AD 614 697

DANIEL H. WAGNER, ASSOCIATES

Operations Research • Mathematics

PAOLI, PENNSYLVANIA

THEORY OF CUMULATIVE
DETECTION PROBABILITY

Report to

U. S. Navy Underwater Sound Laboratory

November 10, 1964

COPY 1 OF 1
HARD COPY $. 4.00
MICROFICHE $. 0.75

Under Contract No. N140(70024)74322B

by

DANIEL H. WAGNER, ASSOCIATES
Paoli, Pennsylvania

By Edward P. Loane
Edward P. Loane

Henry R. Richardson

Edward S. Boylan

Approved Daniel H. Wagner

DDC
APR 15 1965
TISIA B

PROCESSING COPY

ARCHIVE COPY

ABSTRACT

Cumulative detection probability, cdp, is defined as probability of at least one success in n trials. "Success" means that the (signal) stochastic process exceeds a given threshold. Exact formulas or approximations for cdp are given in the cases where the process being sampled in the trials is two-state Markov, Gaussian, "step," and "step-plus-jitter." In the two-state Markov case, taken largely from others, k-success formulas are also given. Finding cdp is equivalent to finding the distribution of the maximum of a sequence of random variables and to finding a cumulative multivariate distribution.

TABLE OF CONTENTS

ABSTRACT		
PREFACE		iii
SUMMARY		v
CHAPTER I	INTRODUCTION	1
	Background of the Problem	2
	Stochastic Process Model for cdp	3
CHAPTER II	TWO-STATE MARKOV PROCESSES	7
	Stationary Two-State Markov Processes	7
	Non-Stationary Two-State Markov Processes	17
	Continuous-Parameter Two-State Markov Processes	18
	Independent Processes	19
CHAPTER III	GAUSSIAN PROCESSES	25
	Approximation Method	25
	Error Estimates	33
CHAPTER IV	STEP PROCESSES	41
	Definitions	41
	Unimodal Threshold Theorems	43
	General Thresholds -- A Recursive Scheme	48
CHAPTER V	STEP-PLUS-JITTER PROCESSES	53
	Definitions	54
	Exact Recursive Formula for cdp	55
	The 2-Dependent Approximation	56
	The m-dependent Approximation	67
REFERENCES		71

TABLE OF CONTENTS (Continued)

APPENDIX A SOME FUNDAMENTALS OF STOCHASTIC PROCESSES

APPENDIX B LITERATURE SURVEY ON EVALUATING THE MULTIVARIATE NORMAL DISTRIBUTION

APPENDIX C CORRELATION THEOREMS

PREFACE

This is a report to the U. S. Navy Underwater Sound Laboratory under Contract Number N140(70024)74322B as part of an investigation of optimal search procedures for surface ship sonar. The report is related to other reports on the project as follows: A previous report, reference [a], presented stochastic process inferences from empirical sonar data; these inferred processes underlie most of the theoretical analysis in the present report. A forthcoming report will present results on optimal search procedures; the analysis will depend heavily upon some of the methods discussed herein.

We wish to acknowledge the splendid direction and cooperation we have received during this work from Mr. Carleton S. Walker of the System Planning Staff at USNUSL, and through him, from Mr. Stanley A. Peterson, Associate Technical Director for Systems Development. We further acknowledge the valuable assistance of our colleague, Mr. David C. Bossard, who reviewed previous work on two-state Markov processes and contributed Theorems II-1 and II-2 in Chapter II.

One of the authors, Professor Edward S. Boylan, has returned to Rutgers, The State University, after participating in this investigation during the summer of 1964.

SUMMARY

This report presents methods of computing cumulative detection probability (cdp), i.e., the probability of at least one successful detection in n trials. This is equivalent to the problem of computing the (cumulative) distribution of an n-dimensional random variable, and to the problem of computing the distribution of the maximum of an n-term sequence of random variables. In this summary we describe the stochastic process model used and the nature of the results for various types of processes.

Stochastic Process Detection Model

We fix a continuous-parameter stochastic (signal) process X^* (which associates to time t a random variable X_t^*) and a threshold function T^*. These determine a success-failure process D^*: Detection at time t means $X_t^* > T_t^*$ or equivalently, $D_t^* = 1$; no detection means $X_t^* \leq T_t^*$, i.e., $D_t^* = 0$. If X^* is sampled (observed) at $t_1, \ldots, t_n$, denote $X_{t_i}^* = X_i$, $T_{t_i}^* = T_i$, and $D_{t_i}^* = D_i$. Define

$$p_{t_i}^* = p_i = \Pr\{X_i > T_i\}$$

and

$$\text{cdp} = P_n = \Pr\{X_i > T_i \quad \text{for some } i,\ 1 \leq i \leq n\}.$$

The problem is to compute P_n under various assumptions on X^* or X. In Chapter I, the model is developed in more detail and background of the problem is given.

Two-State Processes

Two-state Markov processes are investigated in Chapter II. Attention is confined to the discrete-parameter success-failure process D (and briefly to D^*), ignoring

X^* and X. It is assumed that D is Markovian. Most of the results are taken from previous literature.

If D is stationary (i.e., independent of i, $p_i = p$, and $r_i = r$, the correlation between D_i and D_{i+1}), then D is characterized by a single 2 x 2 transition matrix (II-1). The transition probabilities are expressed in terms of p and r, and conversely. Formula (II-2) gives the probability of exactly k successes in n trials (for $k = 0, 1, 2, \ldots$) by means of a generating function. These are added to form probabilities of at least k successes (P_n if $k = 1$), explicitly given in formulas (II-3) for $k = 1$ and $k = 2$. (Other than in Chapter II, the report is restricted to $k = 1$.) Also given is the probability of at least two consecutive successes. These probabilities are presented graphically in Figures II-1 through II-6 for various r, n, p, and for $k = 1, 2$.

If D is non-stationary, then the situation for $k = 1$ is only slightly more complicated. The transition probabilities are given by formulas (II-4) in terms of p_i and r_i which now depend on i; P_n is given by (II-5).

When the number of observations n in a fixed interval (with D stationary) is made to approach infinity, $P_\infty = \lim_{n \to \infty} P_n$ is given by Theorem II-1. This represents cdp for a continuously observed process; P_∞ is graphed in Figure II-7.

For a sequence of independent trials, a generating function is given in Theorem II-2 for the probability of exactly k successes. This is applied explicitly to $k = 1, 2, 3$, and $k = 0$ is elementary. The formulas reduce to the binomial distribution if p_i is independent of i.

Gaussian Processes

If X^* is Gaussian, then computation of P_n is equivalent to computation of a multivariate normal distribution. This is investigated in Chapter III--previous literature is reviewed in Appendix A.

The principal method offered in this case is a convenient approximation based on the assumption (used in Chapter II) that the discrete-parameter success-failure process D is Markovian. Computation aids (constructed from bivariate normal tables) are presented in Figures III-1 through III-4, and application is illustrated by a tabulated example.

For the purpose of investigating the error in this approximation to P_n, Theorem III-1 reduces the problem for $n = 3$ to a one-dimensional integration, under the assumption that (X_1, X_2, X_3) is Markovian as well as Gaussian; formula (III-2)

gives a corresponding result for $n=5$, requiring bivariate rather than univariate normal tables for inputs to the integration. Theorem III-2 shows that under the assumption that X is Markovian, the approximation to P_3 is no less than the true P_3, and an exact formula is given for the error. A bound on this error expression is given by Theorem III-3.

Examples of the error in the approximation are presented in Table III-1, primarily for $n=3$ or 5 and X Markovian. The errors in P_n are of the order of .01 or less. Under these conditions, the approximation appears to be very good, and the errors do not appear to grow substantially as n increases from 3 to 5. Neither stationarity in X nor constant T appear important for accuracy. However, in a special case where X is not Markovian, the approximation is rated fair at best.

Step Process

A step process X^* is one in which each realization is a step function, with the jumps occurring in a Poisson process, and with the sampled values before and after a jump being independent draws from a fixed distribution. For this type of process, cdp is investigated in Chapter IV.

An exact formula for P_n is given in Theorem IV-1 in the "unimodal" case, wherein the threshold is non-increasing prior to a minimum point and is non-decreasing thereafter. The corollary gives a neat form of this formula for the case where the observations are uniformly spaced. The latter formula is used in Theorem IV-2 to pass to a limit to obtain P_∞ -- this includes the constant-threshold case of Theorem II-1 as a special case. The monotonic case of this theorem was known previously.

Without the assumption of unimodal thresholds, convenient exact formulas have not been found in closed form; however, an exact recursive procedure is given for P_n. The required number of multiplications is in the order of n^2, and thus the recursive procedure appears more efficient than a straightforward approach, which would require over 2^{n-1} multiplications.

Step-Plus-Jitter Process

A step-plus-jitter process X^* is the sum of a step process Z^* and a jitter process J^*; it is assumed that J^* is independent of Z^* and of itself. The cdp problem for X^* of this type is investigated in Chapter V.

The assumption of unimodal thresholds is no longer helpful. An exact (but laborious) method of computing P_n is given by recursive formulas in the same vein as for the pure step process of Chapter IV.

An approximation to cdp is derived by means of an "m-independent" process, wherein each observation depends only on the outcomes of the preceding m trials. The parameters of the m-independent process are chosen to make it resemble the step-plus-jitter process.

This is applied to obtain an approximation to P_n, by means of a recursive procedure (which requires much less computation than the exact procedure). The accuracy of this approximation is illustrated for $m=2$ in Figures V-1, V-2, and V-3. Applied to a pure step process, the approximation is very accurate. When jitter is present (of a particular uniformly-distributed type), the approximation appears to be good, providing there is a high probability of jumps between observations; if this probability is not much higher than 1/3 (the lowest value for which the approximation can be used with $m=2$), then the accuracy is not as good, but it is still much better than an estimate based on an independence assumption.

Miscellaneous

Appendix A surveys various fundamentals relevant to stochastic processes. Definitions and elementary properties are given for n-dimensional random variables and their distributions, correlation, stochastic processes with emphasis on Markov processes, and related concepts.

Previous publications on computing the multivariate Gaussian distribution are reviewed in Appendix B. Surprisingly little appears to be known on the problem. Formulas for Gaussian cdp are cited in some very special cases, which do not appear to be of much practical interest for present purposes.

Appendix C presents some statements on correlation. Theorem C-1 gives the autocorrelation of the sum of two stochastic processes. Theorem C-2 gives some basic facts on correlation between two two-state random variables. Theorem C-3 gives an inequality comparing the correlation between two Gaussian variables with the correlation between the corresponding two success-failure variables obtained by thresholding; the latter correlation is smaller. Theorem C-4 shows that the autocorrelation of a step process equals that of the associated success-failure process obtained by thresholding; both autocorrelations are exponential decay.

THEORY OF CUMULATIVE DETECTION PROBABILITY

CHAPTER I

INTRODUCTION

This report is addressed to the theoretical aspects of the problem of computing cumulative detection probability, cdp, i.e., the probability of at least one success in a sequence of detection trials. This problem has arisen in several ways during an investigation of optimal sonar search procedures, conducted for the Underwater Sound Laboratory. Common practice has been to treat successive detection trials (e.g., pings) as being statistically independent. One quickly sees that when this assumption is not satisfied (as is usually the case), then absolute estimates of cdp can be greatly erroneous. Moreover, relative comparisons and selections of optimal search parameters may also be substantially in error, although this may not be as obvious. The purpose of the report is to provide theoretical tools which can be used to avoid such erroneous assumptions of independence.

The substance of this report is largely mathematical in character, and will be presented in the language of probability theory. To find this report usable, the reader will need facility with elementary probability theory, and will need some knowledge of stochastic process concepts. In particular, an understanding of the elementary fundamentals of a Markov process (or Markov chain) will be necessary. Appendix A surveys the most relevant fundamentals, with emphasis on Markov processes, but, of course, it is not a substitute for a text.

It is useful to bear in mind that the problem of computing cdp is equivalent to the problem of computing the distribution of the maximum of a sequence of random variables, and to the problem of computing a (cumulative) multivariate distribution. These equivalences assist in bringing results from the literature to bear on the cdp problem, and may in turn enable cdp research to contribute to these equivalent problems. It is surprising how little is known, apparently, about computing the multivariate <u>normal</u> distribution, for example.

In the remainder of this introduction we review the historical background of the cdp problem, and relate the detection problem to a basic mathematical model: thresholding on a stochastic process. Thereafter, problems of detection per se

arise only in abstraction. The four succeeding chapters are addressed to finding cdp when the underlying processes are respectively two-state Markov, Gaussian, "step," and "step-plus-jitter." The appendices survey fundamentals of stochastic processes, review the literature on multivariate normal distributions, and present some theorems on correlation which are useful for cdp purposes.

Background of the Problem

The historical background of the problem of computing cdp stems from early work on search theory, envisioning radar and sonar detection, notably various OEG reports, and derived publications, by B. O. Koopman. We particularly cite references [b], [c], [d], [e], and [f]. In general, the problem of computing cdp has been one in which both empirical knowledge and theoretical tools have been inadequate.

The first three of these references are addressed to general theory of search, while reference [f] is an analysis of sonar data. Among other things, they develop sweep widths as areas under lateral range curves. Each point on a lateral range curve is a cdp, viz., the probability that a target passing at given lateral range will be detected at some point during the pass or, alternatively, at some point prior to reaching closest approach. The possibility that fluctuations in conditions affecting detection will occur during a pass is recognized, but is generally not treated explicitly in references [b], [c], [d], and [f]; where glimpses are accumulated, they are done on the basis of glimpse-to-glimpse independence. In some cases, adjustment is made via a curve-fitting parameter (in fitting a theoretical model to empirical data), the adjustment having an effect of partially compensating for the error in the independence assumption.

Reference [e] by Koopman is a theoretical treatment of cumulative success in a sequence of correlated trials -- the basic model used is that success-failure as a two-state process is a Markov chain (Koopman has also developed this model in greater depth in reference [g], among other papers). Some of the results will be included here for completeness. Moreover, a useful approximation akin to this model is presented in Chapter III.

More recent progress in theoretical tools has been offered by references [h] and [i]. Reference [h], Appendix B, presented a model for time fluctuations of acoustics in sonar detection (presumed to be primarily propagation loss) known as a "(λ, σ)-process." This was applied to cumulative detection and approach problems in Appendix C of the same reference (these Appendices B and C were due to J. D. Kettelle Jr.). For purposes of incorporating detection capability in cumulative

fashion in simulation of undersea warfare, APL has used a random walk model described in reference [i]. This is a special type of Markov chain in which one may transition only to an adjacent state, from among a finite set of states. Evidently sea state is the fluctuating parameter primarily envisioned. Suitably adapted to operational data, the models of references [e], [h], and [i] could be useful improvements over the assumption of independent glimpses.

More recent progress has also been made in empirical knowledge of sequential correlation effects in detection parameters, although the situation still leaves a great deal to be desired. References [j] and [k] are careful analyses respectively of closely spaced radar observations on an air target and closely spaced sonar propagation losses in convergence-zone paths. Reference [a] is a report under the present project which presents statistical analyses of BRASS II sonar data, emphasizing sequential correlation behavior in bottom loss and echo-to-reverberation ratio. Of the existing gaps in empirical knowledge of the sonar aspects of the problem, human operator effects appear to be by far the most important (recognized a decade ago in reference [f]). However, knowledge of correlation behavior in propagation loss and background interference certainly has room for improvement; in particular, there are no known data in this vein on direct-path sonar propagation loss.

It is the aim of the present report to contribute theoretical tools to computation of cdp. The methods developed are motivated largely by the models used to describe the empirical sonar data in reference [a]. In reference [a] it was inferred that under certain circumstances, the signal process could be of a Gaussian nature and in other circumstances "step-plus-jitter." The step-plus-jitter process has as a special case a step process, of which the (λ, σ)-process mentioned above is an example. From this derives the interest in the types of processes treated in Chapters III, IV, and V. We note again that reference [a] does not include operator effects -- when these are included, the signal process is highly conjectural, and the simpler two-state Markov processes considered in Chapter II may be as plausible as the others.

Stochastic Process Model for cdp

We suppose that a target is moving on an arbitrary track relative to a sensor device being used to detect the target. The sensor operates in discrete glimpses. The separation between glimpses could arise, for example, from the fact that the sensor is an active sonar and must wait for sound travel between pings, or from the sensor being either a sonar or a radar transmitting directionally in sweeps, illuminating the target at most once each sweep. Passive sonar detection would lead to a somewhat different model (glimpses would be overlapping time "windows"), but it is possible that results in terms of discrete glimpses as developed herein might also be useful in passive problems.

The outcome of the ith glimpse is a random variable X_i, regarded as signal and often measured in decibels. This random variable may take on values from a continuum of real numbers and possesses a probability distribution over these values. The sequence of glimpses thus gives rise to a sequence $(X_1, X_2, \ldots)$ of random variables, and this sequence is termed a "stochastic process" with a "discrete parameter." This term is defined more explicitly in Appendix A (and more generally and abstractly in textbooks referenced there).

It is convenient to consider that, more generally, a glimpse is potentially available to the sensor at any time and to denote by X_t^* the signal random variable which would result from a glimpse at time t. This determines a stochastic process X^* with "continuous parameter"; the process associates with each non-negative number t, a random variable X_t^*. In practice, X^* will be sampled at discrete times $t_1, \ldots, t_n$, giving rise to a discrete-parameter process $(X_{t_1}^*, \ldots, X_{t_n}^*)$, of the sort mentioned in the preceding paragraph, which we denote $(X_1, \ldots, X_n)$ for brevity. In the body of the report, asterisks will distinguish continuous-parameter terminology from discrete-parameter terminology -- we usually deal with the latter. In the appendices, this distinction will not be made.

One usually thinks of the parameter of the signal process X^* as being time (usually denoted t), although it could just as well be distance. This matter does not concern us in the present report. We note that in the stochastic process inferences from empirical data in reference [a], one could not discern from the data whether the processes were time-dependent or distance-dependent.

To relate the signal random variable to detection, we introduce a detection threshold function T^*. If X^* is sampled (observed) at the point t, then detection ensues if and only if the sampled value of X_t^* exceeds T_t^*. This gives rise to a new, two-state, process D^* which is called the success-failure process:

$$D_t^* = \begin{Bmatrix} 0 & \text{if } X_t^* \leq T_t^* \\ 1 & \text{if } X_t^* > T_t^* \end{Bmatrix} .$$

A success (detection) occurs whenever $D_t^* = 1$, and a failure occurs whenever $D_t^* = 0$. We define p_t^*, the single-trial success (detection) probability (also called unconditional success probability) for a trial (glimpse) at the point t, by

$$p_t^* = \Pr \{ D_t^* = 1 \}$$

$$= \Pr \{ X_t^* > T_t^* \} .$$

(The operator Pr is read, "the probability of the event.")

Suppose again that the continuous-parameter process X^* is sampled at $t_1, \ldots, t_n$. Then we denote $T^*_{t_i}$, $D^*_{t_i}$, and $p^*_{t_i}$ by T_i, D_i, and p_i respectively for $i = 1, \ldots, n$. We now define P_n, the <u>cumulative detection probability</u> (abbreviated "cdp") in the n trials by

$$P_n = \Pr\{D_i = 1 \quad \text{for some } i,\ 1 \le i \le n\}$$

$$= \Pr\{X_i > T_i \text{ for some } i,\ 1 \le i \le n\}.$$

Computing P_n is the problem to which this report is addressed. In two cases, Theorems II-1 and IV-2, we will treat an extension of P_n applying directly to the continuous-parameter process X^*; this cdp is, for a given interval $[0, u]$,

$$P^*_\infty = \Pr\{X^*_t > T^*_t \quad \text{for some } t,\ 0 \le t \le u\}$$

$$= 1 - \Pr\{\sup_{0 \le t \le u} (X^*_t - T^*_t) \le 0\}.$$

There is a certain amount of arbitrariness in setting the detection threshold function T^* -- different thresholds might be used for different reference levels in the signal process X^*, and the reference levels need not be constant. The most convenient choice in this regard would be to incorporate into T^* the <u>deterministic</u> (i.e., systematic) variations during an encounter, while X^* reflects the <u>random</u> variations about a constant mean, i.e., the mean of the random variable X^* is the same for all t. In this case, it is plausible that the process X^* will be "stationary" in the sense defined in Appendix A, in addition to having a stationary mean.

As an example of variability in thresholds, suppose the target is passing on a straight track with a stated closest approach to the sensor. During the sequence of glimpses, the range will decrease monotonically, will reach a minimum, and then will increase monotonically thereafter. Suppose that parameters other than range (e.g., environmental and operator effects) which affect detection have the same statistical behavior on each glimpse, on each pair of consecutive glimpses, etc.

These effects other than range can be conveniently reflected in a stationary stochastic process X^*, while the deterministic effects of range are reflected in the detection threshold T^*, decreasing with range. This example has a particular type of threshold variation which is useful in Chapter IV: the detection function is referred to as "unimodal" (although it is actually the single-trial success probability p_t^* which increases to a maximum and decreases thereafter). Another example of systematic effect on detection, for some detection systems, is the effect of target aspect (plausibly again unimodal).

Throughout the body of the report, we fix a continuous-parameter stochastic process X^*, a threshold function T^*, a single-trial success probability function p^*, and the associated two-state success-failure process D^*, all as defined in general terms above. In Chapters III, IV, and V, the process X^* is specialized to be respectively Gaussian, "step," and "step-plus-jitter," to be defined later. Chapter II does not deal with X^* explicitly, but assumes that the two-state process D^* is Markovian. In all cases, X^* is sampled (i.e., observed or trials are made) at $t_1, \ldots, t_n$, giving rise to the discrete-parameter quantities X_i, T_i, D_i, and p_i, for $i = 1, \ldots, n$. The central problem is to find the corresponding P_n.

We note two conventions: If q is a probability, then $1 - q$ will be denoted by $\bar{q}$. The end of a proof is marked: #####

CHAPTER II

TWO-STATE MARKOV PROCESSES

In this chapter we assume that D^*, the success-failure process being sampled, is a two-state Markov process, and we present formulas for cdp in this case. In practice, the two-state nature of the process would presumably arise by thresholding on a multi-state process X^*, as described in Chapter I. However, the source of the two-state process does not concern us in this chapter.

A two-state Markov process is possibly the simplest type of non-independent process, for cdp purposes. For this reason, it will be a convenient type of an assumption in practical applications. Moreover, the principal technique given in Chapter III for computing cdp in Gaussian processes, will be an approximation which is closely related to cdp in two-state Markov processes.

In this chapter, we will be able to give formulas for the probability of at least k successes, and exactly k successes, in n trials. Elsewhere in the report, we are restricted to $k=1$.

The first section, which contains most of the discussion, is addressed to the stationary case, wherein the single-trial probability of success is constant. The non-stationary case is discussed briefly in the second section. Limiting values of cdp, as the number of trials in a given interval becomes infinite, are given in the third section. The final section reviews independent processes which may be regarded as a special case of Markov processes.

Most of the results of this chapter are taken from Koopman, reference [e], and Thiess, reference [l]. Theorems II-1 and II-2 were supplied by our colleague, D. C. Bossard.

Stationary Two-State Markov Processes

In this section we investigate the simplest type of two-state Markov process, in which p, the single-trial probability of success, does not change from trial to trial. Correlation between successive trials also does not change. This is a stationary process. (See Appendix A for definitions of these terms.)

The process in this case is characterized by a matrix of four transition probabilities given as follows, each entry being the probability that if the jth trial results in the state of the row heading, then the $j+1^{st}$ trial results in the column heading:

$$\begin{array}{c} \\ \text{Success} \\ \text{Failure} \end{array} \begin{array}{c} \begin{array}{cc} \text{Success} & \text{Failure} \end{array} \\ \begin{pmatrix} a & \bar{a} \\ b & \bar{b} \end{pmatrix} \end{array} \qquad \text{(II-1)}$$

The assumption that the process is Markovian means simply, by definition, that these probabilities do not depend on outcomes of trials <u>prior</u> to the jth trial. The assumption of stationarity means simply that these probabilities are independent of j. Processes of this type, as well as non-stationary cases, are explored in detail by Koopman in references [e] and [g]. Additional discussion of Markov processes is given in Appendix A.

Let r be the correlation coefficient (see Appendix A and Theorem C-2 of Appendix C) between successive trials, and p be the probability that a random trial results in success. We then have (reference [e]):

$$r = a - b$$

$$a = p + \bar{p}\, r$$

$$b = p - p\, r.$$

We thus see that p and r suffice to describe the situation; note, however, that we must have $r \geq -p/\bar{p}$ and $r \geq -\bar{p}/p$, since a and b are probabilities. The correlation coefficient between trials h units apart is r^h.

Let $R_n^{(k)}$ be the probability of <u>exactly</u> k successes in n trials. Let $P_n^{(k)}$ be the probability of <u>at least</u> k successes in n trials -- we are primarily interested in k=1, and k=1 will be understood if the superscript is omitted. Then $R_n^{(0)}$ is the probability of n consecutive failures:

$$R_n^{(0)} = \bar{p}\, \bar{b}^{\,n-1}.$$

For $k \neq 0$, $R_n^{(k)}$ is given by (see Thiess, reference [1]):

$$R_n^{(k)} = \sum_{m=k}^{n-1} \nu_m^{(k)} \bar{a} \bar{b}^{n-m-1} + \nu_n^{(k)} ,$$

where the $\nu_m^{(k)}$'s are found from

$$x \frac{bx+p(1-x)}{1-\bar{b}x} \left(ax + \frac{\bar{a}bx^2}{1-\bar{b}x}\right)^{k-1} = \sum_{m=k}^{\infty} \nu_m^{(k)} x^m . \qquad \text{(II-2)}$$

The left member as a function of x is a "generating function" for the probabilities $R_n^{(k)}$. Koopman gives an alternative generating function formulation as formula (20), Section 5, of reference [e]. Formula (II-2) is slightly more general than the corresponding expression in reference [1], since it allows for a probability of p that the first event will be a success rather than requiring that the first event be a failure. We quote the results for $k=0$ and $k=1$:

$$R_n^{(0)} = \bar{p} \bar{b}^{n-1}$$

$$R_n^{(1)} = p \bar{a} \bar{b}^{n-2} + (n-2) \bar{p} \bar{a} b \bar{b}^{n-3} + \bar{p} b \bar{b}^{n-2} .$$

From these we can calculate the $P_n^{(1)}$ and $P_n^{(2)}$:

$$P_n^{(1)} = 1 - R_n^{(0)}$$

$$P_n^{(2)} = 1 - R_n^{(0)} - R_n^{(1)} . \qquad \text{(II-3)}$$

We can also express these probabilities in terms of r:

$$R_n^{(0)} = \bar{p}\,(\bar{p} + p\,r)^{n-1}$$

$$R_n^{(1)} = p\,\bar{p}\,(1-r)\,(\bar{p} + p\,r)^{n-3}\,(2r + n\bar{p}\,(1-r)).$$

If $r = 0$, then $R_n^{(k)}$ reduces to the familiar binomial form (see last section).

Reference [1] also gives the probability $S_n^{(2)}$ of receiving two consecutive successes under the same conditions as above:

$$S_n^{(2)} = p\,a + \bar{p}\,\left(1 - \frac{\xi^{n-1}\,(\xi + b) - \zeta^{n-1}\,(\zeta + b)}{\xi - \zeta}\right),$$

where

$$\zeta = \frac{1}{2}\,(\bar{b} - \sqrt{\bar{b}^2 + 4\,\bar{a}\,b})$$

$$\xi = \frac{1}{2}\,(\bar{b} + \sqrt{\bar{b}^2 + 4\,\bar{a}\,b}).$$

Figures II-1 ($p = 0.3$), II-2 ($p = 0.5$) and II-3 ($p = 0.8$) give the probabilities $P_n^{(1)}$ and $P_n^{(2)}$ for $n = 2$, 3, 4, and 5. In Figure II-1, $S_4^{(2)}$ is shown by a dashed line; of course, $S_2^{(2)} = P_2^{(2)}$.

We note from Figure II-1 that the effect of correlation increases with the number of trials (as expected) and that it increases as the unconditional probability, p, decreases. The effect on the probability $P_4^{(1)}$ of at least one success in 4 trials is opposite to the effect on $P_4^{(2)}$ and $S_4^{(2)}$, where at least two and at least two successive detections are respectively required; the former probability decreases as r increases, while the latter probabilities generally increase, reaching a maximum for $r < 1$ in some cases.

The effect of correlation on $P_n = P_n^{(1)}$ is shown in a different form in Figures II-4, II-5, and II-6. For various values of p, the number of trials required to yield various levels of P_n versus the correlation coefficient r is plotted. One notes that the number of trials required increases monotonically with r, the increase becoming very rapid as r approaches unity.

FIGURE II-1

CUMULATIVE PROBABILITIES IN STATIONARY TWO-STATE MARKOV PROCESSES

(p = 0.3)

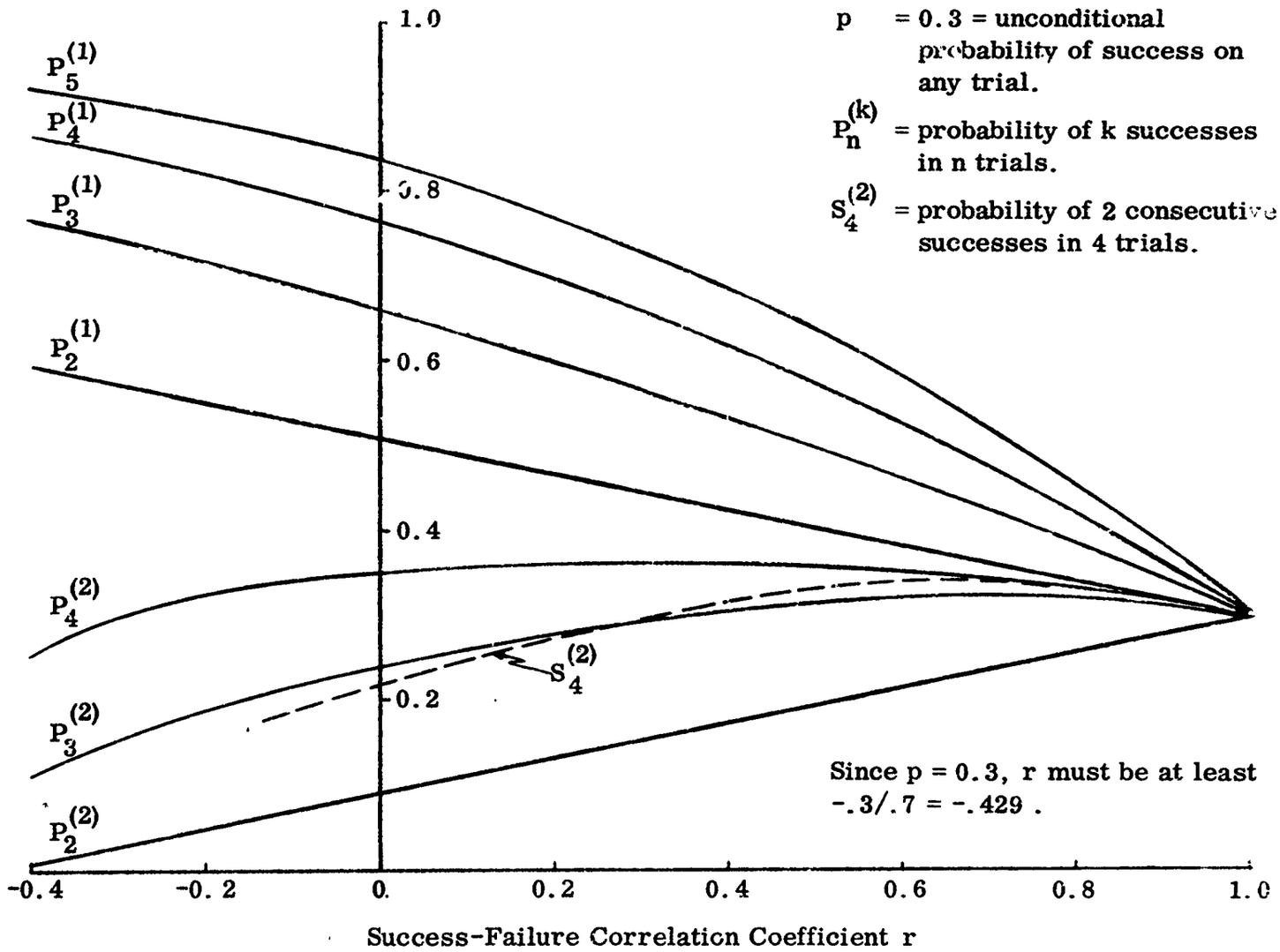


FIGURE II-2

CUMULATIVE PROBABILITIES IN STATIONARY TWO-STATE MARKOV PROCESSES

(p = 0.5)

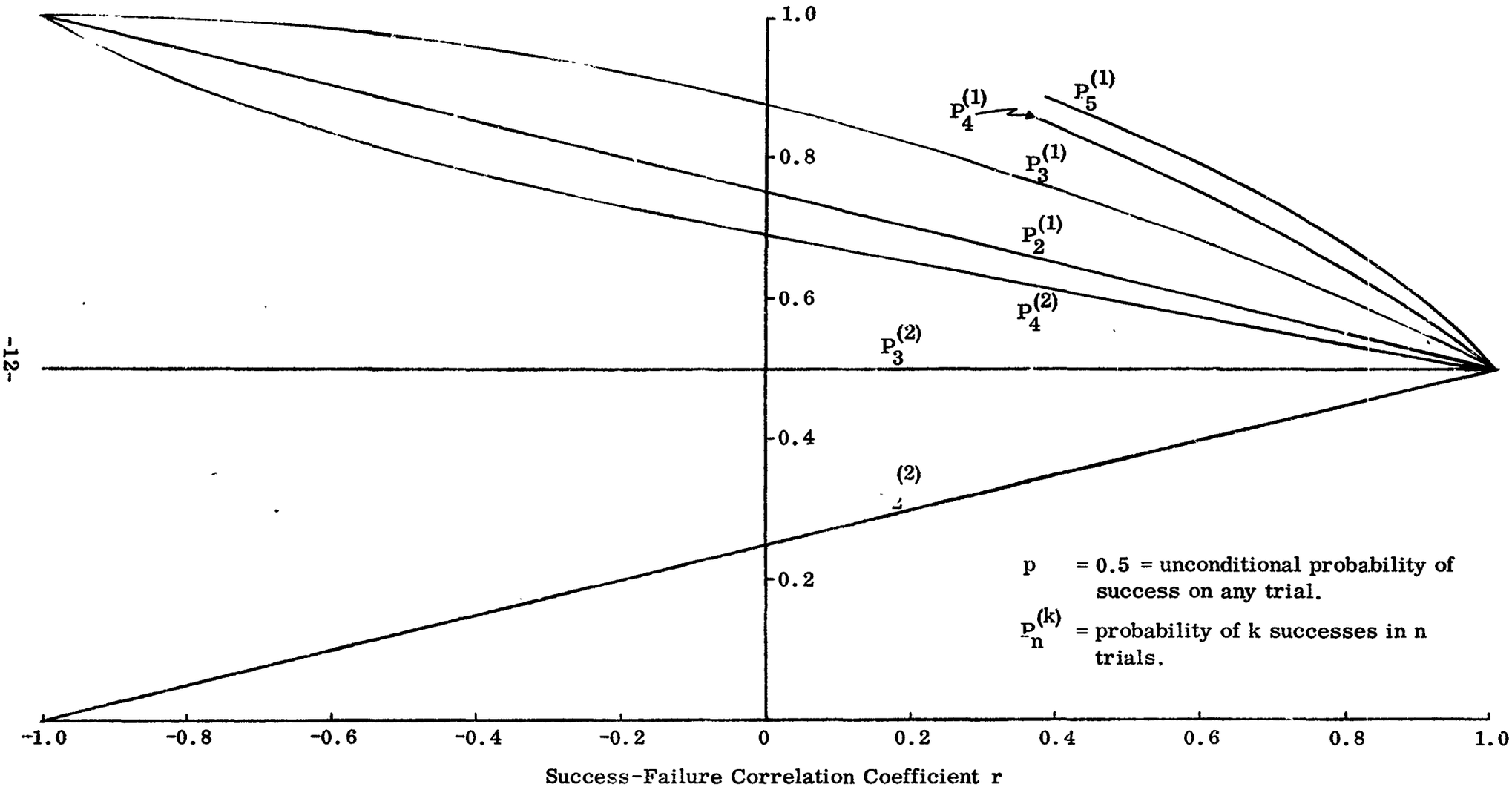


FIGURE II-3

CUMULATIVE PROBABILITIES IN STATIONARY TWO-STATE MARKOV PROCESSES

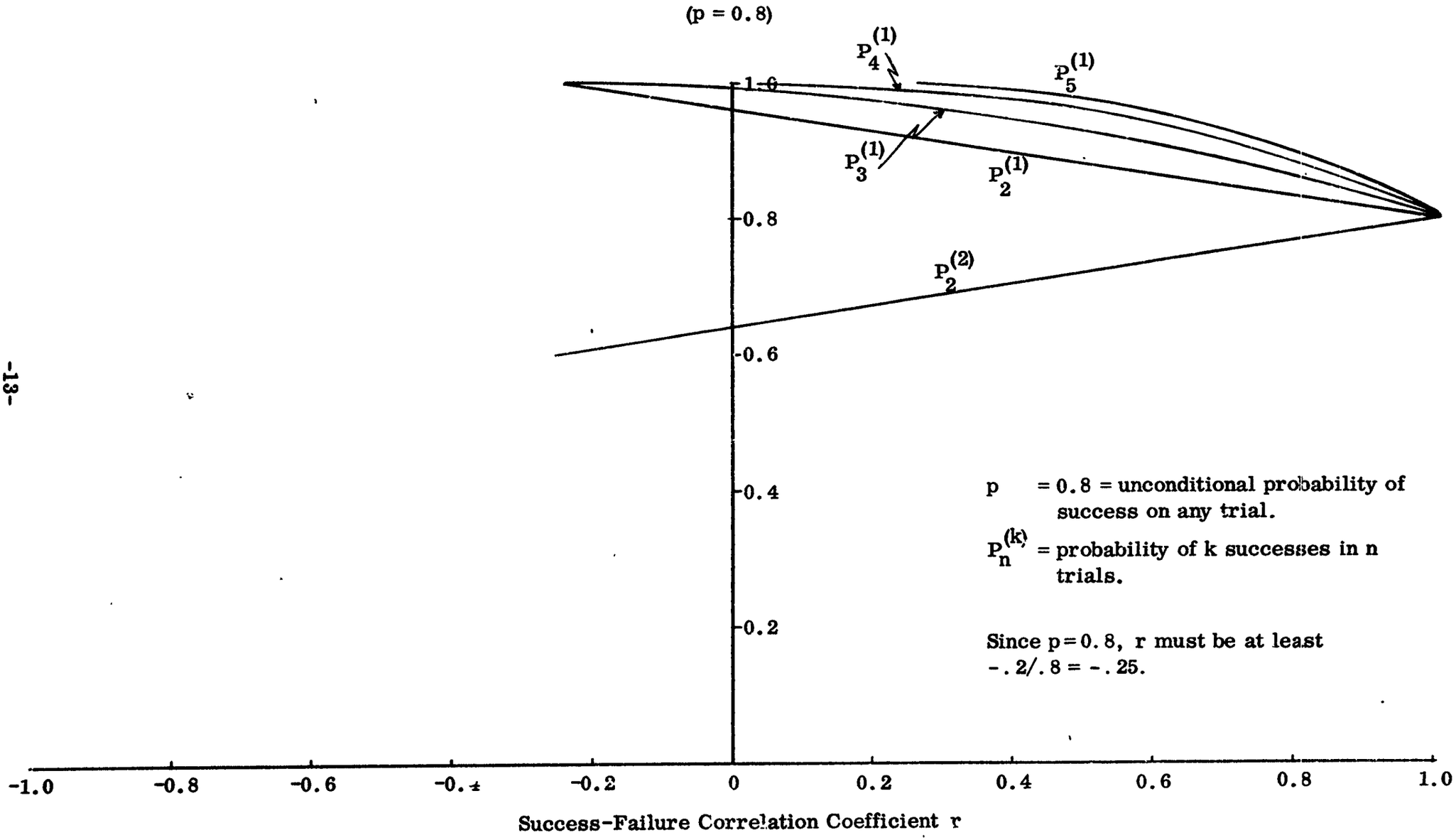


FIGURE II-4

NUMBER OF TRIALS REQUIRED FOR GIVEN cdp

($p = 0.3$)

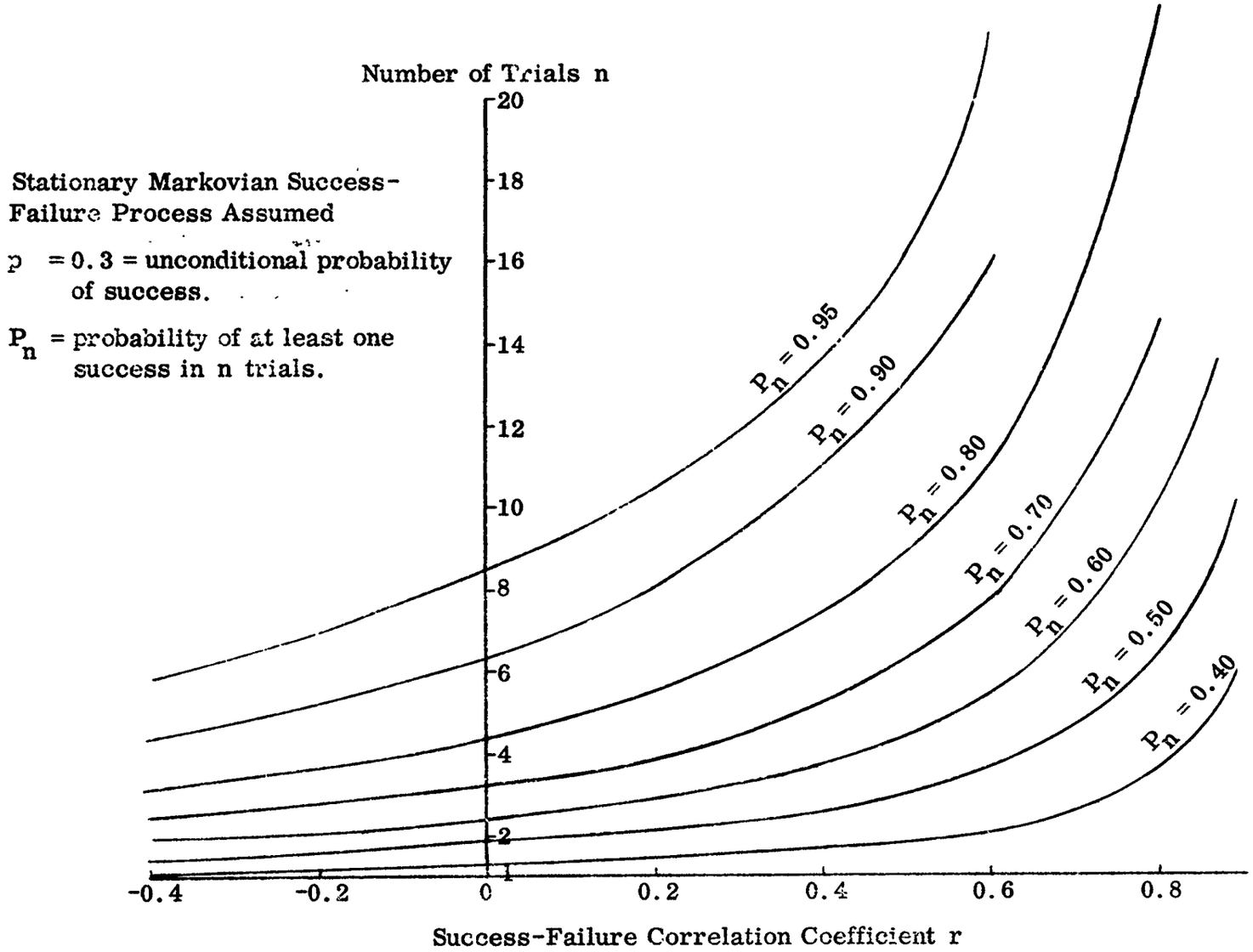


FIGURE II-5

NUMBER OF TRIALS REQUIRED FOR GIVEN cdp

(p = 0.5)

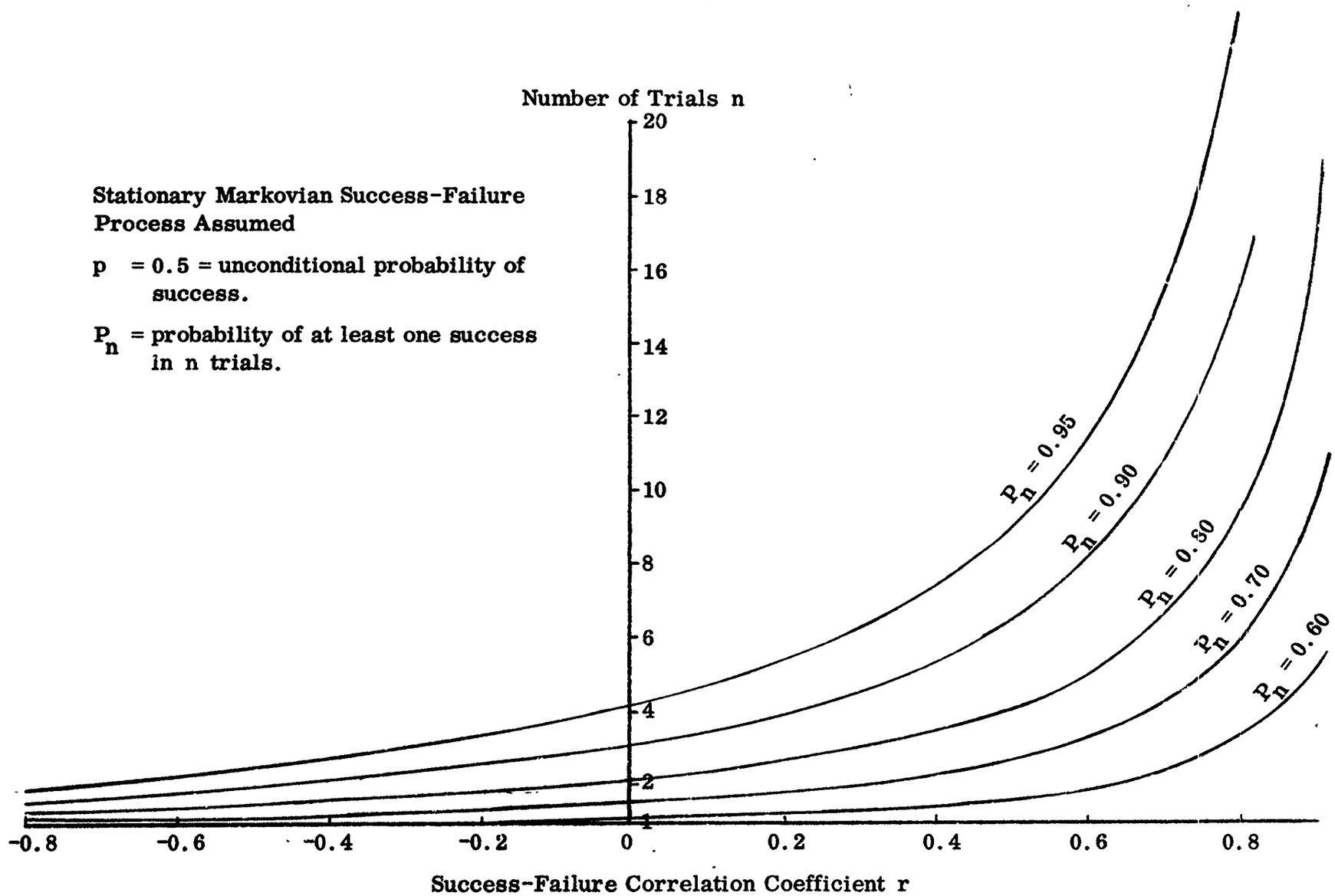
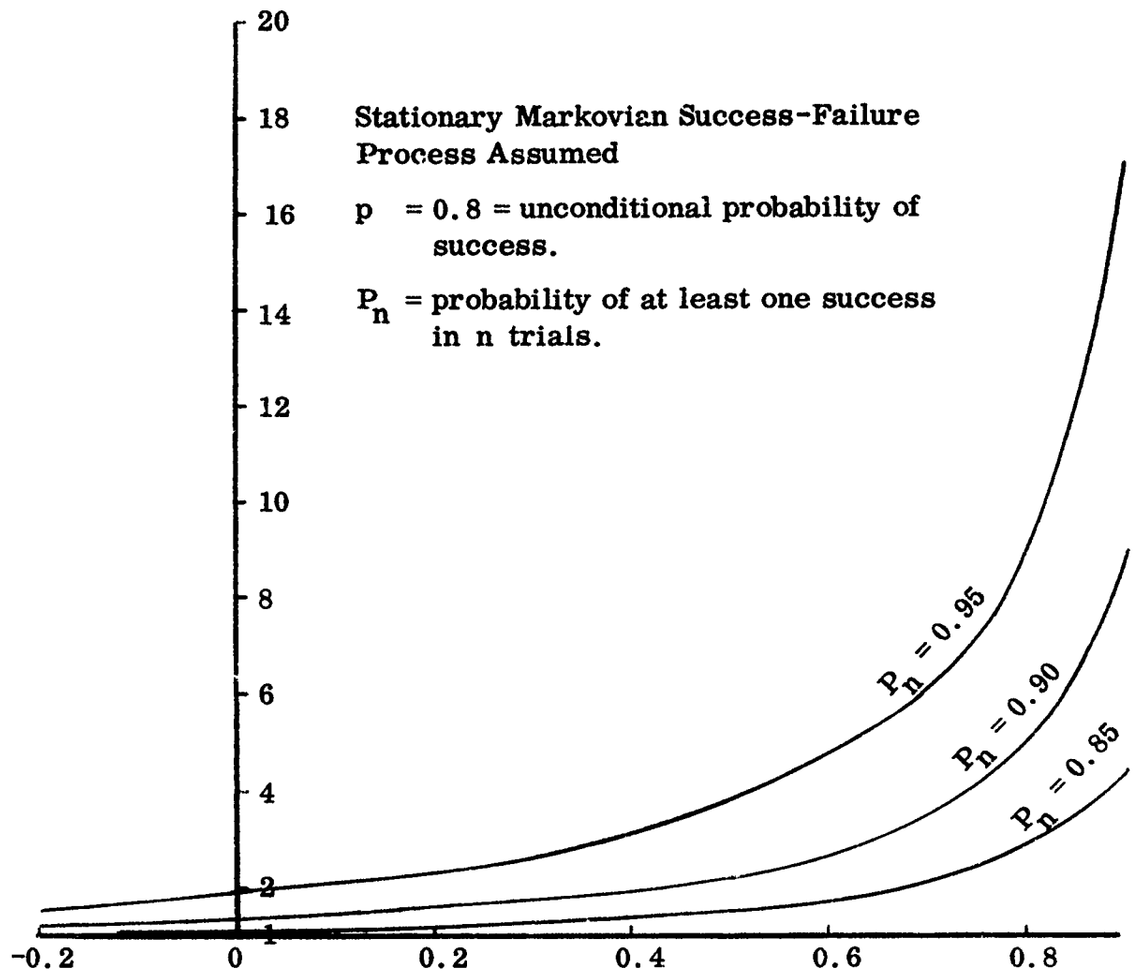


FIGURE II-6

NUMBER OF TRIALS REQUIRED FOR GIVEN cdp

(p = 0.8)



Non-Stationary Two-State Markov Processes

In the non-stationary case, the quantities a, b, p, and r are permitted to change from trial to trial, and are subscripted by the trial number accordingly. One generalization is that the correlation coefficient between the jth trial and the $(j+h)^{th}$ trial is

$$\prod_{i=j}^{j+h-1} r_i,$$

where r_i is the correlation between the ith and $(i+1)^{st}$ trial.

The transition probabilities on the jth trial are given by (page 5, reference [e]):

$$a_j = p_j + r_j \bar{p}_{j-1} \sqrt{\frac{p_j \bar{p}_j}{p_{j-1} \bar{p}_{j-1}}}$$

and (II-4)

$$b_j = p_j - r_j p_{j-1} \sqrt{\frac{p_j \bar{p}_j}{p_{j-1} \bar{p}_{j-1}}}.$$

The cdp is given by

$$P_n = P_n(1) = 1 - \bar{p}_1 \prod_{j=2}^{n} \bar{b}_j. \qquad \text{(II-5)}$$

We shall make use of formula (II-5) in devising an approximation to cdp in a Gaussian process in Chapter III.

Reference [e] gives difference equations for the probability of exactly k successes in n trials in this non-stationary case.

Continuous-Parameter Two-State Markov Processes

In practice we customarily achieve a sequence of trials, described previously as a two-state Markov process, by sequentially sampling a continuous process. (The discrete-parameter process is usually called a Markov chain.) We now state some elementary facts regarding the continuous version of the process, confining attention to the stationary case. For present purposes, we define a continuous-parameter two-state process D to be stationary Markov if every discrete-parameter process obtained by sampling D at uniform points is stationary Markov.

It has been noted that the correlation between trials separated by h is r^h, where r is the correlation between successive trials. The corresponding statement in a continuous-parameter process is that the autocorrelation function associated with the process is of the form $e^{-\lambda h}$, for some fixed $\lambda > 0$.

Suppose that D is a continuous stationary two-state Markov process, and p is the unconditional probability of success. Let $\lambda > 0$ be as above. Suppose that D is sampled at the $n+1$ points, it/n for $i = 0, \ldots, n$, uniformly spaced in the interval $[0, t]$. Then the limiting value of P_n is given by the following theorem:

Theorem II-1. Under the above conditions,

$$P_\infty = \lim_{n \to \infty} P_n = 1 - \bar{p}\, e^{-\lambda p t}.$$

Proof. From formula (II-3) and the given autocorrelation function, we have

$$1 - P_n = \bar{p}\, [\bar{p} + p\, e^{-\lambda t/n}]^{n-1}.$$

The proof is completed by letting n approach infinity.*

#####

* Throughout the report, the end of a proof will be denoted by: #####.

Theorem II-1 is a special case of a more general result given in Chapter IV as Theorem IV-2.

A comparison of this limiting case, i.e., when the number of trials approaches infinity, with cases for a finite number of trials is illustrated in Figure II-7. One notes that in general, with 10 or more trials, cdp is not much smaller than the limiting cdp ($n=\infty$), and hence this formula for the limiting cdp may be a useful approximation when the number of trials is fairly large.

A different type of limiting situation is treated by Koopman in reference [g]: the sequence of trial sequences is not contained in a fixed interval. Under various conditions on limits, generating functions and explicit formulas are developed for the probability of exactly k successes in the limiting case.

Independent Processes

We conclude this chapter with a brief discussion of independent processes (which may be regarded as a special type of Markov process).

A stochastic process X is independent if X_i and X_j are independent random variables whenever $i \neq j$. Since there are no correlation problems in independent processes, there is no loss of generality, for cdp purposes, in confining attention to two-state success-failure versions of independent processes, i.e., the D process in the notation of Chapter I. A two-state Markov process specializes to one of this type if $a_j = b_j$ in formulas (II-4).

Let p_i be the unconditional probability of success in the ith trial of an independent process, for $i=1,\ldots,n$. Let $P_n^{(k)}$ be the probability of at least k successes in the n trials, and let $R_n^{(k)}$ be the probability of exactly k successes. Then

$$1 - P_n^{(1)} = \prod_{i=1}^{n} (1-p_i) = R_n^{(0)}. \qquad \text{(II-6)}$$

If $p_i = p_o$ for $i=1,\ldots,n$, then

$$1 - P_n^{(1)} = (1-p_o)^n \qquad \text{(II-7}$$

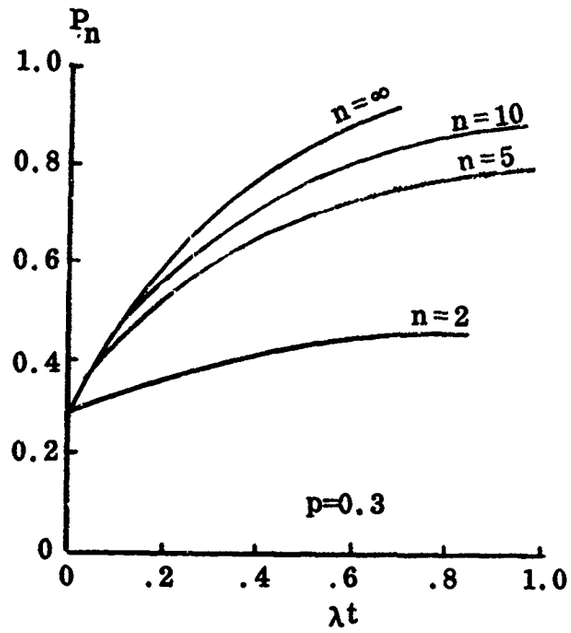
$$R_n^{(k)} = p_o^{\,k} (1-p_o)^{n-k} \binom{n}{k} \qquad \text{(II-8)}$$

and

$$P_n^{(k)} = \sum_{j=k}^{n} R_n^{(j)} = 1 - \sum_{j=0}^{k-1} R_n^{(j)}. \qquad \text{(II-9)}$$

FIGURE II-7

EFFECT OF NUMBER OF TRIALS ON TWO-STATE MARKOV cdp



Assumptions:

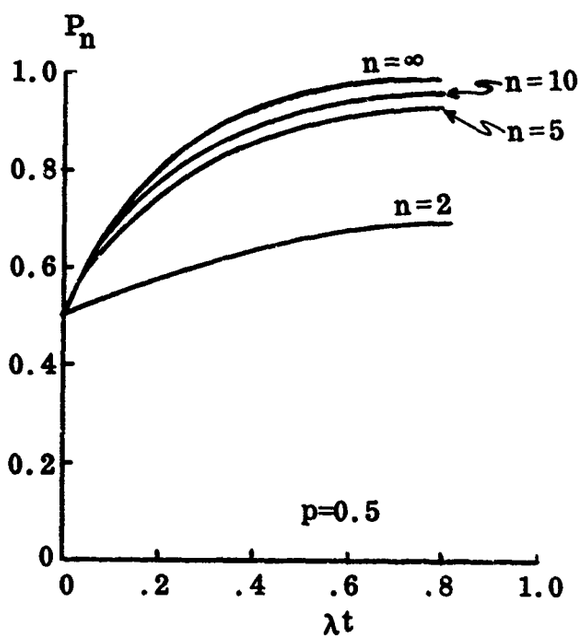
Success-failure process is stationary Markov

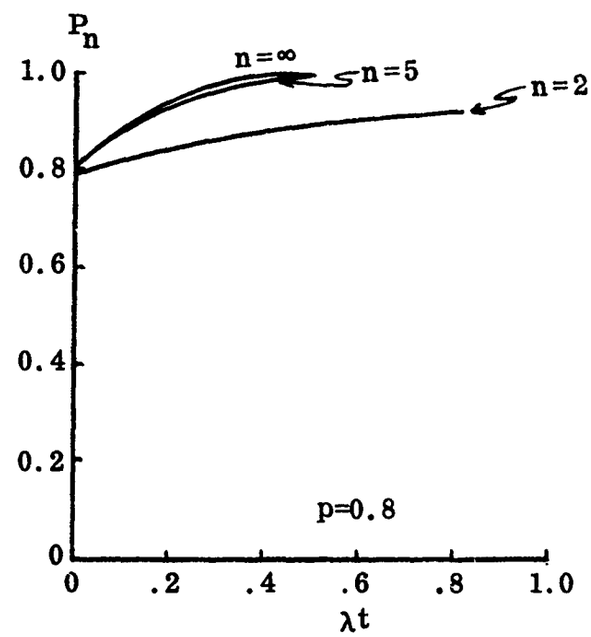
p = unconditional probability of success (constant)

n = number of trials in interval [o,t]

P_n = probability of at least one success in n trials

$e^{-\lambda h}$ = autocorrelation between trials separated by h





Formulas (II-6) through (II-9) are well-known; formulas (II-8) and (II-9) are tabulated in reference [m]. Formulas (II-6) and (II-7) are often applied erroneously (when the process is not independent), which is the motivation for this report.

The following theorem provides a means of finding $R_n^{(k)}$ in general:

<u>Theorem II-2</u>. With $p_1, \ldots, p_n$ and $R_n^{(k)}$ defined as above, we have

$$R_n^{(k)} = \frac{(-1)^k}{k!} \frac{\partial^k}{\partial t^k} \Lambda_n(p_1, \ldots, p_n, t) \Big|_{t=1} \quad \text{for } k = 0, \ldots, n,$$

where

$$\Lambda_n(p_1, \ldots, p_n, t) = \prod_{i=1}^{n} (1 - p_i t);$$

Λ_n may be thought of as a generating function for $R_n^{(k)}$.

<u>Proof.</u> We have the following recursive equation, writing $R_n^{(k)}$ as $R_n^{(k)}(p_1, \ldots, p_n)$:

$$R_n^{(k)}(p_1, \ldots, p_n) = \frac{1}{k} \sum_{i=1}^{n} p_i R_{n-1}^{(k-1)}(p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n).$$

It is also easily shown that

$$\frac{(-1)^k}{k!} \frac{\partial^k}{\partial t^k} \Lambda_n(p_1, \ldots, p_n, t) = \frac{(-1)^k}{k!} \sum_{i=1}^{n} \frac{\partial^{k-1}}{\partial t^{k-1}} \Lambda_{n-1}(p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n, t) \cdot (-p_i)$$

$$= \frac{1}{k} \sum_{i=1}^{n} p_i \cdot \frac{(-1)^{k-1}}{(k-1)!} \frac{\partial^{k-1}}{\partial t^k} \Lambda_{n-1}(p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n, t),$$

so that

$$\frac{(-1)^k}{k!} \frac{\partial^k}{\partial t^k} \Lambda_n$$

satisfies the same recursive equation as $R_n^{(k)}$.

Now, since

$$R_n^{(0)}(p_1, \ldots, p_n) = \prod_{i=1}^{n} (1 - p_i) = \Lambda_n(\mu_1, \ldots, p_n, t)$$

and

$$R_n^{(1)}(p_1, \ldots, p_n) = \sum_{i=1}^{n} \frac{p_i}{(1 - p_i)} \cdot \prod_{j=1}^{n} (1 - p_j)$$

$$= (-1) \left[\frac{\partial}{\partial t} \Lambda_n(p_1, \ldots, p_n, t)\right] \Bigg|_{t=1},$$

the theorem follows by induction on $k+n$.

#####

We illustrate Theorem II-2 by application to $k=1, 2, 3$. We have

$$\log \Lambda_n = \Sigma \log (1 - p_i t)$$

$$\frac{\partial \Lambda_n}{\partial t} = \Lambda_n \Sigma \frac{1}{1 - p_i t}$$

$$\frac{\partial^2 \Lambda_n}{\partial t^2} = - \frac{\partial \Lambda_n}{\partial t} \Sigma \frac{p_i}{1 - p_i t} - \Lambda_n \Sigma \frac{p_i^2}{(1 - p_i t)^2}$$

$$\frac{\partial^3 \Lambda_n}{\partial t^3} = - \frac{\partial^2 \Lambda_n}{\partial t^2} \Sigma \frac{p_i}{1 - p_i t} - 2 \frac{\partial \Lambda_n}{\partial t} \Sigma \frac{p_i^2}{(1 - p_i t)^2} - \Lambda_n \Sigma \frac{2 p_i^3}{(1 - p_i t)^3} .$$

Therefore, by Theorem II-2, letting $\omega_i = p_i / \bar{p}_i$ for $i = 1, \ldots, n$,

$$R_n^{(0)} = \Pi \bar{p}_i$$

$$R_n^{(1)} = \Pi \bar{p}_i \cdot \Sigma \omega_i$$

$$R_n^{(2)} = \Pi \bar{p}_i \cdot \{ (\Sigma \omega_i)^2 - \Sigma \omega_i^2 \}$$

$$R_n^{(3)} = \frac{1}{3} \Pi \bar{p}_i \cdot \{ 3 (\Sigma \omega_i) (\Sigma \omega_i^2) - (\Sigma \omega_i)^3 + 2 \Sigma \omega_i^3 \} .$$

These suffice to give $P_n^{(k)}$ for $k = 1, \ldots, 4$.

CHAPTER III

GAUSSIAN PROCESSES

In this chapter we assume that the underlying continuous-parameter process X^* is Gaussian. By this is meant simply that the joint distribution of any sequence of observation of random variables $(X_1, \ldots, X_n)$ is multivariate normal (see Appendix A).

In view of the definition of a Gaussian process, the problem of computing cdp,

$$P_n = 1 - \Pr\{X_i \leq T_1, \ldots, X_n < T_n\},$$

is equivalent to evaluating the multivariate normal distribution. For this reason, we have included in Appendix B, a survey of the literature on computing this distribution.

The main method of this chapter is a convenient approximation described in the first section, with computation aids in Figures III-1 through III-4. The computation method is illustrated by an example. The approximation is based on the assumption that the success-failure process D is Markovian. The error introduced by this assumption is investigated in the second section; insofar as it has been tested (for $n=3$ and $n=5$ trials -- in Table III-1), the approximation is very accurate, providing the discrete-parameter observation process X is Markovian (also true if X^* is Markovian). The approximation is less accurate in some special cases where X is not Markovian.

Approximation Method

Our principal method for computing cdp, P_n, in a Gaussian process will be an approximation method presented in this section. The method is based on the assumption that the success-failure process D is Markovian. It is intended that this approximation be primarily used when X is Markovian -- it does not follow that D is Markovian (e.g. see Rosenblatt, reference [n]), which is the source of error in the approximation. Errors in P_n will be investigated in the next section.

Under the assumption that D is Markovian, $\bar{P}_n$ is given by formula (II-5) of Chapter II:

$$\bar{P}_n = \bar{p}_1 \prod_{i=2}^{n} \bar{b}_i ,$$

where $\bar{b}_i$ is the probability that the ith trial fails, given failure on the $(i-1)^{st}$ trial; as before, p_i is the probability of success on the ith trial.

Define $z_{i,i-1}$ to be the probability that both the ith and $(i-1)^{st}$ trials fail, for $i = 2, \ldots, n$. Then by definition of conditional probability (or by applying formula (II-4) of Chapter II and Theorem C-2(ii) of Appendix C),

$$\bar{b}_i = \frac{z_{i,i-1}}{\bar{p}_{i-1}} .$$

The formula for $\bar{P}_n$ can thus be written:

$$\begin{aligned} \bar{P}_n &= \prod_{i=1}^{n} \bar{p}_i \cdot \prod_{i=2}^{n} \frac{z_{i,i-1}}{\bar{p}_i \, \bar{p}_{i-1}} \\ &= \prod_{i=1}^{n} \bar{p}_i \cdot \prod_{i=2}^{n} \gamma_{i,i-1} , \end{aligned} \qquad \text{(III-1)}$$

where

$$\gamma_{i,i-1} = \frac{z_{i,i-1}}{\bar{p}_i \, \bar{p}_{i-1}} .$$

Note that the product of the $\gamma_{i,i-1}$'s represents a multiplicative correction to the value of $\bar{P}_n$ which would be obtained in the presence of independent trials. For this reason, we will refer to $\gamma_{i,i-1}$ as a <u>correction factor</u>; the double subscript refers to the fact that it depends on behavior in both the ith and $(i-1)^{st}$ trials.

The Gaussian nature of X has nothing to do with formula (III-1). However, assuming, as we do, that X is Gaussian, $z_{i,i-1}$ is conveniently obtained from the bivariate normal tables, reference [D]. Using these tables, curves have been prepared for computational convenience, and presented as Figures III-1 through III-4, giving $\log_{10} \gamma_{i,i-1}$. The inputs to these curves pertain to the multi-state process X rather than the two-state process D:

(1) Thresholds T_i and T_{i-1} measured in standard deviations from the means of X_i and X_{i-1}. The thresholds are separately used to determine p_i and p_{i-1}.

(2) The correlation $\rho_{i,i-1}$ between X_i and X_{i-1}. This is <u>not</u> the same as the correlation r_{i-1} (not needed) between D_i and D_{i-1}, as shown by Theorem C-3 of Appendix C.

Separate curves have been prepared for $\rho_{i,i-1} = .3, .5, .7, .9$. At the end points

$$\log_{10} \gamma_{i,i-1} = 0 \qquad \text{when } \rho_{i,i-1} = 0$$

and

$$\log_{10} \gamma_{i,i-1} = -\log_{10} \max \{ \bar{p}_i , \bar{p}_{i-1} \} \quad \text{when } \rho_{i,i-1} = 1 .$$

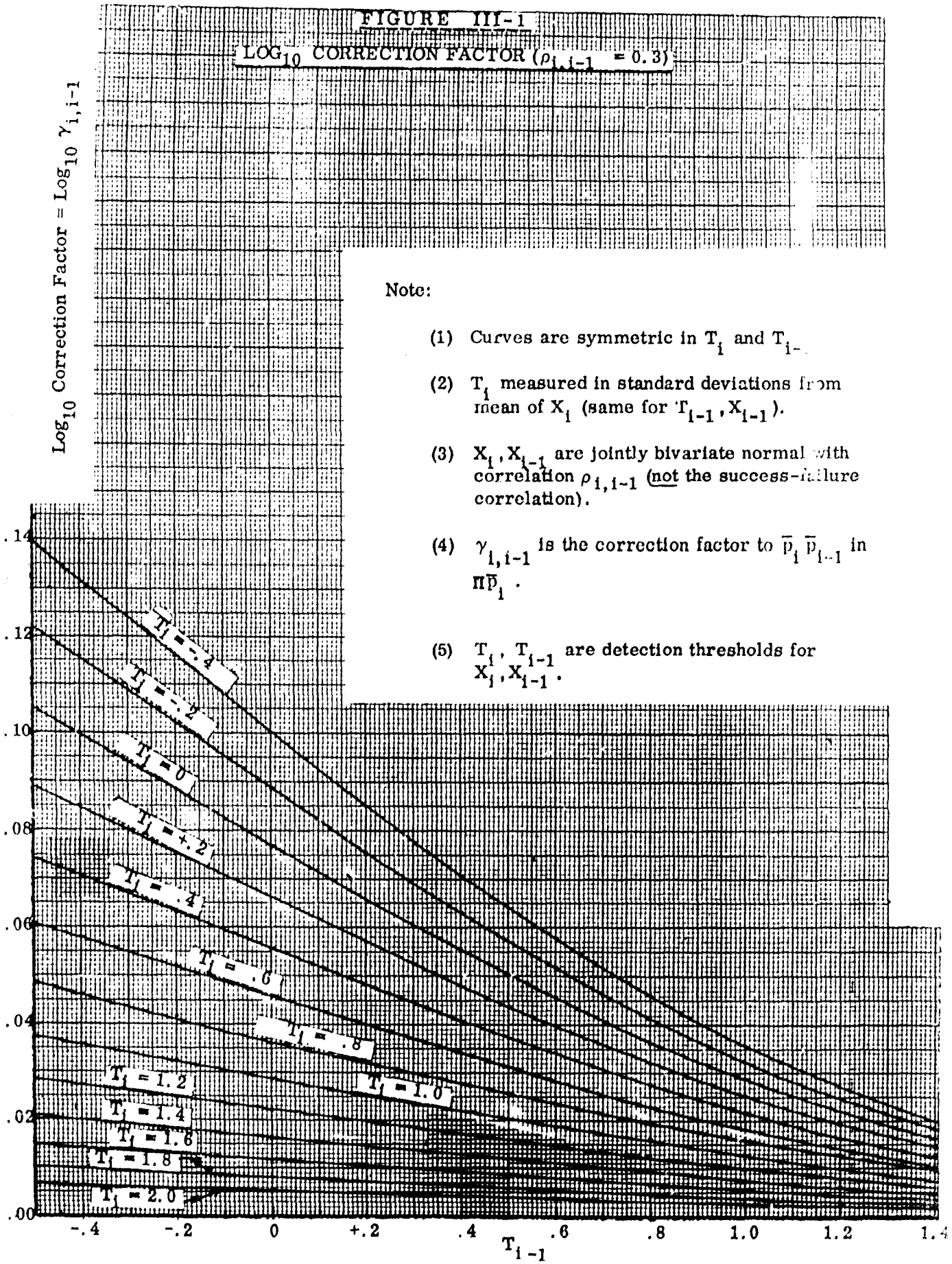
Linear interpolation of $\log_{10} \gamma_i$ with respect to $\rho_{i,i-1}$ is quite accurate. The figures are symmetric in T_i and T_{i-1}.

<u>Example:</u> We give an example to illustrate the procedure. Let X^* be a stationary Markov (Gaussian) process with common mean zero, common standard deviation of 5 db, and autocorrelation between any $X^*_{t_1}$ and $X^*_{t_2}$

$$\alpha (t_2 - t_1) = e^{-.69 \, | t_2 - t_1 |} ,$$

FIGURE III-1

LOG_{10} CORRECTION FACTOR ($\rho_{i,i-1} = 0.3$)



Note:

(1) Curves are symmetric in T_i and T_{i-1}.

(2) T_i measured in standard deviations from mean of X_i (same for T_{i-1}, X_{i-1}).

(3) X_i, X_{i-1} are jointly bivariate normal with correlation $\rho_{i,i-1}$ (not the success-failure correlation).

(4) $\gamma_{i,i-1}$ is the correction factor to $\bar{p}_i\ \bar{p}_{i-1}$ in $\Pi\bar{p}_i$.

(5) T_i, T_{i-1} are detection thresholds for X_i, X_{i-1}.

FIGURE III-2

LOG_{10} CORRECTION FACTOR ($\rho_{i,i-1} = 0.5$)

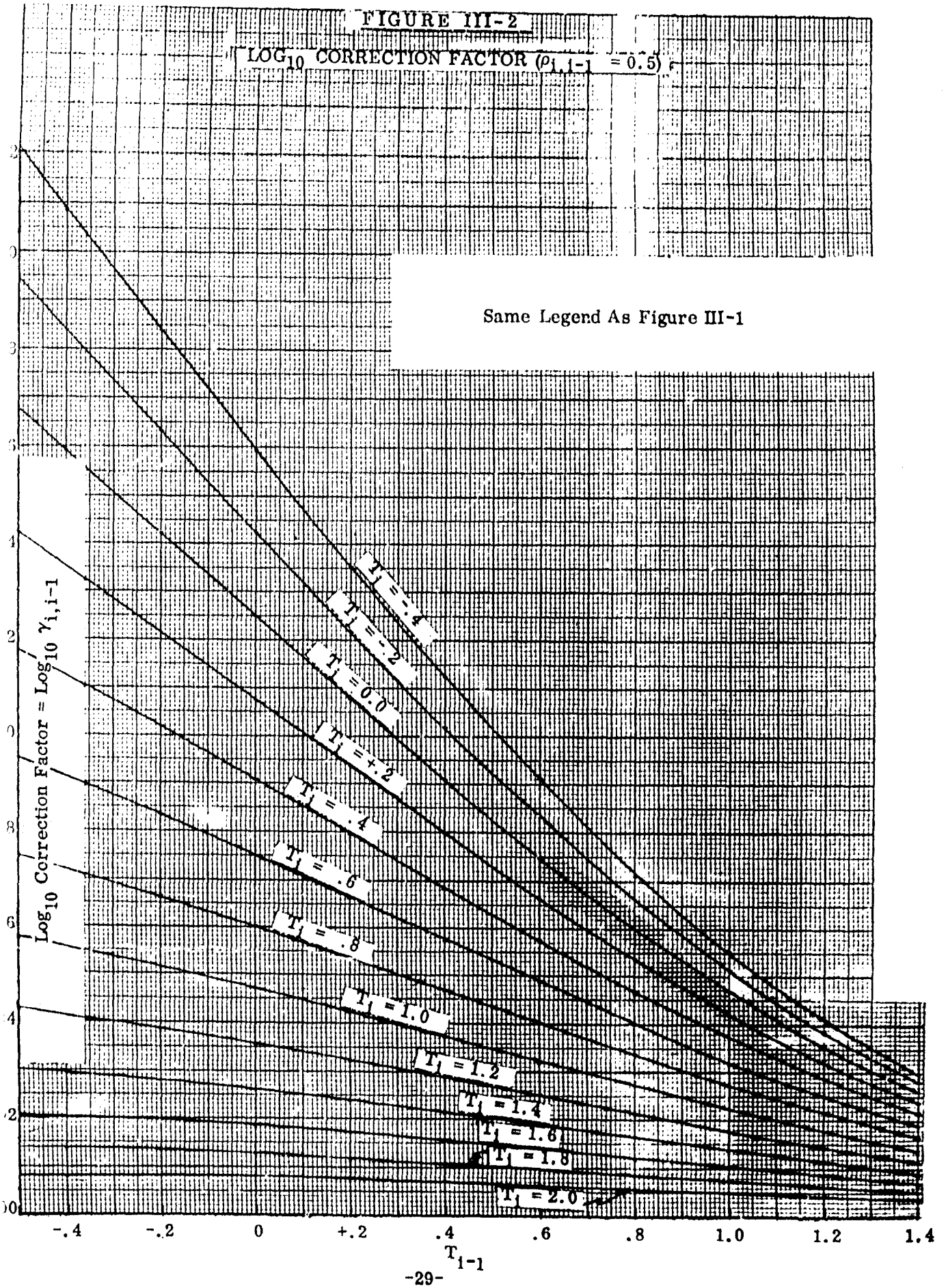
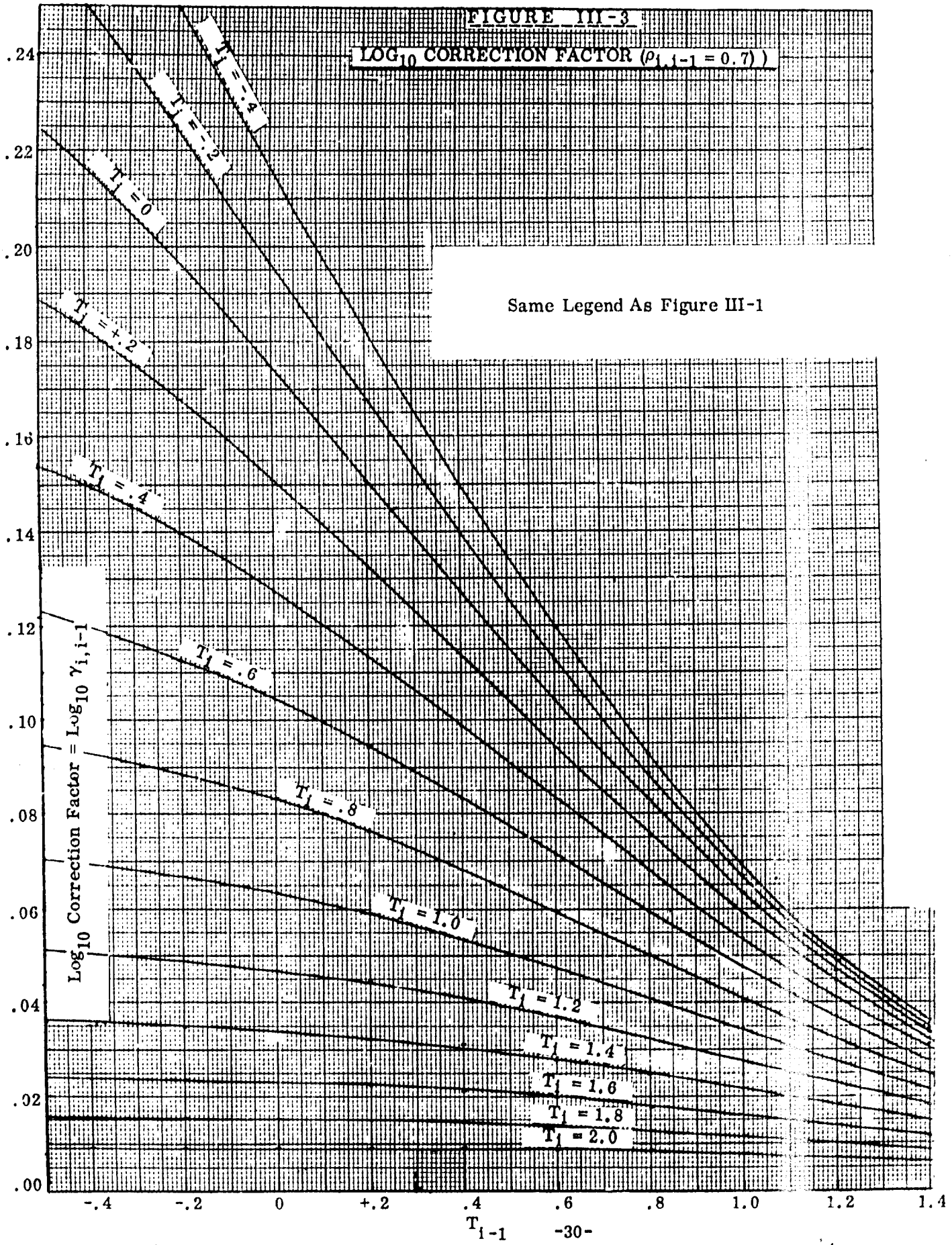


FIGURE III-3

LOG_{10} CORRECTION FACTOR ($\rho_{i,i-1} = 0.7$)



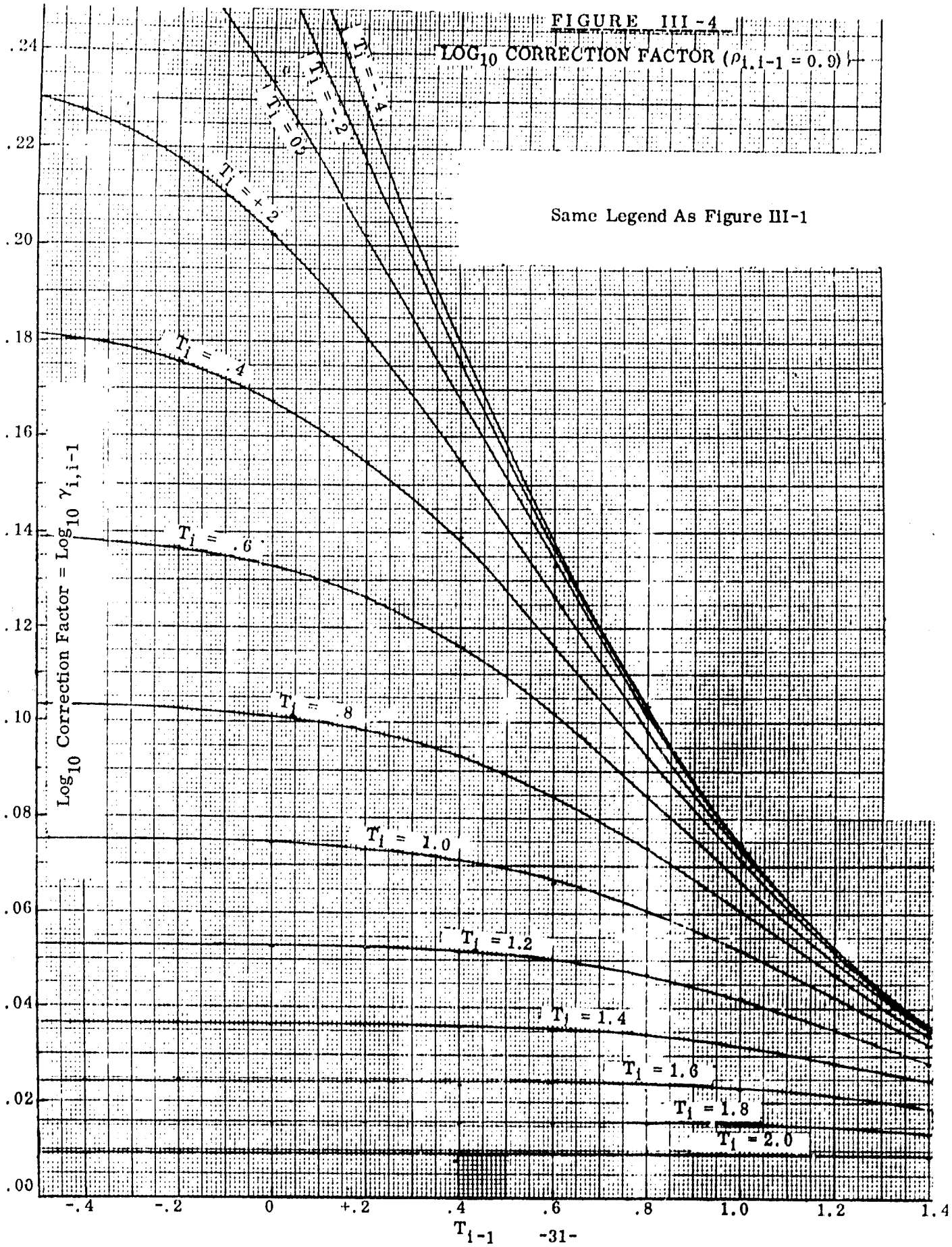


FIGURE III-4
LOG10 CORRECTION FACTOR ($\rho_{i,i-1}$ = 0.9)
Same Legend As Figure III-1
T_i = -.4
T_i = -.2
T_i = 0
T_i = +.2
T_i = .4
T_i = .6
T_i = .8
T_i = 1.0
T_i = 1.2
T_i = 1.4
T_i = 1.6
T_i = 1.8
T_i = 2.0
Log10 Correction Factor = Log10 $\gamma_{i,i-1}$
.24
.22
.20
.18
.16
.14
.12
.10
.08
.06
.04
.02
.00
-.4
-.2
0
+.2
.4
.6
.8
1.0
1.2
1.4
T_{i-1}

i.e., correlation is 0.5 between variables one time unit apart. Let X^* be sampled at times $t_1 = 0$, $t_2 = 2$, $t_3 = 3$, and $t_4 = 5$, i.e., four non-uniformly spaced trials. The process $(X_1, \ldots, X_4) = (X^*_{t_1}, \ldots, X^*_{t_4})$ is also Markovian, but is not stationary since

$$\rho_{21} = \alpha(t_2 - t_1) = .5^2 \neq .5 = \rho_{32} .$$

(It can be shown that the corresponding $(D_1, \ldots, D_4)$ is not Markovian, which is what makes the method only approximate.) Let the detection thresholds (in db) be 6.5, 2.5, 0, and 4.0 respectively.

The following table shows in tabular form all the steps necessary to calculate $\bar{P}_n$ in this situation:

Trial index i	1	2	3	4
Trial time t_i	0	2	3	5
Threshold T(in σ's from mean)	1.3	.5	.0	.8
Single-trial failure probability $\bar{p}_i$	.9032	.6915	.5000	.7881
Correlation $\rho_{i,i-1}$		.25	.50	.25
$\log_{10} \gamma_{i,i-1}$		.011	.083	.030
$\log_{10} \bar{p}_i$	.956-1	.840-1	.699-1	.897-1

Then $\log_{10} \bar{P}_n$ is the sum of the bottom two rows, i.e.

$$\log_{10} \overline{P}_n = 3.516\text{-}4 = .516\text{-}1$$

$$\overline{P}_n = .328 .$$

If one assumed the trials were independent, then one would obtain

$$\log_{10} \overline{P}_n = .392\text{-}1$$

$$\overline{P}_n = .247 .$$

Error Estimates

In this section we investigate the error in the approximation method of the preceding section. The error is introduced by the assumption that the success-failure process D is Markovian. For this purpose we assume that the discrete-parameter Gaussian process X is Markovian. When this condition is not satisfied, no information has been obtained on error estimates, except for the very special case of formula (A-4) of Appendix A. Even in the Markov case our error estimates are restricted primarily to the case where n=3 or 5 trials. General statements relevant to these error estimates are embodied in the three theorems below. At the conclusion of the theorems numerical examples are given in Table III-1.

Theorem III-1. Let (X_1, X_2, X_3) be a Gaussian Markov process and assume $\mu_i = 0$ and $\sigma_i = 1$ for $i = 1, 2, 3$. Let ρ_{ij} for $i \neq j$ be the value of the correlation between the random variables X_i and X_j, and assume that $0 < \rho_{ij} < 1$. Then

$$P_3 = \Pr\{ X_1 \leq T_1 \text{ and } X_2 \leq T_2 \text{ and } X_3 \leq T_3 \}$$

$$= \int_{-\infty}^{T_2} \Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1 - \rho_{21}^2}}\right) \cdot \Phi\left(\frac{T_3 - u\rho_{32}}{\sqrt{1 - \rho_{32}^2}}\right) d\,\Phi(u) ,$$

where Φ is the standard normal distribution function (see Appendix A).

Proof. It can be shown (see reference [o]), that the conditional distribution of X_j, given that $X_i = u$, is simply normal with mean $u\rho_{ij}$ and variance $1 - \rho_{ij}^2$. By hypothesis (X_1, X_2, X_3) is Gaussian Markov, and therefore X_1 and X_3 will be independent, given a particular value for X_2. Thus

$$\bar{P}_3 = \int_{-\infty}^{T_2} \Pr\{ X_i \le T_i, \text{ for } i = 1,3 \mid X_2 = u \}$$

$$= \int_{-\infty}^{T_2} \Pr\{ X_1 \le T_1 \mid X_2 = u \} \cdot \Pr\{ X_3 \le T_3 \mid X_2 = u \} \, d\Phi(u)$$

$$= \int_{-\infty}^{T_2} \Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1 - \rho_{21}^2}}\right) \Phi\left(\frac{T_3 - u\rho_{32}}{\sqrt{1 - \rho_{32}^2}}\right) d\Phi(u) . \quad \#\#\#\#\#$$

The same method can be used in more general fashion, e.g. if $(X_1, \ldots, X_5)$ is Gaussian Markov, then

$$\bar{P}_5 = \Pr\{ X_i \le T_i \text{ for } 1 \le i \le 5 \} = \int_{-\infty}^{T_3} \Pr\{ X_1 \le T_1 \text{ and } X_2 \le T_2 \mid X_3 = u \} \cdot \Pr\{ X_4 \le T_4 \text{ and } X_5 \le T_5 \mid X_3 = u \} \, d\Phi(u) . \tag{III-2}$$

Since the conditional distributions in the integrand are bivariate normal, one has a more difficult time in carrying out the numerical integration; however the bivariate normal distribution is tabulated in reference [p] .

Theorem III-2. Let (X_1, X_2, X_3), μ_i, σ_i^2, ρ_{ij} and P_3 be as in Theorem III-1. Let $\hat{P}_3$ be the approximate cdp obtained by assuming the success-failure process to be Markovian (formula III-1). Then the error in cdp is

$$\hat{P}_3 - P_3 = \frac{1}{\Phi(T_2)} \int_{-\infty}^{T_2} \int_{-\infty}^{T_2} \Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) \left[\Phi\left(\frac{T_3 - u\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right) - \Phi\left(\frac{T_3 - v\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right)\right] d\Phi(u)\, d\Phi(v) \qquad \text{(III-3)}$$

and

$$\hat{P}_3 \geq P_3 .$$

<u>Proof</u>. Under the Markov assumption on (D_1, D_2, D_3), the success-failure process,

$$1 - \hat{P}_3 = \Pr\{D_1 = D_2 = D_3 = 0\}$$

$$= \Pr\{D_1 = 0 \mid D_2 = 0\} \ \Pr\{D_3 = 0 \mid D_2 = 0\} \ \Pr\{D_2 = 0\}$$

$$= \Pr\{X_2 \leq T_2\} \ \Pr\{X_1 \leq T_1 \mid X_2 \leq T_2\} \ \Pr\{X_3 \leq T_3 \mid X_2 \leq T_2\}$$

$$= \Phi(T_2) \left[\int_{-\infty}^{T_2} \Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) d\Phi(u) \ \frac{1}{\Phi(T_2)} \right] \left[\int_{-\infty}^{T_2} \Phi\left(\frac{T_3 - u\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right) d\Phi(u) \ \frac{1}{\Phi(T_2)} \right]$$

$$= \frac{1}{\Phi(T_2)} \int_{-\infty}^{T_2} \int_{-\infty}^{T_2} \Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) \Phi\left(\frac{T_3 - v\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right) d\Phi(u)\, d\Phi(v) . \qquad \text{(III-4)}$$

Using Theorem III-1, we have

$$1 - P_3 = \int_{-\infty}^{T_2} \Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) \Phi\left(\frac{T_3 - u\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right) d\Phi(u)$$

$$= \frac{1}{\Phi(T_3)} \int_{-\infty}^{T_2} \int_{-\infty}^{T_2} \Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) \Phi\left(\frac{T_3 - u\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right) d\Phi(u)\, d\Phi(v) . \qquad \text{(III-5)}$$

Equation (III-3) follows from (III-4) and (III-5). It remains to show that $\hat{P}_3 \geq P_3$. If we denote the factor in square brackets in the right-hand side of (III-3) by $B(u,v)$, then clearly $B(u,v) = -B(v,u) > 0$ for $v > u$. Since the product measure of Φ itself is symmetric about the line $v = u$,

$$\hat{P}_3 - P_3 = \frac{1}{\Phi(T_2)} \int_{-\infty}^{T_2} \int_{-\infty}^{T_2} \left[\Phi\left(\frac{T_1 - u\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) - \Phi\left(\frac{T_1 - v\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) \right] B(u,v)\, d\Phi(u)\, d\Phi(v). \qquad \text{(III-6)}$$

But this integration is over the region where $v \geq u$, and in this case both factors in the integrand of formula (III-6) are non-negative. Hence $\hat{P}_3 \geq P_3$. #####

We now turn to the problem of obtaining an upper bound on the error $\hat{P}_3 - P_3$ in Theorem III-2.

<u>Theorem III-3</u>. Assume the hypothesis and notation of Theorem III-2 and also that the detection thresholds are constant, i.e. $T_i = T_o$. Then

$$\hat{P}_3 - P_3 \leq \frac{1}{\pi\sqrt{1-\rho_{21}^2}\sqrt{1-\rho_{32}^2}} \exp\left[-\frac{T_o^2}{2} \left(\frac{1-\rho_{21}}{1+\rho_{21}} + \frac{1-\rho_{32}}{1+\rho_{32}} \right) \right] \text{ for } T_o \geq 0$$

and

$$\hat{P}_3 - P_3 \leq \Phi(T_o) \quad \text{for } T_o < 0.$$

Proof. Let $T_o \geq 0$. Then $\Phi(T_o) \geq \frac{1}{2}$, and hence $1/\Phi(T_o) \leq 2$. When $v \geq u$ we obtain, by the Mean Value Theorem,

$$\Phi\left(\frac{T_o - u\rho_{ij}}{\sqrt{1-\rho_{ij}^2}}\right) - \Phi\left(\frac{T_o - v\rho_{ij}}{\sqrt{1-\rho_{ij}^2}}\right) = \frac{(v-u)}{\sqrt{1-\rho_{ij}^2}} \; e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} ,$$

where x is some number such that

$$\frac{T_o - v\rho_{ij}}{\sqrt{1-\rho_{ij}^2}} \leq x \leq \frac{T_o - u\rho_{ij}}{\sqrt{1-\rho_{ij}^2}} .$$

If $-\infty \leq v \leq T_o$, then the smallest value that x may assume is

$$\frac{T_o - T_o\rho_{ij}}{\sqrt{1-\rho_{ij}^2}} = \frac{T_o(1-\rho_{ij})}{\sqrt{1-\rho_{ij}^2}} = \frac{T_o\sqrt{1-\rho_{ij}}}{\sqrt{1+\rho_{ij}}} .$$

Now using equation (III-6) in the proof of Theorem IV-2, we see that for $T_o \geq 0$,

$$\hat{P}_3 - P_3 \leq \frac{1}{2\pi} \int_{-\infty}^{T_o} \int_{-\infty}^{v} \frac{(v-u)^2}{\sqrt{1-\rho_{21}^2}\sqrt{1-\rho_{32}^2}} \exp\left[-\frac{T_o^2}{2}\left(\frac{1-\rho_{21}}{1+\rho_{21}} + \frac{1-\rho_{32}}{1+\rho_{32}}\right)\right] d\Phi(u)\, d\Phi(v)$$

$$\leq \frac{1}{2\pi\sqrt{1-\rho_{21}^2}\sqrt{1-\rho_{32}^2}} \exp\left[-\frac{T_o^2}{2}\left(\frac{1-\rho_{21}}{1-\rho_{21}} + \frac{1-\rho_{32}}{1+\rho_{32}}\right)\right] \int_{-\infty}^{T_o} \int_{-\infty}^{v} (v-u)^2\, d\Phi(u)\, d\Phi(v)$$

$$\leq \frac{1}{\pi\sqrt{1-\rho_{21}^2}\sqrt{1-\rho_{32}^2}} \exp\left[-\frac{T_o^2}{2}\left(\frac{1-\rho_{21}}{1+\rho_{21}} + \frac{1-\rho_{32}}{1+\rho_{32}}\right)\right] .$$

For the case $T_o < 0$, we use equation (III-3) of Theorem III-2. Then,

$$\hat{P}_3 - P_3 \leq \frac{1}{\Phi(\hat{T})} \int_{-\infty}^{T_o} \int_{-\infty}^{T_o} \left| \Phi\left(\frac{T_o - u\,\rho_{21}}{\sqrt{1-\rho_{21}^2}}\right) \right| \left| \left[\Phi\left(\frac{T_o - u\,\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right) - \Phi\left(\frac{T_o - v\,\rho_{32}}{\sqrt{1-\rho_{32}^2}}\right)\right] \right| d\Phi(u)\, d\Phi(v)$$

$$\leq \frac{1}{\Phi(\hat{T})} \int_{-\infty}^{T_o} \int_{-\infty}^{T_o} d\Phi(u)\, d\Phi(v) = \Phi(\hat{T}).$$

#####

In order to test the accuracy of the approximation, the true P_n and approximate P_n were calculated and are displayed in Table III-1, for $n = 3$ or 5, and for various values of $\rho_{i,i-1}$ and T_i. The primary methods of calculating P_n were Theorem III-1 and formula (III-2), in cases where the Gaussian process $X = (X_1, \ldots, X_n)$ is Markovian. An additional method for X Markovian is mentioned in the notes. Some special cases where X is not Markovian were tested by use of formula (B-4) of Appendix B.

In most cases in Table III-1, X is stationary, while in some cases it is not, as noted. In most cases the threshold sequence is constant, but not in others.

For the cases where X is Markovian, the errors in P_n are in the order of .01. The errors are greater for P_n close to .5 than for P_n close to 1.0, as might be expected. The errors for $n = 5$ are not substantially different from those for $n = 3$. Note particularly example (11), which is both non-stationary and has non-constant thresholds -- the error is .003, which is very small considering that $P_n = .562$.

Note that the assumed correlations are in the range .4 to .7. In general, one expects less error for very high correlation or very low (absolute) correlation, since the approximation is perfect if (a) $\rho = 0$ throughout X or (b) $\rho = 1$ throughout X and the single-trial success probabilities are unimodal (i.e., non-decreasing prior to some point and non-increasing thereafter).

For the last three cases of Table III-1, X is not Markovian and the errors are somewhat higher, .068 for $n = 5$ and $n = 10$, and the approximation could not be rated better than fair. Nevertheless, the approximation is still a better estimate than would be obtained by assuming independence (being lower than the latter).

We conclude that the approximation is very good for at least five trials in case X is Gaussian Markov. This would be implied if X^* is Gaussian Markov, but not conversely. Stationarity and constant thresholds do not appear important. For X not Markovian, it is plausible that the approximation will be fair in many cases of interest -- certainly examples are at hand where the error is very large, especially as a percentage of $1 - P_n$.

TABLE III-1

EXAMPLES OF ERROR IN APPROXIMATING GAUSSIAN cdp

	No. of trials	Correlations	Thresholds (in σ's from mean)	True cdp	Approximation cdp	Error
	n	$\rho_{i,i-1}$	T_i	P_n	$\hat{P}_n$	$\hat{P}_n - P_n$
			- Markov Gaussian Process -			
(1)	5	.6	1.0	.443	.453	.010
(2)	5	.6	.5	.687	.696	.009
(3)	5	.4	1.0	.501	.505	.004
(4)	3	$1/\sqrt{2}$	0	.708	.721	.013
(5)	3	$1/\sqrt{5}$	0	.784	.790	.006
(6)	3	$1/\sqrt{2}$	$\sqrt{2}$	$.153 \pm .013$	.164	.011
(7)	3	$1/\sqrt{2}$	$\frac{1}{2}\sqrt{2}$	$.416 \pm .020$	.425	.009
(8)	3	$1/\sqrt{5}$	$\frac{1}{2}\sqrt{2}$	$.483 \pm .014$	.486	.003
(9)	3	$1/\sqrt{5}$	$-1/\sqrt{5}$	$.911 \pm .006$	.912	.001
(10)	3	$1/\sqrt{5}$	$-2/\sqrt{5}$	$.972 \pm .003$	.972	.000
(11)	3	.6, .36	(1.0, 0, 1.0)	.562	.565	.003
(12)	3	.7	(.4, 0, .4)	.611	.614	.003
(13)	3	.7	(-.2, 0, +.4)	.703	.709	.006
			- Gaussian Process Not Markov -			
(14)	2	$\left\{\begin{matrix}\rho_{ii} = 1 \\ \rho_{ij} = \frac{1}{2}, i \neq j\end{matrix}\right\}$	0	.667	.667	----
(15)	5		0	.833	.901	.068
(16)	10		0	.919	.987	.068

Notes

(1) In examples (4) through (11), P_n was computed by Theorem III-1 -- the "$\pm$" in (6) through (10) indicates the computed upper and lower Riemann sums in performing the numerical integration. In examples (1), (2), and (3), formula (III-2) was used. In (12) and (13), a more tedious calculation of P_n was performed: the three Gaussian distributions were each grouped into fourteen states and the two 14 x 14 transition matrices were multiplied. In examples (14), (15), and (16), formula (B-4) of Appendix B was used for P_n.

(2) In all cases, the approximation method using Figures III-1 through III-4 was used for $\hat{P}_n$.

(3) In example (11), (X_1, X_2, X_3) is stationary, since the two correlations are not equal. In (14), (15), and (16), X is again not stationary. In all other cases, X is stationary.

(4) The thresholds are not constant in examples (11), (12), and (13), but are constant elsewhere.

CHAPTER IV

STEP PROCESSES

In this chapter the underlying continuous-parameter stochastic process X^* is assumed to be a step process, defined in the first section. Intuitively, one may think of the sample functions of the step process as being step-functions, i.e., constant over a period of time until a jump occurs. After the jump occurs, a new value for the process is chosen which is independent of the values before the jump. The occurrence of jumps is a Poisson process. Pure step processes of this form are probably too idealized to be realized in physical detection processes -- their principal usefulness is as a first approximation to the step-plus-jitter processes considered in Chapter V.

In Theorems IV-1 and IV-2, it is assumed that the threshold function is "unimodal." Theorem IV-1 then gives a formula for cdp when the continuous-parameter process is sampled discretely. The corollary to this theorem gives a neater form of the cdp formula, in the case where the time between observations is constant. Theorem IV-2 passes the result of Theorem IV-1 to a limiting case, where the step process is observed continuously.

The final section deals with an arbitrary threshold function. In this case convenient exact formulas for cdp are not available. However, a recursive procedure has been devised which is suitable when the number of observations is not large. The number of multiplications required to compute cdp for a sequence of n observations using the recursive relations is of the order of n^2.

Definitions

The definition of a step process requires the specification of two auxiliary processes Y and N. The process Y consists of a sequence of independent, identically-distributed random variables Y_i for $i = 0, 1, 2, \ldots$. The common distribution function of the random variables Y_i is called the location-after-jump distribution, and will be denoted by K. The process N, called the aggregate-jump

process, is specified to be a Poisson* process with intensity λ (see reference [q] for an excellent elementary discussion of this process): N_t is the number of jumps which occur from time zero up to and including time t. The numbers of jumps occurring in disjoint intervals are independent. Moreover, the probability of any number of jumps occurring in a given interval depends only upon the length of the interval and not upon its location in time. The expected number of jumps in time t is λt and

$$\Pr\{N_t = m\} = e^{-\lambda t} \frac{(\lambda t)^m}{m!} \quad \text{for } m = 0, 1, \ldots.$$

In particular, the probability of no jump in any interval of length s is $e^{-\lambda s}$.

The <u>step</u> process, X^*, is now defined by the following composition:

$$X_t^* = Y_{N_t} \quad \text{for } s \geq 0.$$

When the location-after-jump distribution is normal with standard deviation σ, X^* has been referred to as a "(λ, σ)-process" in Appendix B (originated by J. D. Kettelle Jr.) of reference [h].

Clearly, X^* simply describes a special type of sampling scheme where a new sample is drawn from the distribution K whenever a jump occurs (as indicated by the process N). Each new sample is independent of those previously drawn.

A threshold function is said to be <u>unimodal</u> if it is non-increasing prior to some parameter value (not necessarily unique) at which the threshold is minimal and therefore non-decreasing. Actually, the term "unimodal" would be more aptly applied to the corresponding single-trial success probabilities, p_t, since these rise to a <u>maximum</u> and are non-increasing thereafter.

It is shown in Theorem C-4 of Appendix C that the autocorrelation between observations separated by s in a step process is $e^{-\lambda s}$. We note, without proof, that a step-process is Markovian -- we do not use this fact per se.

---

* It should be noted that in the physical process of reference [a], which provides a principal motivation for considering step processes, the occurrence of jumps fits much better to an Erlang process of order two or three than to a Poisson process (Erlang of order one).

Unimodal Threshold Theorems

In this section, we assume the threshold function T^* is unimodal. Theorem IV-1 will give the formula for cdp $= P_n$ for the case where the process X^* is observed at n arbitrary discrete points in time, and Theorem IV-2 will extend this result to the continuous-parameter case. We repeat the convention made in Chapter I that the symbols X, T, and p denote respectively the process, threshold function, and single-trial success function obtained by sampling the continuous-parameter process X^* at discrete points in time $t_1, \ldots, t_n$. Note that if T^* is unimodal, then T will also be unimodal.

Theorem IV-1. Let X be the discrete parameter process obtained by observing X^* at times $t_1, \ldots, t_n$, and β_i (for $i = 2, 3, \ldots, n$) be the probability of at least one jump in an interval of length $t_i - t_{i-1}$. Thus

$$\beta_i = 1 - e^{-\lambda (t_i - t_{i-1})}.$$

Let c be an observation corresponding to a minimum threshold T_c. Then P_n is given by*

$$P_n = 1 - (1 - p_c) \prod_{i=1}^{c-1} (1 - \beta_{i+1} p_i) \prod_{i=c+1}^{n} (1 - \beta_i p_i). \qquad \text{(IV-1)}$$

Proof. We make use of the following elementary fact about conditional probabilities: Let A, B, and C be arbitrary events and let $\bar{C}$ be the complementary event, "not C." Suppose, in addition, that B is independent of both C and $\bar{C}$. Then

$$\Pr\{A|B\} = \Pr\{A|B \cap C\} \Pr\{C\} + \Pr\{A|B \cap \bar{C}\} \Pr\{\bar{C}\}.$$

---

* When the upper limit of the symbol Π is less than the lower limit, we follow the convention that the product is then equal to one.

Let F_i correspond to the event $X_i \leq T_i$, and L_i correspond to the event "at least one jump occurs between t_{i-1} and t_i." The probability $\bar{P}_n$ may then be written as

$$\bar{P}_n = \prod_{i=1}^{n} d_i, \tag{IV-2}$$

where

$$d_i = \Pr\{ F_i \mid F_1 \cap F_2 \ldots \cap F_{i-1} \} \quad \text{for } i > 1$$

$$d_1 = \Pr\{ F_1 \} = \bar{p}_1 .$$

Noting that L_i is independent of $F_1 \cap \ldots \cap F_{i-1}$, we may use the above remark and write

$$\begin{aligned} d_i = {} & \Pr\{ F_i \mid F_1 \cap F_2 \cap \ldots \cap F_{i-1} \cap L_i \} \ \Pr\{ L_i \} \\ & + \Pr\{ F_i \mid F_1 \cap F_2 \cap \ldots \cap F_{i-1} \cap \bar{L}_i \} \ \Pr\{ \bar{L}_i \} . \end{aligned} \tag{IV-3}$$

Now, if $i \leq c$, then

$$d_i = \bar{p}_i \beta_i + \frac{\bar{p}_i}{\bar{p}_{i-1}} \bar{\beta}_i = \frac{\bar{p}_i}{\bar{p}_{i-1}} (1 - \beta_i p_{i-1}). \tag{IV-4}$$

The first term arises from the fact that when a jump occurs in the interval between t_i and t_{i-1}, the event F_i is independent of the past. Thus the first summand in equation (IV-3) is just $\bar{p}_i \beta_i$. When no jump occurs in the interval between t_i and t_{i-1}, we must take the past behavior of the process into account. The conditioning in (IV-3) states that, among other things, the event F_{i-1} has occurred. This means

that $X_{i-1} \leq T_{i-1}$, and consequently $X_i \leq T_{i-1}$, since there has not been a jump between t_{i-1} and t_i. Now, since the thresholds do not get any smaller as we proceed backwards in time, we gain no new information from the condition that events $F_1, \ldots, F_{i-2}$ have occurred. Thus

$$\Pr\{ F_i \mid F_1 \cap \ldots F_{i-1} \cap \bar{L}_i \} = \Pr\{ X_i \leq T_i \mid X_i \leq T_{i-1} \} = \frac{\bar{p}_i}{\bar{p}_{i-1}},$$

whence formula (IV-4) follows.

The situation is slightly simpler whenever $i > c$. In this case,

$$d_i = \bar{p}_i \beta_i + \bar{\beta}_i = 1 - \beta_i p_i. \qquad \text{(IV-5)}$$

As before, the first summand of (IV-3) becomes $\bar{p}_i \beta_i$, since the event F_i is independent of the past when a jump occurs between t_{i-1} and t_i. When no jump occurs between t_{i-1} and t_i, we have $\Pr\{ F_i \mid F_1 \cap \ldots \cap F_{i-1} \} = 1$, since, in this case, $X_{i-1} = X_i$ and the conditioning states that $X_{i-1} \leq T_{i-1}$. But $i > c$ implies that $T_{i-1} \leq T_i$, and, therefore, $X_i \leq T_i$. Thus when $i > c$, the second summand in (IV-3) is given by $\bar{\beta}_i$, and this establishes (IV-5). Combining (IV-4) and (IV-5) by means of (IV-2), we obtain the conclusion.

#####

Corollary. Suppose the observations are uniformly spaced, that is,

$$t_i = (i-1)\,\delta \quad \text{for } i = 1, 2, \ldots, n.$$

Then the formula for P_n is

$$P_n = 1 - \frac{1 - p_c}{1 - \beta\, p_c} \prod_{i=1}^{n} (1 - \beta\, p_i),$$

where β is the probability of at least one jump in an interval of length δ, i.e., $\beta = 1 - e^{-\lambda\delta}$.

In the next theorem we proceed to compute P_∞, the limiting value of cdp when it is assumed that the process X^* can be continuously observed over an entire interval $[0,t]$. Such a probability might be expressed formally as

$$P_\infty = 1 - \Pr\{ \sup_{0 \le s \le t} X_s^* - T_s^* \le 0 \};$$

we define its value to be $\lim_{n\to\infty} P_n$, where P_n is the cdp resulting from taking n uniformly-spaced observations in the interval $[0,t]$, with the first observation at 0 and the last at t.

(It can be shown that even when the observations are not uniformly spaced, the same limit will be obtained as long as the first and last observations are at times 0 and t respectively, and the maximum time between observations approaches zero.)

Recall that p_t^* was defined to be the probability of a success in a trial at time t, and define s_o to be a point such that $T_{s_o}^*$ is a minimum threshold.

Theorem IV-2.* Let the unimodal threshold function T^* be a continuous function. Then the limiting cdp, P_∞, for continuous observations over the interval $[0,t]$ is given by

$$P_\infty = 1 - (1 - p_{s_o}^*) \exp\left[-\lambda \int_0^t p_s^* \, ds\right],$$

where λ is the intensity of the Poisson aggregate-jump process.

Proof. Let the interval $[0,t]$ be divided uniformly into $n-1$ intervals of length δ -- the observations taking place at the points $t_i = (i-1)\delta$. Now, using

---

* In the more special case when the thresholds are monotonic, this formula was obtained earlier by J. D. Kettelle Jr. (see reference [h], page B-4) using a differential equation approach. Both results include Theorem II-1 of Chapter II, which assumes that the thresholds are constant.

the corollary to Theorem IV-1, we have

$$\bar{P}_n = \frac{1-p_c}{1-\beta\, p_c} \prod_{i=1}^{n} (1-\beta\, p_i),$$

where

$$\beta = 1-e^{-\lambda\delta} = \lambda\delta + o(\delta).$$

If there is more than one candidate for c, we choose c so that t_c is at minimum distance from s_o. Clearly then, $t_c \to s_o$ as $n \to \infty$. We compute

$$\bar{P}_\infty = \lim_{n\to\infty} \bar{P}_n$$

by the use of logarithms. Now

$$\ln \bar{P}_n = \ln(1-p_c) - \ln(1-\beta\, p_c) + \sum_{i=1}^{n} \ln(1-\beta\, p_i)$$

$$= \ln(1-p_c) - \ln(1-\beta\, p_c) + \sum_{i=1}^{n} \ln(1-\lambda\, p_i\, \delta + o(\delta))$$

$$= \ln(1-p_c) - \ln(1-\beta\, p_c) + \sum_{i=1}^{n} (-\lambda\, p_i\, \delta) + n\cdot o(\delta).$$

Finally, since $\lim_{n\to\infty} \beta = 0$ and $n = 1 + t/\delta$, we have

$$\ln \bar{P}_\infty = \lim_{n\to\infty} \ln \bar{P}_n = \ln(1-p^*_{s_o}) - \lambda \int_0^t p^*_s\, ds.$$

Here, the fact that p_s^* is non-decreasing and then non-increasing ensures that it is continuous almost everywhere and hence Riemann-integrable. Further $\lim_{n\to\infty} p_c = p_{s_0}^*$ since $p_c = 1 - K(T_{t_c})$, $T_{t_c} \to T_{s_0}$ from the right (since T is continuous and T_{s_0} is a minimum threshold), and K is a distribution function and hence right continuous.

#####

General Thresholds -- A Recursive Scheme

One might attempt to compute cdp for n discrete observations of a step process by simply making a list of all the different configurations of jumps which are possible in the n-1 intervals of time between n observations, and then compute the non-detection probability for each configuration. The result of adding these together, weighted by the probability of each configuration of jumps would then give $1 - P_n$. However, if we denote the event "at least one jump" by a 1 and the event "no jumps" by a 0, then it is clear that there are $2^{(n-1)}$ ways the jumps may occur. (We are not particularly interested in "how many" jumps occur in an interval, since the net effect of any number of jumps is to cause us to take another independent sample from the distribution K.) Clearly, the above method of computing cdp would become tedious for large n. The recursive scheme to be described is a considerable improvement over such a straightforward approach.

As before, let $\{t_i\}$ be a sequence of observations, and β_i be the probability of at least one jump occurring in an interval of length $t_i - t_{i-1}$. In addition, let

$$q_{k,j} = \Pr\{X_i \leq \min_{j \leq i \leq k} T_i\} = K(\min_{j \leq i \leq k} T_i) \quad \text{for } k \geq j.$$

Note that $q_{j,j} = \bar{p}_j$.

We shall compute $\bar{P}_n$ for $n = 1, 2, 3, 4$ in order to demonstrate the method of computing $\bar{P}_{n+1}$ in terms of $\bar{P}_n$. Table IV-1 shows this computation. The symbol $Q_m(k)$ denotes the contribution to $\bar{P}_m$ from all configurations of jumps ending in k zeros. When $k = 0$, we, of course, refer to those configurations ending with a 1. The case $n = 1$ is trivial, i.e., $\bar{P}_1 = q_{1,1} = \Pr\{X_1 \leq T_1\}$. An inspection of Table IV-1 shows that,

TABLE IV-1

CONTRIBUTIONS TO cdp -- RECURSIVE SCHEME

Configuration of Jumps		Contribution to $\bar{P}_n$	
$n=2$	1	$\beta_2\ q_{2,2}\ q_{1,1}$	$Q_2^{(0)}$
	0	$\bar{\beta}_2\ q_{2,1}$	$Q_2^{(1)}$
$n=3$	1 1	$\beta_3\ \beta_2\ q_{3,3}\ q_{2,2}\ q_{1,1}$	$Q_3^{(0)}$
	0 1	$\beta_3\ \bar{\beta}_2\ q_{3,3}\ q_{2,1}$	
	1 0	$\bar{\beta}_3\ \beta_2\ q_{3,2}\ q_{1,1}$	$Q_3^{(1)}$
	0 0	$\bar{\beta}_3\ \bar{\beta}_2\ q_{3,1}$	$Q_3^{(2)}$
$n=4$	1 1 1	$\beta_4\ \beta_3\ \beta_2\ q_{4,4}\ q_{3,3}\ q_{2,2}\ q_{1,1}$	$Q_4^{(0)}$
	0 1 1	$\beta_4\ \beta_3\ \bar{\beta}_2\ q_{4,4}\ q_{3,3}\ q_{2,1}$	
	1 0 1	$\beta_4\ \bar{\beta}_3\ \beta_2\ q_{4,4}\ q_{3,2}\ q_{1,1}$	
	0 0 1	$\beta_4\ \bar{\beta}_3\ \bar{\beta}_2\ q_{4,4}\ q_{3,1}$	
	1 1 0	$\bar{\beta}_4\ \beta_3\ \beta_2\ q_{4,3}\ q_{2,2}\ q_{1,1}$	$Q_4^{(1)}$
	0 1 0	$\bar{\beta}_4\ \beta_3\ \bar{\beta}_2\ q_{4,3}\ q_{2,1}$	
	1 0 0	$\bar{\beta}_4\ \bar{\beta}_3\ \beta_2\ q_{4,2}\ q_{1,1}$	$Q_4^{(2)}$
	0 0 0	$\bar{\beta}_4\ \bar{\beta}_3\ \bar{\beta}_2\ q_{4,1}$	$Q_4^{(3)}$

$$Q_3^{(0)} = \beta_3 \; q_{3,3} \; \bar{P}_2$$

$$Q_3^{(1)} = \bar{\beta}_3 \frac{q_{3,2}}{q_{2,2}} Q_2^{(0)}$$

$$Q_3^{(2)} = \bar{\beta}_3 \frac{q_{3,1}}{q_{2,1}} Q_2^{(1)}.$$

Moreover,

$$Q_4^{(0)} = \beta_4 \; q_{4,4} \; \bar{P}_3$$

$$Q_4^{(1)} = \bar{\beta}_4 \frac{q_{4,3}}{q_{3,3}} Q_3^{(0)}$$

$$Q_4^{(2)} = \bar{\beta}_4 \frac{q_{4,2}}{q_{3,2}} Q_3^{(1)}$$

$$Q_4^{(3)} = \bar{\beta}_4 \frac{q_{4,1}}{q_{3,1}} Q_3^{(2)}.$$

The general recursive formulas are given by the following set of equations:

$$\bar{P}_n = Q_n^{(0)} + Q_n^{(1)} + \ldots + Q_n^{(n-1)}$$

$$Q_n^{(0)} = \beta_n \; q_{n,n} \; \bar{P}_{n-1}$$

$$Q_n^{(1)} = \bar{\beta}_n \frac{q_{n,n-1}}{q_{n-1,n-1}} Q_{n-1}^{(0)}$$

$$Q_n^{(2)} = \bar{\beta}_n \frac{q_{n,n-2}}{q_{n-1,n-2}} Q_{n-1}^{(1)}$$

$$\vdots$$

$$Q_n^{(n-1)} = \bar{\beta}_n \frac{q_{n,1}}{q_{n-1,1}} Q_{n-1}^{(n-2)}.$$

The procedure begins at $n = 2$, where

$$Q_2^{(0)} = \beta_2 \; q_{2,2} \; q_{1,1}$$

and

$$Q_2^{(1)} = \bar{\beta}_2 \; q_{2,1}.$$

Assuming that the ratios involving $q_{i,j}$ have been computed beforehand, it takes about $2n$ multiplications to go from $\bar{P}_{n-1}$ to $\bar{P}_n$. Thus, in order to compute $\bar{P}_n$ starting from $\bar{P}_2$ we must make

$$2 \sum_{i=3}^{n} i = n(n+1) - 6$$

multiplications. Clearly this is better than the "straightforward approach" which would involve more than 2^{n-1} multiplications. Certain special threshold functions will also reduce the amount of calculation.

CHAPTER V

STEP-PLUS-JITTER PROCESSES

In this chapter we assume that the continuous-parameter process X^* is a "step-plus-jitter" process. This is a step process perturbed by slight disturbances which we call "jitter." Precise definitions will be given in the next section. The computation of cdp, P_n, will be discussed only for the case of a finite sequence of observations of the continuous-parameter process, i.e., no limiting cases will be considered.

Recall that the problem of finding non-recursive formulas for cdp for an ordinary step process was quite difficult except when the threshold function T was unimodal. The presence of jitter further complicates matters, and in fact no non-recursive formulas have been found, even in the case where the threshold function is constant. However, recursive relations similar to those of Chapter IV are derived which are valid for an arbitrary threshold function and for any finite sequence of observations. These are given in the second section. The main disadvantage of these relations is that they depend upon a number of integral expressions which may be difficult to evaluate. Even assuming the values of the integrals are known, the number of multiplications to compute cdp, P_n, is of the order of n^2.

In order to avoid these difficulties, an approximation to the step-plus-jitter process by a special "m-dependent" (see Appendix A) process is discussed in the last two sections. When the threshold function is constant and the observations are uniformly spaced, cdp may be computed for the approximate step-plus-jitter process by recursive relations which greatly reduce the number of integrals to be evaluated. In addition, the computation of cdp for the approximation requires merely order of n multiplications. The last section deals with the general m-dependent approximation, while the section preceding it considers in detail the special 2-dependent case. Here numerical calculations as shown in Figures V-1 through V-3 indicate that the approximation is most satisfactory when there is a high probability of a jump occuring between observations.

Definitions

A step-plus-jitter process X^* is defined to be the sum of two stochastic processes Z^* and J^*, where Z^* is a step process, and J^* (called jitter) is a process of independent, identically-distributed random variables, which is assumed to be statistically independent of Z^*; the common distribution function of the random variables J_t^* will be denoted by G.

We recall a step process, Z^*, is defined by the equation

$$Z_t^* = Y_{N_t},$$

where Y is a process of independent identically-distributed random variables, and N is a Poisson process describing the occurrence of jumps.

As usual, let X, Z, and J be the discrete-parameter process resulting from making observations at the points $t_1, \ldots, t_n$. For any such sequence of observations, we let

$$f_{k,j} = \Pr\{X_j \leq T_j,\ X_{j+1} \leq T_{j+1}, \ldots, X_k \leq T_k \mid \text{no jumps between } j \text{ and } k\}.$$

The quantity $f_{k,j}$ may be computed by

$$\begin{aligned} f_{k,j} &= \Pr\{Z_j + J_j \leq T_j, Z_j + J_{j+1} \leq T_{j+1}, \ldots, Z_j + J_j \leq T_k\} \\ &= \Pr\{J_j \leq T_j - Z_j,\ J_{j+1} \leq T_{j+1} - Z_j, \ldots, J_k \leq T_k - Z_j\} \qquad \text{(V-1)} \\ &= \int_{-\infty}^{\infty} \prod_{i=j}^{k} G(T_i - x)\, dK(x), \end{aligned}$$

where G is defined above and K is the location-after-jump distribution of the step process. Note that when the threshold function T is constant ($T \equiv T_0$), $f_{k,j}$ depends only upon k-j. In this case, we denote $f_{k,j}$ by f_{k-j} and

$$f_{k-j} = \int_{-\infty}^{\infty} G^{k-j+1}(T_0 - x) \; dK(x). \qquad \text{(V-2)}$$

The autocorrelation associated with the step-plus-jitter process $Z^* + J^*$ is

$$\alpha(s) = \frac{\sigma_K^2}{\sigma_K^2 + \sigma_G^2} \; e^{-\lambda s} \text{ for } s \neq 0,$$

$$\alpha(0) = 1,$$

where σ_K^2 and σ_G^2 are respectively the variance of K and G, and λ is the intensity of the Poisson process N. This is seen by applying Theorems C-2 and C-4 of Appendix C. Note that α is discontinuous at 0 (assuming, as we do, that $\sigma_G^2 \neq 0$, i.e., the jitter is not constant).

Exact Recursive Formula for cdp

In this section, recursive relations will be displayed which allow one to compute the exact value of cdp, P_n, whenever the underlying continuous-parameter process is step-plus-jitter. These relations are identical in form (with a substitution noted below) to the relations appearing in the last section of Chapter IV. They apply with arbitrary threshold function.

Again, we use 1 and 0 to represent the occurrence or non-occurrence of one or more jumps in a given interval and β_j for $j \geq 2$ is the probability of at least one jump occurring in an interval of length $t_j - t_{j-1}$. The symbol $Q_n^{(k)}$ is used to denote the contribution to $\bar{P}_n$ from all configurations of jumps ending in k zeros. The only specific change which has been made in going from the relations of Chapter IV to those of this section is to replace $q_{k,j}$ with $f_{k,j}$ defined in the preceding section.

The relations are now given by the following set of equations:

$$\bar{P}_n = Q_n^{(0)} + Q_n^{(1)} + \ldots + Q_n^{(n-1)}$$

$$Q_n^{(0)} = \beta_n \, f_{n,n} \, \overline{P}_{n-1}$$

$$Q_n^{(1)} = \overline{\beta}_n \, \frac{f_{n,n-1}}{f_{n-1,n-1}} \, Q_{n-1}^{(0)}$$

$$Q_n^{(2)} = \overline{\beta}_n \, \frac{f_{n,n-2}}{f_{n-1,n-2}} \, Q_{n-1}^{(1)}$$

$$\cdots$$

$$Q_n^{(n-1)} = \overline{\beta}_n \, \frac{f_{n,1}}{f_{n-1,1}} \, Q_{n-1}^{(n-2)} \, .$$

The procedure begins at n=2, where

$$Q_2^{(0)} = \beta_2 \, f_{1,1}$$

and

$$Q_2^{(1)} = \overline{\beta}_2 \, f_{2,1} \, .$$

The reader may verify these equations by consulting Table IV-1 of Chapter IV, keeping in mind that the symbols $q_{k,j}$ should be replaced by $f_{k,j}$. Note that multiplications in order of n^2 are required to compute P_n, <u>assuming</u> that the required ratios of the $f_{k,j}$ are calculated in advance. In view of equation (V-1), the calculation of $f_{k,j}$ could be a formidable task. The type of approximation introduced in the next two sections substantially reduces the number of $f_{k,j}$ to be computed.

The 2-Dependent Approximation

In this section we discuss in detail an approximation to the step-plus-jitter process by a 2-dependent process when the threshold function is constant, and the observations are uniformly spaced (i.e., $t_i = (i-1)\,\delta$). This is a special case of the general m-dependent approximation discussed in the next section. We have singled

out the 2-dependent approximation because it is the simplest non-trivial example of the general method. Appendix A defines m-dependent processes, which are used in these approximations.

As we have seen throughout this report, the presence of correlation between random variables makes the problem of computing cdp vastly more difficult than was the case for independent random variables. In fact, the literature survey in Appendix B showed for the class of Gaussian processes, there are no finite procedures or tabulations available for the computation (under general conditions) of cdp for as few as three trials. For the two-dependent approximation $\hat{X}$, however, the random variables $\hat{X}_i$ and $\hat{X}_j$ are independent whenever $|i-j|>2$. This manifests itself in the fact that the recursive formulas obtained in this section for the approximation will require only order n multiplications for the computation of cdp--in contrast to order n^2 multiplications to compute the cdp for the actual step-plus-jitter process.

As was noted earlier, whenever the threshold function is constant, $f_{k,j}$ depends only upon k-j and is denoted by f_{k-j}.

The approximate step-plus-jitter process is obtained by constructing a new aggregate jump process Γ, which will resemble the discretely observed Poisson process with the exception that at least one or more jumps must occur in every three or more successive intervals.

Let

$$\Gamma_i = \sum_{l=2}^{i} \eta_l \quad \text{for } i>1 \tag{V-3}$$

$$\Gamma_1 = 0,$$

where η_l may assume only the values zero or one. We will consider $\eta_l = 1$ to mean that one jump has occurred in the interval the $(l-1)^{st}$ and l^{th} observation while $\eta_l=0$ will mean that no jump has occurred in this interval. The process η is defined to be a stationary 2-<u>stage</u> Markov process (not to be confused with a 2-<u>dependent</u> process--(see Appendix A) with two states, 0 and 1. The transition function Ψ,

$$\Psi(i,j;k) = \Pr\{\ \eta_{l+2} = k \,|\, \eta_l = i \text{ and } \eta_{l+1} = j\ \} \text{ for any } l>2,$$

is specified to be of the form

$$\Psi(i,j;1) = \chi,$$

whenever i and j are not both zero, and

$$\Psi(0,0;1) = 1.$$

We have chosen this particular specification in order to have the process, η, resemble a sequence of independent, identically distributed random variables. The constant χ (a conditional probability) is, at this point, unspecified. In fact, the proper selection of χ is the critical part of the approximation.

The initial distribution ψ,

$$\psi(i,j) = \Pr\{\eta_2 = i, \eta_3 = j\},$$

is taken to be the stationary initial distribution found by solving equations

$$\psi(0,j)\,\Psi(0,j;k) + \psi(1,j)\,\Psi(1,j;k) = \psi(j,k) \text{ for } 0 \leq j,\ k \leq 1.$$

The solution is, of course, expressed in terms of χ and is given by

$$\psi(0,0) = \frac{\bar{\chi}^2}{\chi^2 - 3\chi + 3}$$

$$\psi(1,0) = \frac{\bar{\chi}}{\chi^2 - 3\chi + 3}$$

$$\psi(0,1) = \frac{\overline{\chi}}{\chi^2 - 3\chi + 3}$$

$$\psi(1,1) = \frac{\chi}{\chi^2 - 3\chi + 3} .$$

Note that

$$\Pr\{\eta_1 = 1\} = \psi(1,0) + \psi(1,1) = \frac{1}{\chi^2 - 3\chi + 3} .$$

In order to choose χ, we equate $\Pr\{\eta_1 = 1\}$ with the probability of at least one jump in an interval of length δ for a Poisson process, obtaining the relation

$$1 - e^{-\lambda\delta} = \frac{1}{\chi^2 - 3\chi + 3} , \qquad \text{(V-4)}$$

or equivalently

$$(\chi^2 - 3\chi + 3) = \frac{e^{\lambda\delta}}{e^{\lambda\delta} - 1} .$$

Here λ is the intensity of the Poisson process N. Note that (V-4) implies that, $1 - e^{-\lambda\delta}$, the probability of at least one jump in an interval of length δ, must be at least 1/3 for a solution to exist. Since $\chi^2 - 3\chi + 3$ is strictly decreasing for χ between 0 and 1, a solution in this domain will be unique when it exists.

The η process has now been completely specified, and hence so has the aggregate jump process Γ which was defined in terms of η by equation (V-3).

Finally, the approximate step-plus-jitter process, $\hat{X}$, is defined to be

$$\hat{X} = \hat{Z} + J ,$$

where

$$\hat{Z} = Y_{\Gamma}.$$

The process $\hat{X}$ is 2-dependent because if $i-j>2$, then $\hat{X}_i$ and $\hat{X}_j$ are independent, since Γ will have increased at least once between time i and time j. Also, the stationarity of $\hat{X}$ is a direct consequence of the stationarity of η

Recursive relations will be derived for computing cdp for $\hat{X}$, under the assumption that the threshold function T is constant ($T \equiv T_0$). To this end, Table V-1 gives the array of possible realizations of η for the case where n=5 (i.e., l=1,2,3,4). Note that

$$\Pr\{X_j \leq T_0, \ldots, X_k \leq T_0 \mid \text{no jumps indicated by } \Gamma \text{ between j and k}\}$$

$$= \Pr\{\hat{Z}_j + J_j \leq T_0, \ldots, \hat{Z}_j + J_k \leq T_0\}$$

$$= \int_{-\infty}^{\infty} G^{k-j+1}(T_0 - x)\, dK(x) = f_{k-j} \quad ,$$

by equation (V-2).

Each row in Table V-1 represents a different realization and the expression following each is the contribution to the nondetection probability of that particular path.

The recursive relationships for cdp, P_n, are given by writing $\overline{P}_n$ in the form

$$\overline{P}_n = \psi(1,1)\, M_n^{(11)} + \psi(1,0) M_n^{(10)} + \psi(0,1) M_n^{(01)} + \psi(0,0)\, M_n^{(00)},$$

where $\psi(i,j)\, M_n^{(ij)}$ is the contribution to $\overline{P}_n$ from all the cases where the first two values of the realizations of η are i,j. We may express $M_n^{(ij)}$ in terms of $M_{n-1}^{(ij)}$ by the relations

TABLE V-1

CONTRIBUTIONS TO cdp -- RECURSIVE SCHEME

Jump Configuration				Contribution to $\overline{P}_n$	
1	1	1	1	$\psi(1,1)\ \Psi(1,1;1)\ \Psi(1,1;1)$	$f_0f_0f_0f_0f_0$
1	1	1	0	$\psi(1,1)\ \Psi(1,1;1)\ \Psi(1,1;0)$	$f_0f_0f_0f_1$
1	1	0	1	$\psi(1,1)\ \Psi(1,1;0)\ \Psi(1,0;1)$	$f_0f_0f_1f_0$
1	1	0	0	$\psi(1,1)\ \Psi(1,1;0)\ \Psi(1,0;0)$	$f_0f_0f_2$
1	0	1	1	$\psi(1,0)\ \Psi(1,0;1)\ \Psi(0,1;1)$	$f_0f_1f_0f_0$
1	0	1	0	$\psi(1,0)\ \Psi(1,0;1)\ \Psi(0,1;0)$	$f_0f_1f_1$
1	0	0	1	$\psi(1,0)\ \Psi(1,0;0)\ \Psi(0,0;1)$	$f_0f_2f_0$
0	1	1	1	$\psi(0,1)\ \Psi(0,1;1)\ \Psi(1,1;1)$	$f_1f_0f_0f_0$
0	1	1	0	$\psi(0,1)\ \Psi(0,1;1)\ \Psi(1,1;0)$	$f_1f_0f_1$
0	1	0	1	$\psi(0,1)\ \Psi(0,1;0)\ \Psi(1,0;1)$	$f_1f_1f_0$
0	1	0	0	$\psi(0,1)\ \Psi(0,1;0)\ \Psi(1,0;0)$	f_1f_2
0	0	1	1	$\psi(0,0)\ \Psi(0,0;1)\ \Psi(0,1;1)$	$f_2f_0f_0$
0	0	1	0	$\psi(0,0)\ \Psi(0,0;1)\ \Psi(0,1;0)$	f_2f_1

$$M_n^{(1j)} = f_0\{\Psi(1,j;0)\, M_{n-1}^{(j0)} + \Psi(1,j;1)\, M_{n-1}^{(j1)}\}$$

$$M_n^{(01)} = \frac{f_1}{f_0}\ \{\Psi(0,1;0)\, M_{n-1}^{(10)} + \Psi(0,1;1)\, M_{n-1}^{(11)}\}$$

$$M_n^{(00)} = \frac{f_2}{f_1}\ \Psi(0,0;1)\, M_{n-1}^{(01)} = \frac{f_2}{f_1}\ M_{n-1}^{(01)}\ ,$$

and the initial values are

$$M_3^{(11)} = f_0^{\,3}$$

$$M_3^{(10)} = f_0 f_1$$

$$M_3^{(01)} = f_1 f_0$$

$$M_3^{(00)} = f_2\ .$$

One may easily see that these relations are valid by inspecting Table V-1 for the case of n=5.

Note that, in general, one never needs to use the values of f_j for $j>2$. This is what makes the approximation convenient since the exact expression for $\overline{P}_n$ would involve all values $f_0, \ldots, f_{n-1}$. Once the ratios f_2/f_1 and f_1/f_0 have been computed and $M_n^{(ij)}$ are known, it requires only 10 multiplications to obtain all $M_{n+1}^{(ij)}$. Thus it only requires 10 (n-3) +4 multiplications to obtain P_n from the given values of $M_3^{(ij)}$. Also one need not compute P_n separately at each stage.

Figure V-1 shows the non-detection probability $\overline{P}_n$ for a <u>step process</u> Z (no jitter present) and the corresponding approximate step process $\hat{Z}$ versus the number of trials n. The single-trial probability of failure is taken to be .8. The comparison is made for a high probability (.865) of one or more jumps occurring between trials and a low jump probability (394). Note that .865 and .394 correspond to $\lambda\delta=2$ and $\lambda\delta=\frac{1}{2}$, where $\lambda\delta$ is the expected number of jumps between trials for the Poisson process. Also note that .394 is close to the limiting value of 1/3 -- the smallest value for which the approximation is valid. (A line

FIGURE V-1

COMPARISON OF cdp FOR STEP AND APPROXIMATE STEP PROCESSES

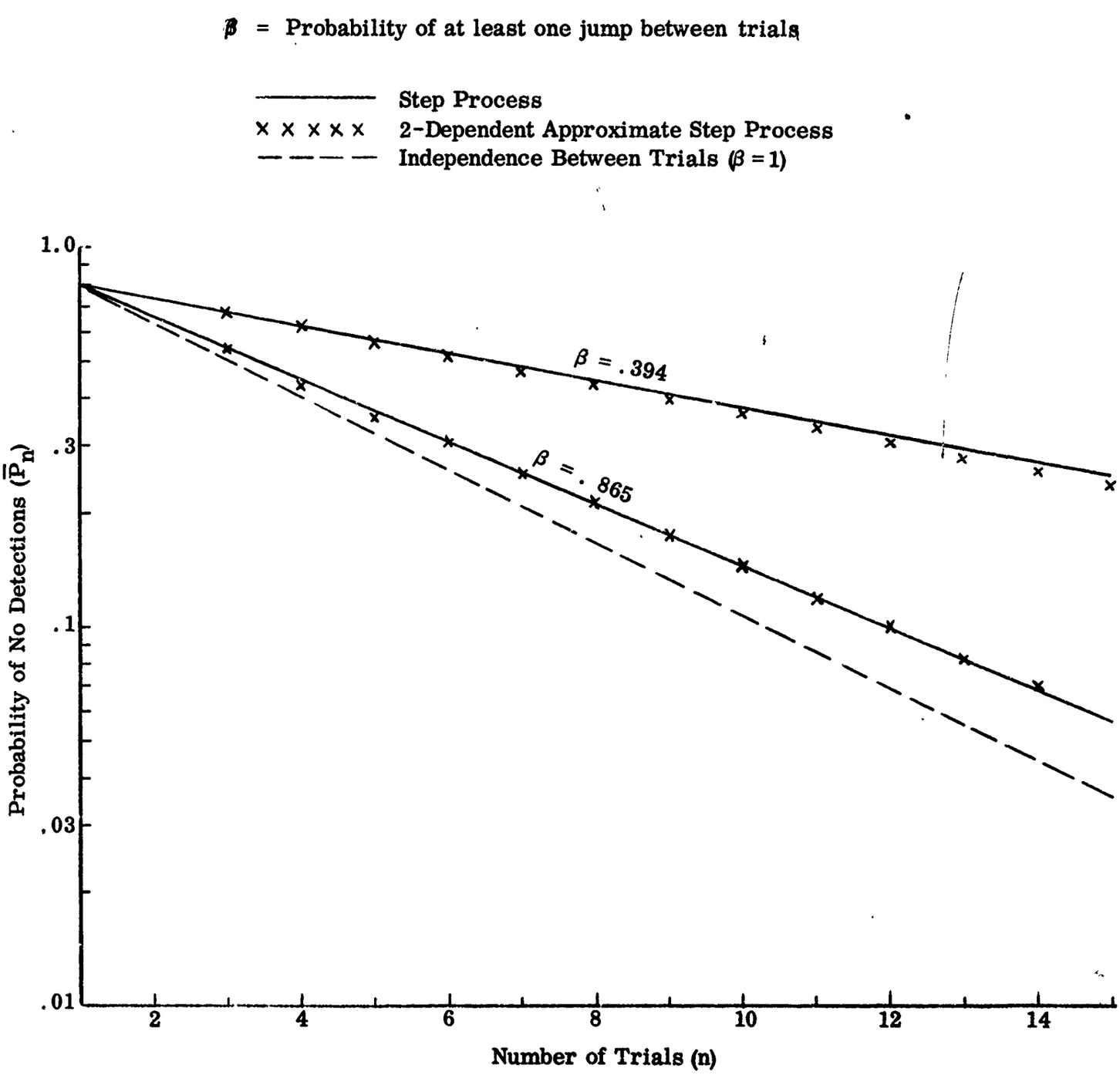


FIGURE V-2

COMPARISON OF cdp FOR STEP-PLUS-JITTER AND APPROXIMATE STEP-PLUS-JITTER PROCESSES (HIGH JUMP PROBABILITY)

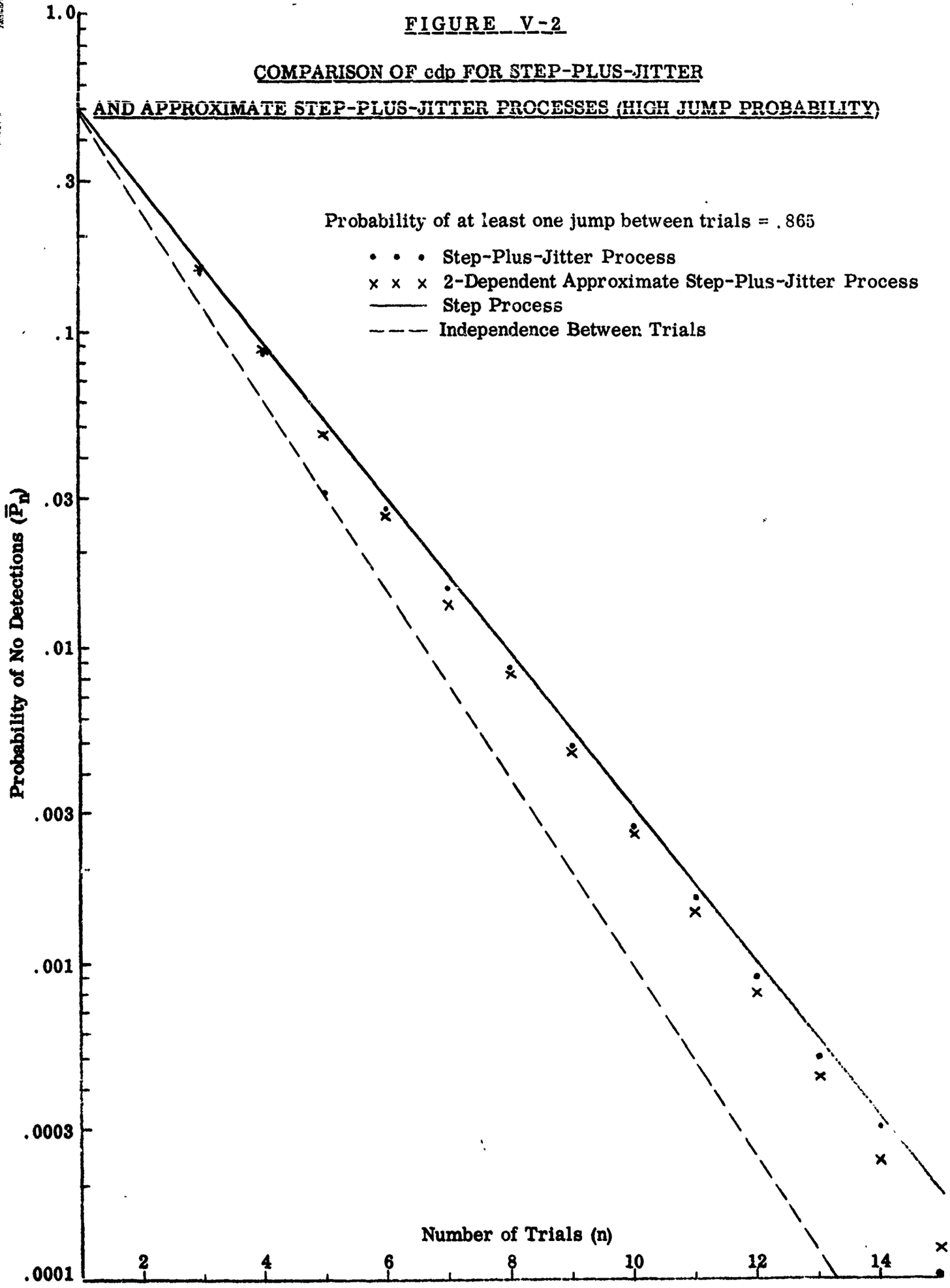
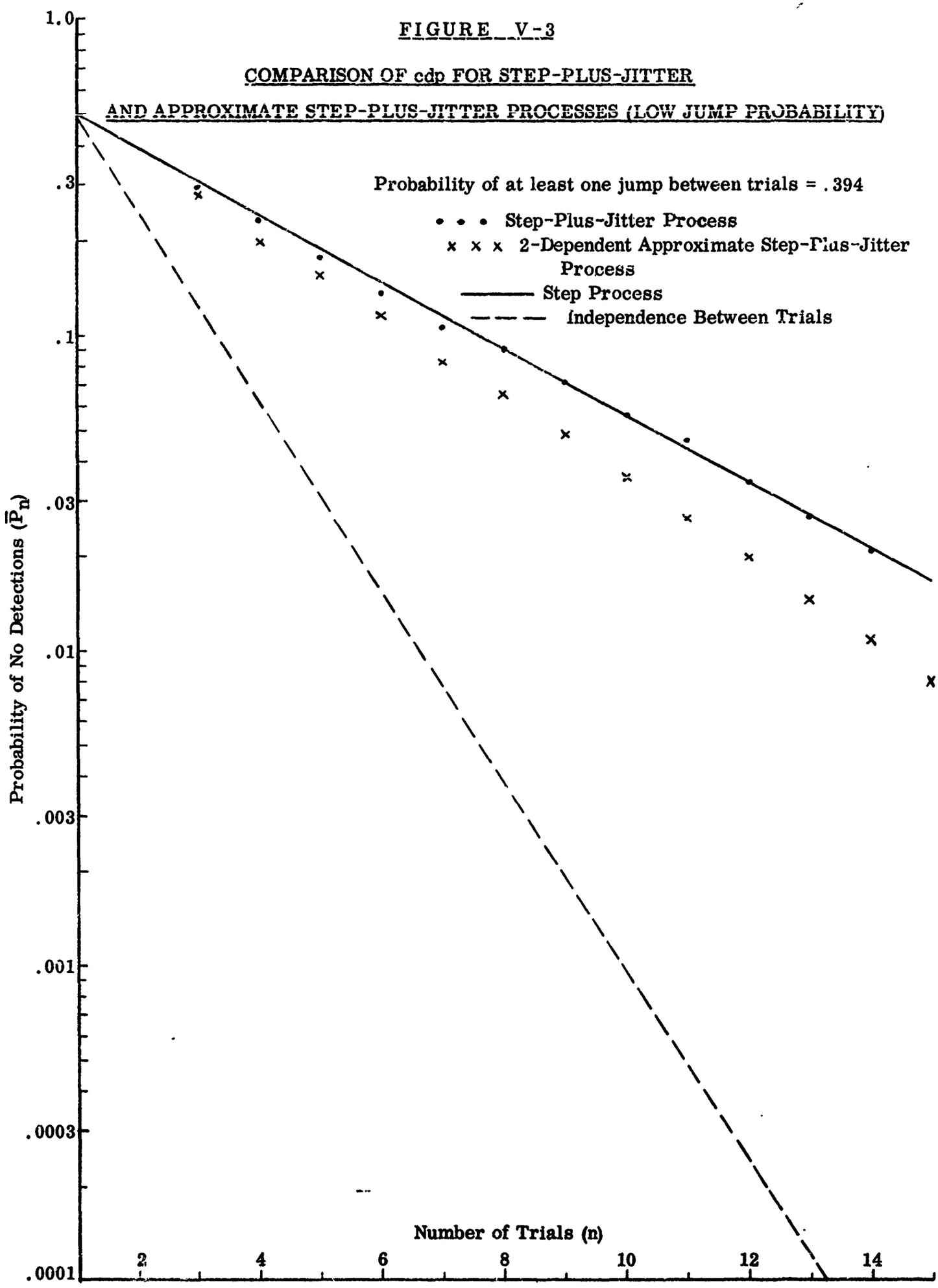


FIGURE V-3

COMPARISON OF cdp FOR STEP-PLUS-JITTER AND APPROXIMATE STEP-PLUS-JITTER PROCESSES (LOW JUMP PROBABILITY)



representing the value of $1-P_n$ under the assumption of independence between trials has been included in the figure for reference.) The formula for unimodal thresholds (Theorem IV-1, Chapter IV) was used to compute P_n for the step process. The figure shows clearly that even in the worse of the two cases (jump probability = .394), the percentage error does not exceed about 6%.

In order to test the approximation when jitter is actually present, we have made a comparison for the following special type of step-plus-jitter process. (The results appear in Figures V-3 and V-4). Let the "location-after-jump" distribution K be given by

$$K(x)=\left\{\begin{array}{ll} 0 \text{ when } x < -a & \\ \dfrac{x+a}{2a} & \text{when } -a \leq x \leq a \\ 1 \text{ when } x > a & \end{array}\right\},$$

and the jitter distribution G be given by

$$G(x)=\left\{\begin{array}{ll} 0 \text{ when } x < -c & \\ \dfrac{x+c}{2c} & \text{when } -c \leq x \leq c \\ 1 \text{ when } x > c & \end{array}\right\}.$$

Both K and G have thus been taken to be uniform distributions where the mass of K is located between -a and a and the mass of G is located between -c and c. We shall choose c to be smaller than a, as the word "jitter" implies. Now

$$f_m = \int G^{m+1}(T_0 - x)\, dK(x)$$

$$= \frac{1}{2a} \int_{-a}^{a} G^{m+1}(T_0 - x)\, dx \text{ for } m \geq 0,$$

and, assuming that $T_0 = 0$,

$$f_m = \frac{1}{2a} \int_{-a}^{a} G^{m+1}(-x)dx = \frac{a-c}{2a} + \frac{c}{(m+2)a} .$$

Note that as the jitter becomes smaller and smaller, i.e., $c \to 0$, $f_m \to \frac{1}{2}$ which corresponds to the value of f_m for the pure step process. Figures V-2 and V-3 show a comparison between $\bar{P}_n$ as computed for the step-plus-jitter process and for the approximate step-plus-jitter process in the case just described. In each figure we have included for comparison, lines representing $\bar{P}_n$ for the case of independence between trials and the case of the pure step process (i.e., without the jitter present). Figure V-2 assumes that the probability of at least one jump between trials is .865 and Figure V-3 assumes that the jump probability is .394. (These are the two cases that were shown in Figure V-1 for the case of a pure step process.) The figures indicate clearly that the approximation is much better when the probability of one or more jumps in an interval is high. This is to be expected since in this case one would expect Γ to be a better approximation to the Poisson process. Also, in the cases considered, the values of $\bar{P}_n$ as computed for the pure step process are surprisingly close to those computed for the actual step-plus-jitter process. One must be careful in drawing general conclusions, however, since the G and K distributions are of a very special type. Finally, note what a poor approximation is given by the assumption of independence between trials.

The m-dependent Approximation

In this section, the results of the preceding section are generalized to the case of an m-dependent approximation. The method of approach will be exactly the same, i.e., the Poisson arrival process is approximated by a multi-stage Markov process Γ. The approximate step-plus-jitter process is then defined in terms of Γ. On intuitive grounds, one would expect the approximation to become better with increasing m.

Once again assume that

$$\Gamma_i = \sum_{l=2}^{i} \eta_l , \text{ for } i > 1, \qquad \text{(V-5)}$$

$$\Gamma_1 = 0 ,$$

where η may assume only the values zero and one. We now define η to be a stationary m-stage Markov process. That is, the transition function Ψ depends on the past

m positions of the process. Let

$$\Psi(i_1, i_2, \ldots, i_m; j) = \Pr\{\eta_l = j \mid \eta_{l-1} = i_m, \ldots, \eta_{l-m} = i_1\}$$

and

$$\psi(i_1, \ldots, i_m) = \Pr\{\eta_2 = i_1, \eta_3 = i_2, \ldots, \eta_{m+1} = i_m\},$$

where the i's and j are either zero or one. The transition behavior of η will once again be chosen in such a way as to make η resemble a sequence of independent identically-distributed random variables, the exception being that no sequences with more than m successive zeros will be permitted. Thus, the transition function Ψ is defined by

$$\Psi(i_1, \ldots, i_m; 1) = \chi \text{ for } i_1 + \ldots + i_k \neq 0,$$

$$\Psi(0, \ldots, 0; 1) = 1.$$

The equations defining a stationary initial distribution ψ are given by

$$\sum_{k=0}^{1} \psi(k, i_1, i_2, \ldots, i_{m-1})\, \Psi(k, i_1, i_2, \ldots, i_{k-1}; i_k) = \psi(i_1, \ldots i_m).$$

The existence of such a distribution is assured by the fact that the m-tuples $\eta_l, \eta_{l+1}, \ldots, \eta_{l+m-1})$ form a vector-valued Markov process (reference [s], page 89).

Finally, we equate $\Pr\{\eta_l = 1\}$ with the probability of at least one jump in an interval of length δ for the Poisson process, and solve for the appropriate value of χ. This step is expressed by ($\sum_i$ denotes the summing over all subscripted i's).

$$\Pr\{\eta_l = 1\} = \sum_i \psi(1, i_1, \ldots, i_{m-1}) = 1 - e^{-\lambda\delta}.$$

The η process is now completely specified, and hence so is Γ by equation (V-5). The approximate step-plus-jitter process $\hat{X}$ is then defined by

$$\hat{X} = \hat{Z} + J,$$

where

$$\hat{Z} = \hat{Y}_{\Gamma}.$$

The same reasoning as in the last section shows that $\hat{X}$ is stationary and m-dependent.

The recursive relations for computing cdp are again found by expressing $\overline{P}_n$ in the form

$$\begin{aligned}
\overline{P}_n = & \sum_i \psi(1, i_1, \ldots, i_{m-1})\, M_n^{(1\, i_1 \cdots i_{m-1})} \\
& + \sum_i \psi(0, 1, i_1, \ldots, i_{m-2})\, M_n^{(0 1 i_1 \cdots i_{m-2})} \\
& + \quad \cdots \qquad\qquad\qquad\

$$M_n^{(01i_1\cdots i_{m-2})} = \frac{f_1}{f_0}\left(\Psi(0,1,i\ \ldots i_{m-2};0)M_{n-1}^{(1i_1\cdots i_{m-2}0)} + \Psi(0,1,i\ ,\ldots,i_{m-2};1)M_{n-1}^{(1i_1\cdots i_{m-2}1)}\right)$$

$$M_n^{(00\ldots1)} = \frac{f_{m-1}}{f_{m-2}}\left(\Psi(0,0,\ldots,1;0)\ M_{n-1}^{(00\ldots10)} + \Psi\ (0,0,\ldots,1;1)\ M_{n-1}^{(00\ldots11)}\right)$$

$$M_n^{(00\ldots0)} = \frac{f_m}{f_{m-1}}\ \Psi(0,0,\ldots,0;1)\ M_{n-1}^{(00\ldots01)} = \frac{f_m}{f_{m-1}}\ M_{n-1}^{(00\ldots01)}\ . \qquad \text{(V-7)}$$

These relations may be verified by examining the possible ways the jumps may occur.

The general scheme described by equations V-6 and V-7 is applicable only for finding $\overline{P}_n$ where $n>m$. A modification would, of course, be possible, but this would involve more labor than actually computing the true $\overline{P}_n$ for $n\leq m$. As it is, we still have to specify the values of $M_{m+1}^{(i_1\cdots i_m)}$ in order to initiate the recursion. These are given by noting that

$$M_{m+1}^{(i_1\cdots i_m)} = \Pr\{P_{m+1}\ |\ \eta_2=i_1,\ \eta_3=i_2,\ldots,\eta_{m+1}=i_m\},$$

which may be found by inspection. For example, if $m=6$, then

$$M_7^{(101001)} = f_0f_1f_2f_0.$$

The number of multiplications needed to find M_n(given M_{n-1} and the ratios f_{i+1}/f_j for $j=0,\ldots,m-1$) is

$$3\ (2^m-1)\ +1. \qquad \text{(V-8)}$$

It is important that n does not appear in (V-8) since this implies that the number of multiplications required to compute any $\overline{P}_n$ is just of the order of n. However, the fact that (V-8) increases rapidly with m suggests that only the approximations using small values of m would be practical for most applications.

REFERENCES

[a] Statistical Analysis of Bottom-Bounce Data from BRASS II Cruises, Daniel H. Wagner, Associates Report to U. S. Navy Underwater Sound Laboratory, by Daniel H. Wagner, Bevier Hasbrouck, and Barry Belkin, dated July 15, 1964, Confidential.

[b] Search and Screening, OEG Report 56, by B. O. Koopman, 1946, Unclassified.

[c] The Theory of Search: I. Kinematic Bases, by B. O. Koopman, "Operations Research," Vol. 4, No. 3, June 1956, pp. 324-346.

[d] The Theory of Search: II. Target Detection, by B. O. Koopman, "Operations Research," Vol. 4, No. 5, October 1956, pp. 503-531.

[e] Probabilities in a Sequence of Correlated Events, OEG Report 59, by B. O. Koopman, dated 29 April 1949, Unclassified.

[f] Characteristics of Sonar Performance as Indicated by Analysis of 1951 OPDEVFOR Trial Data, OEG Report 69, by B. O. Koopman, dated 18 May 1953, Confidential.

[g] A Law of Small Numbers in Markoff Chains, by B. O. Koopman, "Transactions of the American Mathematical Society," Vol. 70, No. 2, March 1951, pp. 277-290.

[h] Influence of Noise on SSK Effectiveness, Kettelle & Wagner Report to David Taylor Model Basin, dated December 31, 1960, Secret.

[i] Simulating Detection Capabilities by Means of Random Walks, Applied Physics Laboratory Planning Analysis Internal Note PAIN-30, by R. E. Jenkins, dated 24 July 1961.

[j] Echo Variability and the Formulation of a Radar Theory, OEG Study 644-A, by C. F. Kent, dated 9 March 1961, Unclassified.

[k] Time Variability of Underwater Acoustic Signals, NEL Report 885, by E. P. Anderson, dated 7 January 1959, Confidential.

[l] Intermittence as a Stochastic Process, by F. B. Thiess, "Proceedings of the Working Conference on the Theory of Reliability," New York University Conference Center, April 17-19, 1957.

[m] Tables of the Binomial Probability Distribution, U. S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series 6.

[n] Functions of a Markov Process That Are Markovian, by M. Rosenblatt, "Journal of Mathematics and Mechanics," Vol. 8, 1959.

[o] An Introduction to Multivariate Statistical Analysis, by T. W. Anderson, John Wiley & Sons, 1958.

[p] Tables of the Bivariate Normal Distribution Function and Related Functions, U. S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series 50.

[q] An Introduction to Probability Theory and Its Applications, by W. Feller, (Volume 1), John Wiley & Sons, Inc., 1950.

[r] Mathematical Methods of Statistics, by Harald Cramér, Princeton University Press, 1946.

[s] Stochastic Processes, by J. L. Doob, John Wiley & Sons, Inc., 1953.

[t] Random Processes, by M. Rosenblatt, Oxford University Press, 1962.

[u] Probability Integrals of Multivariate Normal and Multivariate t, by S. S. Gupta, "Annals of Mathematical Statistics," Vol. 34, 1963.

[v] Bibliography on the Multivariate Normal Integrals and Related Topics, by S. S. Gupta, "Annals of Mathematical Statistics," Vol. 34, 1963.

[w] A Note on the Evaluation of the Multivariate Normal Integral, by F. N. Davis, "Biometrika," Vol. 40, 1953, pp. 458-459.

[x] On the Range of Partial Sums of a Finite Number of Independent Normal Variates, by A. A. Anis and E. H. Lloyd, "Biometrika," Vol. 40, 1953, pp. 35-42.

[y] Approximations in the Probability Integral and Certain Percentage Points of a Multivariate Analogue of Student's t Distribution, by C. W. Dunnett and M. Sobel, "Biometrika," Vol. 42, 1955, pp. 258-260.

[z] On the Evaluation of the Probability Integral of a Multivariate Normal Distribution, by S. John, "Sankhya," Vol. 21, 1959, pp. 367-370.

[aa] A Reduction Formula for Normal Multivariate Integrals, by R. L. Plackett, "Biometrika," Vol. 41, 1954, pp. 351-360.

[bb] The One-Sided Barrier Problem for Gaussian Noise, by D. Slepian, "Bell System Tech I.," Vol. 41, 1962.

[cc] Limit Theorems for the Maximum Term in Stationary Sequences, by S. M. Berman, "Annals of Mathematical Statistics," Vol. 2, 1964.

APPENDIX A

SOME FUNDAMENTALS OF STOCHASTIC PROCESSES

This appendix discusses some elementary concepts in stochastic processes relevant to the body of the report. Although the measure-theoretic approach to probability theory has provided a solid foundation for the subject and has led to many new and important advances, an effort has been made here to present the necessary concepts without resort to measure theory. This discussion is certainly no substitute for a text. For an advanced treatment of multivariate distributions, see, for example, Cramér, reference [r]. For the more advanced subject, stochastic processes, see either Doob, reference [s], or Rosenblatt, reference [t].

The first section discusses random variables and distributions. General stochastic processes are discussed in the second section, while the third section is addressed to the important class of Markov processes.

Random Variables

An n-dimensional random variable* $X = (X_1, \ldots, X_n)$ may be regarded as a real vector which describes the outcome of some experiment which may be repeated (at least conceptually) a large number of times under uniform conditions. Associated with the random variable X is a function H of n variables, called its joint distribution function.* A function H can fill this role if, and only if it satisfies the following conditions (see reference [r], page 79).

(1) $0 \leq H(x_1, \ldots, x_n) \leq 1$

(2) $\lim_{x_i \to \infty} H(x_1, \ldots, x_n) = 0$ for $i = 1, \ldots, n$

(3) $\lim_{x_1, \ldots, x_k \to \infty} H(x_1, \ldots, x_k) = 1$

* Such an n-tuple of one-dimensional random variables is also a stochastic process (next section). The two terms refer to the same object from different viewpoints, both useful.

(4) H is non-decreasing and continuous to the right in each variable x_i

(5) The differences

$$H(x_1 + y_1, \ldots, x_n + y_n) - H(x_1, x_2 + y_2, \ldots, x_n + y_n)$$

$$- \ldots - H(x_1 + y_1, \ldots, x_{n-1} + y_{n-1}, x_n)$$

$$+ \ldots + (-1)^n H(x_1, \ldots, x_n)$$

must all be non-negative.

This function specifies the stochastic behavior of the random variable X in the sense that the probability that X_i not exceed x_i for $i = 1, \ldots, n$ is defined to be $H(x_1, \ldots, x_n)$. In symbols,

$$\Pr\{X_i \leq x_i \quad \text{for } i = 1, \ldots, n\} = H(x_1, \ldots, x_n).$$

Thus 1 - H is a cdp.

The set function Pr may be extended to a much larger class of sets than just the particular n-dimensional semi-infinite intervals $X_i \leq x_i$ for $i = 1, \ldots, n$. See reference [r] for the details of this construction.

The conditional probability that X be a member of A, given X is a member of B, is defined to be

$$\Pr\{X \in A \mid X \in B\} = \frac{\Pr\{X \in A \cap B\}}{\Pr\{X \in B\}}.$$

Suppose now that g is a real-valued function of n variables and integrable in the Lebesgue-Stieltjes sense (see reference [r]) with respect to H. Then $g(X_1, \ldots, X_n)$ is a random variable; its <u>expectation</u> is defined to be

$$E\{g(X_1, \ldots, X_n)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, \ldots, x_n)\, dH(x_1, \ldots, x_n).$$

The mean, μ_i, and the variance, σ_i^2, of X_i are defined by

$$\mu_i = E\{X_i\}$$

$$\sigma_i^2 = E\{(X_i - \mu_i)^2\},$$

when the expectations are finite (to which case we restrict ourselves). The square root, σ_i, of the variance of X_i is called its standard deviation.

The correlation coefficient ρ_{ij} between X_i and X_j is defined by

$$\rho_{ij} = \frac{E\{(X_i - \mu_i)(X_j - \mu_j)\}}{\sigma_i \sigma_j},$$

providing $\sigma_i \sigma_j \neq 0$. The concept of correlation permeates this report (Appendix C presents some miscellaneous correlation results). Note that $-1 \leq \rho_{ij} \leq 1$. If $\rho_{ij} > 0$, then X_i and X_j tend to be high together or low together, and the magnitude of ρ_{ij} is an index of this tendency. If $\rho_{ij} < 0$ (probably not important in cdp problems), then X_i tends to be high when X_j is low, and vice versa. If X_i and X_j are independent (defined below), then $\rho_{ij} = 0$, but not necessarily conversely.

The correlation coefficient between two two-state (i.e. success-failure) random variables will be denoted by r rather than ρ. It is important to realize that the correlation between success-failure variables obtained by thresholding multi-state variables does not, in general, have the same value as the correlation between the latter, as shown, for example, by Theorem C-3, Appendix C--for a non-trivial case where the two are the same, see Theorem C-4.

In the one dimensional case (n=1), the distribution function H is just a non-decreasing right-continuous function such that

$$\lim_{x \to \infty} H(x) = 1$$

$$\lim_{x \to -\infty} H(x) = 0.$$

If H is absolutely continuous, then it is expressible in the form

$$H(x) = \int_{-\infty}^{x} h(y)\, dy,$$

where h is called the <u>probability density function</u> associated with the random variable; at points x where H is differentiable (which is almost everywhere), $h(x)=H'(x)$.

A one-dimensional random variable X is said to be <u>normally distributed</u> when it has a density function h of the form

$$h(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2};$$

here σ^2 is the variance and μ the mean of X. ' density function of the normally distributed random variable with $\mu=0$ and $\sigma^2=$ called the <u>standard normal density</u> and is denoted by φ. In this case, the corresponding distribution function is called the <u>standard normal distribution</u> and is denoted by Φ.

In general, if a joint distribution function $H(x_1,\ldots,x_n)$ may be expressed in the form

$$H(x_1,\ldots,x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} h(y_1,\ldots,y_n)\, d_{y_1}\ldots d_{y_n},$$

then h is called the <u>joint density</u> function.

The random variables $X_1,\ldots,X_n$ are said to be <u>multivariate normal</u> whenever their joint distribution function H has the following density:

$$h(x) = \frac{1}{(2\pi)^{n/2}|V|^{\frac{1}{2}}} \exp\{-\tfrac{1}{2}(x-\mu)\, V^{-1}\, (x-\mu)'\}, \qquad \text{(A-1)}$$

where $x=(x_1,\ldots,x_n)$, $\mu=(\mu_1,\ldots,\mu_n)$ and V is the non-negative definite symmetric matrix of covariances $V_{ij}=\sigma_i\sigma_j\rho_{ij}$ (the prime denotes the transpose)--again μ_i and σ_i^2 are the mean and variance of X_i respectively. The symbol $|V|$ denotes the determinant of V. Strictly speaking, the distribution thus defined is called non-singular multivariate normal, owing to the nonsingularity of V. A more general definition of multivariate random variables is available which applies even when $|V| = 0$. However, in this case, the variates do not have a density function and we will not consider this case further (see reference [r], page 312). Note that $V_{ii}=\sigma_i^2$ and hence, if we wished, we could express the correlations in terms of the covariances by

$$\rho_{ij}=\frac{V_{ij}}{\sqrt{V_{ii}\,V_{jj}}}.$$

When n=2 and X_1 and X_2 both have mean zero and variance one, formula (A-1) becomes

$$h(x_1,x_2)=\frac{1}{2\pi\sqrt{1-\rho^2}}\quad\exp\left[-\frac{x_1^2-2\rho\,x_1x_2+x_2^2}{2(1-\rho^2)}\right],$$

where ρ is the correlation coefficient between X_1 and X_2. This is the bivariate normal distribution -- it is tabulated in reference [p], which is very useful in the methods of Chapter III. While there are numerous applications of the multivariate normal throughout the statistical literature, our present interest is in its use in the description and study of Gaussian processes which are defined in the next section.

In concluding this section, we define the important concept of independent random variables. This is most easily done in terms of marginal distributions. As before, let $X=(X_1,\ldots,X_n)$; then the marginal distribution H_i of X_i is defined by

$$H_i(x)=\Pr\{X_i\le x\}=H(\infty,\infty,\ldots,x,\ldots\infty),$$

where x is the ith argument of the joint distribution function H of X. The random variables $X_1,\ldots,X_n$ are then said to be independent if

$$H(x_1, \ldots, x_n) = H_1(x_1)\, H_2(x_2) \ldots H_n(x_n).$$

If X_1 and X_2 are independent, then we can express the distribution function H of their sum, $X_1 + X_2$, by the formula

$$H(x) = \int_{-\infty}^{\infty} H_1(x-y)\, d\, H_2(y),$$

where H is called the <u>convolution</u> of H_1 and H_2 (used in Chapter IV).

<u>Stochastic Processes</u>

A <u>stochastic process</u> X is simply defined to be an indexed collection of random variables X_t. We restrict each X_t to be one-dimensional unless noted otherwise. The set of indices t is called the <u>parameter set</u> of the process. In this report the parameter set is restricted to be either the half line $[0, \infty)$ in which case the process is said to have a <u>continuous parameter</u> or the positive integers, in which case the process is said to have a <u>discrete parameter</u>. In order to make this definition precise, one must specify the joint distribution of every finite sub-sequence $(X_{t_1}, \ldots, X_{t_n})$ of the random variables in a consistent manner. Let $H_{t_1 t_2 \cdots t_n}$ denote the joint distribution function of $(X_{t_1}, \ldots, X_{t_n})$; then the following conditions must be satisfied:

(1) If κ is a permutation of the integers $1, \ldots, n$, then

$$H_{t_{\kappa_1} t_{\kappa_2} \ldots t_{\kappa_n}}(x_{\kappa_1}, \ldots, x_{\kappa_n}) = H_{t_1, \ldots, t_n}(x_1, \ldots, x_n).$$

(2) If $1 \leq j \leq n$ and $x_{j+1}, \ldots, x_n$ approach infinity, then

$$H_{t_1 t_2 \ldots t_j}(x_1, \ldots, x_j) = H_{t_1 t_2 \ldots t_n}(x_1, \ldots, x_j, \infty, \ldots, \infty).$$

The set of possible values that the random variables may assume is called the state space of the process When the random variables can only assume a finite or a countable number of values, the state space is said to be discrete, and when the random variables can assume an uncountable number of values, the state space is said to be continuous. The underlying stochastic processes X^* discussed in Chapters III, IV, and V all have a continuous state space, while the two-state process D discussed in Chapter II has a discrete state space.

A Gaussian process (treated in Chapter III) is simply defined to be a stochastic process for which all the joint distributions $H_{t_1 t_2 \ldots t_n}$ are multivariate normal.

A process X is defined to be stationary if and only if

$$\Pr\{X_{t_i} \leq x_i \text{ for } i=1,\ldots,k\} = \Pr\{X_{t_i+s} \leq x_i \text{ for } i=1,\ldots,k\},$$

that is,

$$H_{t_1 \ldots t_k}(x_1,\ldots,x_k) = H_{t_1+s,\, t_2+s,\, \ldots,\, t_k+s}(x_1,\ldots,x_k).$$

One may express the content of these equations by saying that the stochastic behavior of the process is invariant under displacements in time. If the process is stationary, then, assuming that the means and covariances are defined,

(1) the mean and variance of X_t are independent of t; and

(2) the correlation $\rho_{s,t}$ between X_s and X_t depends only upon $|s-t|$.

Associated with any stationary process is the autocorrelation function α, which is defined by

$$\alpha(s) = \rho_{t,s+t},$$

whenever both t and s+t are parameters of the process. If X and W are stationary stochastic processes, with the same parameter set, then their cross-correlation function ρ_{XW} is defined by

$$\rho_{XW}(s) = \frac{E\{(X_{s+t}-E\{X_{s+t}\})(W_t-E\{W_t\})\}}{\sigma_X \sigma_W},$$

where σ_X^2 and σ_W^2 are the common variances (not zero) of the X and W processes respectively (cross-correlation arises in this report only in Theorem C-1 of Appendix C). Note that when X is a discrete-parameter process, the domain of α will be the entire set of integers, and when X is a continuous-parameter process, the domain of α and ρ_{XW} will be the entire set of real numbers. Also note that $\alpha(s) = \alpha(-s)$.

A discrete parameter process X is said to be m-dependent if X_i and X_j are independent whenever $|i-j| > m$. This type of process was useful in approximating the "step-plus-jitter" process in Chapter V. The autocorrelation function α for a stationary m-dependent process has the property that $\alpha(s) = 0$ for $|s| > m$.

Markov Processes

Finally we discuss the important class of Markov processes. A Markov process may be thought of as a process where future behavior is independent of the past, given exact knowledge of the present. We may express this formally as

$$\Pr\{A \cap B \mid X_t = x\} = \Pr\{A \mid X_t = x\} \Pr\{B \mid X_t = x\},$$

where A is an event defined by a condition on random variables with subscripts less than t and B is an event defined by a condition on random variables with subscripts greater than t. An equivalent formulation of this statement, known as the Markov property, is that

$$\Pr\{X_{t_{n+1}} \leq y \mid X_{t_n} = x_n, \ldots, X_{t_1} = x_1\} = \Pr\{X_{t_{n+1}} \leq y \mid X_{t_n} = x_n\}, \qquad \text{(A-2)}$$

for any $t_1 < t_2 < \ldots < t_{n+1}$.

Suppose for the time being that both the state space and the parameter of the process are discrete (this case is often referred to as a Markov chain). For convenience, we denote the states by the positive integers, $i=1,2,\ldots$. Let the jth transition matrix (or transition function) Ψ_j and the initial probability vector ψ be given by

$$\Psi_j(k,h) = \Pr\{X_j = h \mid X_{j-1} = k\}$$

and

$$\psi(k) = \Pr\{X_1 = k\}$$

respectively. Using the Markov property (as expressed by equation (A-2)),joint probabilities may be computed by

$$\Pr\{X_1 = k_1, \ldots, X_n = k_n\} = \psi(k_1)\,\Psi_2(k_1,k_2)\ldots\Psi_n\,(k_{n-1},k_n). \qquad \text{(A-3)}$$

From formula (A-3) one obtains formula (II-5) of Chapter II directly.

Defining, for $l > i$,

$$\Psi_{i,l}(k,h) = \Pr\{X_l = h \mid X_i = k\},$$

the matrix $\Psi_{i,l}$ gives the probabilities of transition from time i to time l and is computable by the matrix product

$$\Psi_{i,l} = \prod_{j=i+1}^{l} \Psi_j.$$

Multiplication of transition matrices can be useful in cdp problems by use of "trapping states." To illustrate, we give an alternative derivation of formula (II-5) of Chapter II in this fashion. Replace the transition matrix of Chapter II, by

$$\begin{array}{c c} & \begin{array}{cc} \text{success} & \text{failure} \end{array} \\ \begin{array}{c} \text{success} \\ \text{failure} \end{array} & \begin{pmatrix} 1 & 0 \\ b_j & \bar{b}_j \end{pmatrix} \end{array}.$$

Here "success" is a trapping state--once entered it cannot be left. If one multiplies these matrices for $j=2,\ldots,n$, then the lower right entry, $\Pi_{j=2}^{n} \bar{b}_j$, of the product is the probability that $D_2=\ldots=D_n=0$, given $D_1=0$. Thus

$$\bar{P}_n = \bar{p}_1 \prod_{j=2}^{n} \bar{b}_j ,$$

which is the formula (II-5). If the trapping state had not been used, then by multiplying transition matrices we would merely obtain relations between the first and final trials without regard for outcomes of intermediate trials.

If the transition matrices Ψ_j are independent of j, i.e., $\Psi_j = \Psi$, then the process is said to have a <u>stationary transition mechanism</u>. This will be the case, for example, when the process is stationary. However, it is quite possible for the process to be non-stationary and still have a stationary transition mechanism.

In Chapter II we have considered, in detail, the success-failure process D as being two-state Markov. There it was noted that the transition matrices Ψ_j could be conveniently expressed in terms of the correlation r_{j-1} between D_j and D_{j-1} and the single-trial probabilities p_j and p_{j-1}.

One apparent generalization of the Markov process is the notion of an m-stage <u>Markov process</u>. Here, the conditional probabilities satisfy a condition analogous to (A-2) namely,

$$\Pr\{X_{n+1}=k_{n+1} \mid X_n=k_n, \ldots, X_1=k_1\} = \Pr\{X_{n+1}=k_{n+1} \mid X_n=k_n, \ldots, X_{n-m+1}=k_{n-m+1}\}.$$

Intuitively, this means that the probability of moving into a new state is independent of history of the process more than m-steps into the past. (These processes are also called "multiple" Markov processes, reference [s], page 89, or "m-step" Markov processes, reference [t], page 60.) It is worth mentioning that the associated vector-valued process, $(W_1, W_2, \ldots)$, where

$$W_i = (X_i, X_{i+1}, \ldots, X_{i+m-1}) \text{ for } i=1,2,\ldots$$

is just an ordinary Markov process, and it is sometimes useful to use this fact. We use m-stage Markov processes in Chapter V in connection with an approximation to the "step-plus-jitter" process.

Suppose now that the Markov process has a continuous parameter. (We still assume the state space to be discrete.) Let the matrix $\Psi_{s,t}$ be defined by

$$\Psi_{s,t}(k,h) = \Pr\{X_t=h \mid X_s=k\},$$

for each pair of parameter values s, t such that $t>s$. As before, the Markov property (expressed by equation (A-2))allows us to compute joint probabilities by the relation

$$\Pr\{X_{t_1}=k_1, \ldots, X_{t_n}=k_n\} = \Pr\{X_{t_1}=k_1\} \Psi_{t_1,t_2}(k_1,k_2) \ldots \Psi_{t_{n-1},t_n}(k_{n-1},k_n).$$

The transition matrices $\Psi_{s,t}$ satisfy the equation

$$\Psi_{s,t} = \Psi_{s,\tau} \Psi_{\tau,t},$$

analogous to (A-3) for a Markov chain. When $\Psi_{s,t}$ depends only upon (t-s), the transition mechanism is said to be stationary. Again, this will be the case whenever the process itself is stationary. The Poisson process N, used in the definition of "step"

and "step-plus-jitter" processes, is an example of a discrete-state-space Markov process with a continuous parameter. Its transition matrices are given by (reference [t], page 122):

$$\Psi_{s,t}(k,h) = \begin{cases} \dfrac{[\lambda(t-s)]^{(h-k)} e^{-\lambda(t-s)}}{(h-k)!} & \text{for } h \geq k \\ 0 & \text{for } h < k. \end{cases}$$

Here the state space is assumed to be the non-negative integers $k=0, 1, \ldots,$ and $N_0 = 0$. Clearly, even though the transition mechanism is stationary, the process is not, since

$$\Pr\{N_t = k\} = \frac{[\lambda t]^k}{k!} e^{-\lambda t},$$

which depends upon t.

A discussion similar to the above may be carried out for the case of a continuous state space (see reference [t]). It is worth noting that the step processes in Chapter IV are continuous-state-space Markov processes (this fact is not used explicitly, however), and that in Chapter III, Theorems III-2 through III-4 are based on the assumption of a Gaussian Markov process. It should also be noted that the autocorrelation function, α, of a large class of stationary Markov processes (including step processes and stationary Markov Gaussian processes) is of exponential form, i.e., $\alpha(s) = e^{-\lambda s}$ (see reference [s], page 234).

APPENDIX B

LITERATURE SURVEY ON EVALUATING THE MULTIVARIATE NORMAL DISTRIBUTION

The purpose of this appendix is to survey the previously published work which has direct bearing on the problem of computing cdp for the class of Gaussian processes. As noted in Chapter I, this problem is equivalent to evaluating the multivariate normal distribution function. The paper of Gupta, reference [v], together with his extensive bibliography, reference [u], provide an excellent review of previous results on this subject and provide the basis for the discussion in this appendix.

As elsewhere, we denote by P_n the value of cdp relative to a fixed threshold function T. That is, cdp is given by

$$P_n = 1 - \Pr\{X_i \leq T_i \text{ for } i=1,\ldots,n\},$$

where $X_1,\ldots,X_n$ are multivariate normal. Unless noted otherwise, the random variables X_i will be assumed to have mean zero and variance one. This assumption causes no loss in generality, since if this is not the case, we may consider the new random variables $\hat{X}_i$, defined by

$$\hat{X}_i = \frac{X_i - \mu_i}{\sigma_i},$$

where μ_i and σ_i^2 are the mean and variance of X_i. The random variable $\hat{X}_i$ will then

have mean zero and variance one. Moreover, the correlation between $\hat{X}_i$ and $\hat{X}_j$ is the same as the correlation between X_i and X_j.

Most results relating to the exact computation of $\bar{P}_n$ fall into the following three categories:

(1) attempts to express $\bar{P}_n$ in a closed form involving tabulated functions (usually associated with some very special selection of the numbers T_i, e.g., $T_i = 0$ for $i=1,\ldots,n$.);

(2) the computation of $\bar{P}_n$ in terms of an infinite series expansion or the expression of $\bar{P}_n$ as a single definite integral; and

(3) attempts to give reduction formulas for $\bar{P}_n$, i.e., methods of expressing $\bar{P}_n$ in terms of the multivariate normal distribution for fewer than n random variables.

Clearly, results in the first category would be most satisfactory for our purposes if they were sufficiently general. However, the results in reference [v] in this category, which we now discuss, are restricted to the special case where $T_i = 0$ for $1 \leq i \leq n$. Even then there are no general closed form expressions available for $\bar{P}_n$ when $n>3$. The expressions for n=2 and n=3 are given by

$$P_2 = \frac{1}{4} + \frac{(\text{arc sin } \rho)}{2\pi} \tag{B-1}$$

and

$$P_3 = \frac{1}{8} + \frac{1}{4\pi} (\text{arc sin } \rho_{12} + \text{arc sin } \rho_{13} + \text{arc sin } \rho_{23}), \tag{B-2}$$

where ρ_{ij} is the correlation coefficient between X_i and X_j, (see reference [v], page 801). For n odd the following recursive relation is given by Davis, reference [w]:

$$\bar{P}_n = \tfrac{1}{2} [1 - \Sigma \bar{P}_1 + \Sigma \bar{P}_2 + \ldots + (-1)^{n-1} \Sigma \bar{P}_{n-1}]; \tag{B-3}$$

here $\bar{P}_j$ represents the probability that all of a given selection of j of the X's are negative, and the summation is extended over all $\binom{n}{j}$ combinations. In fact, formula (B-2) was computed making use of (B-1) and (B-3). For n larger than 3, however, the intermediate values of $\bar{P}_j$ for j even must be known, and at the present time these are not available in closed form.

If the correlation matrix is restricted to certain special types, then formulas for $\bar{P}_n$ may be found for n>3. As a case in point, suppose all the correlations between distinct random variables are equal to $\frac{1}{2}$, i.e., $\rho_{ij} = \frac{1}{2}$ for $i \neq j$. Then one has (reference [v])

$$\bar{P}_n = \frac{1}{n+1} . \qquad \text{(B-4)}$$

Another special case is where the inverse $V^{-1} = (w_{ij})$ of the covariance matrix is such that

$$w_{ij} = \left\{ \begin{array}{l} 2 \text{ if } i=j \\ -1 \text{ is } |i-j| = 1 \\ 0 \text{ otherwise} \end{array} \right\} .$$

Here, Anis and Lloyd have proved that (reference [x])

$$\bar{P}_n = (n+1)^{-3/2} .$$

Note that here the random variables X_i are <u>not</u> assumed to have variance 1, in fact for n=2, each random variable X_i has variance 2/3.

While the aforementioned results are interesting from a theoretical point of view, they are much too restrictive to be of any use at all in computing cdp in practical situations.

Results in the second category are somewhat less restrictive but, of course, their application must deal with problems of convergence or numerical integration. An example would be the following formula, reference [y], (we are no longer restricted to $T_i = 0, i=1,\ldots,n$):

$$\Pr\{X_i \le T_i \text{ for } i = 1,\ldots,n\} = \int_{-\infty}^{\infty} \left[\prod_{i=1}^{n} \Phi\left(\frac{T_i - \alpha_i y}{(1-\alpha_i^2)^{\frac{1}{2}}}\right)\right] \varphi(y)\, dy, \qquad \text{(B-5)}$$

providing the correlations ρ_{ij} are given by $\rho_{ij} = \alpha_i \alpha_j$ for $i \neq j$. Here Φ is the standard normal distribution function and φ is the standard normal density function. For more general types of correlation, there are expansions for $\overline{P}_n$ in terms of generalized tetrachoric series, reference [v]. However, these are quite complicated and converge very slowly for high values of $|\rho_{ij}|$.

The results in the third category are rather sparse and have the disadvantage of reducing the given problem to another which may still not have a satisfactory solution. For example, John, reference [z], obtains a formula for evaluating the probability integrals $\overline{P}_n$ in n-dimensions if those of (n-1)-dimensions are available. Plackett, reference [aa], also gives a reduction formula applicable for arbitrary n, but its application is quite laborious when n is large (see page 804 of reference [v]).

In addition to exact expressions, there are a number of results in finding bounds on $\overline{P}_n$ The most useful result along these lines for present purposes is the result of Slepian, reference [bb], which shows how P_n changes with the correlation ρ_{ij}. The precise statement of this theorem is as follows:

<u>Theorem B-1 (Slepian)</u>. Let $(X_1,\ldots,X_n)$ and $(\hat{X}_1,\ldots,\hat{X}_n)$ be multivariate normal with means zero, variances one, and correlations respectively ρ_{ij} and $\hat{\rho}_{ij}$. Then if $\rho_{ij} \ge \hat{\rho}_{ij}$ for all i, j the inequality

$$\Pr\{X_i \le T_i \text{ for } 1 \le i \le n\} \ge \Pr\{\hat{X}_i \le T_i \text{ for } 1 \le i \le n\}$$

holds for arbitrary $T_1,\ldots,T_n$.

We now give the following two applications of this theorem to the approximation of cdp.

Application 1. One can show, using Slepian's method, (see Berman, reference [cc]) that, when $T_i \equiv T_0$, for $i=1,\ldots,n$,

$$\frac{\partial \overline{P}_n}{\rho_{kl}} \leq \frac{1}{2\pi} \quad (1-\rho_{kl}^2)^{-\frac{1}{2}} \exp\left[- \frac{T_0^2}{(1+\rho_{kl})}\right]. \tag{B-6}$$

For convenience, let the right hand side of (B-6) be denoted by $\theta\ (T_0, \rho_{kl})$. Suppose that (as in the stationary case) ρ_{ij} depends only on $|i-j|$, and define $\rho_j = \rho_{1+j,1}$. Using the law of the mean, we have

$$\left| \Pr\{X_i \leq T_0 \text{ for } 1 \leq i \leq n\} - \Pr\{\hat{X}_i \leq T_0 \text{ for } 1 \leq i \leq n\} \right| \leq \sum_{j=1}^{n-1} |\rho_j| (n-j)\ \theta\ (T_0, |\rho_j|),$$

where $\hat{X}_i$ are independent and normally distributed. Since

$$\Pr\{\hat{X}_i \leq T_0 \text{ for } 1 \leq i \leq n\} = \Phi(T_0)^n,$$

where Φ is the cumulative normal distribution, we have a bound on $\overline{P}_n$.

Application 2. One can use Slepian's theorem in still another way. If

$$\rho_{ij} = \rho \text{ for } i \neq j, 1 \leq i,\ j \leq n,$$

then it can be shown (reference [v]) that the desired cdp is

$$\Pr\{X_i \le T_i, \text{ for } 1 \le i \le n\} = \int_{-\infty}^{\infty} [\prod_{i=1}^{n} \Phi(\frac{T_i - \rho^{\frac{1}{2}} y}{(1-\rho)^{\frac{1}{2}}})] d\Phi(y),$$

where Φ is the normal distribution function. This is a special case of equation (B-5) given earlier. Therefore, if ρ_{ij} are not constant, let

$$\rho = \min_{i \neq j} \rho_{ij} \text{ and } \rho^* = \max_{i \neq j} \rho_{ij},$$

and then

$$\int_{-\infty}^{\infty} \prod_{i=1}^{n} \Phi(\frac{T_i - \rho^{\frac{1}{2}} y}{(1-\rho)^{\frac{1}{2}}}) \, d\Phi(y) \le \Pr\{X_i \le T_i \text{ for } 1 \le i \le n\} \le \int_{-\infty}^{\infty} \prod_{i=1}^{n} \Phi(\frac{T_i - \rho^{*\frac{1}{2}} y}{(1-\rho^*)^{\frac{1}{2}}}) d\Phi(y).$$

See reference [v] for more details. If the $T_i \equiv T_0$ for $i=1, \ldots, n$, then we have the slightly simpler formula

$$\int_{-\infty}^{\infty} \Phi^n (\frac{T_0 - \rho^{\frac{1}{2}} y}{(1-\rho)^{\frac{1}{2}}}) \, d\Phi(y) \le \Pr\{X_i \le T_0 \text{ for all } i, 1 \le i \le n\} \le \int_{-\infty}^{\infty} \Phi^n (\frac{\hat{T} - \rho^{*\frac{1}{2}} y}{(1-\rho^*)^{\frac{1}{2}}}) d\Phi(y).$$

Tables of values for integrals of this latter type have been made for selected values of T_0, ρ (or ρ^*), and n in reference [v].

APPENDIX C

CORRELATION THEOREMS

In this appendix we present four theorems concerning correlation, which are relevant to cdp calculations.

Theorem C-1 gives a formula for the autocorrelation of the sum of two stochastic processes. Theorem C-2 presents elementary facts on correlation between two-state random variables. Theorems C-3 and C-4 relate the autocorrelation of a multi-state process to that of a two-state success-failure process obtained by thresholding.

Autocorrelation of a Sum

Theorem C-1 presents a formula for the autocorrelation function of a stationary stochastic process which is the sum of two (cross-correlated) stationary stochastic processes. In general this theorem is useful in synthesizing empirical estimates of correlation from estimates on additive component processes. Note that if there is no cross correlation between X and W, then a simpler statement of the theorem is that the autocovariance (variance times autocorrelation) of X +W is the sum of the autocovariances of X and W.

Theorem C-1. If X and W are stochastic processes with constant means μ_1 and μ_2, constant standard deviations σ_1 and σ_2, and autocorrelation functions α_1 and α_2 respectively, then the autocorrelation of their sum is

$$\alpha(s) = \frac{\sigma_1^2 \alpha_1(s) + \sigma_2^2 \alpha_2(s) + \sigma_1 \sigma_2 [\rho(s) + \rho(-s)]}{\sigma_1^2 + \sigma_2^2 + 2\sigma_1 \sigma_2 (\rho(0))},$$

where ρ is the cross-correlation function between X and W.

Proof. We have, denoting the variance of X + W by σ^2,

$$\sigma^2 \alpha(s) = E\{(X+W)_t \cdot (X+W)_{t+s}\} - (\mu_1 + \mu_2)^2$$

$$= E\{X_t X_{t+s}\} - \mu_1^2 + E\{W_t W_{t+s}\} - \mu_2^2$$

$$+ E\{X_t W_{t+s}\} - \mu_1 \mu_2 + E\{X_{t+s} W_t\} - \mu_1 \mu_2 .$$

It is easily shown that

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + 2\sigma_1 \sigma_2 \rho(0).$$

Since $E\{X_{t+s} W_t\} = E\{X_t W_{t-s}\}$, the theorem follows. #####

Correlation Between Two-State Variables

The following theorem presents three statements concerning the correlation of random variables over a two point sample space. Statement (ii) is used in proving Theorem C-3.

Theorem C-2. Let D_1 and D_2 be two random variables over the same two-point state space $\{0,1\}$. Let

$$p_1 = Pr\{D_1=1\}, \; p_2 = Pr\{D_2=1\}$$

and

$$z = Pr\{D_1 = 0 \text{ and } D_2 = 0\}.$$

Then

(i) $z \leq \bar{p}_1$, $z \leq \bar{p}_2$, and $\bar{p}_1 + \bar{p}_2 \leq 1 + z$;

(ii) if $0 < p_1, p_2 < 1$, then the correlation coefficient r between D_1 and D_2 is

$$r = \frac{z - \bar{p}_1\bar{p}_2}{\sqrt{p_1\bar{p}_1 p_2\bar{p}_2}};$$

(iii) if $0 < p_1, p_2 < 1$, then

$$-\sqrt{\frac{\bar{p}_1\bar{p}_2}{p_1 p_2}} \leq r \leq \sqrt{\frac{\bar{p}_1 p_2}{p_1\bar{p}_2}}$$

and

$$-\sqrt{\frac{p_1 p_2}{\bar{p}_1\bar{p}_2}} \leq r \leq \sqrt{\frac{p_1\bar{p}_2}{\bar{p}_1 p_2}}.$$

Note that the quantities in the conclusions do not depend on the state designation -- they could be arbitrary numbers and the same results hold.

Proof. Statement (i) is obvious. Proof of (ii) is straightforward computation, noting that when z, p_1, and p_2 are given, the other three joint probabilities between D_1 and D_2 are determined. Statement (iii) follows from (ii) and (i). #####

Correlation in Thresholded Processes

Theorem C-3 provides an inequality comparison between the autocorrelation of a stationary Gaussian process with that of an associated success-failure process, obtained by thresholding. Since any two observations of a Gaussian process are random variables whose joint distribution is bivariate normal, random variables rather than processes are treated in the theorem. One implication of the theorem is that the success-failure process associated with a Markov Gaussian process need not be Markov (e.g. the special case where T equals the means and equality holds-- if r is exponential, $\sin^{-1} r$ is not exponential).

Theorem C-3. Let X_1 and X_2 be a pair of random variables whose joint distribution is bivariate normal with correlation coefficient ρ. Let D_1 and D_2 be formed by collapsing X_1 and X_2 into the two states 0 and 1 by use of a threshold T:

$$D_1 = 1 \text{ if and only if } X_1 > T$$

$$D_2 = 1 \text{ if and only if } X_2 > T.$$

Let r be the correlation coefficient between D_1 and D_2. Then

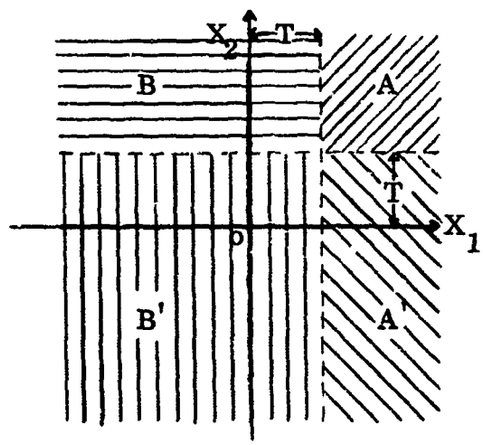
$$|r| \leq \frac{2}{\pi} \sin^{-1} |\rho| \leq |\rho|,$$

and

$$r>0 \text{ if and only if } \rho>0.$$

If $T = E\{X_1\} = E\{X_2\}$, then $|r| = (2/\pi) \sin^{-1} |\rho|$ (in this case r is sometimes called "clipper correlation."

<u>Proof.</u> We may just as well assume that X_1 and X_2 each have zero mean and unit variance. The threshold T defines the following regions A, B, A', B', in the plane:



For L a rectangular subset of the plane, let

$$F(L) = \iint_L \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\right] dxdy.$$

We then have, by Theorem C-2(ii),

$$r = 1 - \frac{F(B)}{F(A \cup B)\,[1-F(A \cup B)]}.$$

Let

$$\varphi(x) = \frac{}{\sqrt{2\pi}} e^{-x^2/2}, \qquad \Psi(z) = \int_{-z}^{z} \varphi(x)dx,$$

and

$$k = \sqrt{\frac{1-\rho}{1+\rho}}.$$

Then it is straightforward to show that

$$\frac{d}{dT} F(B) = -\varphi(T)\,\Psi(kT)$$

and

$$\frac{d}{dT} F(A \cup B) = -\varphi(T).$$

It follows that

$$\frac{dr}{dT} = \frac{\varphi(T)}{F(A \cup B)\,[1-F(A \cup B)]}\,[\Psi(kT) - (1-r)\,\Psi(T)]. \qquad \text{(C-1)}$$

Letting

$$\Omega(T,\rho) = \frac{\Psi(kT)}{\Psi(T)},$$

we see from formula (C-1) that

$$\frac{dr}{dT} > 0 \text{ if and only if } r > 1-\Omega(T,\rho).$$

We show, by contradiction, that if $\rho > 0$, and $T \geq 0$ (which implies $1-\Omega(T,\rho)$ is positive and decreasing with T), then $dr/dT < 0$. Let $\rho > 0$ and assume that, for some

$T_0 \geq 0$ and corresponding r_0.

$$\left.\frac{dr}{dT}\right|_{T_0} > 0.$$

Then $r_0 > 1-\Omega(T,\rho)$ and, since this inequality will hold for all $T > T_0$, we have $r > r_0$ for $T > T_0$. It then follows from equation (C-1) that $dr/dT \to \infty$ as $T \to \infty$, since

$$\lim_{T\to\infty} \frac{\varphi T}{F(A \cup B)} = \infty$$

by l'hospital's rule. This implies $r > 1$, a contradiction. Therefore,

$$\rho > 0 \text{ and } T \geq 0 \text{ implies } \frac{dr}{dT} \leq 0. \tag{C-2}$$

In a similar fashion, if $r \leq 0$ when $\rho > 0$ and $T \geq 0$, then $dr/dT \to \infty$ as $T \to \infty$; thus $r < -1$, in contradiction. Therefore,

$$\rho > 0 \text{ and } T \geq 0 \text{ implies } r > 0. \tag{C-3}$$

By similar contradiction arguments,

$$\rho < 0 \text{ and } T \geq 0 \text{ implies } r < 0 \text{ and } \frac{dr}{dT} > 0, \tag{C-4}$$

for, in this case, $1 - \Omega(T,\rho)$ is negative and increasing with T.

Now, by symmetry, r corresponding to T. equals r corresponding to -T; thus $|r|$ obtains it maximum at T=0 and relations counterpart to (C-2), (C-3), and (C-4) hold for $T \leq 0$.

When T=0,

$$F(B) = \frac{1}{4}\left(1 - \frac{2}{\pi}\sin^{-1}\rho\right)$$

$$F(A \cup B) = \frac{1}{2}$$

and

$$r = \frac{2}{\pi}\sin^{-1}\rho .$$

The conclusion now follows. #####

The bound on $|r|$ given by Theorem C-3 is tabulated as follows:

If $\lvert\rho\rvert$	=	0	.1	.2	.3	.5	1.0
Then $\lvert r\rvert \leq \frac{2}{\pi}\sin^{-1}\lvert\rho\rvert$	=	0	.06	.13	.19	.33	1.0

The final theorem again deals with a two-state process, this time obtained by thresholding a step process rather than a Gaussian process. The result here is more easily applied, since one finds that the autocorrelation of the two processes, the step process and the success-failure process, are equal.

Theorem C-4. Let X be a step process, with $1/\lambda$ as the mean separation between jumps. Let D be the two state (success-failure) process formed by a grouping of all values of X above a threshold T into success, and all other values of X into failure. Then the autocorrelation function associated with these processes, α_X and α_D respectively, are equal:

$$\alpha_X(s) = \alpha_D(s) = e^{-\lambda s} \text{ for } s > 0.$$

Proof. By Theorem C-2 the autocorrelation associated with D is

$$\alpha_D(s) = \frac{u(s) - p^2}{p(1-p)},$$

where

$$p = \Pr\{X > T\}$$

and

$$u(s) = \Pr\{X_t > T \text{ and } X_{t+s} > T\}.$$

Since X is a step process,

$$u(s) = p[e^{-\lambda s} + (1-e^{-\lambda s})\, p],$$

where $e^{-\lambda s}$ is the probability of no jump in X during a separation s. Therefore,

$$\alpha_D(s) = \frac{p[e^{-\lambda s} + (1-e^{-\lambda s})\, p] - p^2}{p(1-p)} = e^{-\lambda s}.$$

Now consider the autocorrelation associated with X. Let H, μ, and σ^2 be the distribution function, mean, and variance of X_t (same for all t). Then

$$E\{X_t X_{t+s}\} = \int_{-\infty}^{\infty} [x^2 e^{-\lambda s} + (1 - e^{\lambda s})\, x \int_{-\infty}^{\infty} y\, dH(y)]\, dH(x)$$

$$= \int_{-\infty}^{\infty} [x^2 e^{-\lambda s} + (1 - e^{-\lambda s})\, x\mu]\, dH(x)$$

$$= e^{-\lambda s}(\sigma^2 + \mu^2) + \mu^2(1-e^{-\lambda s})$$

$$= e^{-\lambda s}\sigma^2 + \mu^2;$$

so that

$$\alpha_X(s) = \frac{E\{(X_t-\mu)\,(X_{t+s}-\mu)\}}{\sigma^2} = \frac{E\{X_tX_s\} - \mu^2}{\sigma^2} = e^{-\lambda s}.$$

Thus, $\alpha_X = \alpha_D$. #####