POSSIBLE INHOMOGENEITY OF THE VARIANCE OF THE FREEZE HAZARD DISTRIBUTION OVER ARIZONA

PAUL C. KANGIESER

Weather Bureau State Climatologist, Phoenix, Ariz.

[Manuscript received June 5, 1961; revised December 1, 1961]

ABSTRACT

Previous investigations indicate that the freeze date variance varies slowly with geographical factors over parts of the United States. Evidence is presented here which indicates that the variation may be greater over the State of Arizona.

1. INTRODUCTION

In 1958, Thom and Shaw [1] carried out a climatological analysis of freeze data for Iowa. They defined a "*t*-degree freeze" as " . . . the occurrence of a minimum temperature of $t°$ or lower." The last date of occurrence in the spring and the first date of occurrence in the fall were investigated, using thresholds of 32°, 28°, 24°, 20°, and 16° F. They arrived at a number of interesting and useful conclusions concerning the distribution of these dates. They found, for example, that for the State of Iowa, the distribution of these dates could be considered normal, and that, for a given station and threshold temperature, the spring series could be considered independent from the fall series. Tests on Arizona data indicate that both properties are also true here.

Thom and Shaw also found that the Iowa data showed very little difference, from station to station, in the variance of the dates. This was true for both the spring and fall series, and for each threshold. Frederick, Johnson, and MacDonald [2] found in New York that it was even possible to pool variances for all thresholds and use the same average value for each threshold. This property of homogeneity, even within threshold groups, seems to be largely lacking in Arizona data, particularly in the spring series. It is emphasized at the outset, however, that if this is true, there is no reason to question any of the conclusions drawn by Thom and Shaw in their excellent study. Geographically, Iowa and Arizona are very different. The range of elevation in Iowa, for example, is about 1,200 feet while in Arizona it is about 12,000 feet. While elevation may not account for all of the variation, there is some evidence that it may be a factor. The purpose of this paper is to present the results of statistical tests for homogeneity of variance within each of five temperature threshold groups and for the elevation factor in Arizona.

2. DATA

Several years ago, the National Weather Records Center in Asheville, N.C. computed the mean and variance of the freeze hazard distribution for selected stations in all States. The thresholds used were 32°, 28°, 24°, 20°, and 16°F., and both the first date of occurrence in the fall and the last date of occurrence in the spring were included. The data used in this study were essentially the above data for Arizona, with the following exclusions and changes: (1) only stations with at least 15 years of record were included, (2) only stations for which a given threshold temperature occurred during every year were used, and (3) the beginning of "fall" was counted as August 1 instead of July 1. The latter change affected only three items in the original tabulation; namely, the 32° threshold at Alpine and the 32° and 28° thresholds at Fort Valley. The change was made because the large "break" between dates of occurrence of these thresholds in the summertime usually comes between the end of July and early in September, at these two high-elevation stations.

3. TEST FOR HOMOGENEITY OF VARIANCE

As mentioned earlier, it was found (in agreement with Thom and Shaw's results for Iowa) that there was little reason to doubt that the Arizona spring and fall freeze dates are independent. Also, each appears to be normally distributed. To test for homogeneity of variance within each threshold group, a number of tests are available. The test chosen was Bartlett's,[1] which has good power against a variety of alternatives ([3] and [4]). The hypothesis tested is that the variances of k normally distributed

[1] In this study Bartlett's test was used because the number of degrees of freedom (number of years of record) varied widely among stations. If a sufficiently large number of stations with the same length of record is available, a much simpler test (Cochran's test) may be used (see [1]).

populations are equal. The samples may be of any size, but should be independent. The samples are of size n_i, where $\Sigma n_i = N$. The variance of the ith sample is S_i^2. Let

$$M = (N-k) \ln S_p^2 - \Sigma[(n_i-1) \ln S_i^2] \tag{1}$$

$$S_p^2 = \frac{\Sigma(n_i-1)S_i^2}{N-k} \tag{2}$$

$$A = \frac{1}{3(k-1)}\left[\Sigma\left(\frac{1}{n_i-1}\right) - \frac{1}{N-k}\right] \tag{3}$$

$$f_1 = k-1 \tag{4}$$

$$f_2 = \frac{k+1}{A^2} \tag{5}$$

$$b = \frac{f_2}{1-A+(2/f_2)} \tag{6}$$

The sampling distribution of $F = f_2M/[f_1(b-M)]$ is approximately $F(f_1, f_2)$. F was computed from the observed k samples. A 5 percent level of significance was chosen. If the observed F is larger than $F\ (f_1, f_2)$, as read from standard tables of the F-distribution, then we shall reject at the 5 percent level of significance the hypothesis that the k populations have equal variances. The results are shown in table 1. All F values for the spring thresholds are significantly larger than $F\ (f_1, f_2)$ at the 5 percent level. While the fall values are rather large, only one is large enough to be significant.

We have assumed in making the tests that the k samples are independent, which is not true in this case. The effect of dependence between samples is to lower the value of k to an "effective" value, say k_i. If the mean correlation between the series of dates at all stations (all possible pairs considered) is $\bar{r}$, then k_i may be estimated from the formula

$$k_i = \frac{k}{1+(k-1)\bar{r}} \tag{7}$$

In order to get a rough idea of the value of $\bar{r}$ for the 32° spring threshold, values of the linear correlation coefficient were computed using the 10 possible pairs of combinations for the five stations: Phoenix, Prescott, Flagstaff, Tucson, and Winslow.

The individual values of r ranged from 0.32 to −0.32 with an overall average of 0.02. The five stations are

TABLE 1.—*Results of test for homogeneity of variance within each freeze hazard threshold group*

	Spring					Fall				
Threshold (° F.)	32	28	24	20	16	32	28	24	20	16
F	3.11	2.86	2.18	2.41	2.20	1.38	1.51	1.15	1.41	2.26
$F(f_1,f_2)$	1.44	1.52	1.60	1.69	1.94	1.47	1.57	1.62	1.69	1.94
k	34	25	19	15	9	30	21	18	15	9
$\bar{r}$ (critical)	0.55	0.42	0.15	0.21	0.10	------	------	------	------	0.12

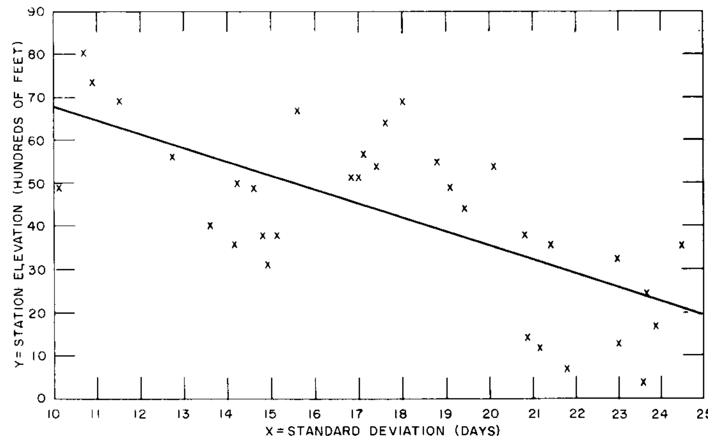

FIGURE 1.—Relationship between standard deviation (days) of the date of last occurrence in spring of a temperature of 32° F. or less and station elevation (hundreds of feet). All cases from NWRC summarization for Arizona stations with 15 or more years of record and for which threshold temperature occurred every year. (WBAS Tucson omitted.)

widely separated, so it seemed advisable to make similar computations for nearby stations. The two closest stations in the north are Flagstaff and Fort Valley, for which $r=0.39$. The closest pair in the south are Bisbee and Douglas Smelter, for which $r=0.08$. It can be shown, using formula (7) and a table for the F-distribution that the value of $\bar{r}$ would have to be greater than 0.55 for the 32° spring threshold to become non-significant. The values of $\bar{r}$ which, if exceeded, would make the other thresholds non-significant are shown in table 1 in the row labeled "$\bar{r}$ (critical)."

Because of the nature of the data, it is impossible to reach firm conclusions as a result of the statistical tests. However, for the spring thresholds at least, the validity of the assumption of equal variance appears questionable enough that pooling them seems inadvisable.

4. TEST FOR ELEVATION FACTOR

It was mentioned above that there is some evidence that elevation may be a factor in producing these differences. Figure 1 is a graph of the station elevation against the standard deviation of the last date in the spring with a temperature of 32° F. or lower. The straight line was fitted by least squares. If the coordinate variables were unrelated the slope of this line would be zero. By making some assumptions about the data, we can test whether or not the slope of the line is different from zero by using the statistic

$$t = \frac{(b-0)S_x\sqrt{N-1}}{S_{yx}} \tag{8}$$

where N is the sample size, b is the sample value of the slope of the regression line, S_{yx} is the so-called "standard error of estimate," and S_x is the sample standard deviation of the variable X. If the distribution of Y for each X is normal, with the same variance and the same mean, then the sampling distribution of the above statistic is a t-distribution with $N-2$ degrees of freedom.

In this case $N=34$, so we will reject (at the 5 percent level) the hypothesis that the slope of the line is zero if t is greater than 2.04 or less than -2.04. The sample value of t was -5.26. This is a highly significant value, since it would still be significant if the "effective" degrees of freedom were reduced from 32 to 2, or the "effective" number of cases reduced from 34 to 4 to account for dependence in the data.

5. CONCLUSIONS

In States where it has not been demonstrated that "freeze"-date variances are relatively uniform, the homogeneity of the variances should be tested before using the average as a common value for all stations. The need for such tests may be especially important in States where rugged terrain produces large elevation differences between stations.

REFERENCES

1. H. C. S. Thom and R. H. Shaw, "Climatological Analysis of Freeze Data for Iowa," *Monthly Weather Review*, vol. 86, No. 7, July 1958, pp. 251–257.
2. R. H. Frederick, E. C. Johnson, and H. A. MacDonald, "Spring and Fall Freezing Temperatures in New York State," *Cornell Miscellaneous Bulletin 33*, New York State College of Agriculture, 1959.
3. W. J. Dixon and F. J. Massey, Jr., *Introduction to Statistical Analysis*, McGraw-Hill Book Co., Inc., 1st Ed., 1951, pp. 90–91.
4. G. E. P. Box, "A General Distribution Theory for a Class of Likelihood Criteria," *Biometrika*, vol. 36, 1949, pp. 317–346.