NOTES AND CORRESPONDENCE

On the Conditional Distribution of Daily Precipitation Amounts

EDWIN H. CHIN AND JOHN F. MILLER

Office of Hydrology, National Weather Service, NOAA, Silver Spring, MD 20910

4 December 1978 and 11 April 1980

ABSTRACT

Possible conditional dependence of the distribution of daily precipitation amounts on the occurrence of precipitation on the preceding day is examined. Test results based on 25 years of data at 30 stations in the conterminous United States show that such dependence exists only in the winter season over the Pacific Northwest.

The use of the Markov chain model to represent the daily precipitation occurrence process has become a standard procedure. There have been relatively few attempts at modeling the conditional dependence of the daily precipitation amount. Katz (1977) proposed a chain-dependent model for daily precipitation. Denoting the amount of precipitation on the nth day by X_n, he made two modeling assumptions: (i) the X_n's are conditionally independent given the precipitation occurrence process on the nth day, and (ii) the distribution of X_n depends on whether the preceding day was rainy or not. However, in testing the conditional distribution of X_n with data from one station, he found that there was little difference between the two distribution functions.

Using mainly data from tropical stations, Buishand (1978) showed that the mean daily rainfall amount on a wet day was related to its position in a sequence of wet days. However, the F-test for the equality of means he used was valid only if the precipitation amounts were normally distributed. Buishand stated that nearly the same results were obtained using the cube roots of the data.

The objective of the present study is to examine the possible conditional dependence of the distribution of daily precipitation amounts using a larger data set. The data consisted of daily precipitation amount x in January–February and July–August for the period 1949 through 1973 at 30 National Weather Service stations from various climatic regions in the conterminous United States.

Let $F_i(x)$, $i = d, r$, denote the cumulative distribution function of daily precipitation amount x. The subscript i indicates whether the preceding day was dry (d) or wet (r). The null hypothesis under test is

$$H_0\colon F_d(x) = F_r(x) \quad \text{for all } x \tag{1}$$

against the alternate hypothesis

$$H_1\colon F_d(x) \neq F_r(x) \quad \text{for some } x. \tag{2}$$

Here, the x's are positive precipitation amounts from some continuous distribution. We also retain Katz's modeling assumption (i). To avoid making substantial assumptions regarding the unknown distribution functions of precipitation amounts in the process of hypothesis testing, it is proper to use nonparametric methods. Among the set of nonparametric tests, the Smirnov test is responsive to any kind of difference between the distribution functions from which the two samples were drawn.

The two-tailed Smirnov test has its test statistic D defined as

$$D_{d,r} = \mathrm{Max}\left|F_d(x) - F_r(x)\right|,$$

or the maximum vertical deviation between two (empirical) distribution functions. Tables are available that give critical values of the sampling distribution of D under the null hypothesis at various levels of significance α for different combinations of sample sizes.

The results of testing (1) using a two-tailed Smirnov test are shown in Table 1. Columns with significance level α exceeding 0.10 were obtained using an extrapolation procedure. Since the uncertainties in the extrapolation procedure increase with α, and an α value of 0.50 or greater indicates a complete lack of meaningful difference between the two samples, cases with $\alpha > 0.50$ were simply

TABLE 1. Two-tailed Smirnov test for H_0: $F_d(x) = F_r(x)$.

Station		n	$\bar{x}$	s	D^*	Significance level
January–February						
Boston, MA	x_d	317	0.364	0.461	0.1094	0.062
	x_r	263	0.277	0.358		
Philadelphia, PA	x_d	285	0.284	0.352	0.0397	>0.500
	x_r	222	0.282	0.325		
Columbia, SC	x_d	253	0.370	0.438	0.1118	0.094
	x_r	235	0.454	0.582		
Tampa, FL	x_d	210	0.385	0.513	0.0686	>0.500
	x_r	111	0.407	0.483		
Cleveland, OH	x_d	322	0.169	0.239	0.0808	0.170
	x_r	474	0.147	0.228		
Louisville, KY	x_d	302	0.292	0.388	0.0990	0.128
	x_r	260	0.363	0.435		
Birmingham, AL	x_d	273	0.426	0.476	0.0563	>0.500
	x_r	270	0.540	0.782		
Peoria, IL	x_d	250	0.174	0.360	0.0571	>0.500
	x_r	184	0.203	0.303		
Baton Rouge, LA	x_d	261	0.457	0.576	0.0588	>0.500
	x_r	207	0.522	0.646		
Shreveport, LA	x_d	237	0.395	0.511	0.0551	>0.500
	x_r	219	0.411	0.531		
Minneapolis, MN	x_d	226	0.111	0.178	0.0345	>0.500
	x_r	166	0.097	0.123		
Des Moines, IA	x_d	225	0.152	0.226	0.0554	>0.500
	x_r	153	0.146	0.258		
Springfield, MO	x_d	227	0.201	0.307	0.0912	0.350
	x_r	195	0.279	0.484		
Houston, TX	x_d	246	0.299	0.451	0.1113	0.112
	x_r	218	0.471	0.709		
Casper, WY	x_d	227	0.071	0.078	0.0239	>0.500
	x_r	137	0.069	0.071		
Denver, CO	x_d	174	0.111	0.143	0.0255	>0.500
	x_r	119	0.107	0.157		
Albuquerque, NM	x_d	114	0.096	0.104	0.0468	>0.500
	x_r	54	0.096	0.133		
Salt Lake City, UT	x_d	255	0.140	0.165	0.0548	>0.500
	x_r	197	0.149	0.177		
Idaho Falls, ID	x_d	240	0.069	0.125	0.0780	0.490
	x_r	226	0.089	0.121		
Boise, ID	x_d	252	0.093	0.119	0.1138	0.049
	x_r	332	0.134	0.160		
Spokane, WA	x_d	262	0.131	0.154	0.1185	0.023
	x_r	411	0.172	0.172		
Olympia, WA	x_d	195	0.206	0.226	0.2342	0.000
	x_r	777	0.423	0.481		
Pendleton, OR	x_d	265	0.089	0.109	0.1605	0.001
	x_r	320	0.138	0.154		
Burns, OR	x_d	254	0.101	0.127	0.1353	0.014
	x_r	299	0.157	0.183		

TABLE 1. (*Continued*)

Station		n	$\bar{x}$	s	D^*	Significance level
January–February						
Eugene, OR	x_d	232	0.274	0.351	0.1762	0.000
	x_r	624	0.441	0.558		
Medford, OR	x_d	250	0.160	0.231	0.1257	0.016
	x_r	399	0.264	0.354		
Eureka, CA	x_d	213	0.318	0.363	0.1284	0.012
	x_r	549	0.464	0.546		
San Francisco, CA	x_d	191	0.284	0.371	0.1600	0.004
	x_r	322	0.430	0.496		
Fresno, CA	x_d	183	0.222	0.295	0.1276	0.098
	x_r	184	0.284	0.349		
Los Angeles, CA	x_d	136	0.401	0.567	0.0878	>0.500
	x_r	132	0.502	0.691		
July–August						
Boston, MA	x_d	292	0.289	0.387	0.0649	>0.500
	x_r	189	0.353	0.743		
Philadelphia, PA	x_d	267	0.372	0.511	0.0919	0.318
	x_r	184	0.540	0.741		
Columbia, SC	x_d	279	0.438	0.616	0.0998	0.112
	x_r	299	0.594	0.841		
Tampa, FL	x_d	329	0.462	0.659	0.0610	0.478
	x_r	482	0.526	0.764		
Cleveland, OH	x_d	295	0.343	0.453	0.0471	>0.500
	x_r	196	0.310	0.406		
Louisville, KY	x_d	277	0.386	0.559	0.0618	>0.500
	x_r	195	0.349	0.400		
Birmingham, AL	x_d	291	0.385	0.492	0.0657	>0.500
	x_r	288	0.448	0.606		
Peoria, IL	x_d	263	0.419	0.552	0.0700	>0.500
	x_r	156	0.487	0.627		
Baton Rouge, LA	x_d	301	0.454	0.598	0.0566	>0.500
	x_r	309	0.474	0.581		
Shreveport, LA	x_d	216	0.414	0.648	0.0738	>0.500
	x_r	142	0.431	0.609		
Minneapolis, MN	x_d	300	0.369	0.472	0.0578	>0.500
	x_r	181	0.374	0.520		
Des Moines, IA	x_d	278	0.377	0.499	0.0596	>0.500
	x_r	182	0.369	0.538		
Springfield, MO	x_d	253	0.347	0.483	0.0692	>0.500
	x_r	168	0.440	0.673		
Houston, TX	x_d	241	0.411	0.504	0.0504	>0.500
	x_r	209	0.480	0.818		
Casper, WY	x_d	207	0.121	0.160	0.0284	>0.500
	x_r	100	0.126	0.204		
Denver, CO	x_d	259	0.159	0.250	0.0293	>0.500
	x_r	189	0.199	0.338		
Albuquerque, NM	x_d	267	0.156	0.200	0.0308	>0.500
	x_r	194	0.149	0.218		

TABLE 1. (*Continued*)

Station		n	$\bar{x}$	s	D*	Significance level
			July–August			
Salt Lake City, UT	x_d	153	0.143	0.243	0.2080	0.019
	x_r	83	0.244	0.347		
Idaho Falls, ID	x_d	127	0.154	0.236	0.0955	>0.500
	x_r	65	0.140	0.144		
Boise, ID	x_d	79	0.093	0.150	0.2133	0.218
	x_r	34	0.154	0.193		
Spokane, WA	x_d	143	0.107	0.122	0.1175	>0.500
	x_r	65	0.148	0.212		
Olympia, WA	x_d	134	0.136	0.153	0.1215	0.147
	x_r	257	0.213	0.252		
Pendleton, OR	x_d	97	0.083	0.097	0.1490	>0.500
	x_r	44	0.128	0.136		
Burns, OR	x_d	101	0.122	0.204	0.2228	0.126
	x_r	38	0.161	0.190		
Eugene, OR	x_d	70	0.129	0.195	0.1642	0.398
	x_r	53	0.220	0.279		
Medford, OR	x_d	61	0.130	0.160	0.1847	>0.500
	x_r	30	0.188	0.241		
Eureka, CA	x_d	87	0.060	0.110	0.2363	0.200
	x_r	27	0.183	0.264		
San Francisco, CA	x_d	18	0.044	0.066		
	x_r	1	0.020	0.000		
Fresno, CA	x_d	10	0.050	0.075		
	x_r	1	0.010	0.000		
Los Angeles, CA	x_d	18	0.051	0.057		
	x_r	2	0.015	0.007		

* Observed.

designated as such. The July–August cases for San Francisco, Fresno and Los Angeles were deleted because the sample sizes were too small to be meaningful. Results were considered significant if the test indicated a significance level < 0.05. Otherwise, the null hypothesis was not rejected. From Table 1, it was evident that hypothesis (1) is, in general, valid. For most stations, there was essentially no significant difference between the distribution functions $F_d(x)$ and $F_r(x)$. However, there were some notable exceptions. For example, for Salt Lake City, the July–August case indicated rejection of the null hypothesis. But additional tests on precipitation data at neighboring stations did not show a similar effect. This suggested that the summer singularity at Salt Lake City might have been caused by sampling fluctuations.

The most important exceptions occurred in January–February in the Pacific Northwest, covering Washington, Oregon, western Idaho and northern California. As the semi-permanent Aleutian low intensifies and moves southward in the fall and winter, this region is frequently visited by frontal storms coming in from the Pacific Ocean. These disturbances may become so frequent in winter that successive storms overlap one another. Therefore, the total precipitation for a day preceded by a rainy day could be significantly greater than if preceded by a dry day. It should be noted that such a phenomenon is not limited to the coastal zone but extends past the Cascade Range to stations such as Spokane and Pendleton, and further eastward over the Blue Mountains to western Idaho. The greatest difference between the distribution functions $F_d(x)$ and $F_r(x)$ occurred in the winter at Olympia, Washington.

Compared to the well-confirmed Markov dependence nature of daily precipitation occurrences, the evidence of conditional dependence of daily precipitation amounts on previous precipitation occurrence is less substantial. The present study used 25 years of daily precipitation data at 30 stations in the contiguous United States, from which it is concluded that except in the winter season in the Pacific Northwest, the distribution of daily precipitation amounts does not depend on whether the preceding day was wet or dry.

Acknowledgment. We are grateful to Gerald F. Cotton of NOAA, Air Resources Laboratories, for valuable discussions.

REFERENCES

Buishand, T. A., 1978: Some remarks on the use of daily rainfall models. *J. Hydrol.*, **36**, 295–308.

Katz, R. W., 1977: Precipitation as a chain-dependent process. *J. Appl. Meteor.*, **16**, 671–676.