SHORTER CONTRIBUTION

551.501.45

ACCURACY OF MONTHLY MEAN TEMPERATURES COMPUTED FROM DAILY READINGS IN WHOLE DEGREES

By J.V. Maher

Central Office, Bureau of Meteorology, Melbourne
(Manuscript received February 1967)

The problem: If daily temperatures are reported telegraphically in whole degrees, is it statistically correct to mean these in lots of 30 or 31 and to quote the answer to the nearest .1°F

It will be assumed that normal rounding off procedures are followed, i.e. when the tenths digit is 1, 2, 3 or 4 the rounding off is downward, for 6, 7, 8 and 9 it is upward, and for 5 the rounding is to the odd, that is, half upward, half downward. Thus there is a probability of 1/10 for tenths digits of 0, 1, 2, 3, 4, 6, 7, 8 and 9 and 1/20 for 5 (up) and 5 (down).

If the effect in tenths of a degree of the occurrence of 0, 1 .... 9 in rounding off is considered, we have the following probabilities of occurrence:

-4, -3, -2, -1, -0. +1, +2, +3, +4 - (all 1/10)
+5 - (1/20)
-5 - (1/20)

The standard deviation of a population such as this is 2.915 tenths of a degree.

It follows from the central limit theorem that the distribution of a series of mean values of samples of 30 from this population is approximately normal, with the same mean (zero) and standard deviation

$$\frac{2.915}{\sqrt{30}} = .532 \text{ (tenths of a degree)} = .0532 \text{ degrees.}$$

Assuming normality, 68 per cent of the means of the samples of 30 will be in error by .0532 or less. The standard normal tables also show that 82.6 per cent of values can be expected to be +0.05 or less and thus 65 per cent between -0.050 and +0.050. These values would round off to zero.

Similarly, taking 0.15 as the boundary between errors of 0.1 and errors rounded off to 0.2 or more, we find from standard normal tables that 99.76 per cent of errors may be expected to be less than +0.2 (rounded off to the nearest tenth of a degree) and thus 99.52 per cent between -0.2 and +0.2, leaving .48 per cent greater than 0.2 regardless of sign. The maximum error would be ± 0.5°F but would be extremely rare.

As a rough check on the above, from a sample of 540 temperatures 234 sets of 30 were selected at random and the means of each set were calculated on a 1004 Data Processor with and without rounding off to the nearest degree, and the errors examined. These results are compared in Table 1 with the theoretical results obtained above.

Table 1. Comparison of the distribution of errors in means of 30 temperatures obtained from practical tests and theoretical considerations.

Error in mean	Zero	± 0.1°F	±0.2°F or more
	%	%	%
Theoretical	66	33½	½
Practical	60	38½	1½

There were certain faults in the practical experiment. For example, the distribution of decimal digits in the sample of temperatures was not as specified. They were distributed as shown in Table 2.

Table 2. Distribution of decimal digits in the sample temperature data

Decimal Digit	0	1	2	3	4	5 odd	5 even	6	7	8	9	Total
Frequency	50	44	60	60	59	18	24	71	55	49	50	540

Using the same 234 means of 30 values, the means of each of the groups of 30 whole number temperatures were rounded off to the nearest degree and compared with the unrounded means of the same (but unrounded) groups of 30 temperatures. The errors are shown in Table 3.

Table 3. Distribution of errors involved in using rounded-off means of groups of 30 rounded-off temperatures.

Error (Tenths of ^{o}F)	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
Frequency (per cent)	1.7	6.8	8.5	9.0	10.2	10.7	11.2	8.5	8.1	9.0	11.2	4.7	.4

The maximum error would be 1.0^{o}F but would be extremely rare, it theoretical probability being one in $2^{29} \times 10^{31}$ for a sample of 30 temperatures.

Conclusion: The distribution of errors justifies the quotation of monthly means of daily whole degree values of temperature in tenths of a degree in preference to whole degrees.