Probability Model for Droughts in Ancient China

D. A. MOOLEY
Indian Institute of Tropical Meteorology, Pune 411 005, India

21 April 1980 and 11 August 1980

ABSTRACT

A statistical analysis has been made of the occurrence of drought over east-central China during the periods 725 BC–426 BC and 200 BC–1 BC as reported by Wang (1979) from ancient Chinese chronicles. The occurrence of drought is random and the number of droughts in a decade follows the Poisson as well as the binomial distribution. Changes in the mean frequency of drought are not found to be significant.

1. Introduction

Wang (1979) has presented nicely the useful meteorological records which he extracted from three ancient Chinese chronicles, covering the period 2187 BC–3 AD and the area between 30–40°N and 106–120°E, i.e., the central part of east China. The purpose of this note is to present a statistical analysis of droughts revealed by these records for periods when the records were continuous and free from gaps. This analysis includes randomness of occurrence of drought, a suitable probability model for drought occurrence, and significant variation, if any, in the mean.

2. Scrutiny of data

A careful examination of the data tabulated from the chronicles shows that there are two periods which are suitable for the present analysis. These are 194 BC–3 AD and 722 BC–426 BC. The former period covers the meteorological records from the chronicles of the single Han Dynasty, while the latter period covers the records from the chronicles of Dukes of Lu (722–503 BC) and Dukes of Gin (511–426 BC). In the record for the periods 425–201 BC and 2187–730 BC there are wide gaps. Hence, the statistical analysis was confined to the periods 200–1 BC and 725–426 BC only. Regarding the records of droughts, it has been mentioned by Wang that these pertain to widespread dryness of large areal extent and therefore are climatologically important. Droughts described in the records as severe and droughts as inferred from descriptions, such as drying of rivers or no rain for some months, have been considered in the present analysis. During the period 200–1 BC there were 18 droughts, all of which have been described as severe. The period 725–426 BC had 11 droughts which have been described as severe and 9 droughts as inferred from drying of rivers and absence of rain in some months, making up 20 droughts in all during the period. The years in which these droughts occurred are indicated in Table 1.

3. Examination of data for randomness of occurrence of droughts

The occurrence of drought was examined for randomness. The WMO Working Group on Climatic Fluctuations [Dzerdzeevskii (1966)], has recommended the application of the Mann-Kendall rank statistic for testing randomness against trend. This test has been applied to the time interval between successive years of drought. The values of τ, the test statistic obtained, are -0.047 and -0.177 for the periods 725–426 BC and 200–1 BC, respectively, and the corresponding values of τ significant at the 5% level are ± 0.329 and ± 0.350. The test does not suggest any trend in the time interval series. Another test for randomness which was applied to the time interval series is Swed and Eisenhart's (1943) test of runs above and below the median. This test neither suggests any trend nor any oscillation. However, this test lacks power.

Cox (1970) has given a test for trends which can be applied to a series of occurrence and non-occurrence, as in the present case of droughts. The test statistic is T_0, which is the sum of the ranks of the years of drought when the years of the period under consideration are in the time sequence. The values of the statistics computed

TABLE 1. Years of drought during the periods 725–426 BC and 200–1 BC

Period (BC)	Years (BC)
725–426 BC	663, 657, 639, 625, 617, 614, 602, 568, 565, 545, 539, 536, 526, 518, 503, 494, 492, 469, 462, 431
200–1 BC	190, 177, 158, 147, 142, 129, 124, 120, 110, 109, 100, 98, 92, 81, 71, 61, 14, 13

0027-0644/81/010191-03$04.75
© 1981 American Meteorological Society

TABLE 2. Results of the test for trend as given by Cox (1970).

Period (BC)	T_0	$E(T_0)$	$g_1(T_0)$	$g_2(T_0)$	$\frac{(T_0 - E(T_0))}{SD}$	$P(T > T_0)$
725–426	3456	3010	0.000	−0.060	1.190	0.12
200–1	1682	1809	0.000	−0.067	−0.542	0.71

Note: $E(T_0)$ denotes mean or expected value of T_0, and $g_1(T_0)$ and $g_2(T_0)$ denote Fisher's measures of skewness and kurtosis of the distribution of T_0. SD is standard deviation.

in connection with the test are given in Table 2. This test also does not suggest any trend.

Since positive association, covering year-to-year serial correlation is possible, a first-order Markov chain model would be the natural alternative. Transition matrices were obtained for the two periods and a chi-square test was applied for independence. The test does not suggest any dependence or serial correlation. The results of the test are given in Table 3.

Thus, the results of the tests applied indicate that the occurrence of drought during the two periods considered was random in a time continuum.

4. Suitable probability model for drought occurrence

Successive droughts are generally separated by rather long intervals and, as such, the occurrences of drought can be treated as independent events. As

TABLE 3. Transition matrix and chi-square test for serial correlation in the series of occurrence and non-occurrence of drought.

Period: 725–426 BC

		This year: No drought	This year: Drought
Last year	No drought	259 (260.3)	20 (18.7)
	Drought	20 (18.7)	0 (1.3)

Chi-square statistic = 1.49 (df = 1)
$P(\chi^2 > 1.49) = 0.23$

Period: 200–1 AD

		This year: No drought	This year: Drought
Last year	No drought	165 (164.6)	16 (16.4)
	Drought	16 (16.4)	2 (1.6)

Chi-square statistic = 0.12 (df = 1)
$P(\chi^2 > 0.12) = 0.74$

Note: Figures in parentheses give the cell frequencies on the assumption of independence. df denotes degrees of freedom.

TABLE 4. Fit of the Poisson and the binomial distribution to y, the number of years of drought in a decade over the east central China.

Total period considered (BC)	y	f_0	f_p	f_b
725–426	0	15	15.40	15.05
	1	11	10.27	10.75
	2	3	3.42	3.45
	≥3	1	0.91	0.75
200–1	0	8	8.13	7.79
	1	7	7.32	7.70
	2	4	3.29	3.43
	≥3	1	1.26	1.08

Note: f_0 is observed frequency, and f_p and f_b are frequencies on Poisson and binomial hypotheses, respectively.

seen from a large number of trials ($n = 200$ and $n = 300$) the probability p of the occurrence of drought is low (<0.1), and np is finite (being 20 when $n = 300$ and 18 when $n = 200$). In this situation, a Poisson distribution is expected to show a good fit. However, since a Poisson distribution is a limiting case of the binomial distribution when p is low and n is large, it is possible that in the situation of transition from binomial to a Poisson distribution, both the distributions might show good fit. In view of this position, both the distributions have been fitted to the number of droughts in a 10-year period. The fit is illustrated by Table 4 which gives the observed frequencies, and the frequencies on the hypotheses of the Poisson and the binomial distributions. The fit appears to be very good for both the distributions.

The fit of both the distributions has been tested for significance by applying the test of variance as suggested by Cochran (1954) who found this test to be more powerful than the chi-square test. The test statistic for the test of variance is

$$\chi_\nu^2 = \frac{(y_i - 10p)^2}{10p} \quad \text{for a Poisson distribution,} \quad (1)$$

$$\chi_\nu^2 = \frac{(y_i - 10p)^2}{10p(1 - p)} \quad \text{for a binomial distribution,} \quad (2)$$

where y_i is the number of drought years in the ith 10-year period, p the probability of the occurrence of a drought year, and ν the number of degrees of freedom is $n - 1$, n being the number of decades. χ_ν^2 is referred to chi-square table with $n - 1$ degrees of freedom. In addition, chi-square test has also been applied for testing the goodness-of-fit of the Poisson and the binomial distributions. The results of both the tests for fit of both the distributions are given in Table 5.

Considering the results of both the tests, it can be stated that the fit of both the distributions is very good and that there is little difference between the two distributions. It may be mentioned that as long as p is near 0.1, as it is in this study, the difference

TABLE 5. Results of variance test and chi-square test for Poisson and binomial distributions applied to the number of droughts in a decade over the east central China.

Test	Distribution	Total period (BC)	Value of test statistic χ^2	Degrees of freedom	Observed significance level ($x^2 > \chi^2$)
Variance	Poisson	725–426	28.00	29	0.52
Variance	Poisson	200–1	17.55	19	0.55
Variance	binomial	725–426	30.00	29	0.42
Variance	binomial	200–1	19.30	19	0.44
Chi-square	Poisson	725–426	0.09	1	0.77
Chi-square	Poisson	200–1	0.06	1	0.82
Chi-square	binomial	725–426	0.02	1	0.89
Chi-square	binomial	200–1	0.12	1	0.74

between the Poisson and the binomial distributions is very small. As p increases beyond 0.1, the difference between the two distributions increases.

In a study of droughts over India during the period 1771–1978 AD, Mooley and Pant (1979) found that the number of droughts in a 5- or 10-year period follows a Poisson probability distribution.

Before we consider application of the Poisson or the binomial model to obtain the probabilities of one, two or three droughts in a 10-year period, let us examine whether the means over sub-periods show any significant variation. Cochran (1954) has suggested the use of the test statistic

$$\chi^2 = \frac{N_1 N_2}{N_1 + N_2} \frac{(\bar{y}_1 - \bar{y}_2)^2}{\bar{y}}, \quad \text{df} = 1, \tag{3}$$

where N is the total period in years, $N_1 + N_2 = N$, $\bar{y}_1$, $\bar{y}_2$ and $\bar{y}$ are the means based on N_1, N_2 and N years of data, and df is degrees of freedom. Significance of χ^2 can be judged by referring to chi-square tables for 1 degree of freedom. Changes in the mean after the first half of the periods 200–1 BC and 725–426 BC changes from 725–426 BC to 200–1 BC and from 725–506 BC (the period covered by the chronicles of the Dukes of Lu) to 505–426 BC (the period covered by the chronicles of the Dukes of Gin) were tested for significance. The results of these tests are given in Table 6. It can be seen that none of the changes in the mean are significant.

TABLE 6. Test for significance of the change in the mean.

Details of change	Test statistic χ^2 (df = 1)	$P(x^2 > \chi^2)$ df = 1
i) Period 200–101 BC to period 100–1 BC	0.22	0.67
ii) Period 725–576 BC to period 575–426 BC	1.80	0.20
iii) Period 725–426 BC to period 200–1 BC	0.83	0.39
iv) Period 725–506 BC of Dukes of Lu to 505–426 BC of Dukes of Gin	0.11	0.74

In view of the fact that the mean over different periods has not changed significantly, we can use the Poisson or the binomial model with this mean to obtain the probabilities of one, two or three droughts in a 10-year period. These probabilities could be utilized for planning funds for ameliorating the lot of the people struck by the calamity of drought.

5. Concluding remarks

The number of droughts in a decade over the east central China is found to follow the Poisson as well as the binomial distribution. Changes in the mean frequency of droughts are not observed to be significant and hence any of these two models with the mean provided by the data could be used to obtain the probabilities, of one, two or three droughts in a decade. These probabilities could be used for planning funds for mitigating the sufferings of the people resulting from the droughts.

Acknowledgments. I am thankful to the Director for his suggestion and for the facilities to pursue this work and to Mrs. S. P. Lakade for typing of the paper.

REFERENCES

Cochran, W. G., 1954: Some methods of strengthening the common tests. *Biometrics*, **10**, 417–451.
Cox, D. R., 1970: *Analysis of Binary Data*. Methuen, 62–64.
Dzerdzeevskii, D., 1966: Report of Working Group on Climatic Fluctuations. Tech. Note No. 79, WMO No. 195, TP 100, 1–80.
Mooley, D. A., and G. B. Pant, 1979: Droughts in India over the last 200 years, their socio-economic impact and remedial measures for them. *Proceedings of the Conference on Climate and History*, Norwich, U.K., Cambridge University Press (in press).
Swed, F. S., and C. Eisenhart, 1943: Tables for testing randomness of grouping in a sequence of alternatives. *Ann. Math. Statis.*, **14**, 66–87.