RANGE ANALYSIS FOR STORAGE PROBLEMS OF PERIODIC-STOCHASTIC PROCESSES

by

JOSE D. SALAS-LA CRUZ

November 1972

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado

57

RANGE ANALYSIS FOR STORAGE PROBLEMS OF PERIODIC-STOCHASTIC PROCESSES

by
Jose D. Salas-La Cruz*

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO

November 1972 No. 57

*Post-Doctoral Research Associate, Department of Civil Engineering, Colorado State University, Fort Collins, Colorado.

TABLE OF CONTENTS

Chapter			Page
I	INTRODUCTION		1
	1.1	General Concepts	1
	1.2	Objective and General Approach in this Investigation	2
II	REVIEW OF LITERATURE		3
	2.1	Analysis of Water Storage Problems by Range	3
	2.2	Water Storage Analyzed by Other Methods	8
III	GENERAL THEORETICAL FORMULATION FOR RANGE OF PERIODIC-STOCHASTIC SERIES		11
	3.1	Stochastic Storage Difference Equation	11
		A. Characteristics of inputs and outputs	11
		B. Autocorrelation and lag cross-correlation functions of non-stationary Markov models	13
	3.2	Partial Sums	14
		A. General characteristics	14
		B. Moments of partial sums	14
		C. Marginal and joint distribution of partial sums	16
	3.3	Surplus, Deficit and Range	17
		A. General characteristics	17
		B. Distribution and moments of surplus, deficit and range	18
IV	EXACT EXPECTED VALUE OF THE RANGE		21
	4.1	Properties of Multivariate Normal Distribution Function	21
	4.2	Expected Value of Surplus of Random Variables with General Covariance Structure	24
		A. The case $n = 1$	24
		B. The case $n = 2$	24
		C. The case $n = 3$	26
	4.3	Expected Value of Range of Independent Random Variables with Changing Standard Deviation	33
	4.4	Expected Values of Range of Equally Dependent Random Variables (Exchangeable Variables)	35

4.5 Expected Values of the Range of First-Order Markov Linearly Dependent Variables 37
4.6 A Note on the Expected Value of Adjusted Range 39

V APPROXIMATE EXPECTED VALUES OF RANGE 42

5.1 Expected Values of Range of Markovian Linear Models with Periodic Autoregression Coefficients 42
5.2 Expected Values of Range of Non-stationary Exchangeable Random Variables 45
5.3 Expected Values of Range of Markov Dependent Random Variables With Periodic Standard Deviation 53

VI VARIANCES OF RANGE 57

6.1 Variance of the Range for Markov Models 57
6.2 Approximate Variance of the Range for Markov Models with Constant Standard Deviation 58
6.3 Approximate Variances of the Range for Markov Models with Periodic Mean and Periodic Standard Deviation 59

VII DESIGN OF DETERMINISTIC-STOCHASTIC STORAGE CAPACITIES 69

7.1 Deterministic and Stochastic Storage 69
7.2 Example of the Application of the Proposed Method 71

VIII CONCLUSIONS 74

REFERENCES 75

APPENDIX 77

PREFACE

Theoretical mathematical treatments of water storage problems in the application of the basic storage differential equation, in which realistic, complex periodic-stochastic processes of inputs and/or outputs and stochastic changes of storage characteristics are taken into consideration, either have not been successful or have been beyond the power of presently available analytical stochastic methods. The usual theoretical treatment has been carried out for relatively simple conditions for storage reservoirs and their inputs and outputs. Simplifications deviate so much from the real world and practical problems, that the planners of and the decision makers related to storage reservoirs have shied away from using the generalized mathematical solutions under these grossly idealized conditions.

The thesis by Jose D. Salas-La Cruz relates to the range analysis of water storage reservoirs with a relatively complex periodic-stochastic input and a simple output. It represents an attempt and successful accomplishment for increasing the power of theoretical treatment of complex hydrologic and water resource storage problems. This piece of work is a continuation of several previous efforts in the analysis of range as the major random variable of storage problems, which have been undertaken within the research project: "Stochastic Processes in Hydrology and Water Resources", sponsored by the U. S. National Science Foundation at Colorado State University, Department of Civil Engineering, Graduate and Research Hydrology and Water Resources Program The continuous analysis of the range, and other random variables related to water storage problems, promise some very significant contributions in the theoretical treatment of water reservoir systems.

When the treatment of storage problems with complex inputs and outputs becomes analytically intractible, the only approach left at present is the use of the experimental statistical (Monte Carlo) method in generating new samples of given sizes for inputs and outputs, with realistic representation of all processes involved. The simulation method permits an assessment of effects of various hydrologic complexities in solving storage problems, at least within the limits of sampling reproduction of the basic processes.

This Hydrology Paper makes a use of both methods, mathematical analytical and data generation, in determining the properties of range when inputs are complex periodic-stochastic processes. A huge gap exists at present between the mathematical theoretical solutions of water storage problems, derived under over-simplifying assumptions, and the solutions which would be obtained with realistic physical conditions of inputs, outputs, and stochastic changes inside the storage capacities. Continuous attempts are needed to make bridges between the mathematical analysis of storage problems with realistic assumptions and true solutions which would be obtained under these realistic physical conditions. The progress in finding theoretical solutions for reservoir problems may be fastest by combining the use of all methods available in obtaining the probabilistic properties of range and other random variable related to storage problems.

The results presented in this paper explain how the realistic inputs affect the key parameters of the range, with the range conceived as the needed storage capacity for regulating the inputs (given in the form of various generated samples) to produce given simple outputs for given regulating time intervals. Particularly, it is shown how the periodicity in the mean, in the standard deviation and in the autocorrelation coefficients of stochastic components of runoff input series with intervals smaller than the year, affect the expected range and the variance of the range. The data generation method can be a very useful procedure for showing planners and operators of reservoirs that the theoretical analyses of storage problems have a realistic relationship with current practical problems of design and operation of storage capacities.

September 1972

Vujica Yevjevich
Professor-in-charge of
Hydrology and Water Resources
Program, Department of
Civil Engineering

ACKNOWLEDGMENT

This paper is the Ph.D. dissertation submitted and defended by the writer in the Civil Engineering Department of Colorado State University during the spring 1972. The author wishes to extend his sincere gratitude to his adviser and major professor, Dr. V. Yevjevich, Professor of Civil Engineering for his guidance, counsel and help. Special thanks are given to Dr. D. C. Boes, Associate Professor of Statistics, whose guidance and suggestions have been invaluable in the author's work. Thanks are also expressed to Dr. H. J. Morel-Seytoux, Associate Professor of Civil Engineering, Dr. M. M. Siddiqui, Professor of Statistics and Dr. E. V. Richardson, Professor of Civil Engineering for their assistance.

Financial support in the form of research graduate assistantship, obtained under the project "Stochastic Processes in Hydrology," sponsored by the U.S. National Science Foundation, Grant No. GK-11444, is gratefully acknowledged.

ABSTRACT

The storage problem of within-the-year water fluctuations is the main topic of this paper. The storage difference equation which relates inputs, outputs and storage is used for formulating the mathematical problem. This leads to the problem of determining the expected values and variances of the range or adjusted range of cumulative departures from the population and sample mean, respectively.

Using the univariate, bivariate and trivariate normal distribution functions for the marginal and joint distributions of the partial sums, the exact expressions of the expected range are derived for $n = 1,2$ and 3. From these general expressions, particular cases of the expected range of independent and linearly dependent variables are derived. Based on these derived exact equations of the expected range, approximate equations are derived for higher values of n.

The expected value of the adjusted range of inputs equally dependent (exchangeable variables) and outputs equal to a percentage of the mean inflow, is shown to be expressed in the same way as the expected value of the unadjusted range of exchangeable random variables. This result is relevant in hydrology because when one is interested on overyear storage design and the assumption of independence of streamflow events is sufficiently accurate and the regulation or development is expressed as a fraction of the sample mean inflow, then the expected value of the storage for a given number of years is given by the expected adjusted range which now may be computed exactly by the derived equation.

The variance of the range was derived mathematically for the case of Markov first-order linearly dependent normal random variables for the case of $n = 1$ and 2. For the case of higher values n and periodic standard deviation, approximate equations are obtained by using the data generation method.

Based on mathematical approximations derived for the expected range and assuming a Markov first-order linear dependence structure of the stochastic part of monthly streamflows, a design method is developed by which the total storage is made up of two parts: (a) a deterministic storage which is a function of the standard deviation of the periodic monthly mean μ_τ and on the mean and standard deviation of the periodic monthly standard deviation σ_τ ; and (b) a stochastic storage which is a function of the mean and standard deviation of the periodic monthly standard deviation σ_τ and of the first serial correlation coefficient ρ .

Jose D. Salas-La Cruz
Civil Engineering Department
Colorado State University
Fort Collins, Colorado 80521
May, 1972

CHAPTER I

INTRODUCTION

1.1 General Concepts

Water is always controlled and regulated by a water resource system to serve a wide variety of uses. For example, water is regulated for urban use, irrigation, hydropower, navigation, recreation, water quality control, flood control, and so on. These uses may be either competitive or complementary to various degrees. This does not make the problem of design and operation of a water resource system with reservoirs a simple task.

As one example of competition, release of water for irrigation or municipal supply may impair recreational uses at the reservoir and power production. An example of complementary use may be the case of flood control with low flow augmentation. Water conflicts usually are compromised in project design. That is, trade-offs are considered in allocating the supply for different uses, which in turn require an estimate of alternative designs of a water resource system.

One of the most importants aspects of water resource systems is water regulation by reservoirs. It basically represents man's interference with the hydrologic cycle in an attempt to "balance" supply and demand. In other words, one often needs to smooth out the peaks and lows of streamflow so as to obtain a greater beneficial use of water resources.

The design of a water resource system must be viewed within the context of hydrologic risk and hydrologic and economic uncertainties. The stochastic nature of inputs and outputs of a water resource system is the reason for considering the hydrologic risk and uncertainties. The economic uncertainties are also present because the discount rate and other economic parameters are subject to uncertain changes over time. This risk and all uncertainties make it necessary to consider alternative designs to achieve developments that are optimal.

Within the past two decades, the methods for planning, design and operation of water resource systems have been changing from the use of "rules of thumb" and "engineering judgment" to a more formal type of analysis based on mathematical models. Approaches to be used in design of storage capacities may be classified into three methods: empirical, experimental (simulation or data generation), and analytical (mathematical), (Yevjevich, 1972)*.

The empirical method, known as the Rippl's diagram or mass curve is still the most commonly used method for analyzing the relationship between reservoir input, output and storage capacity. This method assumes that both input and output are known functions of time and produce the storage capacity required for no water shortage to occur during the period considered for analysis. However, the reliability of results of this analysis, based on a single sequence of hydrologic events or historical record, is limited, because it is unlikely that the same flow sequence will occur again during the life of a reservoir. In other words, another sequence of hydrologic events will require a storage capacity different from that found by using the historical record. Another disadvantage of this empirical method is in the length of historical records, which is likely to be quite different from the economic life of a dam. Since the required storage capacity for a given regulation rule increases with an increase of the length of record, the estimated capacity based on a historical record will be different from that based on the economic life of the project.

Because of the stochastic nature of streamflows and water uses, one cannot speak of the storage capacity of a reservoir in a deterministic sense. In reality, the needed capacity for a given sample size is a random variable, and it is therefore necessary to consider statistical measures such as the expected values and variances of the distribution of this variable in the design of the finite capacity of a reservoir. The data generation method approaches this problem by generating either a large number of samples of the project life size or large samples of data. This method is called, in mathematical statistics and probability theory, the Monte Carlo method. It uses independent random numbers of empirical or theoretical probability distribution functions, the time dependence structure and adds the periodic

***Name and date in parenthesis refer to the author's name and date of publication given in the bibliography.**

components when they are present in a series. This method enables one to determine approximately the moments and probability distribution functions of random variables related to storage problems.

The mathematical method consists of finding by exact, asymptotic or approximate derivations the properties of various variables related to storage capacity design, such as the mean, variance and other parameters of surplus, deficit and range. Exact general expressions for some of these properties of the range, with the range definition based on the cumulative departures from the mean, have been derived in the past only for the case of independent and identically distributed normal random variables. Similar properties are not known when the random variables are dependent and have non-stationarities.

The complexity of reservoir capacity designs depends on the type of required or proposed regulation. For example, if the regulation is of the overyear storage type, the analysis is based on annual streamflows and a given degree of river development or draft, which are usually given as a percentage of the mean inflow. In dealing with annual streamflows, the assumption of independence of events is in many cases sufficiently accurate. However, in other cases, the serial correlation between the values is significant, with Markov or linear autoregression models widely used for describing the dependence, (Yevjevich, 1964; Fiering, 1967). In many cases, annual streamflows are stationary stochastic processes; therefore the properties of the random variable of storage capacity may be derived either from exact or from approximate equations.

If the within-the-year water fluctuations are considered in the design of the reservoir storage capacity, then the analysis is usually made either with monthly, weekly or daily streamflows, or with monthly, weekly or daily outflows. In dealing with monthly values of streamflows, a non-stationary stochastic process must be considered, since time series show periodicities in the mean, standard deviation and often also in autocorrelation coefficients, besides the time dependence structure of stationary stochastic components, (Thomas and Fiering, 1962; Roesner and Yevjevich, 1966; Yevjevich, 1971). Time series of monthly outflows of reservoirs, as water use time series, also show some characteristics similar to the monthly streamflows, (Salas-LaCruz and Yevjevich, 1972). The need to deal with non-stationary series of inflows and outflows makes the general mathematical treatment of storage problems extremely complex.

1.2 Objective and General Approach in this Investigation

The storage problem of within-the-year water fluctuations is the topic of this paper. Therefore, mathematical models of monthly streamflow series are used. The main objective of this investigation is to determine mathematical equations for the expected value and variance of storage capacity needed, measured by the range values, which can be used in the design of a reservoir.

The storage difference equation which relates inputs, outputs and storage is used for formulating the mathematical problem. This leads to the problem of determining the expected values and variances of the range or adjusted range of cumulative departures from the population mean and sample mean, respectively.

Using the univariate, bivariate and trivariate normal distribution functions for the marginal and joint distributions of the partial sums, the exact expressions of the expected range are derived for $n = 1,2$ and 3. Based on these exact expressions, approximate equations are derived for the expected range for higher values of n.

The variance of the range was derived mathematically for the case of Markov first-order linearly dependent normal random variables for the case of $n = 1$ and 2. For the case of higher values of n and the standard deviation periodic, approximate equations are obtained by using the data generation method.

Based on mathematical approximations derived for the expected range and assuming a Markov first-order linear dependence structure of the stochastic part of monthly streamflows, a design method is developed by which the total storage capacity is made up of two parts: (a) a deterministic storage which is a function of the standard deviation of the periodic monthly mean and of the mean and standard deviation of the periodic monthly standard deviation; and (b) a stochastic storage which is a function of the mean and standard deviation of the periodic standard deviation and of the first serial correlation coefficient.

CHAPTER II

REVIEW OF LITERATURE

Empirical, simulation (experimental) and analytical methods have been used in the past in dealing with the analysis of reservoir storage design and operation. The empirical method proposed by W. Rippl, (1883), and somewhat modified later by many other authors, has been the most commonly used. With the development of the digital computer in the past 15 years, experimental simulation or data generation methods became attractive. Finally, mathematical analytical methods using the probability theory, mathematical statistics and stochastic process analysis have also been attempted by many authors during the last two decades, in efforts to solve the water storage differential equations under various conditions.

From a theoretical point of view, previous investigations of water storage problems may be broadly classified into two categories:

(1) Studies of reservoirs by assuming an infinite storage capacity. A great deal of research has been done along this line, and the concepts of the surplus, deficit and range of cumulative or partial sums were mainly analyzed under this assumption. The problem is, given the inflow and outflow characteristics, to find the moments and distribution of the storage capacity of a reservoir which, starting with any initial water level, would not run either empty or full in the following n years.

(2) Studies of reservoirs by assuming a finite storage capacity. The finite size of the storage capacity of the reservoir is given, and by assuming the inflow characteristics and the operating rules which determine the outflows, the problem is to find the time dependent probability function of storage levels, their limiting distribution, probabilities of water overflow and probabilities of emptiness of the finite reservoir.

Since this study considers the reservoir storage problem by assuming an infinite storage, a detailed review of previous research concerning the statistical properties of the range and adjusted range comprises the first part of this chapter. The second part presents only a review of the investigations followed mainly by P. A. P. Moran, N. U. Prabu, W. B. Langbein, E. H. Lloyd, and R. Jeng.

2.1 Analysis of Water Storage Problems by Range

Let x_i be a sequence of random variables and assume that $E(x_i) = 0$, and

$$\begin{aligned} S_i &= x_1 + x_2 + \dots + x_i \ ; i = 1,2,\dots ; n \\ M_n &= \max (0, S_1, S_2, \dots, S_n) \\ m_n &= \min (0, S_1, S_2, \dots, S_n) \\ R_n &= M_n - m_n \ . \end{aligned} \tag{2.1}$$

The random variable S_i is called the cumulative or partial sum, M_n the maximum partial sum or surplus, m_n the minimum partial sum or deficit and R_n the range of the partial sums.

In many applications, especially for small values of n , it is necessary to modify the above definitions; that is, each component of the partial sum is corrected for the estimated sample mean $\overline{x}_n$. Therefore, the above random variables will take the form

$$\begin{aligned} S_i^* &= S_i - \frac{i}{n} S_n \\ M_n^* &= \max (0, S_1^*, S_2^*, \dots, S_n^*) \\ m_n^* &= \min (0, S_1^*, S_2^*, \dots, S_n^*) \\ R_n^* &= M_n^* - m_n^* \end{aligned} \tag{2.2}$$

where S_i^* is called the adjusted partial sum, M_n^* the adjusted maximum partial sum or adjusted surplus, m_n^* the adjusted minimum partial sum or adjusted deficit and R_n^* the adjusted range. Both types of the above random variables, unadjusted and adjusted, are graphically shown in Figs. 2.1 and 2.2, respectively.

The distributions of M_n , M_n^* , m_n , m_n^* , R_n , and R_n^* are of interest in the theory of water storage and reservoir design. Assume a reservoir is of an infinite capacity which receives during every year a random streamflow input either of a symmetric or a skewed probability density function and releases the population mean discharge μ or the sample mean $\overline{x}_n$. The probability that, starting with an initial water level, the reservoir will not run dry in the following n years is given by the distribution function of R_n or R_n^*. In general, finding these exact distribution functions is a difficult mathematical problem even for cases of independent normal random inputs. Therefore, one tries to approximate these distributions by finding either their exact expected values for finite values of n or their asymptotic expected values.

After Rippl (1883) introduced the mass curve method for analyzing the relationship between the inputs, outputs, and storage capacity of a reservoir, several engineers tried to improve it. A. Hazen (1914), realizing the shortcomings of Rippl's approach, used standardized streamflow values of several rivers in order to increase the length of the historical records. He was able to test different reservoir storage capacities and evaluate the number of periods of water shortage occurring with each size. Subsequently, C. E. Sudler (1927) for the first time generated synthetic sequences by writing historical records on cards, shuffling them and then drawing a series of cards to represent a sequence of flows. Sudler's attempts were the first to use an experimental approach to approximate the stochastic nature of reservoir design and thus replaced the Rippl's and Hazen's empirical approach.

H. E. Hurst (1951), in computing the storage required for the Great Lakes of the Nile River Basin, was the first to apply more formally the concepts of probability theory to the storage problem. His method made a statistical interpretation of Rippl's approach by estimating the mean adjusted range of cumulative departures of streamflow records. He specifically used the binomial expansion for approximating the normal probability density function, and, with some concepts of combinatorial analysis, he derived the asymptotic expected adjusted range as

$$E\{R_n^*\} = \sigma\sqrt{\frac{n\pi}{2}}\,, \qquad 2.3$$

in which σ is the standard deviation and n is the length of record.

Hurst also analyzed a large number of records of annual values of natural phenomena such as rainfall, temperature, water levels, riverflows and so on. From the plots of the rescaled mean adjusted range $\overline{R}_n/\sigma_n$ against the observation length n, Hurst concluded that the observed adjusted ranges do not increase as the square root of n, but as a higher power n^c, with a mean value of c of 0.729 and a standard deviation of 0.092.

Hurst's findings led many hydrologists to propose stochastic models to account for high and low frequency effects in order to reproduce the departures from the square root law, usually called the Hurst phenomenon. However, even though Hurst analyzed a large number of records, these departures from the square root law, to the understanding of the writer, do not represent a conclusive characteristic of streamflow processes. Fiering (1967) clearly says Hurst's results are the outcome of "a jumble of distributions, record lengths, correlations and processes." Another weakness of Hurst's findings is that his slopes are based on estimated mean adjusted ranges which are highly uncertain, especially for values of $n \geq 100$. For example, for the records of around 1000 years, the mean adjusted range for n = 100 was computed by averaging 10 values, for n = 500 by averaging 2 values, and for n = 1,000 there is only one value. How can his slopes be the evidence of low frequency effects if the mean values were estimated over such small samples? The writer considers that the Hurst's results should be accepted with caution before trying to reproduce slopes which may not really represent natural characteristic of streamflow. If in the future, with more available records, Hurst's findings are substantiated, then the use of stochastic models which could reproduce slopes higher than 0.5 for n very large may be necessary, particularly if one is interested in designing reservoirs for periods of time greater than 100 years (Fiering, 1967).

W. Feller (1951) found the general expression of the probability density function of the range R(t) in continuous time. Feller assumed independent normal random variables and approximated the discrete random variables S_i with a continuously changing normal variable S(t), with mean zero and variance t. Thus, the moments of R(t) constitute the asymptotic moments of the discrete variable R_n. In particular, he obtained the asymptotic mean and asymptotic variance of the range as

$$E\{R_n\} \doteq 2\sqrt{\frac{2n}{\pi}} \approx 1.5958\, n^{1/2}\,, \qquad 2.4$$

and

$$\mathrm{Var}\{R_n\} \doteq 4n(\log 2 - 2/\pi) \approx 0.2181\, n\,. \qquad 2.5$$

By approximating the discrete random variables S_i^* with a continuously changing variable $S^*(t)$, Feller also found the expression of the exact distribution of the adjusted range $R^*(T)$ in continuous time. In particular, the asymptotic mean and asymptotic variance of R_n^* are given as

$$E\{R_n^*\} \doteq \sqrt{\frac{n\pi}{2}} \approx 1.2533\, n^{1/2} \quad , \qquad 2.6$$

and

$$\mathrm{Var}\{R_n^*\} \doteq \frac{\pi}{2}\left(\frac{\pi}{3} - 1\right) \approx 0.0741\, n \, . \qquad 2.7$$

These theoretical results also apply for cases in which the underlying distribution of the original random variables are not normal, since for large values of n the partial sums S_n or S_n^* are asymptotically normally distributed.

A. A. Anis and E. H. Lloyd (1953) gave the exact expected value of the maximum of the partial sums $S_1, S_2, \ldots S_n$ of independent normal variables with mean zero and variance unity, in the form

$$E\{M_n\} = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{n-1} i^{-1/2} \qquad 2.8$$

which leads to the expected value of the range

$$E\{R_n\} = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n-1} i^{-1/2} \, . \qquad 2.9$$

Equation 2.9 gives the asymptotic expected value of $2\sqrt{2n/\pi}$ in agreement with Feller's results.

Subsequently, A. A. Anis (1955) published the exact second moment of the maximum of the partial sums $S_1, S_2, \ldots, S_n$, for independent standard normal random variables. His equation for $n \geqslant 2$ is

$$E\{M_n^2\} = \frac{1}{2}(n+1)$$

$$+ \frac{1}{2\pi} \sum_{i=1}^{n-2} \sum_{j=1}^{i} \left[j(i-j+1) \right]^{-1/2} , \qquad 2.10$$

which gives an asymptotic second moment equal to

$$E\{M_n^2\} \doteq n - \frac{2+\sqrt{2}}{\pi}\, n^{1/2} \, . \qquad 2.11$$

A. A. Anis (1956) presented a recurrence relationship for obtaining the numerical evaluation of all the moments of the maximum of the partial sums, $S_1, S_2, \ldots, S_n$, of independent standard normal variables as

$$E\{M_{n+1}^r\} = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{n-1} i^{-1/2}\, E\{M_{n-i+1}^{r-1}\} \cdot$$

$$+ (r-1)\, n\, E\{M_{n+1}^{r-2}\} - \frac{1}{2}(r-1) \sum_{i=1}^{n-1} E\{M_{i+1}^{r-2}\} \, . \qquad 2.12$$

for $n \geqslant 2$ and $r \geqslant 3$. Therefore, by using the first two moments as given by Eqs. 2.8 and 2.10, higher order moments may be obtained from Eq. 2.12.

F. Spitzer (1956), using combinatorial analysis, published a more general result than previously obtained. Considering a sequence of independent and identically distributed random variables and $S_j = x_1 + x_2 \ldots + x_j$ and $M_j = \max(0, S_1, S_2, \ldots, S_j)$, and

$$S_j^+ = \max(0, S_j) , \qquad 2.13$$

Spitzer derived the identity

$$\sum_{j=0}^{\infty} \Phi_j(t)\, z^j = \exp\left[\sum_{j=1}^{\infty} j^{-1}\, \psi_j(t)\, z^j \right] . \qquad 2.14$$

where $\Phi_j(t)$ and $\psi_j(t)$ are the characteristic functions of M_j and S_j^+, respectively, that is

$$\Phi_j(t) = E\{\exp(i\, t\, M_j)\} \qquad 2.15$$

$$\psi_j(t) = E\{\exp(i\, t\, S_j^+)\} \qquad 2.16$$

Spitzer's equation (Eq. 2.14) is general and valid for independent and identically distributed random variables of any distribution function. From this identity, the moments of the surplus M_n may be directly obtained. For the first moment, differentiating Eq. 2.14 with respect to t, and setting $t = 0$, then

$$\sum_{j=1}^{\infty} \Phi_j'(0)\, z^j = [\sum_{j=1}^{\infty} j^{-1}\ \psi_j'(0)\, z^j]$$

$$\exp [\sum_{j=1}^{\infty} j^{-1}\ \psi_j(0)] z^j ,$$

and

$$\sum_{j=1}^{\infty} \Phi_j'(0)\, z^j = [\sum_{j=1}^{\infty} j^{-1}\ \psi_j'(0)\, z^j]\ (1 - z)^{-1} .$$

Since from Eqs. 2.15 and 2.16

$$\Phi_j'(0) = i\, E\,(M_j) \quad \text{and} \quad \psi_j'(0) = i\, E\,(S_j^+)$$

then the first moment of the surplus is

$$E\{M_n\} = \sum_{i=1}^{n} i^{-1}\ E\{S_i^+\} . \qquad 2.17$$

Similarly, differentiating Eq. 2.14 twice with respect to t and setting $t = 0$, then

$$\sum_{j=1}^{\infty} \Phi_j''(0)\, z^j = (1 - z)^{-1}\ \{\sum_{j=1}^{\infty} j^{-1}\ \psi_j''(0)\, z^j$$

$$+ [\sum_{j=1}^{\infty} j^{-1}\ \psi_j'(0)\, z^j]^2\ \} .$$

Since $\Phi_j''(0) = -\,E(M_j^2)$ and $\psi_j''(0) = -\,E(S_j^{+2})$, then the second moment of the maximum for $n \geqslant 2$ is

$$E\{M_n^2\} = \sum_{i=1}^{n} i^{-1}\ E(S_i^{+\,2})$$

$$+ \sum_{i=2}^{n} \sum_{j=1}^{i-1} j^{-1}\,(i-j)^{-1}\ E(S_j^+)\,E(S_{i-j}^+) . \qquad 2.18$$

Equations 2.17 and 2.18 are generally valid for independent and identically distributed random variables of any distribution function. Specifically, for the case of normal random variables with mean zero and variance σ^2, the partial sums S_i are also normally distributed with mean zero and variance $\mathrm{Var}\, S_i = i\sigma^2$. The expected value of S_i^+, is

$$E\{S_i^+\} = E\{\frac{1}{2}\,[S_i + |\,S_i\,|\,]\} = \int_0^{\infty} S_i\, f\,(S_i)\, dS_i$$

$$E\{S_i^+\} = \frac{1}{\sqrt{2\pi}}\,[\mathrm{Var}\ S_i\,]^{1/2} . \qquad 2.19$$

Similarly, the second moment of S_i^+ is

$$E\{S_i^{+2}\} = \frac{1}{2}\,E\{S_i^2\} + \frac{1}{2}\,E\{S_i\,|\,S_i\,|\}$$

Since for a symmetric distribution $E(S_i|S_i|) = 0$, then

$$E\{S_i^{+2}\} = \frac{1}{2}\,\mathrm{Var}\{S_i\} . \qquad 2.20$$

Substitution of Eqs. 2.19 and 2.20 into 2.17 and 2.18 leads to the expected value and second moment of the maximum of partial sums for the case of independent normal random variables. This substitution then results in:

$$E\{M_n\} = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{n} i^{-1}\,[\mathrm{Var}\{S_i\}]^{1/2} \qquad 2.21$$

and

$$E\{M_n^2\} = \frac{1}{2} \sum_{i=1}^{n} i^{-1}\ \mathrm{Var}\{S_i\}$$

$$+ \frac{1}{(2\pi)} \sum_{i=2}^{n} \sum_{j=1}^{i-1} j^{-1}\,(i-j)^{-1}\ [\mathrm{Var}\{S_j\}\,\mathrm{Var}\{S_{i-j}\}\,]^{1/2} . \qquad 2.22$$

Therefore the expected value of the range may be written as

$$E\{R_n\} = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1}\,[\mathrm{Var}\{S_i\}]^{1/2} . \qquad 2.23$$

For the particular case of standard normal random variables, the Eqs. 2.21 and 2.22 are in agreement with Eqs. 2.8 and 2.10 derived by A. A. Anis.

M. E. Solari and A. A. Anis (1957) derived the exact expected value and the second moment of the maximum of the adjusted partial sums for independent and standard normal random variables as

$$E\{M_n^*\} = \frac{1}{2}\sqrt{\frac{n}{2\pi}} \sum_{i=1}^{n} i^{-1/2} (n-i)^{1/2} , \qquad 2.24$$

and

$$E\{M_n^{*2}\} = \frac{1}{6}\left[\frac{n^2-1}{n} + \frac{\sqrt{n}}{2\pi} \sum_{i=2}^{n-1} \sum_{j=1}^{i-1} \frac{i(2i-n)}{\sqrt{j^3(n-i)(i-j)^3}}\right] , \qquad 2.25$$

which lead to asymptotic values of $\sqrt{n\pi/2}/2$ and $n/2 - \sqrt{n}$ respectively.

N. U. Prabu (1965), reviewing Moran's model for the storage, gave a non-explicit solution of the probability generating function of the maximum partial sum M_n for independent random variables, in both discrete and continuous time. M_n is defined as

$$M_{n+1} = \max(0, M_n + x_n - m) , \quad n = 0, 1, 2, \ldots , \qquad 2.26$$

with x_n the random input m the constant outflow.

For the case of input x_n of a discrete distribution function, with the probability generating function

$$K(\theta) = E\{\theta^{x_n}\}, |\theta| < 1 ,$$

Prabu gives

$$\sum_{0}^{\infty} t^n E\{\theta^{M_n}\} = \frac{1}{\theta^m - tK(\theta)} \prod_{r=1}^{m} \left(\frac{\theta - \xi_r}{1 - \xi_r}\right),$$

$$(|t| < 1 , |\theta| \leq 1) , \qquad 2.27$$

where $\xi_1, \xi_2, \ldots, \xi_m$ are the roots of the functional equation $\xi^m = tK(\xi)$, such that $|\xi_r| < 1$. If $m > m_1 = E(x_n)$, then the limiting distribution of M_n as $n \to \infty$ exists, and its probability generating function becomes

$$U(\theta) = \frac{(m - m_1)(1-\theta)}{K(\theta) - \theta^m} \prod_{r=1}^{m-1} \left(\frac{\theta - \alpha_r}{1 - \alpha_r}\right) , \qquad 2.28$$

where $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ are the roots of the equation $\alpha^m = K(\alpha)$ within the unit circle.

For the case of input x_n having a continuous distribution $K(x) = P\{x_n \leq x\}$ and the partial sum, defined as $S_n = x_o + x_1 + \ldots + x_{n-1}$, the distribution function $K_n(x) = P\{S_n \leq x\}$, $K_1(x) = K(x)$; the probability generating function of M_n is

$$\sum_{0}^{\infty} t^n E\{e^{-\theta M_n}\} = \exp\left[\sum_{1}^{\infty} \frac{t^n}{n} K_n(nm) + \sum_{1}^{\infty} \frac{(te^{\theta m})^n}{n!} \int_{nm}^{\infty} e^{-\theta x} \, dK_n(x)\right]$$

$$(|t| < 1 , R_e(\theta) > 0) . \qquad 2.29$$

Furthermore, if $m > m_1 = E(x_n)$, the limiting storage function is

$$E\{e^{-\theta M_n}\} = \exp\left[-\sum_{1}^{\infty} n^{-1} \int_{0}^{\infty} (1 - e^{-\theta x}) \, dK_n(x + nm)\right] ,$$

$$[R_e(\theta) > 0] . \qquad 2.30$$

V. Yevjevich (1965) gives a detailed analysis of applications of surplus, deficit and range in hydrology. He made a comparison of the empirical, data generation and analytical methods of obtaining statistical properties of surplus, deficit and range for values of $n = 2$ and $n = 3$. Using the data generation approach, he found the mean, variance, skewness coefficient and the distribution of the unadjusted and adjusted surplus and range for a first-order Markov process for values of n up to 50 and various values of ρ.

M. J. Melentijevich (1965) investigated the case of the range when the output is linearly dependent on storage. Using the data generation method, he gives approximate equations for the expected value and variance of the range. Approximating the storage difference equation in discrete time by the continuity equation in continuous time, and using S. Chandrasekhar's (1954) method and the Fokker-Planck partial differential equations, he also found the probability density function of the cumulative sums.

P. Sutabutra (1967) investigated the reservoir design problem for within-the-year regulation assuming a constant standard deviation for variables at various positions during the year and the first-order Markov linear model for the stochastic part of the monthly streamflow data. He separated the total storage into a deterministic storage, as a function of the periodic means of the inflow and outflow series only, and a stochastic storage, as the expected value of the range for the first order Markov model. Based on his simulation, he suggested that the expected range for the first-order Markov model may be expressed as an approximation by

$$E\{R_n\} = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1} [\mathrm{Var}\{S_i\}]^{1/2} , \qquad 2.31$$

which is the same as $E\{R_n\}$ given by Eq. 2.23.

V. Yevjevich (1967) using the data generation approach, also suggested that the expected range of linearly dependent normal variables may be expressed by Eq. 2.31. He specifically analyzed the cases of the first and second-order Markov models and the simple moving average scheme. The expected values of the range computed by Eq. 2.31 gave a close approximation to the values obtained by his simulation.

O. Ditlevsen (1969) found the asymptotic distribution function of the maximum of a stationary stochastic process in continuous time by considering the partial sums in continuous time as

$$S(t) = \int_0^t [x(t) - E(x)]\, dt , \qquad 2.32$$

and the maximum of the process $S(t)$ in continuous time defined as

$$\eta(T) = \sup_{0 \leq t \leq T} \int_0^t x(t)\, dt . \qquad 2.33$$

Assuming the case of a standard normal process, Ditlevsen found that asymptotically as $T \to \infty$,

$$F_{\eta(T)}(u) \doteq 2\Phi\{\frac{u}{[\mathrm{var}\ \omega(T)]^{1/2}}\} - 1 , \qquad 2.34$$

where

$$\omega(T) = \int_0^T x(t)\, dt .$$

J. M. Mejia (1971), using the asymptotic distribution of $\eta(T)$ as given by Ditlevsen, derived the asymptotic expected value of $\eta(T)$ or the asymptotic expected value of the range $E\{R(T)\} = 2E\{\eta(T)\}$ as

$$E\{R(T)\} \doteq \frac{4}{\sqrt{2\pi}} [\mathrm{Var}\ \omega(T)]^{1/2} \qquad 2.35$$

where the variance of $\omega(T)$ is given by

$$\mathrm{Var}\ \omega(T) = 2\, A(T)\, [T - G(T)] \qquad 2.36$$

with

$$A(T) = \int_0^T \rho(u)\, du \qquad 2.37$$

and

$$G(T) = \frac{1}{A(T)} \int_0^T u\, \rho(u)\, du , \qquad 2.38$$

where $\rho(u)$ is the autocorrelation function of the continuous stationary process $x(t)$.

2.2 Water Storage Analyzed by Other Methods

P. A. P. Moran (1954) applied the probability theory to the problem of finite water storage. Moran's model was formulated in discrete time, so that the process occurs at discrete series of time intervals. The following assumptions are made:

(1) The water input x_t is a continuous, independent and identically distributed random variable. This input is assumed to occur during the "wet season" and is stored until the "dry season" when it is released.

(2) The reservoir has a finite capacity K, and the storage at time n before the input

x flows into the reservoir is Z_t .If $Z_t + x_t > K$, an amount $Z_t + x_t - K$ will overflow, but if $Z_t + x_t \leq K$, there will be no overflow. The reservoir now contains a quantity $\min(K, Z_t + x_t)$.

(3) At time n + 1, an amount of water $m(< K)$ if $Z_t + x_t \geqslant m$ or $Z_t + x_t$ if $Z_t + x_t < m$ is released from the reservoir. The release is thus $Y_t = \min(m, Z_t + x_t)$.

From these assumptions, the storage function Z_t satisfies the recurrence relation

$$Z_{t+1} = \min(K, Z_t + x_t) - \min(m, Z_t + x_t) \quad 2.39$$

so that the random variable Z_t forms a homogeneous Markov chain.

Considering the case in which the inputs have a discrete probability distribution with $P\{x_t = j\} = g_j$, ($j = 0,1,2,...$), the Markov chain Z_t has a finite number of states 0, 1, 2, ..., K-m. Let its transition probabilities be denoted by

$$P_{ij}^{(n)} = P\{Z_t = j \mid Z_o = i\} \qquad 2.40$$

$$(i, j = 0, 1, \ldots, K-m, n \geqslant 1);$$

furthermore, let $P_{ij}^{(0)} = 1$ or 0 depend on whether $i = j$ or $i \neq j$, and also denote $P_{ij}^{(1)} = P_{ij}$. From the recurrence relation of Eq. 2.39, Moran found that the transition probability matrix $P = (P_{ij})$ may be written as

$$P = \begin{array}{c|cccccc} i \backslash j & 0 & 1 & 2 & \cdots & k-m-1 & k-m \\ \hline 0 & G_m & g_{m+1} & g_{m+2} & \cdots & g_{k-1} & h_k \\ 1 & G_{m-1} & g_m & g_{m+1} & \cdots & g_{k-2} & h_{k-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ m & G_o & g_1 & g_2 & \cdots & g_{k-m-1} & h_{k-m} \\ m+1 & 0 & g_o & g_1 & \cdots & g_{k-m-2} & h_{k-m-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ k-m & 0 & 0 & 0 & \cdots & g_{m-1} & h_m \end{array} \qquad 2.41$$

where $G_i = g_0 + g_1 + ... + g_i$, $h_i = g_i + g_{i+1} + ...$, $(i \geqslant 0)$, and it is assumed that $m < K/2$. From the above transition probability matrix, it follows

$$\sum_{n=2}^{\infty} P_{ij}^{(n)} z^n = z^2 Q_i (I - zP)^{-1} R_j , \qquad 2.42$$

where $Q_i = (P_{io}, P_{i1}, ..., P_{i,K-m})$,I is the identity matrix and $R_j = (P_{oj}, P_{1j}, ..., P_{K-m,j})^T$.

The distribution of the stationary storage was also obtained by Moran while N. U. Prabhu (1958) derived the exact solution when the inputs have geometric, negative binomial and Poisson distributions. Subsequently, A. Ghosal (1960), following Moran's storage theory, analyzed the problem of emptiness with overflow and before overflow, finding the expected values of the wet periods.

W. B. Langbein (1958) presented an application of queuing theory to the storage problem. The analogy of queuing theory with the storage problem is as follows. The inflow to the reservoir represents the arrivals, the impounded water is the queue, and the regulated outflows represent the departures. Langbein developed a procedure for determining the frequency distribution of storage, the frequency of spills, the frequency that the reservoir may be empty and the frequency distribution of reservoir outflows. He presented two kinds of solutions. The first solution was algebraic, applicable only to a linear service function and normal inflows, and the second solution gave a method termed "probability routing" when service functions are non-linear and inflows are non-normal. His procedure also allows the analysis considering monthly inflows and outflow demands.

E. H. Lloyd (1963) extended Moran's model of finite reservoirs so as to take into account the serial correlation of inflows. The assumption is made that the dependence structure of this sequence may be approximated by a homogeneous Markov chain. Using bivariate Markov processes as the joint distribution of storage and inflows, he derived the limiting distribution of storage. In another study (1963), Lloyd obtained the explicit expressions for the distribution of reservoir levels in terms of the correlation coefficient between consecutive inflows. The probabilities of emptiness and spill-over are also given. Subsequently, E. H. Lloyd and S. Odoom (1964) analyzed the case of seasonal inflows. A simple case, a two-seasonal year with three-valued input distributions, is given. The main modification they made to the non-seasonal model was to assume different distribution of inflows in each season.

R. Jeng (1967) found the probability density function of water levels in a finite storage for inflows with independent increments and outflow equal to the mean of inflow. He assumed that the inflow process was independent of storage and that the inflow varies extremely rapidly compared to variations of the storage. Under the above assumptions the storage process is a case of a one-dimensional diffusion process with zero drift, in the presence of two reflecting barriers at 0 and K with K the finite storage capacity. Using the method of image points, Jeng derived the time dependent probability density function of the water levels or storage, and also found its limiting distribution function as $t \to \infty$.

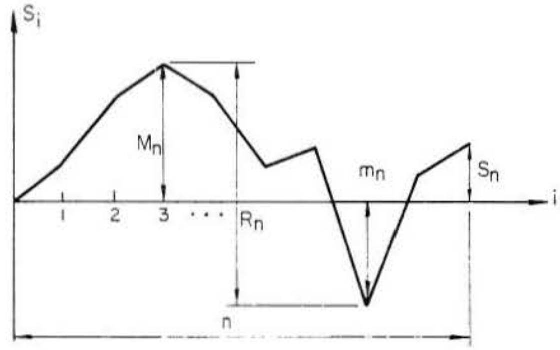

Fig. 2.1 Definition of the maximum partial sum, M_n (surplus), the minimum partial sum, m_n (deficit), and the range, R_n .

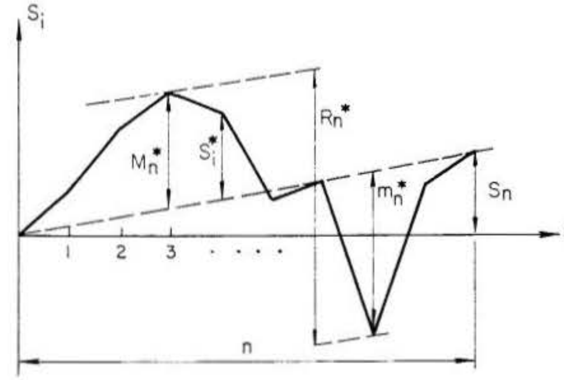

Fig. 2.2 Definition of adjusted partial sum, S_i^*, the adjusted maximum partial sum, M_n^* (adjusted surplus), the adjusted minimum partial sum, m_n^* (adjusted deficit), and the adjusted range R_n^* .

CHAPTER III

GENERAL THEORETICAL FORMULATION FOR RANGE OF PERIODIC-STOCHASTIC SERIES

A general mathematical formulation is presented in this chapter for analyzing the range problem of periodic-stochastic inputs and outputs. General characteristics of inputs and outputs commonly used in hydrology are reviewed, and some autocovariance and/or autocorrelation functions are derived for use in the following chapters. Subsequently, the general characteristics, moments and distributions of partial sums, surplus, deficit and range are reviewed.

3.1 Stochastic Storage Difference Equation

The basic relationship between inflow, outflow and storage is expressed by the difference equation

$$x_t - y_t = \frac{\Delta S}{\Delta t} \qquad 3.1$$

where x_t and y_t are the inflow and outflow respectively, and S is the storage of the reservoir. Considering the time increment of t equal to one, the above equation may be expressed as

$$x_t - y_t = S_t - S_{t-1} ,$$

or

$$S_t = S_{t-1} + (x_t - y_t) . \qquad 3.2$$

Equation 3.2 constitutes the general stochastic storage difference equation whose solution is expressed in terms of moments and probability distribution, since x_t and y_t are in general random variables. The solution of the Eq. 3.2 depends in general on the complexity of input and output, x_t and y_t respectively. They may be independent identically distributed random variables, independent but not identically distributed, dependent stationary and dependent non-stationary random variables.

A. Characteristics of inputs and outputs. In general, inputs and outputs show periodic and stochastic components and may be described by mathematical models of the form,

$$x_{p,\tau} = \mu_\tau + \sigma_\tau z_{p,\tau} , \qquad 3.3$$

$$z_{p,\tau} = \sum_{j=1}^{m} \alpha_{j,\tau-j} z_{p,\tau-j} + k_{m,\tau} \epsilon_{p,\tau} , \qquad 3.4$$

and

$$k_{m,\tau} = [1 - \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_{i,\tau-i} \alpha_{j,\tau-j} \rho_{|i-j|,\tau-\ell}]^{1/2}$$
$$[\ell = \max(i,j)] , \qquad 3.5$$

where $\tau = 1,2,...,\omega$, with ω the annual cycle (of 12 months, 52 weeks, or 365 days), $p = 1,2,...,n$, with n the number of years of record, $x_{p,\tau}$ represents the input or output series, μ_τ and σ_τ are the periodic mean and standard deviation, $\alpha_{j,\tau-j}$ are the periodic autoregression coefficients which are functions of the periodic auto correlation coefficients $\rho_{j,\tau-j}$, $z_{p,\tau}$ is a m-th order non-stationary Markov process, and $\epsilon_{p,\tau}$ is a second-order stationary and independent stochastic component.

By Fourier analysis, the periodicities in the mean, standard deviation and autocorrelation coefficients may be represented by

$$\nu_\tau = \bar{\nu}_\tau + \sum_{j=1}^{m} [A_j \cos(2\pi f_j \tau) + B_j \sin(2\pi f_j \tau)] \qquad 3.6$$

where ν_τ may represent μ_τ, σ_τ or $\rho_{k,\tau}$; $\bar{\nu}_\tau$ is the mean of ν_τ, m is the number of significant harmonics, A_j and B_j the Fourier coefficients and f_j is the frequency of the harmonic j. The estimation from the sample of the periodicities μ_τ, σ_τ, and $\rho_{k,\tau}$, and the estimation of Fourier coefficients are given elsewhere (Yevjevich, 1972).

The periodic autoregression coefficients $\alpha_{j,\tau-j}$ of the m-th order Markov model $z_{p,\tau}$ of Eq. 3.4 may be obtained by taking the expectation of the product of $z_{p,\tau}$ and $z_{p,\tau-k}$ as

$$E\{z_{p,\tau} z_{p,\tau-k}\} = \sum_{j=1}^{m} \alpha_{j,\tau-j} E\{z_{p,\tau-k} z_{p,\tau-j}\}$$
$$+ k_{m,\tau} E\{ z_{p,\tau-k} \epsilon_{p,\tau} \}.$$

Since $z_{p,\tau-k}$ and $\epsilon_{p,\tau}$ are mutually independent, with means zero and variances unity, it follows that

$$\rho_{k,\tau-k} = \sum_{j=1}^{m} \alpha_{j,\tau-j} \rho_{|j-k|,\tau-\ell} \qquad 3.7$$

with $\ell = \max(j,k), k = 1,2,\ldots,m$, the first subscript of ρ denoting the lag and the second the position in time. This expression is a system of m equations with m unknowns, $\alpha_{j,\tau-j}, j = 1,2,\ldots,m$, which may be solved as a function of autocorrelation coefficients, $\rho_{k,t-k}$. As may be noted, Eq. 3.7 is general and may be simplified to the well known recursive equation for the m-th order stationary Markov model, or with constant autoregression coefficients.

Since the first, second and third-order Markov models are most commonly used in hydrology, the autoregression coefficients for these non-stationary models can be derived from 3.7 and are

(1) For the first-order Markov model, m = 1

$$\alpha_{1,\tau-1} = \rho_{1,\tau-1} \; ; \tag{3.8}$$

(2) For the second-order Markov model, m = 2

$$\alpha_{1,\tau-1} = \frac{\rho_{1,\tau-1} - \rho_{1,\tau-2}\rho_{2,\tau-2}}{1 - \rho^2_{1,\tau-2}} \tag{3.9}$$

and

$$\alpha_{2,\tau-2} = \frac{\rho_{2,\tau-2} - \rho_{1,\tau-1}\rho_{1,\tau-2}}{1 - \rho^2_{1,\tau-2}} \tag{3.10}$$

and

(3) For the third-order Markov model, m = 3,

$$\alpha_{1,\tau-1} = \frac{\rho_{1,\tau-1}(1 - \rho^2_{1,\tau-3}) + \rho_{1,\tau-3}\rho_{1,\tau-2}\rho_{3,\tau-3} - \rho_{1,\tau-2}\rho_{2,\tau-2} - \rho_{2,\tau-3}\rho_{3,\tau-3}}{1 + 2\rho_{1,\tau-2}\rho_{2,\tau-3}\rho_{1,\tau-3} - \rho^2_{1,\tau-3} - \rho^2_{1,\tau-2} - \rho^2_{2,\tau-3}} +$$

$$+ \frac{\rho_{1,\tau-3}\rho_{2,\tau-2}\rho_{2,\tau-3}}{1 + 2\rho_{1,\tau-2}\rho_{2,\tau-3}\rho_{1,\tau-3} - \rho^2_{1,\tau-3} - \rho^2_{1,\tau-2} - \rho^2_{2,\tau-3}} \tag{3.11}$$

$$\alpha_{2,\tau-2} = \frac{\rho_{2,\tau-2}(1 - \rho^2_{2,\tau-3}) + \rho_{1,\tau-2}\rho_{2,\tau-3}\rho_{3,\tau-3} - \rho_{1,\tau-2}\rho_{1,\tau-1} - \rho_{1,\tau-3}\rho_{3,\tau-3}}{1 + 2\rho_{1,\tau-2}\rho_{2,\tau-3}\rho_{1,\tau-3} - \rho^2_{1,\tau-3} - \rho^2_{1,\tau-2} - \rho^2_{2,\tau-3}} +$$

$$+ \frac{\rho_{1,\tau-3}\rho_{2,\tau-3}\rho_{1,\tau-1}}{1 + 2\rho_{1,\tau-2}\rho_{2,\tau-3}\rho_{1,\tau-3} - \rho^2_{1,\tau-3} - \rho^2_{1,\tau-2} - \rho^2_{2,\tau-3}} \tag{3.12}$$

and

$$\alpha_{3,\tau-3} = \frac{\rho_{3,\tau-3}(1 - \rho^2_{1,\tau-2}) + \rho_{1,\tau-3}\rho_{1,\tau-2}\rho_{1,\tau-1} - \rho_{1,\tau-3}\rho_{2,\tau-2} - \rho_{2,\tau-3}\rho_{1,\tau-1}}{1 + 2\rho_{1,\tau-2}\rho_{2,\tau-3}\rho_{1,\tau-3} - \rho^2_{1,\tau-3} - \rho^2_{1,\tau-2} - \rho^2_{2,\tau-3}} +$$

$$+ \frac{\rho_{1,\tau-2}\rho_{2,\tau-2}\rho_{2,\tau-3}}{1 + 2\rho_{1,\tau-2}\rho_{2,\tau-3}\rho_{1,\tau-3} - \rho^2_{1,\tau-3} - \rho^2_{1,\tau-2} - \rho^2_{2,\tau-3}} \tag{3.13}$$

B. Autocorrelation and lag cross-correlation functions of non-stationary Markov models. Since the m-th order Markov model, as given by Eq. 3.4, is non-stationary, its covariance structure depends on the lag k and the time position t. With the subscript (p,τ) of z and ϵ variables changed to t for simplicity of notation and assuming $E\{z_t\} = 0$, then

$$\mathrm{cov}\{z_t, z_{t+k}\} = E\{z_t z_{t+k}\}.$$

Taking the expectation of the product $z_t z_{t+k}$ with z_t given by Eq. 3.4, it follows

$$\mathrm{cov}\{z_t, z_{t+k}\} = \sum_{j=1}^{m} \alpha_{j,t+k-j} \, \mathrm{cov}\{z_t, z_{t+k-j}\} .$$

Since Var z_t is constant and equal to unity the autocorrelation function for the positive lags becomes

$$\rho(k,t) = \sum_{j=1}^{m} \alpha_{j,t+k-j} \, \rho(k-j,t) \qquad (k > m), \qquad 3.14$$

where $\rho(k,t)$ and $\rho(k\text{-}j,t)$ are the two-dimensional autocorrelation functions of the lags and the positions. Similarly, the autocorrelation function for the negative lags becomes

$$\rho(k,t) = \sum_{j=1}^{m} \alpha_{j,t-j} \, \rho(-k-j, t+k) \qquad (k < -m), \qquad 3.15$$

with $\rho(o,t) = 1$, and $\rho(k,t)$ for $|k| \leqslant m$ estimated directly from data.

Equations 3.14 and 3.15 may be used recursively to obtain the autocorrelation function of the m-th order non-stationary Markov process z_t for any lag $|k| > m$ and at any time t. In particular, for the first-order Markov model, Eqs. 3.14 and 3.15 may be simplified as

$$\rho(k,t) = \prod_{i=1}^{k} \rho_{1,t+k-i} \qquad (k > 1),$$

and

$$\rho(k,t) = \prod_{i=1}^{k} \rho_{1,t-i} \qquad (k < -1), \qquad 3.16$$

with $\rho(o,t) = 1$. In the case of the stationary first-order Markov model with the coefficient of correlation $\rho_{1,t}$ a constant for every t, the above equations simplify to the well known expression $\rho(k,t) = \rho_1^k$.

For higher-order Markov models, say $m \geqslant 2$, the autocorrelation function may be obtained from the following iteration equations:

For the second-order Markov model, $m = 2$,

$$\rho(k,t) = \alpha_{1,t+k-1} \, \rho(k-1,t) + \alpha_{2,t+k-2} \, \rho(k-2,t) \qquad (k > 2) \qquad 3.17$$

with $\rho(1,t)$ and $\rho(2,t)$ replaced by $\rho_{1,t}$ and $\rho_{2,t}$ respectively, and

$$\rho(k,t) = \alpha_{1,t-1} \, \rho(-k-1, t+k) + \alpha_{2,t-2} \, \rho(-k-2, t+k) \qquad (k < -2) \qquad 3.18$$

with $\rho(-1,t+k)$ and $\rho(-2,t+k)$ replaced by $\rho_{1,t+k-1}$ and $\rho_{2,t+k-2}$ respectively.

For the third-order Markov model, $m = 3$,

$$\rho(k,t) = \alpha_{1,t+k-1} \, \rho(k-1,t) + \alpha_{2,t+k-2} \, \rho(k-2,t) + \alpha_{3,t+k-3} \, \rho(k-3,t) \qquad (k > 3) \qquad 3.19$$

with $\rho(1,t)$, $\rho(2,t)$, and $\rho(3,t)$ replaced by $\rho_{1,t}$, $\rho_{2,t}$ and $\rho_{3,t}$ respectively, and

$$\rho(k,t) = \alpha_{1,t-1} \, \rho(-k-1, t+k) + \alpha_{2,t-2} \, \rho(-k-2, t+k) + \alpha_{3,t-3} \, \rho(-k-3, t+k) \qquad (k < -3) \qquad 3.20$$

with $\rho(-1,t+k)$, $\rho(-2,t+k)$ and $\rho(-3,t+k)$ replaced by $\rho_{1,t+k-1}$, $\rho_{2,t+k-2}$ and $\rho_{3,t+k-3}$ respectively.

3.2 Partial Sums

A. General characteristics. By using Eq. 3.2 and assuming $S_o = 0$, the following sequence of partial sums is formed.

$$
\begin{aligned}
S_o &= 0 && = 0 \\
S_1 &= (x_1 - y_1) && = S_1(x) - S_1(y) \\
S_2 &= (x_1 - y_1) + (x_2 - y_2) && = S_2(x) - S_2(y) \\
&\vdots && \vdots \\
S_i &= (x_1 - y_1) + \ldots + (x_i - y_i) && = S_i(x) - S_i(y) \\
&\vdots && \vdots \\
S_n &= (x_1 - y_1) + \ldots + (x_n - y_n) && = S_n(x) - S_n(y)
\end{aligned}
\qquad 3.21
$$

where $S_i(x)$ and $S_i(y)$ denote the partial sums $x_1 + x_2 \ldots + x_i$ and $y_1 + y_2 + \ldots + y_i$, respectively. Equation 3.21 is a general representation of the partial sums, and according to the characteristics of the output y_t, for instance $y_t = \mu_x$ or $y_t = \overline{x}_n$, it may represent a sequence of unadjusted or adjusted partial sums, respectively, as are defined in Eqs. 2.1 and 2.2 of Chapter II.

Considering the general model for periodic-stochastic inputs and outputs as in Eqs. 3.3, 3.4 and 3.5, and replacing the subscript (p,τ) by t, then

$$x_t = \mu_t(x) + \sigma_t(x)\, z_t(x) \qquad 3.22$$

and

$$y_t = \mu_t(y) + \sigma_t(y)\, z_t(y) \qquad 3.23$$

with the periodic μ and σ and the z variable as defined previously. Therefore, the general term S_i of the partial sum of Eq. 3.21 may be represented by

$$S_i = \sum_{t=1}^{i} [\mu_t(x) - \mu_t(y)] + \sum_{t=1}^{i} [\sigma_t(x)\, z_t(x) - \sigma_t(y)\, z_t(y)] \,. \qquad 3.24$$

For subsequent use related to the expected values and variance of the range, it will be necessary to know the moments, and marginal and joint distribution functions of the partial sums $S_o, S_1, S_2 \ldots, S_n$.

B. Moments of partial sums. Equation 3.24 has the expected value of S_i

$$E\{S_i\} = \sum_{t=1}^{i} [\mu_t(x) - \mu_t(y)] \,. \qquad 3.25$$

For inputs and outputs stationary in the mean, Eq. 3.25 simplifies to $E\{S_i\} = 0$.

The variance of $S_i(x) = x_1 + x_2 + \ldots + x_i$ is

$$\text{Var } S_i = \sum_{t=1}^{i} \sum_{u=1}^{i} \text{cov}\,\{x_t, x_u\} \,, \qquad 3.26$$

in which the general covariance of x_t is

$$
\begin{aligned}
\text{cov}\,\{x_t, x_u\} &= E\,\{x_t\, x_u\} \\
- E\{x_t\} E\{x_u\} &= E\,\{[\mu_t(x) + \sigma_t(x)\, z_t(x)] \\
&\quad [\mu_u(x) + \sigma_u(x)\, z_u(x)]\} - \mu_t(x)\, \mu_u(x) \,,
\end{aligned}
$$

which simplifies to

$$
\begin{aligned}
\text{cov}\,\{x_t, x_u\} &= \sigma_t(x)\, \sigma_u(x)\, E\{z_t(x)\, z_u(x)\} \\
&= \sigma_t(x)\, \sigma_u(x)\, \rho_{z(x)}(u - t, t) \,,
\end{aligned}
\qquad 3.27
$$

where $\rho_{z(x)}(u\text{-}t,t)$ is the autocorrelation function of the Markov process $z_t(x)$ given in general by Eqs. 3.14 and 3.15. Substitution of Eq. 3.27 into 3.26 leads to

$$\text{Var } S_i(x) = \sum_{t=1}^{i} \sum_{u=1}^{i} \sigma_t(x)\, \sigma_u(x)\, \rho_{z(x)}(u - t, t). \qquad 3.28$$

Similarly,

$$\text{Var } S_i(y) = \sum_{t=1}^{i} \sum_{u=1}^{i} \sigma_t(y)\, \sigma_u(y)\, \rho_{z(y)}(u - t, t) \,. \qquad 3.29$$

The covariance function between $S_i(x)$ and $S_i(y)$ is

$$\text{cov}\,\{S_i(x), S_i(y)\} = \sum_{t=1}^{i} \sum_{u=1}^{i} \text{cov}\{x_t, y_u\} \qquad 3.30$$

with the general covariance of x_t and y_u

$$\mathrm{cov}\{x_t, y_u\} = E\{x_t\, y_u\} - E\{x_t\} E\{y_u\}$$

$$= E\{[\mu_t(x) + \sigma_t(x)\, z_t(x)]\, [\mu_u(y) + \sigma_u(y)\, z_u(y)]\}$$

$$- \mu_t(x)\, \mu_u(y)$$

which simplifies to

$$\mathrm{cov}\{x_t, y_u\}$$

$$= \sigma_t(x)\, \sigma_u(y)\, \rho_{z(x)\,.\,z(y)}(u-t,t); \qquad 3.31$$

with $\rho_{z(x)z(y)}(u\text{-}t,t)$ the lag cross-correlation function of the two non-stationary Markov processes $z_t(x)$ and $z_u(y)$. Substitution of Eq. 3.31 into Eq. 3.30 leads to

$$\mathrm{cov}\{S_i(x), S_i(y)\}$$

$$= \sum_{t=1}^{i} \sum_{u=1}^{i} \sigma_t(x)\, \sigma_u(y)\, \rho_{z(x).\,z(y)}(u-t,t)\,. \qquad 3.32$$

Since the variance of the partial sum $S_i = S_i(x) - S_i(y)$ may be expressed by

$$\mathrm{Var}\ S_i = \mathrm{Var}\ S_i(x) + \mathrm{Var}\ S_i(y)$$

$$- 2\,\mathrm{cov}\{S_i(x), S_i(y)\},$$

and using Eqs. 3.28, 3.29 and 3.32 then

$$\mathrm{Var}\ S_i = \sum_{t=1}^{i} \sum_{u=1}^{i} [\sigma_t(x)\, \sigma_u(x)\, \rho_{z(x)}(u-t,t)$$

$$+ \sigma_t(y)\, \sigma_u(y)\, \rho_{z(y)}(u-t,t)$$

$$- 2\, \sigma_t(x)\, \sigma_u(y)\, \rho_{z(x).\,z(y)}(u-t,t)] \qquad 3.33$$

Equation 3.33 represents the general expression for the variance of the partial sums S_i for the general case of stochastic difference equations of inputs and outputs. For subsequent applications, simplified inputs and outputs are used, so that Eq. 3.33 simplifies as

(1) For x_t independent and $y_t = \mu_x$, with μ_x the general mean of x_t, then

$$\mathrm{Var}\ S_i = \sum_{t=1}^{i} \sigma_t^2(x)\ ; \qquad 3.34$$

(2) For x_t an independent identically distributed variable with the variance σ^2 and $y_t = \mu_x$

$$\mathrm{Var}\ S_i = i\,\sigma^2\ ; \qquad 3.35$$

(3) For x_t with $\sigma_t^2(x)$ the variance of x_t; the first-order non-stationary Markov model and $y_t = \mu_x$

$$\mathrm{Var}\ S_i = \sum_{t=1}^{i} \sigma_t^2(x)$$

$$+ 2 \sum_{t=1}^{i-1} \sum_{u=1}^{i-t} \sigma_t(x)\, \sigma_{t+u}(x) \prod_{k=1}^{u} \rho_{1,t+u-k}\ ; \qquad 3.36$$

(4) For x_t the first-order Markov model with constant variance σ^2 but the periodic first autocorrelation coefficient and $y_t = \mu_x$

$$\mathrm{Var}\ S_i$$

$$= \sigma^2\ [\, i + 2 \sum_{t=1}^{i-1} \sum_{u=1}^{i-t} \prod_{k=1}^{u} \rho_{1,t+u-k}]\ ; \qquad 3.37$$

(5) For x_t the first-order stationary Markov model and $y_t = \mu_x$,

$$\mathrm{Var}\ S_i$$

$$= \frac{\sigma^2}{(1-\rho)^2}\ [\,(1-\rho^2)\, i - 2\rho\,(1-\rho^i)]\ ; \qquad 3.38$$

(6) For x_t the m-th order non-stationary Markov model and $y_t = \mu_x$

$$\mathrm{Var}\ S_i$$

$$= \sum_{t=1}^{i} \sigma_t^2(x) + 2 \sum_{t=1}^{i-1} \sum_{u=1}^{i-t} \sigma_t(x)\, \sigma_{t+u}(x)$$

$$\sum_{j=1}^{m} \alpha_{j,t+u-j}\, \rho_{z(x)}(u-j,t)\ ; \qquad 3.39$$

with $\rho_{z(x)}(u\text{-}t,t)$ given by Eq. 3.14;

(7) For x_t equally correlated ($\rho_{ij} = \rho$), with a periodic standard deviation and $y_t = \mu_x$,

$$\mathrm{Var}\ S_i$$

$$= \sum_{t=1}^{i} \sigma_t^2(x) + 2\,\rho \sum_{t=1}^{i-1} \sum_{u=1}^{i-t} \sigma_t(x)\, \sigma_{t+u}(x)\ ; \qquad 3.40$$

(8) For x_t independent and $y_t = \alpha\bar{x}_n$(*)

Var S_i^*

$$= \left(\frac{n-2i\alpha}{n}\right) \sum_{t=1}^{i} \sigma_t^2(x) + \left(\frac{i\alpha}{n}\right)^2 \sum_{t=1}^{n} \sigma_t^2(x) \ ; \qquad 3.41$$

(9) For x_t second-order stationary and independent, and $y_t = \alpha\bar{x}_n$

$$\text{Var } S_i^* = \sigma^2 \left[i - \frac{i^2\alpha}{n}(2-\alpha) \right] \ ; \qquad 3.42$$

(10) For x_t the first-order stationary Markov model and $y_t = \alpha\bar{x}_n$

$$\text{Var } S_i^* = \left(\frac{n-2i\alpha}{n}\right) \text{Var } S_i + \left(\frac{i\alpha}{n}\right)^2 \text{Var } S_n - \frac{2i\alpha\sigma^2}{n} \frac{\rho(1-\rho^{i-1})(1-\rho^{n-i})}{(1-\rho)^2} \ ; \qquad 3.43$$

with Var$\{S_i\}$ and Var$\{S_n\}$ given by Eq. 3.38, and

(11) For x_t equally correlated with a periodic standard deviation and $y_t = \alpha\bar{x}_n$

$$\text{Var } S_i^* = \left(\frac{n-2i\alpha}{n}\right) \text{Var } S_i + \left(\frac{i\alpha}{n}\right)^2 \text{Var } S_n - \frac{2i\alpha\rho}{n} \sum_{j=1}^{n-i} \sum_{t=1}^{i} \sigma_t(x)\,\sigma_{i+j}(x) \ . \qquad 3.44$$

(*) In the case when $y_t = \alpha\bar{x}_n$ with $\bar{x}_n$ the sample mean and α the level of development, the partial sums are called the adjusted partial sums and are denoted by S_i^*

C. Marginal and joint distribution of partial sums. The distribution function of the random variable S_i depends on distributions of x_t and y_t, which in turn depend on distributions of $z_t(x)$ and $z_t(y)$, respectively. If $z_t(x)$ and $z_t(y)$ are normally distributed with mean zero and variance unity, then $x_t \sim N[\mu_t(x), \sigma_t(x)]$ and $y_t \sim N[\mu_t(y), \sigma_t(y)]$. Since the sum of normal variables is also normal, the distribution of S_i is normal, with the expected value and variance given by Eqs. 3.25 and 3.33, respectively. In case the input x_t is an independent non-normal random variable and the output is $y_t = \mu_x$, the distribution of S_i is asymptotically normal for large values of i.

Since the distribution of the partial sum S_i is normal, then the joint distribution function of the sequence of partial sums $S_1, S_2, \ldots, S_i$ is multivariate normal, with means and variances given by Eqs. 3.25 and 3.33, respectively, and the autocovariance structure dependent on the means, variances and autocovariances of the components of the partial sums $(x_t - y_t)$.

For example, in the case of independent identically distributed (i.i.d.) inputs and $y_t = \mu_x$, S_i has zero mean and variance equal to $i\sigma^2$. It is easy to show for this case that the autocorrelation function of the sequence $S_1, S_2, \ldots, S_i$ is

$$\rho(k, i) = \left(\frac{i}{i+k}\right)^{1/2}, \quad \text{for } k \geqslant 0,$$

and

$$\rho(k, i) = \left(\frac{i+k}{i}\right)^{1/2}, \quad \text{for } k \leqslant 0, \qquad 3.45$$

where k denotes the lag, and i refers to the partial sum considered.

For the case of a stationary input of the first-order Markov model and the output $y_t = \mu_x$, S_i has zero mean and variance given by Eq. 3.38. Then the autocorrelation function of the sequence $S_1, S_2, \ldots, S_i$ is

$$\rho(k,i) = \frac{[(1-\rho^2)\,i - \rho\,(1-\rho^i)\,(1+\rho^k)]}{[(1-\rho^2)\,i - 2\rho\,(1-\rho^i)]^{1/2}[(1-\rho^2)\,(i+k) - 2\rho\,(1-\rho^{i+k})]^{1/2}}, \quad \text{for } k \geqslant 0$$

and 3.46

$$\rho(k,i) = \frac{[(1-\rho^2)\,(i+k) - \rho\,(1-\rho^{i+k})\,(1+\rho^{-k})]}{[(1-\rho^2)\,(i+k) - 2\rho\,(1-\rho^{i+k})]^{1/2}\,[(1-\rho^2)\,i - 2\rho\,(1-\rho^i)]^{1/2}}, \quad \text{for } k \leqslant 0$$

The sequence of random variables $S_1, S_2, \ldots, S_i$, constitutes a non-stationary process, even for the simplest case of independent identically distributed (i.i.d.), inputs, and outputs $y_t = \mu_x$. Although the mean is zero for all i's, the variance depends on i, and the autocorrelation function depends not only on the lag k, but also on i. This makes it difficult, in general, to find the properties of the maximum, minimum or the range of this sequence of partial sums for a sample of size n.

3.3 Surplus, Deficit and Range

A. General characteristics. The maximum (surplus), minimum (deficit) and range are defined in Chapter II as

$$M_n = \max(0, S_1, S_2, \ldots, S_n),$$
$$m_n = \min(0, S_1, S_2, \ldots, S_n),$$
$$R_n = M_n - m_n$$

with M_n defined as above as always positive increasing and m_n as always negative decreasing functions, while R_n is a non-decreasing function of n.

In some cases, (A. A. Anis and E. H. Lloyd, 1953; A. A. Anis, 1955 and A. A. Anis, 1956), the maximum and minimum are defined as

$$M'_n = \max(S_1, S_2, \ldots, S_n)$$
$$m'_n = \min(S_1, S_2, \ldots, S_n)$$

and the range as

$$R'_n = M'_n - m'_n$$

In this case, M'_n, m'_n and R'_n may take on either positive or negative values, although M'_n and R'_n are the increasing functions and m'_n a decreasing function as n increases.

Following E. H. Lloyd (1967), the relations between M_n and M'_n, m_n and m'_n, and R_n and R'_n may be derived as follows:

M'_n may be written as

$$M'_n = \max(S_1, S_2, \ldots, S_n)$$
$$= \max(0, S_2 - S_1, S_3 - S_1, \ldots, S_n - S_1) + S_1,$$

or

$$M'_n = \max\{0, (x_2 - y_2), (x_2 - y_2) + (x_3 - y_3), \ldots, (x_2 - y_2) + \ldots + (x_n - y_n)\} + S_1.$$

Let $w_i = x_{i+1} - y_{i+1}$; then the above expression may be written as

$$M'_n = \max\{0, w_1, w_1 + w_2, \ldots, w_1 + w_2 + \ldots + w_n\} + S_1,$$

and let $S'_i = w_1 + w_2 + w_3, \ldots, + w_i$, then

$$M'_n = \max\{0, S'_1, S'_2, \ldots, S'_{n-1}\} + S_1$$

At this point the assumption of the process $w_i = x_{i+1} - y_{i+1}$ being stationary is necessary. In this case, the distribution of S'_i is the same as the distribution of S_i. Therefore, the distribution of M'_n will depend on the distribution of M_{n-1} and S'_1.

Assuming that $E\{S_1\} = 0$, the expected value and variance of M_n' become

$$E\{M_n'\} = E\{M_{n-1}\}, \qquad 3.47$$

and

$$\mathrm{Var}\{M_n'\} = \mathrm{Var}\{M_{n-1}\} + \mathrm{Var}\{S_1\} + 2\,\mathrm{Cov}\{S_1, M_{n-1}\}. \qquad 3.48$$

Similarly, it may be shown that

$$E\{m_n'\} = E\{m_{n-1}\} \qquad 3.49$$

and

$$\mathrm{Var}\{m_n'\} = \mathrm{Var}\{m_{n-1}\} + \mathrm{Var}\{S_1\} + 2\,\mathrm{Cov}\{S_1, m_{n-1}\}. \qquad 3.50$$

The range R_n' may also be written as

$$R_n' = \max(0, S_1', S_2', \ldots, S_{n-1}') - \min(0, S_2', \ldots, S_{n-1}');$$

therefore

$$E\{R_n'\} = E\{R_{n-1}\}, \qquad 3.51$$

and

$$\mathrm{Var}\{R_n'\} = \mathrm{Var}\{R_{n-1}\}. \qquad 3.52$$

These final equations make it possible to compare the results obtained by A. A. Anis based on the sequence $S_1, S_2, \ldots, S_n$ with other results, for example those of Spitzer, based on the sequence $S_o, S_1, \ldots, S_n$ with $S_o = 0$.

B. Distribution and moments of surplus, deficit and range. Consider $F(M_n)$ and $F(m_n)$ to be the cumulative distribution functions of the surplus M_n and deficit m_n, respectively, that is

$$F(M_n) = P\{M_n \leqslant s\} \text{ and } F(m_n) = P\{m_n \leqslant s\}$$

Consider furthermore that M_n and m_n are defined as $M_n = \max(S_i, S_2, \ldots, S_n)$ and $m_n = \min(S_1, S_2, \ldots, S_n)$.

Therefore,

$$F(M_n) = P\{S_1 \leqslant s, S_2 \leqslant s, \ldots, S_n \leqslant s\}$$

or

$$F(M_n) = \int_{-\infty}^{s}, \ldots, \int_{-\infty}^{s} f(S_1, S_2, \ldots, S_n)\, dS_1\, dS_2 \ldots dS_n \qquad 3.53$$

The joint density function of $S_1, S_2, \ldots, S_n$ may be expressed as

$$f(S_1, S_2, \ldots, S_n) = f(S_1)\, f(S_2 \mid S_1)\, f(S_3 \mid S_1, S_2) \ldots f(S_n \mid S_1, S_2, \ldots, S_{n-1})$$

Therefore Eq. 3.53 becomes

$$F(M_n) = \int_{-\infty}^{s} \ldots \int_{-\infty}^{s} f(S_1)\, f(S_2 \mid S_1)\, f(S_3 \mid S_1, S_2) \ldots f(S_n \mid S_1, S_2, \ldots, S_{n-1})\, dS_1\, dS_2 \ldots dS_n \qquad 3.54$$

This equation constitutes a general expression for the distribution function of the maximum of the partial sums $S_1, S_2, \ldots, S_n$. However, unless the distribution function of S_i and their respective conditional distributions are very simple, an explicit solution for $F(M_n)$ is not possible. The best result obtained regarding the distribution of M_n was that of Spitzer (1956) which relates the characteristic functions of M_n and $S_i^+ = \max(0, S_i)$, for the case of i.i.d. variables.

Similarly, the distribution function of, m_n may in general be expressed as

$$F(m_n) = P\{m_n \leqslant s\} = 1 - P\{m_n > s\}$$

or

$$F(m_n) = 1 - P\{S_1 > s, S_2 > s, \ldots, S_n > s\}. \qquad 3.55$$

Let $Y = -m_n$, then

$$P\{Y \leqslant s\} = P\{-m_n \leqslant s\} = P\{m_n \geqslant -s\}$$

or

$$F(-m_n) = P\{S_1 \geq -s, S_2 \geq -s, \ldots, S_n \geq -s\}$$

$$F(-m_n) = \int_{-s}^{\infty} \cdots \int_{-s}^{\infty} f(S_1, S_2, \ldots, S_n)\, dS_1\, dS_2 \ldots dS_n$$

Let us consider the change of variables $s_i = -w_i$; then $F(-m_n)$ may be expressed as

$$F(-m_n) = \int_{s}^{-\infty} \cdots \int_{s}^{-\infty} f(-S_1, -S_2, \ldots, -S_n)\, d(-S_1)\, d(-S_2) \ldots d(-S_n)$$

or

$$F(-m_n) = \int_{-\infty}^{s} \cdots \int_{-\infty}^{s} f(-S_1, -S_2, \ldots, -S_n)\, dS_1\, dS_2 \ldots dS_n \qquad 3.56$$

Let us further consider at this point that the input random variables are i.i.d. with a symmetrical density function, and that the output is $y_t = \mu_x$. The joint distribution function of the sequence $S_1, S_2, \ldots, S_n$ is also symmetric,

$$f(S_1, S_2, \ldots, S_n) = f(-S_1, -S_2, \ldots, -S_n)$$

in which case Eq. 3.56 takes the form

$$F(-m_n) = \int_{-\infty}^{s} \cdots \int_{-\infty}^{s} f(S_1, S_2, \ldots, S_n)\, dS_1\, dS_2 \ldots dS_n\,. \qquad 3.57$$

Finally, comparing Eqs. 3.53 and 3.57, then

$$F(M_n) = F(-m_n) \qquad 3.58$$

This result is useful because the moments of the maximum and the minimum of partial sums may be shown to be related as

$$E\{M_n^r\} = (-1)^r\, E\{m_n^r\}\,, \qquad 3.59$$

and in particular the mean and variances are related as

$$E\{M_n\} = -E\{m_n\} \qquad 3.60$$

and

$$\mathrm{Var}\{M_n\} = \mathrm{Var}\{m_n\}\,. \qquad 3.61$$

The distribution function of the range R_n depends on the joint distribution of M_n and m_n. That is

$$F(R_n) = P\{R_n \leq r\} = P\{M_n - m_n \leq r\}\,,$$

or

$$F(R_n) = \int_{-\infty}^{\infty} P\{M_n - m_n \leq r \mid M_n\}\, f(M_n)\, dM_n\,,$$

or

$$F(R_n) = \int_{-\infty}^{\infty} P\{m_n \geq M_n - r \mid M_n\}\, f(M_n)\, dM_n\,;$$

since $P\{m_n \geq M_n - r|M_n\} = 1 - P\{m_n \leq M_n - r|M_n\}$, then $F(R_n)$ may be expressed as

$$F(R_n) = 1 - \int_{-\infty}^{\infty} P\{m_n \leq M_n - r \mid M_n)\, f(M_n)\, dM_n \qquad 3.62$$

The problem is that finding explicitly the joint distribution of M_n and m_n is very difficult, because even the marginal distributions of M_n and m_n cannot be represented in explicit form. V. Yevjevich (1965) found by numerical integration the distribution functions of the surplus, M_n, deficit m_n and range R_n for the case of inputs i.i.d. normal variables and output $y_t = \mu_x$ for values of n of 1, 2, and 3.

The moments of the range, surplus and deficit are related as follows;

$$E\{R_n\} = E\{M_n\} - E\{m_n\} \qquad 3.63$$

For the particular case in which the distribution of components of the partial sums is symmetrical, Eq. 3.60 applies, so that

$$E\{R_n\} = 2E\{M_n\}\,. \qquad 3.64$$

Similarly, the variance of the range is

$$\mathrm{Var}\{R_n\} = \mathrm{Var}\{M_n\} + \mathrm{Var}\{m_n\} - 2\,\mathrm{Cov}\{M_n, m_n\}$$

or

$$\mathrm{Var}\{R_n\} = \mathrm{Var}\{M_n\} + \mathrm{Var}\{m_n\} - 2\,\mathrm{Var}^{1/2}\{M_n\}\mathrm{Var}^{1/2}\{m_n\}\,\rho(M_n, m_n) \qquad 3.65$$

where $\rho(M_n, m_n)$ is the correlation between M_n and m_n as functions of n. For the particular case of symmetric distribution of the components of partial sums, Eq. 3.61 applies; therefore, Eq. 3.65 is simplified to

$$\mathrm{Var}\{R_n\} = 2\,\mathrm{Var}\{M_n\}\,[1 - \rho(M_n, m_n)] \qquad 3.66$$

CHAPTER IV

EXACT EXPECTED VALUE OF THE RANGE

The theoretical expected values of the range for $n = 1,2$, and 3 are developed in this chapter, considering in general that the joint distribution function of the sequence of partial sums is multivariate normal. In particular, the univariate, bivariate and trivariate normal distributions are used to derive the expected values of the maxima M_1, M_2, and M_3, which in turn lead to the expected values of the range R_1, R_2, and R_3. Some of the characteristics of these distributions are reviewed, derived and subsequently used in this chapter.

4.1 Properties of Multivariate Normal Distribution Function

Following A. M. Mood and F. A. Graybill (1963), let $W_1, W_2, ..., W_n$ be an n-dimensional random variable which is designated as elements of an $n \times 1$ random vector W by

$$W = \begin{pmatrix} W_1 \\ W_2 \\ \cdot \\ \cdot \\ \cdot \\ W_n \end{pmatrix}$$

This random vector is distributed as an n-variate normal if the joint probability density of $W_1, W_2, ..., W_n$ is

$$f(W) = f(W_1, W_2, \ldots, W_n)$$

$$= \frac{1}{(2\pi)^{n/2} \, |C|^{1/2}} \exp\{-\frac{1}{2}(W-\mu)C^{-1}(W-\mu)^T\} \qquad 4.1$$

where C is a positive definite symmetric matrix. Its elements are constants and is the covariance matrix, μ is an $n \times 1$ vector whose elements μ_i are the expected values of the random variables W_i, which are constants, and C^{-1} denoting the inverse matrix of C and $(W\text{-}\mu)^T$ representing the transpose of the matrix $(W\text{-}\mu)$. The covariance matrix C is explicitly given as

$$C = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{n1} & \sigma_{n2} & & \sigma_{nn} \end{pmatrix} \qquad 4.2$$

in which the element σ_{ij} represents the covariance of random variables W_i and W_j, equal to

$$\sigma_{ij} = \sqrt{\sigma_{ii}\,\sigma_{jj}} \;\; \rho_{ij}, \qquad 4.3$$

with σ_{ii} and σ_{jj} the variances of W_i and W_j respectively and ρ_{ij} their correlation coefficient. It may be shown for the n-variate normal random vector W that the marginal distribution of any W_i is normal with mean μ_i and variance σ_{ii}.

Another important point concerns the conditional distributions. Let the $n \times 1$ random vector W, the $n \times 1$ vector μ and the matrix C be partitioned as follows:

$$W = \begin{pmatrix} W_1^* \\ W_2^* \end{pmatrix}, \mu = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}, \text{ and } C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \qquad 4.4$$

with

$$W_1^* = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_k \end{pmatrix} \qquad U_1 = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix}$$

$$\text{and } C_{11} = \begin{pmatrix} \sigma_{11}, & \sigma_{12}, \ldots, & \sigma_{1k} \\ \sigma_{21}, & \sigma_{22}, \ldots, & \sigma_{2k} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \sigma_{k1}, & \sigma_{k2}, \ldots, & \sigma_{kk} \end{pmatrix} \qquad 4.5$$

The conditional distribution of W_1^* given W_2^* is the k-variate normal with the mean

$$U_1^* = U_1 + C_{12} C_{22}^{-1} (W_2^* - U_2) \ , \qquad 4.6$$

and the covariance matrix

$$C_{11.2} = C_{11} - C_{12} C_{22}^{-1} C_{21} \ , \qquad 4.7$$

in which $C_{11.2}$ denotes the covariance matrix of W_1^* given W_2^*. The partial correlation of W_i and $W_j (i,j < k)$, given $W_{k+1}, ..., W_n$, is defined by

$$\rho_{ij.(k+1)......n} = \frac{\sigma_{ij.(k+1)........n}}{\sqrt{\sigma_{ii.(k+1)....n}\ \sigma_{jj.(k+1)....n}}} \ . \qquad 4.8$$

For the particular cases of n = 1,2, and 3, the joint and conditional density functions are given in explicit forms:

(a) For n = 1, the univariate density function is

$$f(X) = \frac{1}{\sqrt{2\pi}\ \sigma_x} \exp\{ -\frac{1}{2} (\frac{X-\mu_x}{\sigma_x})^2 \} , \qquad 4.9$$

with μ_x and σ_x the mean and standard deviation respectively.

(b) For n = 2, the bivariate normal density function is

$$f(X,Y) = \frac{1}{(2\pi)\sigma_x \sigma_y (1-\rho_{xy}^2)^{1/2}} \exp\{ -\frac{1}{2(1-\rho_{xy}^2)} [(\frac{X-\mu_x}{\sigma_x})^2 - 2(\frac{X-\mu_x}{\sigma_x})(\frac{Y-\mu_y}{\sigma_y}) + (\frac{Y-\mu_y}{\sigma_y})^2] \} \qquad 4.10$$

while the conditional density function of X given Y is

$$f(X \mid Y) = \frac{1}{\sqrt{2\pi}\ \sigma_x (1-\rho_{xy}^2)} \exp\{ -\frac{1}{2\sigma_x^2 (1-\rho_{xy}^2)} [X - \mu_x - \frac{\rho_{xy}\sigma_x}{\sigma_y} (Y-\mu_y)]^2 \} . \qquad 4.11$$

(c) For n = 3, and assuming that $\mu_x = \mu_y = \mu_z = 0$, the trivariate normal density function is

$$f(X,Y,Z) = \frac{1}{(2\pi)^{3/4} | C |^{1/2}} \exp\{ -\frac{1}{2 | C |} [c_1 X^2 + c_2 Y^2 + c_3 Z^2 + 2c_4 XY + 2c_5 XZ + 2c_6 YZ] \} , \qquad 4.12$$

where

$$\begin{array}{lll} c_1 = \sigma_{yy}\sigma_{zz} - \sigma_{yz}^2 & , & c_4 = \sigma_{xz}\sigma_{yz} - \sigma_{xy}\sigma_{zz} , \\ c_2 = \sigma_{xx}\sigma_{zz} - \sigma_{xz}^2 & , & c_5 = \sigma_{xy}\sigma_{yz} \quad \sigma_{yy}\sigma_{xz} , \\ c_3 = \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2 & , & c_6 = \sigma_{xy}\sigma_{xz} - \sigma_{xx}\sigma_{yz} , \end{array} \qquad 4.13$$

and

$$C = \begin{pmatrix} \sigma_{xx}, \sigma_{xy}, \sigma_{xz} \\ \sigma_{yx}, \sigma_{yy}, \sigma_{yz} \\ \sigma_{zx}, \sigma_{zy}, \sigma_{zz} \end{pmatrix} = \begin{pmatrix} \sigma_x^2, \sigma_{xy}, \sigma_{xz} \\ \sigma_{yx}, \sigma_y^2, \sigma_{yz} \\ \sigma_{zx}, \sigma_{zy}, \sigma_z^2 \end{pmatrix} \qquad 4.14$$

Consider the three-dimensional vectors

$$W = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad \mu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and C of Eq. 4.14, and the partition

$$W = \begin{pmatrix} W_1^* \\ W_2^* \end{pmatrix} \quad \text{where} \quad W_1^* = \begin{pmatrix} X \\ Y \end{pmatrix} \quad \text{and} \quad W_2^* = (Z) \ .$$

From Eq. 4.6 and since $U_1 = U_2 = 0$, the conditional distribution of X and Y given Z has the mean $U_1^* = C_{12}C_{22}^{-1}\ W_2^*$. Since

$$C_{12} = \begin{pmatrix} \sigma_{xz} \\ \sigma_{yz} \end{pmatrix} \qquad C_{22} = (\sigma_z^2)$$

then

$$U_1^* = U_{xy.z}^* = \begin{pmatrix} \mu_{x.z} \\ \mu_{y.z} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_x}{\sigma_z}\rho_{xz}Z \\ \frac{\sigma_y}{\sigma_z}\rho_{yz}Z \end{pmatrix}. \qquad 4.15$$

Similarly, since the matrices C_{11} and C_{21} are given by

$$C_{11} = \begin{pmatrix} \sigma_x^2 & , & \sigma_{xy} \\ \sigma_{yx} & , & \sigma_y^2 \end{pmatrix}, C_{21} = (\sigma_{xz}, \sigma_{yz})$$

and using Eq. 4.7, the covariance matrix, $C_{11.2}$, denoted now as $C_{xy.z}$, is

$$= \begin{pmatrix} \sigma_x^2 - \sigma_{xz}^2/\sigma_z^2 & , & \sigma_{xy} - \sigma_{xz}\sigma_{yz}/\sigma_z^2 \\ \sigma_{xy} - \sigma_{xz}\sigma_{yz}/\sigma_z^2 & , & \sigma_y^2 - \sigma_{yz}^2/\sigma_z^2 \end{pmatrix} \qquad 4.16$$

which may also be expressed as

$$C_{xy.z} = \begin{pmatrix} \sigma_{x.z}^2 & , \sigma_{xy.z} \\ \sigma_{yx.z} & , \sigma_{y.z}^2 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_x^2(1-\rho_{xz}^2) & , \sigma_x\sigma_y(\rho_{xy} - \rho_{xz}\rho_{yz}) \\ \sigma_x\sigma_y(\rho_{xy} - \rho_{xz}\rho_{yz}), & \sigma_y^2(1-\rho_{yz}^2) \end{pmatrix} \qquad 4.17$$

where the ρ's denote the correlation coefficients between the indicated random variables. Therefore, by using Eq. 4.10, the conditional distribution of X,,and ,Y given Z, is

$$f(X,Y|Z) = \frac{1}{2\pi\sqrt{C_{xy.z}}}$$

$$\exp\{-\frac{1}{2(1-\rho_{xy.z}^2)}\ [(\frac{X-\mu_{x.z}}{\sigma_{x.z}})^2 - 2\rho_{xy.z}$$

$$(\frac{X-\mu_{x.z}}{\sigma_{x.z}})(\frac{Y-\mu_{y.z}}{\sigma_{y.z}}) + (\frac{Y-\mu_{y.z}}{\sigma_{y.z}})^2]\}, \qquad 4.18$$

with $\rho_{xy.z} = \sigma_{xy.z}/\sigma_{x.z}\sigma_{y.z}$ and $\sigma_{xy.z}, \sigma_{x.z}$ and $\sigma_{y.z}$ given by Eq. 4.17.

Similarly, the conditional distribution function of X and Z given Y is

$$f(X,Z|Y) = \frac{1}{2\pi\sqrt{C_{xz.y}}}$$

$$\exp\{-\frac{1}{2(1-\rho_{xz.y}^2)}\ [(\frac{X-\mu_{x.y}}{\sigma_{x.y}})^2 - 2\rho_{xz.y}$$

$$(\frac{X-\mu_{x.y}}{\sigma_{x.y}})(\frac{Z-\mu_{z.y}}{\sigma_{z.y}}) + (\frac{Z-\mu_{z.y}}{\sigma_{z.y}})^2]\}, \qquad 4.19$$

with $\rho_{xz.y} = \sigma_{xz.y}/\sigma_{x.y}\sigma_{z.y}$ and the matrices of mean and covariance given by

$$U^*_{xz.y} = \begin{pmatrix} \mu_{x.y} \\ \mu_{z.y} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_x}{\sigma_y}\rho_{xy}Y \\ \frac{\sigma_z}{\sigma_y}\rho_{zy}Y \end{pmatrix}, \qquad 4.20$$

and

$$C_{xz.y} = \begin{pmatrix} \sigma_{x.y}^2, \sigma_{xz.y} \\ \sigma_{zx.y}, \sigma_{z.y}^2 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_x^2(1-\rho_{xy}^2) & , \sigma_x\sigma_z(\rho_{xz} - \rho_{xy}\rho_{yz}) \\ \sigma_x\sigma_z(\rho_{xz} - \rho_{xy}\rho_{yz}), & \sigma_z^2(1-\rho_{yz}^2) \end{pmatrix}, \qquad 4.21$$

Finally, the conditional distribution function of Y and Z given X is

$$f(Y,Z|X) = \frac{1}{2\pi\sqrt{C_{xy.z}}}$$

$$\exp\{-\frac{1}{2(1-\rho_{xy.z}^2)}\ [(\frac{Y-\mu_{y.x}}{\sigma_{y.x}})^2 2\rho_{yz:x}$$

$$(\frac{Y-\mu_{y.x}}{\sigma_{y.x}})(\frac{Z-\mu_{z.x}}{\sigma_{z.x}}) + (\frac{Z-\mu_{z.x}}{\sigma_{z.x}})^2]\}, \qquad 4.22$$

with $\rho_{yz.x} = \sigma_{yz.x}/\sigma_{y.x}\sigma_{z.x}$ and the matrices of mean and covariance given by

$$U^*_{yz.x} = \begin{pmatrix} \mu_{y.x} \\ \mu_{z.x} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_y}{\sigma_x}\rho_{xy}X \\ \frac{\sigma_z}{\sigma_x}\rho_{xz}X \end{pmatrix} \qquad 4.23$$

and

$$C_{yz.x} = \begin{pmatrix} \sigma^2_{y.x}, \sigma_{yz.x} \\ \sigma_{zy.x}, \sigma^2_{z.x} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2_y (1-\rho^2_{xy}) & , \sigma_y \sigma_z (\rho_{yz} - \rho_{xy}\rho_{xz}) \\ \sigma_y \sigma_z (\rho_{yz} - \rho_{xy}\rho_{xz}), & \sigma^2_z (1-\rho^2_{xz}) \end{pmatrix} \quad 4.24$$

4.2 Expected Value of Surplus of Random Variables with General Covariance Structure

The following mathematical derivations deal with the expected value of the maximum of partial sums for $n = 1$, 2, and 3. They are performed in general so that the expected values obtained may be used for both the unadjusted and adjusted partial sums. The assumption is made that the departures $(x_t - y_t)$ are normally distributed with mean zero so that the distribution of the partial sums is also normal with mean zero. In order to simplify the mathematical derivations the following notation is introduced:

$$X = S_1 = (x_1 - y_1)$$
$$Y = S_2 = (x_1 - y_1) + (x_2 - y_2), \text{ and}$$
$$Z = S_3 = (x_1 - y_1) + (x_2 - y_2) + (x_3 - y_3) \ . \quad 4.25$$

A. The case n = 1. According to the above notation the maximum M_1 is defined as

$$M_1 = \max(0, X)$$

Then

(1) $M_1 = 0$ if $X < 0$

(2) $M_1 = X$ if $X > 0$

The expected value of M_1 is

$$E\{M_1\} = E\{X\} = \int_0^\infty X f(X) \, dX$$

Since X is normally distributed, $f(X)$ is defined by Eq. 4.9, so that

$$E\{M_1\} = \frac{1}{\sqrt{2\pi}} \sigma_x \quad 4.26$$

Since for a symmetric distribution, Eq. 3.64 applies, then the expected value of the range is

$$E\{R_1\} = \sqrt{\frac{2}{\pi}} \, [\text{Var } X]^{1/2} \quad 4.27$$

B. The case n = 2. In this case the maximum M_2 is defined as

$$M_2 = \max(0, X, Y)$$

Then

(1) $M_2 = 0$ for $X \leqslant 0$, $Y \leqslant 0$

(2) $M_2 = X$ for $X > 0$, $Y < X$

(3) $M_2 = Y$ for $X < Y$, $Y > 0$.

The expected value of M_2 is

$$E\{M_2\} = E\{X\} + E\{Y\} , \quad 4.28$$

where

$$E\{X\} = \int_0^\infty \int_{-\infty}^X X f(X,Y) \, dY \, dX , \quad 4.29$$

and

$$E\{Y\} = \int_0^\infty \int_{-\infty}^Y Y f(X,Y) \, dX \, dY \ . \quad 4.30$$

Since the above two integrals are symmetric, the solution of only one is necessary. Therefore, for solving $E\{X\}$ let us use the conditional distribution of X given, Y , so that

$$E\{X\} = \int_0^\infty \int_{-\infty}^X X f(X \mid Y) f(Y) \, dY \, dX ,$$

which, separated into two integrals, gives

$$E\{X\} = \int_{-\infty}^0 f(Y) \int_0^\infty X f(X \mid Y) \, dX \, dY$$
$$+ \int_0^\infty f(Y) \int_Y^\infty X f(X \mid Y) \, dX \, dY , \quad 4.31$$

where $f(Y)$ and $f(X|Y)$ are given by Eqs. 4.9 and 4.11 respectively, with, μ_x and μ_y, equal to zero. For convenience, the conditional density function is expressed by

$$f(X \mid Y) = \frac{1}{\sqrt{2\pi}\, a} \exp\{-\frac{1}{2a^2}(X - bY)^2\}$$

where

$$a = \sigma_x (1 - \rho^2_{xy}) \text{ and } b = \rho_{xy} \, \sigma_x / \sigma_y \ .$$

With the above expression for $f(X|Y)$, the inside integrals of Eq. 4.31, denoted from now on by I, are

$$I = \int_{\ell}^{\infty} X f(X \mid Y)\, dX$$

$$= \int_{\ell}^{\infty} X \frac{1}{\sqrt{2\pi}\, a} \exp\{-\frac{1}{2a^2} (X - bY)^2\} dX$$

whose solution is equal to

$$I = \frac{a}{\sqrt{2\pi}} \exp\{-\frac{1}{2} (\frac{\ell - bY}{a})^2\}$$

$$+ bY [1 - \Phi(\frac{\ell - bY}{a})] , \qquad 4.32$$

with $\Phi(.)$ denoting the univariate normal cumulative distribution function.

For the first inside integral of Eq. 4.31, denoted by I_1, $\ell = 0$, so that Eq. 4.32 gives

$$I_1 = \frac{a}{\sqrt{2\pi}} \exp\{-\frac{1}{2} (\frac{b}{a})^2 Y^2\} + bY \Phi(\frac{b}{a} Y) . \qquad 4.33$$

For the second inside integral of Eq. 4.31, $\ell = Y$, so that Eq. 4.32 produces

$$I_2 = \frac{a}{\sqrt{2\pi}}$$

$$\exp\{-\frac{1}{2} (\frac{1-b}{a})^2 Y^2\} + bY \Phi[-\frac{(1-b)}{a} Y] . \qquad 4.34$$

Substitution of Eqs. 4.33 and 4.34 into Eq. 4.31 leads to

$$E\{X\} = \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{0} f(Y) \exp$$

$$\{-\frac{1}{2} (\frac{b}{a})^2 Y^2\} dY + \frac{a}{\sqrt{2\pi}} \int_{0}^{\infty} f(Y) \exp$$

$$\{-\frac{1}{2} (\frac{1-b}{a})^2 Y^2\} dY +$$

$$+ b \int_{-\infty}^{0} Y f(Y) \Phi(\frac{bY}{a}) dY + b \int_{0}^{\infty} Y f(Y) \Phi$$

$$[-\frac{(1-b)}{a} Y] dY = E\{X\} = \frac{a}{\sqrt{2\pi}} [\int_{-\infty}^{0} f(Y) \exp$$

$$\{-\frac{1}{2} (\frac{b}{a})^2 Y^2\} dY + \int_{-\infty}^{0} f(Y)$$

$$\exp\{-\frac{1}{2} (\frac{1-b}{a})^2 Y^2\} dY] + b [\int_{-\infty}^{0} Y f(Y)$$

$$\Phi(\frac{bY}{a}) dY - \int_{-\infty}^{0} Y f(Y) \Phi\{\frac{(1-b)}{a} Y\} dY] . \qquad 4.35$$

The above expression basically contains the following two types of integrals

$$I_3 = \int_{-\infty}^{0} f(Y) \exp\{-\frac{1}{2} c^2 Y^2\} dY$$

$$\text{and} \quad I_4 = \int_{-\infty}^{0} Y f(Y) \Phi(cY)\, dY$$

with $c = b/a$ for the first and third integrals and, $c = (1-b)/a$ for the second and fourth integrals of Eq. 4.35. The solutions of these integrals are:

$$I_3 = \int_{-\infty}^{0} f(Y) \exp\{-\frac{1}{2} c^2 Y^2\} dY$$

$$= \frac{1}{2(c^2 \sigma_y^2 + 1)^{1/2}} , \qquad 4.36$$

and

$$I_4 = \int_{-\infty}^{0} Y f(Y) \Phi(cY)\, dY$$

$$= \frac{\sigma_y}{2\sqrt{2\pi}} [-1 + \frac{c\, \sigma_y}{(c^2 \sigma_y^2 + 1)^{1/2}}] . \qquad 4.37$$

Substitution of Eqs. 4.36 and 4.37 into Eq. 4.35 leads to

$$E\{X\} = \frac{a^2}{2\sqrt{2\pi}} \{\frac{1}{[a^2 + b^2 \sigma_y^2]^{1/2}}$$

$$+ \frac{1}{[a^2 + (1-b)^2 \sigma_y^2]^{1/2}}\} + \frac{b\, \sigma_y^2}{2\sqrt{2\pi}}$$

$$\{\frac{b}{[a^2 + b^2 \sigma_y^2]^{1/2}} - \frac{(1-b)}{[a^2 + (1-b)^2 \sigma_y^2]^{1/2}}\} ,$$

Finally, replacing the constants a and b by $\sigma_x(1-\rho_{xy}^2)^{1/2}$ and $\rho_{xy}\sigma_x/\sigma_y$ respectively, the above equation becomes

$$E\{X\} = \frac{1}{2\sqrt{2\pi}} \frac{\sigma_x}{[\mathrm{Var}(Y - X)]^{1/2}}$$

$$\{\sigma_x - \rho_{xy} \sigma_y + [\mathrm{Var}(Y - X)]^{1/2}\} . \qquad 4.38$$

Since the integral $E\{Y\}$ of Eq. 4.30 is of the same type as $E\{X\}$ of Eq. 4.29, Eq. 4.38 by making the corresponding replacements becomes

$$E\{Y\} = \frac{1}{2\sqrt{2\pi}} \frac{\sigma_y}{[\mathrm{Var}(Y-X)]^{\frac{1}{2}}} \{\sigma_y - \rho_{xy}\sigma_x + [\mathrm{Var}\,(Y-X)]^{\frac{1}{2}}\}. \qquad 4.39$$

Substitution of Eqs. 4.38 and 4.39 into Eq. 4.28 leads to

$$E\{M_2\} = \frac{1}{\sqrt{2\pi}} \{\frac{1}{2}[\mathrm{Var}\,X]^{\frac{1}{2}} + \frac{1}{2}[\mathrm{Var}\,Y]^{\frac{1}{2}} + \frac{1}{2}[\mathrm{Var}\,(Y-X)]^{\frac{1}{2}}\}. \qquad 4.40$$

Consequently, the expected value of the range is

$$E\{R_2\} = \sqrt{\frac{2}{\pi}} \{\frac{1}{2}[\mathrm{Var}\,X]^{\frac{1}{2}} + \frac{1}{2}[\mathrm{Var}\,Y]^{\frac{1}{2}} + \frac{1}{2}[\mathrm{Var}\,(Y-X)]^{\frac{1}{2}}\}. \qquad 4.41$$

C. The case n = 3. The maximum M_3 is defined as $M_3 = \max(0, S_1, S_2, S_3) = \max(0, X, Y, Z)$, or

(1) $M_3 = 0$, for $X \leq 0$, $Y \leq 0$, $Z \leq 0$

(2) $M_3 = X$, for $X > 0$, $Y < X$, $Z < X$

(3) $M_3 = Y$, for $X < Y$, $Y > 0$, $Z < Y$

(4) $M_3 = Z$, for $X < Z$, $Y < Z$, $Z > 0$

Therefore, the expected value of M_3 may be written as

$$E\{M_3\} = E\{X\} + E\{Y\} + E\{Z\}, \qquad 4.42$$

where

$$E\{X\} = \int_0^{\infty} \int_{-\infty}^{X} \int_{-\infty}^{X} X\, f(X,Y,Z)\, dY\, dZ\, dX, \qquad 4.43$$

$$E\{Y\} = \int_0^{\infty} \int_{-\infty}^{Y} \int_{-\infty}^{Y} Y\, f(X,Y,Z)\, dX\, dZ\, dY, \qquad 4.44$$

and

$$E\{Z\} = \int_0^{\infty} \int_{-\infty}^{Y} \int_{-\infty}^{Y} Z\, f(X,Y,Z)\, dX\, dY\, dZ. \qquad 4.45$$

Using the conditional density functions, the above integrals become

$$E\{X\} = \int_0^{\infty} \int_{-\infty}^{X} \int_{-\infty}^{X} X\, f(X)\, f(Y,Z \mid X)\, dY\, dZ\, dX, \qquad 4.46$$

$$E\{Y\} = \int_0^{\infty} \int_{-\infty}^{Y} \int_{-\infty}^{Y} Y\, f(Y)\, f(X,Z \mid Y)\, dX\, dZ\, dY, \qquad 4.47$$

and

$$E\{Z\} = \int_0^{\infty} \int_{-\infty}^{Z} \int_{-\infty}^{Z} Z\, f(Z)\, f(X,Y \mid Z)\, dX\, dY\, dZ \qquad 4.48$$

where $f(X,Y|Z)$, $f(X,Z|Y)$ and $f(Y,Z|X)$ are given by Eqs. 4.18, 4.19 and 4.22 respectively.

Solution of the integral $E\{X\}$ of Eq. 4.46. By making the following change in the conditional density function $f(Y,Z|X)$ of Eq. 4.22,

$$k_1 = 1 - \rho_{yz.x}^2, \quad k_x = (2\pi\, \sigma_{y.x}\sigma_{z.x}\sqrt{k_1})^{-1} \qquad 4.49$$

and

$$u = \frac{Y - \sigma_y \rho_{xy} X/\sigma_x}{\sigma_{y.x}}, \quad \text{and } \sigma = \frac{Z - \sigma_z \rho_{xz} X/\sigma_x}{\sigma_{z.x}}, \qquad 4.50$$

$f(Y,Z|X)$ becomes

$$f(Y,Z \mid X) = k_x \exp\{-\frac{1}{2k_1}(u^2 - 2\rho_{yz.x} uv + v^2)\},$$

and the integral $E\{X\}$ of Eq. 4.46 is expressed as

$$E\{X\} = k_x \sigma_{y.x}\sigma_{z.x} \int_0^{\infty} \int_{-\infty}^{c_2 X} \int_{-\infty}^{c_1 X} f(X) \exp\{-\frac{1}{2k_1}(u^2 - 2\rho_{yz.x} uv + v^2)\}\, du\, dv\, dX \qquad 4.51$$

in which

$$c_1 = \frac{\sigma_x - \sigma_y \rho_{xy}}{\sigma_x \sigma_{y.x}} \qquad c_2 = \frac{\sigma_x - \sigma_z \rho_{xz}}{\sigma_x \sigma_{z.x}} \qquad 4.52$$

The constants c_1 and c_2 are usually negative for the linear dependence between the components of the partial sums. They are equal to zero for the case of independence (see Appendix). Therefore, the solution that follows is for $c_1 \leqslant 0$ and $c_2 \leqslant 0$.

Replacing $-c_1$ by b_1, and $-c_2$ by b_2 the triple integral of Eq. 4.51 is graphically shown in Fig. 4.1.

In order to integrate first in X, Eq. 4.51 is separated into two integrals as

$$E\{X\} = k_x \sigma_{y.x} \sigma_{z.x} \Big[\int_{-\infty}^{0} \int_{-\infty}^{0} \exp\{-\frac{1}{2k_1}(u^2 - 2\rho_{yz.x} uv + v^2)\} \int_{0}^{-v/b_2} X\, f(X)\, dX\, du\, dv - \int_{-\infty}^{0} \int_{b_1 v/b_2}^{0} \exp\{-\frac{1}{2k_1} (u^2 - 2\rho_{yz.x} uv + v^2)\} \int_{-u/b_1}^{-v/b_2} X\, f(X)\, dX\, du\, dv \Big] \qquad 4.53$$

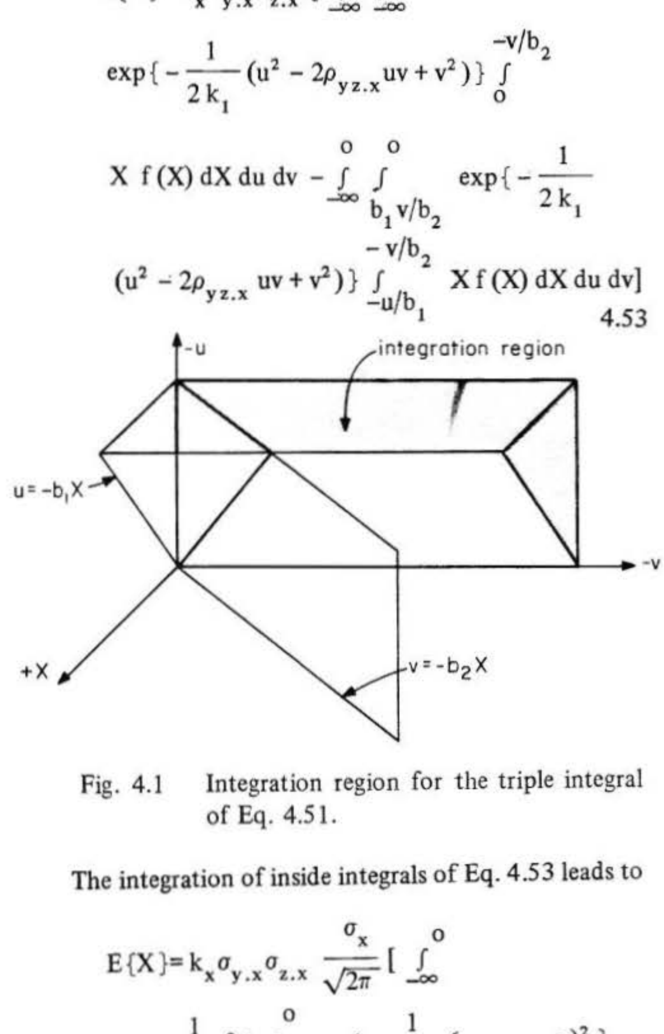

Fig. 4.1 Integration region for the triple integral of Eq. 4.51.

The integration of inside integrals of Eq. 4.53 leads to

$$E\{X\} = k_x \sigma_{y.x} \sigma_{z.x} \frac{\sigma_x}{\sqrt{2\pi}} \Big[\int_{-\infty}^{0} \exp\{-\frac{1}{2} v^2\} \int_{-\infty}^{0} \exp\{-\frac{1}{2k_1}(u - \rho_{yz.x} v)^2\}\, du\, dv + \int_{-\infty}^{0} \exp\{-\frac{1}{2} k_2 v^2\} \int_{b_1 v/b_2}^{0} \exp\{-\frac{1}{2k_1} (u - \rho_{yz.x} v)^2\}\, du\, dv - \int_{-\infty}^{0} \exp\{-\frac{1}{2} k_2 v^2\} \int_{-\infty}^{0} \exp\{-\frac{1}{2k_1}(u - \rho_{yz.x} v)^2\}\, du\, dv - \int_{-\infty}^{0} \exp\{-\frac{1}{2} k_3 v^2\} \int_{b_1 v/b_2}^{0} \exp\{-\frac{1}{2}(\sqrt{k_4}\, u - \frac{\rho_{yz.x}}{k_1 \sqrt{k_4}} v)^2\}\, du\, dv \Big] \qquad 4.54$$

in which the constants k_1 and k_x are given by Eq. 4.49 and k_2, k_3, and k_4 are

$$k_2 = \frac{1 + (b_2 \sigma_x)^2}{(b_2 \sigma_x)^2}, \quad k_3 = \frac{1 + (b_1 \sigma_x)^2}{k_1 + (b_1 \sigma_x)^2}, \quad k_4 = \frac{k_1 + (b_1 \sigma_x)^2}{k_1 (b_1 \sigma_x)^2} \qquad 4.55$$

Integrals of Eq. 4.54 have the general form

$$I = \int_{-\infty}^{0} \exp\{-\frac{1}{2} a_1 v^2\} \int_{a_2 v}^{0} \exp\{-\frac{1}{2a_3} (a_4 u - a_5 v)^2\}\, du\, dv \qquad 4.56$$

and their solution depends on the lower limit of the inside integral. Therefore, in order to find $E\{X\}$, and subsequently $E\{Y\}$ and $E\{Z\}$, the following cases of Eq. 4.56 were first solved:

(a) For $0 < a_2 < \infty$

$$I = (2\pi) \frac{\sqrt{a_3}}{\sqrt{a_1}\, a_4} \Big[\frac{1}{2\pi} \arctan \Big(\frac{a_2 a_4 - a_5}{\sqrt{a_1} \sqrt{a_3}}\Big) + \frac{1}{2\pi} \arctan \Big(\frac{a_5}{\sqrt{a_1} \sqrt{a_3}}\Big) \Big]. \qquad 4.57$$

(b) For $a_2 = \infty$,

$$I = (2\pi) \frac{\sqrt{a_3}}{\sqrt{a_1} \sqrt{a_4}} \Big[-\frac{1}{4} + \frac{1}{2\pi} \arctan \Big(\frac{a_5}{\sqrt{a_1} \sqrt{a_3}}\Big) \Big]. \qquad 4.58$$

(c) For $-\infty < a_2 < 0$

$$I = (2\pi)\frac{\sqrt{a_3}}{\sqrt{a_1}\;a_4}\left[-\frac{1}{2\pi}\arctan\left(\frac{|a_2|\,a_4 + a_5}{\sqrt{a_1}\sqrt{a_3}}\right) + \frac{1}{2\pi}\arctan\left(\frac{a_5}{\sqrt{a_1}\sqrt{a_3}}\right)\right] . \qquad 4.59$$

and

(d) For $a_2 = -\infty$

$$I = (2\pi)\frac{\sqrt{a_3}}{\sqrt{a_1}\;a_4}\left[\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{a_5}{\sqrt{a_1}\sqrt{a_3}}\right)\right] , \qquad 4.60$$

in which the angles are reduced to the first quadrant and measured counter-clockwise.

The first integral of Eq. 4.54, denoted by I_1, with $a_1 = 1$, $a_2 = -\infty$, $a_3 = k_1$, $a_4 = 1$ and $a_5 = \rho_{yz.x}$, is obtained from Eq. 4.60 as

$$I_1 = (2\pi)\sqrt{k_1}\left[\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{\rho_{yz.x}}{\sqrt{k_1}}\right)\right] . \qquad 4.61$$

The second integral of Eq. 4.54, denoted by I_2 with $a_1 = k_2$, $a_2 = b_1/b_2$, $a_3 = k_1$, $a_4 = 1$ and $a_5 = \rho_{yz.x}$, is obtained from Eq. 4.57, as

$$I_2 = (2\pi)\frac{\sqrt{k_1}}{\sqrt{k_2}}\left[\frac{1}{2\pi}\arctan\left(\frac{b_1 - b_2\,\rho_{yz.x}}{b_2\sqrt{k_1}\sqrt{k_2}}\right) + \frac{1}{2\pi}\arctan\left(\frac{\rho_{yz.x}}{\sqrt{k_1}\sqrt{k_2}}\right)\right] \qquad 4.62$$

The third integral of Eq. 4.54, denoted by I_3, with $a_1 = k_2$, $a_2 = -\infty$, $a_3 = k_1$, $a_4 = 1$ and $a_5 = \rho_{yz.x}$, is obtained from Eq. 4.60, as

$$I_3 = (2\pi)\frac{\sqrt{k_1}}{\sqrt{k_2}}\left[\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{\rho_{yz.x}}{\sqrt{k_1}\sqrt{k_2}}\right)\right] . \qquad 4.63$$

Finally, the fourth integral of Eq. 4.54, denoted by I_4, with $a_1 = k_3$, $a_2 = b_1/b_2$, $a_3 = 1$, $a_4 = \sqrt{k_4}$ and $a_5 = \rho_{yz.x}/k_1\sqrt{k_4}$, is obtained from Eq. 4.57 as

$$I_4 = \frac{(2\pi)\;\sqrt{k_1}}{\sqrt{k_1}\sqrt{k_3}\sqrt{k_4}}\left[\frac{1}{2\pi}\arctan\left(\frac{b_1 k_1 k_4 - b_2\,\rho_{yz.x}}{b_2 k_1\sqrt{k_3}\sqrt{k_4}}\right) + \frac{1}{2\pi}\arctan\left(\frac{\rho_{yz.x}}{k_1\sqrt{k_3}\sqrt{k_4}}\right)\right] . \qquad 4.64$$

Substituting Eqs. 4.61 through 4.64 into Eq. 4.54, and since Eq. 4.49 gives $k_x\sigma_{y.x}\sigma_{z.x}\sqrt{k_1} = 1/(2\pi)$, it follows that:

$$E\{X\} = \frac{\sigma_x}{\sqrt{2\pi}}\left\{\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{\rho_{yz.x}}{\sqrt{k_1}}\right) - \frac{1}{\sqrt{k_2}}\left[\frac{1}{4} - \frac{1}{2\pi}\arctan\left(\frac{b_1 - b_2\,\rho_{yz.x}}{b_2\sqrt{k_1}\sqrt{k_2}}\right)\right] - \frac{1}{\sqrt{k_1}\sqrt{k_3}\sqrt{k_4}}\left[\frac{1}{2\pi}\arctan\left(\frac{b_1 k_1 k_4 - b_2\,\rho_{yz.x}}{b_2 k_1\sqrt{k_3}\sqrt{k_4}}\right) + \frac{1}{2\pi}\arctan\left(\frac{\rho_{yz.x}}{k_1\sqrt{k_3}\sqrt{k_4}}\right)\right]\right\} \qquad 4.65$$

Solution of the integral $E\{Y\}$ of Eq. 4.47. Following a similar change of variables as in the case of the integral $E\{X\}$, Eq. 4.47 becomes

$$E\{Y\} = k_y\sigma_{x.y}\sigma_{z.y}\int_0^\infty\int_{-\infty}^{c_2'Y}\int_{-\infty}^{c_1'Y} Y\,f(Y)\exp\left\{-\frac{1}{2k_1'}\left(u^2 - 2\rho_{xz.y}uv + v^2\right)\right\}du\,dv\,dY \qquad 4.66$$

in which

$$k_1' = 1 - \rho_{xz.y}^2 \quad , \quad k_y = (2\pi\,\sigma_{x.y}\sigma_{z.y}\sqrt{k_1'}\,)^{-1} \qquad 4.67$$

$$b_1' = c_1' = \frac{\sigma_y - \sigma_x\,\rho_{xy}}{\sigma_y\,\sigma_{x.y}} ,$$

$$\text{and } b_2' = -c_2' = \frac{\sigma_y - \sigma_z\,\rho_{yz}}{\sigma_y\,\sigma_{z.y}} , \qquad 4.68$$

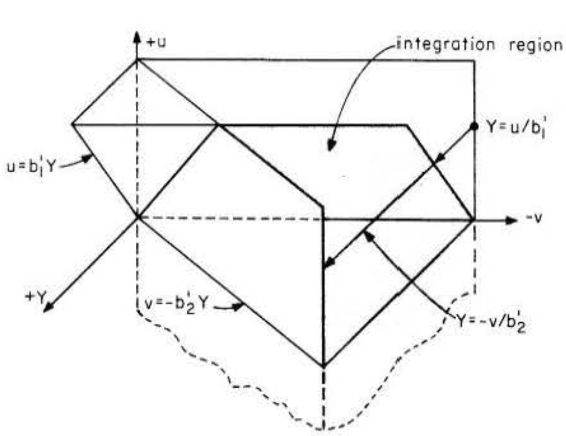

Fig. 4.2 Integration region for the triple integral of Eq. 4.66.

with the constants $c_1' > 0$ and, $c_2' \leqslant 0$ (see Appendix). The integration region of $E\{Y\}$ of Eq. 4.66 is graphically shown in Fig. 4.2.

In order to integrate first in Y, Eq. 4.66 is separated into two integrals, see Fig. 4.2, as

$$E\{Y\} = k_y\, \sigma_{x.y}\, \sigma_{z.y} \Big[\int_{-\infty}^{0} \int_{-\infty}^{0} \exp\{-\frac{1}{2k_1'}(u^2 - 2\rho_{xz.y}\, uv + v^2)\} \int_{0}^{-v/b_2'} Y f(Y)\, dY\, du\, dv + \int_{-\infty}^{0} \int_{0}^{-b_1' v/b_2'} \exp\{-\frac{1}{2k_1'}(u^2 - 2\rho_{xz.y}\, uv + v^2)\} \int_{u/b_1'}^{-v/b_2'} Y f(Y)\, dY\, du\, dv\Big] \; . \qquad 4.69$$

The integration of the inside integrals of Eq. 4.69 leads to the following four integrals:

$$E\{Y\} = k_y \sigma_{x.y} \sigma_{z.y} \frac{\sigma_y}{\sqrt{2\pi}} \Big[\int_{-\infty}^{0} \exp\{-\frac{1}{2}v^2\} \int_{-\infty}^{0} \exp\{-\frac{1}{2k_1'}(u - \rho_{xz.y}v)^2\} du\, dv + \int_{-\infty}^{0} \exp\{-\frac{1}{2}k_2' v^2\} \int_{-b_1' v/b_2'}^{0} \exp\{-\frac{1}{2k_1'}(u - \rho_{xz.y}v)^2\} du\, dv - \int_{-\infty}^{0} \exp\{-\frac{1}{2}k_2' v^2\} \int_{-\infty}^{0} \exp\{-\frac{1}{2k_1'}(u - \rho_{xz.y}v)^2\} du\, dv - \int_{-\infty}^{0} \exp\{-\frac{1}{2}k_3' v^2\} \int_{-b_1' v/b_2'}^{0} \exp\{-\frac{1}{2}(\sqrt{k_4'}\, u - \frac{\rho_{xz.y}}{k_1'\sqrt{k_4'}} v)^2\} du\, dv \Big] \; , \qquad 4.70$$

where the constants k_1' and k_y are given by Eq. 4.67 and k_2', k_3', and k_4' are

$$k_2' = \frac{1 + (b_2'\sigma_y)^2}{(b_2'\,\sigma_y)^2} \; , \quad k_3' = \frac{1 + (b_1'\sigma_y)^2}{k_1' + (b_1'\sigma_y)^2} \; ,$$

$$\text{and} \quad k_4' = \frac{k_1' + (b_1'\sigma_y)^2}{k_1'(b_1'\sigma_y)^2} \; .$$

Since the four integrals of Eq. 4.70 are of the same type as those of Eq. 4.56, their solutions, given by Eqs. 4.57 to 4.60 will also be used here.

The first integral of Eq. 4.70, denoted by I_1', with $a_1 = 1$, $a_2 = -\infty$, $a_3 = k_1'$, $a_4 = 1$, and $a_5 = \rho_{xz.y}$, is obtained by using Eq. 4.60 as

$$I_1' = (2\pi)\sqrt{k_1'}\; [\frac{1}{4} + \frac{1}{2\pi}\arctan\,(\frac{\rho_{xz.y}}{\sqrt{k_1'}})] \; . \qquad 4.72$$

The second integral of Eq. 4.70, denoted by I_2', with $a_1 = k_2'$, $a_2 = -b_1'/b_2'$, $a_3 = k_1'$, $a_4 = 1$, and $a_5 = \rho_{xz.y}$, is obtained from Eq. 4.60 as

$$I_2' = (2\pi)\frac{\sqrt{k_1'}}{\sqrt{k_2'}}\; [-\frac{1}{2\pi}\arctan\,(\frac{b_1' + b_2'\,\rho_{xz.y}}{b_2'\sqrt{k_1'}\sqrt{k_2'}}) + \frac{1}{2\pi}\arctan\,(\frac{\rho_{xz.y}}{\sqrt{k_1'}\sqrt{k_2'}})] \; . \qquad 4.73$$

The third integral of Eq. 4.70, denoted by I'_3, with $a_1 = k'_2$, $a_2 = -\infty$, $a_3 = k'_1$, $a_4 = \infty$, and $a_5 = \rho_{xz.y}$, is obtained from Eq. 4.60 as

$$I'_3 = (2\pi)\,\frac{\sqrt{k'_1}}{\sqrt{k'_2}}\,\left[\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{\rho_{xz.y}}{\sqrt{k'_1}\,\sqrt{k'_2}}\right)\right] . \qquad 4.74$$

Finally, the fourth integral of Eq. 4.70, denoted by I'_4, with $a_1 = k'_3$, $a_2 = -b'_1/b'_2$, $a_3 = 1$, $a_4 = \sqrt{k'_4}$, and $a_5 = \rho_{xz.y}/k'_1\sqrt{k'_4}$, is obtained from Eq. 4.59 as

$$I'_4 = \frac{(2\pi)\sqrt{k'_1}}{\sqrt{k'_1}\,\sqrt{k'_3}\,\sqrt{k'_4}}\left[-\frac{1}{2\pi}\arctan\left(\frac{b'_1 k'_1 k'_4 + b'_2 \rho_{xz.y}}{b'_2 k'_1 \sqrt{k'_3}\sqrt{k'_4}}\right) + \frac{1}{2\pi}\arctan\left(\frac{\rho_{xz.y}}{k'_1\sqrt{k'_3}\sqrt{k'_4}}\right)\right] . \qquad 4.75$$

Substituting Eqs. 4.72 through 4.75 into Eq. 4.70, and since Eq. 4.67 gives $k_y \sigma_{x.y}\sigma_{z.y}\sqrt{k_1}' = 1/(2\pi)$, it follows that:

$$E\{Y\} = \frac{\sigma_y}{\sqrt{2\pi}}\left\{\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{\rho_{xz.y}}{\sqrt{k'_1}}\right) - \frac{1}{\sqrt{k'_2}}\left[\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{b'_1 + b'_2\rho_{xz.y}}{b'_2\sqrt{k'_1}\sqrt{k'_2}}\right)\right] + \frac{1}{\sqrt{k'_1}\sqrt{k'_3}\sqrt{k'_4}}\left[\frac{1}{2\pi}\arctan\left(\frac{b'_1 k'_1 k'_4 + b'_2\rho_{xz.y}}{b'_2 k'_1\sqrt{k'_3}\sqrt{k'_4}}\right) - \frac{1}{2\pi}\arctan\left(\frac{\rho_{xz.y}}{k'_1\sqrt{k'_3}\sqrt{k'_4}}\right)\right]\right\} . \qquad 4.76$$

Solution of the integral $E\{Z\}$ of Eq. 4.48. Following a similar change of variables as in the case of the integral $E\{X\}$, Eq. 4.48 becomes

$$E\{Z\} = k_z \sigma_{x.z}\sigma_{y.z}\int_0^{\infty}\int_{-\infty}^{c''_2 Z}\int_{-\infty}^{c''_1 Z} Z\, f(Z)\,\exp\left\{-\frac{1}{2k''_1}\left(u^2 - 2\rho_{xy.z}uv + v^2\right)\right\}du\,dv\,dZ , \qquad 4.77$$

where

$$k''_1 = 1 - \rho^2_{xy.z} , \quad k_z = (2\pi\,\sigma_{x.z}\sigma_{y.z}\sqrt{k''_1})^{-1} \qquad 4.78$$

$$b''_1 = c''_1 = \frac{\sigma_z - \sigma_x\rho_{xz}}{\sigma_z\sigma_{x.z}} , \quad b''_2 = c''_2 = \frac{\sigma_z - \sigma_y\rho_{yz}}{\sigma_z\sigma_{y.z}} \qquad 4.79$$

with the constants $c_1 > 0$ and, $c_2 > 0$ (see Appendix).

The integration region of $E\{Z\}$ of Eq. 4.77 is graphically shown in Fig. 4.3.

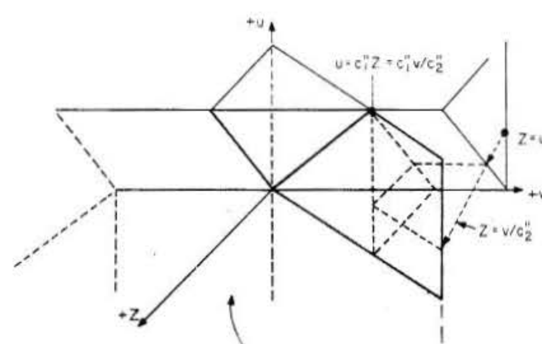

Fig. 4.3 Integration region for the triple integral of Eq. 4.77.

In order to integrate first in Z, Eq. 4.66 is separated into five integrals, see Fig. 4.3, as follows:

$$E\{Z\} = k_z\,\sigma_{x.z}\,\sigma_{y.z}\,[\int_{-\infty}^{0}\int_{-\infty}^{0} \exp\{-\frac{1}{2k_1''}(u^2 - 2\rho_{xy.z}\,uv + v^2)\}\int_{0}^{\infty} Z\,f(Z)\,dZ\,du\,dv +$$

$$+ \int_{-\infty}^{0}\int_{0}^{\infty} \exp\{-\frac{1}{2k_1''}(u^2 - 2\rho_{xy.z}\,uv + v^2)\}\int_{u/b_1''}^{\infty} Z\,f(Z)\,dZ\,du\,dv +$$

$$+ \int_{0}^{\infty}\int_{-\infty}^{0} \exp\{-\frac{1}{2k_1''}(u^2 - 2\rho_{xy.z}\,uv + v^2)\}\int_{v/b_2''}^{\infty} Z\,f(Z)\,dZ\,du\,dv +$$

$$+ \int_{0}^{\infty}\int_{0}^{\infty} \exp\{-\frac{1}{2k_1''}(u^2 - 2\rho_{xy.z}\,uv + v^2)\}\int_{u/b_1''}^{\infty} Z\,f(Z)\,dZ\,du\,dv -$$

$$- \int_{0}^{\infty}\int_{0}^{b_1''v/b_2''} \exp\{-\frac{1}{2k_1''}(u^2 - 2\rho_{xy.z}\,uv + v^2)\}\int_{u/b_1''}^{v/b_2''} Z\,f(Z)\,dZ\,du\,dv] \ . \qquad 4.80$$

The integration of the inside integrals of Eq. 4.80 leads to

$$E\{Z\} = k_z\,\sigma_{x.z}\,\sigma_{y.z}\,\frac{\sigma_z}{\sqrt{2\pi}}\,[\int_{-\infty}^{0}\exp\{-\frac{1}{2}v^2\}\int_{-\infty}^{0}\exp\{-\frac{1}{2k_1''}(u-\rho_{xy.z}\,v)^2\}du\,dv -$$

$$- \int_{-\infty}^{0}\exp\{-\frac{1}{2}k_3''v^2\}\int_{\infty}^{0}\exp\{-\frac{1}{2}(\sqrt{k_4''}\,u - \frac{\rho_{xy.z}}{k_1''\sqrt{k_4''}}\,v)^2\}du\,dv -$$

$$- \int_{-\infty}^{0}\exp\{-\frac{1}{2}k_2''v^2\}\int_{\infty}^{0}\exp\{-\frac{1}{2k_1''}(u-\rho_{xy.z}v)^2\}du\,dv +$$

$$+ \int_{-\infty}^{0}\exp\{-\frac{1}{2}k_3''v^2\}\int_{-\infty}^{0}\exp\{-\frac{1}{2}(\sqrt{k_4''}\,u - \frac{\rho_{xy.z}}{k_1''\sqrt{k_4''}}\,v)^2\}du\,dv +$$

$$+ \int_{-\infty}^{0}\exp\{-\frac{1}{2}k_2''v^2\}\int_{b_1''v/b_2''}^{0}\exp\{-\frac{1}{2k_1''}(u-\rho_{xy.z}\,v)^2\}du\,dv -$$

$$- \int_{-\infty}^{0}\exp\{-\frac{1}{2}k_3''v^2\}\int_{b_1''v/b_2''}^{0}\exp\{-\frac{1}{2}(\sqrt{k_4''}\,u - \frac{\rho_{xy.z}}{k_1''\sqrt{k_4''}}\,v)^2\}du\,dv] \ . \qquad 4.81$$

where the constants k_1'' and k_z are given by Eq. 4.78, and k_2'', k_3'', and k_4'' are

$$k_2'' = \frac{1 + (b_2'' \sigma_z)^2}{(b_2'' \sigma_z)^2} \,, \qquad 4.82$$

$$k_3'' = \frac{1 + (b_1'' \sigma_z)^2}{k_1'' + (b_1'' \sigma_z)^2} \,, \text{ and } \; k_4'' = \frac{k_1'' + (b_1'' \sigma_z)^2}{k_1'' (b_1'' \sigma_z)^2} \;.$$

Since all the integrals of Eq. 4.81 are of the same type as those of Eq. 4.56, their solutions as given by Eqs. 4.57 through 4.60 are used here.

The first integral of Eq. 4.81, denoted by I_1'', with $a_1 = 1$, $a_2 = -\infty$, $a_3 = k_1''$, $a_4 = 1$, and $a_5 = \rho_{xy.z}$ is obtained from Eq. 4.60 as

$$I_1'' = (2\pi)\sqrt{k_1''}\ [\frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{\sqrt{k_1''}})] \,. \qquad 4.83$$

The second integral of Eq. 4.81, denoted by I_2'', with $a_1 = k_3''$, $a_2 = \infty$, $a_3 = 1$, $a_4 = \sqrt{k_4''}$ and $a_5 = \rho_{xy.z}/k_1''\sqrt{k_4''}$ is obtained from Eq. 4.58 as

$$I_2'' = \frac{(2\pi)\ \sqrt{k_1''}}{\sqrt{k_1''}\ \sqrt{k_3''}\ \sqrt{k_4''}}\ [-\frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{k_1''\sqrt{k_3''}\ \sqrt{k_4''}})] \,. \qquad 4.84$$

The third integral of Eq. 4.81, denoted by I_3'', with $a_1 = k_2''$, $a_2 = \infty$, $a_3 = k_1''$, $a_4 = 1$, and $a_5 = \rho_{xy.z}$ is obtained from Eq. 4.58 as

$$I_3'' = (2\pi)\ \frac{\sqrt{k_1''}}{\sqrt{k_2''}}\ [-\frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{\sqrt{k_1''}\ \sqrt{k_2''}})] \,. \qquad 4.85$$

The fourth integral of Eq. 4.81, denoted by I_4'', with $a_1 = k_3''$, $a_2 = -\infty$, $a_3 = 1$, $a_4 = \sqrt{k_4''}$, and $a_5 = \rho_{xy.z}/k_1''\sqrt{k_4''}$ is obtained from Eq. 4.60 as

$$I_4'' = \frac{(2\pi)\ \sqrt{k_1''}}{\sqrt{k_1''}\ \sqrt{k_3''}\ \sqrt{k_4''}}\ [\frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{k_1''\ \sqrt{k_3''}\ \sqrt{k_4''}})] \,, \qquad 4.86$$

The fifth integral of Eq. 4.81, denoted by I_5'', with $a_1 = k_2''$, $a_2 = b_1''/b_2''$, $a_3 = k_1''$, $a_4 = 1$, and $a_5 = \rho_{xy.z}$ is obtained from Eq. 4.57 as

$$I_5'' = (2\pi)\ \frac{\sqrt{k_1''}}{\sqrt{k_2''}}\ [\frac{1}{2\pi} \arctan (\frac{b_1'' - b_2''\rho_{xy.z}}{b_2''\sqrt{k_1''}\ \sqrt{k_2''}}) + \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{\sqrt{k_1''}\ \sqrt{k_2''}})] \,. \qquad 4.87$$

Finally the sixth integral of Eq. 4.81, denoted by I_6'', with $a_1 = k_3''$, $a_2 = b_1''/b_2''$, $a_3 = 1$, $a_4 = \sqrt{k_4''}$ and $a_5 = \rho_{xy.z}/k_1''\sqrt{k_4''}$ is obtained from Eq. 4.57 as

$$I_6'' = \frac{(2\pi)\ \sqrt{k_1''}}{\sqrt{k_1''}\ \sqrt{k_3''}\ \sqrt{k_4''}}\ [\frac{1}{2\pi} \arctan (\frac{b_1''\,k_1''\,k_4'' - b_2''\rho_{xy.z}}{b_2''\,k_1''\sqrt{k_3''}\ \sqrt{k_4''}}) + \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{k_1''\sqrt{k_3''}\ \sqrt{k_4''}})] \,. \qquad 4.88$$

Substituting Eqs. 4.83 through 4.88 into 4.81, and since Eq. 4.78 gives $k_z \sigma_{x.z} \sigma_{y.z} \sqrt{k_1''} = 1/(2\pi)$, then

$$\begin{aligned} E\{Z\} = \frac{\sigma_z}{\sqrt{2\pi}}\ \{ & \frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{\sqrt{k_1''}}) + \frac{1}{\sqrt{k_2''}}\ [\frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{b_1'' - b_2''\rho_{xy.z}}{b_2''\sqrt{k_1''}\sqrt{k_2''}})] + \\ & + \frac{1}{\sqrt{k_1''}\sqrt{k_3''}\sqrt{k_4''}}\ [\frac{1}{2} - \frac{1}{2\pi} \arctan (\frac{\rho_{xy.z}}{k_1''\sqrt{k_3''}\sqrt{k_4''}}) - \frac{1}{2\pi} \arctan (\frac{b_1''\,k_1''\,k_4'' - b_2''\rho_{xy.z}}{b_2''\,k_1''\sqrt{k_3''}\sqrt{k_4''}})]\ \} \,. \end{aligned} \qquad 4.89$$

Substituting the derived expected values $E\{X\}$, $E\{Y\}$ and $E\{Z\}$ as given by Eqs. 4.65, 4.76 and 4.89, respectively, into Eq. 4.42 gives the expected value of the maximum M_3 and consequently the expected value of the range R_3.

4.3 Expected Value of Range of Independent Random Variables with Changing Standard Deviation

The expected value of ranges R_1, R_2, and R_3 for independent components of partial sums are derived here based on the above derived general expressions.

For $n = 1$, Eq. 4.27 holds without any modification.

For $n = 2$, the difference $Y - X$ of Eq. 4.25 is $x_2 - y_2$; therefore, $\mathrm{Var}\{Y\text{-}X\} = \mathrm{Var}(x_2 - y_2) = \sigma_2^2$. Furthermore, Eqs. (7) and (8) of the Appendix give $\mathrm{Var}\ X = \sigma_1^2$ and $\mathrm{Var}\ Y = \sigma_1^2 + \sigma_2^2$, so that Eq. 4.41 gives the expected value of the range R_2 as

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\left[\frac{1}{2}\sigma_1 + \frac{1}{2}\sigma_2 + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)^{1/2}\right] . \quad 4.90$$

For the particular case of i.i.d. random variables, $[\mathrm{Var}\ X]^{1/2} = \sigma_1 = \sigma_2$, so that Eq. 4.90 becomes

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\{[\mathrm{Var}\ X]^{1/2} + \frac{1}{2}[\mathrm{Var}\ Y]^{1/2}\}.$$

By using the notation $S_1 = X$ and $S_2 = Y$, finally

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\{[\mathrm{Var}\ S_1]^{1/2} + \frac{1}{2}[\mathrm{Var}\ S_2]^{1/2}\} \quad 4.91$$

which is in agreement with Spitzer's formula given by Eq. 2.23. For the particular case of the standard normal variable, Eq. 4.91 further simplifies to

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\left[1 + \frac{1}{\sqrt{2}}\right] , \quad 4.92$$

in agreement with Anis' and Lloyd's formula given by Eq. 2.9.

For $n = 3$, the expected values of X, Y, and Z are first evaluated as given by Eqs. 4.65, 4.76, and 4.89, respectively.

Evaluation of $E\{X\}$ of Eq. 4.65: Substitution of $\rho_{yz.x}$ of Eq. (17), and constants k_1 and k_2, and k_4 of Eqs. (19) and (20) of the Appendix leads to

$$\frac{1}{\sqrt{k_2}} = 0 , \quad \frac{1}{\sqrt{k_4}} = 0 \quad \text{and} \quad \frac{\rho_{yz.x}}{\sqrt{k_1}} = \frac{\sigma_2}{\sigma_3} ,$$

which substituted into Eq. 4.65 give

$$E\{X\} = \frac{\sigma_1}{\sqrt{2\pi}}\left[\frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{\sigma_2}{\sigma_3}\right)\right] . \quad 4.93$$

Evaluation of $E\{Y\}$ of Eq. 4.76: Substitution of $\rho_{xz.y}$ of Eq. (17), and constants b_1' and b_2', k_1' and k_2', and k_3' and k_4' of Eqs. (21), (22), and (23) of the Appendix leads to

$$\frac{1}{\sqrt{k_2'}} = 0 , \quad \frac{\rho_{xz.y}}{\sqrt{k_1'}} = 0 ,$$

$$\frac{1}{\sqrt{k_1'}\sqrt{k_3'}\sqrt{k_4'}} = \frac{\sigma_2}{(\sigma_1^2 + \sigma_2^2)^{1/2}}$$

and

$$\frac{b_1' k_1' k_4' + b_2' \rho_{xz.y}}{b_2' k_1' \sqrt{k_3'}\sqrt{k_4'}} = \infty$$

which, substituted into Eq. 4.76, gives

$$E\{Y\} = \frac{\sigma_y}{\sqrt{2\pi}}\left[\frac{1}{4} + \frac{1}{4}\frac{\sigma_2}{(\sigma_1^2 + \sigma_2^2)^{1/2}}\right] .$$

Since Eq. (8) of Appendix, gives $\sigma_y = (\sigma_1^2 + \sigma_2^2)^{1/2}$, then

$$E\{Y\} = \frac{1}{\sqrt{2\pi}}\left[\frac{1}{4}\sigma_2 + \frac{1}{4}(\sigma_1^2 + \sigma_2^2)^{1/2}\right] . \quad 4.94$$

Evaluation of $E\{Z\}$ of Eq. 4.89: Substitution of $\rho_{xy.z}$ of Eq. (17) and constants b_1'' and b_2'', k_1'' and k_2'', and k_3'' and k_4'' of Eqs. (24), (25), and (26) of the Appendix leads to

$$\frac{\rho_{xy.z}}{\sqrt{k_1''}} = \frac{\sigma_1\sigma_3}{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} ,$$

$$\frac{1}{\sqrt{k_2''}} = \frac{\sigma_3}{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}}$$

$$\frac{b_1'' - b_2'' \rho_{xy.z}}{b_2'' \sqrt{k_1''}\sqrt{k_2''}} = \frac{\sigma_2}{\sigma_1}, \quad \frac{1}{\sqrt{k_1''}\sqrt{k_3''}\sqrt{k_4''}}$$

$$= \frac{(\sigma_2^2 + \sigma_3^2)^{1/2}}{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \quad \frac{\rho_{xy.z}}{k_1'' \sqrt{k_3''}\sqrt{k_4''}}$$

$$= \frac{\sigma_1 \sigma_3 (\sigma_2^2 + \sigma_3^2)^{1/2}}{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}}, \quad \frac{b_1'' k_1'' k_4'' = b_2'' \rho_{xy.z}}{b_2'' k_1'' \sqrt{k_3''}\sqrt{k_4''}}$$

$$= \frac{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)}{\sigma_1 \sigma_3 (\sigma_2^2 + \sigma_3^2)^{1/2}}.$$

Substituting the above expressions into Eq. 4.89, then

$$E\{Z\} = \frac{\sigma_z}{\sqrt{2\pi}} \{ \frac{1}{4} + \frac{1}{2\pi} \arctan \left(\frac{\sigma_1 \sigma_3}{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \right) + \frac{\sigma_3}{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \left[\frac{1}{4} + \frac{1}{2\pi} \arctan \left(\frac{\sigma_2}{\sigma_1} \right) \right] + \frac{(\sigma_2^2 + \sigma_3^2)^{1/2}}{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \left[\frac{1}{2} - \frac{1}{2\pi} \arctan \left(\frac{\sigma_1 \sigma_3 (\sigma_2^2 + \sigma_3^2)^{1/2}}{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)} \right) - \frac{1}{2\pi} \arctan \left(\frac{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)}{\sigma_1 \sigma_3 (\sigma_2^2 + \sigma_3^2)^{1/2}} \right) \right] \}.$$

Since Eq. (9) of the Appendix gives $\sigma_z = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}$, the above equation simplifies to

$$E\{Z\} = \frac{1}{\sqrt{2\pi}} [\frac{1}{4} \sigma_3 + \frac{1}{4}(\sigma_2^2 + \sigma_3^2)^{1/2} + \frac{1}{4}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2} + \sigma_3 \frac{1}{2\pi} \arctan \left(\frac{\sigma_2}{\sigma_1} \right) + (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2} \frac{1}{2\pi} \arctan \left(\frac{\sigma_1 \sigma_3}{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \right)]. \tag{4.95}$$

Substituting Eqs. 4.93, 4.94, and 4.95 into Eq. 4.42, the expected value of the maximum M_3 becomes

$$E\{M_3\} = \frac{1}{\sqrt{2\pi}} \{ \frac{1}{4} (\sigma_1 + \sigma_2 + \sigma_3) + \frac{1}{4} [(\sigma_1^2 + \sigma_2^2)^{1/2} + (\sigma_2^2 + \sigma_3^2)^{1/2} + (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}] + \sigma_1 \frac{1}{2\pi} \arctan \left(\frac{\sigma_2}{\sigma_3} \right) + \sigma_3 \frac{1}{2\pi} \arctan \left(\frac{\sigma_2}{\sigma_1} \right) + (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2} \frac{1}{2\pi} \arctan \left(\frac{\sigma_1 \sigma_3}{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \right) \}.$$

Consequently, the expected value of the range R_3 is given by

$$E\{R_3\} = \sqrt{\frac{2}{\pi}} \{ \frac{1}{4} (\sigma_1 + \sigma_2 + \sigma_3) + \frac{1}{4} [(\sigma_1^2 + \sigma_2^2)^{1/2} + (\sigma_2^2 + \sigma_3^2)^{1/2} + (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}] + \sigma_1 \frac{1}{2\pi} \arctan \left(\frac{\sigma_2}{\sigma_3} \right) + \sigma_3 \frac{1}{2\pi} \arctan \left(\frac{\sigma_2}{\sigma_1} \right) + (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2} \frac{1}{2\pi} \arctan \left(\frac{\sigma_1 \sigma_3}{\sigma_2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \right) \}. \tag{4.96}$$

For the particular case of i.i.d. random variables, or $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$,

$$[\mathrm{Var}\ X]^{1/2} = [\mathrm{Var}\ S_1]^{1/2} = \sigma_1$$

$$[\mathrm{Var}\ Y]^{1/2} = [\mathrm{Var}\ S_2]^{1/2} = (\sigma_1^2 + \sigma_2^2)^{1/2} = (\sigma_2^2 + \sigma_3^2)^{1/2}$$

$$[\mathrm{Var}\ Y]^{1/2} = [\mathrm{Var}\ S_3]^{1/2} = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}$$

Eq. 4.96 takes the form

$$E\{R_3\} = \sqrt{\frac{2}{\pi}} \{ [\mathrm{Var}\ S_1]^{1/2} + \frac{1}{2} [\mathrm{Var}\ S_2]^{1/2} + \frac{1}{3} [\mathrm{Var}\ S_3]^{1/2} \}, \tag{4.97}$$

which is in agreement with Spitzer's formula given by Eq. 2.23. For the particular case of the standard normal variable, Eq. 4.97 simplifies to

$$E\{R_3\} = \sqrt{\frac{2}{\pi}} [1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}}], \tag{4.98}$$

which is in agreement with Anis' and Lloyd's formula given by Eq. 2.9.

4.4 Expected Values of Range of Equally Dependent Random Variables (Exchangeable Variables)

Exchangeable random variables have the property that the variances are the same, and the correlation between any two variables is also the same (M. Loeve, 1960). The expected range of this type of variables is of importance, especially when deriving the expected adjusted range as given in section 4.6 of this chapter.

Following D. B. Owen and G. P. Steck (1962), exchangeable variables may be generated by

$$x_t = \sqrt{\rho}\ \epsilon_o + \sqrt{1-\rho}\ \epsilon_t\ , \quad 0 \leqslant \rho < 1\ , \qquad 4.99$$

in which ϵ_o and ϵ_t are independent normal random variables with mean zero and variance unity, with $E\{x_t\} = 0$, $Var\{x_t\} = 1$, and $E\{x_t\, x_{t+u}\} = \rho$.

For $n = 1$, Eq. 4.27 holds without modification. For $n = 2$ the difference Y-X of Eq. 4.25 is equal to x_2-y_2. Since an equal variance is assumed, then

$$Var\ \{Y - X\} = Var\ \{x_2 - y_2\} = \sigma^2$$

Because Eq. (29) of Appendix gives Var $X = \sigma^2$, Eq. 4.41 becomes

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\ \{[Var\ X]^{1/2} + \frac{1}{2}\ [Var\ Y]^{1/2}\}.$$

With the notation $S_1 = X$ and $S_2 = Y$, finally,

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\ \{[Var\ S_1]^{1/2} + \frac{1}{2}[Var\ S_2]^{1/2}\}\ . \qquad 4.100$$

By using Eqs. (29) and (30) of the Appendix, the explicit equation for the expected value of R_2 becomes

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\ \sigma[\ 1 + \frac{1}{\sqrt{2}}\ (1+\rho)^{1/2}\ ]\ . \qquad 4.101$$

For $n = 3$, the expected values of X, Y and Z are first evaluated as given by Eqs. 4.65, 4.76 and 4.89, respectively. Evaluation of E{ X } of Eq. 4.65: Substitution of $\rho_{yz.x}$ of Eq. (39), and constants b_1 and b_2, k_1 and k_3 and k_4 of Eqs. (40), (41), and (42) of the Appendix, leads to

$$\frac{\rho_{yz.x}}{\sqrt{k_1}} = (1+2\rho)^{1/2}\ , \qquad \frac{1}{\sqrt{k_2}} = \frac{\sqrt{2}\ \rho}{(1+\rho)^{1/2}}\ ,$$

$$\frac{b_1 - b_2\,\rho_{yz.x}}{b_2\sqrt{k_1}\sqrt{k_2}} = 0\ ,$$

$$\frac{1}{\sqrt{k_1}\ \sqrt{k_2}\ \sqrt{k_3}} = \rho\ ,$$

$$\frac{b_1 k_1 k_4 - b_2\,\rho_{yz.x}}{b_2 k_1\sqrt{k_3}\ \sqrt{k_4}} = \frac{(1-\rho)(1+2\rho)^{1/2}}{2\rho}\ ,$$

and $$\frac{\rho_{yz.x}}{k_1\sqrt{k_3}\sqrt{k_4}} = \rho\,(1+2\rho)^{1/2}\ .$$

By substituting these expressions into Eq. 4.65, it becomes

$$E\{X\} = \frac{\sigma}{\sqrt{2\pi}}\ \{\frac{1}{4} + \frac{1}{2\pi}$$

$$\arctan\ [(1+2\rho)^{1/2}] - \frac{\sqrt{2}}{4}\ \frac{\rho}{(1+\rho)^{1/2}}\ -$$

$$-\rho\ [\frac{1}{2\pi}\ \arctan\ (\frac{(1-\rho)(1+2\rho)^{1/2}}{2\rho})$$

$$+\frac{1}{2\pi}\ \arctan\,(\rho\,(1+2\rho)^{1/2})]\ \}.$$

After simplifying, we finally have

$$E\{X\} = \frac{\sigma}{\sqrt{2\pi}}\{\frac{1}{4} - \frac{\sqrt{2}}{4}\frac{\rho}{(1+\rho)^{1/2}}$$

$$+\frac{1}{2\pi}\arctan\ [(1+2\rho)^{1/2}]$$

$$-\rho\ \frac{1}{2\pi}\ \arctan\ [\frac{(1+2\rho)^{1/2}}{\rho}]\ \}\ . \qquad 4.102$$

Evaluation of E{Y} of Eq. 4.76: Substitution of $\rho_{xz.y}$, of Eq. (39), and constants b_1' and b_2', k_1' and k_2', and k_3' and k_4' of Eqs. (43), (44) and (45) of the Appendix, lead to

$$\frac{\rho_{xz.y}}{\sqrt{k_1'}} = 0\ , \qquad \frac{1}{\sqrt{k_2'}} = \frac{\sqrt{2}\ \rho}{(1+\rho)^{1/2}}\ ,$$

$$\frac{b_1' + b_2'\,\rho_{xz.y}}{b_2'\sqrt{k_1'}\sqrt{k_2'}} = (1+2\rho)^{1/2}\ ,$$

$$\frac{1}{\sqrt{k_1'}\sqrt{k_3'}\sqrt{k_4'}} = \frac{(1+\rho)^{1/2}}{\sqrt{2}} ,$$

$$\frac{b_1' k_1' k_4' + b_2' \rho_{xz.y}}{b_2' k_1' \sqrt{k_3'}\sqrt{k_4'}} = \frac{(1+2\rho)^{1/2}}{\rho} ,$$

and $$\frac{\rho_{xz.y}}{k_1' \sqrt{k_3'}\sqrt{k_4'}} = 0 .$$

By substituting these expressions into Eq. 4.76, it becomes

$$E\{Y\} = \frac{\sigma_y}{\sqrt{2\pi}}$$

$$\{\frac{1}{4} - \frac{\sqrt{2}\ \rho}{(1+\rho)^{1/2}} [\frac{1}{4} + \frac{1}{2\pi} \arctan (1+2\rho)^{1/2}] +$$

$$+ \frac{(1+\rho)^{1/2}}{\sqrt{2}} \frac{1}{2\pi} \arctan [\frac{(1+2\rho)^{1/2}}{\rho}] \} .$$

Since Eq. (30) of the Appendix gives $\sigma_y = \sqrt{2}\ \sigma\ (1+\rho)^{1/2}$, then

$$E\{Y\} = \frac{\sigma}{\sqrt{2\pi}} \{\frac{\sqrt{2}}{4} (1+\rho)^{1/2} - \frac{1}{2} \rho$$

$$+ (1+\rho) \frac{1}{2\pi} \arctan [\frac{(1+2\rho)^{1/2}}{\rho}] - 2\rho$$

$$\frac{1}{2\pi} \arctan (1+2\rho)^{1/2} \} . \qquad 4.103$$

Evaluation of $E\{Z\}$ of Eq. 4.89: Substitution of $\rho_{xy.z}$ of Eq. (39) and constants b_1'' and b_2'', k_1'' and k_2'', and k_3'' and k_4'' of Eqs. (46), (47), and (48) of the Appendix, leads to

$$\frac{\rho_{xy.z}}{\sqrt{k_1''}} = \frac{1}{\sqrt{3}} , \qquad \frac{1}{\sqrt{k_2''}} = \frac{(1+2\rho)^{1/2}}{\sqrt{3}}$$

$$\frac{b_1'' - b_2'' \rho_{xy.z}}{b_2'' \sqrt{k_1''}\sqrt{k_2''}} = (1+2\rho)^{1/2} ,$$

$$\frac{1}{\sqrt{k_1''}\sqrt{k_3''}\sqrt{k_4''}} = \frac{\sqrt{2}\ (1+2\rho)^{1/2}}{\sqrt{3}(1+\rho)^{1/2}}$$

$$\frac{b_1'' k_1'' k_4'' - b_2'' \rho_{xy.z}}{b_2'' k_1'' \sqrt{k_3''}\sqrt{k_4''}} = \frac{3(1+\rho)^{1/2}}{\sqrt{2}(1+2\rho)^{1/2}} ,$$

$$\frac{\rho_{xy.z}}{k_1'' \sqrt{k_3''}\sqrt{k_4''}} = \frac{\sqrt{2}(1+2\rho)^{1/2}}{3(1+\rho)^{1/2}}$$

By substituting these expressions into Eq. 4.89, it becomes

$$E\{Z\} = \frac{\sigma_z}{\sqrt{2\pi}} \{\frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{1}{\sqrt{3}})$$

$$+ \frac{(1+2\rho)^{1/2}}{\sqrt{3}} [\frac{1}{4} + \frac{1}{2\pi} \arctan (1+2\rho)^{1/2}] +$$

$$+ \frac{\sqrt{2}(1+2\rho)^{1/2}}{\sqrt{3}(1+\rho)^{1/2}} [\frac{1}{2} - \frac{1}{2\pi}$$

$$\arctan (\frac{\sqrt{2}(1+2\rho)^{1/2}}{3(1+\rho)^{1/2}})$$

$$- \frac{1}{2\pi} \arctan (\frac{3(1+\rho)^{1/2}}{\sqrt{2}(1+2\rho)^{1/2}})] \}$$

Since Eq. (31) of Appendix gives $\sigma_z = \sqrt{3}\ \sigma(1+2\rho)^{1/2}$, the above equation simplifies to

$$E\{Z\} = \frac{\sigma}{\sqrt{2\pi}}$$

$$\{(1+2\rho) [\frac{1}{4} + \frac{1}{2\pi} \arctan (1+2\rho)^{1/2}] +$$

$$+ \frac{\sqrt{2}}{4} \frac{(1+2\rho)}{(1+\rho)^{1/2}} + \frac{1}{\sqrt{3}} (1+2\rho)^{1/2} \} . \qquad 4.104$$

Substituting Eqs. 4.102, 4.103, and 4.104 into Eq. 4.42 gives the expected value of the maximum M_3 as

$$E\{M_3\} = \frac{\sigma}{\sqrt{2\pi}}$$

$$[1 + \frac{1}{\sqrt{2}} (1+\rho)^{1/2} + \frac{1}{\sqrt{3}} (1+2\rho)^{1/2}] .$$

Consequently the expected value of the range R_3 becomes

$$E\{R_3\} = \sqrt{\frac{2}{\pi}}\ \sigma\ [\ 1$$

$$+ \frac{1}{\sqrt{2}}\ (1+\rho)^{1/2} + \frac{1}{\sqrt{3}}\ (1+2\rho)^{1/2}\ ]\ . \qquad 4.105$$

Equations (29), (30), and (31) of the Appendix give $[\text{Var } X]^{1/2} = \sigma, [\text{Var } Y]^{1/2} = \sigma\sqrt{2}(1+\rho)^{1/2}$, and $[\text{Var } Z]^{1/2} = \sigma\sqrt{3}(1+2\rho)^{1/2}$, and a substitution of $S_1 = X$, $S_2 = Y$, and $S_3 = Z$, as indicated by Eq. 4.25, leads to

$$E\{R_3\} = \sqrt{\frac{2}{\pi}}\ \{[\text{Var } S_1]^{1/2}$$

$$+ \frac{1}{2}\ [\text{Var } S_2]^{1/2} + \frac{1}{3}\ [\text{Var } S_3]^{1/2}\ \}\ . \qquad 4.106$$

In summary, the expected value of the range for n = 1, 2 and 3 of exchangeable random variables are:

$$E\{R_1\} = \sqrt{\frac{2}{\pi}}\ [\text{Var } S_1]^{1/2}\ ,$$

$$E\{R_2\} = \sqrt{\frac{2}{\pi}}\ \{[\text{Var } S_1]^{1/2} + \frac{1}{2}\ [\text{Var } S_2]^{1/2}\ \}\ ,$$

$$E\{R_3\} = \sqrt{\frac{2}{\pi}}\ \{[\text{Var } S_1]^{1/2}$$

$$+ \frac{1}{2}\ [\text{Var } S_2]^{1/2} + \frac{1}{3}\ [\text{Var } S_3]^{1/2}\ \}\ .$$

As a conclusion, the general expression for the expected range of exchangeable random variables can be written as

$$E\{R_n\} = \sqrt{\frac{2}{\pi}}\ \sum_{i=1}^{n}\ i^{-1}\ [\text{Var } S_i]^{1/2}\ , \qquad 4.107$$

in agreement with Spitzer's formula (Eq. 2.23).

4.5 Expected Values of the Range of First-Order Markov Linearly Dependent Variables

The exact expected values of the range for n = 1,2, and 3 are given here for the case of a stationary first-order Markov model.

For n = 1, Eq. 4.27 holds also without modification. For n = 2, Eqs. 4.100 and 4.101, valid for exchangeable random variables, are also valid in this case because only two random variables are considered.

For n = 3 the expected values of X,Y, and Z given by Eqs. 4.65, 4.76 and 4.89, respectively, are first evaluated.

Evaluation of E{X} of Eq. 4.65: Substitution of $\rho_{yz.x}$ of Eq. (59), and constants b_1 and b_2, k_1 and k_2, and k_3 and k_4 of Eqs. (60), (61), and (62) of the Appendix, leads to

$$\frac{\rho_{yz.x}}{\sqrt{k_1}} = (1+\rho)\qquad , \qquad \frac{1}{\sqrt{k_2}} = \frac{\rho(1+\rho)^{1/2}}{\sqrt{2}}$$

$$\frac{b_1 - b_2\,\rho_{yz.x}}{b_2\sqrt{k_1}\,\sqrt{k_2}} = \frac{\rho}{\sqrt{2}(1+\rho)^{1/2}}\ ,$$

$$\frac{1}{\sqrt{k_1}\,\sqrt{k_3}\,\sqrt{k_4}} = \rho$$

$$\frac{b_1 k_1 k_4 - b_2\,\rho_{yz.x}}{b_2\,k_1\sqrt{k_3}\,\sqrt{k_4}} = \frac{1}{\rho(1+\rho)}\ ,$$

$$\frac{\rho_{yz.x}}{k_1\sqrt{k_3}\,\sqrt{k_4}} = \rho(1+\rho)$$

By substituting these expressions into Eq. 4.65, it becomes

$$E\{X\} = \frac{\sigma_x}{\sqrt{2\pi}}\ \{\frac{1}{4} + \frac{1}{2\pi}\arctan\ (1+\rho)$$

$$-\frac{\rho(1+\rho)^{1/2}}{\sqrt{2}}\ [\frac{1}{4} - \frac{1}{2\pi}\ \arctan\ (\frac{\rho}{\sqrt{2}(1+\rho)^{1/2}})] -$$

$$-\ \rho[\frac{1}{2\pi}\ \arctan\ (\frac{1}{\rho(1+\rho)}) + \frac{1}{2\pi}\ \arctan\ (\rho(1+\rho))]\ \}$$

which simplifies further to

$$E\{X\} = \frac{\sigma}{\sqrt{2\pi}}\ \{\frac{1}{4}(1-\rho) - \frac{1}{4}\ \frac{\rho(1+\rho)^{1/2}}{\sqrt{2}}$$

$$+ \frac{1}{2\pi}\ \arctan\ (1+\rho)\ +$$

$$+ \frac{\rho(1+\rho)^{1/2}}{\sqrt{2}}\ \frac{1}{2\pi}\ \arctan\ [\frac{\rho}{\sqrt{2}(1+\rho)^{1/2}}]\ \} \qquad 4.108$$

Evaluation of $E\{Y\}$ of Eq. 4.76: Substitution of $\rho_{xz.y}$ of Eq. (59), and constants b_1' and b_2', k_1', and k_2', and k_3' and k_4' of Eqs. (63), (64), and (65) of the Appendix, leads to

$$\frac{\rho_{xz.y}}{\sqrt{k_1'}} = -\frac{\rho}{\sqrt{2}(1+\rho)^{1/2}} \quad , \quad \frac{1}{\sqrt{k_2'}} = \frac{\rho(1+\rho)^{1/2}}{\sqrt{2}}$$

$$\frac{b_1' + b_2' \rho_{xz.y}}{b_2'\sqrt{k_1'}\sqrt{k_2'}} = (1+\rho) \quad , \quad \frac{1}{\sqrt{k_1'}\sqrt{k_3'}\sqrt{k_4'}} = \frac{(1+\rho)^{1/2}}{\sqrt{2}}$$

$$\frac{b_1' k_1' k_4' + b_2' \rho_{xz.y}}{b_2' k_1' \sqrt{k_3'}\sqrt{k_4'}} = \frac{2}{\rho} , \quad \frac{\rho_{xz.y}}{k_1' \sqrt{k_3'}\sqrt{k_4'}} = -\frac{\rho}{2}$$

By substituting these expressions into Eq. 4.76, it becomes

$$E\{Y\} = \frac{\sigma_y}{\sqrt{2\pi}} \{\frac{1}{4} + \frac{1}{2\pi} \arctan [-\frac{\rho}{\sqrt{2}\,(1+\rho)^{1/2}}]$$

$$-\frac{\rho\,(1+\rho)^{1/2}}{\sqrt{2}} [\frac{1}{4} + \frac{1}{2\pi} \arctan (1+\rho)] +$$

$$+\frac{(1+\rho)^{1/2}}{\sqrt{2}} [\frac{1}{2\pi} \arctan (\frac{2}{\rho}) - \frac{1}{2\pi}$$

$$\arctan (-\frac{\rho}{2})]\}$$

Since Eq. (50) of the Appendix gives $\sigma_y = \sigma\sqrt{2}\,(1+\rho)^{1/2}$, this expression further simplifies to

$$E\{Y\} = \frac{\sigma}{\sqrt{2\pi}} \{\frac{1}{4}(1-\rho^2)$$

$$+\frac{\sqrt{2}}{4}(1+\rho)^{1/2} - \rho(1+\rho)\frac{1}{2\pi} \arctan (1+\rho) -$$

$$-\sqrt{2}(1+\rho)^{1/2} \frac{1}{2\pi} \arctan [\frac{\rho}{\sqrt{2}(1+\rho)^{1/2}}]\} \qquad 4.109$$

Evaluation of $E\{Z\}$ of Eq. 4.89: Substitution of $\rho_{xy.z}$ of Eq. (59), and constants b_1'' and b_2'', k_1'' and k_2'', and k_3'' and k_4'' of Eqs. (66), (67) and (68) of the Appendix, leads to

$$\frac{\rho_{xy.z}}{\sqrt{k_1''}} = \frac{(1+\rho)^2}{(3+4\rho+2\rho^2)^{1/2}} \quad , \quad \frac{1}{\sqrt{k_2''}} = \frac{(1+\rho+\rho^2)}{(3+4\rho+2\rho^2)^{1/2}} ,$$

$$\frac{b_1'' - b_2'' \rho_{xy.z}}{b_2''\sqrt{k_1''}\sqrt{k_2''}} = (1+\rho) \quad ,$$

$$\frac{1}{\sqrt{k_1''}\sqrt{k_3''}\sqrt{k_4''}} = \frac{(1+\rho)^{1/2}(2+\rho)}{\sqrt{2}(3+4\rho+2\rho^2)^{1/2}} ,$$

$$\frac{b_1'' k_1'' k_4'' - b_2'' \rho_{xy.z}}{b_2'' k_1'' \sqrt{k_3''}\sqrt{k_4''}} = \frac{(1+\rho)^{1/2}(3+2\rho+\rho^2)}{\sqrt{2}(1+\rho+\rho^2)} ,$$

$$\frac{\rho_{xy.z}}{k_1'' \sqrt{k_3''}\sqrt{k_4''}} = \frac{(1+\rho)^{5/2}(2+\rho)}{\sqrt{2}(3+4\rho+2\rho^2)} .$$

By substituting these expressions into Eq. 4.89, it becomes

$$E\{Z\} = \frac{\sigma_z}{\sqrt{2\pi}} \{\frac{1}{4} + \frac{1}{2\pi} \arctan [\frac{(1+\rho)^2}{(3+4\rho+2\rho^2)^{1/2}}]$$

$$+\frac{(1+\rho+\rho^2)}{(3+4\rho+2\rho^2)^{1/2}} [\frac{1}{4} +$$

$$+\frac{1}{2\pi} \arctan (1+\rho)] + \frac{(1+\rho)^{1/2}(2+\rho)}{\sqrt{2}(3+4\rho+2\rho^2)^{1/2}} [\frac{1}{2}$$

$$-\frac{1}{2\pi} \arctan (\frac{(1+\rho)^{5/2}(2+\rho)}{\sqrt{2}(3+4\rho+2\rho^2)})$$

$$-\frac{1}{2\pi} \arctan (\frac{(1+\rho)^{1/2}(3+2\rho+\rho^2)}{\sqrt{2}\,(1+\rho+\rho^2)})]\}$$

After further simplification and since Eq. (51) of the Appendix gives $\sigma_z = \sigma(3+4\rho+2\rho^2)^{1/2}$, then

$$E\{Z\} = \frac{\sigma}{\sqrt{2\pi}} \{(1+\rho+\rho^2)[\frac{1}{4} + \frac{1}{2\pi} \arctan (1+\rho)]$$

$$+\frac{(2+\rho)(1+\rho)^{1/2}}{\sqrt{2}} \frac{1}{2\pi} \arctan [\frac{\sqrt{2}(1+\rho)^{1/2}}{\rho}] +$$

$$+(3+4\rho+2\rho^2)^{1/2} [\frac{1}{4} + \frac{1}{2\pi} \arctan (\frac{(1+\rho)^2}{(3+4\rho+2\rho^2)^{1/2}})]\} \qquad 4.110$$

Substituting Eqs. 4.108, 4.109, and 4.110 into Eq. 4.42 gives the expected value of the maximum M_3 as

$$E\{M_3\} = \frac{\sigma}{\sqrt{2\pi}}\{[\frac{3}{4} + 2\ \frac{1}{2\pi}\ \arctan(1+\rho)]$$

$$+\sqrt{2}(1+\rho)^{1/2}\ [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{2+2\rho-\rho^2}{2\sqrt{2}\,\rho(1+\rho)^{1/2}})]\ +$$

$$+(3+4\rho+2\rho^2)^{1/2}\ [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{(1+\rho)^2}{(3+4\rho+2\rho^2)^{1/2}})]\ \}$$

Consequently, the expected value of the range R_3 is

$$E\{R_3\} = \sqrt{\frac{2}{\pi}}\ \sigma\ \{[\frac{3}{4} + 2\ \frac{1}{2\pi}\ \arctan(1+\rho)]$$

$$+\sqrt{2}(1+\rho)^{1/2}\ [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{2+2\rho-\rho^2}{2\sqrt{2}\,\rho(1+\rho)^{1/2}})]\ +$$

$$+(3+4\rho+2\rho^2)^{1/2}\ [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{(1+\rho)^2}{(3+4\rho+2\rho^2)^{1/2}})\ ] \qquad 4.111$$

Equations (49), (50), and (51) of the Appendix give $[\mathrm{Var}\ X]^{1/2} = \sigma$, $[\mathrm{Var}\ Y]^{1/2} = \sigma\sqrt{2}(1+\rho)^{1/2}$, and $[\mathrm{Var}\ Z]^{1/2} = \sigma(3+4\rho+2\rho^2)^{1/2}$. A substitution of $S_1 = X$, $S_2 = Y$, and, $S_3 = Z$, as indicated by Eq. 4.25, leads to

$$E\{R_3\} = \sqrt{\frac{2}{\pi}}\ \{\ [\frac{3}{4} + \frac{2}{2\pi}$$

$$\arctan(1+\rho)\ ]\ [\mathrm{Var}\ S_1]^{1/2}\ +$$

$$+\ [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{2+2\rho-\rho^2}{2\sqrt{2}\,\rho\,(1+\rho)^{1/2}})\ ]$$

$$[\mathrm{Var}\ S_2]^{1/2} + [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{(1+\rho)^2}{(3+4\rho+2\rho^2)^{1/2}})\ ]$$

$$[\mathrm{Var}\ S_3]^{1/2}\ \},$$

or

$$E\{R_3\} = \sqrt{\frac{2}{\pi}}\ \{c_1(\rho)\ [\mathrm{Var}\ S_1]^{1/2}$$

$$+\,c_2(\rho)\ \frac{1}{2}\ [\mathrm{Var}\ S_2]^{1/2} + c_3(\rho)\ \frac{1}{3}\ [\mathrm{Var}\ S_3]^{1/2}\ \}, \qquad 4.112$$

with $c_1(\rho) = [\frac{3}{4} + \frac{2}{2\pi}\ \arctan(1+\rho)\ ]$,

$$c_2(\rho) = 2\ [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{2+2\rho-\rho^2}{2\sqrt{2}\,\rho\,(1+\rho)^{1/2}})\ ],$$

and

$$c_3(\rho) = 3\ [\frac{1}{4} + \frac{1}{2\pi}\ \arctan(\frac{(1+\rho)^2}{(3+4\rho+2\rho^2)^{1/2}})\ ].$$

For the particular case of $\rho = 0$, $c_1(\rho) = c_2(\rho) = c_3(\rho) = 1$, then Eq. 4.112 simplifies to

$$E\{R_3\} = \sqrt{\frac{2}{\pi}}\ \{\,[\mathrm{Var}\ S_1]^{1/2}$$

$$+\ \frac{1}{2}\ [\mathrm{Var}\ S_2]^{1/2} + \frac{1}{3}\ [\mathrm{Var}\ S_3]^{1/2}\ \}$$

in agreement with Spitzer's equation (Eq. 2.23).

4.6 A Note on the Expected Value of Adjusted Range

The expected values of adjusted range of exchangeable random variables are shown to be given by the same formula as for the expected values of the range of a transformed variable which also shows the property of exchangeability.

Let us assume the inputs are exchangeable variables, as defined in Section 4.4, while the outputs are equal to $\alpha\bar{x}_n$, with $0 < \alpha \leqslant 1$ and $\bar{x}_n$ the sample mean. Then the adjusted partial sums, as given in general by Eq. 3.2 are

$$\begin{aligned}
S^*_o &= 0, \\
S^*_1 &= S^*_o + (x_1 - \alpha\bar{x}_n), \\
S^*_2 &= S^*_1 + (x_2 - \alpha\bar{x}_n), \qquad 4.113 \\
&\vdots \\
S^*_n &= S^*_{n-1} + (x_n - \alpha\bar{x}_n).
\end{aligned}$$

By using the transformation

$$w_t = x_t - \alpha\,\bar{x}_n\ , \qquad 4.114$$

this new process, w_t, has the expected value

$$E\{w_t\} = E\{x_t\} - \alpha\,E\{\bar{x}_n\} = 0, \qquad 4.115$$

and the variance, using Eq. 4.114, is

$$\mathrm{Var}\{w_t\} = \mathrm{Var}\{x_t\} + \alpha^2\,\mathrm{Var}\{\overline{x}_n\} - 2\alpha\,\mathrm{cov}\{x_t, \overline{x}_n\}. \qquad 4.116$$

Because the variance of the sample, $\overline{x}_n$ is

$$\mathrm{Var}\{\overline{x}_n\} = \frac{1}{n^2}\,\mathrm{Var}\{S_n\} = \frac{1}{n^2}\,[\,n\,\sigma^2 + 2\sum_{i=1}^{n-1}\sum_{j=1}^{n-i}\mathrm{cov}\{x_i, x_{i+j}\}] ,$$

and since the original process x_i has equal autocorrelation coefficients, with $\mathrm{Cov}\{x_i, x_{i+j}\} = \sigma^2\rho$, the above equation becomes

$$\mathrm{Var}\{\overline{x}_n\} = \frac{\sigma^2}{n}\,[\,1 + (n-1)\,\rho\,] . \qquad 4.117$$

The covariance of x_t and $\overline{x}_n$ is

$$\mathrm{Cov}\{x_t, \overline{x}_n\} = \frac{1}{n}\,E\{x_t \sum_{i=1}^{n} x_i\}$$

$$= \frac{\sigma^2}{n}\,[\,1 + (n-1)\,\rho\,] . \qquad 4.118$$

Substituting Eqs. 4.117 and 4.118 into Eq. 4.116 leads to

$$\mathrm{Var}\{w_t\} = \frac{\sigma^2}{n}\,\{\,n + \alpha\,(\alpha - 2)\,[\,1 + (n-1)\,\rho\,]\,\}. \qquad 4.119$$

The covariance of the process w_t is

$$\mathrm{Cov}\{w_t, w_{t+k}\} = E\{x_t\,x_{t+k}\} + \alpha^2\,E\{\overline{x}_n^{\,2}\} - \alpha\,E\{x_t\,\overline{x}_n\} - \alpha\,E\{x_{t+k}\,\overline{x}_n\} .$$

Substituting Eqs. 4.117 and 4.118 into the above expression leads to

$$\mathrm{Cov}\{w_t, w_{t+k}\} = \frac{\sigma^2}{n}\,\{\,n\rho + \alpha\,(\alpha - 2)\,[\,1 + (n-1)\,\rho\,]\,\}. \qquad 4.120$$

Therefore, the autocorrelation function of, w_t is

$$\rho\,(w_t) = \frac{n\rho + \alpha\,(\alpha - 2)\,[\,1 + (n-1)\,\rho\,]}{n + \alpha\,(\alpha - 2)\,[\,1 + (n-1)\,\rho\,]} . \qquad 4.121$$

Equations 4.115, 4.119, and 4.121 show the process w_t to be second-order stationary and to have equal autocorrelation coefficients, independent of the lag k, that is, w_t is a sequence of exchangeable random variables. This property shown by the components of the adjusted partial sums is important, because, as shown in section 4.4, the expected value of the range of a sequence of partial sums whose components are exchangeable random variables may be obtained by using Eq. 4.107.

For the sequence of adjusted partial sums S_0^*, S_1^*, S_2^*, ..., S_n^*, the expected value of the adjusted range is

$$E\{R_n^*\} = \sqrt{\frac{2}{\pi}}\,\sum_{i=1}^{n} i^{-1}\,[\mathrm{Var}\,S_i^*]^{1/2} . \qquad 4.122$$

In the case of independent standard normal variables and $\alpha = 1$, Eq. 4.1 simplifies to the equation given by Solari and Anis (1957). For computing the variance of S_i for this case, Eq. 2.2 gives the general terms S_i^*, expressed by $S_i^* = S_i - iS_n/n$, so that

$$\mathrm{Var}\{S_i^*\} = \mathrm{Var}\{S_i\} + (\frac{i}{n})^2\,\mathrm{Var}\{S_n\} - 2\,\frac{i}{n}\,\mathrm{Cov}\{S_i, S_n\} .$$

For i.i.d. and standard normal variables, $\mathrm{Var}\{S_i\} = i$, $\mathrm{Var}\{S_n\} = n$, and $\mathrm{Cov}\{S_i, S_n\} = i$, so that $S_n = i$, so that

$$\mathrm{Var}\{S_i^*\} = \frac{i}{n}\,(n - i) . \qquad 4.123$$

Substituting Eq. 4.123 into Eq. 4.122 gives

$$E\{R_n^*\} = \sqrt{\frac{2}{\pi}}\,\sum_{i=1}^{n}\frac{(n-i)^{1/2}}{i^{1/2}\,n^{1/2}}$$

$$= \sqrt{\frac{n}{2\pi}}\,\sum_{i=1}^{n}\frac{2(n-i)^{1/2}}{n\,i^{1/2}} . \qquad 4.124$$

From Eq. 2.24, the expected value of the adjusted range, given by Solari and Anis, is

$$E\{R_n^*\} = \sqrt{\frac{n}{2\pi}}\,\sum_{i=1}^{n} i^{-1/2}\,(n - i)^{-1/2} . \qquad 4.125$$

To show that the summations in both Eqs. 4.124 and 4.125 are the same, write

$$\sum_{i=1}^{n} \frac{2(n-i)^{1/2}}{n\, i^{1/2}} = \sum_{i=1}^{n} i^{-1/2}\,(n-i)^{1/2} \; .$$

Changing variables $n - i = j$ on the left-hand side, then

$$\sum_{j=1}^{n} \frac{2\, j^{1/2}}{n\,(n-j)^{1/2}} = \sum_{i=1}^{n} i^{-1/2}\,(n-i)^{-1/2} \; .$$

Separating the left-hand summation into two parts and passing one to the right-hand side gives

$$\sum_{i=1}^{n} \frac{i^{1/2}}{n\,(n-i)^{1/2}} = \sum_{i=1}^{n} \frac{1}{i^{1/2}\,(n-i)^{1/2}} - \sum_{i=1}^{n} \frac{i^{-1/2}}{n\,(n-i)^{1/2}} \; ,$$

$$\sum_{i=1}^{n} \frac{i^{1/2}}{n(n-i)^{1/2}} = \sum_{i=1}^{n} \left[\frac{1}{i^{1/2}\,(n-i)^{1/2}} - \frac{i^{1/2}}{n\,(n-i)^{1/2}} \right] \; ,$$

and

$$\sum_{i=1}^{n} \frac{i^{1/2}}{n\,(n-i)^{1/2}} = \sum_{i=1}^{n} \frac{(n-i)^{1/2}}{n\; i^{1/2}} = \sum_{i=1}^{n} \frac{i^{1/2}}{n(n-i)^{1/2}} \; ,$$

which proves that Eqs. 4.124 and 4.125 are identical.

The conclusion of this analysis is that the expected values of adjusted range of exchangeable random variables may be expressed in the same way as the formula for the expected value of unadjusted range. Equation 4.122 is, therefore, valid when input is either independent, or dependent with equal autocorrelation coefficients (exchangeables), while the output is equal to a percentage of the mean inflow, that is, $y_t = \alpha\bar{x}_n$, with α being the level of development.

The above result is relevant in hydrology because when one is interested in overyear storage design, and the assumption of independence of streamflow events is sufficiently accurate and the degree of regulation or development is expressed as a fraction of the sample mean inflow, the expected value of the storage in a given number of years is given by the expected adjusted range which now can be computed exactly by Eq. 4.122. This equation is of mathematical interest as well, because it also gives the expected adjusted range when the original variables have the property of exchangeability.

CHAPTER V

APPROXIMATE EXPECTED VALUES OF RANGE

The exact expected values of range for $n = 1,2$, and 3 are derived in Chapter IV, considering the univariate, bivariate, and trivariate normal distribution functions for the partial sums S_1 , S_2 and S_3 . Based on the exact expected values of range for $n = 1,2$, and 3, the computer simulation or the data generation method is used in this chapter to obtain the approximated equations of the expected values of range for large values of n . In particular, the following cases are studied: the Markov models with periodic autoregression coefficients, the non-stationary exchangeable random variables, and the Markov models with periodic standard deviation.

5.1 Expected Values of Range of Markovian Linear Models with Periodic Autoregression Coefficients

Considering the general model given by Eq. 3.3, it is assumed that $\mu_\tau = 0$ and $\sigma_\tau = \sigma =$ a constant. The Markovian models considered in this section are of the form

$$x_{p,\tau} = \sigma\, z_{p,\tau} = \sigma\, [\sum_{j=1}^{m} \alpha_{j,\tau-j}\, z_{p,\tau-j} + k_{m,\tau}\, \epsilon_{p,\tau}]$$

with $k_{m,\tau}$ given by Eq. 3.5.

V. Yevjevich (1967) gives an approximate equation for the expected values of ranges of linearly dependent normal variables. In particular, he uses the first and second-order Markov models with constant autoregression coefficients and moving average schemes. The same equation was used by P. Sutabutra (1967) for the first-order Markov model.

The same equation is used in this section for approximating the expected value of ranges of Markovian models with periodic autoregression coefficients, or

$$E\{R_n\} = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1}\, [\mathrm{Var}\, S_i]^{1/2} \qquad 5.1$$

The approximation of the above proposed equation is checked in general by the data generation method, for various values of n . For the particular case of $n = 3$ and the first-order Markov model, a comparison is made between the expected values of range given by the exact Eq. 4.112 and by the approximate Eq. 5.1. The results of this comparison are given in Table 5.1. This table shows a high closeness of expected values obtained by both equations where the percentage relative differences are less than 0.09 for all cases of ρ analyzed.

TABLE 5.1 COMPARISON OF THE EXPECTED VALUE OF RANGE FOR $n=3$, GIVEN BY THE EXACT EQ. 4.112 AND THE APPROXIMATED EQ. 5.1, FOR THE FIRST-ORDER MARKOV MODEL.

ρ	Expected range for $n=3$		Difference	Relative
	Exact Equation 4.112 (1)	Approximated Equation 5.1 (2)	(2)-(1)	Error in Percentage
0.0	1.822728	1.822728	0.000000	0.0000
0.1	1.881283	1.881455	0.000172	0.0092
0.2	1.939242	1.939801	0.000559	0.0288
0.3	1.996763	1.997770	0.001007	0.0504
0.4	2.053957	2.055367	0.001410	0.0687
0.5	2.110908	2.112601	0.001693	0.0802
0.6	2.167675	2.169480	0.001805	0.0833
0.7	2.224303	2.226013	0.001710	0.0769
0.8	2.280826	2.282211	0.001385	0.0607
0.9	2.337268	2.338085	0.000817	0.0349

Equation 3.39 gives the general expression of the variance of the partial sum S_i for the m-th order Markov linear model with a periodic standard deviation and periodic autoregression coefficients. In the case of a constant standard deviation, Eq. 3.39 simplifies to

$$\mathrm{Var}\ \{S_i\} = \sigma^2\ [i + 2 \sum_{t=1}^{i-1} \sum_{u=1}^{i-t} \sum_{j=1}^{m} \alpha_{j,t+u-j}\, \rho_{z(x)}\, (u-j,t)] \quad ,$$

where $\alpha_{j,\tau}$ are the periodic autoregression coefficients which may be computed by the solution of a system of m linear equations as given by Eq. 3.7. For the particular cases of the first, second, and third-order Markov models, these coefficients can be computed directly from Eqs. 3.8 to 3.13. The periodic autocorrelation function $\rho_{z(x)}(u\text{-}j,t)$ be computed by using the recursive Eq. 3.14.

Substituting the above equation for Var S_i into Eq. 5.1 the expected value of range of the m-th order Markov model with a constant variance and periodic autoregression coefficients becomes

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}}\, \sigma \sum_{i=1}^{n} i^{-1} \, [i + 2 \sum_{t=1}^{i-1} \sum_{u=1}^{i-t} \sum_{j=1}^{m}$$

$$\alpha_{j,t+u-j}\, \rho_{z(x)}(u - j,t)]^{1/2} \qquad 5.2$$

For the particular case of the constant autoregression coefficients, Eq. 5.2 simplifies to

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1} \, [i + 2 \sum_{u=1}^{i-1}$$

$$(i-u) \sum_{j=1}^{m} \alpha_j\, \rho_{z(x)}(u - j)]^{1/2} \qquad 5.3$$

which is identical to the equation given by V. Yevjevich (1967).

An explicit expression of $E\{R_n\}$ for the case of the first-order Markov model with periodic autocorrelation coefficients may be obtained by using the variance of S_i given in Eq. 3.37, so that Eq. 5.2 becomes

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}}\, \sigma \sum_{i=1}^{n} i^{-1}$$

$$[i + 2 \sum_{t=1}^{i-1} \sum_{u=1}^{i-t} \prod_{k=1}^{u} \rho_{1,t+k-1}]^{1/2} \quad , \qquad 5.4$$

where $\rho_{1,\tau}$ is the first periodic autocorrelation coefficient, which may in general be represented by the harmonic function as given by Eq. 3.6.

In the case of a constant first autocorrelation coefficient, that is, $\rho_{1,\tau} = \rho$, Eq. 5.4 simplifies to

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}}\, \sigma(1-\rho)^{-2} \sum_{i=1}^{n} i^{-1}$$

$$[(1-\rho^2)\, i - 2\rho(1-\rho^i)]^{1/2} \quad , \qquad 5.5$$

which is in agreement with the equation given by P. Sutabutra (1967). It may also be shown that, for the case of $\rho_{1,\tau} = 0$, Eqs. 5.2 through 5.5 simplify to Eq. 2.9 for i.i.d. normal random variables given by Anis and Lloyd (1953).

The validity of the Eqs. 5.2 through 5.5 were tested by the data generation method. The first, second, and third-order Markov models were the only models tested since they are the most commonly used in hydrology. In all cases, 2000 sequences of normal independent random numbers were generated, and the respective Markov dependence was then introduced. The mean ranges for values of n up to 60 were obtained by averaging the computed ranges of 2000 samples.

For the first-order Markov model, the following cases were analyzed:

(a) $\bar{\rho}_{1,\tau} = 0.60$, $s(\rho_{1,\tau}) = 0.00$

(b) $\bar{\rho}_{1,\tau} = 0.60$, $s(\rho_{1,\tau}) = 0.102$

(c) $\bar{\rho}_{1,\tau} = 0.60$, $s(\rho_{1,\tau}) = 0.207$

where $\bar{\rho}_{1,\tau}$ and $s(\rho_{1,\tau})$ represent the mean and standard deviation of the periodic first autocorrelation coefficient, respectively. The results obtained are presented in Figs. 5.1 through 5.5 showing the mean ranges of simulated samples and the values obtained by Eq. 5.4 or Eq. 5.5 for values of n up to 60. In all cases, the agreement between the mean ranges of simulated samples and those computed by Eq. 5.4 or Eq. 5.5 are very good. Figure 5.5 gives a comparison of the cases studied. It shows that after a transition period, which is around one cycle or 12 units, the expected ranges of n increase with the increase of the standard deviation of $\rho_{1,\tau}$.

For the second-order Markov model, the cases analyzed are given in Table 5.2

TABLE 5.2 CASES ANALYZED FOR THE SECOND-ORDER MARKOV MODELS.

Lag k	Mean $\bar{\rho}_{k,\tau}$	Standard Deviation $s(\rho_{k,\tau})$	
		(a)	(b)
1	0.60	0.0	0.102
2	0.45	0.0	0.102

The results for the mean ranges of simulated samples and those obtained from Eq. 5.2 are shown in Figs. 5.6 and 5.7 for values of n up to 60. In both cases, the agreements are very good.

For the third-order Markov model, the cases analyzed are given in Table 5.3.

TABLE 5.3 CASES ANALYZED FOR THE THIRD-ORDER MARKOV MODELS.

Lag k	Mean $\bar{\rho}_{k,\tau}$	Standard Deviation $s(\rho_{k,\tau})$ (a)	Standard Deviation $s(\rho_{k,\tau})$ (b)
1	0.60	0.00	0.102
2	0.45	0.00	0.102
3	0.30	0.00	0.102

Figures 5.8 and 5.9 show the results for the mean ranges of simulated samples and those computed by Eq. 5.2 for values of n up to 60. In both cases the agreement is very good.

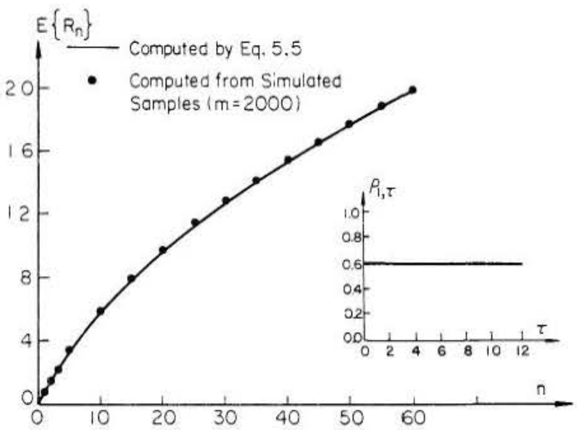

Fig. 5.1 Mean range obtained from simulated samples and the expected values of range computed by Eq. 5.5, for the first-order Markov model with a constant autocorrelation coefficient.

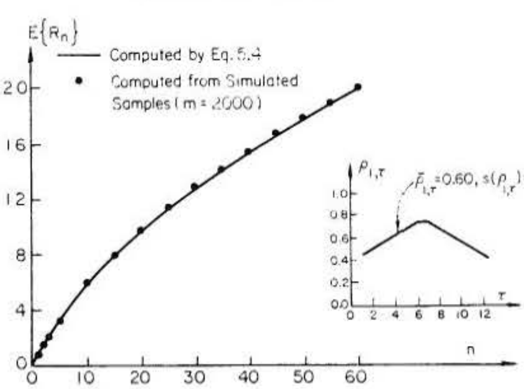

Fig. 5.2 Mean range obtained from simulated samples and the expected values of range computed by Eq. 5.4, for the first-order Markov model with the periodic autocorrelation coefficient.

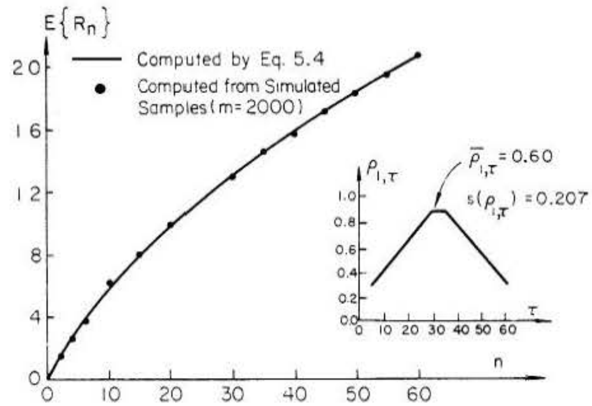

Fig. 5.3 Mean range obtained from simulated samples and the expected values of range computed by Eq. 5.4, for the first-order Markov model with the periodic autocorrelation coefficient.

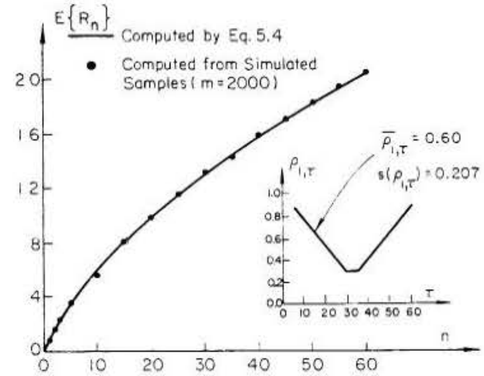

Fig. 5.4 Mean range obtained from simulated samples and the Expected values of range computed by Eq. 5.4, for the first-order Markov model with the periodic autocorrelation coefficient.

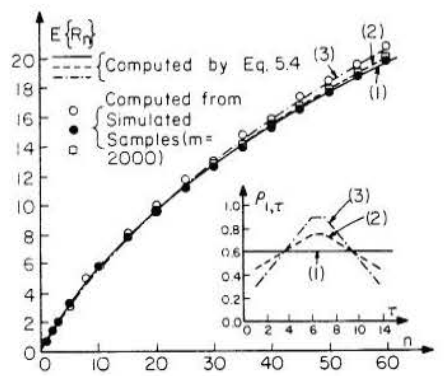

Fig. 5.5 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.4, for first-order Markov models with $\bar{\rho}_{1,\tau}$ = 0.60, and (1) $s(\rho_{1,\tau})$ = 0.0, (2) $s(\rho_{1,\tau})$ = 0.102, and (3) $s(\rho_{1,\tau})$ = 0.207.

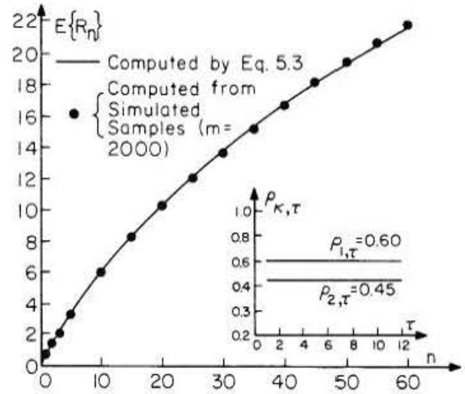

Fig. 5.6 Mean range obtained from simulated samples and the Expected values of range computed by Eq. 5.3, for the second-order Markov model with constant autocorrelation coefficients.

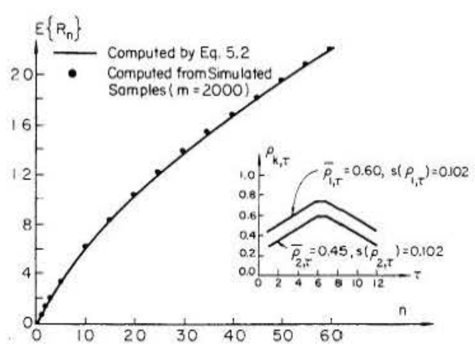

Fig. 5.7 Mean range obtained from simulated samples and the Expected values of range computed by Eq. 5.2, for the second-order Markov model with periodic autocorrelation coefficients.

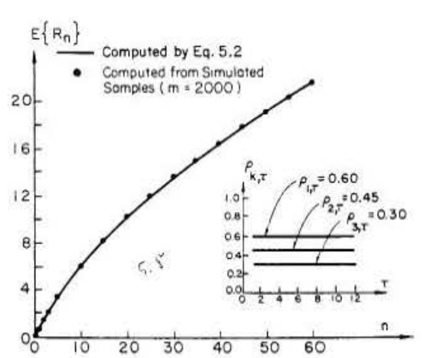

Fig. 5.8 Mean range obtained from simulated samples and the Expected values of range computed by Eq. 5.3, for the third-order Markov model with constant autocorrelation coefficients.

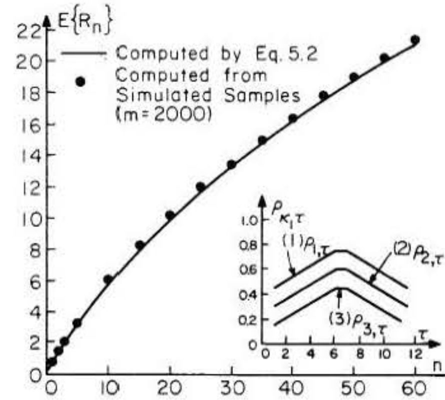

Fig. 5.9 Mean range obtained from simulated samples and the Expected values of range computed by Eq. 5.2, for the third-order Markov model with (1) $\bar{\rho}_{1,\tau} = 0.60$ and $s(\rho_{1,\tau}) = 0.102$, (2) $\bar{\rho}_{2,\tau} = 0.45$ and $s(\rho_{2,\tau}) = 0.102$, and (3) $\bar{\rho}_{3,\tau} = 0.30$ and $s(\rho_{3,\tau}) = 0.102$.

The results obtained above lead to the conclusion that Eq. 5.1 and the derived Eqs. 5.2 through 5.5 are very good approximations of the true expected value of the range for Markov models with periodic autoregression coefficients.

5.2 Expected Values of Range of Non-stationary Exchangeable Random Variables

Non-stationary exchangeable random variables are defined for the purposes of this study as variables which have standard deviation changing with t, but which have equal autocorrelation coefficients. For example, σ_t may be an increasing, a decreasing or a periodic function of t, while the correlation ρ_{ij} between x_i and x_j for $t = i$ and $t = j$ is constant and equal to ρ for any i and j. This kind of variable may be generated by

$$x_t = \sigma_t (\sqrt{\rho}\ \epsilon_o + \sqrt{1-\rho}\ \epsilon_t), 0 \leqslant \rho < 1 \qquad 5.6$$

where ϵ_o and ϵ_t are independent normal variables with mean zero and variance one, both uncorrelated. It follows that $E\{x_t\} = 0$, $Var\{x_t\} = \sigma_t^2$, and $Cov\{x_t, x_{t+u}\} = \sigma_t\ \sigma_{t+u}\rho$. For the particular case of $\rho = 0$, Eq. 5.6 leads to independent variables with changing standard deviations with t.

An approximate equation is proposed in this study for the expected range of the above defined non-stationary exchangeable random variables, as

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} \frac{i^{-1}}{\binom{n}{i}} \sum_{j=1}^{\binom{n}{i}} [Var\{S_i\}_j]^{1/2}, \qquad 5.7$$

where $(S_i)_j$ in this case denotes the j-th sum of size i out of $\binom{n}{i}$ possible sums. In other words, for given values of n and i, there are $\binom{n}{i}$ possible ways in which S_i may be formed. For example, for the case of n = 3, Eq. 5.7 takes the form

$$E\{R_3\} = \sqrt{\frac{2}{\pi}} \{(\mathrm{Var}\, S_1)^{1/2} + \frac{1}{6} [(\mathrm{Var}\, S_2)_1^{1/2} + (\mathrm{Var}\, S_2)_2^{1/2} + (\mathrm{Var}\, S_2)_3^{1/2}] + \frac{1}{3} (\mathrm{Var}\, S_3)^{1/2}\} ,$$

which, in terms of the components of the partial sums, becomes

$$E\{R_3\} = \sqrt{\frac{2}{\pi}} \{(\mathrm{Var}\, x_1)^{1/2} + \frac{1}{6} [(\mathrm{Var}\, \{x_1 + x_2\})^{1/2} + (\mathrm{Var}\, \{x_1 + x_3\})^{1/2} + (\mathrm{Var}\, \{x_2 + x_3\})^{1/2}] + \frac{1}{3} (\mathrm{Var}\, \{x_1 + x_2 + x_3\})^{1/2} \} .$$

For the particular case of i.i.d. random variables, Eq. 5.7 simplifies to

$$E\{R_n\} = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1} [\mathrm{Var}\, S_i]^{1/2}$$

which is in agreement with Spitzer's equation given as Eq. 2.23 in Chapter II.

The degree of approximation by Eq. 5.7 to the exact expected values of range is checked by the data generation method for various values of ρ and n. For the particular case of $\rho = 0$ and n = 3, a comparison is made between the exact expected value of range given by Eq. 4.96 and expected values computed by Eq. 5.7. The results of this comparison are given in Table 5.4 for various combinations of σ_1, σ_2, and σ_3. This table shows that Eq. 5.7 gives a good approximation to the exact expected values of range. The differences relative to the exact values are less than 0.75 percent in all cases analyzed.

The validity of Eq. 5.7 is also tested for increasing, decreasing and periodic functions of the standard deviation σ_t for various values of n. For the first case, σ_t was made increasing from 1 to 12, and for the second case it was made decreasing from 12 to 1. The results of the comparison of the mean ranges obtained from simulated samples and those given by Eq. 5.7 are shown in Figs. 5.10 and 5.11 for values of n up to 12. They are also given in Table 5.5.

For the case of periodic standard deviation σ_τ, several cases were analyzed by using the model of Eq. 5.6. These cases are given in Table 5.6.

For cases shown in Table 5.6, the mean ranges obtained from simulated samples and those computed by Eq. 5.7 are shown in Figs. 5.12,5.13 and 5.14. They are also shown in Tables 5.7, 5.8 and 5.9. These results lead to the conclusion that Eq. 5.7 gives a high degree of approximation to the expected values of range of non-stationary exchangeable random variables.

Figure 5.15 shows a comparison of the expected values of range of i.i.d. random variables (with $\sigma = 10$) and independent variables with periodic standard deviation (with $\bar{\sigma}_\tau = 10$ and $s(\sigma_\tau) = 6.87$). The basic characteristic of this comparison is that the mean ranges of variables with

TABLE 5.4. COMPARISON OF EXACT EXPECTED VALUES OF RANGE FOR n=3, GIVEN BY EQ. 4.96 AND THE APPROXIMATE VALUES COMPUTED BY EQ. 5.7 FOR THE CASE OF INDEPENDENT VARIABLES WITH STANDARD DEVIATIONS VARYING WITH t.

Test No.	Standard Deviations			Expected Range n=3		Difference	Relative Error in Percentage
	σ_1	σ_2	σ_3	Eq. 4.96 (1)	Eq. 5.7 (2)	(2)-(1)	
1	1.0	1.0	1.0	1.822728	1.822728	0.000000	0.000
2	1.0	1.0	10.0	8.705911	8.738561	+0.032650	+0.375
3	1.0	10.0	1.0	8.803861	8.738561	-0.065300	-0.740
4	10.0	1.0	1.0	8.705911	8.738561	+0.032650	+0.375
5	10.0	10.0	1.0	13.937151	13.909359	-0.027792	-0.200
6	10.0	1.0	10.0	13.853776	13.909359	+0.055583	+0.401
7	1.0	10.0	10.0	13.937151	13.909359	-0.027792	-0.200
8	1.0	10.0	100.0	84.199965	84.251436	+0.051471	+0.061
9	1.0	100.0	10.0	84.365130	84.251436	-0.113694	-0.135
10	100.0	10.0	1.0	84.199965	84.251436	+0.051471	+0.061

TABLE 5.5 COMPARISON OF SIMULATED MEAN RANGE AND APPROXIMATED EXPECTED RANGE OF EQ. 5.7 FOR INDEPENDENT RANDOM VARIABLES WITH INCREASING AND DECREASING STANDARD DEVIATION.

n	Mean Range					
	For Increasing σ_τ			For Decreasing σ_τ		
	Simulated m=2000	By Equation 5.7	Difference in %	Simulated m=2000	By Equation 5.7	Difference in %
1	0.775	0.798	2.88	9.296	9.575	2.92
2	2.052	2.089	0.48	15.294	15.670	2.40
3	3.743	3.788	1.19	19.625	20.077	2.25
4	5.858	5.840	0.31	23.100	23.398	1.27
5	8.276	8.207	0.84	25.677	25.931	0.98
6	10.948	10.861	0.80	27.674	27.855	0.65
7	13.976	13.779	1.43	29.158	29.290	0.45
8	17.087	16.944	0.84	30.244	30.327	0.27
9	20.510	20.343	0.82	30.997	31.038	0.13
10	24.403	23.961	1.84	31.496	31.486	0.03
11	28.069	27.791	1.00	31.732	31.729	0.01
12	32.272	31.821	1.42	31.820	31.821	0.003

TABLE 5.6 CASES ANALYZED FOR THE NON-STATIONARY EXCHANGEABLE RANDOM VARIABLES.

Correlation Coefficient	Periodic Standard Deviation σ_τ								
	(a)			(b)			(c)		
ρ	Period ω	$\bar{\sigma}_\tau$	$s(\sigma_\tau)$	Period ω	$\bar{\sigma}_\tau$	$s(\sigma_\tau)$	Period ω	$\bar{\sigma}_\tau$	$s(\sigma_\tau)$
0.0	12	5.0	2.79	12	10.0	6.87	6	5.0	3.28
0.3	12	5.0	2.79	12	10.0	6.87	6	5.0	3.28
0.6	12	5.0	2.79	12	10.0	6.87			
0.9	12	5.0	2.79	12	10.0	6.87			

periodic standard deviation is higher than those with a constant standard deviation. The differences between them increases as n increases.

The plot of the mean range against n for the case of a periodic σ_τ shows that it is an increasing periodic function with the same period as that of σ_t, but with a shift in phase. The maximum amplitude of the mean range is located three units forward with respect to the position of maximum amplitude of the periodic σ_τ. This characteristic is valid only for the particular case analyzed here, that is, with symmetric periodic function σ_τ. For cases of asymmetric or more complex functions σ_τ, the characteristics of the periodic mean range vary accordingly.

The use of Eq. 5.7 in approximating the mean range obtained from simulated samples of non-stationary exchangeable random variables is very good. For large values of n, say $n > 20$, however, the computation takes too much computer time. Therefore, two ways of solving this problem have been developed as described below.

Equation 5.7 requires that, for given values of n and i , the average of the standard deviation of all the possible sums of size i must be computed. Instead of following that route, one can take a random sample of size, say 100, out of all the possible sums of size i and then take the average over the sample size. This can be done easily in a digital computer. For practical use of this procedure, a compromise should be made between the accuracy of results and the amount of computer time required, both of which depend on the size of the sample considered. Figure 5.16 shows an example of application of this procedure for the case of independent random variables with $\overline{\sigma}_\tau = 5.00$ and $s(\sigma_\tau) = 2.79$. The number of sums, as the sample size, in this case was selected as m = 50.

TABLE 5.7 COMPARISON OF SIMULATED MEAN RANGE AND APPROXIMATED EXPECTED RANGE OF EQ. 5.7 FOR NON-STATIONARY EXCHANGEABLE RANDOM VARIABLES. CASE OF $\bar{\sigma}_\tau = 5.0$ AND $s(\sigma_\tau) = 2.79$.

n	Correlation Coefficient											
	ρ=0.0			ρ=0.30			ρ=0.60			ρ=0.90		
	Simulated m=2000	By Equation 5.7	Difference in %	Simulated m=2000	By Equation 5.7	Difference in %	Simulated m=2000	By Equation 5.7	Difference in %	Simulated m=2000	By Equation 5.7	Difference in %
1	1.530	1.596	4.13	1.579	1.596	1.06	1.584	1.596	0.75	1.594	1.5946	0.12
2	2.991	3.072	2.64	3.227	3.247	0.62	3.408	3.403	0.15	3.564	3.545	0.53
3	4.835	4.923	1.79	5.474	5.433	0.75	5.927	5.871	0.95	6.302	6.261	0.65
4	7.489	7.446	0.58	8.518	8.492	0.31	9.403	9.366	0.39	10.156	10.134	0.22
5	10.897	10.873	0.22	12.625	12.719	0.74	14.207	14.231	0.17	15.586	15.550	0.23
6	15.733	15.725	0.05	18.587	18.737	0.80	21.144	21.173	0.14	23.327	23.285	0.18
7	19.942	19.886	0.28	24.288	24.378	0.37	27.945	27.930	0.05	31.082	30.978	0.33
8	22.286	22.053	1.06	27.973	27.932	0.15	32.557	32.468	0.27	36.461	36.321	0.38
9	23.533	23.256	1.19	30.462	30.298	0.54	35.762	35.626	0.38	40.284	40.119	0.41
10	24.236	23.919	1.32	31.946	31.867	0.25	37.858	37.794	0.17	42.893	42.768	0.29
11	24.614	24.302	1.28	33.005	32.944	0.18	39.383	39.323	0.15	44.777	44.655	0.27
12	24.878	24.568	1.26	33.824	33.788	0.11	40.594	40.537	0.14	46.281	46.164	0.25
13	25.107	24.827	1.13	34.577	34.632	0.16	41.745	41.752	0.01	47.766	47.672	0.20
14	25.411	25.182	0.91	35.676	35.709	0.09	43.317	43.282	0.08	49.693	49.560	0.27
15	25.892	25.773	0.46	37.286	37.273	0.03	45.553	45.453	0.22	52.385	52.211	0.33

TABLE 5.8 COMPARISON OF SIMULATED MEAN RANGE AND APPROXIMATED EXPECTED RANGE OF EQ. 5.7 FOR NON-STATIONARY EXCHANGEABLE RANDOM VARIABLES. CASE OF $\bar{\sigma} = 10.0$ AND $s(\sigma_\tau) = 6.87$.

n	Correlation Coefficient											
	ρ=0.0			ρ=0.3			ρ=0.6			ρ=0.9		
	Simulated m=1000	By Equation 5.7	Difference in %	Simulated m=1000	By Equation 5.7	Difference in %	Simulated m=1000	By Equation 5.7	Difference in %	Simulated m=1000	By Equation 5.7	Difference in %
1	1.530	1.596	4.13	1.578	1.596	1.13	1.589	1.596	0.44	1.600	1.596	0.25
2	3.732	3.802	1.84	3.930	3.998	1.70	4.136	4.175	0.93	4.331	4.337	0.14
3	7.226	7.282	0.77	7.931	7.933	0.02	8.526	8.502	0.28	9.069	9.015	0.60
4	12.997	12.852	1.13	14.275	14.359	0.58	15.696	15.648	0.31	16.946	16.796	0.89
5	22.296	22.216	0.36	25.195	25.224	0.11	27.861	27.767	0.34	30.287	30.017	0.90
6	33.108	32.900	0.63	38.253	38.298	0.12	42.926	42.758	0.39	47.029	46.663	0.78
7	42.298	42.004	0.70	50.544	50.496	0.09	57.498	57.317	0.31	63.722	63.211	0.81
8	48.251	47.700	1.15	59.437	59.341	0.16	68.636	68.438	0.29	76.741	76.208	0.70
9	50.735	50.050	1.37	63.980	64.019	0.06	74.811	74.724	0.12	84.350	83.796	0.66
10	51.862	51.088	1.51	66.794	66.635	0.24	78.742	78.402	0.43	89.034	88.328	0.80
11	52.321	51.545	1.50	68.245	68.085	0.23	80.847	80.510	0.42	91.686	90.963	0.79
12	52.524	51.748	1.50	69.048	68.883	0.24	82.018	81.699	0.39	93.168	92.465	0.76
13	52.687	51.945	1.43	69.812	69.680	0.19	83.181	82.888	0.35	94.679	93.968	0.76
14	53.056	52.371	1.31	71.278	71.128	0.21	85.303	84.998	0.36	97.352	96.604	0.77
15	53.810	53.312	0.93	73.846	73.743	0.14	88.955	88.684	0.30	101.864	101.138	0.72

TABLE 5.9 COMPARISON OF SIMULATED MEAN RANGE AND APPROXIMATED EXPECTED RANGE OF EQ. 5.7 FOR NON-STATIONARY EXCHANGEABLE RANDOM VARIABLES. CASE OF $\bar{\sigma}_\tau = 5.0$ AND $s(\sigma_\tau) = 3.28$.

n	Correlation Coefficient					
	ρ=0.00			ρ=0.30		
	Simulated m=1000	By Equation 5.7	Difference in %	Simulated m=1000	By Equation 5.7	Difference in %
1	0.788	0.798	1.25	0.779	0.798	2.38
2	4.369	4.428	1.33	4.565	4.542	0.51
3	9.900	9.992	0.92	10.593	10.656	0.59
4	14.576	14.415	1.12	16.244	16.077	1.04
5	16.407	16.076	2.06	18.906	18.627	1.50
6	16.610	16.253	2.20	19.307	19.043	1.39
7	16.799	16.417	2.33	19.713	19.455	1.33
8	18.291	17.894	2.22	22.181	21.937	1.11
9	21.260	21.226	0.16	26.696	26.802	0.39
10	24.060	24.255	0.80	31.424	31.520	0.30
11	25.237	25.377	0.55	33.605	33.856	0.74
12	25.405	25.480	0.29	33.997	34.257	0.76
13	25.526	25.580	0.21	34.388	34.657	0.78
14	26.563	26.642	0.30	36.635	36.981	0.93
15	28.852	29.259	1.39	41.252	41.531	0.66
16	31.070	31.722	2.05	45.606	46.014	0.89
17	32.101	32.619	1.59	47.885	48.284	0.83
18	32.200	----	--	48.239	----	--

In using the procedure just outlined, Eq. 5.7 takes the form

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} \frac{i^{-1}}{m} \sum_{j=1}^{m} [\text{Var}\{S_i\}_j]^{1/2} \quad , \qquad 5.8$$

where m denotes the sample size of the sums computed, and the subscript j denotes a particular realization of the sum of size i , taken at random.

Another procedure has been developed in this study for obtaining the approximate mean range of independent variables with standard deviation varying with t . This procedure is based on the exact expected range of i.i.d. random variables and an equivalent standard deviation $\hat{\sigma}_n$ of the n variables considered.

The proposed equation is

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}}\ \hat{\sigma}_n \sum_{i=1}^{n} i^{-1/2} \qquad 5.9$$

with $\hat{\sigma}_n$ defined by

$$\hat{\sigma}_n = \sqrt{\frac{1}{n} \sum_{\tau=1}^{n} \sigma_\tau^2} \quad . \qquad 5.10$$

The idea behind this procedure is that by multiplying the function $\hat{\sigma}_n$, as given by Eq. 5.10, by the exact mean range of i.i.d. random variables, the effect of the changing standard deviation may be accounted for.

In the particular case of a periodic standard deviation σ_τ , with $\tau = 1,2,...,\omega$, with ω ,the main cycle (for example, one year) and considering p the number of cycles (for example, the number of years), then Eqs. 5.9 and 5.10 are combined as

$$E\{R_n\} \doteq \sqrt{\frac{2}{\pi}} \sqrt{\frac{1}{\omega} \sum_{\tau=1}^{\omega} \sigma_\tau^2} \sum_{i=1}^{n} i^{-1/2} \qquad 5.11$$

which is valid only for values of $n = p\omega$, say for n = 12, 24, 36,..., 12p , with p an integer, and ω equal to 12 months. Notice that, for the particular case of i.i.d. random variables with $\sigma_\tau = \sigma$, the above equations simplify to Eq. 2.23.

The validity of this procedure for obtaining the approximate mean range of independent random variables with standard deviations varying with t was tested by comparing the mean ranges obtained directly by simulation with those computed by Eq. 5.9. The first two tests considered the cases of standard deviations increasing and decreasing with t . For this, 250 sequences of random numbers, each of size 600, were generated by increasing or decreasing (according to the case) their standard deviation every 50 generated numbers. These standard deviations varied from 1 to 12 and from 12 to 1 for the increasing and decreasing cases, respectively. The results of these tests are shown in Fig. 5.17 for values of n up to 600 .

Two cases of periodic standard deviations with cycles of 12 months were also tested. The results of these tests are shown in Fig. 5.18 for the mean ranges of n up to 600. For all cases analyzed, the agreement between the mean ranges obtained by simulation and those computed by Eq. 5.9 are very good for both small and large values of n . It is interesting to observe in Fig. 5.18 that the increasing periodic mean range may be reproduced by considering the equivalent periodic function $\hat{\sigma}_n$, as given by Eq. 5.10.

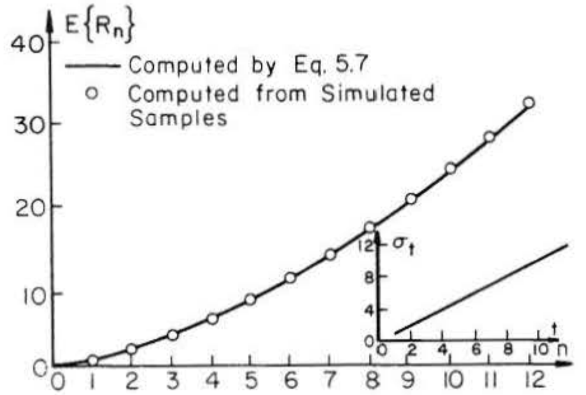

Fig. 5.10 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.7, for independent random variables with standard deviation increasing with t.

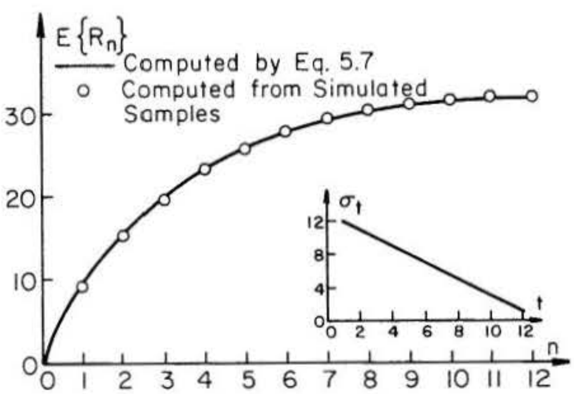

Fig. 5.11 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.7, for independent random variables with standard deviation decreasing with t .

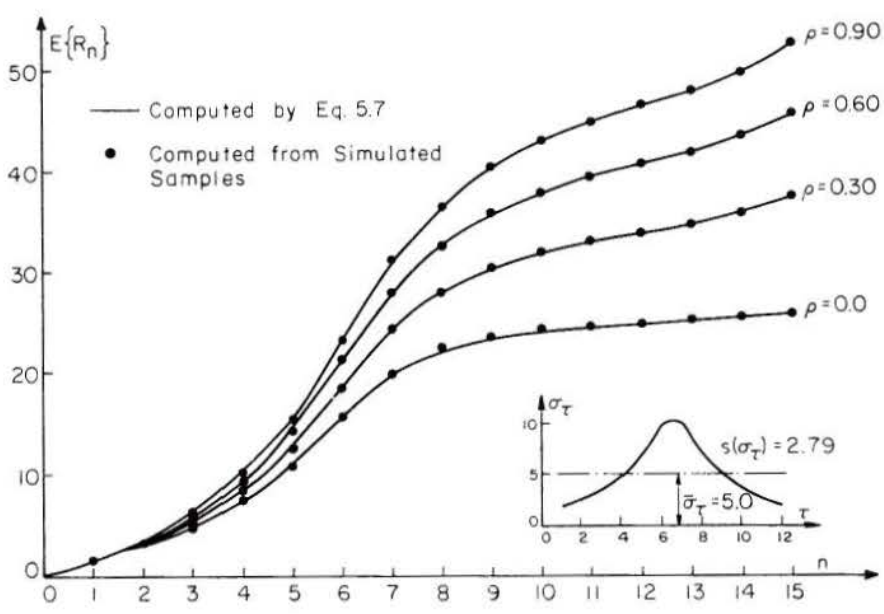

Fig. 5.12 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.7 for non-stationary exchangeable random variables of Eq. 5.6.

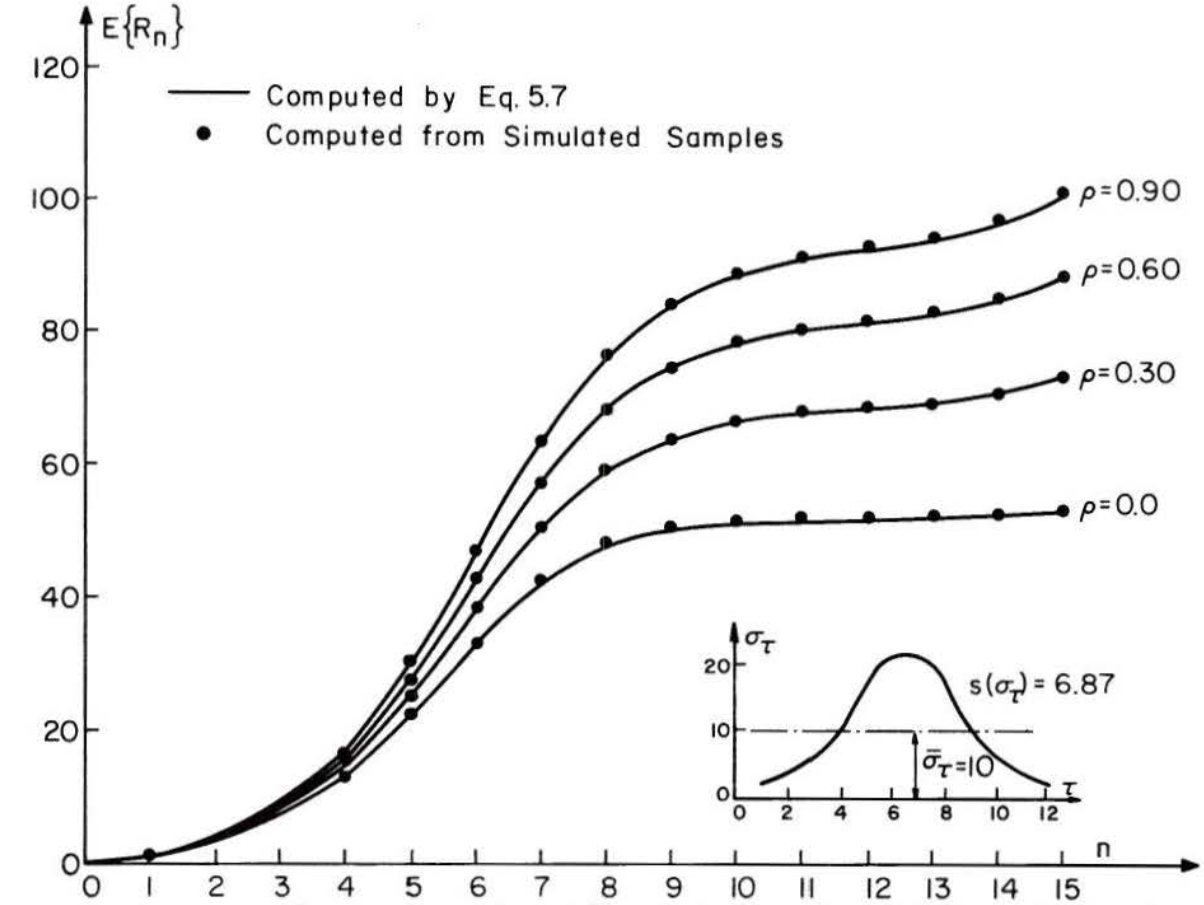

Fig. 5.13 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.7, for non-stationary exchangeable random variables of Eq. 5.6.

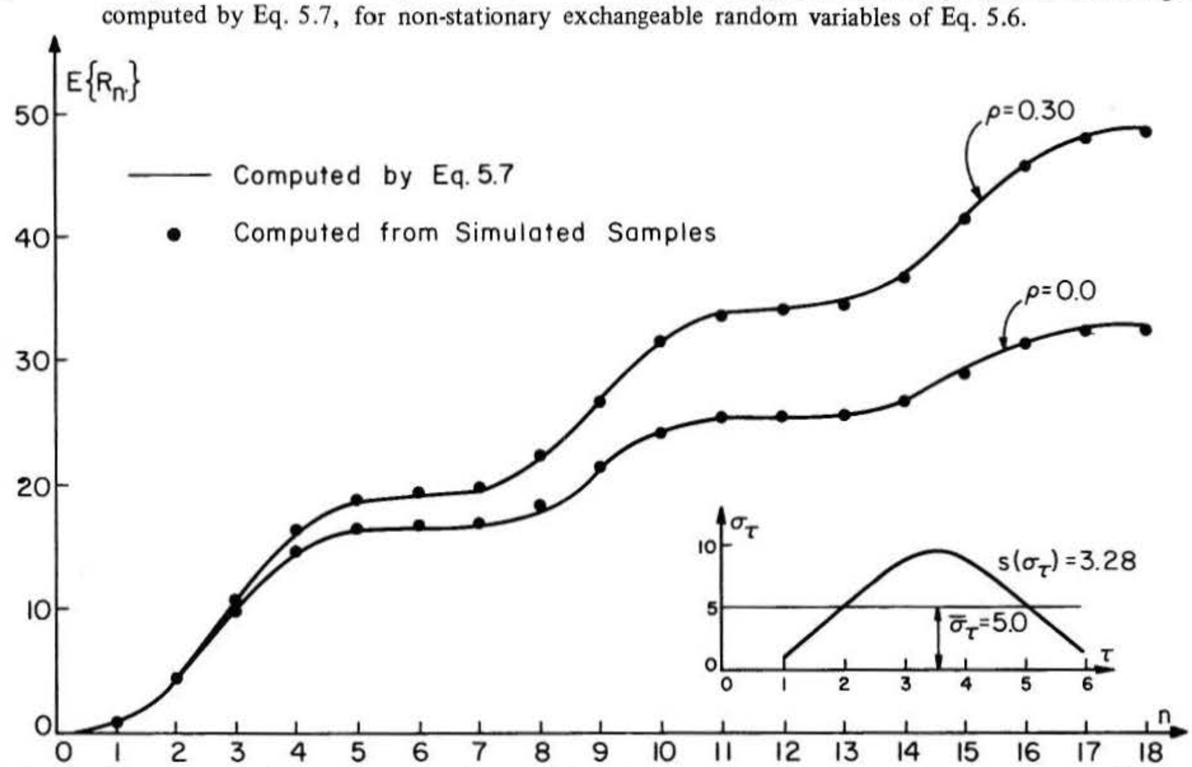

Fig. 5.14 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.7, for non-stationary exchangeable random variables of Eq. 5.6.

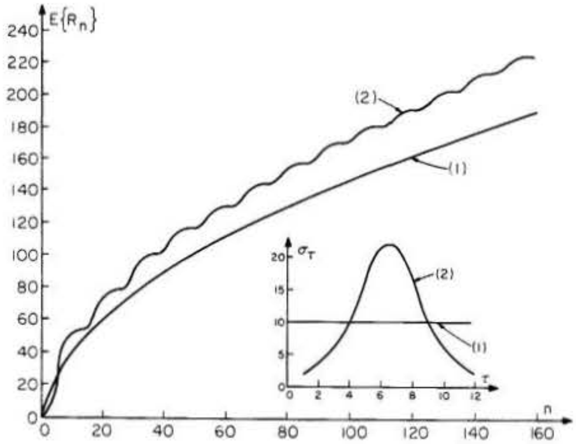

Fig. 5.15 Comparison of the Expected values of range for (1) i.i.d. variables with $\sigma = 10$, and (2) variables with periodic standard deviation with $\overline{\sigma}_\tau = 10$ and $s(\sigma_\tau) = 6.87$.

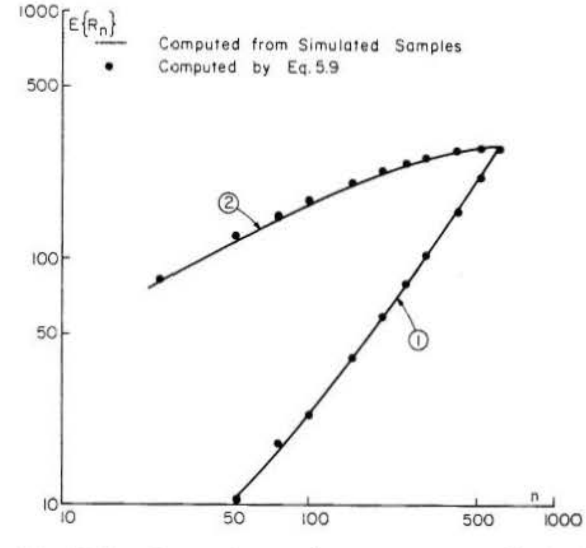

Fig. 5.17 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.9, for independent variables with standard deviation increasing with t (1), and decreasing with t (2) .

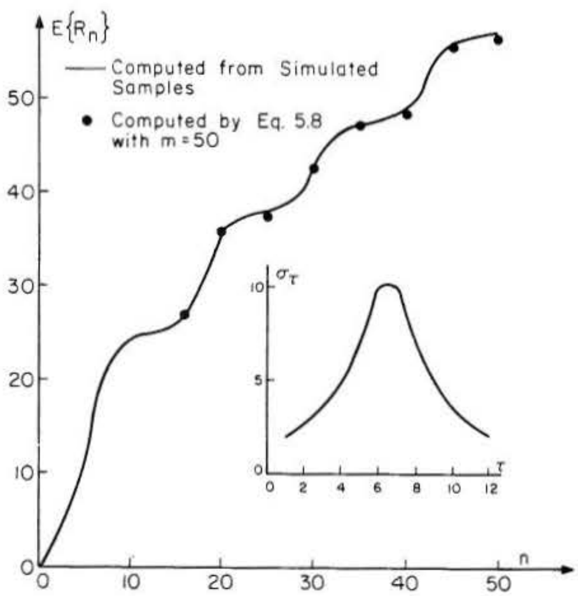

Fig. 5.16 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.8, for independent variables with periodic standard deviation, with $\overline{\sigma}_\tau = 5$ and $s(\sigma_\tau) = 2.79$.

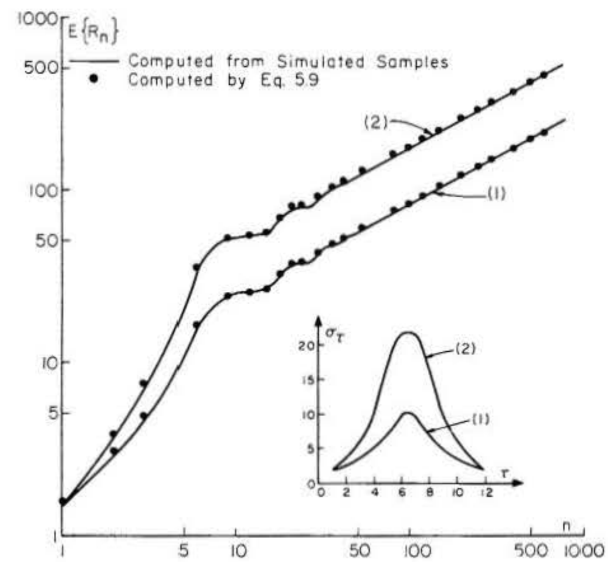

Fig. 5.18 Comparison of mean ranges obtained from simulated samples and the Expected values of range computed by Eq. 5.9, for two cases of independent variables with periodic standard deviation. (1) $\overline{\sigma}_\tau = 5$ and $s(\sigma_\tau) = 2.79$, and (2) $\overline{\sigma}_\tau = 10$ and $s(\sigma_\tau) = 6.87$.

5.3 Expected Values of Range of Markov Dependent Random Variables With Periodic Standard Deviation

The use of Eqs. 5.7 and 5.9 for approximating the expected values of range of Markov dependent random variables with a periodic standard deviation did not give satisfactory results. Another procedure was developed for the particular case of Markov models with the constant autoregression coefficients. Let us first discuss some characteristics related to the expected values of range of this kind of models.

Figure 5.19 shows the plot of mean ranges obtained from simulated samples of the first-order Markov model with a periodic standard deviation for n up to 60. These mean ranges are increasing periodic functions, with the same period as that of σ_τ and maximum amplitudes which are three units out of phase with respect to σ_τ. This last characteristic refers to the particular case of σ_τ considered. Figure 5.19 shows the mean ranges for the case of $\bar{\sigma}_\tau = 5.0$, $s(\sigma_\tau) = 2.79$, and ρ values of 0.0, 0.3, 0.6, and 0.9. It also shows the mean range for the case of a constant $\sigma = 5$. As in the case of stationary Markov models, the mean range for a particular n increases as ρ increases, for Markov models with periodic standard deviation.

The expected values of range of Markov models with a periodic standard deviation are expressed as

$$E\{R_n\} = f(\bar{\sigma}_\tau, s(\sigma_\tau), \rho) \quad . \qquad 5.12$$

where $\bar{\sigma}_\tau$ and $s(\sigma_\tau)$ denote the mean and standard deviation of the periodic standard deviation and ρ is the first autocorrelation coefficient which defines the dependence. With the above notation, four functions are defined as follows,

$$f_1 = f_1(1,0,0) = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1/2} \quad , \qquad 5.13$$

$$f_2 = f_2(1,0,\rho) \doteq \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1} \; [\mathrm{Var}\, S_i]^{1/2} \quad , \qquad 5.14$$

$$f_3 = f_3(\bar{\sigma}_\tau, s(\sigma_\tau), 0) \doteq \sqrt{\frac{2}{\pi}}\, \hat{\sigma}_n \sum_{i=1}^{n} i^{-1/2} \quad , \qquad 5.15$$

and

$$f_4 = f_4(\bar{\sigma}_\tau, s(\sigma_\tau), \rho) \quad . \qquad 5.16$$

That is, f_1 denotes the expected values of range of i.i.d. random variables with variance unity and is exactly that given by Eq. 2.23; f_2 denotes the expected values of range of Markov models with variance unity and the first autocorrelation coefficient ρ, which, as described in section 5.1, may be approximated by Eq. 5.5; f_3 denotes the expected values of range of independent variables with a periodic standard deviation, which, as described in section 5.2, may be approximated by Eqs. 5.7, 5.8, or 5.9, (in Eq. 5.15, f_3 is approximated by Eq. 5.9); finally, f_4 denotes the expected values of range of the Markov model with a periodic standard deviation.

The basic hypothesis in approximating the expected values of range of Markov models with periodic standard deviation, denoted by f_4, may be expressed mathematically as

$$f_2(1,0,\rho) - f_1(1,0,0) \doteq \frac{1}{\bar{\sigma}_\tau} [f_4(\bar{\sigma}_\tau, s(\sigma_\tau), \rho) - f_3(\bar{\sigma}_\tau, s(\sigma_\tau), 0)] \qquad 5.17$$

which is also shown schematically in Fig. 5.20.

The idea behind the above hypothesis is that the effects of dependence due to ρ and non-stationarity due to a periodic σ_τ may be separated. In other words, one can go from the function $f_1(1,0,0)$ to $f_3(\bar{\sigma}_\tau, s(\sigma_\tau), 0)$ by using the procedures developed in the previous section 5.2. Then the function $f_4(\bar{\sigma}_\tau, s(\sigma_\tau), \rho)$ will be obtained by superimposing the effect of ρ as in the stationary case.

The validity of the above hypothesis of Eq. 5.17 was tested by computer simulation for $\rho = 0.60$ and for two cases of periodic σ: $\bar{\sigma}_\tau = 5.0$, $s(\sigma_\tau) = 2.79$, and $\bar{\sigma}_\tau = 10.0$, $s(\sigma_\tau) = 6.87$. The effect of $\rho = 0.60$ for the stationary and non-stationary cases, as expressed by Eq. 5.17, are shown for the above two cases in Fig. 5.21 and Tables 5.10 and 5.11 for n up to 600. The results obtained are very good, especially for n greater than 10.

Based on the hypothesis expressed by Eq. 5.17, the proposed approximation to the expected values of range of Markov models with periodic standard deviation is

$$E\{R_n\} \doteq \frac{2}{\pi} \{\hat{\sigma}_n \sum_{i=1}^{n} i^{-1/2} + \bar{\sigma}_\tau \left[\sum_{i=1}^{n} i^{-1} (\mathrm{Var}\, S_i)^{1/2} - \sum_{i=1}^{n} i^{-1/2}\right]\} \qquad 5.18$$

where $\hat{\sigma}_n$ is given by Eq. 5.10 and Var S_i by Eq. 3.38. It should be noted that the function $f_3(\bar{\sigma}_\tau, s(\sigma_\tau), 0)$ was approximated in Eq. 5.15 by Eq. 5.9. However, better accuracy is obtained if f_3 is approximated by Eq. 5.7 or Eq. 5.8.

Equation 5.18 was used for computing the approximated mean ranges of the two cases of Markov models: (a) $\rho = 0.60$, $\bar{\sigma}_\tau = 5.0$, and $s(\sigma_\tau) = 2.79$, and (b) $\rho = 0.60$, $\bar{\sigma}_\tau = 10.0$, and $s(\sigma_\tau) = 6.87$. These mean ranges were compared with those directly obtained by simulation, and the agreement between them is very good, as shown in Fig 5.22 and Tables 5.12 and 5.13.

A hypothesis similar to that expressed by Eq. 5.17 may be extended to cases of higher order Markov models or even to Markov models with periodic autoregression coefficients. In such cases, the equations developed in section 5.1 should be useful.

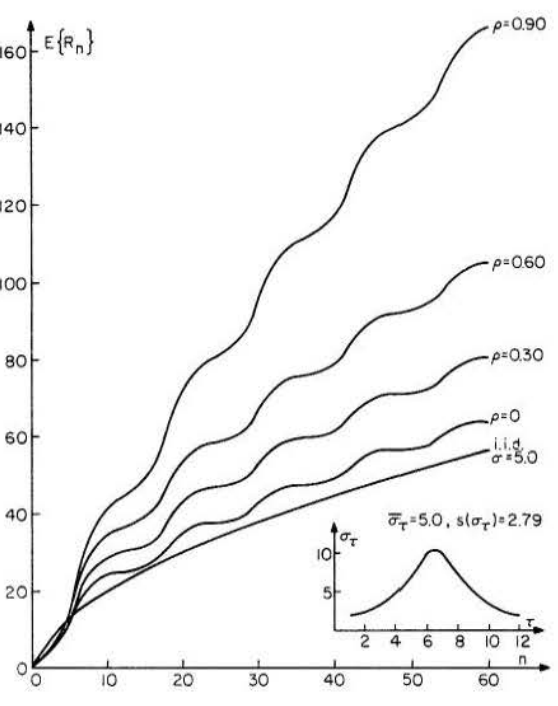

Fig. 5.19 Mean ranges obtained from simulated samples for the Markov model $x_{p,\tau} = \sigma_\tau(\rho x_{p,\tau-1} + \sqrt{1-\rho^2}\ \epsilon_{p,\tau})$ with periodic standard deviation σ_τ and constant first autocorrelation coefficient ρ.

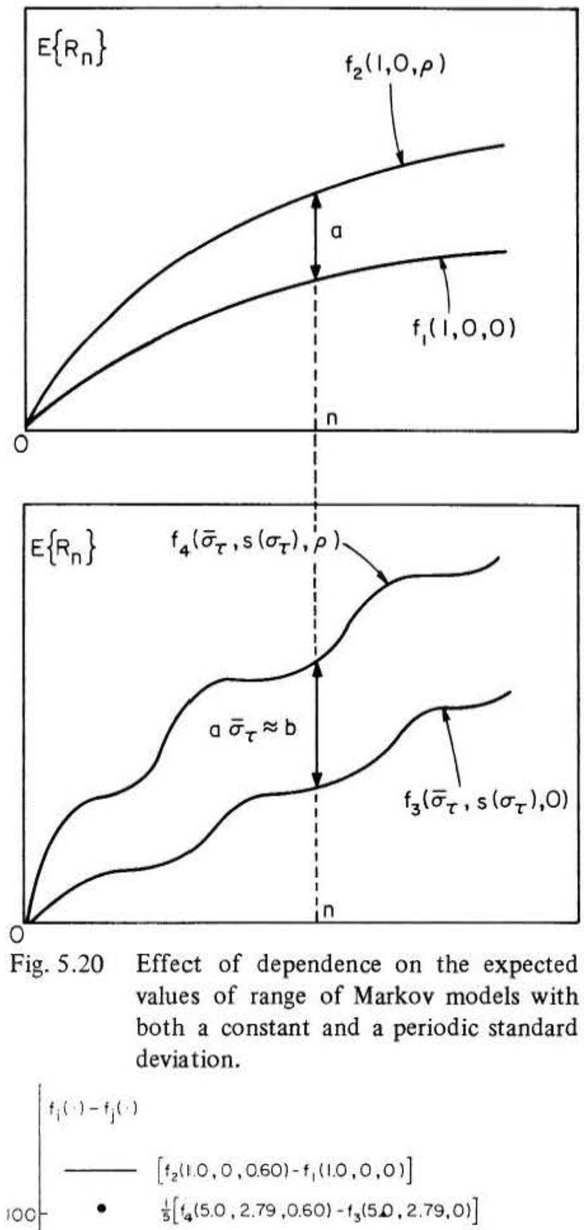

Fig. 5.20 Effect of dependence on the expected values of range of Markov models with both a constant and a periodic standard deviation.

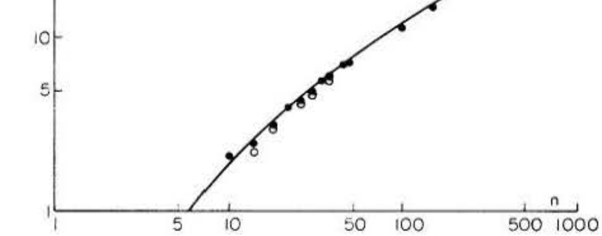

Fig. 5.21 Comparison of the effect of dependence on the mean range, for two cases of Markov models with both a constant and a periodic standard deviation.

TABLE 5.10 COMPARISON OF THE EFFECT OF DEPENDENCE ON THE MEAN RANGE, FOR MARKOV MODELS WITH CONSTANT AND PERIODIC STANDARD DEVIATION. CASE OF $\bar{\sigma}_\tau = 5$, $s(\sigma_\tau) = 2.79$ AND $\rho = 0.60$.

n	Mean Range By Simulation		Difference	Standardized Difference	Difference
	$f_4 = f_4(\bar{\sigma}_\tau, s(\sigma_\tau), \rho)$	$f_3 = f_3(\bar{\sigma}_\tau, s(\sigma_\tau), 0)$	$f_4 - f_3$	$\frac{1}{\bar{\sigma}_\tau}(f_4 - f_3)$	$f_2(1,0,\rho) - f_1(1,0,0)$
6	20.157	15.733	4.424	0.885	1.002
10	34.682	24.236	10.447	2.089	1.843
14	37.734	25.411	12.323	2.464	2.606
18	46.928	31.225	15.703	3.140	3.300
22	57.241	37.321	19.920	3.984	3.939
26	59.681	38.267	21.414	4.282	4.534
30	66.757	42.347	24.410	4.882	5.092
34	75.196	47.091	28.105	5.621	5.619
38	77.224	47.905	29.319	5.864	6.120
42	83.736	51.967	31.769	6.354	6.598
46	91.313	56.430	34.883	6.976	7.056
50	92.958	57.062	35.896	7.179	7.497
100	138.197	80.926	57.271	11.454	12.031
150	177.602	102.776	74.826	14.965	15.556
200	213.196	121.759	91.437	18.287	18.541
250	240.101	135.943	104.158	20.832	21.180
300	264.541	148.922	115.619	23.124	23.570
350	288.376	162.417	125.959	25.192	25.770
400	313.256	175.542	137.714	27.543	27.819
450	336.302	188.450	147.852	29.570	29.746
500	356.882	198.909	157.973	31.595	31.569
550	374.815	209.233	165.582	33.116	33.303
600	392.443	218.893	173.550	34.710	34.961

TABLE 5.11 COMPARISON OF THE EFFECT OF DEPENDENCE ON THE MEAN RANGE, FOR MARKOV MODELS WITH CONSTANT AND PERIODIC STANDARD DEVIATION. CASE OF $\bar{\sigma}_\tau = 10$, $s(\sigma_\tau) = 6.87$ AND $\rho = 0.60$.

n	Simulated Mean Range		Difference	Standardized Difference	Difference
	$f_4 = f_4(\bar{\sigma}_\tau, s(\sigma_\tau), \rho)$	$f_3 = f_3(\bar{\sigma}_\tau, s(\sigma_\tau), 0)$	$f_4 - f_3$	$\frac{1}{\bar{\sigma}_\tau}(f_4 - f_3)$	$f_2(1,0,\rho) - f_1(1,0,0)$
6	40.853	33.110	7.743	0.774	1.002
10	71.683	51.882	19.801	1.980	1.843
14	74.896	53.081	21.815	2.181	2.606
18	95.917	66.096	29.821	2.982	3.300
22	118.680	79.719	38.961	3.896	3.939
26	121.424	80.660	40.764	4.076	4.534
30	136.366	89.946	46.420	4.642	5.092
34	154.982	100.422	54.560	5.456	5.619
38	157.340	101.230	56.110	5.611	6.120
42	171.839	110.273	61.566	6.157	6.598
46	189.909	120.212	69.696	6.970	7.056
50	191.865	120.834	71.031	7.103	7.497
100	285.500	171.580	113.920	11.392	12.031
150	367.300	219.120	148.180	14.818	15.556
200	443.160	259.870	183.290	18.329	18.541
250	498.820	290.220	208.600	20.860	21.180
300	549.670	317.660	232.010	23.201	23.570
350	599.700	346.750	252.950	25.295	25.770
400	649.430	373.560	275.870	27.587	27.819
450	695.200	400.650	294.550	29.455	29.746
500	738.020	422.580	315.440	31.544	31.569
550	775.890	445.580	330.310	33.031	33.303
600	812.880	466.680	346.200	34.620	34.961

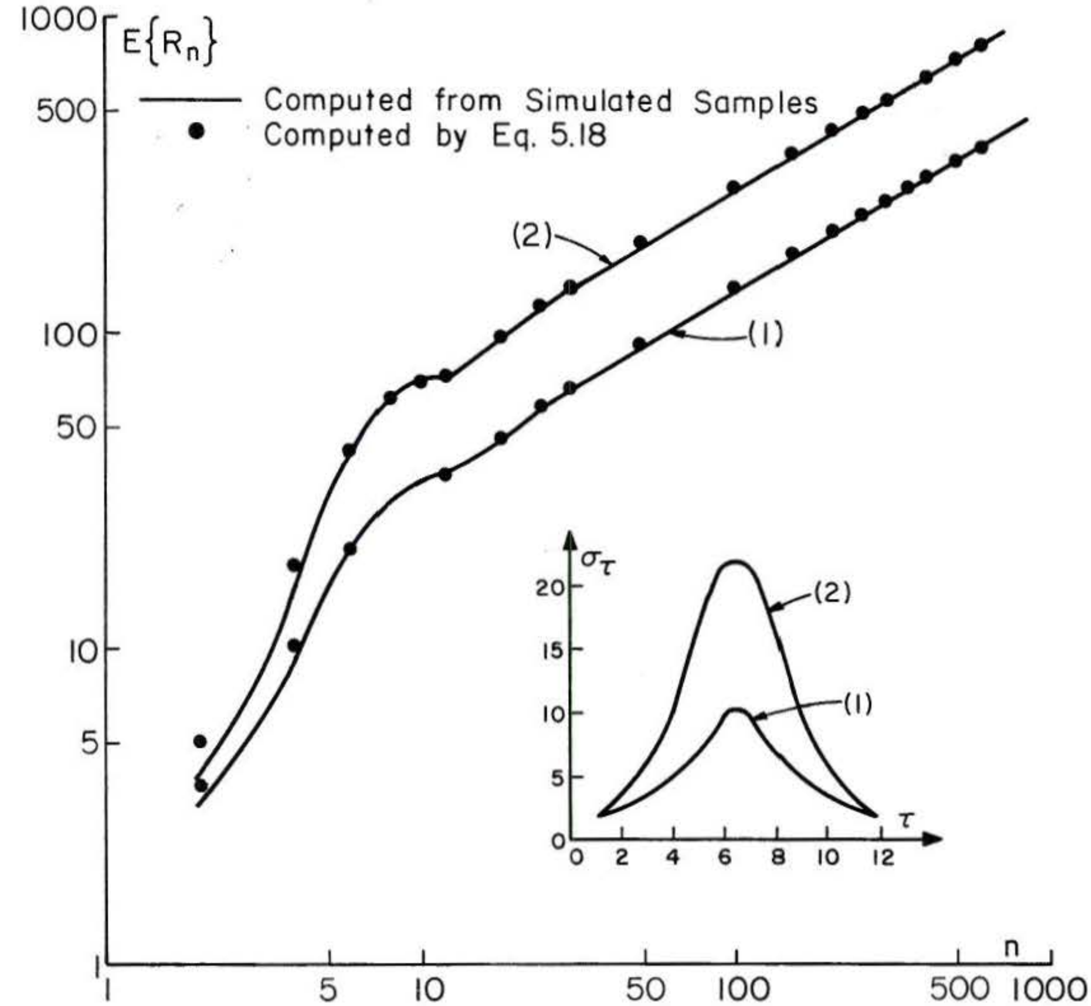

Fig. 5.22 Comparison of mean ranges obtained from simulated samples and the expected values of range computed by Eq. 5.18, for two cases of Markov models with ρ = 0.60 and with periodic standard deviation. (1) $\overline{\sigma}_\tau = 5$ and $s(\sigma_\tau) = 2.79$, and (2) $\overline{\sigma}_\tau = 10$ and $s(\sigma_\tau) = 6.87$.

TABLE 5.12 COMPARISON BETWEEN THE MEAN RANGES OBTAINED BY SIMULATION AND THOSE COMPUTED BY EQ. 5.18, FOR MARKOV MODELS WITH PERIODIC STANDARD DEVIATION. CASE OF $\overline{\sigma}_\tau = 5$, $s(\sigma_\tau) = 2.79$ AND $\rho = 0.60$.

n	f_2-f_1	$\overline{\sigma}_\tau(f_2-f_1)$	$f_3(\overline{\sigma}_\tau, s(\sigma_\tau), 0)$	Computed By Equation 5.18 $f_4(\overline{\sigma}_\tau, s(\sigma_\tau), \rho)$	Simulated $f_4(\overline{\sigma}_\tau, s(\sigma_\tau), \rho)$
2	0.149	0.747	2.991	3.738	3.320
4	0.562	2.809	7.489	10.298	9.013
6	1.002	5.011	15.733	20.744	20.157
8	1.432	7.161	22.286	29.447	30.477
10	1.843	9.215	24.236	33.451	34.682
12	2.234	11.169	24.878	36.047	36.392
18	3.300	16.500	31.225	47.725	46.928
24	4.242	21.207	37.857	59.064	58.564
30	5.092	25.460	42.347	67.807	66.757
50	7.497	37.484	57.062	94.546	92.958
100	12.031	60.155	80.926	141.081	138.197
150	15.556	77.780	102.776	180.556	177.602
200	18.541	92.705	121.759	214.464	213.196
250	21.180	105.900	135.943	241.843	240.101
300	23.570	117.850	148.922	266.772	264.541
350	25.770	128.850	162.417	291.267	288.376
400	27.819	139.095	175.542	314.637	313.256
450	29.746	148.730	188.450	337.180	336.302
500	31.569	157.845	198.909	356.754	356.882
550	33.303	166.515	209.233	375.748	374.815
600	34.961	174.805	218.893	393.698	392.443

TABLE 5.13 COMPARISON BETWEEN THE MEAN RANGES OBTAINED BY SIMULATION AND THOSE COMPUTED BY EQ. 5.18, FOR MARKOV MODELS WITH PERIODIC STANDARD DEVIATION. CASE OF $\sigma_\tau = 10$, $s(\sigma_\tau) = 6.87$ AND $\rho = 0.60$.

n	f_2-f_1	$\sigma_\tau(f_2-f_1)$	$f_3(\sigma_\tau, s(\sigma_\tau), 0)$	Computed By Equation 5.18 $f_4(\sigma_\tau, s(\sigma_\tau), \rho)$	Simulated $f_4(\overline{\sigma}_\tau, s(\sigma_\tau), \rho)$
2	0.149	1.494	3.732	5.226	4.015
4	0.562	5.617	13.007	18.624	15.372
6	1.002	10.023	33.110	43.133	40.853
8	1.432	14.322	48.281	62.603	63.997
10	1.843	18.430	51.882	70.312	71.683
12	2.234	22.337	52.546	74.883	73.463
18	3.300	33.000	66.096	99.095	95.917
24	4.242	42.416	80.273	122.689	120.250
30	5.092	50.920	89.946	140.866	136.366
50	7.497	74.970	120.834	195.804	191.865
100	12.031	120.310	171.582	291.892	285.498
150	15.556	155.560	219.123	374.683	367.304
200	18.541	185.410	259.867	445.277	443.163
250	21.180	211.800	290.217	502.017	498.825
300	23.570	235.700	317.659	553.359	549.667
350	25.770	257.700	346.755	604.455	599.701
400	27.819	278.190	373.558	651.748	649.433
450	29.746	297.460	400.655	698.115	695.196
500	31.569	315.690	422.580	738.270	738.017
550	33.303	333.030	445.585	778.615	775.890
600	34.961	349.610	466.682	816.292	812.881

CHAPTER VI

VARIANCES OF RANGE

The exact variance of the range for any finite value of n is not known even for the case of i.i.d. normal variables. The exact variance of the range for the case of stationary Markov models is derived in the first section of this chapter for n of 1 and 2. For higher values of n, the mathematical derivation becomes extremely cumbersome. Therefore, in these cases, and for Markov models with periodic standard deviation, approximate equations are obtained using the data generation method.

6.1 Variance of the Range for Markov Models

The general type of the first-order Markov model is used here,

$$z_t = \rho\, z_{t-1} + \epsilon_t \quad , \qquad 6.1$$

where ρ is the first autocorrelation coefficient of the process z_t and ϵ_t is an i.i.d. variable uncorrelated with z_{t-1}. It is assumed that $E\{z_t\} = E\{\epsilon_t\} = 0$, and $E\{z_t^2\} = 1$, and $E\{\epsilon_t^2\} = (1 - \rho^2)$.

In this case, the partial sums S_0, S_1, and S_2 are

$$\begin{aligned} S_0 &= 0 && = 0 \\ S_1 &= z_1 && = X \\ S_2 &= (1+\rho)z_1 + \epsilon_2 && = (1+\rho)X + Y \end{aligned} \qquad 6.2$$

where for simplicity of derivation the new symbols $X = z_1$ and $Y = \epsilon_2$ are introduced.

For $n = 1$, $R_1 = \max(0,S_1) - \min(0,S_1)$, so that

$R_1 = S_1$ for $S_1 > 0$, and $R_1 = -S_1$ for $S_1 < 0$, or $R_1 = |S_1|$ for $-\infty < S_1 < \infty$.

The second moment of R_1 is

$$E\{R_1^2\} = E\{|S_1|^2\} = E\{S_1^2\} = \sigma_x^2 \qquad 6.3$$

where σ_x denotes the standard deviation of $S_1 = X$.

From Eq. 4.27, the expected value of R_1 is $E\{R_1\} = \sqrt{2/\pi}\ \sigma_x$. Therefore, the variance of R_1 becomes

$$\mathrm{Var}\{R_1\} = E\{R_1^2\} - E^2\{R_1\}$$

$$\mathrm{Var}\{R_1\} = \sigma_x^2\left(1 - \frac{2}{\pi}\right) \qquad 6.4$$

For $n = 2$, $R_2 = \max(0,S_1,S_2) - \min(0,S_1,S_2)$, so that

$R_2 = S_2 - S_1$	for	$S_1 < 0 < S_2$,
$R_2 = -(S_2 - S_1)$	for	$S_2 < 0 < S_1$,
$R_2 = S_2$	for	$0 < S_1 < S_2$,
$R_2 = -S_2$	for	$S_2 < S_1 < 0$,
$R_2 = S_1$	for	$0 < S_2 < S_1$,
$R_2 = -S_1$	for	$S_1 < S_2 < 0$,

which in terms of the variables X and Y, given by Eq. 6.2, become

$R_2 = [(1+\rho)X + Y]$ for $X > 0$, $\rho X + Y > 0$,

$R_2 = -[(1+\rho)X + Y]$ for $X < 0$, $\rho X + Y < 0$,

$R_2 = (\rho X + Y)$ for $X < 0$, $(1+\rho)X + Y > 0$,

$R_2 = -(\rho X + Y)$ for $X > 0$, $(1+\rho)X + Y < 0$,

$R_2 = X$ for $(1+\rho)X + Y > 0$, $\rho X + Y < 0$,

and $R_2 = -X$ for $(1+\rho)X + Y < 0$, $X + Y > 0$.

Because of symmetric regions of integration, the second moment of R_2 is

$$E\{R_2^2\} = 2E\{[(1+\rho)X + Y]^2\} + 2E\{(-\rho X - Y)^2\} + 2E\{X^2\} \qquad 6.5$$

where the moments shown in Eq. 6.5 may be expressed as

$$\begin{aligned} E\{(1+\rho)X+Y\} = {} & (1+\rho)^2 \int_0^\infty \int_{-\rho X}^\infty X^2 f(X) f(Y)\, dY dX + \\ & + 2(1+\rho) \int_0^\infty \int_{-\rho X}^\infty XY\, f(X) f(Y)\, dY dX \\ & + \int_0^\infty \int_{-\rho X}^\infty Y^2 f(X) f(Y)\, dY dX , \end{aligned} \qquad 6.6$$

$$\begin{aligned} E\{(-\rho X - Y)^2\} = {} & \rho^2 \int_0^\infty \int_{-\infty}^{-(1+\rho)X} X^2 f(X) f(Y)\, dY dX + \\ & + 2\rho \int_0^\infty \int_{-\infty}^{-(1+\rho)X} XY\, f(X) f(Y)\, dY dX \\ & + \int_0^\infty \int_{-\infty}^{-(1+\rho)X} Y^2 f(X) f(Y)\, dY dX , \end{aligned} \qquad 6.7$$

and

$$E\{X^2\} = \int_0^{\infty} \int_{-(1+\rho)X}^{-\rho X} X^2 \; f(X) \, f(Y) \, dY dX \quad 6.8$$

with f(X) and f(Y) the density functions given by Eq. 4.9.

The integrals of Eqs. 6.6, 6.7, and 6.8 are equal to

$$\int_0^{\infty} \int_{-\rho X}^{\infty} X^2 \; f(X) \, f(Y) \, dY dX = \frac{1}{2} \; \sigma_x^2$$

$$+ \frac{\rho \, \sigma_x^3 \; \sigma_y}{(2\pi)(\sigma_y^2 + \rho^2 \sigma_x^2)} - \frac{\sigma_x^2}{2\pi} \arctan\left(\frac{\sigma_x}{\rho \, \sigma_x}\right) \quad , \quad 6.9$$

$$\int_0^{\infty} \int_{-\rho X}^{\infty} XY \; f(X) \, f(Y) \, dY dX = \frac{\sigma_x \; \sigma_y^3}{(2\pi)(\sigma_y^2 + \rho^2 \; \sigma_x^2)} \quad 6.10$$

$$\int_0^{\infty} \int_{-\rho X}^{\infty} Y^2 \; f(X) \, f(Y) \, dY dX = \frac{1}{4} \; \sigma_y^2$$

$$- \frac{\rho \; \sigma_x \; \sigma_y{}^3}{(2\pi)(\sigma_y^2 + \rho^2 \; \sigma_x^2)} + \frac{\sigma_y{}^2}{2\pi} \arctan\left(\rho \frac{\sigma_x}{\sigma_y}\right) \quad , \quad 6.11$$

$$\int_0^{\infty} \int_{-\infty}^{-(1+\rho)X} X^2 \; f(X) \, f(Y) \, dY dX =$$

$$- \frac{(1+\rho) \; \sigma_x^3 \; \sigma_y}{(2\pi)[\sigma_y^2 + (1+\rho)^2 \; \sigma_x^2]} + \frac{\sigma_x{}^2}{2\pi} \arctan\left[\frac{\sigma_y}{(1+\rho)\sigma_x}\right] \quad 6.12$$

$$\int_0^{\infty} \int_{-\infty}^{-(1+\rho)X} XY \; f(X) \, f(Y) \, dY dX =$$

$$- \frac{\sigma_x \; \sigma_y^3}{(2\pi)[\sigma_y^2 + (1+\rho)^2 \; \sigma_x^2]} \quad , \quad 6.13$$

$$\int_0^{\infty} \int_{-\infty}^{-(1+\rho)X} Y^2 \; f(X) \, f(Y) \, dY dX = \frac{1}{4} \; \sigma_y^2$$

$$+ \frac{(1+\rho) \; \sigma_x \; \sigma_y^3}{(2\pi)[\sigma_y^2 + (1+\rho)^2 \; \sigma_x^2]} - \frac{\sigma_y^2}{2\pi} \arctan\left[\frac{(1+\rho)\sigma_x}{\sigma_y}\right] \quad , \quad 6.14$$

and

$$\int_0^{\infty} \int_{-(1+\rho)X}^{-\rho X} X^2 \; f(X) \, f(Y) \, dY dX =$$

$$\frac{\sigma_x^3 \; \sigma_y}{(2\pi)} \left\{ \frac{(1+\rho)}{[\sigma_y^2 + (1+\rho)^2 \; \sigma_x^2]} - \frac{\rho}{(\sigma_y^2 + \rho^2 \; \sigma_x^2)} \right\} +$$

$$+ \sigma_x^2 \; \left\{ \frac{1}{2\pi} \arctan\left(\frac{\sigma_y}{\rho \sigma_x}\right) - \frac{1}{2\pi} \arctan\left[\frac{\sigma_y}{(1+\rho) \; \sigma_x}\right] \right\} \quad . \quad 6.15$$

Substituting Eqs. 6.9 through 6.11 into Eq. 6.6, Eqs. 6.12 through 6.14 into Eq. 6.7, and Eq. 6.15 into Eq. 6.8, gives

$$E\{ (1+\rho) \, X + Y \} = (1+\rho)^2 \; \sigma_x^2 \; \left[\frac{1}{2} + \frac{\rho \; \sigma_x \; \sigma_y}{(2\pi)(\sigma_y^2 + \rho^2 \; \sigma_x^2)}\right.$$

$$\left. - \frac{1}{2\pi} \arctan\left(\frac{\sigma_y}{\rho \; \sigma_x}\right)\right] + \frac{(2+\rho) \; \sigma_x \; \sigma_y^3}{2\pi \; (\sigma_y^2 + \rho^2 \; \sigma_x^2)}$$

$$+ \sigma_y^2 \; \left[\frac{1}{4} + \frac{1}{2\pi} \arctan\left(\frac{\rho \; \sigma_x}{\sigma_y}\right)\right] \quad , \quad 6.16$$

$$E\{ (-\rho X - Y)^2 \} = \rho^2 \sigma_x^2 \; \left\{ - \frac{(1+\rho) \; \sigma_x \; \sigma_y}{2\pi \; [\sigma_y^2 + (1+\rho)^2 \; \sigma_x^2]}\right.$$

$$\left. + \frac{1}{2\pi} \arctan\left[\frac{\sigma_y}{(1+\rho) \; \sigma_x}\right] \right\} + \frac{(1-\rho) \; \sigma_x \; \sigma_y^3}{2\pi \; [\sigma_y^2 + (1+\rho)^2 \; \sigma_x^2]}$$

$$+ \sigma_y^2 \; \left\{ \frac{1}{4} - \frac{1}{2\pi} \arctan\left[\frac{(1+\rho) \; \sigma_x}{\sigma_y}\right] \right\} \quad , \quad 6.17$$

and

$$E\{X^2\} = - \frac{\rho \; \sigma_x^3 \; \sigma_y}{2\pi \; (\sigma_y^2 + \rho^2 \; \sigma_x^2)}$$

$$+ \frac{(1+\rho) \; \sigma_x^3 \; \sigma_y}{2\pi \; [\sigma_y^2 + (1+\rho)^2 \; \sigma_x^2]} + \sigma_x^2 \; \left\{ \frac{1}{2\pi} \operatorname{arc\,tan}\left(\frac{\sigma_y}{\rho \; \sigma_x}\right)\right.$$

$$\left. - \frac{1}{2\pi} \arctan\left[\frac{\sigma_y}{(1+\rho) \; \sigma_x}\right] \right\} \quad . \quad 6.18$$

Substituting Eqs. 6.16 through 6.18 into Eq. 6.5, and since $\sigma_x^2 = 1$, and $\sigma_y^2 = 1 - \rho^2$, the second moment of the range R_2 becomes

$$E\{R_2^2\} = 2\,(1+\rho) + \frac{3(1-\rho^2)^{1/2}}{\pi}$$

$$- (1+2\rho) \frac{1}{2\pi} \arctan\left[\frac{(1-\rho)^{1/2}}{(1+\rho)^{1/2}}\right] \quad . \quad 6.19$$

Since the first moment of R_2 is given by Eq. 4.101, the variance of R_2 becomes

$$\mathrm{Var}\{R_n\} = 2(\,1 + \rho)$$

$$- \frac{2}{\pi} + \frac{(\,1 + \rho)^{1/2}}{\pi} \; [\,3\,(\,1 - \rho)^{1/2} - (\,1 + \rho)^{1/2}$$

$$- 2\sqrt{2}\,] - (1 + 2\rho) \arctan\left[\frac{(\,1 - \rho)^{1/2}}{(\,1 + \rho)^{1/2}}\right] \quad . \quad 6.20$$

6.2 Approximate Variance of the Range for Markov Models with Constant Standard Deviation

In this section, the results of the simulation approach are presented for obtaining the variance of the range for Markov models with constant standard deviation. First, however, a sensitivity analysis was performed to see the effect of the periodicity in the autocorrelation coefficients on the magnitude of the variance of the range.

For the first-order Markov model, as given by Eq. 3.4 for $m = 1$, the variance of the range was computed for n up to 60 and for a periodic first autocorrelation coefficient. Figure 6.1 gives the plot of $\mathrm{Var}\{R_n\}$ against n for $\overline{\rho}_{1,\tau} = 0.6$ and for three values of $s(\rho_{1,\tau})$, 0.0, 0.102 and 0.207. This figure shows that the periodicity in $\rho_{1,\tau}$ increases the variance of the range as the value of $s(\rho_{1,\tau})$ increases. It also shows that the increase in $\mathrm{Var}\{R_n\}$ is augmented as n increases. No attempt was made to quantify these experienced increases of $\mathrm{Var}\{R_n\}$ for particular values of $\overline{\rho}_{1,\tau}$ and $s(\rho_{1,\tau})$.

For the second and third-order Markov models, no appreciable differences are found between the variance of the range obtained with constant and periodic autocorrelation coefficients. The results obtained in these cases are shown in Figs. 6.2 and 6.3 for the second and third-order Markov models, respectively.

Experimental curves are obtained by simulation for the variance of the range of the first and second-order Markov models with constant autoregression coefficients. The plot of the values of $\mathrm{Var}\{R_n\}$ against n suggests that a straight line fit is good in cases of $n \geqslant 6$. Therefore, the variance of the range was approximated by

$$\mathrm{Var}\{R_n\} = \sigma^2 \,[\,A + B\,n\,] \;, \qquad 6.21$$

where σ is the constant standard deviation and the linear regression coefficients A and B are functions of the autoregression coefficients of the Markov model considered.

For the first-order Markov model with $\sigma = 1$, Fig. 6.4 shows the plot of $\mathrm{Var}\{R_n\}$ against n for n up to 50 and for various values of ρ. The straight line fit to values of $\mathrm{Var}\{R_n\}$ obtained from simulated samples is shown to be a good approximation. Table 6.1 also gives the values of the simulated and fitted variance of the range for various values of n and ρ. The linear regression parameters of Eq. 6.21 are given in Table 6.2 for various values of ρ. They are also shown in Fig. 6.5, which may be particularly useful for finding the A and B values for ρ not explicitly obtained.

P. Sutabutra (1967) suggested another empirical equation to approximate the variance of the range of first-order Markov models, namely

$$\mathrm{Var}\{R_n\} = C\,(\,n,\rho)\;\sum_{i=1}^{n} i^{-1}\;\mathrm{Var}\{S_i\}\,, \qquad 6.22$$

with

$$C\,(n,\rho) = 0.2181\,(\,1 + 0.4\rho + 0.4\rho^2)\,(\,1 + \frac{1.5}{n}\,)\;. \qquad 6.23$$

A comparison was made between the percentage relative errors obtained in using Eqs. 6.21 and 6.22 for approximating the variance of the range. The results of this comparison are shown in Table 6.3 and indicate that the Eq. 6.21 gives a better fit to the simulated variances of the range, decreasing the errors considerably with respect to those obtained by Eq. 6.22.

For the second-order Markov model, the simulated and fitted $\mathrm{Var}\{R_n\}$ against n are shown in Figs. 6.6, 6.7 and 6.8 for n up to 100 and for various values of ρ_1 and ρ_2, the first and second autocorrelation coefficients, respectively. The straight line fit in this case is also very good, and the respective linear regression coefficients A and B of Eq. 6.21 are given in Table 6.4.

6.3 Approximate Variances of the Range for Markov Models with Periodic Mean and Periodic Standard Deviation

In this section, the variance of the range is obtained by computer simulation for the general case

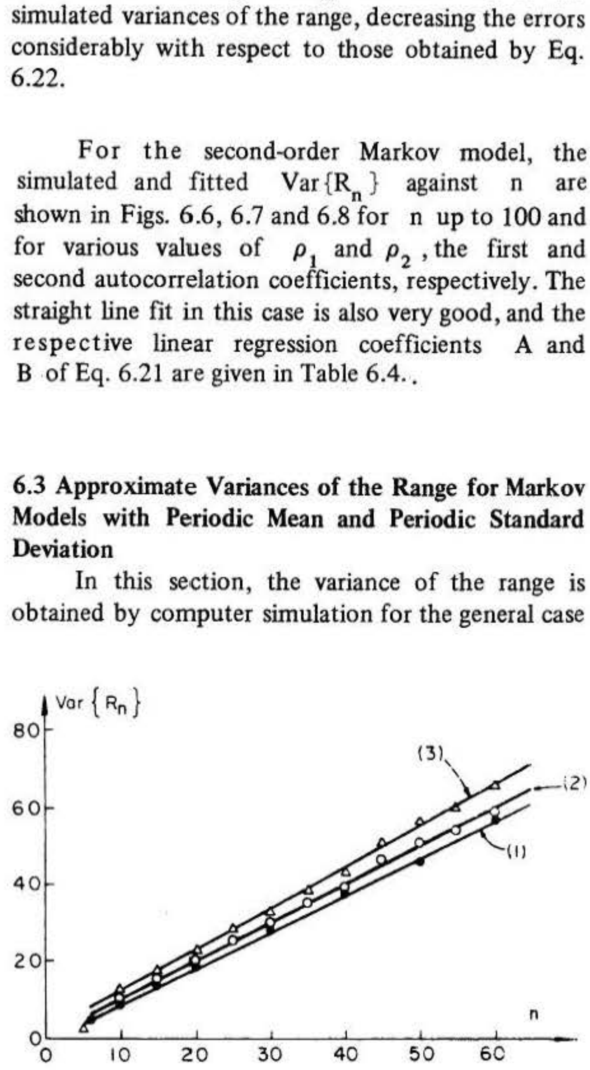

Fig. 6.1 Variance of the range for the first-order Markov model with constant and periodic first autocorrelation coefficient with $\overline{\rho}_{1,\tau} = 0.60$ and (1) $s(\rho_{1,\tau}) = 0.0$, (2) $s(\rho_{1,\tau}) = 0.102$, and (3) $s(\rho_{1,\tau}) = 0.207$.

of Markov models with periodic mean and periodic standard deviation. From Eqs. 3.3, 3.4 and 3.5

$$x_{p,\tau} = \mu_\tau + \sigma_\tau \left[\rho\, z_{p,\tau-1} + \sqrt{1-\rho^2}\; \epsilon_{p,\tau} \right] , \qquad 6.24$$

with μ_τ , σ_τ , ρ , $z_{p,\tau}$ and $\epsilon_{p,\tau}$ defined as in Section 3.1. In obtaining the variance of the range, it is assumed that the output y_t of Eq. 3.2 is $\overline{\mu_\tau}$.

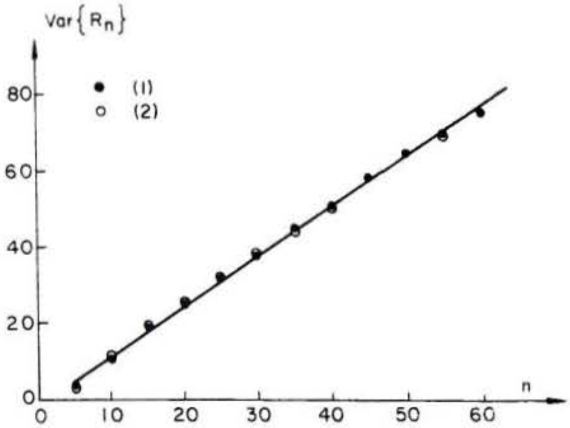

Fig. 6.2 Variance of the range for the second-order Markov model with constant and periodic first and second autocorrelation coefficients. (1) $\rho_{1,\tau} = \rho_1 = 0.60$ and $\rho_{2,\tau} = \rho_2 = 0.45$, and (2) $\overline{\rho}_{1,\tau} = 0.60$, $\overline{\rho}_{2,\tau} = 0.45$ and $s(\rho_{k,\tau}) = 0.102$ for $k = 1$ and 2 .

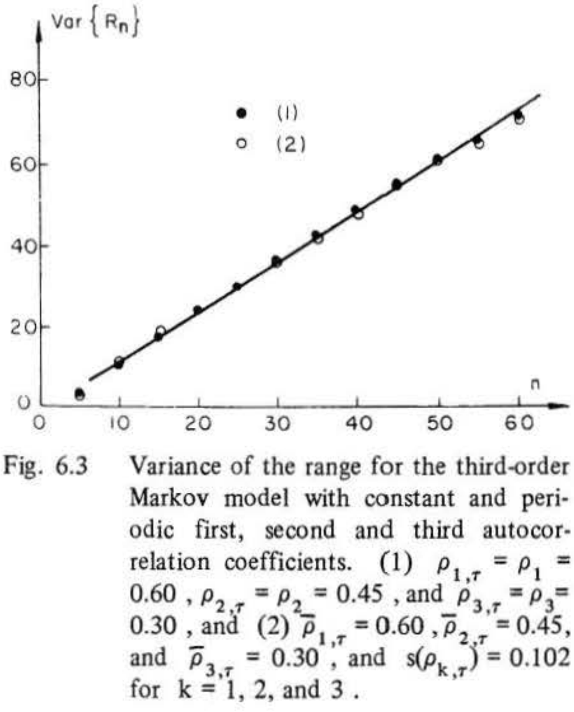

Fig. 6.3 Variance of the range for the third-order Markov model with constant and periodic first, second and third autocorrelation coefficients. (1) $\rho_{1,\tau} = \rho_1 = 0.60$, $\rho_{2,\tau} = \rho_2 = 0.45$, and $\rho_{3,\tau} = \rho_3 = 0.30$, and (2) $\overline{\rho}_{1,\tau} = 0.60$, $\overline{\rho}_{2,\tau} = 0.45$, and $\overline{\rho}_{3,\tau} = 0.30$, and $s(\rho_{k,\tau}) = 0.102$ for $k = 1, 2$, and 3 .

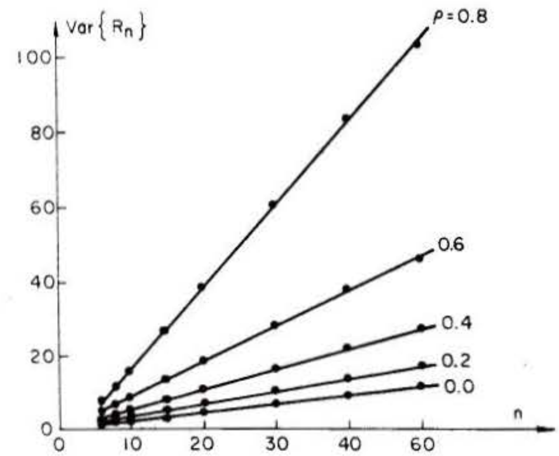

Fig. 6.4 Variance of the range obtained from simulated samples and fitted linear function of Eq. 6.21, for the first-order Markov model with constant first autocorrelation coefficient ρ .

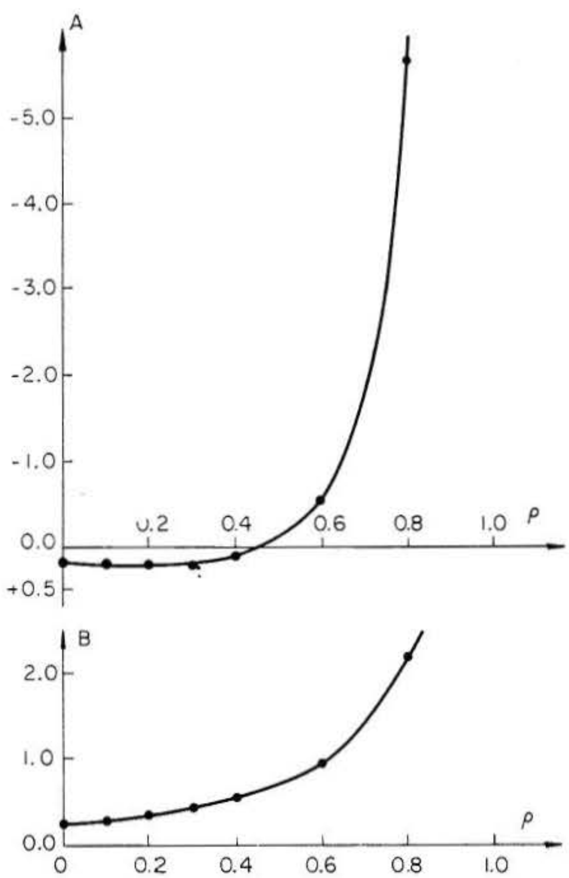

Fig. 6.5 Regression coefficients of fitted linear function (Eq. 6.21) to variance of the range of the first-order Markov model.

TABLE 6.1 COMPARISON OF VARIANCES OF THE RANGE OBTAINED FROM SIMULATED SAMPLES AND BY EQ. 6.21, FOR THE FIRST-ORDER MARKOV MODEL OF EQ. 3.4 WITH CONSTANT ρ, FOR n UP TO 50 AND VARIOUS VALUES OF ρ.

	$\rho = 0.0$		$\rho = 0.10$(*)		$\rho = 0.20$(*)		$\rho = 0.30$		$\rho = 0.40$(*)		$\rho = 0.60$(*)		$\rho = 0.80$(*)	
n	Simulated	Equation	Simulated	Equation	Simulated	Equation	Simulated	Equation	Simulated	Equation	Simulated	Equation	Simulated	Equation
6	1.4783	1.5996	1.8054	1.8586	2.1670	2.2503	2.6344	2.8824	3.1657	3.3341	4.7701	5.1157	7.5608	7.4800
8	1.9562	2.0672	2.3769	2.4091	2.8840	2.9296	3.5247	3.7667	4.3250	4.4067	6.7870	6.9964	11.6119	11.8627
10	2.5070	2.5348	2.9779	2.9597	3.6300	3.6090	4.5914	4.6511	5.5132	5.4792	8.8479	8.8770	15.9342	16.2455
15	3.8849	3.7038	4.3127	4.3360	5.2914	5.3072	7.2030	6.8620	8.2048	8.1607	13.7095	13.5785	27.0679	27.2023
20	4.9867	4.8728	5.7013	5.7124	7.0060	7.0055	9.3578	9.0729	10.8047	10.8421	18.4344	18.2800	38.1203	38.1591
30	7.2892	7.2108	8.5668	8.4652	10.5303	10.4021	13.6938	13.4947	16.4264	16.2049	28.0569	27.6830	60.5044	60.0728
40	9.5269	9.5488	11.4259	11.2179	14.0707	13.7987	17.8166	17.9165	22.0310	21.5678	37.9329	37.0861	83.6533	81.9865
50	11.7952	11.8868	13.7626	13.9706	16.9184	17.1952	22.1630	22.3383	26.4554	26.9306	45.5670	46.4891	102.4562	103.9001

(*) For these values of ρ, the $Var(R_n)$ obtained by simulation were taken from P. Sutabutra (1967).

TABLE 6.2 PARAMETERS OF LINEAR REGRESSION FOR THE VARIANCE OF THE RANGE OF THE FIRST-ORDER MARKOV MODEL OF EQ. 3.4.

	Values of ρ						
	0.0	0.1	0.2	0.3	0.4	0.6	0.8
A	0.19676	0.20693	0.21238	0.22929	0.11639	-0.52607	-5.66821
B	0.23380	0.27527	0.33966	0.44218	0.53629	0.94030	2.19137
Standard Error Of Regr. Coeff.	0.00285	0.00305	0.00402	0.00606	0.00696	0.01323	0.02190
Correlation Coefficient	0.99956	0.99963	0.99958	0.99944	0.99950	0.99941	0.99970

TABLE 6.3 COMPARISON OF PERCENTAGE RELATIVE ERRORS OBTAINED IN USING EQS. 6.21 AND 6.22 FOR COMPUTING THE VARIANCES OF THE RANGE OF THE FIRST-ORDER MARKOV MODELS.

	RELATIVE ERRORS IN PERCENTAGE									
	$\rho = 0.10$		$\rho = 0.20$		$\rho = 0.40$		$\rho = 0.60$		$\rho = 0.80$	
n	Equation 6.22	Equation 6.21	Equation 6.22	Equation 6.21	Equation 6.22	Equation 6.21	Equation 6.22	Equation 6.21	Equation 6.22	Equation 6.21
6	-6.088	-2.862	-4.677	-3.702	-2.096	-5.051	+1.682	-6.758	+1.220	+1.080
8	-3.745	-1.336	-2.690	-1.560	-0.864	-1.854	+2.516	-2.991	+1.485	-2.115
10	-1.321	+0.618	-0.750	+0.585	-0.006	+0.620	+2.362	-0.327	+14.901	-1.916
15	-1.824	-0.539	-2.147	-0.299	-2.853	+0.540	-1.657	+0.965	+10.696	-0.494
20	-1.229	-0.194	-2.246	+0.006	-4.542	-0.345	-5.140	+0.845	+5.438	-0.102
30	+0.352	+1.201	-1.611	+1.217	-5.870	+1.367	-8.653	+1.351	-1.468	+0.718
40	+1.061	+1.854	-1.268	+1.971	-6.374	+2.148	-10.133	+2.284	-4.956	+2.033
50	-2.238	-1.489	-4.986	-1.610	-10.750	-1.765	-15.344	-1.983	-11.207	-1.390
Average Absolute Error %	2.232	1.262	2.547	1.369	4.169	1.711	5.936	2.188	6.421	1.231

TABLE 6.4 REGRESSION COEFFICIENTS OF LINEAR FUNCTION FIT TO VARIANCES OF THE RANGE OF THE SECOND-ORDER MARKOV MODEL.

	$\rho_1=0.40$			$\rho_1=0.60$			$\rho_1=0.80$	
	$\rho_2=0.10$	$\rho_2=0.20$	$\rho_2=0.30$	$\rho_2=0.25$	$\rho_2=0.30$	$\rho_2=0.40$	$\rho_2=0.40$	$\rho_2=0.50$
A	-0.33983	-0.66148	-1.21142	-0.12678	-0.51548	-1.86086	3.98548	2.20953
B	0.47315	0.60505	0.77663	0.65861	0.77870	1.08101	0.38242	0.90797
Standard Error Of Regr. Coeff.	0.00426	0.00519	0.00617	0.00609	0.00670	0.00860	0.00837	0.01021
Correlation Coefficient	0.99931	0.99938	0.99946	0.99927	0.99937	0.99946	0.99595	0.99893

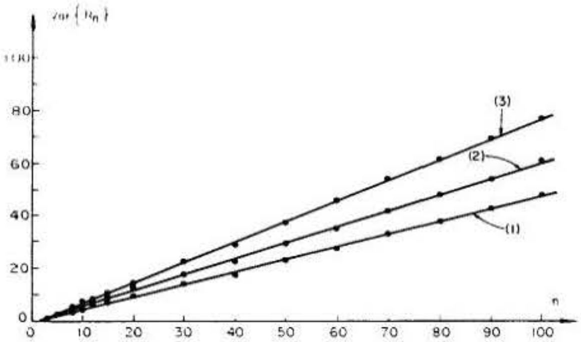

Fig. 6.6 Variance of the range obtained from simulated samples and fitted linear function of Eq. 6.21, for the second-order Markov model with constant autocorrelation coefficients. Cases of $\rho_1 = 0.40$ and (1) $\rho_2 = 0.10$, (2) $\rho_2 = 0.20$, and (3) $\rho_2 = 0.30$.

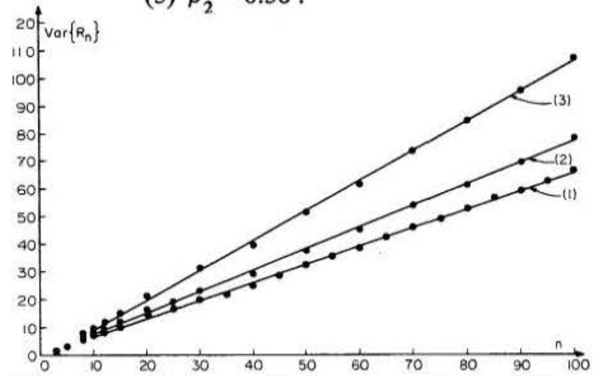

Fig. 6.7 Variance of the range obtained from simulated samples and fitted linear function of Eq. 6.21, for the second-order Markov model with constant autocorrelation coefficients. Cases of $\rho_1 = 0.60$ and (1) $\rho_2 = 0.25$ (2) $\rho_2 = 0.30$, and (3) $\rho_2 = 0.40$.

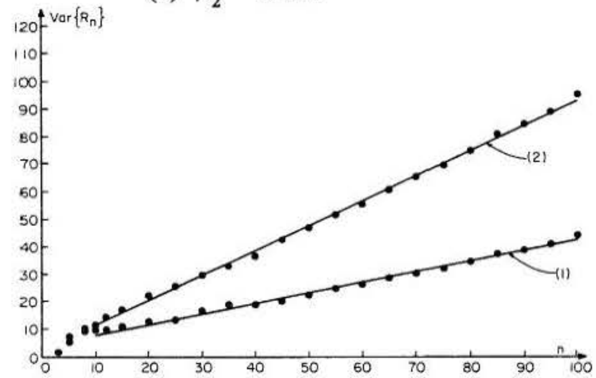

Fig. 6.8 Variance of the range obtained from simulated samples and fitted linear func- of Eq. 6.21, for the second-order Markov model with constant autocorrelation coefficients. Cases of $\rho_1 = 0.80$ and (1) $\rho_2 = 0.40$ and (2) $\rho_2 = 0.50$.

In general, whenever periodicity exists in parameters of the components of the models representing the inputs and outputs, the resulting variance of the range is also a periodic function. The first simulation was performed to see whether the characteristics of Var$\{R_n\}$, when μ_τ and σ_τ are periodic functions, (see Fig. 6.9) depart significantly from the stationary cases. These curves are shown in Fig. 6.10, where the mean and standard deviation of μ_τ are $\overline{\mu}_\tau = 20$ and $s(\mu_\tau) = 12.40$, and the mean and standard deviation of σ_τ are $\overline{\sigma}_\tau = 5.0$ and $s(\sigma_\tau) = 2.79$. For these cases, Fig. 6.10 shows, for $\rho = 0.0$ and $\rho = 0.60$, the variance of the range against n for values of n up to 60. This figure shows also how complex, Var$\{R_n\}$ becomes whenever one uses models with periodic functions.

A general characteristic presented by Fig. 6.10 is that after a transition region the variance of the range becomes a non-decreasing function of n , because the effect of periodicities on Var$\{R_n\}$ decreases with n . This characteristics differs from that of the expected range for which, as will be shown in Chapter VII, the expected range is always a non-decreasing function for all values of n . Figure 6.10 also shows that Var$\{R_n\}$ is a periodic function with its phases and amplitudes dependent on the periodic functions μ_τ and σ_τ . The plot also shows that the amplitudes of the periodic function Var$\{R_n\}$ decrease as n becomes large. Similarly, as in the case of the variance of range for stationary Markov models, the effect of dependence, in this case of periodic μ_τ and σ_τ , is considerable.

Strictly speaking, the variance of the range for models of the type of Eq. 6.24 depends on amplitudes and phases of periodic functions μ_τ and σ_τ as well as on ρ . If one considers the Fourier fit of μ_τ and σ_τ , as suggested by Eq. 3.6, the number of parameters to consider for determining the variance of the range becomes excessive. Therefore, the approach in this study is to look for other parameters which are functions of μ_τ and σ_τ , such as the standard deviation $s(\mu_\tau)$ and the mean and standard deviation $\overline{\sigma}_\tau$ and $s(\sigma_\tau)$. By choosing only the parameters $s(\mu_\tau)$, $\overline{\sigma}_\tau$ and $s(\sigma_\tau)$ as representative of μ_τ and σ_τ , one mainly neglects the influence of their phases. In order to see how great this influence is on the variance of the range, a sensitivity analysis was performed with $s(\mu_\tau) = 12.40$ and for two phases, and with $\overline{\sigma}_\tau = 10$, $s(\sigma_\tau) = 6.87$ and for three phases. These functions, μ_τ and σ_τ , are shown in Fig. 6.11.

Five different combinations of symmetric and skewed μ_τ and σ_τ, as shown in Fig. 6.11 were considered, and in all cases the first autocorrelation coefficients was $\rho = 0.60$. The variances of the range obtained in these 5 cases are shown in Fig. 6.12. This figure shows that, basically, the influence of the different phases of μ_τ and σ_τ is significant only in the transition region. Beyond this region or for $n > 50$, they all tend to converge to approximately the same variances. Therefore, for all practical purposes, the influence of phases in μ_τ and σ_τ may be neglected for larger n. Subsequently, all the analysis is based on symmetric functions μ_τ and σ_τ, and the only parameters used to define μ_τ and σ_τ are $s(\mu_\tau)$, and $\overline{\sigma}_\tau$ and $s(\sigma_\tau)$. The different functions of μ_τ and σ_τ considered afterward are shown in Figs. 6.13 and 6.14.

Another characteristic observed from the analysis of the computer simulated results is that, for given values of $\overline{\sigma}_\tau$, $s(\sigma_\tau)$ and ρ, the influence of μ_τ is significant only in the transition region. For $n > 50$, the variances of the range tend to converge to approximately the same values. Table 6.5 gives a comparison of variances obtained for values of n up to 350 for the cases of $\overline{\sigma}_\tau = 20$, $s(\sigma_\tau) = 0$ and 14.22, $\rho = 0$, and $s(\mu_\tau) = 0$ and $s(\mu_\tau) = 190.96$. Table 6.6 gives the comparison for the same case as above except that $\rho = 0.60$. These comparisons are also shown in Figs. 6.15, 6.16, and 6.17. The results of this analysis lead to the conclusion that for large values of n, say $n > 50$, the variance of the range for the general case of Markov models with a periodic mean μ_τ, and a periodic standard deviation σ_τ, depends only on $\overline{\sigma}_\tau$, $s(\sigma_\tau)$ and ρ. That is,

$$\mathrm{Var}\{R_n\} = f(\overline{\sigma}_\tau, s(\sigma_\tau), \rho) \quad . \qquad 6.25$$

The restriction on n for the validity of Eq. 6.25, for all practical purposes is not important, because, whenever one considers models with periodic components, one is dealing with, say, with monthly or weekly values and so only the variances of the range for large values of n are of interest.

The variances of the range for values of n up to 350 and various values of $\overline{\sigma}_\tau$, $s(\sigma_\tau)$ and ρ are obtained and are presented in Tables 6.7, 6.8, and 6.9. They are also shown in Figs. 6.18 through 6.26. In all cases analyzed, the plot in arithmetic scale of $\mathrm{Var}\{R_n\}$ against n follows approximately a straight line. For the particular cases of $s(\sigma_\tau) = 0$ and $\rho = 0$, the values presented in the respective tables and figures were obtained by using Feller's asymptotic formula, given by Eq. 2.5. For the cases of $s(\sigma_\tau) = 0$ and $\rho \neq 0$, they were obtained by using the empirical results of application of Eq. 6.21.

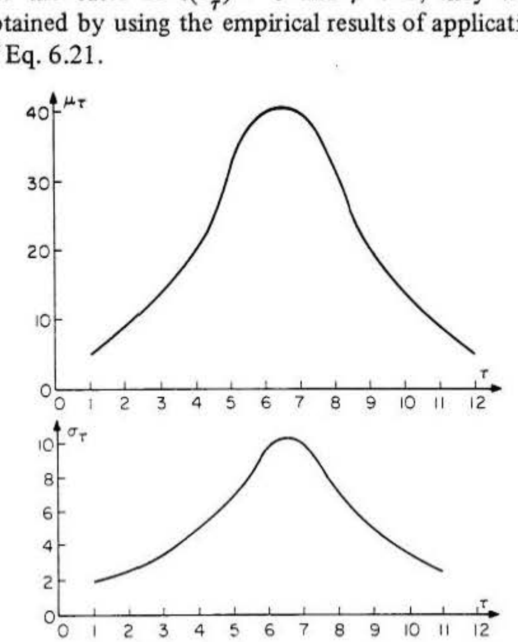

Fig. 6.9 Periodic mean μ_τ, with $\overline{\mu}_\tau = 20$ and $s(\mu_\tau) = 12.40$, and periodic standard deviation σ_τ, with $\overline{\sigma}_\tau = 5$ and $s(\sigma_\tau) = 2.79$, considered when $\mathrm{Var}\{R_n\}$ of Fig. 6.10 are obtained by simulation.

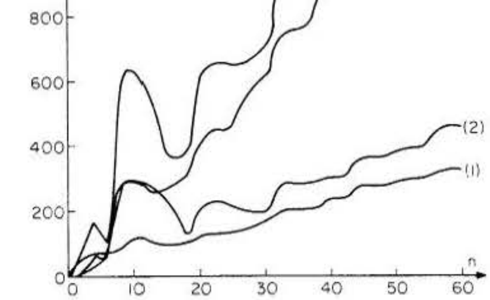

Fig. 6.10 Variance of the range obtained from simulated samples for first-order Markov models with $\overline{\mu}_\tau = 20$ and $s(\mu_\tau) = 12.40$, and with $\overline{\sigma}_\tau = 5$ and (1) $s(\sigma_\tau) = 0.0$ and $\rho = 0.0$, (2) $s(\sigma_\tau) = 2.79$ and $\rho = 0.0$, (3) $s(\sigma_\tau) = 0.0$ and $\rho = 0.60$, and (4) $s(\sigma_\tau) = 2.79$ and $\rho = 0.60$.

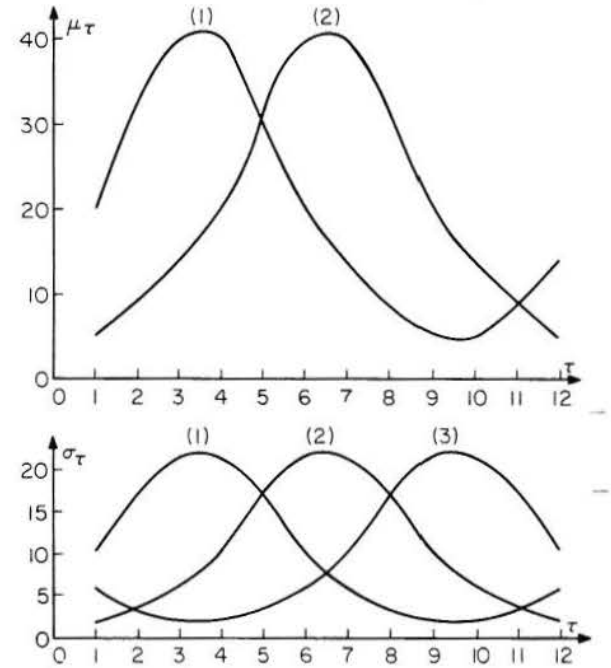

Fig. 6.11 Periodic mean μ_τ with $\bar{\mu}_\tau = 20$ and $s(\mu_\tau) = 12.40$ for two different phases (upper graph) and periodic standard deviation σ_τ with $\bar{\sigma}_\tau = 10$ and $s(\sigma_\tau) = 6.87$ for three different phases (lower graph). These μ_τ and σ_τ are used in obtaining variances of Fig. 6.12.

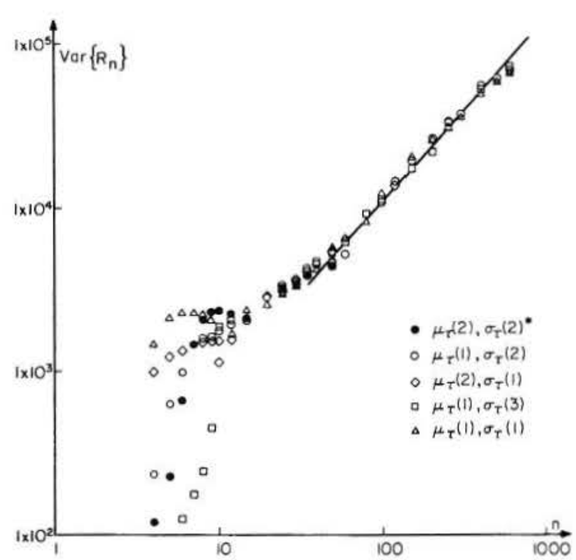

Fig. 6.12 Variance of the range obtained from simulated samples for $s(\mu_\tau) = 12.40$, $\bar{\sigma}_\tau = 10$, $s(\sigma_\tau) = 6.87$ and $\rho = 0.60$, and five combinations of phases of μ_τ and σ_τ. (*number in parenthesis refer to types of μ_τ and σ_τ indicated in in Fig. 6.11).

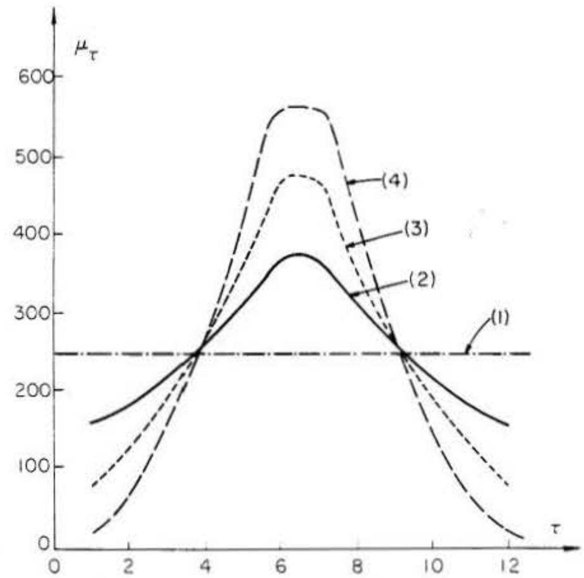

Fig. 6.13 Four different periodic mean μ_τ used in part of this chapter and Chapter VII. They have $\bar{\mu}_\tau = 250$ and $s(\mu_\tau)$ equal to (1) 0.0, (2) 73.03, (3) 134.04, and (4) 190.96.

TABLE 6.5 COMPARISON OF THE VARIANCE OF THE RANGE FOR MODELS WITH $s(\mu_\tau) = 0$ AND $s(\mu_\tau) = 190.96$ IN CASE OF $\rho = 0$, AND BOTH A CONSTANT AND A PERIODIC STANDARD DEVIATION.

n	$\bar{\sigma}_\tau=20, s(\sigma_\tau)=0, \rho=0.0$		$\bar{\sigma}_\tau=20, s(\sigma_\tau)=14.22, \rho=0.0$	
	$s(\mu_\tau)=0$	$s(\mu_\tau)=190.96$	$s(\mu_\tau)=0$	$s(\mu_\tau)=190.96$
1	129.67	348.32	5.19	13.93
3	276.42	971.03	56.33	175.35
6	684.07	668.10	1285.21	1217.20
10	1543.10	2430.48	2305.41	7190.78
15	1837.27	1571.58	2346.62	2562.93
20	2132.30	1727.13	3183.43	3417.79
30	2705.41	2066.21	3905.72	3484.86
40	3422.04	3399.94	4766.91	5098.71
50	3845.00	3750.00	4827.00	5147.00
75	6069.00	6329.00	8612.00	9325.00
100	8523.00	8513.00	13913.00	14648.00
150	14340.00	14403.00	21683.00	23996.00
200	20313.00	20572.00	28446.00	28852.00
250	26300.00	27188.00	36932.00	36514.00
300	29952.00	29944.00	43053.00	41357.00
350	35259.00	34672.00	52066.00	51619.00

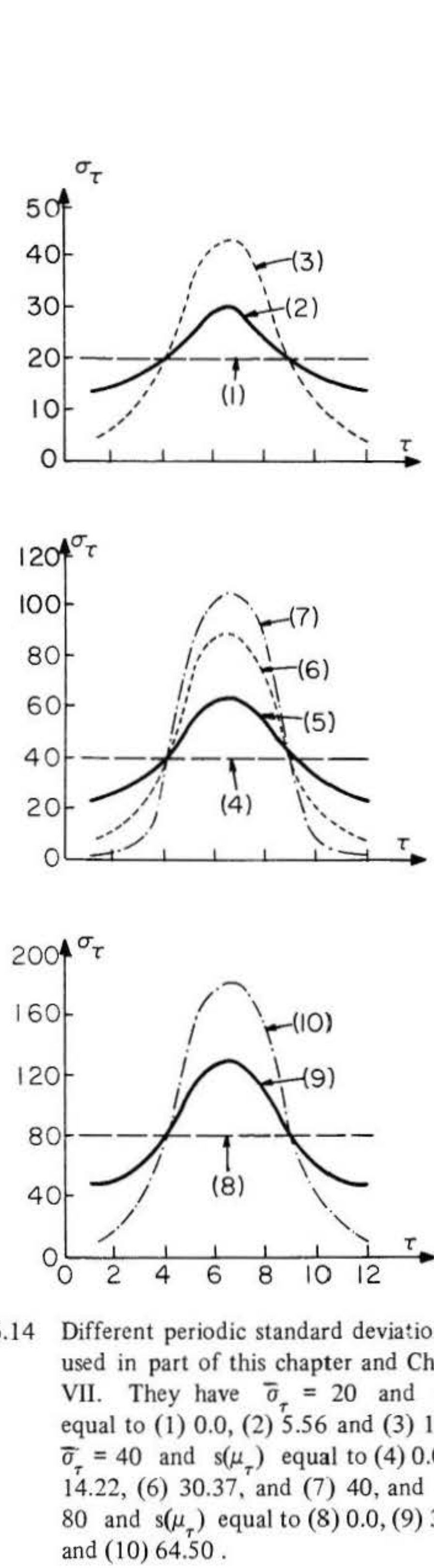

Fig. 6.14 Different periodic standard deviation σ_τ used in part of this chapter and Chapter VII. They have $\overline{\sigma}_\tau = 20$ and $s(\mu_\tau)$ equal to (1) 0.0, (2) 5.56 and (3) 14.22, $\overline{\sigma}_\tau = 40$ and $s(\mu_\tau)$ equal to (4) 0.0, (5) 14.22, (6) 30.37, and (7) 40, and $\overline{\sigma}_\tau$ = 80 and $s(\mu_\tau)$ equal to (8) 0.0, (9) 30.37 and (10) 64.50 .

TABLE 6.6 COMPARISON OF THE VARIANCE OF THE RANGE FOR MODELS WITH $s(\mu_\tau) = 0$ AND $s(\mu_\tau) = 190.96$ IN CASE OF $\rho = 0.60$ AND BOTH A CONSTANT AND A PERIODIC STANDARD DEVIATION.

n	$\overline{\sigma}_\tau=20, s(\sigma_\tau)=0, \rho=0.60$		$\overline{\sigma}_\tau=20, s(\sigma_\tau)=14.22, \rho=0.60$	
	$s(\mu_\tau)=0$	$s(\mu_\tau)=190.96$	$s(\mu_\tau)=0$	$s(\mu_\tau)=190.96$
1	129.67	348.32	5.19	13.93
3	599.00	1986.98	100.90	304.91
6	2014.22	1278.43	3012.91	2133.38
10	4581.92	6857.56	8279.99	20605.91
15	7358.97	4148.05	9146.76	6656.39
20	8835.53	5868.06	12798.75	10803.38
30	11188.38	7159.14	14875.94	11583.52
40	14343.70	11688.07	17908.69	16041.51
50	16012.00	14284.00	18215.00	16705.00
75	23970.00	24103.00	28489.00	29698.00
100	33899.00	32673.00	44957.00	45094.00
150	56693.00	55238.00	77096.00	77269.00
200	79837.00	80796.00	104280.00	106030.00
250	104991.00	106130.00	132695.00	132174.00
300	120032.00	118882.00	146796.00	148425.00
350	139950.00	138456.00	169876.00	171444.00

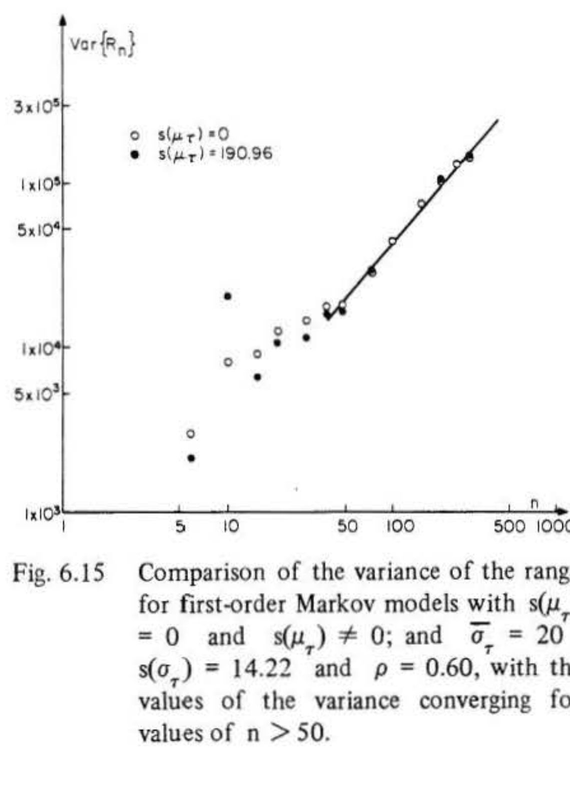

Fig. 6.15 Comparison of the variance of the range for first-order Markov models with $s(\mu_\tau) = 0$ and $s(\mu_\tau) \neq 0$; and $\overline{\sigma}_\tau = 20$, $s(\sigma_\tau) = 14.22$ and $\rho = 0.60$, with the values of the variance converging for values of n > 50.

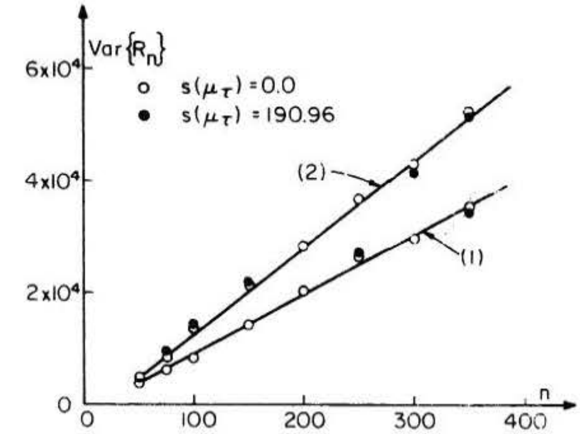

Fig. 6.16 Comparison of variances of the range for models with $s(\mu_\tau) = 0$ and $s(\mu_\tau) \neq 0$. Cases of $\rho = 0$, and both constant σ_τ with (1) $\sigma_\tau = 20$, and periodic σ_τ with (2) $\overline{\sigma}_\tau = 20$ and $s(\sigma_\tau) = 14.22$.

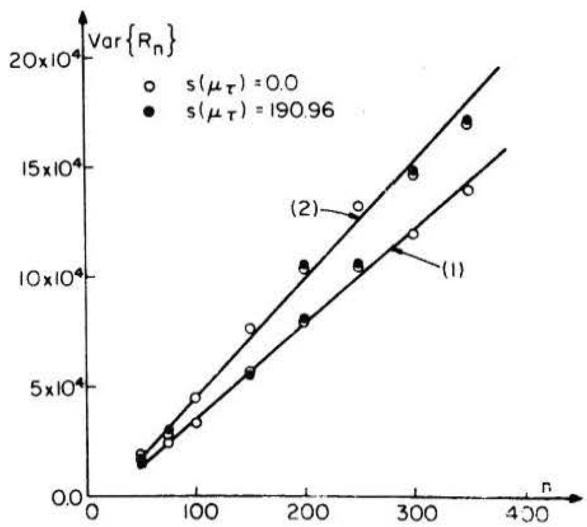

Fig. 6.17 Comparison of variances of the range for first-order Markov models with $s(\mu_\tau) = 0$ and $s(\mu_\tau) \neq 0$. Cases of $\rho = 0.60$, and both constant σ_τ with (1) $\sigma_\tau = 20$, and periodic σ_τ with (2) $\overline{\sigma}_\tau = 20$ and $s(\sigma_\tau) = 14.22$.

TABLE 6.7 VARIANCE OF THE RANGE FOR MARKOV MODELS WITH PERIODIC STANDARD DEVIATION. CASES OF $\overline{\sigma} = 20$ AND THREE VALUES OF $s(\sigma_\tau)$.

n	Var{R_n}: $\overline{\sigma}=20$, $s(\sigma_\tau)=0.$			$\overline{\sigma}=20$, $s(\sigma_\tau)=5.56$			$\overline{\sigma}=20$, $s(\sigma_\tau)=14.22$		
	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$
50	4360	8936	18596	3529	6764	15381	4827	8829	18215
75	6540	13358	27999	6098	11099	23254	8612	15234	28489
100	8720	17780	37402	9634	17594	35764	13913	24746	44957
150	13080	26624	56208	15469	28954	61987	21683	39743	77096
200	17440	35468	75014	21767	40963	87886	28446	52219	104280
250	21800	44312	93819	28219	52690	112411	36932	66814	132695
300	26160	53156	112626	32392	59979	127295	43053	75890	146796
350	30520	62000	131432	37762	69691	146332	52066	91049	169876

TABLE 6.8 VARIANCE OF THE RANGE FOR MARKOV MODELS WITH PERIODIC STANDARD DEVIATION. CASES OF $\overline{\sigma}_\tau = 40$ AND THREE VALUES OF $s(\sigma_\tau)$.

n	Var{R_n}: $\overline{\sigma}_\tau=40$, $s(\sigma_\tau)=0.$			$\overline{\sigma}_\tau=40$, $s(\sigma_\tau)=14.22$			$\overline{\sigma}_\tau=40$, $s(\sigma_\tau)=40.0$		
	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$
50	17440	35743	74382	14547	27782	62526	26032	45260	85227
75	26160	53431	111994	25494	46089	94512	43687	75815	134958
100	34880	71119	149606	40876	74125	147997	70189	122480	212159
150	52320	106495	224830	64929	121190	256388	111187	201578	368682
200	69760	141871	300054	89614	169499	363897	142795	260791	482370
250	87200	177247	375278	116904	219562	463106	178178	323102	599622
300	104640	212623	450502	135368	250202	524378	215522	370493	655538
350	122080	247999	525726	158209	290363	600772	265648	448575	764821

TABLE 6.9 VARIANCE OF THE RANGE FOR MARKOV MODELS WITH PERIODIC STANDARD DEVIATION. CASES OF $\overline{\sigma}_\tau = 80$ AND THREE VALUES OF $s(\sigma_\tau)$.

	Var{R_n}: $\overline{\sigma}_\tau=80$, $s(\sigma_\tau)=0.0$			$\overline{\sigma}_\tau=80$, $s(\sigma_\tau)=30.37$			$\overline{\sigma}_\tau=80$, $s(\sigma_\tau)=64.50$		
	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$	$\rho=0.0$	$\rho=0.30$	$\rho=0.60$
50	69760	142971	297529	58422	111440	250830	85592	153305	306865
75	104640	213724	447977	102753	186023	380107	148168	261040	483966
100	139520	284475	598425	166088	300651	597657	238804	422323	758703
150	209280	425979	899321	265143	492915	1036389	375548	687228	1306274
200	279040	567484	1200217	364149	686504	1470519	487246	894681	1744844
250	348800	708987	1501113	473063	888691	1870963	627225	1137926	2207363
300	418560	850492	1802009	548731	1013787	2115339	742078	1293311	2433665
350	488320	991996	2102905	642178	1173561	2420344	902639	1559591	2819232

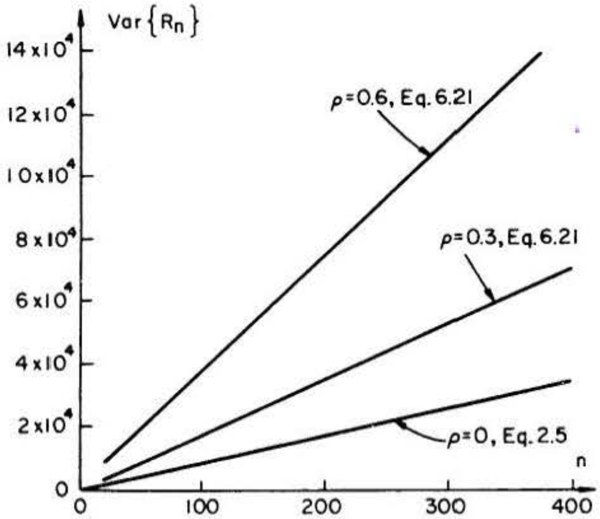

Fig. 6.18 Variance of the range for Markov models with constant standard deviation. Cases of $\sigma_\tau = 20$ and $\rho = 0.0, 0.3,$ and 0.6 .

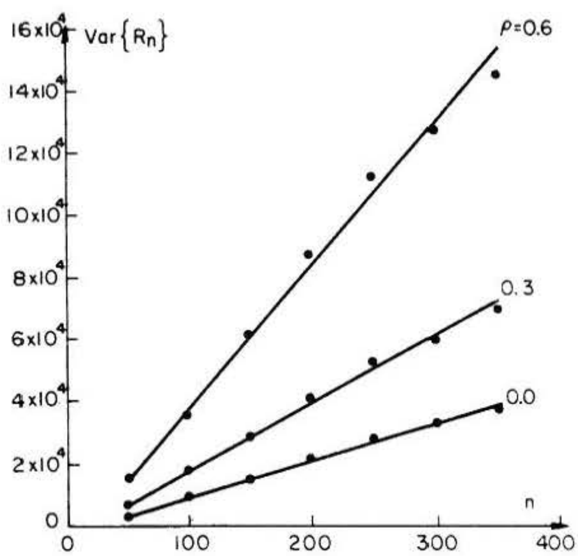

Fig. 6.19 Variance of the range for first-order Markov models with periodic standard deviation. Cases of $\overline{\sigma}_\tau = 20$, $s(\sigma_\tau) = 5.56$ $\rho = 0.0, 0.3,$ and 0.6 .

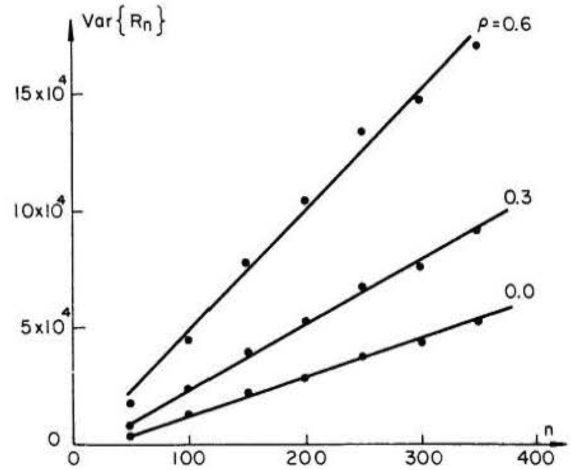

Fig. 6.20 Variance of the range for first-order Markov models with periodic standard deivation. Cases of $\overline{\sigma}_\tau = 20$, $s(\sigma_\tau) =$ 14.22 and $\rho = 0.0, 0.3,$ and 0.6 .

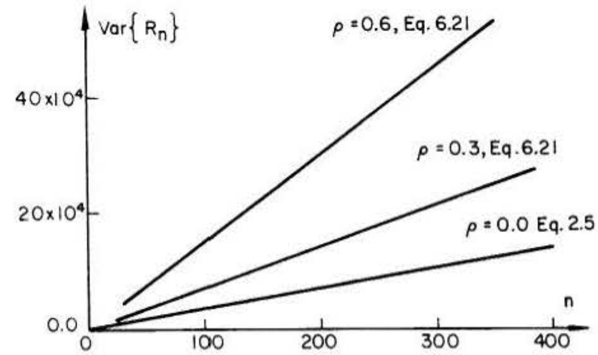

Fig. 6.21 Variance of the range for first-order Markov models with constant standard deviation. Cases of $\sigma_\tau = 40$ and $\rho = 0.0,$ $0.3,$ and 0.6 .

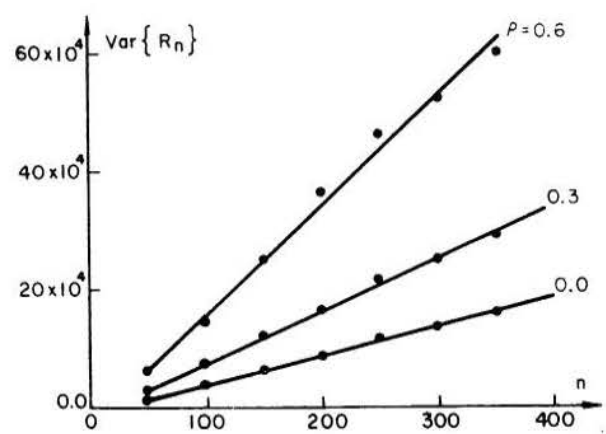

Fig. 6.22 Variance of the range for first-order Markov models with periodic standard deviation. Cases of $\overline{\sigma}_\tau = 40$, $s(\sigma_\tau) =$ 14.22 and $\rho = 0.0, 0.3,$ and 0.6 .

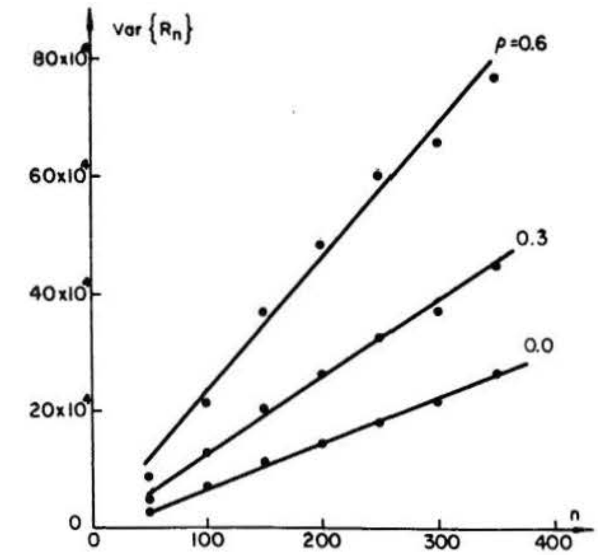

Fig. 6.23 Variance of the range for first-order Markov models with periodic standard deviation. Cases of $\overline{\sigma}_\tau = 40$, $s(\sigma_\tau) = 40$ and $\rho = 0.0, 0.3,$ and 0.6 .

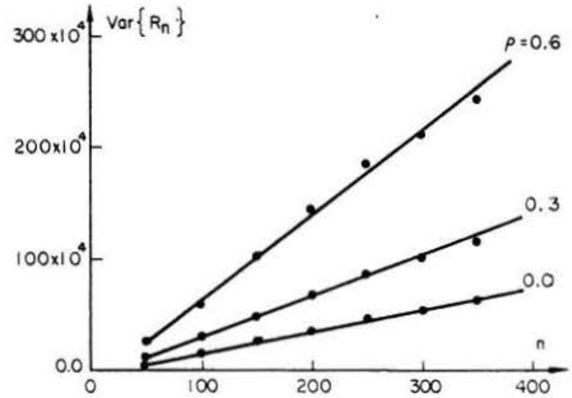

Fig. 6.25 Variance of the range for first-order Markov models with periodic standard deviation. Cases of $\overline{\sigma}_\tau = 80$, $s(\sigma_\tau) = 30.37$ and $\rho = 0.0, 0.3,$ and 0.6 .

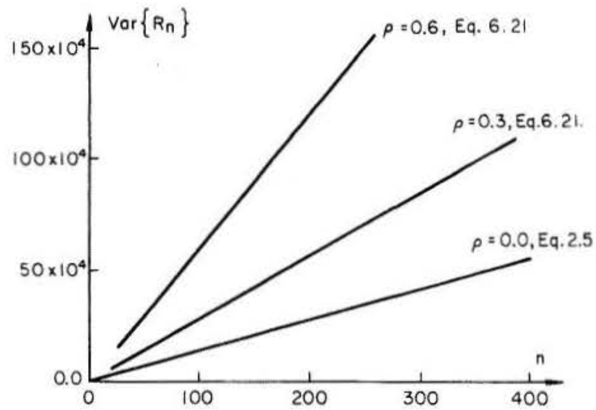

Fig. 6.24 Variance of the range for first-order Markov models with constand standard deviation. Cases of $\sigma_\tau = 80$ and $\rho = 0.0,$ $0.3,$ and 0.6 .

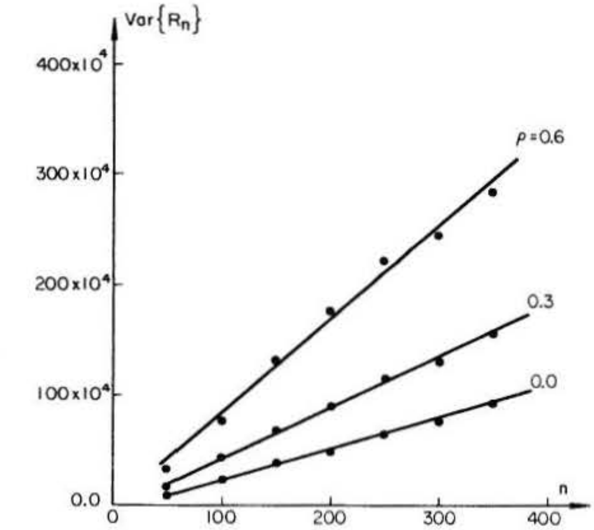

Fig. 6.26 Variance of the range for first-order Markov models with periodic standard deviation. Cases of $\overline{\sigma}_\tau = 80$, $s(\sigma_\tau) = 64.50$ and $\rho = 0.0, 0.3,$ and 0.6 .

CHAPTER VII

DESIGN OF DETERMINISTIC-STOCHASTIC STORAGE CAPACITIES

This chapter deals with determining the storage capacity of a reservoir when the within-the-year inflow fluctuations are considered. The analysis is based on the approximate expected values of the range developed in Chapter V and on some further results described herein. The main assumption is that the inputs are described by a Markov model with periodic mean μ_τ and periodic standard deviation σ_τ as represented by Eq. 6.24, and the output is equal to the mean input $\bar{\mu}_\tau$.

7.1 Deterministic and Stochastic Storage

First, a sensitivity analysis is performed to see the effect of each component μ_τ, σ_τ and ρ on the expected value of the range. The functions μ_τ and σ_τ used here are those previously shown in Fig. 6.9. Figure 7.1 shows the expected range for the following cases:

(1) i.i.d. variables with $\sigma = 5.0$,

(2) independent variables with $\bar{\sigma}_\tau = 5.0$ and $s(\sigma_\tau) = 2.79$,

(3) periodic function μ_τ only, without randomness,

(4) μ_τ with $s(\mu_\tau) = 12.40$, σ_τ with $\bar{\sigma}_\tau = 5.0$ and $s(\sigma_\tau) = 0.0$, and $\rho = 0.0$,

(5) μ_τ with $s(\mu_\tau) = 12.40$, σ_τ with $\bar{\sigma}_\tau = 5.0$ and $s(\sigma_\tau) = 2.79$, and $\rho = 0.0$,

(6) μ_τ with $s(\mu_\tau) = 12.40$, σ_τ with $\bar{\sigma}_\tau = 5.0$ and $s(\sigma_\tau) = 0.0$, and $\rho = 0.60$, and

(7) μ_τ with $s(\mu_\tau) = 12.40$, σ_τ with $\bar{\sigma}_\tau = 5.0$ and $s(\sigma_\tau) = 2.79$, and $\rho = 0.60$.

The results shown in Fig. 7.1 are important, giving a good idea of the influence of each component on the expected range. For the case of i.i.d. random variables with $\sigma = 5.0$, a well-known increasing smooth curve is shown. Then, for periodic σ_τ with $\bar{\sigma}_\tau = 5.0$ and $s(\sigma_\tau) = 2.79$, the expected range is a periodic non-decreasing function of n with a period equal to the period of σ_τ and with decreasing amplitudes as n becomes large. The expected range after the transition region is greater than the expected range of the case of a constant standard deviation. For case (3) the function μ_τ has no random part. The range in this case increases from zero up to a maximum value of 64 at $n = 8$ and remains constant for all greater values of n. Cases (4) and (5) give for $\rho = 0$ the effect of the periodic function σ_τ combined with the function μ_τ. In these cases, the expected range is again greater when σ_τ is periodic than when σ_τ is constant. The same result is given for cases (6) and (7) for $\rho = 0.60$. A general characteristic shown by cases (4) through (7) is that they are all periodic functions with a period equal to one half of the period of μ_τ. This result defers from case (2) in which the period shown by the expected range was the same as that of σ_τ. Figure 7.1 also shows that the effect of dependence, determined in this case by ρ, is considerable.

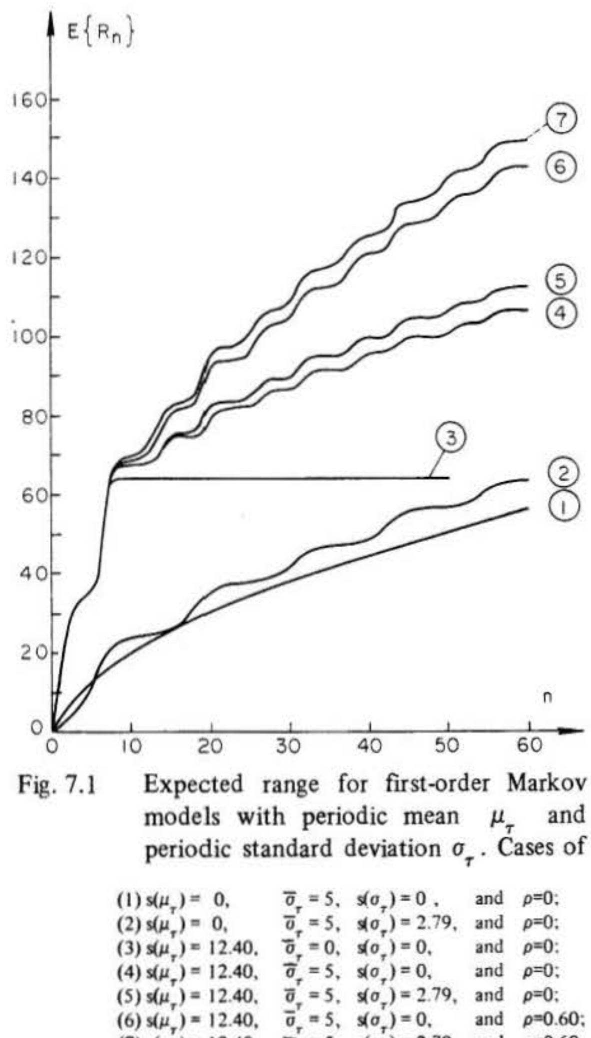

Fig. 7.1 Expected range for first-order Markov models with periodic mean μ_τ and periodic standard deviation σ_τ. Cases of

(1) $s(\mu_\tau) = 0$, $\bar{\sigma}_\tau = 5$, $s(\sigma_\tau) = 0$, and $\rho = 0$;
(2) $s(\mu_\tau) = 0$, $\bar{\sigma}_\tau = 5$, $s(\sigma_\tau) = 2.79$, and $\rho = 0$;
(3) $s(\mu_\tau) = 12.40$, $\bar{\sigma}_\tau = 0$, $s(\sigma_\tau) = 0$, and $\rho = 0$;
(4) $s(\mu_\tau) = 12.40$, $\bar{\sigma}_\tau = 5$, $s(\sigma_\tau) = 0$, and $\rho = 0$;
(5) $s(\mu_\tau) = 12.40$, $\bar{\sigma}_\tau = 5$, $s(\sigma_\tau) = 2.79$, and $\rho = 0$;
(6) $s(\mu_\tau) = 12.40$, $\bar{\sigma}_\tau = 5$, $s(\sigma_\tau) = 0$, and $\rho = 0.60$;
(7) $s(\mu_\tau) = 12.40$, $\bar{\sigma}_\tau = 5$, $s(\sigma_\tau) = 2.79$, and $\rho = 0.60$.

The long term effect of the phases of μ_τ and σ_τ is analyzed with $s(\mu_\tau) = 12.40$ and two phases, and $\bar{\sigma}_\tau = 10$, and with $s(\sigma_\tau) = 6.87$ and three phases and $\rho = 0.60$. As in the case of the variance of the range, five different combinations of symmetric and skewed μ_τ and σ_τ were used as shown previously in Fig. 6.11. The expected ranges obtained for the five cases considered are shown in Fig. 7.2. These results lead to the conclusion that the influence of the phases of μ_τ and σ_τ is significant only in the transition region. Beyond this region, or say for $n > 50$, the expected ranges tend to converge to approximately the same values. Therefore, for all practical purposes, the effects of the phases of μ_τ and σ_τ are neglected, and, subsequently, the analyses are made for symmetric functions of μ_τ and σ_τ only.

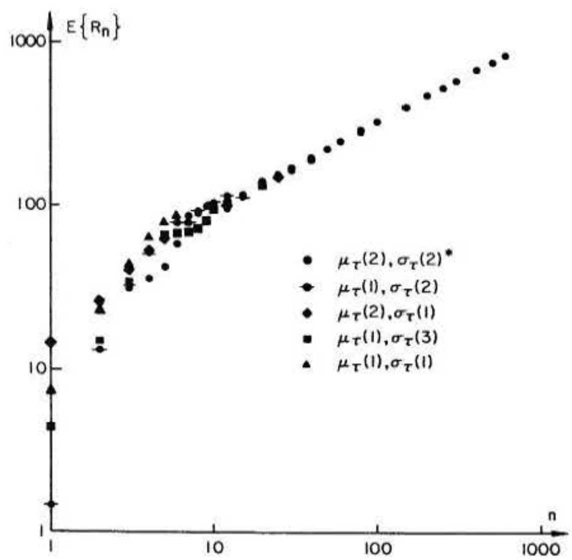

Fig. 7.2 Expected range obtained from simulated samples for first-order Markov models with $s(\mu_\tau) = 12.40$, $\bar{\sigma}_\tau = 10$, $s(\sigma_\tau) = 6.87$ and $\rho = 0.60$ for five different combinations of phases of μ_τ and σ_τ. (*numbers in parenthesis refer to types of μ_τ and σ_τ indicated in Fig. 6.11).

In determining the storage capacity of a reservoir for within-the-year regulation on the mean flow $\bar{\mu}_\tau$, and for inputs of the Markov models type with periodic mean μ_τ and periodic standard deviation σ_τ, the expected storage, given by the expected range of cumulative departures from the mean $\bar{\mu}_\tau$, is divided into two parts: (1) A deterministic storage which is a function of the standard deviation of μ_τ and the mean and standard deviation of σ_τ, and (2) A stochastic storage which is a function of the mean and standard deviation of σ_τ,, the autocorrelation coefficient ρ, and n. That is,

$$S_T(n) = S_d\,[\,s(\mu_\tau),\ \bar{\sigma}_\tau,\ s(\sigma_\tau)\,] + S_s\,[\,\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho,\ n\,]\ , \qquad 7.1$$

where $S_T(n)$ denotes the total storage required for regulation in n units of time, and $S_d(.)$ and $S_s(.)$ denote the deterministic and stochastic storage functions, respectively. Equation 7.1 is represented graphically in Figs. 7.3 and 7.4.

The hypothesis that the deterministic storage $S_d(.)$ depends only on $s(\mu_\tau)$, $\bar{\sigma}_\tau$ and $s(\sigma_\tau)$ was checked by comparing the expected ranges obtained when μ_τ is considered and when it is not – that is, when $s(\mu_\tau) \neq 0$ and $s(\mu_\tau) = 0$. For example, Fig. 7.3 gives the expected range when $\bar{\sigma}_\tau = 10$, $s(\sigma_\tau) = 6.87$, and $\rho = 0$ for both $s(\mu_\tau) = 12.40$ and $s(\mu_\tau) = 0$. The differences between the expected ranges obtained for these two cases vary around a constant value of 41.96 for n values greater than 50. Figure 7.4 also shows the same case as above except that $\rho = 0.60$. The constant value obtained in this last case is 42.03. These results are also given in Table 7.1. This analysis confirms the postulate of an approximately constant deterministic storage independent of ρ and n for given values of $s(\mu_\tau)$, $\bar{\sigma}_\tau$, and $s(\sigma_\tau)$.

The deterministic storage function $S_d[s(\mu_\tau), \bar{\sigma}_\tau, s(\sigma_\tau)]$ is determined for various values of $s(\mu_\tau)$, $\bar{\sigma}_\tau$, and $s(\sigma_\tau)$. The specific functions μ_τ and σ_τ considered here are shown in Figs. 6.13 and 6.14. Figure 7.5, gives the function $S_d[s(\mu_\tau), \bar{\sigma}_\tau, s(\sigma_\tau)]$ for $s(\mu_\tau) = 73.03$, 134.04 and 190.96, for $\bar{\sigma}_\tau = 20$, 40, and 80, and for $s(\sigma_\tau)$ ranging from 0 to 40. This figure shows that a linear function may be fitted between the values of $S_d(.)$ and $s(\sigma_\tau)$ for particular values of $\bar{\sigma}_\tau$ and $s(\mu_\tau)$. It also shows that the effect of $s(\sigma_\tau)$ is very small so that the function $S_d[s(\mu_\tau), \bar{\sigma}_\tau, s(\sigma_\tau)]$ may be further approximated by a function of only two parameters, namely $s(\mu_\tau)$ and $\bar{\sigma}_\tau$. In this case Figs. 7.6 and 7.7 give a relationship between the deterministic storage function $S_d(.)$ against $\bar{\sigma}_\tau$ and $s(\mu_\tau)$, respectively.

The stochastic storage function $S_s[\bar{\sigma}_\tau, s(\sigma_\tau), \rho, n]$ is determined previously in Chapter V as the expected range of Markov models with periodic standard deviation and is given by Eq. 5.18. Therefore, the total storage $S_T(n)$ of Eq. 7.1 may be approximated by

$$S_T(n) \doteq S_d\,[s\,(\mu_\tau), \bar{\sigma}_\tau\,, s\,(\sigma_\tau)\,] + \sqrt{\frac{2}{\pi}}\,\{\hat{\sigma}_n \sum_{i=1}^{n} i^{-1/2} + \bar{\sigma}_\tau \left[\sum_{i=1}^{n} i^{-1}\,(\mathrm{Var}\,S_i)^{1/2} - \sum_{i=1}^{n} i^{-1/2}\right]\} \qquad 7.2$$

where $\hat{\sigma}_n$ is given by Eq. 5.10 and Var S_i by Eq. 3.38.

7.2 Example of the Application of the Proposed Method

Let us assume that a river has a monthly streamflow which may be described by a Markov model with periodic mean μ_τ and periodic standard deviation σ_τ , with the following values:

Periodic mean: $\bar{\mu}_\tau$ = 200 units, $s(\mu_\tau)$ = 150, the periodic standard deviation:

τ:	1	2	3	4	5	6	7	8	9	10	11	12
σ_τ:	4	7	12	20	34	43	43	34	20	12	7	4

with $\bar{\sigma}_\tau$ = 20 and $s(\sigma_\tau)$ = 14.22 , and with the first autocorrelation coefficient ρ = 0.60.

Assume further that one desires to find the storage capacity for regulating the mean flow $\bar{\mu}_\tau$ = 200 units, which on the average will not run dry or overflow in a period of 20 years – that is, n = 240.

The deterministic storage may be found from Figs. 7.5 through 7.7. Assuming the effect of $s(\sigma_\tau)$ is neglected, then Fig. 7.7 gives a value of S_d = 724 units. The stochastic storage is obtained from Eq. 5.18 in which the function $\hat{\sigma}_n$ is computed by Eq. 5.10. This gives a value of S_s = 970 units for the stochastic storage. Therefore, from Eq. 7.1, the total storage is equal to 1694 units. The variance of this storage may be obtained from Fig. 6.20 which for $\bar{\sigma}_\tau$ = 20 , $s(\sigma_\tau)$ = 14.22 , ρ = 0.60 , and n = 240 gives a value of 124,000 or a standard deviation equal to 352.

It should be noted that the proposed method of separating the total storage into a deterministic and a stochastic part may be extended to higher order Markov models. For these models the deterministic storage function $S_d(.)$ remains the same, while the stochastic storage function depends on several more parameters; that is, in general it will be represented by $S_s[\bar{\sigma}_\tau\,, s(\sigma_\tau)\,, \bar{\rho}_k\,, s(\rho_{k,\tau})]$, with k = 1,2,...,m and m the order of the Markov model considered.

TABLE 7.1 COMPARISON OF THE EXPECTED RANGES FOR MARKOV MODELS WITH ZERO AND PERIODIC MEAN μ_τ .

n	(1) $s(\mu_\tau)$=12.40, $\bar{\sigma}_\tau$=10, $s(\sigma_\tau)$=6.87, ρ=0. (2) $s(\mu_\tau)$=0.0, $\bar{\sigma}_\tau$=10, $s(\sigma_\tau)$=6.87, ρ=0.			(1) $s(\mu_\tau)$=12.40, $\bar{\sigma}_\tau$=10, $s(\sigma_\tau)$=6.87, ρ=0.6 (2) $s(\mu_\tau)$=0.0, $\bar{\sigma}_\tau$=10, $s(\sigma_\tau)$=6.87, ρ=0.6		
	Expected Range		(1)-(2)	Expected Range		(1)-(2)
	(1)	(2)		(1)	(2)	
50	157.47	117.12	40.35	225.83	185.96	39.87
100	214.19	171.58	42.61	328.75	285.50	43.25
150	258.58	219.12	39.46	407.54	367.30	40.24
200	301.25	259.87	41.38	484.38	443.16	41.22
250	331.09	290.22	40.87	539.87	498.82	41.05
300	359.76	317.66	42.10	593.00	549.67	43.33
350	389.83	346.75	43.08	643.22	599.70	43.52
400	417.51	373.56	43.95	693.31	649.43	43.88
450	443.45	400.65	42.80	737.20	695.20	42.00
500	465.49	422.58	42.91	780.62	738.02	42.60
550	487.86	445.58	42.28	817.45	775.89	41.54
600	508.38	466.68	41.70	854.71	812.88	41.83
	Average difference = 41.96			Average difference = 42.03		

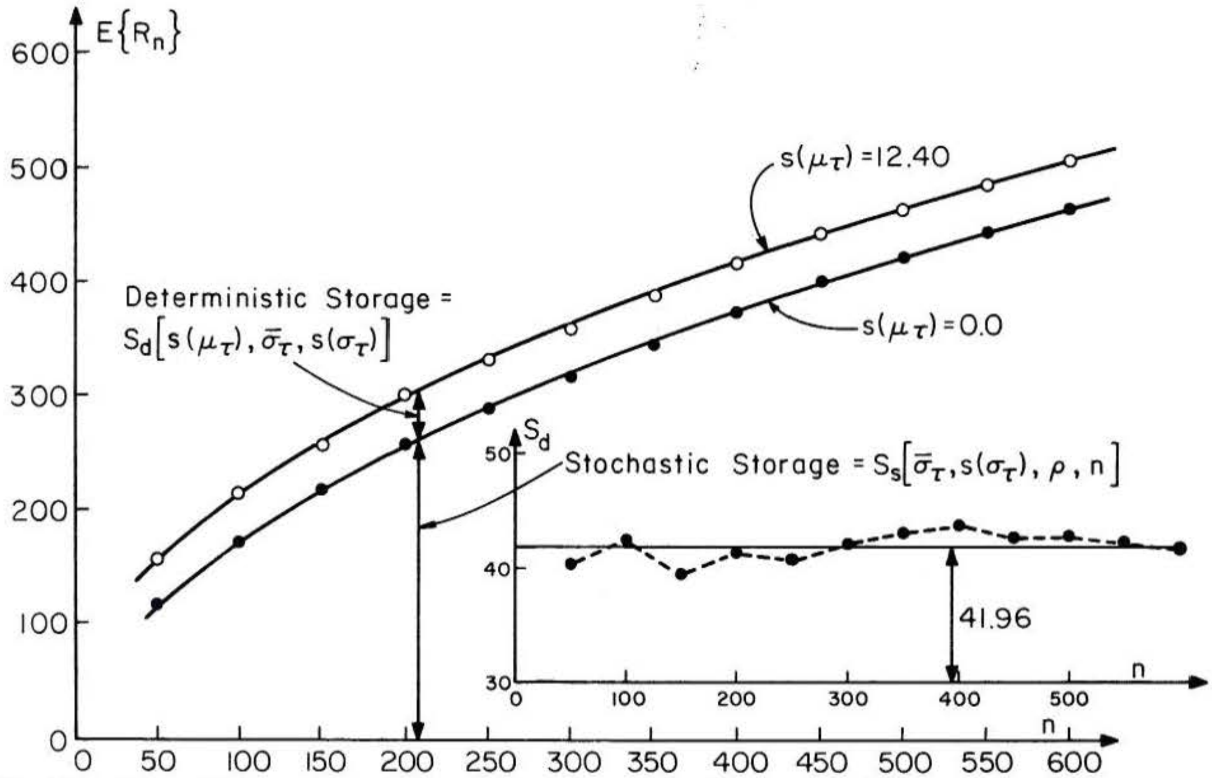

Fig. 7.3 Deterministic and stochastic required storage capacities in case of inputs with periodic mean μ_τ and periodic standard deviation σ_τ with $\bar{\sigma}_\tau = 10$, $s(\sigma_\tau) = 6.87$ and $\rho = 0$.

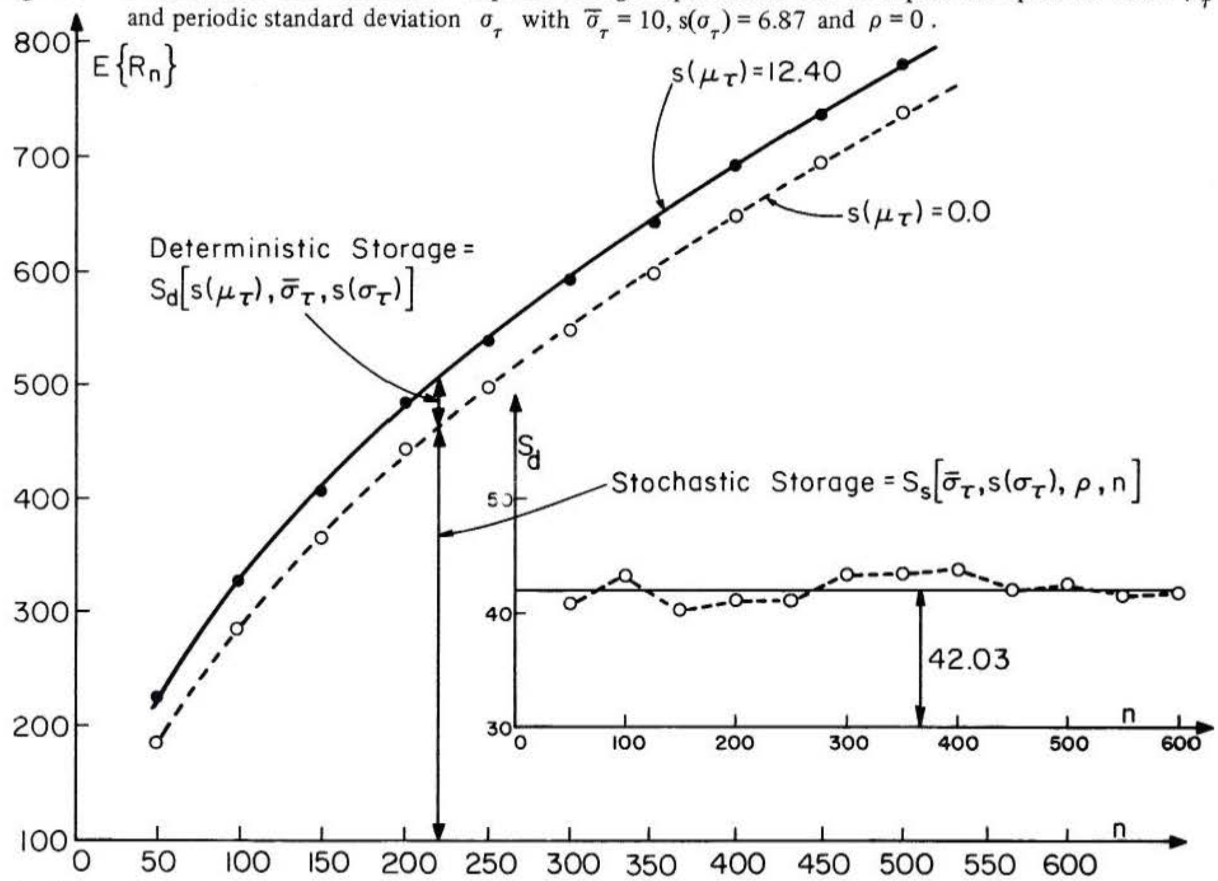

Fig. 7.4 Deterministic and stochastic required storage capacities in case of inputs with periodic mean μ_τ and periodic standard deviation σ_τ with $\bar{\sigma}_\tau = 10$, $s(\sigma_\tau) = 6.87$ and $\rho = 0.60$.

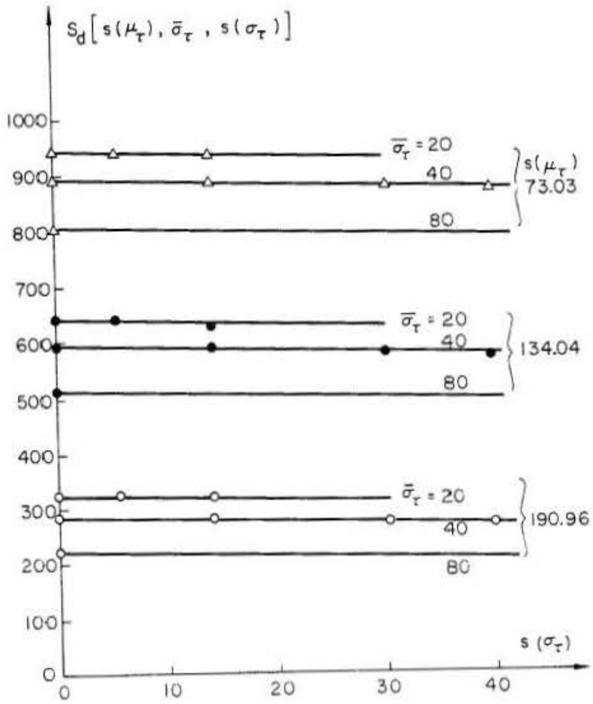

Fig. 7.5 Variation of deterministic storage for various values of $s(\mu_\tau), \bar{\sigma}_\tau$ and $s(\sigma_\tau)$.

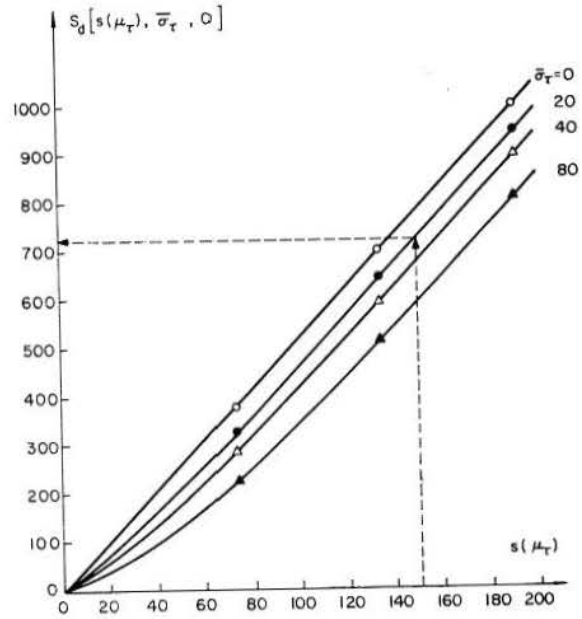

Fig. 7.7 Deterministic storage for the case of $s(\sigma_\tau) = 0$ and various values of $s(\mu_\tau)$ and $\bar{\sigma}_\tau$.

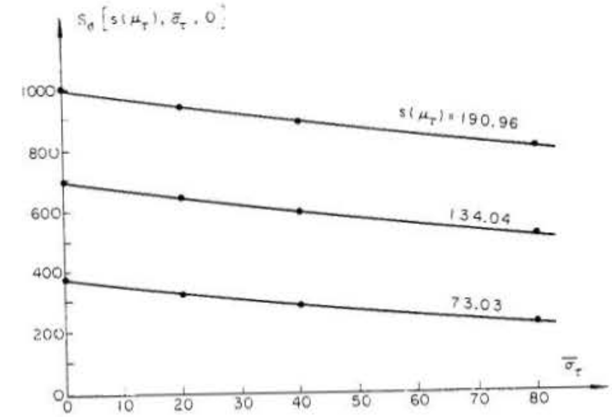

Fig. 7.6 Deterministic storage for the case of $s(\sigma_\tau) = 0$ and various values of $s(\mu_\tau)$ and $\bar{\sigma}_\tau$.

CHAPTER VIII

CONCLUSIONS

The analysis of storage problem considering the within-the-year fluctuations of inflows was the main objective of this study; therefore, mathematical models of monthly values of streamflow were used as examples. The storage difference equation which relates the inputs, outputs, and storage was used for formulating the mathematical problem. This led to the problem of determining the expected values and variances of the range of cumulative departures from the mean.

The main conclusions drawn from this investigation are as follows: (1) Considering that the sequence of partial sums $S_o, S_1, S_2, \ldots, S_n$ follows the general multivariate normal distribution function, the exact expression of the expected value of the surplus $M_n = \max(S_o, S_1, S_2, \ldots, S_n)$ becomes very complex to derive when n is large. For small values of n, namely for $n = 1,2$, and 3, the expected value of the surplus M_n and consequently the expected value of the range R_n were derived in this study.

(2) The derived general expression of the expected value of the range for $n = 1,2$, and 3 permits obtaining the exact expected ranges of stationary and non-stationary inputs. The following cases were derived:

a. Independent random variables with changing standard deviation;
b. Equally dependent random variables, and
c. Markov dependent random variables.

(3) The exact expected values of the range, obtained mathematically, for small values of n such as 1,2, and 3, and the computer simulation approach for larger values of n, can be used to determine the degree of accuracy of approximate equations of the expected range. In this study, approximate equations were obtained for the following cases:

a. General Markov model with constant variance and periodic autoregression coefficients,
b. Non-stationary exchangeable random variables, and
c. Markov dependent random variables with periodic standard deviation and constant autoregression coefficients.

(4) The expected values of the adjusted range of exchangeable random inputs, and outputs equal to a percentage of the mean inflow, may be expressed in the same way as the formula 4.107, valid for the expected range of exchangeable random variables. This result is relevant in hydrology in cases of over-year storage design.

(5) The exact variance of the range was possible to derive for $n = 1$ and 2 for the case of stationary first-order Markov model. The mathematical derivation becomes complex for larger values of n.

(6) Empirical equations, derived by the computer simulation approach, can be used for approximating the variances of the range. In particular, in this study, empirical equations were derived for the variance of the range of the first and second-order Markov models with constant autoregression coefficients. Some empirical curves are also given for cases of non-stationary Markov models.

(7) The total storage capacity required for regulating the mean inflow, when the within-the-year fluctuation of the inflows is taken into account, can be divided into two parts:

a. A deterministic storage which is a function of the standard deviation of μ_τ and the mean and standard deviation of σ_τ. (For these three parameters it is shown that the deterministic storage is practically constant for all n greater than 50.)

b. A stochastic storage which is a function of the mean and standard deviation of σ_τ, of the autocorrelation coefficients of the Markov model considered, and of n.

REFERENCES

Anis, A. A., and Lloyd, E. H., 1953, On the range of partial sums of a finite number of independent random variables: Biometrika, v. 40, p. 35-42.

Anis, A. A., 1955, The variance of the maximum of partial sums of a finite number of independent normal variates: Biometrika, v. 42, p. 96-101.

Anis, A. A., 1956, On the moments of the maximum of partial sums of a finite number of independent normal variates: Biometrika, v. 43, p. 79-84.

Ditlevsen, O., 1969, Approximative extreme value distributions and first passage time probabilities in stationary and related stochastic processes: Personal communication, Denmarks Ingeniorakademi.

Feller, W., 1951, The asymptotic distribution of the range of sums of independent variables: Ann. Math. Statistics, v. 22, p. 427- 432.

Fiering, M. B., 1967, Streamflow Synthesis: Harvard University Press, Cambridge, Massachusetts.

Ghosal, A., 1960, Emptiness in the finite dam: Ann. Math. Statistics, v. 31, p. 803-808.

Hazen, A., 1914, Storage to be provided in impounding reservoirs for municipal water supply: Am. Soc. Civil Engineers Trans., v. 77, p. 1539-1640.

Hurst, H. E., 1951, Long term storage capacities of reservoirs: Am. Soc. Civil Engineers, Trans., v. 116, p. 776-808.

Jeng, R., 1967, Time – dependent solutions for water storage problem: Ph.D. Dissertation, Colorado State University, Fort Collins, Colorado.

Langbein, W. B., 1958, Queuing theory and water storage: Am. Soc. Civil Engineers Proc., v. 83, No. HY5, paper 1811.

Lloyd, E. H., 1963, Reservoirs with serially correlated inflows: Technometrics, v. 5, p. 85-93.

Lloyd, E. H., 1963, A probability theory of reservoirs with serial correlated inputs: Journal of Hydrology, v. 1p. 99-128.

Lloyd, E. H., and Odoom, S., 1964, Probability theory of reservoirs with seasonal inputs: Journal of Hydrology, v. 2, p. 1-10.

Lloyd, E. H., 1967, Stochastic reservoir theory, *in* Advances of Hydroscience: Edited by V. T. Chow, Academic Press, v. 4, p. 281- 339.

Loève, M., 1960, Probability Theory: Second Edition Van Nostrand, Princeton.

Mejia, J. M., 1971, On the generation of multivariate sequences exhibiting the Hurst Phenomenon and some flood frequency analyses: Ph.D. Disseration, Colorado State University, Fort Collins, Colorado.

Melentijevich, M. J., 1965, Characteristics of storage when outflow is dependent upon reservoir volume: Ph.D. Dissertation, Colorado State University, Fort Collins, Colorado.

Mood, A. M., and Graybill, F. A., 1963, Introduction to the Theory of Statistics: McGraw-Hill Series in probability and statistics, Second Edition, McGraw-Hill Book Co., Inc.

Moran, P. A. P., 1954, A probability theory of dams and storage systems: Aust. Jour. Appl. Sci., v. 5, p. 116-124.

Owen, D. B., and Steck, G. P., 1962, Moments of order statistics from the equicorrelated multivariate normal distribution: Ann. Math. Statistics, v. 33, p. 1286-1291.

Prabu, N. U., 1965, Queues and Inventories: Wiley series in probability and mathematical statistics, John Wiley and Sons, Inc., New York.

Prabu, N. U., 1958, Some exact results for the finite dam: Ann. Math. Statistics, v. 29, p. 1234-1243.

Rippl, W., 1883, The capacity of storage reservoirs for water supply: Institution of Civil Engineers, Proc., v. 71, p. 270-278.

Roesner, L. A., and Yevjevich, V. M., 1966, Mathematical models for time series of monthly precipitation and monthly runoff: Hydrology paper No. 15, Colorado State University, Fort Collins, Colorado.

Salas-La Cruz, J. D., and Yevjevich, V. M., 1972, Stochastic structure of water use time series: Forthcoming Hydrology paper, Colorado State University, Fort Collins, Colorado.

Solari, M. E., and Anis, A. A., 1957, The mean and variance of the maximum of the adjusted partial sums of a finite number of independent normal variates: Ann. Math. Statistics, v. 28, p. 706-716.

Spitzer, F., 1956, A combinatorial lemma and its application ot probability theory: Am. Math. Society Trans., v. 82, p. 323-339.

Sudler, C. E., 1927, Storage required for the regulation of streamflow: Am. Soc. Civil Engineers Trans., No. 91, p. 622-704.

Sutabutra, P., 1967, Reservoir storage capacity required when water inflow has a periodic and a stochastic component: Ph.D. Disseration, Colorado State University, Fort Collins, Colorado.

Thomas, H. A., and Fiering, M. B., 1962, Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, *in* Design of Water Resources Systems: Harvard University Press, Cambridge, Massachusetts.

Yevjevich, V. M., 1964, Fluctuation of wet and dry years. Part II. Analysis by serial correlation: Hydrology paper No. 4, Colorado State University, Fort Collins, Colorado.

Yevjevich, V. M., 1965, The application of surplus, deficit and range in hydrology: Hydrology paper No. 10, Colorado State University, Fort Collins, Colorado.

Yevjevich, V. M., 1967, Mean range of linearly dependent normal variables with application to storage problem: Water Resources Research v. 3, No. 3, p. 663-671.

Yevjevich, V. M., 1971, The structure of inputs and outputs of hydrologic systems: United States-Japan Bi-lateral Seminar in Hydrology, Honolulu.

Yevjevich, V. M., 1972, Stochastic Process in Hydrology: Water Resources Publications, Fort Collins, Colorado.

APPENDIX

EVALUATION OF CONSTANTS TO BE USED IN EXPRESSIONS E{X}, E{Y} AND E{Z} OF CHAPTER IV

Let us recall that the maximum M_3 was defined as $M_3 = \max(0,X,Y,Z)$, where

$$S_1 = X = (x_1 - y_1)$$

$$S_2 = Y = (x_1 - y_1) + (x_2 - x_2)$$

$$S_3 = Z = (x_1 - y_1) + (x_2 - y_2) + (x_3 - y_3) ,$$

and let us assume that the departures or components of partial sums $(x_i - y_i)$ are normally distributed with mean zero, changing variance and are linearly dependent.

Therefore the variances of X, Y and Z are given in general as

$$\mathrm{Var}\{X\} = \sigma_x^2 = \sigma_1^2 , \tag{1}$$

$$\mathrm{Var}\{Y\} = \sigma_y^2 = \sigma_1^2 + \sigma_2^2 + 2\,\sigma_1\sigma_2\rho_{12} , \tag{2}$$

$$\mathrm{Var}\{Z\} = \sigma_z = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + 2\,\sigma_1\sigma_2\rho_{12} + 2\,\sigma_1\sigma_3\rho_{13} + 2\,\sigma_2\sigma_3\rho_{23} . \tag{3}$$

The covariances of X and Y, X and Z, and Y and Z may be shown to be

$$\mathrm{Cov}\{X,Y\} = \sigma_1^2 + \sigma_1\sigma_2\rho_{12} \tag{4}$$

$$\mathrm{Cov}\{X,Z\} = \sigma_1^2 + \sigma_1\sigma_2\rho_{12} + \sigma_1\sigma_3\rho_{13} , \tag{5}$$

$$\mathrm{Cov}\{Y,Z\} = \sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho_{12} + \sigma_1\sigma_3\rho_{13} + \sigma_2\sigma_3\rho_{23} . \tag{6}$$

where σ_1, σ_2 and σ_3 denote the standard deviation of the departures $(x_1 - y_1)$, $(x_2 - y_2)$ and $(x_3 - y_3)$ respectively and ρ_{12}, ρ_{13} and ρ_{23} are the correlation coefficients between the indicated components.

A. FOR INDEPENDENT COMPONENTS. In this case $\rho_{12} = \rho_{13} = \rho_{23} = 0$, therefore Eqs. (1) to (6) simplify to

$$\mathrm{Var}\{X\} = \sigma_x^2 = \sigma_1^2 , \tag{7}$$

$$\mathrm{Var}\{Y\} = \sigma_y^2 = \sigma_1^2 + \sigma_2^2 , \tag{8}$$

$$\mathrm{Var}\{Z\} = \sigma_z^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 , \tag{9}$$

$$\mathrm{Cov}\{X,Y\} = \sigma_1^2 , \tag{10}$$

$$\mathrm{Cov}\{X,Z\} = \sigma_1^2 , \tag{11}$$

$$\mathrm{Cov}\{Y,Z\} = \sigma_1^2 + \sigma_2^2 . \tag{12}$$

From the above equations, the correlation coefficients ρ_{xy}, ρ_{xz} and ρ_{yz} are given by

$$\rho_{xy} = \frac{\sigma_1}{(\sigma_1^2 + \sigma_2^2)^{1/2}} ,$$

$$\rho_{xz} = \frac{\sigma_1}{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} , \tag{13}$$

$$\text{and } \rho_{yz} = \frac{(\sigma_1^2 + \sigma_2^2)^{1/2}}{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} .$$

Using the Eqs. 4.17, 4.21 and 4.24, the conditional standard deviations are

$$\sigma_{x.y} = \frac{\sigma_1\sigma_2}{(\sigma_1^2 + \sigma_2^2)^{1/2}} ,$$

$$\sigma_{x.z} = \frac{\sigma_1(\sigma_2^2 + \sigma_3^2)^{1/2}}{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} , \tag{14}$$

$$\sigma_{y.x} = \sigma_2 \ ,$$

$$\sigma_{y.z} = \frac{\sigma_3(\sigma_1{}^2 + \sigma_2{}^2)^{1/2}}{(\sigma_1{}^2 + \sigma_2{}^2 + \sigma_3{}^2)^{1/2}} \ , \qquad (15)$$

$$\sigma_{z.x} = (\sigma_2{}^2 + \sigma_3{}^2)^{1/2} \ , \qquad \sigma_{z.y} = \sigma_3 \ . \qquad (16)$$

Applying Eq. 4.8 to the trivariate case, the partial correlation coefficients $\rho_{xy.z}$, $\rho_{xz.y}$, and $\rho_{yz.x}$ are

$$\rho_{xy.z} = \frac{\sigma_1\sigma_3}{(\sigma_1{}^2 + \sigma_2{}^2)^{1/2}(\sigma_2{}^2 + \sigma_3{}^2)^{1/2}} \ ,$$

$$\rho_{xz.y} = 0 \ , \text{ and } \rho_{yz.x} = \frac{\sigma_2}{(\sigma_2{}^2 + \sigma_3{}^2)^{1/2}} \qquad (17)$$

Substitution of above equations into Eqs. 4.49, 4.52, 4.55, 4.67, 4.68, 4.71, 4.78, 4.79, and 4.82 leads to the following constants:

$$b_1 = -c_1 = 0 \ , \qquad b_2 = -c_2 = 0 \ , \qquad (18)$$

$$k_1 = \frac{\sigma_3^2}{(\sigma_2^2 + \sigma_3^2)} \ , \qquad k_2 = \infty \ , \qquad (19)$$

$$k_3 = \frac{(\sigma_2^2 + \sigma_3^2)}{\sigma_3^2} \ , \qquad k_4 = \infty \ , \qquad (20)$$

$$b_1' = c_1' = \frac{\sigma_2}{\sigma_1(\sigma_1{}^2 + \sigma_2{}^2)^{1/2}} \ ,$$

$$b_2' = -c_2' = 0 \ , \qquad (21)$$

$$k_1' = 1 \ , \qquad k_2' = \infty \ , \qquad (22)$$

$$k_3' = 1 \ , \qquad k_4' = \frac{(\sigma_1^2 + \sigma_2^2)}{\sigma_2^2} \ , \qquad (23)$$

$$b_1'' = c_1'' = \frac{(\sigma_2^2 + \sigma_3^2)^{1/2}}{\sigma_1(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \ ,$$

$$b_2'' = c_2'' = \frac{\sigma_3}{(\sigma_1^2 + \sigma_2^2)^{1/2}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^{1/2}} \qquad (24$$

$$k_1'' = \frac{\sigma_2^2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)}{(\sigma_1^2 + \sigma_2^2)(\sigma_2^2 + \sigma_3^2)} \ ,$$

$$k_2'' = \frac{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)}{\sigma_3^2} \ , \qquad (25)$$

$$k_3'' = \frac{(\sigma_1^2 + \sigma_2^2)(\sigma_2^2 + \sigma_3^2)(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)}{[\sigma_1^2\sigma_2^2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) + (\sigma_1^2 + \sigma_2^2)(\sigma_2^2 + \sigma_3^2)]}$$

$$, \ k_4'' = \frac{[\sigma_1^2\sigma_2^2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) + (\sigma_1^2 + \sigma_2^2)(\sigma_2^2 + \sigma_3^2)]}{\sigma_2^2(\sigma_2^2 + \sigma_3^2)(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)} \qquad (26)$$

B. FOR COMPONENTS WITH EQUAL VARIANCE AND EQUAL DEPENDENCE (exchangeable random variables). In this case,

$$\sigma_1 = \sigma_2 = \sigma_3 = \sigma \ , \qquad (27)$$

and

$$\rho_{12} = \rho_{13} = \rho_{23} = \rho \ . \qquad (28)$$

Therefore Eqs. (1) to (6) simplify to

$$\mathrm{Var}\{X\} = \sigma_x^2 = \sigma^2 \ , \qquad (29)$$

$$\mathrm{Var}\{Y\} = \sigma_y^2 = 2\sigma^2(1+\rho) \ , \qquad (30)$$

$$\mathrm{Var}\{Z\} = \sigma_z^2 = 3\sigma^2(1+2\rho) \ , \qquad (31)$$

$$\mathrm{Cov}\{X,Y\} = \sigma^2(1+\rho) \ , \qquad (32)$$

$$\mathrm{Cov}\{X,Z\} = \sigma^2(1+2\rho) \ , \qquad (33)$$

and

$$\text{Cov}\{Y,Z\} = 2\sigma^2 (1 + 2\rho) . \tag{34}$$

From these equations, the correlation coefficients ρ_{xy}, ρ_{xz}, and ρ_{yz} are

$$\rho_{xy} = \frac{(1+\rho)^{1/2}}{\sqrt{2}} \quad , \quad \rho_{xz} = \frac{(1+2\rho)^{1/2}}{\sqrt{3}}$$

$$\text{, and} \quad \rho_{yz} = \frac{\sqrt{2}(1+2\rho)^{1/2}}{\sqrt{3}(1+\rho)^{1/2}} . \tag{35}$$

Using Eqs. 4.17, 4.21, and 4.24, the conditional standard deviations are

$$\sigma_{x.y} = \frac{\sigma}{\sqrt{2}} (1-\rho)^{1/2} \quad ,$$

$$\sigma_{x.z} = \frac{\sqrt{2}}{\sqrt{3}} \sigma (1-\rho)^{1/2} , \tag{36}$$

$$\sigma_{y.x} = \sigma (1-\rho^2)^{1/2} \quad ,$$

$$\sigma_{y.z} = \frac{\sqrt{2}}{\sqrt{3}} \sigma (1-\rho)^{1/2} , \tag{37}$$

$$\sigma_{z.x} = \sqrt{2}\, \sigma(1-\rho)^{1/2} (1+2\rho)^{1/2} \quad ,$$

$$\text{and} \quad \sigma_{z.y} = \frac{\sigma(1-\rho)^{1/2}(1+2\rho)^{1/2}}{(1+\rho)^{1/2}} . \tag{38}$$

Applying Eq. 4.8 to the trivariate case, the partial correlation coefficients $\rho_{xy.z}$, $\rho_{xz.y}$, and $\rho_{yz.x}$ are

$$\rho_{xy.z} = \frac{1}{2} \quad , \quad \rho_{xz.y} = 0 ,$$

$$\text{and} \quad \rho_{yz.x} = \frac{(1+2\rho)^{1/2}}{\sqrt{2}(1+\rho)^{1/2}} . \tag{39}$$

Substitution of the above equations into Eqs. 4.49, 4.52, 4.55, 4.67, 4.68, 4.71, 4.78, 4.79, and 4.82 leads to the following constants:

$$b_1 = -c_1 = \frac{\rho}{\sigma(1-\rho^2)^{1/2}}$$

$$b_2 = -c_2 = \frac{\sqrt{2}\rho}{\sigma(1-\rho)^{1/2}(1+2\rho)^{1/2}} \quad , \tag{40}$$

$$k_1 = \frac{1}{2(1+\rho)} \quad , \quad k_2 = \frac{(1+\rho)}{2\rho^2} \quad , \tag{41}$$

$$k_3 = \frac{2}{(1-\rho+2\rho^2)} \quad , \quad k_4 = \frac{(1+\rho)(1-\rho+2\rho^2)}{\rho^2} \tag{42}$$

$$b_1' = c_1' = \frac{1}{\sqrt{2}\,\sigma(1-\rho)^{1/2}} \quad ,$$

$$b_2' = -c_2' = \frac{\rho}{\sigma(1-\rho^2)^{1/2}(1+2\rho)^{1/2}} \quad , \tag{43}$$

$$k_1' = 1 \quad , \quad k_2' = \frac{(1+\rho)}{2\rho^2} \quad , \tag{44}$$

$$k_3' = 1 \quad , \quad k_4' = \frac{2}{(1+\rho)} \quad , \tag{45}$$

$$b_1'' = c_1'' = \frac{\sqrt{2}}{\sqrt{3}\,\sigma(1-\rho)^{1/2}} \quad ,$$

$$b_2'' = c_2'' = \frac{1}{\sqrt{2}\sqrt{3}\,\sigma(1-\rho)^{1/2}} \tag{46}$$

$$k_1'' = \frac{3}{4} \quad , \quad k_2'' = \frac{3}{(1+2\rho)} \quad , \tag{47}$$

$$k_3'' = \frac{12(1+\rho)}{(11+13\rho)} \quad , \quad k_4'' = \frac{(11+13\rho)}{6(1+2\rho)} . \tag{48}$$

C. FOR COMPONENTS WITH EQUAL VARIANCE AND MARKOV DEPENDENCE. In this case

$$\sigma_1 = \sigma_2 = \sigma_3 = \sigma ,$$

$$\rho_{12} = \rho_{23} = \rho , \text{ and } \rho_{13} = \rho^2 \quad ,$$

therefore the equations (1) to (6) simplify to

$$\mathrm{Var}\,\{X\} = \sigma_x^2 = \sigma^2 \ , \tag{49}$$

$$\mathrm{Var}\,\{Y\} = \sigma_y^2 = 2\sigma^2(1+\rho) \ , \tag{50}$$

$$\mathrm{Var}\,\{Z\} = \sigma_z^2 = \sigma^2\,(3+4\rho+2\rho^2) \ , \tag{51}$$

$$\mathrm{Cov}\,\{X,Y\} = \sigma^2\,(1+\rho) \ , \tag{52}$$

$$\mathrm{Cov}\,\{X,Z\} = \sigma^2\,(1+\rho+\rho^2), \text{ and} \tag{53}$$

and

$$\mathrm{Cov}\,\{Y,Z\} = \sigma^2\,(2+3\rho+\rho^2) = \sigma^2\,(1+\rho)(2+\rho) \ . \tag{54}$$

From these equations, the correlation coefficients ρ_{xy}, ρ_{xz} and ρ_{yz} are

$$\rho_{xy} = \frac{(1+\rho)^{1/2}}{\sqrt{2}} \ ,$$

$$\rho_{xz} = \frac{(1+\rho+\rho^2)}{(3+4\rho+2\rho^2)^{1/2}} \ ,$$

$$\text{and} \quad \rho_{yz} = \frac{(1+\rho)^{1/2}\,(2+\rho)}{\sqrt{2}\,(3+4\rho+2\rho^2)^{1/2}} \ . \tag{55}$$

Using Eqs. 4.17, 4.21 and 4.24, the conditional standard deviations are

$$\sigma_{x.y} = \frac{\sigma(1-\rho)^{1/2}}{\sqrt{2}} \ ,$$

$$\sigma_{x.z} = \frac{\sigma(1-\rho^2)^{1/2}(2+2\rho+\rho^2)^{1/2}}{(3+4\rho+2\rho^2)^{1/2}} \ , \tag{56}$$

$$\sigma_{y.x} = \sigma(1-\rho^2)^{1/2} \ ,$$

$$\sigma_{y.z} = \frac{\sigma(1-\rho^2)^{1/2}(2+2\rho+\rho^2)^{1/2}}{(3+4\rho+2\rho^2)^{1/2}} \ , \tag{57}$$

$$\sigma_{z.x} = \sigma(1-\rho^2)^{1/2}\,(2+2\rho+\rho^2)^{1/2} \ ,$$

$$\sigma_{z.y} = \frac{\sigma(1-\rho)^{1/2}\,(2+2\rho+\rho^2)^{1/2}}{\sqrt{2}} \ . \tag{58}$$

Applying Eq. 4.8 to the trivariate case, the partial correlation coefficients $\rho_{xy.z}$, $\rho_{xz.y}$ and $\rho_{yz.x}$ are

$$\rho_{xy.z} = \frac{(1+\rho)^2}{(2+2\rho+\rho^2)} \ ,$$

$$\rho_{xz.y} = -\frac{\rho}{(2+2\rho+\rho^2)^{1/2}} \ ,$$

$$\text{and} \quad \rho_{yz.x} = \frac{(1+\rho)}{(2+2\rho+\rho^2)^{1/2}} \ . \tag{59}$$

Substitution of the above equations into Eqs. 4.49, 4.52, 4.55, 4.67, 4.68, 4.71, 4.78, 4.79, and 4.82, leads to the following constants:

$$b_1 = -c_1 = \frac{\rho}{\sigma(1-\rho^2)^{1/2}} \ ,$$

$$b_2 = -c_2 = \frac{\rho(1+\rho)^{1/2}}{\sigma(1-\rho)^{1/2}(2+2\rho+\rho^2)^{1/2}} \ , \tag{60}$$

$$k_1 = \frac{1}{(2+2\rho+\rho^2)} \ ,$$

$$k_2 = \frac{2}{\rho^2(1+\rho)} \ , \tag{61}$$

$$k_3 = \frac{(1-\rho^2)(2+2\rho+\rho^2)}{(1+2\rho^3-2\rho^5-\rho^6)} \ ,$$

$$k_4 = \frac{(1+2\rho^3-2\rho^5-\rho^6)}{\rho^2(1-\rho^2)} \ , \tag{62}$$

$$b_1' = c_1' = \frac{1}{\sqrt{2}\,\sigma(1-\rho)^{1/2}} \ ,$$

$$b_2' = -c_2' = \frac{\rho}{\sigma\sqrt{2}(1-\rho)^{1/2}(2+2\rho+\rho^2)^{1/2}} \ , \tag{63}$$

$$k_1' = \frac{2(1+\rho)}{(2+2\rho+\rho^2)} \qquad k_2' = \frac{2}{\rho^2(1+\rho)} \ , \tag{64}$$

$$k'_3 = \frac{2(2+2\rho+\rho^2)}{(1+\rho)(4+\rho^2)} \quad ,$$

$$k'_4 = \frac{(4+\rho^2)}{2(1+\rho)} \quad , \qquad (65)$$

$$b''_1 = c''_1 = \frac{(1+\rho)^{1/2}(2+\rho)}{\sigma(1-\rho)^{1/2}(2+2\rho+\rho^2)^{1/2}(3+4\rho+2\rho^2)^{1/2}}$$

$$b''_2 = c''_2 = \frac{(1+\rho+\rho^2)}{\sigma(1-\rho^2)^{1/2}(2+2\rho+\rho^2)^{1/2}(3+4\rho+2\rho^2)^{1/2}} \quad , \qquad (66)$$

$$k''_1 = \frac{(3+4\rho+2\rho^2)}{(2+2\rho+\rho^2)^2} \quad ,$$

$$k''_2 = \frac{(3+4\rho+2\rho^2)}{(1+\rho+\rho^2)^2} \quad , \qquad (67)$$

$$k''_3 = \frac{2(3+4\rho+2\rho^2)(2+2\rho+\rho^2)}{(11+25\rho+28\rho^2+18\rho^3+7\rho^4+\rho^5)}$$

$$k''_4 = \frac{(2+2\rho+\rho^2)(11+25\rho+28\rho^2+18\rho^3+7\rho^4+\rho^5)}{(1+\rho)(2+\rho)^2(3+4\rho+2\rho^2)} \qquad (68)$$

KEY WORDS: Range of the cumulative sums, S_t , reservoir design, water storage problems, adjusted range of the cumulative sums, S_t .

ABSTRACT: The storage problem of within-the-year water fluctuations is the main topic of this paper. The storage difference equation which relates inputs, outputs and storage is used for formulating the mathematical problem. This leads to the problem of determining the expected values and variances of the range or adjusted range of cumulative departures from the population and sample mean, respectively.
Using the univariate, bivariate and trivariate normal distribution functions for the marginal and joint distributions of the partial sums, the exact expressions of the expected range are derived for n=1,2, and 3 . From these general expressions, particular cases of the expected range of independent and linearly dependent variables are derived.
The expected value of the adjusted range of inputs equally

KEY WORDS: Range of the cumulative sums, S_t , reservoir design, water storage problems, adjusted range of the cumulative sums, S_t .

ABSTRACT: The storage problem of within-the-year water fluctuations is the main topic of this paper. The storage difference equation which relates inputs, outputs and storage is used for formulating the mathematical problem. This leads to the problem of determining the expected values and variances of the range or adjusted range of cumulative departures from the population and sample mean, respectively.
Using the univariate, bivariate and trivariate normal distribution functions for the marginal and joint distributions of the partial sums, the exact expressions of the expected range are derived for n=1,2, and 3 . From these general expressions, particular cases of the expected range of independent and linearly dependent variables are derived.
The expected value of the adjusted range of inputs equally

KEY WORDS: Range of the cumulative sums, S_t , reservoir design, water storage problems, adjusted range of the cumulative sums, S_t .

ABSTRACT: The storage problem of within-the-year water fluctuations is the main topic of this paper. The storage difference equation which relates inputs, outputs and storage is used for formulating the mathematical problem. This leads to the problem of determining the expected values and variances of the range or adjusted range of cumulative departures from the population and sample mean, respectively.
Using the univariate, bivariate and trivariate normal distribution functions for the marginal and joint distributions of the partial sums, the exact expressions of the expected range are derived for n=1,2, and 3 . From these general expressions, particular cases of the expected range of independent and linearly dependent variables are derived.
The expected value of the adjusted range of inputs equally

KEY WORDS: Range of the cumulative sums, S_t , reservoir design, water storage problems, adjusted range of the cumulative sums, S_t .

ABSTRACT: The storage problem of within-the-year water fluctuations is the main topic of this paper. The storage difference equation which relates inputs, outputs and storage is used for formulating the mathematical problem. This leads to the problem of determining the expected values and variances of the range or adjusted range of cumulative departures from the population and sample mean, respectively.
Using the univariate, bivariate and trivariate normal distribution functions for the marginal and joint distributions of the partial sums, the exact expressions of the expected range are derived for n=1,2, and 3 . From these general expressions, particular cases of the expected range of independent and linearly dependent variables are derived.
The expected value of the adjusted range of inputs equally

dependent (exchangeable variables) and outputs equal to a percentage of the mean inflow, is shown to be expressed in the same way as the expected value of the unadjusted range of exchangeable random variables. This result is relevant in hydrology because when one is interested in overyear storage design and the assumption of independence of streamflow events is sufficiently accurate and the regulation or development is expressed as a fraction of the sample mean inflow, then the expected value of the storage for a given number of years is given by the expected adjusted range which now may be computed exactly by the derived equation.

Jose D. Salas-La Cruz - Hydrology Paper #57
Range Analysis for Storage Problems of Periodic-Stochastic Processes
Colorado State University

dependent (exchangeable variables) and outputs equal to a percentage of the mean inflow, is shown to be expressed in the same way as the expected value of the unadjusted range of exchangeable random variables. This result is relevant in hydrology because when one is interested in overyear storage design and the assumption of independence of streamflow events is sufficiently accurate and the regulation or development is expressed as a fraction of the sample mean inflow, then the expected value of the storage for a given number of years is given by the expected adjusted range which now may be computed exactly by the derived equation.

Jose D. Salas-La Cruz - Hydrology Paper #57
Range Analysis for Storage Problems of Periodic-Stochastic Processes
Colorado State University

dependent (exchangeable variables) and outputs equal to a percentage of the mean inflow, is shown to be expressed in the same way as the expected value of the unadjusted range of exchangeable random variables. This result is relevant in hydrology because when one is interested in overyear storage design and the assumption of independence of streamflow events is sufficiently accurate and the regulation or development is expressed as a fraction of the sample mean inflow, then the expected value of the storage for a given number of years is given by the expected adjusted range which now may be computed exactly by the derived equation.

Jose D. Salas-La Cruz - Hydrology Paper #57
Range Analysis for Storage Problems of Periodic-Stochastic Processes
Colorado State University

dependent (exchangeable variables) and outputs equal to a percentage of the mean inflow, is shown to be expressed in the same way as the expected value of the unadjusted range of exchangeable random variables. This result is relevant in hydrology because when one is interested in overyear storage design and the assumption of independence of streamflow events is sufficiently accurate and the regulation or development is expressed as a fraction of the sample mean inflow, then the expected value of the storage for a given number of years is given by the expected adjusted range which now may be computed exactly by the derived equation.

Jose D. Salas-La Cruz - Hydrology Paper #57
Range Analysis for Storage Problems of Periodic-Stochastic Processes
Colorado State University