RUNS OF PRECIPITATION SERIES

by

Jose Llamas and M. M. Siddiqui

May 1969

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado
RUNS OF PRECIPITATION SERIES

by

Jose Llamas
and
M. M. Siddiqui

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO 80521

May 1969

No. 33
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Abstract</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Subject of this study</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Background of the problem</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td></td>
</tr>
<tr>
<td>A Sequence of Independent Variables</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>3</td>
</tr>
<tr>
<td>2.2 Formulation of the problem</td>
<td>3</td>
</tr>
<tr>
<td>2.3 The independence of ((N_{1j}, S_{1j})) and ((N_{2j}, S_{2j}))</td>
<td>3</td>
</tr>
<tr>
<td>2.4 The distribution of (S_1) in some special cases</td>
<td>4</td>
</tr>
<tr>
<td>2.5 Moment generating functions</td>
<td>4</td>
</tr>
<tr>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Normal and Gamma Distributed Variables</td>
<td>6</td>
</tr>
<tr>
<td>3.1 Normal sequence</td>
<td>6</td>
</tr>
<tr>
<td>3.2 Approximated probability density functions of (I_1) and (I_2)</td>
<td>6</td>
</tr>
<tr>
<td>3.3 Gamma distributed sequence</td>
<td>7</td>
</tr>
<tr>
<td>3.4 Example: Fort Collins, Station No. 53005</td>
<td>9</td>
</tr>
<tr>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Two Mutually Independent Processes</td>
<td>12</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>12</td>
</tr>
<tr>
<td>4.2 Formulation of the problem</td>
<td>12</td>
</tr>
<tr>
<td>4.3 Comment on the dependent case</td>
<td>13</td>
</tr>
<tr>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Monthly Precipitation Series</td>
<td>14</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>14</td>
</tr>
<tr>
<td>5.2 Formulation of the problem</td>
<td>14</td>
</tr>
<tr>
<td>5.3 Properties of runs</td>
<td>15</td>
</tr>
<tr>
<td>5.4 Example</td>
<td>15</td>
</tr>
<tr>
<td>5.5 Explanation of appendixes</td>
<td>15</td>
</tr>
</tbody>
</table>

References | 16 |
Appendix I | 18 |
Appendix II | 20 |
LIST OF FIGURES AND TABLES

Figure

1  Distribution function and probability density function of $S_i$ and $S_2$. 8
2  Distribution function and probability density function of $I_1$ and $I_2$. 8
3  Expected values and variances of $X_{11}$ and $X_{21}$ for normal and gamma assumptions. 10
4  Expected values and variances of $S_1$ and $S_2$ for normal and gamma assumptions. 10
5  Expected values and variances of $I_1$ and $I_2$ for normal and gamma assumptions. 10
6  Graphical representation of the random variables $S_{12}, S_{22}, N_1$ and $N_2$. 12
7  Probability density function of monthly precipitation. Station No. 4.7740. San Diego W.B. APT. 15

Table

1  Improvement of probability density function of $S_i (i = 1, 2).$ 7
2  Improvement of Probability density function of $I_i (i = 1, 2).$ 7
3  Moments of $X_{11}^r$ ($i = 1, 2$) for several values of $r$. 9
4  Moments of $S_i (i = 1, 2)$ for several values of $r$. 9
5  Moments of $I_i (i = 1, 2)$ for several values of $r$. 9
6  Expected values and variances of $X_i$ and $N_i (i = 1, 2)$ for different hypotheses of $(X_n)$. 11
7  Expected values and variances of $S_i$ and $I_i (i = 1, 2)$ for different hypotheses of $(X_n)$. 11
ABSTRACT

Three quantitative measures are introduced for the concepts of "surplus" and "deficit" in hydrologic series. These are: run-length, run-sum and run-intensity. Positive and negative runs of a series are defined in terms of a fixed value, say c, of the variable under consideration, namely precipitation. The distribution function, moments, and other statistical properties of the three variables, run-length, run-sum, and run-intensity, are obtained analytically under the following alternative assumptions on the sequence of annual precipitations:

1. It is independent and normally distributed.
2. It is independent and gamma distributed.

For monthly precipitations, z_t, the series was first standardized by the transformation

\[ x_t = \frac{z_t - \mu_t}{\sigma_t}, \]

where t is of the form t = 12(n-1) + \tau, \tau = 1, ..., 12, n = 1, 2, ..., and where \mu_t and \sigma_t are mean and standard deviation of the series corresponding to the month \tau. Calling "x_t \leq c" as state "0" and "x_t > c" as state "1," the series is then analyzed as a two-state Markov chain with stationary transition probabilities.

Annual precipitation from 27 stations in Colorado, and monthly precipitation from 219 stations in the Western United States are analyzed.
RUNS OF PRECIPITATION SERIES

by

Jose Llamas* and M. M. Siddiqui**

Chapter I

INTRODUCTION

1.1 Subject of this study. The major objective of this study is to carry out the mathematical analysis of some parameters by which the concept of runs of a precipitation series may be defined with reference to the series itself. One of the main problems in water resource projects is to predict accurately the amount of water available during a period of operation and to determine whether or not it will be sufficient. The total amount of water necessary in a given period of time, whether for one particular project or for a number of projects in one region, can be considered as the water demand of that region. Of course, the demand changes from region to region or from country to country. For instance, in arid areas the water demand must be necessarily less than in the humid regions because of different water availability. The same situation is encountered in the agricultural countries as compared to the industrial ones. If in one period of time the supply of water is smaller or greater than the demand, then this period can be considered as dry or wet, respectively.

The concepts of dry or wet periods ought to be taken only in a relative sense so that they depend on a certain level, c. This value, c, can be a constant or a variable according to the characteristics of the water demand. In the case of agricultural projects on a constant surface of land and for the same kind of annual crops, the consumptive use of water is usually constant every year. In the case of urban development, the future requirement of water is related to the growth of the population and to the expected industrial expansion.

1.2 Background of the problem. The problem of runs of a precipitation series has been initiated by Dowker, Siddiqui and Yevjevich [1], and Yevjevich [2]. In these two papers, the authors define a dry or a wet run or a negative or positive run as the period in which the total amount of precipitation is less or greater than a certain constant, c. This constant may correspond to the concept of water demands previously defined. Three main factors may be used in order to characterize a particular negative or positive run: run-length, run-sum, and run-intensity. The run-length of a wet (positive) or a dry (negative) run is the number of terms in a complete positive or negative run, respectively. This is also the duration of a positive or a negative run. This quantity is particularly important in water resource problems because the knowledge of the expected duration of drought or rainfall provides the engineer with the necessary design information.

The run-sum (or the magnitude of a run) is defined as the sum of deviations from a level (water demand) of precipitation over the run-length. These deviations are negative or positive when the run is dry or wet, respectively. In some water resource problems the run-sum is the most important factor. The total capacity of water that must be stored and then supplied depends on the expected run-sum of the future dry negative run. The run-sum of positive or negative runs is directly related to the sizing of reservoir capacities, design and operation of hydroelectric structures, projects of water pollution, sizing of pumps, problems of erosion and sedimentation, and so on.

The third factor characterizing the runs is the run-intensity, which is defined throughout this study as the average intensity or the ratio of run-sum to run-length. This quantity of run-intensity may be used as an index for the classification of regions with respect to precipitation patterns. In this study, the probability distributions of these three quantities will be obtained taking into account several possible cases of the original variable, which is the amount of precipitation in a unit of time.

First, since the unit of time for the precipitation measurement is one year, three different situations are then considered:

(a) One single process of annual precipitation.

(b) Two processes of annual precipitation that are mutually independent.

(c) Two processes of annual precipitation that are dependent.

The term "process" is used in the narrow sense of "stochastic process." It is assumed that any functional dependence on time, such as trend or periodicity, has been removed from any process under consideration. The total amount of annual precipitation is considered as the original random variable.

*Former Ph.D. Graduate of Colorado State University, Civil Engineering Department, Fort Collins, Colorado, presently Adjunct Director of Hydraulic Management, Ministere des Richesses Naturelles, Quebec, P.Q. Canada.

**Professor, Department of Mathematics and Statistics, Colorado State University, Fort Collins, Colorado.
With respect to the original process, two alternative assumptions are made:

(a) The annual precipitations are independent identically distributed normal random variables.

(b) The annual precipitations are independent identically distributed gamma-type random variables.

The hypothesis of independence in annual precipitation is supported by Markovic [3], and physically speaking, seems to be realistic because only some factors of small effects (carry-over of water in river basins, evaporation, etc.) may affect the amount of precipitation of the following year. This hypothesis may be easily verified (autocorrelation test, run test, etc.) before analyzing data for positive and negative runs.

The hypothesis of normality of the annual precipitation is, in fact, one of truncated normality (no negative precipitation) at origin. In the regions where the probability of zero annual precipitation is high (arid or arctic regions), neither the hypothesis of normality nor the hypothesis of gamma distribution are applicable.

From the analysis of the samples from 1141 stations in the western United States, Markovic [3] found that, on the average, the annual precipitations are positively skewed, and the gamma distribution hypothesis is more realistic than the hypothesis of normality.

Second, the three main variables, run-length, run-sum, and run-intensity are analyzed also with monthly precipitation as the random variables. In this case the hypothesis of independence is tested and stated at the beginning of analysis, and no hypothesis of distribution of monthly precipitation is made.

The critical level (water demand) considered in this study is the mean value of the process. In the case of annual precipitation, the second order stationarity of the process is assumed. Therefore, the critical level is assumed to be invariant in time. In the case of monthly precipitation, the stationarity is obtained by standardization of the process.

In order to simplify the algebraic operations, the annual precipitation series is standardized, and in both cases (annual and monthly precipitation) the critical level is assumed to be zero.
Chapter II

A SEQUENCE OF INDEPENDENT VARIABLES

2.1 Introduction. In this chapter, a single process of annual precipitation is analyzed in order to obtain the statistical properties of the three main variables characterizing the positive and negative runs: run-length, run-sum, and run-intensity. This type of single record analysis is necessary before one can study several series and obtain correlation properties of one station with another or of one region with another.

2.2 Formulation of the problem. The problem was formulated by Bowmer, Siddiqui, and Yevjevich [1]. For the sake of ready reference, however, it seems desirable to summarize the essentials of that paper. Some of their results are reported here in a strengthened form, and some new results are also included.

Let \( X_n \), \( n = 1, 2, \ldots \), be independent identically distributed random variables with a common distribution \( F \), which is assumed to be continuous. In the application to be followed after the derivation of theoretical results, \( X_n \) is the total precipitation at a given station during the \( n \)-th year. However, it can also represent the sum of precipitations over several stations in a given region. Also the unit of time may be shorter or longer than a year.

A level, \( c \), in the range of values of \( X_n \) is chosen such that \( 0 < F(c) < 1 \), and the \( n \)-th year is classified as a surplus year if \( X_n > c \) and, in that case, refer to \( X_n - c \) as the surplus. Similarly, the \( n \)-th year is called a deficit year if \( X_n < c \), in which case, \( c - X_n \) is called the deficit. Thus defined, these surpluses and deficits are all positive random variables.

A consecutive sequence of \( k \) surplus years preceded and succeeded by a deficit year is called a positive run-length \( k \), the sum of surpluses \( f(X_n - c) \) over such a run is the positive run-sum, and this run-sum divided by the run-length is called the positive run-intensity. Similar definitions hold for negative runs.

For \( j = 1, 2, \ldots \), let \( N_{1j} \) denote the length of the \( j \)-th negative run-length and \( N_{2j} \) the length of the following positive run-length. If the initial observation, \( X_1 \), is greater than \( c \), the initial positive run is disregarded. Suppose that the \( j \)-th negative run starts with \( X_{1+1} \). Set

\[
S_{1j} = \sum_{k=1}^{N_{1j}} (c - X_{1+k}), \quad I_{1j} = \frac{S_{1j}}{N_{1j}},
\]

\[
S_{2j} = \sum_{k=1}^{N_{2j}} (X_{1+N_{1j}+k} - c), \quad I_{2j} = \frac{S_{2j}}{N_{2j}}.
\]

(2.1)

Then \( S_{1j}, S_{2j} \) are the negative and positive run-sums, respectively, for the \( j \)-th run, and \( I_{1j}, I_{2j} \) are the corresponding run-intensities. The properties of \( N_{1j}, S_{1j}, I_{1j} \), \( i = 1, 2, j = 1, 2, \ldots \) are studied in the further text.

2.3 The independence of \( (N_{1j}, S_{1j}) \) and \( (N_{2j}, S_{2j}) \). For convenience the following notations are introduced:

\[
p = F(c) = P(X_n \leq c), \quad q = 1 - p;
\]

\[
F_1(x) = \frac{F(x - c) - F(c - x)}{F(c)}, \quad \text{if } x \geq 0,
\]

\[
= 0, \quad \text{if } x < 0;
\]

\[
F_2(x) = \frac{F(x + c) - F(c + x)}{1 - F(c)}, \quad \text{if } x \geq 0,
\]

\[
= 0, \quad \text{if } x < 0. \quad (2.2)
\]

Let \( X_{1n}^* \), \( n = 1, 2, \ldots \), be a sequence of independent random variables each with the distribution \( F_1 \) and \( X_{2n}^* \), \( n = 1, 2, \ldots \), another sequence of independent random variables, independent of the sequence \( X_{1n}^* \), with the distribution \( F_2 \). Then

\[
P(X_{1n}^* \leq x) = P(c - X_n \leq x | X_n \leq c) = F_1(x),
\]

\[
P(X_{2n}^* \leq x) = P(X_n - c \leq x | X_n > c) = F_2(x),
\]

\[
P(\sum_{j=1}^{m} X_{1j} \leq x) = P(\sum_{j=1}^{m} (c - X_{j} - c) \leq x, j = 1, \ldots, m) = F_1^m(x),
\]

\[
P(\sum_{j=1}^{m} X_{2j} \leq x) = P(\sum_{j=1}^{m} X_{j} - c \leq x, j = 1, \ldots, m) = F_2^m(x). \quad (2.3)
\]

where, for any distribution function \( H \) and \( m = 1, 2, \ldots, \), \( H^m \) denotes the \( m \)-fold convolution of \( H \) with itself.

First consider the distribution of \( N_{1j} \). If \( X_1 \leq c \), then \( P(N_{11} = k | X_1 \leq c) = P(X_1 \leq c, i = 1, \ldots, k, X_{i+k} > c | X_1 \leq c) = q^{k-1}, k = 1, 2, \ldots \).

If \( X_1 > c \), then

\[
P(N_{11} = k | X_1 > c) = \sum_{j=1}^{m} P(X_1 > c, i = 1, \ldots, j, X_{j+i} \leq c) = q^k \sum_{j=1}^{m} q^j = pq^{k-1}, k = 1, 2, \ldots.
\]

Hence, the unconditional distribution of \( N_{11} \) is
\[
\begin{align*}
P(N_{11} = k) &= P(N_{11} = k | X_1 \leq c) P(X_1 \leq c) + P(N_{11} = k | X_1 > c) P(X_1 > c) = q^k p^{k-1}, \quad k = 1, 2, \ldots, \text{m(2.4)}
\end{align*}
\]

Similarly,
\[
\begin{align*}
P(N_{11} = k_1, N_{21} = k_2) &= P(N_{11} = k_1 | X_1 \leq c) P(X_1 \leq c) \cdot P(N_{21} = k_2 | X_2 \leq c) P(X_2 \leq c) = q_{1}^{k_1} q_{2}^{k_2} \cdot p_{1}^{-k_1-1} p_{2}^{-k_2-1},
\end{align*}
\]
so that
\[
\begin{align*}
P(N_{11} = k_1, N_{21} = k_2) &= P(N_{11} = k_1) P(N_{21} = k_2),
\end{align*}
\]
and \(N_{11}\) and \(N_{21}\) are independent. This argument can be extended to show that \((N_{11}, N_{21}, N_{12}, N_{22}, \ldots)\) are mutually independent. \((N_{11})\) are identically distributed and \((N_{21})\) are identically distributed.

Now, look at the joint distribution of \((N_{11}, S_{11})\)

From (2.3) it follows that
\[
P(S_{11} \leq x | N_{11} = k) = f_{1}^{k}(x),
\]

hence
\[
P(N_{11} = k, S_{11} \leq x) = q^{k-1} f_{1}^{k}(x),
\]

Similar expressions hold for \((N_{21}, S_{21})\).

Finally,
\[
F_{S}(x) = P(S_{11} < x) = \sum_{k=1}^{n} q^{k-1} f_{1}^{k}(x).
\]

Again, one can show that the sequence of vectors \((N_{11}, S_{11})\) is mutually independent and identically distributed with (2.6). This sequence is also independent of \((N_{21}, S_{21})\), which themselves are mutually independent and identically distributed. Since the treatment of one vector sequence is exactly parallel to the other, only one is considered. (In fact \(X_n \leq c\) is equivalent to \(-X_n > -c\) so that a negative run for \(X_n\) at level \(c\), is equivalent to a positive run for \(-S_n\) at level \(-c\).

We choose to concentrate on the negative run \((N_{11}, S_{11})\).

We drop the subscript \(j\) and write it as \((N_{11}, S_{11})\) unless the whole sequence is considered.

2.4 The distribution of \(S_{11}\) in some special cases.

From (2.7), the distribution function, \(F_{S}(x)\) of \(S_{11}\), is directly related to \(F_{1}\) rather than to \(F\). Since \(0 < p < 1, p^{n} \rightarrow 0\), terms after some \(k = n\) may be negligible. For example, if \(p = 1/2, p^{7} < 0.01\) and the series may be truncated at the sixth term with the error of approximation less than one percent uniformly for all \(x\). Actually, since \(F_{k}^{\infty}(x) \leq 1\), then
\[
F_{S}(x) = n \sum_{k=1}^{n} q^{k-1} f_{1}^{k}(x) \leq n q^{n-1} p^{n} = p^{n}
\]

The function, \(F_{S}(x) = F(x, \lambda, r)\), with the density
\[
f(x, \lambda, r) = \frac{\lambda^{r} x^{r-1} e^{-\lambda x}}{\Gamma(r)}, \quad x > 0,
\]

and \(r \geq 0\). If \(r = 1\), then
\[
f_{S}(x) = F_{S}(x) = q \lambda e^{-\lambda x} \left[1 + \lambda x + \frac{(\lambda x)^{2}}{2!} + \ldots\right] = q \lambda e^{-\lambda x}, \quad x > 0,
\]

an exponential distribution. For arbitrary \(r > 0\),
\[
F_{S}(x) = q [F(x, \lambda, r) + p F(x, \lambda, 2r) + p^{2} F(x, \lambda, 3r) + \ldots],
\]

\[
= q \sum_{k=1}^{n} p^{k-1} F(x, \lambda, kr) + R_{n}(x),
\]

where
\[
R_{n}(x) = q [p^{n} F(x, \lambda, (n+1)r) + p^{n+1} F(x, \lambda, (n+2)r) + \ldots]
\]

\[
\leq q p^{n} F(x, \lambda, (n+1)r) [1 + p + p^{2} + \ldots] = p^{n} F(x, \lambda, (n+1)r).\text{(2.10)}
\]

This follows because, for \(x > 0\),
\[
F(x, \lambda, kr) < F(x, \lambda, (k-1)r).
\]

Usually, \(n = 2 \text{ or } 3\) may give a satisfactory approximation.

2.5 Moment generating function. If \(y\) is any random variable with the distribution function, \(J(y)\), and \(E\) is the mathematical expectation operator, then
\[
M_{y}(\theta) = E e^{\theta y} = \int_{-\infty}^{\infty} e^{\theta y} J(y)\text{d}y
\]

is called the moment generating function of \(Y\) or of the distribution, \(J\). \(\theta\) is taken to be a complex number and \(M_{y}(\theta)\) exists at least for \(\text{Re } \theta \leq 0\). If for some \(r = 1, 2, \ldots\), the \(r\)-th moment of \(Y\) exists then it is given by
\[
\mu_{y}^{(r)}(y) = E y^{r} = M_{y}^{(r)}(0),
\]

where \(M_{y}^{(r)}(0)\) is the \(r\)-th derivative of \(M_{y}\) evaluated at \(\theta = 0\).

If \(Y_{1}, Y_{2}, \ldots, Y_{n}\) are independent and \(Y = Y_{1} + \ldots + Y_{n}\), then
\[
M_{Y}(\theta) = M_{1}(\theta) M_{2}(\theta) \ldots M_{n}(\theta).
\]

The function, \(K_{y}(\theta) = \ln M_{y}(\theta)\), is called the cumulant generating function of \(Y\). The \(r\)-th cumulant of \(Y\)
exists if the \( r \)-th moment of \( Y \) exists and is given by
\[
k_r(y) = r\text{-th cumulant of } Y = K_{Y_r}(0) .
\]

Clearly, for \( Y = Y_1 + \ldots + Y_n \), where \( Y_1, \ldots, Y_n \) are independent
\[
k_r(y) = k_r(y_1) + \ldots + k_r(y_n) .
\]

Recall that \( (X_{1,n}, \ldots, X_{n,n}) \) have the common distribution function, \( F_1 \). Set
\[
M_1(v) = E e^{V_{11}^*} = \int_{-\infty}^{\infty} e^{vX} dF_1(x) = \frac{e^{cv}}{p} \int_{-\infty}^{c} e^{vx} d\Phi(x) ,
\]
where \( \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \) is the cumulative distribution function of the standard normal distribution.

\[
K_1(v) = \ln M_1(v) .
\]

From Downer, Siddiqui and Yevjcvich [1],
\[
M_1(u,v) = E e^{uX_{11}^* + vS_1} = \frac{q \exp[u - K_1(v)]}{1-p \exp[u - K_1(v)]} ,
\]
where \( S_1 = \frac{\ln(1-x)}{n} \).

Also
\[
\text{EN}_1 = \frac{1}{q} , \quad \text{var } N_1 = \frac{p}{q^2} ,
\]
\[
\text{ES}_1 = \frac{EX_{11}^*}{q} , \quad \text{var } S_1 = \frac{q \text{var } X_{11}^* + p(\text{EX}_{11}^*)^2}{q^2} .
\]

Cov\( (N_1, S_1) = \frac{p}{q^2} \) \( \text{EX}_{11}^* \)
\[
\text{var } (N_1, S_1) = \frac{p \text{EX}_{11}^*}{\sqrt{pq \text{ var } X_{11}^* + p^2(\text{EX}_{11}^*)^2}}
\] (2.12)

where \( \text{var } X \) is the variance of \( X \), and \( \text{cov}(X, Y) \) is the covariance, and \( \rho(X, Y) \) is the correlation between \( X \) and \( Y \).

The authors just mentioned did not give the moment generating function of \( I_1 = \frac{S_1}{N_1} \) but, in a similar argument,
\[
M_1'(0) = E e^{V_{11}^*} I_1 = E e^v = \sum_{n=1}^{\infty} q^n p^{n-1} (M_1'(\frac{v}{n})) .
\] (2.13)

The evaluation of the moments of \( I_1 \) involves sums of the form
\[
A_1(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} , \quad z = 1, 2, \ldots \quad 0 < z < 1 .
\]

Now,
\[
A_1(z) = 1 + \frac{z}{2} + \frac{z^2}{3} + \ldots = -\frac{1}{z} \ln(1-z) .
\]

Integrating both sides from 0 to \( p \) gives
\[
- \int \frac{p}{z} \ln(1-z) = p + \frac{p^2}{2} + \frac{p^3}{3} + \ldots = \frac{1}{p} A_2(p),
\]
and so on.

Finally,
\[
E I_1 = EX_{11}^* , \quad \text{var } I_1 = \frac{S}{p} \text{ var } X_{11}^* \ln(\frac{1}{q}) .
\] (2.14)

There is little point in giving the algebraic form of higher moments as they can be numerically calculated in a specific situation.
3.1 Normal sequence. Suppose that the original sequence \( \{X_n\} \) is an independent normally distributed sequence with \( \text{EX}_n = 0 \), \( \text{var} \ X_n = 1 \). (If \( \text{EX}_n = \mu \) and \( X_n = \sigma^2 \), then consider the standardized sequence \( (X_n - \mu) / \sigma \) ). Downer, Siddiqui and Yevjevich [1] have studied this situation exhaustively for \( (N_{ij}, S_{ij}) \). Their results are equally applicable to \( (N_{1j}, S_{1j}) \). Now consider the case of \( c = \text{EX}_n = 0 \) to illustrate a method of approximating to the distribution of \( S_{11} \) and \( I_{11} \). Thus

\[
ES_1 = 1.59577, \quad \text{var} \ S_1 = 2.0, \quad ES_1^3 = 18.8615, \quad ES_1^4 = 93.0225
\]

where an approximation to the value of \( \pi \) is used. Also

\[
\begin{align*}
0.79885, & \quad EI_1 = 0.888496, \quad EI_1^3 = 1.263186, \\
0.706437, & \quad EN_1 = 2, \quad \text{var} \ N_1 = 2, \quad \text{cov} (N_1, S_1) = \\
0.59577, & \quad \sigma (N_1, S_1) = 0.798.
\end{align*}
\]

The coefficients of skewness of \( S_1 \) is

\[
C(S_1) = 1.82.
\]

Since this is positive, a gamma distribution is chosen to approximate to \( f_S(x) \). Siddiqui [4] gives the method for this type of approximation.

Let

\[
f(x; g, h) = \frac{g^{-k/2} x^{k-1/2} e^{-x/2g}}{h^k / \Gamma(k/2)} \quad \text{for} \ x > 0
\]

\[
= 0 \quad \text{otherwise}, \tag{3.1}
\]

be the probability density function of a gamma variate, to which the probability density function of \( S_1 \) is approximated.

In (3.1), \( g \) is a scale factor and \( h \) is the effective number of degrees of freedom.

The approximation to the probability density of \( S_1 \) can be improved as follows:

\[
f(x) = f(x; g, h) = \sum_{m=0}^{\infty} \frac{g^m \Gamma(h/2)}{(2g)^m} \frac{1}{\Gamma(m-h/2)} \sum_{j=0}^{\infty} \frac{L_m(c)(-y)^j}{j!}
\]

where \( L_m(c)(y) \) is the Laguerre polynomial of degree \( m \).

\[
(3.2)
\]

\[
L_m(c)(y) = \sum_{j=0}^{m} \frac{m!}{j! (m-j)!} (-y)^j
\]

\[
(3.3)
\]

and

\[
d = E \left( \frac{m-1+q/2}{m-j} \right)^j (2g)^{m-j} \frac{\gamma_j}{J}
\]

where

\[
\gamma_j = E(S_1)^j.
\]

The parameters \( g \) and \( h \) can be computed by the method of moments, i.e., equating the first two moments of the probability density function in Eq. (3.1) with the moments, \( ES_1 \) and \( E(S_1)^2 \), already found. The first two moments of the distribution in Eq. (3.1) are \( gh \) and \( g^2 h(h+2) \). Thus, setting \( gh = ES_1 \) and \( g^2 h(h+2) = ES_1^2 \),

\[
g = 0.626657, \quad h = 2.546482.
\]

Then

\[
f(x; g, h) = 0.816398 e^{-0.797885 x} x^{0.273241}
\]

In this kind of approximation, only the first few polynomials are really important. As a general rule, the order, \( m \), of the last polynomial considered must be such that:

(a) No appreciable oscillations appear in the probability density function.

(b) The coefficient of \( x^m \) must be small in comparison with the coefficients of the terms of lower order.

With those considerations, the probability density function of \( S_1 \) is truncated at \( m = 4 \).

Table 1 shows the different computations. In this table, \( A_m = \frac{\Gamma(h/2)}{\Gamma(m+h/2)} \).

Finally, the probability density function of \( S_1 \) (and \( S_2 \)) is

\[
f_S(x) = 0.816398 e^{-0.797885 x} x^{0.273241} (0.790207 + \ 0.514732 x - 0.265312 x^2 + 0.042132 x^3 - 0.001922 x^4).
\]

3.2 Approximated probability density functions of \( I_1 \) and \( I_2 \). As before, the probability density function of \( I_1 \) (and \( I_2 \)) will be approximated by a function of a gamma variate.

In this case,

\[
\text{Var} \ I_1 = \frac{\gamma}{\Gamma E_1} = 0.157840
\]
TABLE 1

IMPROVEMENT OF PROBABILITY DENSITY FUNCTION OF $S_i$ ($i = 1, 2$)

<table>
<thead>
<tr>
<th>m</th>
<th>$\frac{d_m}{(2g)^m}$</th>
<th>$A_m$</th>
<th>(0.273241)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-0.017781</td>
<td>0.633311</td>
<td>1.579002-2.968478x+1.0441905x$^2$-0.084658x$^3$</td>
</tr>
<tr>
<td>4</td>
<td>-0.192013</td>
<td>0.592816</td>
<td>1.686864-4.228341x+2.22615x$^2$-0.361765x$^3$+0.016887x$^4$</td>
</tr>
</tbody>
</table>

where \( r > 0 \). One can introduce a scale factor \( \lambda \), but it will simply involve multiplying the k-th moment by \( \lambda^k \). Since \( \mu = x, \text{var} x = x \), we consider the moments of the standardized variable

\[
X_1 = \frac{X - r}{\sqrt{r}}
\]

and the sequence \( X_n, n = 1, 2, \ldots \), which are identically distributed.

If \( X_1^{*k} \) and \( X_2^{*k} \) denote the truncated random variables with \( c = 0 \), then

\[
\text{E}X_1^{*k} = r^{-k/2} \text{E}X_1^k
\]

where \( X_1 \) is the variable, \( X \), truncated at \( \mu = r \). Similarly, \( \text{E}X_2^{*k} = r^{-k/2} \text{E}X_2^k \).

Let \( F_1 \) and \( F_2 \) be the distribution functions obtained from Eq. (2.2). Then

TABLE 2

IMPROVEMENT OF PROBABILITY DENSITY FUNCTION OF $I_i$ ($i = 1, 2$)

<table>
<thead>
<tr>
<th>m</th>
<th>$\frac{d_m}{(2g)^m}$</th>
<th>$A_m$</th>
<th>(1.527516)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>0.035602</td>
<td>0.148638</td>
<td>6.727776-25.295999x+22.716197x$^2$-5.297942x$^3$</td>
</tr>
<tr>
<td>4</td>
<td>-0.439633</td>
<td>0.107562</td>
<td>9.296972-46.608005x+62.782081x$^2$-29.284371x$^3$+4.195661x$^4$</td>
</tr>
</tbody>
</table>
Figure 1 Distribution function and probability density function of $S_1$ and $S_2$

\[ F_1(x) = \int_0^x f_1(u) \, du \]
\[ f_1(x) = 0.846396e^{0.797885x+0.273240(0.790207+0.514732x-0.265132x^2}
\quad +0.042152x^3-0.001922x^5) \]

Figure 2 Distribution function and probability density function of $I_1$ and $I_2$

\[ dP_1(x) = \frac{(r-x)^{r-1} e^{-(r-s)}}{r(r) P(r,x)} \, dx, \quad 0 \leq x \leq r \]

where

\[ P(a,x) = \frac{1}{r(a)} \int_0^x t^{a-1} e^{-t} \, dt, \quad a > 0, \quad x > 0, \]

is the incomplete gamma function.

Then

\[ \Gamma(r) P(r,r) E_{11}^k = \int_0^r x^k (r-x)^{r-1} e^{-(r-x)} \, dx = \int_0^r y^{r-1} (r-y)^k e^{-y} \, dy \]

Hence,

\[ E(X_{11}^*)^k = \sum_{j=0}^{k} (-1)^j r^{k-j} \Gamma(r+j) P(r+j,x) \]

In a similar fashion

\[ E(X_{21}^*)^k = \sum_{j=0}^{k} (-1)^j \frac{r^{k-j} \Gamma(r+k-j)}{r^{k/2} \Gamma(r) P(r,x)} \]

\[ E(X_{11}^*)^k = \sum_{j=0}^{k} (-1)^j \frac{r^{k-j} \Gamma(r+k-j)}{r^{k/2} \Gamma(r) P(r,x)} \]

\[ E(X_{21}^*)^k = \sum_{j=0}^{k} (-1)^j \frac{r^{k-j} \Gamma(r+k-j)}{r^{k/2} \Gamma(r) P(r,x)} \]
In the distribution of $X_{11}$ and $X_{21}$,
\[ p = P(x) = P(x, r), \quad q = 1 - p. \]

We define $S_1$ and $S_2$ in terms of the "normalized" variables $X_{11}^*, X_{21}^*$ and then calculate their moments.

The following table shows the values of the first four moments of $X_{11}^*$ and $X_{21}^*$ for several values of $r$. The values of the Incomplete Gamma Function have been taken from K. Pearson [5].

<table>
<thead>
<tr>
<th>r</th>
<th>$EX_{11}^*$</th>
<th>$E(X_{11}^*)^2$</th>
<th>$E(X_{11}^*)^3$</th>
<th>$E(X_{11}^*)^4$</th>
<th>$EX_{21}^*$</th>
<th>$E(X_{21}^*)^2$</th>
<th>$E(X_{21}^*)^3$</th>
<th>$E(X_{21}^*)^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.58198</td>
<td>0.41802</td>
<td>0.32788</td>
<td>0.27027</td>
<td>0.58198</td>
<td>1.16396</td>
<td>3.49104</td>
<td>13.96753</td>
</tr>
<tr>
<td>2</td>
<td>0.64412</td>
<td>0.54456</td>
<td>0.51807</td>
<td>0.53441</td>
<td>0.64412</td>
<td>1.13809</td>
<td>2.89771</td>
<td>9.56150</td>
</tr>
<tr>
<td>4</td>
<td>0.69031</td>
<td>0.65486</td>
<td>0.72614</td>
<td>0.87779</td>
<td>0.69031</td>
<td>1.11027</td>
<td>2.49127</td>
<td>7.06530</td>
</tr>
<tr>
<td>6</td>
<td>0.71055</td>
<td>0.71012</td>
<td>0.84056</td>
<td>1.10099</td>
<td>0.71055</td>
<td>1.09454</td>
<td>2.31406</td>
<td>6.11766</td>
</tr>
</tbody>
</table>

The following tables show the first four moments of $S_1$ and $I_1$ for several values of $r$.

<table>
<thead>
<tr>
<th>r</th>
<th>$ES_1$</th>
<th>$E(S_1)^2$</th>
<th>$E(S_1)^3$</th>
<th>$E(S_1)^4$</th>
<th>$ES_2$</th>
<th>$E(S_2)^2$</th>
<th>$E(S_2)^3$</th>
<th>$E(S_2)^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.58918</td>
<td>4.30027</td>
<td>17.20120</td>
<td>91.63671</td>
<td>0.92068</td>
<td>2.46503</td>
<td>9.89973</td>
<td>53.01123</td>
</tr>
<tr>
<td>2</td>
<td>1.58768</td>
<td>4.33840</td>
<td>17.35705</td>
<td>92.41097</td>
<td>1.08383</td>
<td>2.86815</td>
<td>11.18544</td>
<td>57.70780</td>
</tr>
<tr>
<td>4</td>
<td>1.59252</td>
<td>4.38432</td>
<td>17.63093</td>
<td>94.21243</td>
<td>1.21849</td>
<td>3.24693</td>
<td>12.64762</td>
<td>65.03209</td>
</tr>
<tr>
<td>6</td>
<td>1.59359</td>
<td>4.40703</td>
<td>17.77899</td>
<td>95.26155</td>
<td>1.28230</td>
<td>3.44160</td>
<td>13.46745</td>
<td>69.57820</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>r</th>
<th>$ES_1$</th>
<th>$E(S_1)^2$</th>
<th>$E(S_1)^3$</th>
<th>$E(S_1)^4$</th>
<th>$ES_2$</th>
<th>$E(S_2)^2$</th>
<th>$E(S_2)^3$</th>
<th>$E(S_2)^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.58198</td>
<td>0.38486</td>
<td>0.27422</td>
<td>0.20031</td>
<td>0.58198</td>
<td>0.98911</td>
<td>2.63477</td>
<td>9.77068</td>
</tr>
<tr>
<td>2</td>
<td>0.64412</td>
<td>0.49475</td>
<td>0.42168</td>
<td>0.37902</td>
<td>0.64412</td>
<td>0.96617</td>
<td>2.15717</td>
<td>6.50659</td>
</tr>
<tr>
<td>4</td>
<td>0.69031</td>
<td>0.59059</td>
<td>0.57961</td>
<td>0.60859</td>
<td>0.69031</td>
<td>0.94718</td>
<td>1.85078</td>
<td>4.74713</td>
</tr>
<tr>
<td>6</td>
<td>0.71055</td>
<td>0.63827</td>
<td>0.66659</td>
<td>0.75612</td>
<td>0.71055</td>
<td>0.93751</td>
<td>1.72266</td>
<td>4.09933</td>
</tr>
</tbody>
</table>

Comparing the moments of $S_1$ and $I_1$, $i = 1, 2$, obtained in this way with the same moments as for the normal, it follows that the moments corresponding to the gamma distribution of the original random variable, $X_1$, converge to the moments corresponding to the normal distribution of $X_1$. This convergence is almost independent of $r$ for the moments of $S_1$, but for the other random variables, $S_2$, $I_1$ and $I_2$, both assumptions are similar for large values of $r$ only as shown in Figs. 5, 4 and 5.

3.4 Example: Fort Collins, Station No. 5.3005.

Years of records: $N = 69$
Mean: $\mu = 14.62$
Standard deviation: $\sigma = 4.00$
Equating the mean and variance, it follows
Figure 3  Expected values and variances of $x_{11}^*$ and $x_{21}^*$ for normal and gamma assumptions

Figure 4  Expected values and variances of $S_1$ and $S_2$ for normal and gamma assumptions

Figure 5  Expected values and variances of $I_1$ and $I_2$ for normal and gamma assumptions
TABLE 6  
EXPECTED VALUES AND VARIANCES OF $X_i$ and $N_i$ (i=1,2)  
FOR DIFFERENT HYPOTHESES OF $X_i$  

<table>
<thead>
<tr>
<th>Hypothesis of $X_i$</th>
<th>$EX^*_{i1}$</th>
<th>$EX^*_{i2}$</th>
<th>$VarX^*_{i1}$</th>
<th>$VarX^*_{i2}$</th>
<th>$EN_i$</th>
<th>$EN_2$</th>
<th>$VarN_i$</th>
<th>$VarN_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.79789</td>
<td>0.79789</td>
<td>0.36338</td>
<td>0.36338</td>
<td>2.00000</td>
<td>2.00000</td>
<td>2.00000</td>
<td>2.00000</td>
</tr>
<tr>
<td>Gamma</td>
<td>0.740</td>
<td>0.740</td>
<td>0.245</td>
<td>0.565</td>
<td>2.166</td>
<td>1.857</td>
<td>2.527</td>
<td>1.760</td>
</tr>
<tr>
<td>From the data</td>
<td>0.670</td>
<td>0.910</td>
<td>0.243</td>
<td>0.502</td>
<td>1.850</td>
<td>1.524</td>
<td>0.928</td>
<td>0.725</td>
</tr>
</tbody>
</table>

TABLE 7  
EXPECTED VALUES AND VARIANCES OF $S_i$ AND $I_i$ (i=1,2)  
FOR DIFFERENT HYPOTHESES OF $X_i$  

<table>
<thead>
<tr>
<th>Hypothesis of $X_i$</th>
<th>$ES_1$</th>
<th>$ES_2$</th>
<th>$VarS_1$</th>
<th>$VarS_2$</th>
<th>$EI_1$</th>
<th>$EI_2$</th>
<th>$VarI_1$</th>
<th>$VarI_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>1.59577</td>
<td>1.59577</td>
<td>2.00000</td>
<td>2.00000</td>
<td>0.79789</td>
<td>0.79789</td>
<td>0.25188</td>
<td>0.25188</td>
</tr>
<tr>
<td>Gamma</td>
<td>1.595</td>
<td>1.360</td>
<td>1.880</td>
<td>1.845</td>
<td>0.73800</td>
<td>0.73800</td>
<td>0.160</td>
<td>0.420</td>
</tr>
<tr>
<td>From the data</td>
<td>1.239</td>
<td>1.386</td>
<td>0.816</td>
<td>1.027</td>
<td>0.673</td>
<td>0.957</td>
<td>0.123</td>
<td>0.586</td>
</tr>
</tbody>
</table>

\[ \nu = \frac{\kappa}{\lambda} = 14.62 \]

\[ \sigma^2 = \frac{\kappa}{\lambda^2} = 16. \]

Then

\[ \lambda = 0.9138 \]

\[ r = 15.360 \]

The preceding example shows that the first moment of all random variables, obtained from the data, agrees quite well with the first moment of the theoretical hypothesis (better if the comparison is done with gamma hypothesis). For the random variables, $N_i$, $S_i$, and $I_i$ (i=1,2), the disagreement between the higher moments in both cases, provided by the fact that the sample size and consequently the number of runs, is very small in this example; therefore, the estimation is obviously subject to large sampling fluctuations.
Chapter IV

TWO MUTUALLY INDEPENDENT PROCESSES

4.1 Introduction. In previous chapters, the parameters defining the negative and positive runs of annual precipitation were studied considering one single sequence of original random variables: the total amount of annual precipitation at one station. The concept of runs defined in this way can be generalized to several points in space simultaneously in order to study the behavior of those phenomena in the joint dimensions of time and space. This situation is often encountered in hydrology. For example, if a river is passing through two regions with similar or different meteorological conditions, the expected runs in a downstream storage project depend on the combined pattern of precipitation in both regions. In this case, two different sequences will be required in order to define the process. The same problem can arise in a large watershed in regard to the particular model of precipitation on its main tributaries.

4.2 Formulation of the problem. Consider a sequence of a two-dimensional process, \((X_n, Y_n)\), \(n = 1, 2, \ldots\), where these vectors are mutually independent and have a common distribution function, \(F(X, Y)\). Given two levels, \(c_1\) and \(c_2\), such that \(0 < F(c_1, c_2) < 1\), we have four possible events:

\[
A_n = \{X_n \leq c_1, Y_n \leq c_2\} \quad B_n = \{X_n \leq c_1, Y_n > c_2\}
\]

\[
c_n = \{X_n > c_1, Y_n \leq c_2\} \quad D_n = \{X_n > c_1, Y_n > c_2\}
\]

Of these four, \(A_n\) and \(D_n\) are of interest to us. The \(n\)-th year will be called deficit for both sequences if \(A_n\) occurs and surplus if \(D_n\) occurs. A sequence of \(k\) consecutive \(A\)'s followed and preceded by any other event is a negative run of length, \(k\). A sequence of \(k\) consecutive \(D\)'s followed and preceded by any other event is a positive run of length, \(k\). (For the initial run the requirement of "preceded by" is dropped.) The situation is depicted in Fig. 6.

\[
P(A_n) = F(c_1, c_2) - p, \text{ say}, P(B_n|A_n) = P(D_n) = 1 - p = q.
\]

Thus the distribution of \(N_{11}\) is still given by the formula

\[
P(N_{11} = k) = qp^{k-1}, \quad k = 1, 2, \ldots
\]

The difference is that now there is no guarantee that a negative run will be immediately followed by a positive run. In fact, it is quite possible that a negative run is followed by a few \(B\) and \(C\) type events, which in turn, are followed by another negative run. Also here \(q \neq P(D_n)\). Since the discussion of the positive runs is parallel to that of the negative runs, we omit their mention entirely. We now use the symbols \(S_{n1}\) for \(Z(c - X)\) and \(S_{n2}\) for \(Z(c - Y)\), where the summation is over a (common) negative run.

![Graphical representation of the random variables \(S_{11}, S_{12}, S_{22}, N_1\) and \(N_2\)](image)

When \(\{X_n\}\) is independent of \(\{Y_n\}\), we have

\[
F(c_1, c_2) = F(c_1)G(c_2),
\]

where \(F(x) = P(X_n \leq x)\), \(G(y) = P(Y_n \leq y)\). For example, if both \(X_n\) and \(Y_n\) are standard normal and \(c_1 = c_2 = 0\), then \(F(0,0) = 1/4\). However, we are at liberty to choose \(F\) and \(G\) differently, for instance, \(F\) to be normal and \(G\) to be a gamma distribution.

Now, let

\[
F_{12}(x, y) = P(c_1 - X_1 \leq x, c_2 - Y_1 \leq y | X_1 \leq c_1, Y_1 \leq c_2) = P(c_1 - X_1 \leq x | X_1 \leq c_1) P(c_2 - Y_1 \leq y | Y_1 \leq c_2),
\]

so that the random variables \(\{X_{1n}\}\) and \(\{Y_{1n}\}\), can be defined independently by the truncation of \(F\) and \(G\), respectively. The entire discussion of Chapter II carries through for \(S_{12}\) and \(S_{22}\) except for their covariance properties. We have
\[ \text{ES}_{11} S_{21} = E[E(S_{11}S_{21} | N_{11})] = EN_{1} q^{p} E_{11}^{*} E_{11}^{*} \\
= EN_{1}^{*} E_{11}^{*} E_{11}^{*}, \]
so that

\[ \text{cov}(S_{11}, S_{21}) = \text{var} N_{1} E_{11}^{*} E_{11}^{*} = \frac{p}{q} E_{11}^{*} E_{11}^{*}. \]  

(4.2)

4.3 Comment on the dependent case. From the equation (4.1), it is apparent that no general discussion can be carried very far if \( \{X_n\} \) and \( \{Y_n\} \) are not independent, i.e., when \( F(x, y) \neq F(x)G(y) \). The essential difficulty is in finding the joint distribution of the truncated random variables \( (X_{11}^{*}, Y_{11}^{*}) \).

Even the marginal distribution of \( X_{11}^{*} \) (or of \( Y_{11}^{*} \)) depends on the joint condition \( X_1 \leq c_1, Y_1 \leq c_2 \).
5.1 Introduction. In the three previous chapters the annual precipitation was considered as the basic random variable leading to an objective definition of runs was considered. The series in this form were independent within and strictly stationary.

However, in some cases it is preferable to reduce the length of this original random variable in order to create another process in which the length of observations will be longer. In practical terms this new process offers more advantages, in particular, in the study of those phenomena recurrent in a short period of time. For example, the drought, defined as the negative run over the mean of annual precipitation, does not mean anything to a farmer so long as the precipitation is concentrated in the right period.

In this chapter, the monthly precipitation is the basic random variable. The time series formed by the total precipitation during a month are not stationary because of the seasonal variations. Each series must be considered as a sample of 12 different populations, and some transformations should be necessary in order to bring about stationarity.

5.2 Formulation of the problem. We consider a sequence of monthly precipitation. Let \( P_t, t = \tau + 12(n-1) \) be the total amount of precipitation in the \( \tau \)-th month of the \( n \)-th year. Here, \( \tau = 1, 2, \ldots, 12 \) and \( n = 1, 2, \ldots \). Fix \( \tau \) and set

\[
X_n = \frac{P \tau - u_\tau}{\sigma_\tau}, \quad t = \tau + 12(n-1),
\]

where \( u_\tau \) is the mean value and \( \sigma_\tau \) is the standard deviation for the month \( \tau \). \( X_n, n = 1, 2, \ldots, \) then corresponds to the standardized values of \( P_t \) for the same month of successive years. Clearly, this may be assumed to be either an independent sequence or a mildly dependent stationary sequence. What we assume concerning the dependence is the following. Let

\[
X_n = 1, \quad \text{if} \ X_n > 0,
\]

\[
= 0, \quad \text{if} \ X_n \leq 0.
\]

We assume that the sequence of \( X_n \) forms a two state Markov process with stationary transition probabilities. That is,

\[
P(x_n | x_{n-1}, x_{n-2}, \ldots, x_1) = P(x_n | x_{n-1}) = P(x_2 | x_1). \quad (5.1)
\]

Let

\[
P(x_2 = 0 | x_1 = 0) = 1 - \alpha, \quad P(x_2 = 1 | x_1 = 0) = \alpha
\]

\[
P(x_2 = 0 | x_1 = 1) = \beta, \quad P(x_2 = 1 | x_1 = 1) = 1 - \beta
\]

The transition matrix of the model is

\[
P = \begin{pmatrix}
0 & 1 - \alpha \\
\alpha & \beta \\
1 - \beta & 0
\end{pmatrix}
\]

with equilibrium probabilities

\[
\lim_{n \to \infty} P(x_n = 0) = \pi_0 = \frac{\beta}{\alpha + \beta}, \quad \lim_{n \to \infty} P(x_n = 1) = \pi_1 = \frac{\alpha}{\alpha + \beta}.
\]

We further assume that the initial probability distribution is given by

\[
P(x_1 = 0) = \pi_0, \quad P(x_1 = 1) = \pi_1
\]

so that the chain is stationary. That is,

\[
P(x_n = 0), P(x_n = 1) = \pi_1, \quad \text{for all} \ n = 1, 2, \ldots
\]

Now let \( N_{1j} \) be the \( j \)-th run of 0's and \( N_{2j} \) the following run of 1's. We have

\[
P(N_{1j} = k | x_1 = 0) = P(x_1 = 0, \ldots, x_k = 0, x_{k+1} = 1 | x_1 = 0) = (1 - \alpha)^{k-1} \alpha,
\]

\[
P(N_{1j} = k | x_1 = 1) = \sum_{i=1}^{k} \sum_{j=1}^{j=1} P(x_1 = 1, x_{i+1} = 1, \ldots, x_{j+1} = 0, i = 1, \ldots, k,
\]

\[
x_{j+k+1} = 1 | x_1 = 1) = \sum_{j=1}^{j=1} \beta^{j-1} (1 - \beta)(1 - \alpha)^{k-1}\alpha
\]

\[
= (1 - \alpha)^{k-1} \alpha.
\]

Hence, the unconditional probability

\[
P(N_{1j} = k) = \alpha (1 - \alpha)^{k-1}, \quad k = 1, 2, \ldots
\]

which is the same as Eq. (2.4) with \( p = 1 - \alpha \).

Similarly,

\[
P(N_{2j} = k) = \beta (1 - \beta)^{k-1}, \quad k = 1, 2, \ldots
\]

Let

\[
Z_n = \sum_{j=1}^{n} x_j, \quad Y_n = n - Z_n.
\]

Then \( Z_n \) is the number of surplus months out of \( n \), and
$Y_n$ is the number of deficit months. We know that \[ Y_n \text{ is asymptotically (as } n \rightarrow \infty ) \text{ normally distributed with} \]

\[ EY_n = n \frac{A}{a+B} \quad \text{and } \quad \text{var} \ Y_n = n \frac{A^2 (2a-A)}{(a+B)^2} \]

(Since $Y_n + Z_n = n$, var $(Y_n + Z_n) = 0$ for all $n$.)

5.3 Properties of runs. Defining $S_1$, $I_1$, $S_2$, $I_2$ as before, we note that for $c = 0$, the model outlined above is equivalent to the independent sequence model except that $q = a$, $p = 1 - a$ for $(N_1, S_1, I_1)$ and the same $(p, q)$ will not apply for $(N_2, S_2, I_2)$ unless $\beta = 1 - a$, which is the independent case. Thus, when discussing $N_1, S_1, I_1$, i.e., negative run-length, run-sum and run-intensity, we set $p = 1 - a$, $q = a$, in the formulas (2.11, 2.12 and 2.13).

5.4 Example.

Station 4.7740 San Diego W.B. APT

The probability density function of monthly precipitation at this station is given in Fig. 7. We obtain the following values for the parameters:

\[ a = 0.290 \]
\[ \beta = 0.683 \]

\[ E(S_1) = 1.596 \quad \text{and} \quad E(S_2) = 1.609 \]

\[ E(S_1)^2 = 4.6066 \]
\[ E(S_2)^2 = 5.5628 \]

\[ E(1_1) = 0.4629 \quad \text{and} \quad E(1_2) = 1.0941 \]
\[ E(1_1)^2 = 0.2508 \]
\[ E(1_2)^2 = 2.3969 \]

From the data:

\[ \hat{E}(S_1) = 1.5953 \quad \text{and} \quad \hat{E}(S_1)^2 = 4.2246 \]
\[ \hat{E}(S_2) = 1.5987 \quad \text{and} \quad \hat{E}(S_2)^2 = 5.8756 \]
\[ \hat{E}(1_1) = 0.4629 \quad \text{and} \quad \hat{E}(1_1)^2 = 0.2416 \]
\[ \hat{E}(1_2) = 1.1023 \quad \text{and} \quad \hat{E}(1_2)^2 = 2.2147 \]

5.5 Explanation of appendices. In Appendix I the following tables are provided:

1. Table of incomplete gamma function $P(a, x)$ for $a = 1, k, 14$, and $x = 1, 2, 4, 6, 10$.
2. Data used in example of Chapter III.
3. Locations of precipitation stations in Colorado.
4. Table giving numerical values of means and variances of variables related to runs for the annual precipitation series at stations in Colorado.

Appendix II provides numerical values of parameters discussed in Chapter V, such as $E(X_n^k)$, $\alpha$, $\beta$, $\beta_0$, $\beta_1$ for monthly precipitation series at stations in the Western United States. Areal distribution of these stations is also provided.

Figure 7 Probability density function of monthly precipitation. Station No. 4.7740.
San Diego W.B. APT

15
REFERENCES


3. Markovic, R. D., "Probability Functions of Best Fit to Distributions of Annual Precipitation and Runoff." Colorado State University


## APPENDIX I

### TABLE OF INCOMPLETE GAMMA FUNCTION IN THE FORM P(a,x)

*Used in Chapter III (From K. Pearson [5]*)

<table>
<thead>
<tr>
<th>a</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.63212</td>
<td>0.86498</td>
<td>0.98168</td>
<td>0.99752</td>
<td>0.99995</td>
</tr>
<tr>
<td>2</td>
<td>0.26424</td>
<td>0.59430</td>
<td>0.90745</td>
<td>0.98257</td>
<td>0.99990</td>
</tr>
<tr>
<td>3</td>
<td>0.08030</td>
<td>0.32262</td>
<td>0.76441</td>
<td>0.93761</td>
<td>0.99959</td>
</tr>
<tr>
<td>4</td>
<td>0.01899</td>
<td>0.14317</td>
<td>0.56655</td>
<td>0.84880</td>
<td>0.99668</td>
</tr>
<tr>
<td>5</td>
<td>0.00366</td>
<td>0.05295</td>
<td>0.37099</td>
<td>0.71374</td>
<td>0.98663</td>
</tr>
<tr>
<td>6</td>
<td>0.00060</td>
<td>0.01686</td>
<td>0.21456</td>
<td>0.55412</td>
<td>0.96183</td>
</tr>
<tr>
<td>7</td>
<td>0.00009</td>
<td>0.00483</td>
<td>0.11026</td>
<td>0.39348</td>
<td>0.91243</td>
</tr>
<tr>
<td>8</td>
<td>0.00002</td>
<td>0.00139</td>
<td>0.05067</td>
<td>0.25579</td>
<td>0.85558</td>
</tr>
<tr>
<td>9</td>
<td>0.00001</td>
<td>0.00027</td>
<td>0.02135</td>
<td>0.15252</td>
<td>0.73226</td>
</tr>
<tr>
<td>10</td>
<td>0.00000</td>
<td>0.00005</td>
<td>0.00813</td>
<td>0.08567</td>
<td>0.60625</td>
</tr>
<tr>
<td>11</td>
<td>0.00000</td>
<td>0.00001</td>
<td>0.00274</td>
<td>0.04139</td>
<td>0.47678</td>
</tr>
<tr>
<td>12</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00094</td>
<td>0.02019</td>
<td>0.35680</td>
</tr>
<tr>
<td>13</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00028</td>
<td>0.00883</td>
<td>0.24992</td>
</tr>
<tr>
<td>14</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00008</td>
<td>0.00356</td>
<td>0.16383</td>
</tr>
</tbody>
</table>

## DATA USED IN EXAMPLE OF CHAPTER III

*Fort Collins, Colo. Station No. 53005*

<table>
<thead>
<tr>
<th>Year</th>
<th>P</th>
<th>S</th>
<th>Year</th>
<th>P</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>1891</td>
<td>17.50</td>
<td>.71</td>
<td>1926</td>
<td>13.57</td>
<td>.26</td>
</tr>
<tr>
<td>1892</td>
<td>13.58</td>
<td>-.26</td>
<td>1927</td>
<td>15.77</td>
<td>.28</td>
</tr>
<tr>
<td>1893</td>
<td>5.65</td>
<td>-.22</td>
<td>1928</td>
<td>13.54</td>
<td>-.27</td>
</tr>
<tr>
<td>1894</td>
<td>12.35</td>
<td>-.56</td>
<td>1929</td>
<td>14.08</td>
<td>-.13</td>
</tr>
<tr>
<td>1895</td>
<td>18.07</td>
<td>.86</td>
<td>1930</td>
<td>15.17</td>
<td>.13</td>
</tr>
<tr>
<td>1896</td>
<td>15.76</td>
<td>.23</td>
<td>1931</td>
<td>9.88</td>
<td>-1.18</td>
</tr>
<tr>
<td>1897</td>
<td>15.24</td>
<td>.15</td>
<td>1932</td>
<td>12.80</td>
<td>-.45</td>
</tr>
<tr>
<td>1898</td>
<td>11.03</td>
<td>-.89</td>
<td>1933</td>
<td>15.65</td>
<td>.25</td>
</tr>
<tr>
<td>1899</td>
<td>16.19</td>
<td>-.39</td>
<td>1934</td>
<td>8.87</td>
<td>-1.43</td>
</tr>
<tr>
<td>1900</td>
<td>19.21</td>
<td>1.14</td>
<td>1935</td>
<td>15.95</td>
<td>-.33</td>
</tr>
<tr>
<td>1901</td>
<td>21.17</td>
<td>1.63</td>
<td>1936</td>
<td>11.81</td>
<td>-.70</td>
</tr>
<tr>
<td>1902</td>
<td>18.43</td>
<td>-.95</td>
<td>1937</td>
<td>12.93</td>
<td>-.42</td>
</tr>
<tr>
<td>1903</td>
<td>11.63</td>
<td>-.74</td>
<td>1938</td>
<td>19.72</td>
<td>1.27</td>
</tr>
<tr>
<td>1904</td>
<td>13.13</td>
<td>-.57</td>
<td>1939</td>
<td>7.85</td>
<td>-1.69</td>
</tr>
<tr>
<td>1905</td>
<td>19.85</td>
<td>1.30</td>
<td>1940</td>
<td>13.94</td>
<td>-.17</td>
</tr>
<tr>
<td>1906</td>
<td>19.88</td>
<td>1.31</td>
<td>1941</td>
<td>17.81</td>
<td>.79</td>
</tr>
<tr>
<td>1907</td>
<td>11.64</td>
<td>-.74</td>
<td>1942</td>
<td>21.19</td>
<td>1.63</td>
</tr>
<tr>
<td>1908</td>
<td>17.22</td>
<td>.64</td>
<td>1943</td>
<td>12.27</td>
<td>-.58</td>
</tr>
<tr>
<td>1909</td>
<td>16.24</td>
<td>.40</td>
<td>1944</td>
<td>13.53</td>
<td>-.27</td>
</tr>
<tr>
<td>1910</td>
<td>12.92</td>
<td>-.42</td>
<td>1945</td>
<td>15.73</td>
<td>.27</td>
</tr>
<tr>
<td>1911</td>
<td>10.89</td>
<td>-.93</td>
<td>1946</td>
<td>14.11</td>
<td>-.12</td>
</tr>
<tr>
<td>1912</td>
<td>19.61</td>
<td>1.24</td>
<td>1947</td>
<td>17.95</td>
<td>.83</td>
</tr>
<tr>
<td>1913</td>
<td>15.85</td>
<td>.30</td>
<td>1948</td>
<td>10.85</td>
<td>-1.04</td>
</tr>
<tr>
<td>1914</td>
<td>14.31</td>
<td>-.07</td>
<td>1949</td>
<td>18.79</td>
<td>1.04</td>
</tr>
<tr>
<td>1915</td>
<td>22.79</td>
<td>2.03</td>
<td>1950</td>
<td>12.70</td>
<td>-.48</td>
</tr>
<tr>
<td>1916</td>
<td>13.15</td>
<td>-.36</td>
<td>1951</td>
<td>22.52</td>
<td>1.97</td>
</tr>
<tr>
<td>1917</td>
<td>13.72</td>
<td>-.22</td>
<td>1952</td>
<td>12.74</td>
<td>-.47</td>
</tr>
<tr>
<td>1918</td>
<td>21.79</td>
<td>1.78</td>
<td>1953</td>
<td>11.42</td>
<td>-.80</td>
</tr>
<tr>
<td>1919</td>
<td>10.92</td>
<td>-.92</td>
<td>1954</td>
<td>7.98</td>
<td>-1.66</td>
</tr>
<tr>
<td>1920</td>
<td>11.65</td>
<td>-.74</td>
<td>1955</td>
<td>12.97</td>
<td>-.41</td>
</tr>
<tr>
<td>1921</td>
<td>14.85</td>
<td>.65</td>
<td>1956</td>
<td>12.19</td>
<td>-.60</td>
</tr>
<tr>
<td>1922</td>
<td>9.98</td>
<td>-1.16</td>
<td>1957</td>
<td>19.56</td>
<td>1.23</td>
</tr>
<tr>
<td>1923</td>
<td>27.57</td>
<td>3.23</td>
<td>1958</td>
<td>17.44</td>
<td>.70</td>
</tr>
<tr>
<td>1924</td>
<td>10.64</td>
<td>-.99</td>
<td>1959</td>
<td>14.67</td>
<td>.01</td>
</tr>
<tr>
<td>1925</td>
<td>14.46</td>
<td>-.04</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX I (continued)

In the preceding Table, $P$ means the total amount of annual precipitation in inches, and $S$ is the total amount of annual precipitation in standard measure, i.e., $S = \frac{P - \mu}{\sigma}$ where $\mu = 14.62$ and $\sigma = 4.00$.

Precipitation Stations in Colorado
### APPENDIX I (continued)

#### ANNUAL PRECIPITATION

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Boise</td>
<td>Idaho</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
<td>17.08</td>
</tr>
</tbody>
</table>

**Hypothesis No.**

In the preceding table:

Hypothesis 1 corresponds to the case where the original random variable $X_1$ is assumed to be normally distributed.

In hypothesis 2 the original random variable is assumed to be gamma distributed.

In hypothesis 3 the expected values and variances of all random variables were estimated from the available data, i.e.,

$$ E(Y) = \frac{N}{N} \sum_{i=1}^{N} Y_i $$

$$ E(Y)^2 = \frac{N}{N} \sum_{i=1}^{N} Y_i^2 $$

$$ \text{Var}(Y) = E(Y)^2 - [E(Y)]^2 $$

where $N$ is the number of observations.
Areal distribution of precipitation stations
(After Roesner and Yevjevich)
<table>
<thead>
<tr>
<th>STATION NAME</th>
<th>STATION NO.</th>
<th>Latitude (°)</th>
<th>Longitude (°)</th>
<th>Distance (mi)</th>
<th>Elevation (ft)</th>
<th>Year</th>
<th>Depth (ft)</th>
<th>Temperature (°F)</th>
<th>Humidity (%)</th>
<th>Pressure (in)</th>
<th>Wind Speed (mph)</th>
<th>Wind Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLAYTON</td>
<td>2489</td>
<td>-32.75</td>
<td>151.25</td>
<td>39.2</td>
<td>950</td>
<td>1990</td>
<td>3</td>
<td>72</td>
<td>68</td>
<td>29.92</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>GRAND CAYMAN NATIONAL</td>
<td>24891</td>
<td>-32.68</td>
<td>151.34</td>
<td>40.1</td>
<td>940</td>
<td>1991</td>
<td>4</td>
<td>72</td>
<td>68</td>
<td>29.95</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>PINEVILLE</td>
<td>24892</td>
<td>-32.80</td>
<td>151.29</td>
<td>39.3</td>
<td>940</td>
<td>1992</td>
<td>5</td>
<td>72</td>
<td>68</td>
<td>29.94</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>TULSA UNIVERSITY OF A</td>
<td>24893</td>
<td>-32.79</td>
<td>151.30</td>
<td>39.2</td>
<td>940</td>
<td>1993</td>
<td>6</td>
<td>72</td>
<td>68</td>
<td>29.91</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>TULSA AIRPORT</td>
<td>24894</td>
<td>-32.81</td>
<td>151.32</td>
<td>39.3</td>
<td>940</td>
<td>1994</td>
<td>7</td>
<td>72</td>
<td>68</td>
<td>29.90</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>COMBAT</td>
<td>24895</td>
<td>-32.84</td>
<td>151.31</td>
<td>39.2</td>
<td>940</td>
<td>1995</td>
<td>8</td>
<td>72</td>
<td>68</td>
<td>29.90</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>EUGENE</td>
<td>24896</td>
<td>-32.69</td>
<td>151.44</td>
<td>39.3</td>
<td>940</td>
<td>1996</td>
<td>9</td>
<td>72</td>
<td>68</td>
<td>29.89</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>MOUNTAIN HOME, ARKANSAS</td>
<td>24897</td>
<td>-32.62</td>
<td>151.37</td>
<td>39.4</td>
<td>940</td>
<td>1997</td>
<td>10</td>
<td>72</td>
<td>68</td>
<td>29.88</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>VICKSBURG</td>
<td>24898</td>
<td>-32.81</td>
<td>151.32</td>
<td>39.2</td>
<td>940</td>
<td>1998</td>
<td>11</td>
<td>72</td>
<td>68</td>
<td>29.87</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>LITTLE ROCK</td>
<td>24899</td>
<td>-32.78</td>
<td>151.42</td>
<td>39.3</td>
<td>940</td>
<td>1999</td>
<td>12</td>
<td>72</td>
<td>68</td>
<td>29.86</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>PINE RIDGE</td>
<td>24900</td>
<td>-32.75</td>
<td>151.39</td>
<td>39.2</td>
<td>940</td>
<td>2000</td>
<td>13</td>
<td>72</td>
<td>68</td>
<td>29.85</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>FORT WORTH</td>
<td>24901</td>
<td>-32.85</td>
<td>151.41</td>
<td>39.3</td>
<td>940</td>
<td>2001</td>
<td>14</td>
<td>72</td>
<td>68</td>
<td>29.84</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>LITTLE ROCK</td>
<td>24902</td>
<td>-32.80</td>
<td>151.43</td>
<td>39.3</td>
<td>940</td>
<td>2002</td>
<td>15</td>
<td>72</td>
<td>68</td>
<td>29.83</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>FORT WORTH</td>
<td>24903</td>
<td>-32.80</td>
<td>151.43</td>
<td>39.3</td>
<td>940</td>
<td>2003</td>
<td>16</td>
<td>72</td>
<td>68</td>
<td>29.82</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>FORT WORTH</td>
<td>24904</td>
<td>-32.80</td>
<td>151.43</td>
<td>39.3</td>
<td>940</td>
<td>2004</td>
<td>17</td>
<td>72</td>
<td>68</td>
<td>29.81</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>FORT WORTH</td>
<td>24905</td>
<td>-32.80</td>
<td>151.43</td>
<td>39.3</td>
<td>940</td>
<td>2005</td>
<td>18</td>
<td>72</td>
<td>68</td>
<td>29.80</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>FORT WORTH</td>
<td>24906</td>
<td>-32.80</td>
<td>151.43</td>
<td>39.3</td>
<td>940</td>
<td>2006</td>
<td>19</td>
<td>72</td>
<td>68</td>
<td>29.79</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>FORT WORTH</td>
<td>24907</td>
<td>-32.80</td>
<td>151.43</td>
<td>39.3</td>
<td>940</td>
<td>2007</td>
<td>20</td>
<td>72</td>
<td>68</td>
<td>29.78</td>
<td>18</td>
<td>300</td>
</tr>
</tbody>
</table>

APPENDIX II (continued)
<table>
<thead>
<tr>
<th>STATION NAME</th>
<th>STATION NO.</th>
<th>E(x1)</th>
<th>E(x1)^2</th>
<th>E(x2)</th>
<th>E(x2)^2</th>
<th>x</th>
<th>s</th>
<th>( \bar{x} )</th>
<th>( \sigma )</th>
<th>( \mu )</th>
<th>( \sigma )</th>
<th>E(Dx)</th>
<th>Var(Dx)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPANISH FORK RIVER</td>
<td>42.0119</td>
<td>-0.7038</td>
<td>-0.0732</td>
<td>-0.8957</td>
<td>-1.3974</td>
<td>-0.6286</td>
<td>-0.6372</td>
<td>-0.622</td>
<td>0.5947</td>
<td>0.5947</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>TIDAL</td>
<td>42.0217</td>
<td>-0.2068</td>
<td>-0.3072</td>
<td>-0.880</td>
<td>-1.2757</td>
<td>-0.5332</td>
<td>-0.5332</td>
<td>-0.533</td>
<td>0.5332</td>
<td>0.5332</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>CEDAR LAKE</td>
<td>45.1223</td>
<td>-0.2168</td>
<td>-0.2351</td>
<td>-0.8315</td>
<td>-1.0815</td>
<td>-0.3963</td>
<td>-0.3963</td>
<td>-0.396</td>
<td>0.3963</td>
<td>0.3963</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>CLEARWATER</td>
<td>45.1315</td>
<td>-0.1286</td>
<td>-0.2351</td>
<td>-0.816</td>
<td>-1.0361</td>
<td>-0.3432</td>
<td>-0.3432</td>
<td>-0.343</td>
<td>0.3432</td>
<td>0.3432</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>GOLDENBACK</td>
<td>45.2292</td>
<td>-0.2016</td>
<td>-0.1333</td>
<td>-0.6665</td>
<td>-0.8665</td>
<td>-0.3963</td>
<td>-0.3963</td>
<td>-0.396</td>
<td>0.3963</td>
<td>0.3963</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>HAWK H</td>
<td>45.2928</td>
<td>-0.2278</td>
<td>-0.6539</td>
<td>-0.9231</td>
<td>-1.4516</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>LONGVIEW</td>
<td>45.4759</td>
<td>-0.2147</td>
<td>-0.7349</td>
<td>-0.8582</td>
<td>-1.3529</td>
<td>-0.5629</td>
<td>-0.5629</td>
<td>-0.563</td>
<td>0.5629</td>
<td>0.5629</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>RIMROCK TETON DM</td>
<td>45.7208</td>
<td>-0.2288</td>
<td>-0.6388</td>
<td>-0.9256</td>
<td>-1.4571</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>SEDO WILLO</td>
<td>45.7328</td>
<td>-0.2147</td>
<td>-0.7349</td>
<td>-0.8582</td>
<td>-1.3529</td>
<td>-0.5629</td>
<td>-0.5629</td>
<td>-0.563</td>
<td>0.5629</td>
<td>0.5629</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>WYOMING</td>
<td>45.8207</td>
<td>-0.2288</td>
<td>-0.6388</td>
<td>-0.9256</td>
<td>-1.4571</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>WYOMING</td>
<td>45.8839</td>
<td>-0.2288</td>
<td>-0.6388</td>
<td>-0.9256</td>
<td>-1.4571</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>WYOMING</td>
<td>45.9376</td>
<td>-0.2288</td>
<td>-0.6388</td>
<td>-0.9256</td>
<td>-1.4571</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>WYOMING</td>
<td>45.9675</td>
<td>-0.2288</td>
<td>-0.6388</td>
<td>-0.9256</td>
<td>-1.4571</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>WYOMING</td>
<td>46.2273</td>
<td>-0.2288</td>
<td>-0.6388</td>
<td>-0.9256</td>
<td>-1.4571</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>WYOMING</td>
<td>46.9928</td>
<td>-0.2288</td>
<td>-0.6388</td>
<td>-0.9256</td>
<td>-1.4571</td>
<td>-0.3895</td>
<td>-0.3895</td>
<td>-0.389</td>
<td>0.3895</td>
<td>0.3895</td>
<td>2.0179</td>
<td>0.3111</td>
<td></td>
</tr>
</tbody>
</table>
Key Words: Run-Length, Run-Sum, Run-Intensity, Gamma and Normal Distributions, Moments

Abstract: Three quantitative measures are introduced for the concepts of "surplus" and "deficit" in hydrologic series. These are: run-length, run-sum, and run-intensity. Positive and negative runs of a series are defined in terms of a fixed value, say C, of the variable under consideration, namely precipitation. The distribution function, moments, and other statistical properties of the three variables, run-length, run-sum, and run-intensity, are obtained analytically under the following alternative assumptions on the sequence of annual precipitations:

1. It is independent and normally distributed.
2. It is independent and gamma distributed.

For monthly precipitations, \( X_t \), the series was first standardized by the transformation

\[ x_t = \frac{X_t - \mu}{\sigma} \]

where \( t \) is of the form \( t = 12(n-1) + 1, 12, n = 1, 2, \ldots \), and where \( \mu \) and \( \sigma \) are mean and standard deviation of the series corresponding to the month \( t \). Calling \( x_t \leq C^o \) as state "0" and \( x_t > C^o \) as state "1", the series is then analyzed as a two-state Markov chain with stationary transition probabilities.

Annual precipitation from 27 stations in Colorado, and monthly precipitation from 210 stations in the Western United States are analyzed.

References: Jose Llano and R. M. Siddiqui, Colorado State University Hydrology Paper No. 35 (May 1969), "Runs of Precipitation Series."

Key Words: Run-Length, Run-Sum, Run-Intensity, Gamma and Normal Distributions, Moments

Abstract: Three quantitative measures are introduced for the concepts of "surplus" and "deficit" in hydrologic series. These are: run-length, run-sum, and run-intensity. Positive and negative runs of a series are defined in terms of a fixed value, say C, of the variable under consideration, namely precipitation. The distribution function, moments, and other statistical properties of the three variables, run-length, run-sum, and run-intensity, are obtained analytically under the following alternative assumptions on the sequence of annual precipitations:

1. It is independent and normally distributed.
2. It is independent and gamma distributed.

For monthly precipitations, \( X_t \), the series was first standardized by the transformation

\[ x_t = \frac{X_t - \mu}{\sigma} \]

where \( t \) is of the form \( t = 12(n-1) + 1, 12, n = 1, 2, \ldots \), and where \( \mu \) and \( \sigma \) are mean and standard deviation of the series corresponding to the month \( t \). Calling \( x_t \leq C^o \) as state "0" and \( x_t > C^o \) as state "1", the series is then analyzed as a two-state Markov chain with stationary transition probabilities.

Annual precipitation from 27 stations in Colorado, and monthly precipitation from 210 stations in the Western United States are analyzed.

References: Jose Llano and R. M. Siddiqui, Colorado State University Hydrology Paper No. 35 (May 1969), "Runs of Precipitation Series."
PREVIOUSLY PUBLISHED PAPERS

Colorado State University Hydrology Papers


No. 29 "Estimating Design Floods from Extreme Rainfall," by Frederick C. Bell, July 1968.


Colorado State University Fluid Mechanics Papers

No. 4 "Experiment on Wind Generated Waves on the Water Surface of a Laboratory Channel," by E. J. Plate and C. S. Yang, February 1966.


Colorado State University Hydraulics Papers


No. 2 "Diffusion of Slot Jets with Finite Orifice Length-Width Ratios," by V. Yevjevich, March 1966.