551·577·36 : 519·2

Sequence of months with rain above and below average in Rayalaseema considered in relation to the theory of probability

T. R. SRINIVASAN

Meteorological Office, Poona

(*Received 8 July 1953*)

ABSTRACT. Monthly rainfall over a period of 60 years (1891—1950) in Rayalaseema has been analysed. The averages of the 12 months were determined and the sequences of months with rain above and below average (designated as *A* and *B*) respectively studied. It is found that an empirical relationship of the form $\log F = nR + S$ (where *R* and *S* are constants) exists between *n* the length of the run or sequence and *F* the cumulative frequency of runs of length *n* and more.

After performing the contingency test for the independence of the nature of consecutive months an approach based on the theory of probability is made and a justification for the observed logarithmic relation obtained.

1. Introduction

The average and extreme lengths of wet and dry spells are of importance to a country like India where agricultural operations depend mostly on rains. In this context it is also of interest to study whether the occurrence of a certain spell has any influence on the character of the following spell, *i.e.*, whether the fact that a particular month was *B* (below average) could affect the probability of the next month being *B* as well.

Considerable work has been done along these lines and dates as far back as 1916 when Newnham (1916) studied the 'persistence' of wet and dry days. Gold (1929) discussed whether the run of meteorological events has a real significance or is due to mere chance. He showed that the probable number of runs of length *r* out of *m* events is given by $(m + 3 - r)/2^{r+1}$. An important contribution on this subject is due to Cochran (1936) who extended Gold's formula to the case of events with unequal chances. Beer, Drummond, and Furth (1946) have studied the sequence of wet and dry months at Kew and six other stations in Britain in relation to the theory of probability and this paper is based on their theory.

The present investigation is confined to the rainfall in Rayalaseema which is frequently affected by famine. This area has an annual rainfall of 24·2″, about 60 per cent of which is received during the period June to September.

2. Data

The average of the monthly rainfall at all the raingauge stations in Rayalaseema in a particular year was taken to represent the average monthly rainfall over the area for that year. Based on such data for the 60 years (1891-1950), means were worked out for each of the twelve months. Table 1 shows the monthly and annual means of rainfall for the area for each year together with the number of raingauges. Monthly means were also graphed to detect the presence, if any, of trend in the series. The graphs showed the absence of trend in all the months. Each month was then designated 'above average' *A* or 'below average' *B* according as its rainfall was above or below the mean *M* for that particular month. This eliminated the effect of annual variation. A latitude of two cents was used in obtaining *A*'s and *B*'s and there were in all ten cases out of a total of 720 (60×12) monthly values where the rainfall was within ±2 cents of the mean value. These are denoted by *N* in Table 2. As the aim of the study was to see whether there is 'persistence' (or continuity of) months with rainfall above

TABLE 1

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual	No. of raingauges considered
1891	0·03	0·11	0·11	0·49	0·41	1·95	1·72	2·57	1·69	2·95	0·59	0·17	12·79	48
1892	0·00	0·01	0·01	0·46	0·82	6·03	3·73	8·23	5·15	7·33	0·35	0·51	32·63	48
1893	0·11	0·24	1·31	0·57	2·08	3·50	5·06	2·98	3·37	6·41	4·24	0·00	29·87	48
1894	0·08	0·16	0·09	0·81	1·47	1·05	3·56	7·19	3·38	6·16	1·68	0·04	25·67	48
1895	0·00	0·02	0·00	1·24	1·68	2·29	3·25	2·78	7·73	6·57	0·84	0·99	27·39	48
1896	0·27	0·00	0·03	0·26	1·80	2·03	2·09	3·50	2·38	0·36	2·95	0·20	15·87	48
1897	0·05	0·10	0·08	0·46	1·44	3·17	2·27	4·68	10·60	2·48	0·45	0·17	25·95	48
1898	0·00	0·11	0·04	1·51	1·44	1·89	2·43	1·93	7·84	1·82	3·61	0·80	23·42	48
1899	0·00	0·00	0·04	2·11	1·49	0·68	0·49	2·51	7·93	2·31	0·30	0·01	17·87	48
1900	0·00	0·00	0·01	1·31	1·28	2·32	5·19	1·85	6·19	2·56	0·47	0·45	21·63	48
1901	0·25	1·93	0·01	0·58	2·54	2·42	2·62	2·08	4·03	2·71	2·75	0·59	22·51	49
1902	0·15	0·00	0·03	0·79	1·54	3·21	1·61	3·55	6·53	6·26	1·28	0·76	25·71	49
1903	0·36	0·00	0·00	0·30	2·22	2·65	4·75	4·82	8·09	4·07	7·90	0·94	36·10	49
1904	0·26	0·00	0·10	0·50	2·44	1·65	2·50	0·64	3·21	4·92	0·04	0·12	16·38	49
1905	0·04	0·17	0·46	0·55	1·37	2·83	1·61	8·61	0·81	4·83	0·41	0·01	21·70	52
1906	1·14	0·02	0·04	0·04	0·49	3·54	4·66	5·23	5·55	3·71	0·57	3·93	28·92	52
1907	0·02	0·00	0·15	2·60	0·13	2·15	4·74	1·55	3·76	0·82	2·78	0·70	19·40	52
1908	0·20	0·06	0·41	0·24	1·47	1·46	2·86	1·66	8·93	2·45	0·10	0·02	19·86	52
1909	1·55	0·01	0·03	1·45	2·46	1·61	2·49	8·74	7·32	0·81	0·25	0·01	26·73	52
1910	0·00	0·00	0·05	0·35	1·52	1·69	6·04	6·66	8·81	5·11	3·33	0·00	33·56	52
1911	0·00	0·00	0·03	0·51	2·01	2·35	3·25	2·42	3·80	2·78	1·14	0·34	18·63	48
1912	0·00	0·19	0·02	0·56	0·71	1·39	2·99	4·53	6·59	4·19	4·68	0·00	25·85	48
1913	0·00	0·00	0·00	0·30	2·47	2·96	4·04	0·94	4·74	5·32	0·01	0 78	21·56	48
1914	0·00	0·00	0·02	0·58	1·72	1·86	3·33	5·06	6·09	1·35	1·31	0·25	21·57	49
1915	0·76	0·14	2·19	0·50	1·80	2·22	4·37	2·13	7·84	2·73	5·29	0·06	30·03	49
1916	0·00	0·02	0·00	0·34	2·19	2·27	8·23	6·19	7·07	10·15	3·39	0·01	39·86	49
1917	0·05	2·27	0·34	0·24	1·77	3·54	1·81	5·96	8·44	6·29	3·05	0·09	33·85	49
1918	0·60	0·01	0·17	0·43	2·98	0·86	0·89	2·35	6·04	0·16	4·90	0·37	19·76	49
1919	0·49	0·00	0·49	0·86	1·68	3·12	4·06	0·88	10·59	2·45	3·84	0·30	28·76	48
1920	1·07	0·04	0·02	0·22	1·01	1·53	0·93	2·82	5·23	2·93	0·98	0·00	16·78	49
1921	0·26	0·00	0·02	1·77	0·21	2·52	5·12	1·80	2·32	8·77	2·54	0·03	25·36	50
1922	1·03	0·00	0·00	0·43	1·91	1·08	2·29	2·60	0·99	3·47	6·46	0·08	20·34	50
1923	0·08	0·68	0·80	0·59	1·24	1·64	2·83	0·86	6·42	1·36	0·23	0·12	16·85	50
1924	0·08	0·00	0·04	0·54	1·52	1·45	2·98	3·64	7·33	1·42	3·17	0·06	22·23	50
1925	0·00	0·00	0·05	0·81	4·28	1·17	3·23	4·73	3·42	4·76	2·19	2·48	27·12	50
1926	1·24	0·00	0·12	0·97	1·46	3·50	2·35	1·81	5·80	2·72	0·29	0·03	20·29	50
1927	0·01	0·05	0·01	0·04	0·94	2·78	4·50	2·58	7·09	1·98	4·61	0·00	24·59	51
1928	0·00	1·62	0·30	0·70	0·96	3·00	3·91	3·59	4·03	5·27	0·25	0·39	24·02	49
1929	0·20	0·50	0·03	1·21	1·67	2·11	1·20	1·42	7·55	4·19	2·54	0·42	23·04	49
1930	0·12	0·10	0·28	0·34	3·75	2·95	2·07	1·69	5·48	7·52	1·64	0·81	26·75	49
1931	0·02	0·00	0·10	0·25	1·70	3·83	2·05	1·07	6·35	2·68	2·12	0·46	20·63	49
1932	0·00	0·29	0·00	0·28	1·83	2·05	2·59	6·14	3·43	3·74	3·16	0·04	23·55	49
1933	0·00	0·13	0·20	0·79	2·90	1·23	3·43	7·00	3·65	5·01	1·64	1·56	27·54	49
1934	0·07	0·01	0·00	1·11	0·72	2·81	4·26	1·52	1·46	3·02	2·10	0·06	17·14	49
1935	0·12	0·04	0·06	1·40	0·69	3·37	3·52	8·09	3·33	4·85	0·10	0·77	26·34	49

TABLE 1 (*contd*)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual	No. of rain-gauges considered
1936	0·01	0·78	0·51	0·47	2·29	3·14	2·05	2·36	5·09	1·94	4·45	0·10	23·19	49
1937	0·00	0·47	0·29	3·62	0·88	1·35	4·22	1·43	4·17	5·12	1·08	0·24	22·87	49
1938	0·00	0·09	0·39	0·12	1·94	2·93	2·71	9·53	9·71	0·66	0·11	0·01	28·20	49
1939	0·06	0·00	0·29	1·17	0·43	2·21	2·09	4·54	5·29	5·19	2·81	0·02	24·10	49
1940	0·00	0·00	0·08	1·33	5·25	2·49	2·59	3·78	3·99	6·91	3·05	0·25	29·72	49
1941	0·14	0·38	0·00	0·39	1·23	1·72	1·42	2·55	6·72	3·28	0·92	2·19	20·94	49
1942	0·07	0·03	0·00	1·02	1·19	4·53	1·82	4·36	2·10	1·44	0·99	0·57	18·12	49
1943	0·24	0·01	0·04	0·84	5·67	1·50	2·63	2·33	6·54	6·57	2·31	0·04	28·72	49
1944	0·00	0·35	2·16	0·31	1·16	4·12	5·13	0·94	5·98	6·68	1·96	0·01	28·80	49
1945	0·02	0·00	0·00	0·90	1·63	1·28	5·65	2·77	4·03	1·80	1·52	0·02	19·62	63
1946	0·00	0·18	0·31	0·57	2·14	1·57	2·26	4·12	4·93	2·71	6·71	2·25	27·75	63
1947	0·38	0·04	0·00	0·38	0·70	3·09	4·69	6·12	6·46	3·13	0·28	0·39	25·66	63
1948	0·06	0·00	0·05	1·57	1·45	1·15	2·92	4·04	3·98	3·06	3·51	0·05	21·84	63
1949	0·00	0·00	0·02	0·32	2·48	3·04	3·02	4·57	8·72	3·99	1·19	0·00	27·35	63
1950	0·04	0·15	0·06	0·04	1·65	2·07	2·34	3·08	4·71	4·48	1·26	0·00	19·88	63
Mean	0·20	0·19	0·21	0·76	1·74	2·37	3·16	3·67	5·51	3·85	2·14	0·43	24·2	
% of Annual	0·8	0·8	0·9	3·1	7·2	9·8	13·0	15·1	22·7	15·9	8·8	1·8	—	
Median	0·04	0·02	0·04	0·55	1·60	2·25	2·90	2·99	4·55	3·40	1·55	0·11	24·0	
Std. devn.	0·346	0·441	0·432	0·641	1·048	1·020	1·433	2·255	2·332	2·163	1·849	0·721	5·28	
Coeff. of Variability	173	232	206	84	60	43	36	61	42	56	88	108	21	

TABLE 2

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1891	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
1892	*B*	*B*	*B*	*B*	*B*	*A*	*A*	*A*	*B*	*A*	*B*	*A*
1893	*B*	*A*	*A*	*B*	*A*	*A*	*A*	*B*	*B*	*A*	*A*	*B*
1894	*B*	*B*	*B*	*A*	*B*	*B*	*A*	*A*	*B*	*A*	*B*	*B*
1895	*B*	*B*	*B*	*A*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*A*
1896	*A*	*B*	*B*	*B*	*A*	*B*	*B*	*B*	*B*	*B*	*A*	*B*
1897	*B*	*B*	*B*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*B*	*B*
1898	*B*	*B*	*B*	*A*	*B*	*B*	*B*	*B*	*A*	*B*	*A*	*A*
1899	*B*	*B*	*B*	*A*	*B*	*B*	*B*	*B*	*A*	*B*	*B*	*B*
1900	*B*	*B*	*B*	*A*	*B*	*B*	*A*	*B*	*A*	*B*	*B*	*N*
1901	*A*	*A*	*B*	*B*	*A*	*A*	*B*	*B*	*B*	*B*	*A*	*A*
1902	*B*	*B*	*B*	*A*	*B*	*A*	*B*	*B*	*A*	*A*	*B*	*A*
1903	*A*	*B*	*B*	*B*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
1904	*A*	*B*	*B*	*B*	*A*	*B*	*B*	*B*	*B*	*A*	*B*	*B*
1905	*B*	*N*	*A*	*B*	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*B*

TABLE 2 *(contd)*

Year		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1906		*A*	*B*	*B*	*B*	*B*	*A*	*A*	*A*	*B*	*B*	*B*	*A*
1907		*B*	*B*	*B*	*A*	*B*	*B*	*A*	*B*	*B*	*B*	*A*	*A*
1908		*N*	*B*	*A*	*B*	*B*	*B*	*B*	*B*	*A*	*B*	*B*	*B*
1909		*A*	*B*	*B*	*A*	*A*	*B*	*B*	*A*	*A*	*B*	*B*	*B*
1910		*B*	*B*	*B*	*B*	*B*	*B*	*A*	*A*	*A*	*A*	*A*	*B*
1911		*B*	*B*	*B*	*B*	*A*	*N*	*A*	*B*	*B*	*B*	*B*	*B*
1912		*B*	*N*	*B*	*B*	*B*	*B*	*B*	*A*	*A*	*A*	*A*	*B*
1913		*B*	*B*	*B*	*B*	*A*	*A*	*A*	*B*	*B*	*A*	*B*	*A*
1914		*B*	*B*	*B*	*B*	*N*	*B*	*A*	*A*	*A*	*B*	*B*	*B*
1915		*A*	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*A*	*B*
1916		*B*	*B*	*B*	*B*	*A*	*B*	*A*	*A*	*A*	*A*	*A*	*B*
1917		*B*	*A*	*A*	*B*	*A*	*A*	*B*	*A*	*A*	*B*	*A*	*B*
1918		*A*	*B*	*B*	*B*	*A*	*B*	*B*	*B*	*A*	*B*	*A*	*B*
1919		*A*	*B*	*A*	*A*	*B*	*A*	*A*	*B*	*A*	*B*	*A*	*B*
1920		*A*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*
1921		*A*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*B*	*A*	*A*	*B*
1922		*A*	*B*	*B*	*B*	*A*	*B*	*B*	*B*	*B*	*B*	*A*	*B*
1923		*B*	*A*	*A*	*B*	*B*	*B*	*B*	*B*	*A*	*B*	*B*	*B*
1924		*B*	*B*	*B*	*B*	*B*	*B*	*B*	*B*	*A*	*B*	*A*	*B*
1925		*B*	*B*	*B*	*A*	*A*	*B*	*A*	*A*	*B*	*A*	*A*	*A*
1926		*A*	*B*	*B*	*A*	*B*	*A*	*B*	*B*	*A*	*B*	*B*	*B*
1927		*B*	*B*	*B*	*B*	*B*	*A*	*A*	*B*	*A*	*B*	*A*	*B*
1928		*B*	*A*	*A*	*B*	*B*	*A*	*A*	*B*	*B*	*A*	*B*	*B*
1929		*N*	*A*	*B*	*A*	*B*	*B*	*B*	*B*	*A*	*A*	*A*	*N*
1930		*B*	*B*	*A*	*B*	*A*	*A*	*B*	*B*	*B*	*A*	*B*	*A*
1931		*B*	*B*	*B*	*B*	*B*	*A*	*B*	*B*	*A*	*B*	*B*	*A*
1932		*B*	*A*	*B*	*B*	*A*	*B*	*B*	*A*	*B*	*B*	*A*	*B*
1933		*B*	*B*	*N*	*A*	*A*	*B*	*A*	*A*	*B*	*A*	*B*	*A*
1934		*B*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*B*	*B*	*A*	*B*
1935		*B*	*B*	*B*	*A*	*B*	*A*	*A*	*A*	*B*	*A*	*B*	*A*
1936		*B*	*A*	*A*	*B*	*A*	*A*	*B*	*B*	*B*	*B*	*A*	*B*
1937		*B*	*A*	*A*	*A*	*B*	*B*	*A*	*B*	*B*	*A*	*B*	*B*
1938		*B*	*B*	*A*	*B*	*A*	*A*	*B*	*A*	*A*	*B*	*B*	*B*
1939		*B*	*B*	*A*	*A*	*B*	*B*	*B*	*A*	*B*	*A*	*A*	*B*
1940		*B*	*B*	*B*	*A*	*A*	*A*	*B*	*A*	*B*	*A*	*A*	*B*
1941		*B*	*A*	*B*	*B*	*B*	*B*	*B*	*B*	*A*	*B*	*B*	*A*
1942		*B*	*B*	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*B*	*B*	*A*
1943		*A*	*B*	*B*	*A*	*A*	*B*	*B*	*B*	*A*	*A*	*A*	*B*
1944		*B*	*A*	*A*	*B*	*B*	*A*	*A*	*B*	*A*	*A*	*B*	*B*
1945		*B*	*B*	*B*	*A*	*B*	*B*	*A*	*B*	*B*	*B*	*B*	*B*
1946		*B*	*N*	*A*	*B*	*A*	*B*	*B*	*A*	*B*	*B*	*A*	*A*
1947		*A*	*B*	*B*	*B*	*B*	*A*	*A*	*A*	*A*	*B*	*B*	*B*
1948		*B*	*B*	*B*	*A*	*B*	*B*	*B*	*A*	*B*	*B*	*A*	*B*
1949		*B*	*B*	*B*	*B*	*A*	*A*	*B*	*A*	*A*	*A*	*B*	*B*
1950		*B*	*B*	*B*	*B*	*B*	*A*	*B*	*B*	*B*	*A*	*B*	*B*
Total	*A*	15	11	15	23	23	27	26	24	29	27	27	17
	B	43	46	44	37	36	32	34	36	31	33	33	41
	N	2	3	1	—	1	1	—	—	—	—	—	2

A : Months with rainfall above average

B : Months with rainfall below average

N : Months with rainfall, 'Average ± 0·02'

TABLE 3(a)

Frequencies of *A* and *B* sequences

	Length of run or sequence												
A	1	2	3	4	5	6	7	8	9	10	11	12	months
Observed frequency	105	44	12	4	2	—	—	—	1	—	—	—	
Calculated frequency (Constant probability)	105·9	38·9	14·3	5·2	1·9	0·7	0·3	0·1					
Calculated frequency (Varying probability)	100·6	38·0	15·0	5·9	2·3	0·9	0·3	0·1					
Frequency obtained by stirring	104	44	11	5	2	2							
B	1	2	3	4	5	6	7	8	9	10	11	>12	months
Observed frequency	64	42	24	14	13	7	1	1	1	—	2	1	
Calculated frequency (Constant probability)	61·4	38·9	24·6	15·6	9·9	6·2	3·9	2·5	1·6	1·0	0·6		
Calculated frequency (Varying probability)	60·6	36·6	23·1	15·6	10·2	6·7	4·3	2·7	1·7	1·0	0·6		
Frequency obtained by stirring	69	35	26	14	10	2	6	3	1	2	1	1	

TABLE 3 (b)

Frequencies of *A* and *B* sequences during the wet season (May—November)

	Length of run or sequence							
A	1	2	3	4	5	6	7	months
Observed frequency	73	27	9	3	2	—	1	
Calculated frequency (Constant probability)	71·0	27·4	10·4	3·9	1·4	0·8	0·2	
Calculated frequency (Varying probability)	69·8	27·3	12·6	3·7	1·4	0·5	0·2	
B								
Observed frequency	64	33	11	12	2	—	1	
Calculated frequency (Constant probability)	61·9	31·2	15·3	7·5	3·6	1·6	1·1	
Calculated frequency (Varying probability)	63·3	31·3	14·7	8·0	3·6	1·6	1·1	

or below a value (*viz.* the average M) a latitude more than two cents was not desirable. Table 2 gives the classification of A and B.

Frequencies of A and B sequences of different lengths have been determined from Table 2 and are given in Tables 3 (a) to 3(c). The occurrence of N was regarded as breaking a sequence of like events.

From the above tables we can get the frequency F of a run of at least n months by cumulating from below, *i.e.*, from the longest to the shortest. These cumulative frequencies are given in Table 4 and graphed in Fig. 1 in the logarithmic scale. It may be seen from the graph that $\log F$ is linearly related to n. We have now to seek a theoretical justification for this relationship. Before appealing to the theory of probability to solve this problem we have to establish that the events are independent of one another, *i.e.*, we have to see whether the rainfall above average A or below average B of one particular month influences the character of the succeeding month. The contingency test for independence is employed here. The data were first arranged into four classes, *viz.*, A followed by A, A followed by B, B followed by A and B followed by B and we get

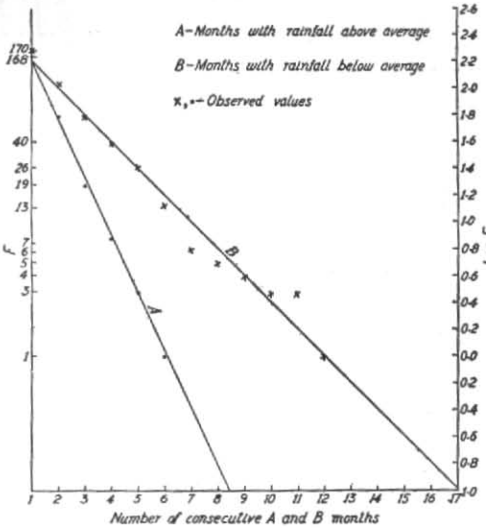

Fig. 1

Current month	Preceding month		
	A	B	Total
A	96	165	261
B	163	276	439
Total	259	441	700

(out of the 720 months there would be obviously 20 months which were either preceded or succeeded by the ten M's and thus we have only 700 months in the 2×2 contingency table above).

TABLE 3 (c)

Frequencies of A and B sequences during the dry season (December—April)

	Length of run or sequence					
A	1	2	3	4	5	months
Observed frequency	48	15	1	—	—	
Calculated frequency (Constant probability)	49·7	11·7	2·7	0·6	0·1	
Calculated frequency (Varying probability)	44·6	11·4	1·8	0·3	0·1	
B						
Observed frequency	28	18	16	11	10	
Calculated frequency (Constant probability)	34·3	22·1	13·0	8·8	11·1	
Calculated frequency (Varying probability)	29·4	16·8	15·4	11·8	11·3	

The X^2 for this table works out as 0·009 for one degree of freedom and is not significant. Thus the character (above average or below average) of a month is not 'affected' by the character of the preceding month. This result is also confirmed by calculating the serial correlation of the first order for the 720 monthly departures. This procedure takes into account both the sign and the magnitude of the departures. The correlation works out as 0·021 which is not significant.

3. Frequency derived from probability theory

Having established that there is no correlation between the amounts of rainfall in successive months we can now proceed to the mathematical treatment of the problem. Applying the theory of probability Beer and others have derived theoretical relationship between n the length of the sequence and $\log F$ where F is the frequency of run whose length is at least n (*i.e.*, F is the integrated frequency). They showed that if p and q be the probability of a month being wet and dry respectively the chance of a sequence of n wet months bounded by dry months is

$$p_{F_n} = q^2 p^n \qquad (1)$$

The corresponding frequency would be $q^2 p^n T$ where T is the total number of months. The probability of a sequence at least n months is

$$P_{F_n} = \sum_{s=n}^{\infty} p_{F_n} = \sum_{s=n}^{\infty} q^2 p^s \rightarrow \frac{q^2 p^n}{q} = q\, p^n$$

and the corresponding frequency $F = q\, p^n\, T$ Thus

$$\log F = \log q + \log T + n \log p \ldots . (2)$$

Starting, therefore, from the theory of random distribution of events, we arrive at the logarithmic law connecting the integrated frequency linearly with n, length of sequence.

From the classification given in the contingency table it is seen that the probability of A (*i.e.*, month with rainfall above average) is

$$p = A/(A+B) = \frac{261}{700} = .373$$

$$q = \frac{439}{700} = .627$$

The frequencies derived from equation (2) and the actual integrated frequency are given below in Table 4. In applying formula (2) T has been taken as 720. The frequencies of B sequences are derived by an interchange of p and q.

TABLE 4

Cumulative frequencies

	Length of run of at least											
	1	2	3	4	5	6	7	8	9	10	11	12 month
A												
Observed	168	63	19	7	3	1	1	1	1			
Constant probability	167·3	61·4	22·5	8·3	3·0	1·1	0·4	0·1				
Variable probability	163·2	62·6	24·6	9·6	3·7	1·4	0·5	0·2	0·1			
B												
Observed	170	106	64	40	26	13	6	5	4	3	3	1
Constant probability	167·3	105·9	67·0	42·4	26·8	17·0	10·7	6·8	4·3	2·7	1·7	1·0
Variable probability	163·5	102·9	66·3	43·2	27·6	17·4	10·7	6·4	3·7	2·0	1·0	0·4

The theoretical lines are drawn in Fig. 1 and the agreement between observation and theory is satisfactory. The theory was then completely tested by calculating the individual frequencies from equation (1). These along with the observed frequencies are given in Table 3 (a) and compare well with them.

Another conclusive test of absence of correlation was performed by rearranging (or stirring) the material. The odd and even months were considered separately and two different A and B sets of frequencies were obtained and then pooled together. These rearranged frequencies compare favourably with the observed frequencies—particularly the A sequences. Thus the "stirring up" process has not altered the frequencies and is, therefore, an evidence of the random distribution of the A's and B's.

4. Frequencies derived from varying probability

In view of the agreement between theory and observation as set out above it was thought worthwhile to make a more detailed approach in which the probability of a month being A or B would not be constant for all the months but would vary with the month. These probabilities can be calculated by considering the number of times (out of 60) each of the 12 months becomes A or B. The M-class being only ten was allocated between A or B in these calculations. We have,

	Jan	Feb	Mar	Apr	May	Jun
A	0·28	0·18	0·25	0·38	0·38	0·45
B	0·72	0·82	0·75	0·62	0·62	0·55

	Jul	Aug	Sep	Oct	Nov	Dec
A	0·43	0·40	0·48	0·45	0·45	0·30
B	0·57	0·60	0·52	0·55	0·55	0·70

Denoting the probabilities of A by $p_1, p_2 \ldots\ldots\ldots p_{12}$ and of B by $q_1, q_2 \ldots\ldots q_{12}$ we proceed to derive the analogue of equations (1) and (2) for the case when the monthly probabilities are variable.

Taking a simple case in which there are two probabilities p_1 and p_2 we have the probability of a run of one month of A (rainfall above average)

$$q_2^2 p_1 + q_1^2 p_2 = P_1 \text{ say,}$$

of two months of A

$$= q_1 p_2 p_1 q_2 + q_2 p_1 p_2 q_1 = 2 p_1 p_2 q_1 q_2 = P_2 \text{ say,}$$

of three months of A

$$= q_1 p_2 p_1 p_2 q_1 + q_2 p_1 p_2 p_1 q_2$$

$$= q_1^2 p_1 p_2^2 + q_2^2 p_1^2 p_2$$

$$= p_1 p_2 (q_1^2 p_2 + q_2^2 p_1)$$

$$= p_1 p_2 P_1$$

Similarly probability of a run of 4 A months would be $p_1 p_2 P_2$ and in general equal to $p_1 p_2 P_{r-2}$. Thus the probabilities of runs greater than 2 are expressible in terms of P_1 and P_2 (*i.e.*, the probabilities of run of 1 and 2 respectively).

For the case with 12 probabilities we have the probability of a run of

1 A month

$$= \sum_{s=1}^{12} q_s p_{s+1} q_{s+2} = P_1$$

2 A months

$$= \sum_{s=1}^{12} q_s p_{s+1} p_{s+2} q_{s+3} = P_2$$

12 A months

$$= \sum_{s=1}^{12} q_s p_{s+1} p_{s+2} \cdots p_{s+12} q_{s+13}$$

$$= P_{12}$$

(in the above equations

$$p_{12+b} = p_b \text{ and } q_{12+b} = q_b \text{)}$$

By the method of induction it can be proved that

$$P_r = (p_1p_2p_3\cdots\cdots p_{12})\, P_{r-12}\,. r > 12$$

Also summing, we get the probability of a run of at least one A month

$$= \sum_{s=1}^{12} P_s + (p_1p_2p_3\cdots p_{12}) \sum_{s=1}^{12} P_s$$

$$+ (p_1p_2\cdots p_{12})^2 \sum_{s=1}^{12} P_s + \cdots\cdots \text{infinity}$$

This series tends in the limit to

$$\frac{\sum_{s=1}^{12} P_s}{1 - p_1p_2p_3p_4\cdots\cdots\cdots p_{12}} \qquad (3)$$

The P's were calculated from the above equations and the frequencies corresponding to each run obtained by multiplying by 60 (the number of years). The frequencies of the sequences are obtained by interchanging the p's and q's. Since $p_1p_2p_3p_4p_5\cdots\cdots\cdots p_{12}$ in (3) is negligible, the integrated frequencies given in Table 3 are simply the addition of the individual frequencies corresponding to variable probability given in Table 2(a). The frequencies calculated from the variable probability do not compare with the observed as favourably as the frequencies coresponding to constant probability. The data were then examined in greater detail by breaking up the twelve months into the dry season (December-April) and the wet season (May-November) and computing the frequency of sequences of various lengths of A and B. The calculated frequencies were obtained from the Cochran's formula *

$$f_{r,m} = Np^r q(2 + q\,\overline{m-r-1})$$

which gives the frequency of length r out of m months in N years, when $r<m$ and p is the probability of a month being A ($p=0.438$ for wet season, $=0.285$ for dry season). When $r=m$ the formula $f_{r,m} = Np^r$ was used.

When the series is considered for the wet and dry seasons separately we find from the calculated values given in Tables 3 (b) and 3 (c), that in the case of B sequences the variable probability method gives a closer representation than the constant probability method. But, in the case of the A sequences the constant probability method gives a closer representation.

In conclusion it may be said that months with rainfall above and below average in Rayalaseema do not persist, *i.e.*, consecutive rainfall amounts are independent. The independence may be due, as Beer and others point out, to the length of the period considered, *i.e.*, a month. It is possible that shorter periods like a week or even a day show some "after effects" or persistence. Work along these lines is in progress.

8. Acknowledgement

The author's grateful thanks are due to all those who kindly helped him in collection, analysis and interpretation of data.

REFERENCES

Beer, A., Drummond, A. J. and Furth, R. (1946). *Quart. J.R. met. Soc.*, **72**, 311, p. 74.

Cochran, W.G. (1936). *J.R. Statist. Soc.*, suppl. 3, p. 49.

Gold, E. (1929). *Quart. J. R. met. Soc.*, **55**, 231, p. 307.

Newnham, E. V. (1916). *Quart. J. R. met. Soc.*, **42**, 179, p. 154.

* The Cochran's formula was employed since the A series (or B series) in each year was not continuous with the corresponding series of a subsequent year