sizes (random variable X_0).

$X_1(t)$ is defined by the same elements as in $X_0(t)$, namely the white stationary sequences of interarrival times $\{\tau_1\}$ and of event sizes $\{X_1\}$. The discontinuities in the realizations follow a renewal process (see also [7,20,34,58,62,63,72] for details about renewal processes) of interarrival time τ_1. Between discontinuities the realization of $X_1(t)$ is constant in time and in the interval $[0,t_{1,1})$, $X_1(t) = X_{1,0}$ (see Fig. (V.2)). If the variable X_1 may take the value x = 0 with finite non zero probability p (see Fig. (V.3a)), then a typical sample function of $X_1(t)$ has the appearance shown in Fig. (V.3b). It may be that when observing $X_1(t)$ with $P(X_1=0) = p \neq 0$ only a portion of the interarrival times can actually be fully observed. Thus, by simply observing $X_1(t)$ the individual

values $\{\gamma_{1,2},\gamma_{1,3}\}$ in Fig. (V.35) cannot be measured but only their sum $\gamma_{1,2}+\gamma_{1,3}$. In other practical situations one may observe $X_1(t)$ and also the full sequence of interarrival times $\{\gamma_1\}$, case in which $X_1(t)$ is said to be observable.

In this chapter it is assumed that the interarrival times $\{\gamma_q\}$ and the magnitudes $\{X_q\}$ of $X_q(t)$ are two mutually independent sets of random variables, q=0,1.

One should recognize that the process $X_1(t)$ becomes a "Borges" process (which is extensively treated in [9]) provided that the variable γ_1 has zero variance and that p = 0.

Assume now that the design of the two-state (failure/survival) system S in Fig. (I.1) of the life time τ is considered.

Consider first the case when S is loaded by $X_q(t)$. Then, the distribution of the extreme random variable

$$y_{q,\tau} = \max_{0\le t\le\tau} \{X_q(t)\} \tag{V.1}$$

has to be calculated for safety checking. The CDF of $y_{q,\tau}$ is found in [57,65] for exponential interarrival time γ_q and in [6] for the case when γ_q is a gamma variable.

Second, consider the situation when S is loaded by the linear combination or the sum of two independent processes:

$$X(t) = X_0(t) + X_1(t) \tag{V.2}$$

Consequently, to check the reliability of S, the CDF of the random variable

$$y_\tau = \max_{0\le t\le\tau} \{X_0(t) + X_1(t)\} = \max_{0\le t\le\tau} \{X(t)\} \qquad (V.3)$$

is needed. Exact and approximate solutions of the distribution of y_τ are given in [10,42] and [57,65], respectively, for the case when the renewal time process of $X_q(t)(q=0,1)$ is Poisson.

All the results presented above rest on the assumption that the probabilistic law of the input variable $X_q(t)$ is perfectly known. The goal of this chapter is to extend some of the probabilistic solutions to the case in which the law of $X_q(t)$ is only partially known; to achieve this aim the results obtained in Chap. III when independent sequences have been considered are used. Thus, following the methodology developed in Sec. II.2, the evaluation of the optimal code strategy, f(x), corresponding to systems of type S and of the optimal design of S are briefly considered for input variables $X_q(t)$ or X(t).

The input of the decision framework calculating f(x) is:

1. The Bayesian description of the metamodel, which consists of a finite set of pairs of models ("model" stands here for distributional form) for the couple $(\mathcal{T}_q, X_q)$. Also, prior PDF's on the unknown parameters contained in the postulated distributions of $(\mathcal{T}_q, X_q)$ are to be given.

2. The Experiment E_q. Two observation vectors $\underline{\mu}^*_{(q)}$ and $\underline{Z}_{(q)}$ corresponding, respectively, to the interarrival time $\{\mathcal{T}_q\}$ and to the magnitude $\{X_q\}$ sequences define the informative experiment E_q of $X_q(t)$. $n^*_{(q)}$ stands for the dimension of $\underline{\mu}^*_{(q)}$. Fig. (V.4) and (V.5) illustrate observationvation vectors corresponding to the various kinds of $X_q(t)$ processes analyzed in this chapter.

To avoid numerical difficulties the observation vector $\underline{\mu}^*_{(q)}$ is

generally replaced by $\underline{\mu}_{(q)}$ of size $n_{(q)}$ (see Fig. (V.4) and Fig. (V.5)) when the time structure of $X_q(t)$ follows a general renewal process or when $P(X_1=0) = p \neq 0$ and $X_1(t)$ is not observable. $\underline{\mu}_{(q)}$ gathers only the observed interarrival times. Consequently, the information such as the time $t_{q,e}$ from the moment when observation starts until the first renewal is observed and the time $t_{q,a}$, from the last observed renewal until the end of observation, is lost. If $n_{(q)}$ is large, the error related to the replacement of $\underline{\mu}^*_{(q)}$ by $\underline{\mu}_{(q)}$ is expected to be negligible. However, this is not generally true when dealing with rare events, i.e. when the number of observed renewals is small.

The computational difficulties mentioned above when the time process of $X_q(t)$ follows a general renewal sequence come from the special form of the likelihood function, $\ell_{\gamma_q}(\underline{\mu}^*_{(q)})$, of γ_q.

$$\ell_{\gamma_q}(\underline{\mu}^*_{(q)}) = f^{(q)}_{EXCESS}(t_{q,e};T_i)\{\prod_{i=1}^{n_{(q)}} f_{\gamma_q}(\mu_{q,i})\ [1-F^{(q)}_{AGE}(t_{q,a};T_f)] \qquad (V.4)$$

where, $f_{\gamma_q}(x)$ is the density of the interarrival time, γ_q, variable, $f^{(q)}_{EXCESS}(x;t)$ is the PDF of the so called excess variable, $\xi^{(q)}_t$, at time t, and
$F^{(q)}_{AGE}(x;t)$ is the CDF of the so called age variable, $A^{(q)}_t$, at the moment t, [52,62,63].
For example, $t_{1,e}$ and $t_{1,a}$ in Fig. (V.4b) are outcomes of the variables $\xi^{(1)}_{T_i}$ and $A^{(1)}_{T_f}$, respectively.

The age variable has the same distribution as the interarrival time γ_q. The CDF, $F_{EXCESS}(x;t)$, of $\xi^{(q)}_t$ satisfies ([63], p.184, [52], p. 193)

the following equation:

$$F^{(q)}_{EXCESS}(x;t) = 1 - F_{\gamma_q}(t+x) + \int_o^t F^{(q)}_{EXCESS}(x;t-s) f_{\gamma_q}(s)ds \qquad (V.5)$$

where, $F_{\gamma_q}(x)$ is the CDF of γ_q.

In the specific case of the exponential interarrival time γ_q of decay parameter $\lambda^{(q)}$, $F^{(q)}_{EXCESS}(x;t)$ can be obtained either as a solution of (IV.5) (see [63], p. 173) or by observing directly ([52], p. 173) that $\xi^{(q)}_t$ exceeds x only if there is no event during the time (t, t+x). Since the Poisson process has stationary independent increments, it is:

$$P(\xi^{(q)}_t > x) = P(N(t+x) - N(t) = 0) =$$

$$= P(N(x) = 0) = e^{-\lambda^{(q)}x} \qquad (V.6)$$

where, N(t) is the Poisson counting process of mean rate $\lambda^{(q)}$.

From Eq. (V.6), and the distribution of the age variable, the likelihood in Eq. (V.4) becomes for γ_q exponential:

$$\ell_{\gamma_q}(\underline{\mu}^*_{(q)}) = [\lambda^{(q)}]^{n_{(q)}+1} e^{-\lambda^{(q)}T} \qquad (V.7)$$

Consequently, when the time process of $X_q(t)$ is Poisson, the number, $n_{(q)}+1$, of observed renewals and the total observation time T are the sufficient statistics of γ_q.

Consider now the case when γ_q is not exponentially distributed. Then, ([63], p. 184, [52], p. 193) when $T_i \to \infty$, the asymptotic excess

variable, $\xi_{\infty}^{(q)} = \lim_{T_i \to \infty} \{\xi_{T_i}^{(q)}\}$, has the density:

$$f_{EXCESS}^{(q)}(x;\infty) = \frac{1}{E[\gamma_q]} [1 - F_{\gamma_q}(x)]$$

From Eq. (V.4) and the asymptotic PDF of $\xi_{\infty}^{(q)}$, it results:

$$\ell_{\gamma_q}(\underline{\mu}_{(q)}^{*}) = \frac{1}{E[\gamma_q]} [1-F_{\gamma_q}(t_{q,e})]\{\prod_{i=1}^{n_{(q)}} f_{\gamma_q}(\mu_{q,i})\} [1-F_{\gamma_q}(t_{q,a})] \qquad (V.8)$$

In general $\ell_{\gamma_q}(\underline{\mu}_{(q)}^{*})$ in Eq. (V.8) is not known in an explicit form. However in the sepecial case of gamma interarrival times, say $\gamma_q \in G(k^{(q)}, \lambda_k^{(q)})$, the asymptotic likelihood in Eq. (V.8) becomes:

$$\ell_{\gamma_q}(\underline{\mu}_{(q)}^{*}) = \frac{1}{k^{(q)}} e^{-\lambda_k^{(q)T}} \sum_{i,j=0}^{k^{(q)}-1} \frac{(t_{q,e})^i}{i!} \frac{(t_{q,a})^j}{j!} (\lambda_k^{(q)})^{n_{(q)}k^{(q)}+i+j+1} \qquad (V.9)$$

i.e. is a finite sum of terms having the form $\lambda^{m_{ij}} e^{-\lambda_k^{(q)}T}$, where $m_{ij} = n_{(q)}k^{(q)}+i+j+1$. Consequently the term (i,j) in $\ell_{\gamma_q}(\underline{\mu}_{(q)}^{*})$ in Eq. (V.9) has the same functional form as the likelihood of a gamma variable corresponding to an observation vector of the type $\underline{\mu}_{(1)}$. Thus, the posterior PDF of $\lambda_k^{(q)}$ is in this case a weighted sum of gamma distributions (see also Eq. (V.37)). As will be seen, for the simplicity of the presentation here, only the informative vector $\underline{\mu}_{(q)}$ is explicitly considered when $k^{(q)} > 1$.

3. The loss function. The cost functions corresponding to S are here the same as in Chap. I (see Eq. (I.4) and Sec. II.2).

The organization of the chapter is the following. Sec. V.1 consi-

ders the design of S loaded by $X_q(t)$ of known probability $p = P(X_1=0)$. Sec. V.1.1 treats the case where p=0. First, Sec. V.1.1.1, the CDF's of $y_{q,\tau}$ and y_τ are found assuming that the probabilistic law of $X_q(t)$ is perfectly known. Also, the probabilistic design of S loaded by $X_q(t)$ and by X(t) for $\mathcal{Y}_q$ of gamma type with integer shape parameter is calculated. Esteva [93] has used this model (with p=0) recently for seismic risk analysis. Sec. V.1.1.2 allows for uncertainty in the law of the input variable when the optimal code strategy and the best design of S are evaluated. Sec. V.1.2 considers the situation where $p \neq 0$ but perfectly known. Both, observable and unobservable $X_1(t)$ processes are analyzed. Sec. V.1.2 extends all the probabilistic and statistical results obtained in Sec. V.1.1 to the more general case when the probability $p = P(X_1=0) > 0$ is known.

Sec. V.2 treats the general case where p is assumed unknown. A general solution for the CDF of $y_{1,\tau}$ is obtained only for exponential interarrival times. An approximate approach is also proposed.

Sec. V.3 summarizes the results of this chapter. Also, the use of the decision framework in Sec. II.2 to model various loads such as earthquakes and tornados is suggested.

V.1 DECISION STRATEGY FOR RENEWAL SEQUENCES WITH RANDOM MAGNITUDES. THE CASE WHEN p IS KNOWN.

The calculation of the optimal level III and level IV design of S subjected to $X_q(t)$ is mainly considered here. Following the development

in Sec. II.2 first the optimal code strategy f(x) is found (see Table (I.1)). Then, f(x) is used to best design S in accordance with the code format. Also the evaluation of the optimal design of S loaded by the sum X(t) in Eq. (V.2) is briefly analyzed. In Eq. (V.2) $X_0(t)$ and $X_1(t)$ are assumed to be mutually independent.

This section considers only the case when the probability $P(X_1=0) = p$ is perfectly known. Sec. V.1.1 deals with the situations where $p = 0$ and Sec. V.1.2 solves the optimal design of S in the more general case when $p \neq 0$ but known.

V.1.1 Design Strategy for Renewal Sequences; p = 0

The optimal design of S loaded by $X_q(t)$ or X(t) when the CDF of X_q is continuous on the real line is analyzed here.

First, Sec. V.1.1.1, it is assumed that the probabilistic law of the input variable is perfectly known. The distributions of the extreme variables $y_{q,\tau}$ and y_τ are obtained assuming that the interarrival time variable γ_q is gamma distributed. Also the CDF of the maximum of the sum of two Borges processes during time τ is evaluated.

Then, Sec. V.1.1.2, generalizes the probabilistic results in Sec. V.1.1.1 to account for statistical uncertainty in the law of the input variable $X_q(t)$. Following Sec. II.2 the optimal code strategy and the best design of S are found.

V.1.1.1 Probabilistic Solution

Let $F^{(q)}(x;\tau)$ be the cumulative distribution function of the extreme variable $y_{q,\tau}$ in Eq. (V.1), i.e.

$$F^{(q)}(x;\tau) = P(y_{q,\tau} \leq x) = P(\max_{0\leq t\leq\tau}\{X_q(t)\} \leq x) \qquad \text{IV.10)}$$

Then, the probability of survival of S designed at X=d during time τ is $F^{(q)}(d;\tau)$.

Assuming that γ_q is exponentially distributed with decay parameter λ, one readily obtained [34,57,65,74] that

$$F^{(q)}(x;\tau) = [F_{x_q}(x)]^q \, e^{-\lambda\tau[1-F_{x_q}(x)]} \qquad \text{(V.11)}$$

where, $F_{x_q}(x)$ is the CDF of X_q, and

$$q = \begin{cases} 0 \ , & \text{for } X_0(t) \text{ in Fig. (V.1)} \\ 1 \ , & \text{for } X_1(t) \text{ in Fig. (V.2)} \end{cases}$$

The generalization of Eq. (V.5) to the case when γ_q is gamma distributed with integer shape parameter k and decay parameter λ, i.e. $\gamma_q \in G(k,\lambda)$, is due to Bickley [6]. The Laplace transform of $F^{(q)}(x;\tau)$ in Eq. (V.4) corresponding to <u>any</u> renewal time structure of $X_q(t)$ is [6]:

$$\mathcal{L}[F^{(q)}(x;\tau)] = \frac{[F_{x_q}(x)]^q}{s} \, \frac{1-G(s)}{1-G(s)F_{x_q}(x)} \qquad \text{(V.12)}$$

where, $G(s) = \mathcal{L}[f_{\gamma_q}(\tau)]$ is the Laplace transform of the PDF of the interarrival time γ_q.

Consequently, the evaluation of $\mathcal{L}[F^{(q)}(x;\tau)]$ depends strictly on the possibility of calculating G(s). For the case when $\gamma_q \in G(k,\lambda)$, one finds

$$G(s) = \frac{\lambda^k}{(s+\lambda)^k} \qquad \text{(V.13)}$$

where, k=1,2,.....; k=1 stands for exponential interarrival times.

From Eq. (V.12) and the value of G(s) in Eq. (V.13) corresponding to gamma interarrival times, it follows that

$$\mathcal{L}[F^{(q)}(x;\tau)] = \frac{[F_{X_q}(x)]^q}{s} \frac{(s+\lambda)^k-\lambda^k}{(s+\lambda)^k-F_{X_q}(x)\lambda^k} \qquad (V.14)$$

where, k=1,2,...

On taking the inverse Laplace transform of the result (V.14) the CDF of the extreme $y_{q,\tau}$ in Eq. (V.1) is obtained. Table (V.15) provides the expressions of $F^{(q)}(x;\tau)$ for $\gamma_q \in G(k,\lambda)$ with k=1,2,3.

Table (V.15)

k	$F^{(q)}(x;\tau)$
1 exponential interarrival times	$[F_X(x)]^q \, e^{-\lambda\tau[1-F_X(x)]}$
2	$D^{2q}[\frac{1+1/D}{2} e^{-\lambda\tau(1-D)} + \frac{1-1/D}{2} e^{-\lambda\tau(1+D)}]$ where, $D = [F_X(x)]^{1/2}$
3	$F^{3q}[\frac{1+1/F+1/F^2}{3} e^{-\lambda\tau(1-F)} - \frac{1/F^2+1/F-2}{3} e^{-\lambda\tau(1+F/2)} \cos\frac{\sqrt{3}}{2}\lambda F\tau - \frac{1/F^2-1/F}{\sqrt{3}} e^{-\lambda\tau(1+F/2)} \sin\frac{\sqrt{3}}{2}\lambda F\tau]$ where, $F = [F_X(x)]^{1/3}$

Note: $F_X(x)$ stands here for either $F_{X_0}(x)$ or $F_{X_1}(x)$.

The evaluation of $F^{(q)}(x;\tau)$ for gamma interarrival times of shape parameter $k > 3$, follows the same procedure. However, the computation of the CDF of the extreme $y_{q,\tau}$ becomes soon prohibitive when k increases.

Approximate expressions of $F^{(q)}(x;\tau)$ for the case when the level x is large are readily obtained from Table (V.15). Noting that F(x), D, $F \to 1$ and consequently $(1-1/D)$, $(1/F^2+1/F-2)$, $1/F^2-1/F) \to 0$ when $x \to \infty$, and replacing exp $(-\varepsilon)$ by $1-\varepsilon$, where $\varepsilon \geq 0$ and small, $F^{(q)}(x;\tau)$ may be approximately evaluated when x is large from

$$F^{(q)}(x;\tau) \simeq 1 - \lambda\tau\{1-[F_X(x)]^{1/k}\} \qquad (V.16)$$

where, $F_X(x)$ is as in Table (V.15) and k is the shape parameter of γ_q.

The result in Eq. (V.16) was evaluated only for k=1,2,3. However, it seems that (V.16) holds for any finite k. Also, it appears from Eq. (V.16) that the process defined by the crossings of $X_q(t)$ at level x, where x is high, follows approximately a Poisson process of mean rate $\lambda\{1-[F_X(x)]^{1/k}\}$, where k is the shape parameter of γ_q assumed here gamma.

Once $F^{(q)}(x;\tau)$ is obtained (from Table (V.15) or Eq. (V.16)) the probabilistic designs at level III, d_{III}, and level IV, d_{IV}, of S loaded by the process $X_q(t)$ are readily obtained. Thus, d_{III} is (see Eq. (I.3)) the P_f - upper fractile of $y_{q,\tau}$, i.e.

$$F^{(q)}(d_{III};\tau) = 1-P_f \qquad (V.17)$$

where, P_f is the (code) target failure probability

and d_{IV} is (see Eq. (I.4)) the minimum point of the expected cost func-

tion

$$E_S[d] = c_S(d) + c_{F,S}[1-F^{(q)}(x;\tau)] \qquad (V.18)$$

corresponding to the structural system S.

Consider now the design of the system S subjected to the process sum X(t) in Eq. (V.2). It is again assumed that the probabilistic laws of $X_0(t)$ and $X_1(t)$ are perfectly known. Only the approximate evaluation of the CDF of the extreme y_τ in Eq. (V.3) due to Peir [65], McGuire and Cornell [57] is considered here. Their result holds only for exponential interarrivals times. To calculate the probability $P(y_\tau \leq x)$, the following variables are first considered:

$$L_1 = \max_{0\leq t\leq\tau} \{X_1(t)\}$$

$$L_2 = \max_{t\in(\text{one holding time of } X_1(\cdot))} \{X_0(t)\}$$

$$L_3 = X_1$$

$$L_4 = \max_{0\leq t\leq\tau} \{X_0(t)\} \qquad (V.19)$$

It was observed [57,65] that the probability $P(y_\tau \leq x)$ can be approximated by any of the following equations:

$$P(y_\tau \leq x) \cong P(L_1+L_2 \leq x)\ P(L_3+L_4 \leq x) \qquad (IV.20)$$

or

$$P(y_\tau \leq x) \cong P(L_1+L_2 \leq x) \tag{V.21}$$

referred to as "CASE I" in [57], or

$$P(y_\tau \leq x) \cong P(L_3+L_4 \leq x) \tag{V.22}$$

referred to as "CASE II" in [57]

where, provided that p = 0, and that $X_q(q=0,1)$ is positive:

$$P(L_1+L_2 \leq x) = \int_0^x \frac{[\lambda_1+\lambda_1^2\tau F_{X_1}(\xi)]f_{X_1}(\xi)e^{-\lambda_1\tau[1-F_{X_1}(\xi)]}}{\lambda_1 + \lambda_0[1-F_{X_0}(x-\xi)]}\, d\xi \tag{V.23}$$

$$P(L_3+L_4 \leq x) = \int_0^x e^{-\lambda_0\tau[1-F_{X_0}(x-\xi)]} f_{X_1}(\xi)d\xi \tag{V.24}$$

From the numerical studies presented in [42,57] it appears that the "CASE I" (see Eq. (V.21) and Eq. (V.23)) is generally better than both solutions in Eq. (V.20) and Eq. (V.22) (see also [34]).

On approximating the CDF, $F_{y_\tau}(x)$, of the variable y_τ by any of the equations (V.20), (V.21), (V.22) and replacing $F^{(q)}(x;\tau)$ in Eq. (V.17) and Eq. (V.18) by $F_{y_\tau}(x)$, the optimal probabilistic level III and level IV designs of S stressed by X(t) in Eq. (V.2) result, respectively.

The extension of Peir and McGuire approximations to the case when $\gamma_q \in G(k,\lambda)$ is straightforward. Thus, the CDF's of the variables L_1 and L_4, and of L_2 conditional on the holding time $\bar{\tau}$ of $X_1(t)$, i.e. $L_2|\bar{\tau}$ are found in Table (V.15) for k=1,2,3. Thus,

$$P(L_1 \leq x) = F^{(1)}(x;\tau)$$

$$P(L_2 \leq x|\bar{\tau}) = F^{(0)}(x;\bar{\tau})$$

$$P(L_4 \leq x) = F^{(0)}(x;\tau) \qquad (V.25)$$

The unconditional CDF of L_2 is:

$$P(L_2 \leq x) = \int_0^\infty P(L_2 \leq x|\bar{\tau}) f_{\gamma_1}(\bar{\tau}) d\bar{\tau} \qquad (V.26)$$

where, $f_{\gamma_1}(\bar{\tau})$ is the desnity of γ_1.

For example, assuming $\gamma_0 \in G(2,\lambda_0)$ and $\gamma_1 \in G(3,\lambda_1)$, the result of the integration in (V.26) is:

$$P(L_2 \leq x) = \frac{\lambda_1^3}{2}\left\{\frac{1+1/D}{[\lambda_0(1-D)+\lambda_1]^3} + \frac{1-1/D}{[\lambda_0(1+D)+\lambda_1]^3}\right\} \qquad (V.27)$$

where, D is here $[F_{X_0}(x)]^{1/2}$.

Consequently, the probabilities $P(L_1+L_2 \leq x)$ and $P(L_3+L_4 \leq x)$ are the output of only one convolution as in the case when γ_q is assumed to be exponentially distributed (see Eq. (V.23) and Eq. (V.24)).

Following the idea in [57,65], the CDF of the maximum of the sum, $y_{B,\tau} = \max_{0\leq t\leq\tau}\{X_{B,0}(t)+X_{B,1}(t)\}$ of two Borges processes during τ (see Fig. (V.6)) may be approximated by the distributions of the so called [57] "CASE I" and "CASE II" random variables, respectively:

$$P\Big(\max_{0\le t\le\tau}\{X_{B,1}(t)\} + \max_{t\epsilon t_{B,1}}\{X_{B,0}(t)\} \le x\Big) =$$

$$= r_{0,1}\int_o^x [F_{X_{B,1}}(x-\xi)]^{r_1} f_{X_{B,0}}(\xi)[F_{X_{B,0}}(\xi)]^{r_{01}-1} d\xi \qquad (V.28)$$

and

$$P(X_{B,1} + \max_{0\le t\le\tau}\{X_{B,0}(t)\} \le x) = \int_o^x [F_{X_{B,0}}(x-\xi)]^{r_0} f_{B,1}(\xi)d\xi \qquad (V.29)$$

where, r_0, r_1 and r_{01} are defined in Fig. (V.6).

$X_{B,0}(t)$ and $X_{B,1}(t)$ in $y_{B,\tau}$ are independent.

The exact expression of the distribution of $y_{B,\tau}$, [9], is:

$$P(y_{B,\tau} \le x) = \{\int_o^x [F_{X_{B,0}}(x-\xi)]^{r_{01}} f_{X_{B,1}}(\xi)d\xi\}^{r_1} \qquad (V.30)$$

The essential assumption in this section was that the probabilistic law of the load $X_q(t)$ is perfectly known. Also, whenever a convolution was written it was considered that the random variables we are dealing with are positive. From Eq. (V.17) and Eq. (V.18) the probabilistic design levels of S corresponding, respectively, to a level III and a level IV code format were calculated. The next section considers the more general situation when at least one component of the law of $X_q(t)$ (i.e. either the time structure γ_q, or the event size X_q) is not perfectly known. In this case the optimal code strategy is inferred from the observation vectors $\underline{\mu}_{(q)}$ (or $\underline{\mu}^*_{(q)}$ provided that γ_q is gamma distributed), $\underline{Z}_{(q)}$ and the code cost functions.

V.1.1.2 Statistical Solution

Consider first the design of S loaded by the partially known process $X_q(t)$. The design of S which corresponds to this case is called statistical design.

In this section it is assumed that P_f in Eq. (V.17) is small and the failure cost $c_{F,S}$ in Eq. (V.18) is large such that the designs d_{III} and d_{IV} of S to be sufficiently high and therefore the approximation in Eq. (III.5) to hold. This assumption permits us to express the simultaneous predictive distribution of the random variable X as a product of the simple predictive distributions of X.

The next three paragraphs evaluate the predictive CDF of $y_{q,\tau}$ in Eq. (V.1) for various informative situations:

(i) known time structure, unknown magnitude

Let $P(N=n;\tau)$ be the probability of having n renewals during the life time τ of S. Then, the CDF of $y_{q,\tau}$ conditional on $\underline{\theta}$, $F^{(q)}(x;\tau/\underline{\theta})$, where $\underline{\theta}$ are the unknown parameters of the CDF, $F_{X_q}(x/\underline{\theta})$, of X_q is [6]:

$$F^{(q)}(x;\tau/\underline{\theta}) = [F_{X_q}(x/\underline{\theta})]^q \sum_{n=1}^{\infty} [F_{X_q}(x/\underline{\theta})]^n \, P(N=n;\tau) \qquad (V.31)$$

For example, when $\gamma_q \in G(1,\lambda)$

$$P(N=n;\tau) = \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau} \qquad (V.32)$$

The predictive CDF, $F^{(q)}(x;\tau/\underline{Z}_{(q)})$, of $y_{q,\tau}$ is obtained by integrating versus $\underline{\theta}$ the probability, $F^{(q)}(x;\tau/\underline{\theta})$ weighted with the posterior PDF of $\underline{\theta}$; it results:

$$F^{(q)}(x;\tau/\underline{Z}_{(q)}) = \sum_{n=1}^{\infty} F_{X_q}(\underline{x}_{(n+q)}/\underline{Z}_{(q)})\ \underline{P}(N=n;\tau) \qquad (V.33)$$

where, $F_{X_q}(\underline{x}_{(n+q)}/\underline{Z}_{(q)})$ is the order (n+q) simultaneous predictive CDF of X_q at $\underline{x}^t_{(n+q)} = \{x, x, \ldots, x\}$ (n+q elements)

Under the assumption that Eq. (III.5) holds, Eq. (V.20) becomes

$$F^{(q)}(x;\tau/\underline{Z}_{(q)}) = [F_{X_q}(x/\underline{Z}_{(q)})]^q \sum_{n=1}^{\infty} [F_{X_q}(x/\underline{Z}_{(q)})]^n\ P(N=n;\tau) \qquad (V.34)$$

where, $F_{X_q}(x/\underline{Z}_{(q)})$ stands for the simple predictive CDF of X_q.

Consequently, provided that the independent approximation in Eq. (III.5) holds, the predictive CDF of $y_{q,\tau}$ in the case (i) has the same functional form as the actual CDF of $y_{q,\tau}$ obtained when the law of X_q is perfectly known (compare Eq. (V.31) for known $\underline{\theta}$ with Eq. (V.34)). Consequently, $F^{(q)}(x;\tau/\underline{Z}_{(q)})$ in Eq. (V.34) is obtained directly from Table (V.15) for the case when k=1,2,3 by simply replacing $F_{X_q}(x)$ by $F_{X_q}(x/\underline{Z}_{(q)})$.

Consider now the calculation of the optimal code strategy. Let $\mathcal{H}_q^{\cdot\cdot} = \{H_\xi^{(q)}\}$ be the operational metamodel of $X_q(t)$, where $H_\xi^{(q)}$ stands for the distributional form of X_q. From Eq. (V.34), the predictive CDF of $y_{q,\tau}$ under $H_\xi^{(q)}$, $F_\xi^{(q)}(x;\tau/\underline{Z}_{(q)})$, is:

$$F^{(q)}(x;\tau/\underline{Z}_{(q)}) = F^{(q)}(x;\tau)\,\big|_{F_{X_q}(x) = F_\xi^{(q)}(x/\underline{Z}_{(q)})} \qquad (V.35)$$

where, $F_\xi^{(q)}(x/\underline{Z}_{(q)})$ is the simple predictive CDF of X_q conditional on $H_\xi^{(q)}$ and for given $\underline{Z}_{(q)}$. $F_\xi^{(q)}(x/\underline{Z}_{(q)})$ is calculated

following the methodology in Chap. III when the prior to posterior analysis for independent sequences was studied.

Once the family of predictive distributions $\{F^{(q)}(x;\tau/\underline{Z}_{(q)})\}_{\xi}$ is found, the decision framework in Sec. II.2 calculating the optimal metamodel, the optimal code strategy and the optimal design of S is directly applicable (see also the paragraph (iii)).

(ii) known magnitude, unknown time structure

Assume that the magnitude X_q of the input variable $X_q(t)$ is perfectly known and that the interarrival time $\gamma_q \in G(k^{(q)},\lambda_k^{(q)})$, where $k^{(q)}$ is known but the decay parameter $\lambda_k^{(q)}$ is not. Let $\underline{\mu}_{(q)}$ (or $\underline{\mu}^*_{(q)}$) in Fig. (V.4) be the observation vector associated with the variable γ_q.

Conditional on $\lambda_k^{(q)}$ the CDF, $F_k^{(q)}(x;\tau/\lambda_k^{(q)})$, of $y_{q,\tau}$ is found for $k^{(q)} = 1,2,3$ in Table (V.15). Consider that the apriori information about $\lambda_k^{(q)}$ (see Sec. III.2.2.2) is quantified by the conjugate prior distribution on $\lambda_k^{(q)}$, here of gamma type with parameters $(s_k^{(q)},\mu_k^{(q)})$.

Provided that the observation vector $\underline{\mu}_{(q)}$ is considered, a posteriori, $\lambda_k^{(q)}$ is gamma distributed with parameters $(a_k^{(q)},b_k^{(q)})$, where:

$$\begin{aligned} a_k^{(q)} &= n_{(q)}k^{(q)} + s_k^{(q)} \\ b_k^{(q)} &= n_{(q)}\hat{\mu}_{(q)} + \mu_k^{(q)} \end{aligned} \qquad (V.36)$$

$$\text{in which } \hat{\mu}_{(q)} = \frac{1}{n_{(q)}} \sum_{i=1}^{n_{(q)}} \mu_{q,i}$$

While the measurement vector $\underline{\mu}^*_{(q)}$ is used, the posterior PDF, $f_k^{(q)}(\lambda_k^{(q)}|\underline{\mu}^*_{(q)})$, of $\lambda_k^{(q)}$ has the form

$$f_k^{(q)}(\lambda_k^{(q)}/\underline{\mu}^*_{(q)}) \propto \sum_{i,j=0}^{k^{(q)}-1} \frac{(t_{q,e})^i}{i!} \frac{(t_{q,a})^j}{j!}$$

$$[\lambda^{n_{(q)}k^{(q)} + s_k^{(q)}+i+j} e^{-(T+\mu_k^{(q)})\lambda_k^{(q)}}] \qquad (V.37)$$

i.e. is a linear combination of gamma distributions of shape parameters $n_{ij} = n_{(q)}k^{(q)}+s_k^{(q)}+i+j+1$ and decay parameter $T + \mu_k^{(q)}$. The sufficient statistic of the likelihood $f_k^{(q)}(\lambda_k^{(q)}/\underline{\mu}^*_{(q)})$ in Eq. (V.37) is $(n_{(q)},T)$.

Clearly, when the renewals follow a Poisson process of mean rate $\lambda_1^{(q)}$ (i.e. $\gamma_q \in G(1,\lambda_1^{(q)})$), $f_1^{(q)}(\lambda_1^{(q)}/\underline{\mu}^*_{(q)})$ in Eq. (V.37) is gamma with parameters $a_1^{(q)} = n_{(q)}+s_1^{(q)}+1$ and $b_1^{(q)} = T+\mu_1^{(q)}$.

The predictive distribution of $y_{q,\tau}$ is obtained by integrating $F_k(x;\tau/\lambda_k^{(q)})$ weighted with the posterior PDF, $f_k^{(q)}(\lambda_k^{(q)}/\underline{\mu}_{(q)})$ or $f_k^{(q)}(\lambda_k^{(q)}/\underline{\mu}^*_{(q)})$, of $\lambda_k^{(q)}$ with respect to $\lambda_k^{(q)}$. Because of $f_k^{(q)}(\lambda_k^{(q)}/\underline{\mu}^*_{(q)})$ is expressed as a weighted sum of gamma distributions, because $f_k^{(q)}(\lambda_k^{(q)}/\underline{\mu}_{(q)})$ is a gamma distribution and because only linear operations are involved when calculating various statistical results, it is sufficient here to present only the solutions corresponding to $\underline{\mu}_{(q)}$. If necessary the results of $\underline{\mu}^*_{(q)}$ can be obtained from a weighted linear combination of the solutions presented.

For example, the CDF of $y_{q,\tau}$, $F_k^{(q)}(x;\tau/\underline{\mu}^*_{(q)})$, conditional on $\underline{\mu}^*_{(q)}$ is a weighted sum (see Eq. (V.37) for weights) of the predictive CDF's of the type in Eq. (V.38).

Thus, the predictive CDF of $y_{q,\tau}$ conditional on $k^{(q)}$ and for given $\underline{\mu}_{(q)}$ is:

$$F_k^{(q)}(x;\tau/\underline{\mu}_{(q)}) = \int_0^\infty F_k^{(q)}(x;\tau/\lambda_k^{(q)}) f_k^{(q)}(\lambda_k^{(q)}/\underline{\mu}_{(q)}) d\lambda_k^{(q)} \qquad (V.38)$$

The solution of the integration in Eq. (V.38) is given in Table (V.39) for $k^{(q)}=1,2,3$.

Table (V.39)

k	$F_k^{(q)}(x;\tau/\underline{\mu}_{(q)})$
1	$[F_X(x)]^q \dfrac{1}{1+\frac{\tau}{b_1}[1-F_X(x)]^{a_1}}$
2	$D^{2q}\{\dfrac{1+1/D}{2[1+\frac{\tau}{b_2}(1-D)]^{a_2}} + \dfrac{1-1/D}{2[1+\frac{\tau}{b_2}(1+D)]^{a_2}}\}$ where, $D = [F_X(x)]^{1/2}$
3	$F^{3q}\{\dfrac{1+1/F+1/F}{3} \dfrac{1}{[1+\frac{\tau}{b_3}(1-F)]^{a_3}} - (\dfrac{1/F^2-1/F}{2j\sqrt{3}} + \dfrac{1/F^2+1/F-2}{6})$ $[\dfrac{1}{[1+\frac{\tau}{b_3}(1+F/2-j\frac{\sqrt{3}}{2}F)]^{a_3}} + \dfrac{1}{1+\frac{\tau}{b_3}(1+F/2+j\frac{\sqrt{3}}{2}F)]^{a_3}}]$ where, $F = [F_X(x)]^{1/3}$, $j = \sqrt{-1}$

Note: $\{F_X(x), a_k, b_k\}$ stands here for either $\{F_{X_0}(x), a_k^{(0)}, b_k^{(0)}\}$ or $\{F_{X_1}(x), a_k^{(1)}, b_k^{(1)}\}$

In this case the operational metamodel is $\mathcal{H}_q^* = \{H_k^{(q)}\}$, where $H_k^{(q)}$ denotes the process $X_q(t)$ of (known) marginal CDF $F_{X_q}(x)$ and gamma interarrival times $\gamma_q \in G(k^{(q)}, \lambda_k^{(q)})$ with $k^{(q)}$ known.

Then, as in the case (i), the predictive CDF $F_k^{(q)}(x;\tau/\underline{\mu}_{(q)})$ is used to find the optimal design of S.

(iii) unknown time structure and magnitude

The design of S loaded by $X_q(t)$ with partially known time structure, γ_q, and event size, X_q, considered here is a typical design situation. For example, γ_0 may model earthquate interarrival times and X_0 their magnitude; clearly neither the PDF of γ_0 nor the PDF of X_0 are perfectly known.

The cases (i) and (ii) are particularization of (iii) and were considered mostly to smooth the transition from the probabilistic to the statistical solution.

Let $\mathcal{H}_q^* = \{H_{k,\xi}^{(q)}\}$ be the operational metamodel of $X_q(t)$, where $H_{k,\xi}^{(q)}$ stands for the process $X_q(t)$ having the marginal distribution of kind $\xi^{(q)}$ and γ_q gamma distributed with the shape parameter $k^{(q)}$.

On combining the results in the paragraphs (i) and (ii) of this section, the predictive CDF, $F_{k,\xi}^{(q)}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)})$, of $y_{q,\tau}$ under $H_{k,\xi}^{(q)}$ and for given $\{\underline{\mu}_{(q)},\underline{Z}_{(q)}\}$ result:

$$F_{k,}^{(q)}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)}) = F_k^{(q)}(x;\tau/\underline{\mu}_{(q)})\Big|_{F_{X_q}(x) = F_\xi^{(q)}(x/\underline{Z}_{(q)})} \qquad (V.40)$$

where, $F_k^{(q)}(x;\tau/\underline{\mu}_{(q)})$ is found in Eq. (V.24) and $F_\xi^{(q)}(x/\underline{Z}_{(q)})$ is the simple predictive CDF of the magnitude sequence $\{X_q\}$ under the distributional form of type ξ and for given observa-

tion $\underline{Z}_{(q)}$.

The predictive CDF, $F_{\mathcal{H}_q}(x;\tau)$, of the Bayesian model of $\mathcal{H}_q \subseteq \mathcal{H}_q^*$ of the load process $X_q(t)$ is:

$$F_{\mathcal{H}_q}(x;\tau) = \sum_{H_{k,\xi} \in \mathcal{H}_q} p_{k,\xi} F_{k,\xi}^{(q)}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)}) \qquad (V.41)$$

where, $F_{k,\xi}^{(q)}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)})$ is obtained in Eq. (V.40), (Eq. (V.35) and Eq. (V.38) are used when the time structure and the magnitude are known, respectively). $p_{k,\xi}$ is the posterior probability of the model $H_{k,\xi}^{(q)}$ conditional on $\mathcal{H}_q$. One may observe that, because of the independence between the time and the magnitude processes defining $X_q(t)$, $p_{k,\xi} = p_k p_\xi$, where p_k and p_ξ are the posterior probabilities of $H_k^{(q)}$ and $H_\xi^{(q)}$, respectively.

From Eq. (V.41) and Eq. (II.31) the optimal design level under $\mathcal{H}_q$ results. Then, by using (II.27) or (II.28) or (II.37) the optimal metamodel $\mathcal{H}_q^{(opt)}$ of $X_q(t)$ is found. From Eq. (II.21) and Eq. (II.25) the optimal level III and level IV code strategies are evaluated. If only pure strategies are considered and the space of pure strategies coincides with $\mathcal{H}_q^{(opt)}$, then the best code strategy, $f(x)$, is the predictive PDF of one of the models in $\mathcal{H}_q^{(opt)}$. This is the code phase.

The designer phase calculates the optimal design level of any specific system S. From Eq. (V.17) and Eq. (V.18) the statistical level III and level IV designs of S are respectively obtained if $F^{(q)}(x;\tau)$ is repaced by $F(x)$, where $F(x)$ is the CDF of the optimal code strategy $f(x)$

(see also Table (I.1)).

The predictive CDF of $y_{q,\tau}$ corresponding to various information levels considered in Sec. V.1.1.1 and Sec. V.1.1.2 are summarized in Table (V.42).

Table (V.42)

The CDF of the extreme $y_{q,\tau}$

		Event size: $X_q \in H_\xi^{(q)}$	
		$\underline{\theta}_\xi^{(q)}$ = known	$\underline{\theta}_\xi^{(q)}$ = unknown
Time structure: $\gamma_q \in H_k^{(q)}$ i.e. $\gamma_q \varepsilon G(k^{(q)}, \lambda_k^{(q)})$	$\lambda_k^{(q)}$ = known	Probabilistic solution: $F^{(q)}(x;\tau)$ in Table (V.15)	Statistical solution: $F_\xi^{(q)}(x;\tau/\underline{Z}_{(q)})$ in Eq. (V.35)
	$\lambda_k^{(q)}$ = unknown	Statistical solution: $F_k^{(q)}(x;\tau/\underline{\mu}_{(q)})$ in Eq. (V.38); see also Table (V.39)	Statistical solution: $F_{k,\xi}^{(q)}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)})$ in Eq. (V.40)

Note: $\underline{\theta}_\xi^{(q)}$ are the parameters of the model $H_\xi^{(q)}$.

Approximate solutions of the predictive CDF of the extreme $y_{q,\tau}$ provided that the level x is high can be obtained either from Table (V.39) observing that $F_X(x)$, D, $F \to 1$ when x increases indefinitely or directly from Eq. (V.16).

Thus, the approximate expression of $F_{k,\xi}^{(q)}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)})$ corresponding to Table (V.39) is:

$$F_{k,\xi}^{(q)}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)}) \cong 1 - \frac{\tau}{b_k^{(q)}} \{1 - [F_{X_q}(x/\underline{Z}_{(q)})]^{1/k^{(q)}}\}^{a_k^{(q)}} \quad (V.43)$$

and to the approximation (V.16) is

$$F^{(q)}_{k,\xi}(x;\tau/\underline{\mu}_{(q)},\underline{Z}_{(q)}) = 1 - E[\lambda^{(q)}_k/\underline{\mu}_{(q)}]\tau\{1 - [F_{X_q}(x/\underline{Z}_{(q)})]^{1/k^{(q)}}\} \tag{V.44}$$

where, (see Eq. (V.36) and Eq. (V.37))

$$E[\lambda^{(q)}_k/\underline{\mu}_{(q)}] = a^{(q)}_k/b^{(q)}_k \tag{V.45}$$

The remainder of this section considers a load combination problem. Assume that the structural system S (see Fig. (I.1)) is loaded simultaneously by the processes $X_0(t)$ and $X_1(t)$, i.e. by the process X(t) in Eq. (V.2). Consider the case when the law of X(t) is not known in parametric sense and assume first that the interarrival times γ_q are exponentially distributed.

The predictive CDF of (L_1+L_2) appears to be analytically inconvenient in practical applications. This is an unfortunate result since [42,57] the "CASE I" approximation seems to work somewhat better than "CASE II". On the other hand the predictive CDF of (L_3+L_4) is readily obtained by direct integration versus the unknown parameters, $\underline{\tilde{\theta}}$, of X(t) of the result in Eq. (V.24) weighted with posterior density of $\underline{\tilde{\theta}}$. It follows, provided that the approximation (III.5) is still accepted, that:

$$P(L_3+L_4 \leq x/\underline{\mu}^*_{(0)},\underline{Z}_{(0)},\underline{Z}_{(1)}) = \int_0^x \frac{f_{X_1}(\xi/\underline{Z}_{(1)})}{1+\frac{\tau}{b^{(0)}_1}[1+F_{X_0}(x-\xi/\underline{Z}_{(0)})]^{a^{(0)}_1}}\, d\xi. \tag{V.46}$$

where, $F_{X_0}(x/\underline{Z}_{(0)})$ and $f_{X_1}(x/\underline{Z}_{(1)})$ are the predictive CDF of X_0

and the predictive PDF of X_1, respectively, and $(a_1^{(0)}, b_1^{(0)})$ are given in the discussion following Eq. (V.37).

When γ_0 is a gamma variable of, say, shape parameter two, (see Eq. (v.25) and Table (V.39)) and γ_1 has any distribution, the predictive CDF of (L_3+L_4) when accepting (III.5) has the form:

$$P(L_3+L_4 \leq x | \underline{\mu}_{(0)}, \underline{Z}_{(0)}, \underline{Z}_{(1)}) = \int_o^x \{ \frac{1+1/D}{2[1+\frac{\tau}{b_2^{(0)}}(1-D)]^{a_2^{(0)}}} + \frac{1-1/D}{2[1+\frac{\tau}{b_2^{(0)}}(1+D)]^{a_2^{(0)}}} \} f_{X_1}(\xi/\underline{Z}_{(1)}) d\xi \quad \text{(V.47)}$$

where, $D = [F_{X_0}(x-\xi/\underline{Z}_{(0)})]^{1/2}$ a

Since at least in the Gaussian case [78] the following is true

$$P(\bigcap_{i=1}^{n} Y_i \leq y_i) \geq \prod_{i=1}^{n} P(Y_i \leq y_i) \quad \text{(V.48)}$$

where, $\underline{Y} \in N(\underline{m}_Y, \underline{\Sigma}_Y)$ and

$$\underline{Y}^t = (Y_1, Y_2, \ldots, Y_n)$$

provided that all the elements γ_{ij} of $\underline{\Sigma}_Y$ are positive, it seems that the effect of using the approximation (III.5) in Eq. (V.46) and Eq. (V.47) is to overestimate the probability that the variable (L_3+L_4) exceeds the level x. Thus, Eq. (III.46) and Eq. (III.47) seem to be conservative. Similar conclusions hold when the predictive CDF of $y_{B,\tau}$ in Eq. (V.28) or

Eq. (V.29) or Eq. (V.30) is calculated.

After the CDF of y_τ is obtained the decision framework in Sec. II.2 (see also the paragraph (iii) provides the optimal code strategy. Then, S is going to be designed in agreement with the code format (see Eq. (I.3) and Eq. (I.4)).

In the Section V.1.1 it was assumed that the cumulative distribution function of X_q has no discontinuity. Sec. V.1.2 considers situations in which the PDF of X_q has a spike at the origin of known area p.

V.1.2 Decision Strategy for Renewal Sequences; $p \neq 0$ but known

At the beginning of this chapter the notion of observable process has been introduced. In the sense of that definition $X_1(t)$ is always observable when $p = 0$ and may or may not be observable while $p \neq 0$.

Two special cases of $X_1(t)$ processes with $p \neq 0$ which are observable are given next:

(a) Assume that p is very small such that $p^2 << p$ and therefore the probability of having two zero values of X_q in the row is negligible. Consequently it may be assumed that the process is observable (see Fig. (V.5)).

(b) Consider the case where $X_1(t)$ models the sustained loads in a portion of a building. Then, any renewal represents a change of occupancy. Assume that any new occupancy may have zero value with probability p. Clearly in this case all the renewals are observed regardless of the values of $X_1(t)$.

The fact that $X_1(t)$ is or is not observable is reflected only in the time measurement vector to be considered when estimating $\mathcal{Y}_q$. Thus,

(see Fig. (V.5)) when $X_1(t)$ is observable and γ_1 is a gamma variable, the predictive CDF corresponding to $\underline{\mu}^*_{(1)}$ can be used by following directly Sec. V.1.1.2. If $X_1(t)$ is not observable and $k^{(1)} > 1$ the analysis is done conditional on $\underline{\mu}_{(1)}$. (The case $k^{(1)} = 1$ is given another solution in Sec. V.2.)

The specific results in this section correspond to the observation vectors $\underline{\mu}_{(1)}$ or $\underline{\tilde{\mu}}^*_{(1)}$ (see Fig. (V.5)). Consider first the derivation of the distribution of the variables $y_{q,\tau}$ and y_τ under the assumption that the probabilistic law of $X_q(t)$ is perfectly known.

Then, following the derivation in Sec. V.1.1 and [6] it may be seen that all probabilistic results summarized in Table (V.15) still hold in the case when X_1 has a finite mass $p \neq 0$ of being zero. However, the equations (V.23) and (V.24) become, respectively, [34]:

$$P(L_1+L_2 \leq x) = P(L_1 \leq x)\frac{\lambda_1}{\lambda_0+\lambda_1} + \int_o^x \frac{P(L_1 \leq x-\xi)\lambda_1\lambda_0 f_{X_0}(\xi)}{\{\lambda_1+\lambda_0[1-F_{X_0}(\xi)]\}^2}\, d\xi \qquad (V.49)$$

and

$$P(L_3+L_4 \leq x) = F_{X_1}(x)e^{-\lambda_0\tau} + \int_o^x F_{X_1}(x-\xi)\lambda_0\tau\, f_{X_0}(\xi)e^{-\lambda_0\tau[1-F_{X_0}(\xi)]}\, d\xi \qquad (V.50)$$

$$\text{where, } P(L_1 \leq x) = F_{X_1}(x)e^{-\lambda_1\tau[1-F_{X_1}(x)]}$$

Consider now the calculation of the predictive CDF of $y_{1,\tau}$ in Eq. (V.1) when $p \neq 0$ and the law of $X_1(t)$ (i.e. the distributional form of γ_1 and X_1) is only partially known. Assume first that the PDF of γ_1 is known (case (i) in Sec. V.1.2) and that the CDF of X_1 (X_1 is assumed here

to be non-negative with probability one) has the form:

$$F_{X_1}(x) = p + (1-p)\,\overline{F}_{X_1}(x) \qquad (V.52)$$

where, $\overline{F}_{X_1}(0) = 0$ and $\overline{F}_{X_1}(x)$ is not known in parametric sense.

The simple predictive CDF of X_1 conditional on the model $H_\xi^{(1)}$ (see case (i) in Sec. V. 1.2.) and on the probability p corresponding to the observation vector $\underline{Z}_{(1)}$ is:

$$F^{(1)}(x/\underline{Z}_{(1)}) = p + (1-p)\,\overline{F}_\xi^{(1)}(x/\underline{Z}_{(1)}) \qquad (V.53)$$

where, $\overline{F}_\xi^{(1)}(x/\underline{Z}_{(1)})$ is the simple predictive CDF of $X_I / X_1 > 0$.

It follows that the result (V.35) still holds if one uses for the predictive CDF of X_1 the equation (V.53). The same thing is valid also for the output of the case (ii), Eq. (V.38), and of the case (iii), Eq. (V.40).

The evaluation of the predictive CDF of the extreme value y_τ associated with the input variable $X(t) = X_0(t) + X_1(t)$ is discussed next. As in Sec. V.1.2 the predictive distribution of the variable L_1+L_2 ("CASE I" in [57]) is obtained only in implicit integral form. The predictive CDF of the variable L_3+L_4 ("CASE II" in [57]) is found by direct computation from Eq. (V.50), the simple predictive probabilities of X_q ($q=0,1$) and the posterior densities of the parameters $\lambda_k^{(q)}$ of γ_q ($q=0,1$):

$$P(L_3+L_4\leq x/\underline{Z}_{(0)},\underline{Z}_{(1)},\underline{\mu}^*_{(0)},\underline{\mu}^*_{(1)},T) = P(L_3+L_4\leq x/\underline{Z}_{(0)},\underline{Z}_{(1)},\underline{\mu}^*_{(0)},T) =$$

$$= \frac{F_{X_1}(x)}{\left(1+\frac{\tau}{b_1^{(0)}}\right)^{a_1^{(0)}}} + a_1^{(0)}\tau \int_0^x \frac{F_{X_1}(x-\xi) f_{X_0}(\xi)}{\left[1+\frac{\tau}{b_1^{(0)}}[1-F_{X_0}(\xi)]\right]^{a_1^{(0)}+1}} d\xi \qquad (V.54)$$

where, $(a_1^{(0)}, b_1^{(0)})$ are obtained in Eq. (V.37),

$F_{X_1}(x)$ is the simple predictive CDF of X_1 in Eq. (V.53),

and $f_{X_0}(x)$, $F_{X_0}(x)$ are simple predictive probabilities of X_0.

One may also observe that the result (V.54) can be generalized to the case when the interarrival times γ_q are gamma distributed(see also Eq. (V.47)).

Once the predictive probabilities of the extreme variables $y_{q,\tau}$ and y_τ are evaluated, the decision framework developed in Sec. II.2 and sketched in Sec. V.1.1 can be directly used to infer the optimal code strategy. Then, the best design of S in Fig. I.1 results.

V.2 DECISION STRATEGY FOR RENEWAL SEQUENCES WITH RANDOM MAGNITUDES. THE CASE WHEN p IS NOT KNOWN.

The problem here is to find the optimal design of S loaded by $X_1(t)$, where $X_1(t)$ has partially known probabilistic law (inclusively the value of $p = P(X_1=0)$). Two solutions are given to this problem.

First assume that the time structure of $X_1(t)$ is of Poisson kind with the mean rate λ. Because the occurences of the "on" holding times of $X_1(t)$ are also Poisson of mean rate $\overline{\lambda} = \lambda(1-p)$, one may directly work with the "on"-process. For example, in Fig. (V.5) the observation vector corresponding to the "on"-process is $\underline{v}^t_{(1)} = \{v_{1,2}, v_{1,2}, v_{1,3}\}$. Also, in this approach it is not necessary to distinguish between observable and unobservable $X_1(t)$ processes.

From Eq. (III.36) the predictive probability of having m "on"-renewals during the future time τ is

$$f(m/\underline{v}_{(1)}) = \frac{\Gamma(a+m)}{m!\Gamma(a)} \left(\frac{b}{b+\tau}\right)^a \left(\frac{\tau}{b+\tau}\right)^m \qquad (V.55)$$

where, (a,b) are given for the Poisson case after Eq. (V.37).

The predictive CDF of the extreme value variable $y_{1,\tau}$ conditional on approximation (III.5) is:

$$P(y_{1,\tau} \leq x | \underline{Z}_{(1)}, \underline{v}_{(1)}) = \sum_{m \geq 0} [F_{X_1}(x|\underline{Z}_{(1)})]^m \frac{\Gamma(a+m)}{m!\Gamma(a)} \left(\frac{b}{b+\tau}\right)^a \left(\frac{\tau}{b+\tau}\right)^m \qquad (V.56)$$

where, $F_{X_1}(x|\underline{Z}_{(1)})$ is the simple predictive CDF of the magnitude variable X_1.

The result in Eq. (V.56) is in the form of an infinite series. However one has to consider only several terms in (V.56) towards evaluating the probability $P(y_{1,\tau} \leq x/\underline{Z}_{(1)}, \underline{v}_{(1)})$. The terms to be taken in Eq. (V.56) are found from the values of $f(m/\underline{v}_{(1)})$ in Eq. (V.55) at various m. Since the mean and the variance of the predictive number of

renewals in τ are:

$$E[m] = a\tau/b$$

$$\text{Var}[m] = \frac{a\tau}{b}(1+\frac{\tau}{b}) \qquad \text{(V.57)}$$

the probability in (V.56) may be approximated by

$$P(y_{1,\tau} \leq x/\underline{Z}_{(1)}, \underline{v}_{(1)}) \cong \sum_{m_1 \leq m \leq m_2} [F_{X_1}(x/\underline{Z}_{(1)})]^m \frac{\Gamma(a+m)}{m!\Gamma(a)} (\frac{b}{b+\tau})^a (\frac{\tau}{b+\tau})^m \qquad \text{(V.58)}$$

where, $m_1 = E[m] + k_1(\text{Var}[m])^{1/2}$

$m_2 = E[m] - k_2(\text{Var}[m])^{1/2}$, and

k_1, k_2 are positive numbers chosen as a function of the precision required in the probability $P(y_{1,\tau} \leq x/\underline{Z}_{(1)}, \underline{v}_{(1)})$.

In general, the number of terms considered in the approximation (V.58) for given accuracy increases with τ and with the number of observed renewals in T (see the form of the variance of the predictive number of renewals in τ, Var[m] in Eq. (V.57)). Unfortunately the extension of this approach to the case when γ_1 is not exponentially distributed lead to numerical difficulties, which come from the fact that the interarrival time variable, $\bar{\gamma}_1$, corresponding to the "on"-process has a different type of distribution than γ_1 has.

The distribution of $\bar{\gamma}_1$ is a geometric mixture of gamma distributions. For example, if $\gamma_1 \in G(2,\lambda)$, then the PDF of $\bar{\gamma}_1$ is:

$$f_{\bar{\gamma}_1}(x) = \frac{1}{2}\lambda(1-p)\{e^{-\lambda(1-\sqrt{p})x} - e^{-\lambda(1+\sqrt{p})x}\} \qquad (V.59)$$

Clearly, the use of the variable $\bar{\gamma}_1$ in our Bayesian decision framework is in general numerically prohibitive. For example, only to evaluate the conjugate prior PDF of (λ,p)corresponding to Eq. (V.59) for $\bar{\gamma}_1$, if it exists, is not an easy task.

Consequently, in the case when γ_1 is not an exponential variable a second approach is proposed instead. The idea here is to estimate first p and then to use the results in Sec. V.1.2 where $X_1(t)$ processes with $p \neq 0$ but known are considered.

The informative experiment E_1 corresponding to the input variable $X_1(t)$ consists of (see Fig. (V.5)):

- the vector $\underline{\mu}_{(1)}$ of observed interarrival times,
- the vector $\underline{Z}_{(1)}$ of observed magnitudes, and
- the total observation time T.

Since in many cases of interest in Structural Reliability (for example $X_1(t)$ may be the model for tornado loads, where γ_1 and $X_1 > 0$ the duration and the intensity of the tornado, respectively) the expected value, $E[\gamma_1]$, of γ_1 is much smaller than T, only the situations in which $T >> E[\gamma_1]$ are analyzed here. It follows, that under this restriction the total number of observed and fictitious renewals of $X_1(t)$ during T is very large.

Consider now an urn having balls of "on" and of "off"kind in the proportion (1-p) and p, respectively and assume that p is not known. The problem is to establish p from the informative experiment which consists here of N drawings with replacement (i.e., after each extraction

the ball is put back in the urn). Let M be the number of "off"-balls out of the total number N of drawings out. Then, if N is large, the estimate

$$\hat{p}(N) = M/N \tag{V.60}$$

of p is expected to have small variance (in fact when $N \to \infty$, $\hat{p}(N)$ converges to the actual value p).

One may observe that the result (V.60) can be used to estimate the probability $p = P(X_1=0)$. In fact the process $X_1(t)$ has two states: the "off"-state while $X_1(t) = 0$ and the "on"-state when $X_1(t) \neq 0$ which occur during any holding time γ_1 with probability p and (1-p), respectively. The average number of holding times (or drawings out) and the average number of holding times having $X_1 = 0$ during T are $T/E[\gamma_1]$ and $T_{off}/E[\gamma_1]$, respectively. T_{off} is the total time in T when $X_1(t)=0$. Consequently, the estimate $\hat{p}(T)$ of p is (Eq. (V.60)):

$$\hat{p}(T) = T_{off}/T \tag{V.61}$$

Thus, provided that $T >> E[\gamma_1]$ is true the estimate $\hat{p}(T)$ in Eq. (V.61) is expected to have small variance. Then the estimate $\hat{p}(T)$ of p is used as if it were the actual value of p. Consequently, all the results of Sec. V.1.2 are valid if one replaces p by the point estimate $\hat{p}(T)$ calculated in Eq. (V.61).

V.3 CONCLUSIONS

The design of the deterministic two-state (failure/survival) structural system S in Fig. (I.1) loaded by the process $X_q(t)$ (see Fig. (V.1) and Fig. (V.2)) or by the combined process X(t) in Eq. (V.3) was analyzed. Also the procedure to find the optimal code strategy corresponding to input variables $X_q(t)$ or X(t) is briefly presented.

First it was noticed that the random variables which control the reliability of S during time τ are the extremes $y_{q,\tau}$ and y_τ. The distribution of these variables have been extensively treated provided that the law of $X_q(t)$ is perfectly known [42,57,65].

The main result in this chapter consists of the generalization of the probabilistic results in [6,57,65] to account for statistical uncertainty in the law of $X_q(t)$. It was obtained that all probabilistic and their corresponding statistical solutions require similar numerical work with only one exception, namely when the predictive probability of the so called "CASE I" variable [57] is considered. This result is valid for any of the $X_1(t)$-processes analyzed in this chapter.

The models considered in this chapter seem to have a broad field of applications. For example, the process $X_0(t)$ was used to model extraordinary loads and $X_1(t)$, p=0, to describe the evolution of the sustained loads in a building [57,65].

The process $X_1(t)$ with $p \neq 0$ may be a good model for, say, wind storms, tornado load and snow load in certain climates. The main objection we have against the present use of $X_q(t)$ to model various loads is related to:

- the Poisson assumption on the arrivals (i.e. γ_q is taken exponentially distributed), and
- the choice of the distributional form of the event size X_q.

The results in this chapter allow one to consider also gamma interarrival times. The choice of the marginal distribution of X_q and of the shape parameter of γ_q can be done in the discrete decision framework developed in Sec. II.2. (Here it was assumed that the action space coincides with the optimal metamodel and that only pure strategies are tested.) The strategy $X_q(t)$ obtained in this way represents the optimal code strategy. The design of any individual system S is then readily obtained in accordance with the code format.

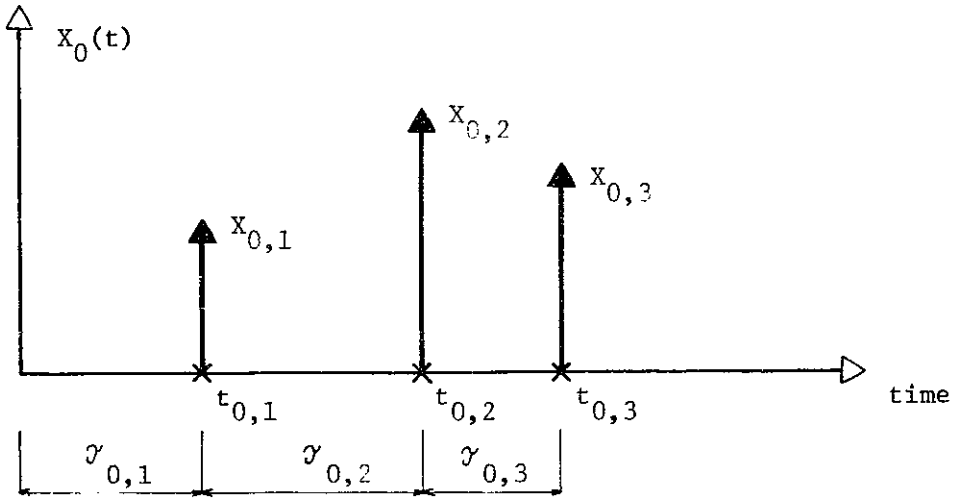


Fig. V.1. Renewal Point Sequence of interarrival time γ_0 and white stationary event size X_0.

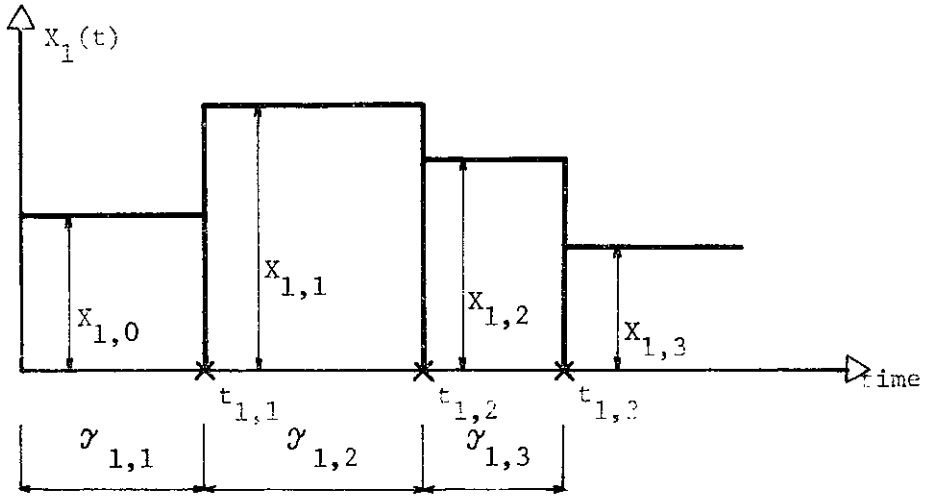
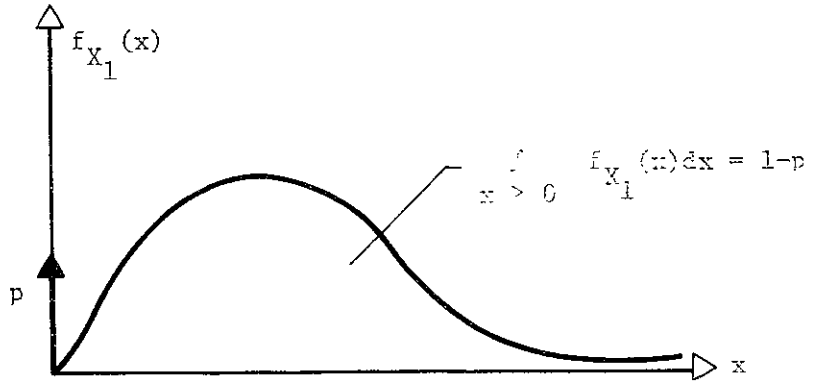
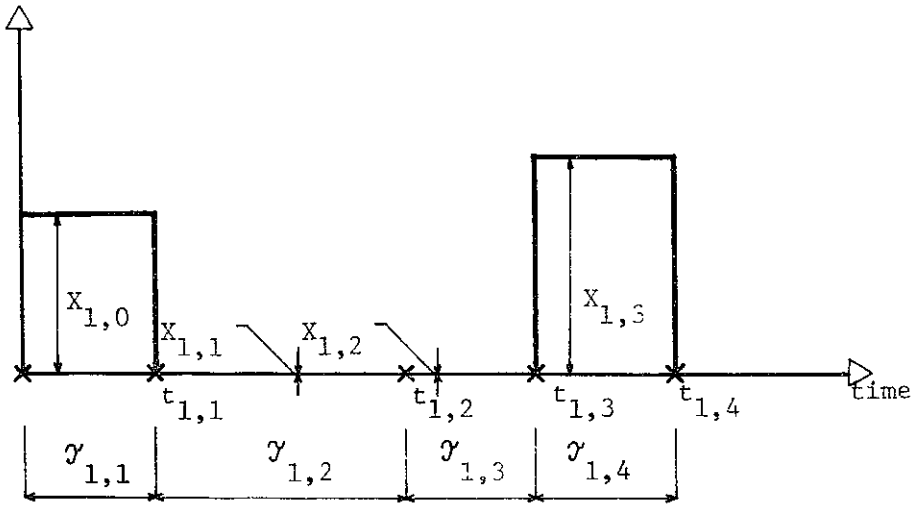


Fig. V.2. Renewal wave sequence of interarrival time γ_1 and white stationary event size X_1. The case when $P(X_1=0) = p$ is zero.

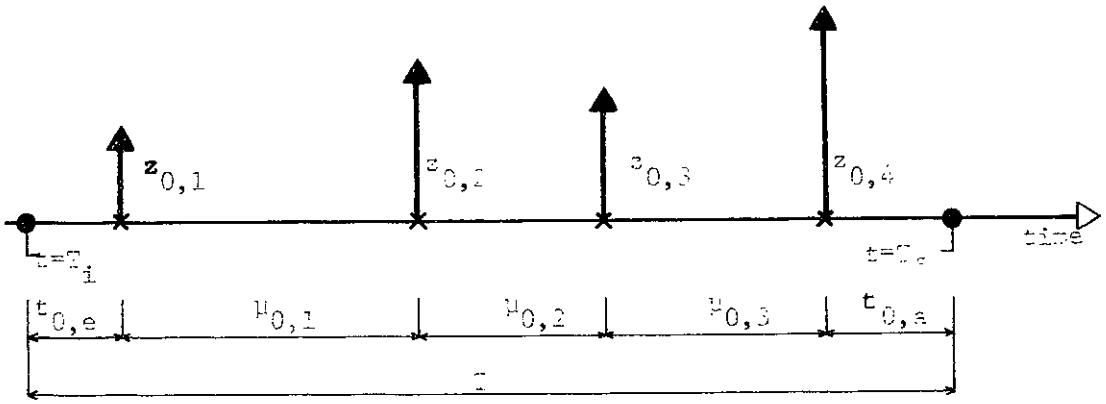


(a) The PDF of X_1; $p = P(X_1=0)$



(b) Hypothetical sample function of $X_1(t)$ when $p = P(X_1=0) \neq 0$

Fig. V.3. Renewal wave sequence; the case when the probability $P(X_1=0) = p$ is not zero.



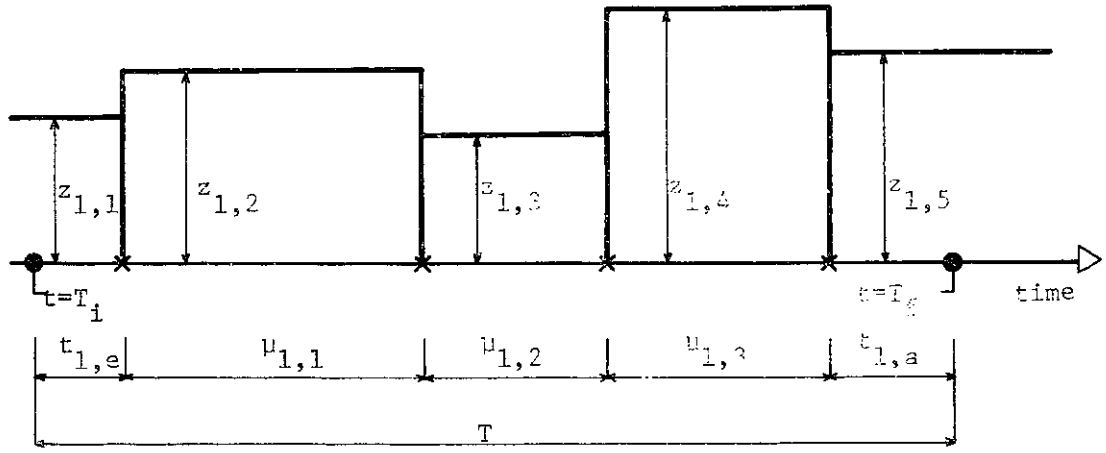
(a) Measurements of $X_0(t)$.

$$\underline{z}^t_{(0)} = \{z_{0,1}, z_{0,2}, z_{0,3}, z_{0,4}\}$$

$$\underline{\mu}^t_{(0)} = \{\mu_{0,1}, \mu_{0,2}, \mu_{0,3}\}$$

$$\underline{\mu}^{*t}_{(0)} = \{t_{o,e}, \underline{\mu}^t_{(0)}, t_{0,a}\}$$

$$n_{(0)} = 3 \ , \ n^*_{(0)} = 5$$



(b) Measurements of $X_1(t)$

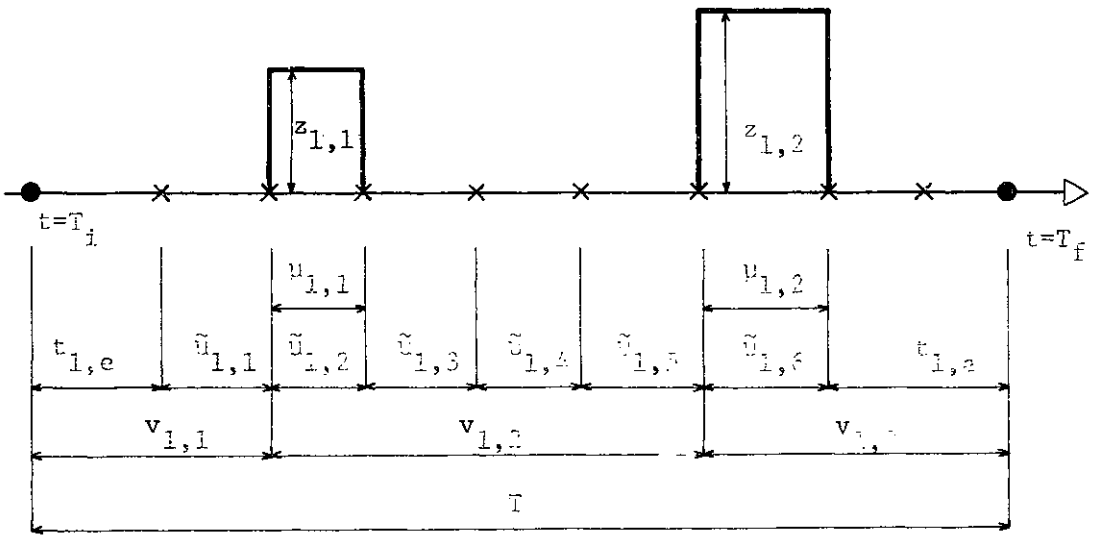
$$\underline{z}^t_{(1)} = \{z_{1,1}, z_{1,2}, z_{1,3}, z_{1,4}, z_{1,5}\}$$

$$\underline{\mu}^t_{(1)} = \{\mu_{1,1}, \mu_{1,2}, \mu_{1,3}\}$$

$$\underline{\mu}^{*t}_{(1)} = \{t_{1,e}, \underline{\mu}^t_{(1)}, t_{1,a}\}$$

$$n_{(1)} = 3 \ , \ n^*_{(1)} = 5$$

Fig. V.4. Observation vectors corresponding to the process $X_q(t)$. The case when $p = 0$.



$$\underline{\mu}^{t}_{(1)} = \{\mu_{1,1}, \mu_{1,2}\}$$

$$\underline{\tilde{u}}^{t} = \{\tilde{u}_{1,1}, \tilde{u}_{1,2}, \cdots, \tilde{u}_{1,6}\}$$

$$\underline{\tilde{\mu}}^{*t}_{(1)} = \{t_{1,e}, \underline{\tilde{\mu}}^{t}_{(1)}, t_{1,a}\}$$

$$n_{(1)} = 2 \ , \ n^{*}_{(1)} = 8$$

$$\underline{v}^{t} = \{v_{1,1}, v_{1,2}, v_{1,3}\}$$

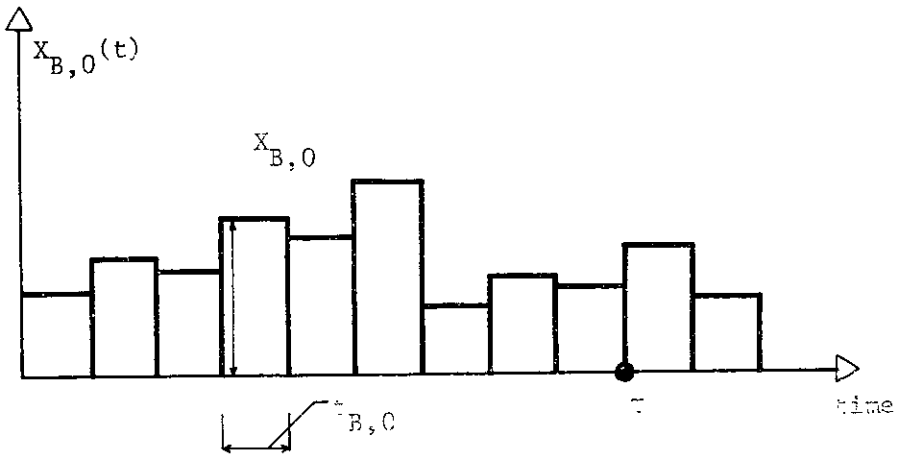
$$n_{(v)} = 3$$

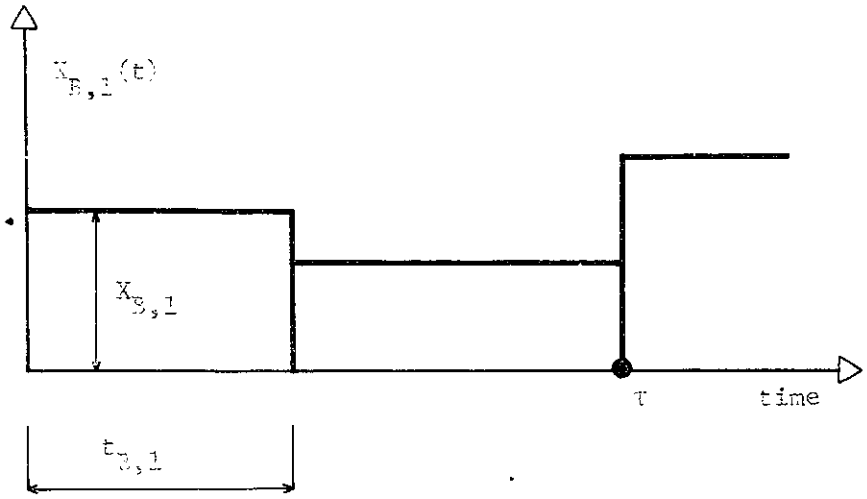
T = total observation time

$$T_{on} = \mu_{1,1} + \mu_{1,2}$$

$$T_{off} = T = T_{on}$$

Fig. V.5. Observation vectors corresponding to the process $X_1(t)$, when $p \neq 0$. The cases when $X_1(t)$ is or is not observable.





$$r_0 = \tau/t_{B,0}$$

$$r_1 = \tau/t_{B,1}$$

$$r_{01} = t_{B,1}/t_{B,0}$$

Fig. V.6. Borges processes.

CHAPTER VI

CONCLUSIONS AND SUGGESTIONS

The problem of calculating the optimal code strategy f(x) (see Table (I.1)) has been considered in this study. Typically, f(x) is chosen to be the best in the sense of statistical inference (for example, f(x) may maximize the likelihood). The optimality condition that f(x) satisfies in the approach developed in this work is of the utility type. Thus, first a set of models, $\mathcal{H}^*$, of the input variable $\underline{X}(t)$ was postulated (see (II.13)). Second, the best (in the utility sense defined by the code loss functions, by the penalty for nonsimplicity of the models $H \epsilon \mathcal{H}^*$ and by the evidence) metamodel, $\mathcal{H}^{(opt)}$, was calculated. Third, the optimal code strategy f(x) was identified as the minimum point of the expected cost $E[f|\mathcal{H}^{(opt)}]$ (see Eq. (II.21) and Eq. (II.25)). The evaluation of $\mathcal{H}^{(opt)}$ and f(x) was done in a Bayesian parametric decision framework. Establishing f(x) concludes the code phase. The designer phase is solved in accorance with the code format (see Eq. (I.3) and Eq. (I.4) for a level III and a level IV code, respectively).

The following are some of the conclusions related to the evaluation of the optimal code strategy f(x):

(1). The optimal metamodel, $\mathcal{H}^{(opt)}$, and the code cost functions determine f(x). Consequently, when evaluating $\mathcal{H}^{(opt)}$ it is necessary to test many models of various probabilistic laws in order to achieve a good compromise between accuracy and simplicity (i.e. $\mathcal{H}^*$ should be a large set). The methodology developed in Sec. II.2 has the capability of testing very many models without significant numerical work.

(2). Because $\mathcal{H}^{(opt)}$ is optimal only conditional on $\mathcal{H}^*$ and because f(x) is selected practically from a finite set of densities (see Eq. (II.15)) the "optimal" code strategy f(x) is only sub optimal in the utility sense defined in Sec. II.2. This drawback can be avoided if, say, a full nonparametrical approach is considered. We think that the extension of the methodology in this work to the case when $\mathcal{H}^*$ is replaced by the space of all models, $\mathcal{H}$, of $\underline{X}(t)$ and A stands for the (infinite) set of all densities will be a very interesting research subject and also a useful practical tool to search for the actual optimal code strategy.

(3). It is obtained (see, the example in Sec. III.3.1) that the design of S in Fig.(I.1) is very sensitive to the particular observation vector at hand when a fully inferential approach is used. This seems not to be the case in the methodology developed in this study. However, additional numerical studies are required.

In this context it seems that the disagreement related to the distribution type of the yearly wind extremes, $X_y(t)$, may simply be the drawback of using inferential approaches to select the best model of $X_y(t)$. To give a complete answer to this problem exten-

sive numerical work has to be performed. As a starting point, say, the evaluation of the optimal design of S corresponding to all the observation vectors in Tables (III.48) and (III.49) and to several failure cost functions is recommended.

While the ingredients of the Bayesian decision framework have been prepared for various kinds of sequences considered in this study, several aside results which may have practical implications have also been obtained:

(1) When designing S for input X(t), where X(t) is a correlated sequence, two situations may be distinguished. First, assume that S is the structural system in Fig. (I.1) loaded by X(t) and that S is going to be designed for long return periods of X(t). Then (see Sec. III.3.2 and Fig. (IV.4))the optimal design, $d^{(opt)}$, of S is not sensitive to the correlation structure of X(t). On the contrary, the functional form of the marginal PDF of X(t) is dominant for determining $d^{(opt)}$. Consequently (Sec. IV.4) in this case of high interest in Structural Reliability one may use only independent sequences to model X(t). Second, consider the design of the reservoir size S, where the input X(t) is the amount of water entering S at time t (see Fig. (I.2)). In this latter problem (see Sec. IV.3.3.1) both the correlation and the PDF shape of X(t) are significant for the size of S.

(2) In the prior to posterior analysis and in the predictive (extreme value) analysis of renewal sequences of random magnitudes (see Chap. V) several analytically simple statistical results for gamma interarrival times have been obtained. These results seem to be of

interest when modeling various loads such as wind storms, tornados or earthquakes [93]. Combined load processes are also considered.

To conclude, we think (based on the numerical evidence we have; see Sec. III.3.1) that the utility approach to model selection is more realistic in Structural Reliability than the present inferential methodology in model selection problems. Clearly, the fact that $\mathcal{H}^*$ is postulated and that $\mathcal{H}^{(opt)}$ is the best only conditional on $\mathcal{H}^*$ is a disadvantage of the proposed decision approach. However, given the large capability of our methodology to check many models when calculating $\mathcal{H}^{(opt)}$, the "optimal" code strategy f(x) is expected to be sufficiently accurate for the purpose of code optimization.

REFERENCES

1. Aitchinson, J., "Bayesian Tolerance Regions", J.R.Statist. Soc.. B, 26, 1964, pp. 161-175.

2. Aitchinson, J., and Sculthorpe, D., "Some Problems of Statistical Prediction", Biometrika, 52, 3 and 4, 469.

3. Astrőm, K., J. and Eykhoff, P., "System Identification - A Survey", Automatica, Vol. 7, pp. 123-162, 1971.

4. Bain, L. J. and Antle, C. E., "EStimation of Parameters in the Weibull Distribution", Technometrics, Vol. 9, No. 4, November, 1967, pp. 621-627.

5. Benjamin, J. R. and Cornell, C. A., Probability, Statistics and Decision for Civil Engineers, McGraw-Hill Book Company, 1970.

6. Bickley, D. T., "Probability Distribution of Maximum Sustained Load", Journal of the Structural Division, ASCE, Vol. 100, No. ST11, November, 1974, p. 2361-2365.

7. Blanc-Lapierre, A. and Fortet, R., Theory of Random Functions, Vol. I, Gordon and Breach, New York, 1965.

8. Bolotin, V. V., Primenenie Metodov Teorii Veroiatnostei i Teorii Nadejnosti v Raschetak Soorujenii, Izdatelstvo Literaturi po Stroitelstvi, Moskov, 1971.

9. Borges, J. F. and Castanheta, M., Structural Safety, 2nd edition, Lisbon, March, 1971.

10. Bosshard, W., "On Stochastic Load Combination", The John Blume Earthquake Engineering Center, Report #16, Stanford U., C. E. Dept., July, 1975.

11. Box, G. E. P., and Tiao, G. C., Bayesian Inference in Statistical Analysis, Addison-Wesley.

12. Box, G. E. P., and Tiao, G. C., "A Bayesian Approach to Some Outlier Problems", Biometrika, 55, pp. 119-129.

13. Brockett, R. W., Finite Dimensional Linear Systems, John Wiley and Sons, Inc., New York, 1970.

14. Cohen, A. C., "Maximum Likelihood Estimation in the Weibull Distribution Based on the Complete and on Censored Samples", Technometrics, Vol. 7, No. 4, November, 1965, pp. 579-588.

15. Cornell, C. A., "Structural Safety Specifications Based on Second-Moment Reliability Analysis, AIBSE Symposium on Concepts of Safety of Structures and Methods of Design, Vol. 4, Final Report, London, 1969, pp. 235-245.

16. Crellin, G. L., "The Philosophy and Mathematics of Bayes' Equation", IEEE. Trans. on Reliability, Vol. R-21, No. 3, August 1972, pp. 131-135.

17. D'Appolito, J. A., "The Evaluation of Kalman Filter Designs for Multisensor Integrated Navigation Systems", The Analytic Sciences Corp., AFAL-TR-70-271, (AD881286), January, 1971.

18. D'Appolito, J. A., and Hutchinson, C. E., "Low Sensitivity Filters for State Estimation in the Presence of Large Parameter Uncertainty", IEEE Trans. A.C., June, 1969, pp. 310-312.

19. Davenport, A. G., "On Statistical Prediction of Structural Performance in the Wind Environment", ASCE National Structural Engineering Meeting, Baltimore, Maryland, April, 19-23, 1971.

20. Davenport, W. B., Probability and Random Processes, McGraw-Hill Book

Company, New York, 1970.

21. DeGroot, M. H., Optimal Statistical Decisions, McGraw-Hill Book Company, New York, 1970.

22. DeGroot, M. H., "Reaching a Consensus", Journal of American Statistical Association, March, 1974, Vol. 69, No. 345, pp. 118-121.

23. Draper, N., and Smith, H., Applied Regression Analysis, Wiley, 1966.

24. Edwards, W., "Special Issue on Revision of Opinions by Men and Man-Mashine Systems. Introduction.", IEEE Trans. of Human Factors in Electronics, March 1966, Vol. HFE-7, No. 1, pp. 1-6.

25. Esteva, L., "Seismic Risk and Seismic Design Decisions", M.I.T., C. E. Dept., March, 1969.

26. Fishman, G. S., Concepts and Methods in Discrete Event Digital Simulation, John Wiley & Sons, New York, 1973.

27. Ferguson, Th. S., Mathematical Statistics. A Decision Theoretic Approach, Academic Press, New York and London, 1967.

28. Ferguson, Th. S., "A Bayesian Analysis of Some Nonparametric Problems", The Annals of Statistics, 1973, Vol. 1, No. 2, pp. 209-230.

29. Foster, M. H., and Martin, M. L., Probability, Confirmation and Simplicity. Readings in Philosophy of Inductive Logic., The Odyssey Press Inc., New York, 1966.

30. Gelb, A. and others, Applied Optimal Estimation, M.I.T. Press, 1974.

31. Glover, K., "Structural Aspects of System Identification", Ph.D. Thesis, E. E. Dept., M.I.T., 1973.

32. Gomes, L. and Vickery, B. J., "On the Prediction of Extreme Wind Speeds from the Parent Distribution", unpublished.

33. Griffin, R. E., and Sage, A. P., "Large and Small Scale Sensitivity Analysis of Optimum Estimation Algorithms", IEEE Trans. A. C., Vol. AC-13, No. 4, August, 1968, pp. 320-328.

34. Grigoriu, M. D., "On the Maximum of the Sum of Random Process Load Models", Internal Project Working Document #1, Structural Loads Analysis and Specifications Project, NSF RANN GI-44060, M.I.T., C. E. Dept., October, 1975.

35. Grigoriu, M. D. and Rodriguez-Iturbe, I., "Prediction and Simulation with Autoregressive Models in Hydrology", to be published.

36. Guttman, I., Statistical Tolerance Regions: Classical and Bayesian, Griffin, London, 1970.

37. Guttman, I., "The Use of the Concept of a Future Observation in Goodness-of-fit Problems", J. R. Statist. Soc. B, 29, pp. 83-100.

38. Halpern, E., "Bayesian Polynomial Regression Analysis", The University of Michigan, Ph.D., 1970, Mathematics.

39. Halpern, E., Polynomial Regression from a Bayesian Approach, Journal of American Statistical Association, March 1973, Vol. 68, No. 341, pp. 137-143.

40. Hampton, J. M., Moore, P. G., and Thomas, H., "Subjective Probability and its Measurements", J. R. Statist. Soc. A, (1973), 136, Part 1, pp. 21-42.

41. Hasofer, A. M. and Lind, N. C., "An Exact and Invariant First-Order Reliability Format", Journal of the Engineering Mechanics Division, ASCE, Vol. 100, No. EM1, February, 1974, pp. 111-121.

42. Hasofer, A. M., "The Time-Dependent Maximum of Floor Live Loads", unpublished.

43. Heffes, H., "The Effect of Erroneous Models on Kalman Filter Response", IEEE Trans. A. C., July, 1966, pp. 541-543.

44. Ho, B. L., and Kalman, R. E., "Effective Construction of Linear State-Variable Models from Input/Output Functions", Regelungstechnik, Heft 12, pp. 545-592, 1966.

45. Hogart, R. M., "Cognitive Processes and the Assessment of Subjective Probability Distributions", Journal of the American Statistical Association, June 1975, Vol. 70, No. 350, pp. 271-294.

46. Hull, J., Moore, P. G., and Thomas, H., "Utility and its Measurement" J. R. Statist. Soc. A, (1973), 136, Part 2, pp. 226-247.

47. Jeffreys, H., Theory of Probability, Oxford, At the Clarendon Press, 1961.

48. Joint-Committee on Structural Safety CEB-CECM-CIB-FIP-IABSE, Sub-Committee for "First order reliability concepts for design codes". Minutes of Meeting held 21 and 22 May 1975, Helsinki.

49. Kailath, Th., "A General Likelihood-Ration Formula for Random Signals in Gaussian Noise", IEEE, Trans. IT, May 1969, pp. 350-361.

50. Kailath, Th., "Likelihood Ratios for Gaussian Processes", IEEE Trans. IT, May 1970, pp . 276-288.

51. Kailath, Th., "A Further Note on a General Likelihood Formula for Random Signals in Gaussian Noise", IEEE Trans. IT, July 1970, pp. 393-396.

52. Karlin, S. and Taylor, H. M., A First Course in Stochastic Processes Second edition, Academic Press, New York, 1975.

53. Kozielecki, J., "Psychological Characteristics of Probabilistic Inference", Acta Psychologica 34, 1970, pp. 480-488.

54. Lawless, J. F., "Approximations to Confidence Intervals for Parameters in the Extreme Value and Weibull Distributions", Biometrika (1974), 61, 1, pp. 123-129.

55. Lenton, R. L., Rodriguez-Iturbe, I., and Schaabe, J. C., "The Estimation of ρ in the First Order Autoregressive Model: A Bayesian Approach", Water Resources Research, Vol. 10, No. 2, April, 1974, pp. 227-241.

56. Lind, N. C., "Theory of Codified Structural Design", Notes pertaining to a set of invited lectures presented to the Academy of Science of Poland, June 5-15, 1972.

57. McGuire, R. K. and Cornell, C. A., "Live Load Effect in Office Buildings", Research Report R73-28, St 365, M.I.T., C. E. Dept., July, 1973.

58. Mihoc, Gh., Cincu, G., and Crain, V., Tearia Probabilitatilor si Statistica Matematica, Ed. Did. si Ped., Bucharest 1970.

59. Nash, R. A., D'Appolito, J. A., and Roy, K. J.,"Error Analysis of Hybrid Inertial Navigation Systems", AIAA Guidance and Control Conference, Paper No. 72-848, Stanford, California, August, 1972.

60. Nishimura, T., "Error Bounds of Continuous Kalman Filters and the Application to Orbit Determination Problems", IEEE Trans. A.C., Vol. AC-12, No. 3, June, 1967, pp. 268-275.

61. Paloheimo, E. and Hannus, M., "Structural Design Based on Weighted Fractiles", Journal of the Structural Division, ASCE, Vol. 100, No. ST7, July, 1974, pp.1367-1378.

62. Papoulis, A., Probability, Random Variables and Stochastic Processes, McGraw-Hill, 1965.

63. Parzen, E., Stochastic Processes, Holden Day, Inc., San Francisco, 1962.

64. Pearson, K., Tables of Incomplete Beta-Function, Cambridge University Press, 1934.

65. Peir, J. C., "A Stochastic Live Load Model for Buildings", Research Report, R71-35, M.I.T., C. E. Dept., September, 1971.

66. Raiffa, H. and Schlaifer, R., Applied Statistical Decision Theory, M.I.T.Press, 1961.

67. Ravindra, M. K., "Structural Code Optimization", University of Waterloo, Ph. D. Thesis, Civil Engineering, 1970.

68. Ravindra, M. K., and Lind, N. C., "Theory of Structural Code Optimization", Journal of the Structural Division, ASCE, Vol. 99, No. ST7, July, 1973, pp. 1541-1553.

69. Rice, S. O., "Mathematical Analysis of Random Noise", Bell System Technical Journal, Vol. 24, 1945, pp. 46-156.

70. Rockette, H.,Antle, C. and Klimko, L. A., "Maximum Likelihood Estimation with the Weibull Model", Journal of American Statistical Association, March, 1974, Vol. 69, No. 345, pp. 246-249.

71. Rosenbluth, E., "Code Specification of Safety and Serviceability", Technical Committee No. 10, Structural Safety and Probabilistic Methods, State of Art Report No. 2A., 1972.

72. Ross, S. M., Introduction to Probability Models, Academic Press, Inc. New York, 1972.

73. Sandi, H., "Loads and Load Combinations", State of Art Report No. 4, Technical Committee No. 10: Structural Safety Probability Methods, ASCE-IABSE Joint Committee, Planning and Design of Tall

Buildings, Fritz Eng. Lab., Lehigh University, Bethlehem, Pa., 1972.

74. Sandi, H., "Safety of Structures Subjected to Temporary Loadings", Rev. Roum. Sci. Techn. - Méc. Appl., Tome 11, No. 5, Bucharest, 1966, pp. 1123-1140.

75. Schweppe, F. C., "Evaluation of Likelihood Functions for Gaussian Signals", IEEE Trans. IT, January 1965, pp. 61-70.

76. Schweppe, F. C., Uncertain Dynamic Systems. Prentice-Hall, Inc., New Jersey, 1973.

77. Sandell, N. R. "6.261 Lectures", M.I.T., E. E. Dept.

78. Sidak, Z., "On Multivariate Normal Probabilities of Rectangles: Their Dependence on Correlations", The Annals of Mathematical Statistics, 1968, Vol. 39, No. 5, pp. 1425-1434.

79. Skyrms, B., An Introduction to Inductive Logic, Dickenson Publ. Co., Belmont, Ca., 1966.

80. Smallwood, R. D., "A Decision Analysis of Model Selection", IEEE Trans SSC, Vol. SSC-4, No. 3, September, 1968, pp. 333-342.

81. Thoman, D. R., Bain, L. J. and Antle, C. E., "Inferences on the Parameters of the Weibull Distribution", Technometrics, Vol. 11, No. 3, August, 1969, pp. 445-460.

82. Thoman, D. R., Bain, L. J., and Antle, C. E., "Maximum Likelihood Estimation, Exact Confidence Intervals for Reliability, and Tolerance Limits in the Weibull Distribution", Technometrics, Vol. 12, No. 2, May, 1970, pp. 363-371

83. Veneziano, D., "Contributions to Second Moment Reliability Theory", Research Report R74-33, M.I.T., April 1974.

84. Veneziano, D., "Statistical Estimation and Prediction in Probabilistic Models with Application to Structural Reliability", Ph.D. Thesis, C. E. Dept., M.I.T., Sept. 1974.

85. Vicens, G. J., Rodriguez-Iturbe, I., and Schaake, J. C., "A Bayesian Framework for the Use of Regional Information in Hydrology", submitted to Water Resources Research, 1974.

86. Vicens, G. J., Rodriguez-Iturbe, I., and Shaake, J. C., "A Bayesian Approach to Modeling of Hydrologic Time Series", Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Report No. 181, March, 1974.

87. Vicens, G. J., Rodriguez-Iturbe, I., and Schaake, J. C., "Bayesian Generation of Synthetic Streamflows", Water Resourses Research, Vol. 11, No. 6, Dec., 1975, pp. 827-838.

88. Wax, N., Selected Papers on Noise and Stochastic Processes, Dover Publications, Inc., New York, 1954.

89. Willems, J. C., "Minimal Realization in State Form from Input/Output Data", Mathematisch Institut Rijksuniversitet Groningen.

90. Winkler, R. L., "The Assessment of Prior Distributions in Bayesian Analysis", American Statistical Association Journal, Sept., 1967, pp. 776-800.

91. Wilks, S. S., Mathematical Statistics, John Wiley and Sons, New York, 1962.

92. Zellner, A., An Introduction to Bayesian Inference in Econometrics, John Wiley & Sons, Inc., New York, 1971.

93. Esteva, L., 1974, "Geology and Probability in the Assessment of Seismic Risk", Proc. Second International Congress of the Interna-

tional Association of Engineering Geologists, Sao Paulo.

94. Wood, E. F., "The Methodology of Bayesian Inference and Decision Making Applied to Extreme Hydrological Events", Ph.D.Thesis, C. E. Dept., M.I.T., January 1974.

BIOGRAPHY

The author was born in Bucharest, Romania, in 1943. He received the title of Engineer from the Institute of Civil Engineering of Bucharest in 1966 and then he worked for one year in a design institute. In February, 1967 he was appointed Assistant at the Institute of Civil Engineering of Bucharest. While holding this position he has studied at the Faculty of Mathematic-Mechanic of Bucharest (1968-1972) and in July 1972 he has received the title of Mathematician.

In February 1974 he attended the Massachusetts Institute of Technology as a graduate student for study toward a doctoral degree.