M.O. 232e.

FOR OFFICIAL USE.

METEOROLOGICAL OFFICE, LONDON.

PROFESSIONAL NOTES NO. 5.

ON THE USE OF THE NORMAL CURVE OF ERRORS IN CLASSIFYING OBSERVATIONS IN METEOROLOGY.

BY

CAPTAIN E. H. CHAPMAN, R.E.

Published by the Authority of the Meteorological Committee.

LONDON:
To be purchased from
THE METEOROLOGICAL OFFICE, Exhibition Road, London, S.W.7

1919.

Price 6d. Net.

ON THE USE OF THE NORMAL CURVE OF ERRORS IN CLASSIFYING OBSERVATIONS IN METEOROLOGY.

By Captain E. H. Chapman, R.E.

It is not possible to eliminate completely the element of chance[1] in dealing with meteorological observations. Accordingly when judging as to whether an observation is unusual or not unusual due regard ought to be paid to the fluctuations which may arise from chance, and from chance alone.

The theory of chance fluctuations attains its greatest simplicity when each observation dealt with is entirely independent of every other, and when the classification is but two-fold.

To take a common and extremely simple example, suppose four pennies are tossed and the resulting number of heads noted. Suppose that they are tossed a large number of times. Then in any one throw of the four pennies we can get either 0, 1, 2, 3 or 4 heads. We should not be surprised if we got four heads, but we should not expect to get this number of heads as often as we should expect to get two heads. In fact the number of times 0, 1, 2, 3, 4 heads would occur in the long run would be proportional to the successive terms of the binomial expansion $(\frac{1}{2} + \frac{1}{2})^4$, *i.e.*, to 1, 4, 6, 4, 1.

The chance of a head at any one throw of a penny is $\frac{1}{2}$. The mean number of heads in a throw of four pennies would tend towards 2. The standard deviation of the number of heads in the four throws would tend towards $\sqrt{\frac{1}{2} \cdot \frac{1}{2} \cdot 4}$,

i.e., to 1

The range of the number of heads in the four throws is from

0 to 4,

i.e., within the limits 2 ± 2,

or $m \pm 2\sigma$,

where m is the mean, and σ the standard deviation of the number of heads in the four throws.

In general, using the conventional terms, if p be the chance of a success in an event, and q the chance of a failure, so that $p+q=1$, then in a series of sets of n events the mean number of successes would tend towards pn, and the standard deviation of the number of successes in the n events would tend towards σ, where $\sigma^2 = pqn$. It is known that most of the observations in cases of this sort lie within the range $m \pm 3\sigma$, where m is the mean, and σ the standard deviation. Unless an observation lies well outside this range we cannot be sure that it has not arisen from chance, and from chance alone.

Examples of the fluctuations due to chance in their simplest form can be obtained from Meteorology. In the Meteorological Office Calendar for 1916 the number of times each particular day of the year was rainless at Kew[2] during the 35 years 1881–1915 is

---

[1] See M.O. 223, V., pp. 13–14.

[2] It should be noted that Kew Observatory referred to in this paper is not at Kew but at Richmond.

ON THE USE OF THE NORMAL CURVE OF ERRORS IN CLASSIFYING OBSERVATIONS IN METEOROLOGY.

By Captain E. H. Chapman, R.E.

It is not possible to eliminate completely the element of chance[1] in dealing with meteorological observations. Accordingly when judging as to whether an observation is unusual or not unusual due regard ought to be paid to the fluctuations which may arise from chance, and from chance alone.

The theory of chance fluctuations attains its greatest simplicity when each observation dealt with is entirely independent of every other, and when the classification is but two-fold.

To take a common and extremely simple example, suppose four pennies are tossed and the resulting number of heads noted. Suppose that they are tossed a large number of times. Then in any one throw of the four pennies we can get either 0, 1, 2, 3 or 4 heads. We should not be surprised if we got four heads, but we should not expect to get this number of heads as often as we should expect to get two heads. In fact the number of times 0, 1, 2, 3, 4 heads would occur in the long run would be proportional to the successive terms of the binomial expansion $(\frac{1}{2} + \frac{1}{2})^4$, *i.e.*, to 1, 4, 6, 4, 1.

The chance of a head at any one throw of a penny is $\frac{1}{2}$. The mean number of heads in a throw of four pennies would tend towards 2. The standard deviation of the number of heads in the four throws would tend towards $\sqrt{\frac{1}{2} \cdot \frac{1}{2} \cdot 4}$,

i.e., to 1

The range of the number of heads in the four throws is from

0 to 4,

i.e., within the limits 2 ± 2,

or $m \pm 2\sigma$,

where m is the mean, and σ the standard deviation of the number of heads in the four throws.

In general, using the conventional terms, if p be the chance of a success in an event, and q the chance of a failure, so that $p+q=1$, then in a series of sets of n events the mean number of successes would tend towards pn, and the standard deviation of the number of successes in the n events would tend towards σ, where $\sigma^2 = pqn$. It is known that most of the observations in cases of this sort lie within the range $m \pm 3\sigma$, where m is the mean, and σ the standard deviation. Unless an observation lies well outside this range we cannot be sure that it has not arisen from chance, and from chance alone.

Examples of the fluctuations due to chance in their simplest form can be obtained from Meteorology. In the Meteorological Office Calendar for 1916 the number of times each particular day of the year was rainless at Kew[2] during the 35 years 1881–1915 is

---

[1] See M.O. 223, V., pp. 13–14.

[2] It should be noted that Kew Observatory referred to in this paper is not at Kew but at Richmond.

given. Thus January 1st was rainless 20 times out of the 35, January 2nd 15, and so on for each day of the year. In these data we have a two-fold classification "rainless day" or "rain day," and we have observations which are totally independent since the rainfall for any particular date of a certain year is unaffected by the rainfall for the same date in any other year.

Consider these Kew data for the 31 days of January. The number of times each day of the month was rainless in the 35 years is in order:—

20, 15, 15, 15, 15, 20, 19, 13, 14, 16, 12, 21, 19, 23, 16, 18, 16, 20, 16, 7, 20, 19, 23, 18, 18, 18, 13, 16, 15, 17, 18.

By addition we have 525 rainless days in the 35 Januaries. Hence the chance p that a day in January will be rainless at Kew is given by $p = \frac{525}{31 \times 35} = \frac{15}{31}$.

The chance q that a day should be a rain day can be got from the relationship $p + q = 1$ which gives $q = \frac{16}{31}$.

We should expect then that any date in January would be rainless on the average $\frac{15}{31} \times 35 = 17$ times in 35 years. The standard deviation here is $\sqrt{\frac{15}{31} \cdot \frac{16}{31} \cdot 35} = 3$. The numbers of rainless days for each date of January at Kew should therefore practically all lie between 17 ± 9, *i.e.*, between 8 and 26. There is only one number outside this range, 7 on the 20th, and this is only just outside the range so that it cannot be said that it is not due to chance fluctuations. At first sight one might have been inclined to think that there was something unusual in the difference between the 14th and 20th of January at Kew. The 14th was rainless 23 times in the 35 years, the 20th was rainless only 7 times. The difference between the 23 and the 7, though considerable, may have arisen entirely through chance fluctuations.

Another illustration of this kind can be given from the same Calendar. At Falmouth the last six days of April were sunless in the 35 years 1881–1915 respectively 5, 5, 3, 1, 6, 4 times. The average for these six days is 4. For the six days then the average chance of a sunless day is p, where $p = \frac{4}{35}$. As before we have $q = \frac{31}{35}$. The standard deviation $\sigma = \sqrt{\frac{4}{35} \cdot \frac{31}{35} \cdot 35} = 2$. The range for the six days is from 1 to 6, which is covered by 4 ± 3, *i.e.*, by $m \pm 1{\cdot}5\sigma$, where m is the mean. Hence there is nothing unusual in the difference between the 28th and 29th of April for which the numbers of sunless days in 35 years are 1 and 6 respectively.

Descriptive words such as unusual, very unusual, etc., are frequently used in Meteorology to describe either single observa-

(21347—12.) Wt. 40659—S.O.P. 521. 600. 5/19. D & S. G. 3.

tions or means of a number of observations. For example, the practice of the Meteorological Office[3] in characterising the weekly values of mean temperature, rainfall, and sunshine for the several districts of the British Isles is shown in the following table:—

TABLE I.

Number of times the value occurs on the average out of 12 possible occasions.	Descriptive adjective for		
	Temperature.	Rainfall.	Sunshine.
1	Very unusual	Very heavy	Very abundant
3	Unusual	Heavy	Abundant
4	Moderate	Moderate	Moderate
3	Deficient	Light	Scanty
1	Very deficient	Very light	Very scanty

This table, however, is not used outside the Weekly Weather Report,[4] and there does not appear to be anything similar to it in use elsewhere in Meteorology.

Summarising this table, and replacing the word moderate by the words not unusual, we have that a value is not unusual if it occurs once out of three possible occasions, whereas a value is unusually large on the one hand, or unusually small on the other hand, if it occurs once out of four possible occasions. A value is described as very unusual (either excess or defect) if it occurs only once out of 12 possible occasions.

If we apply this method of classification to the measurement of stature[5] of 8,585 adult males born in the British Isles we get the following result:—

Very small	Small	Moderate	Tall	Very tall

5′3½″ 5′6½″ 5′8½″ 5′11″

Height in feet and inches.

To take a second example, barometer readings[6] at 9h. at Southampton for the years 1878–1890, 4,748 observations, the method of classification of Table I. gives:—

Very low	Low	Moderate	High	Very high

1000 1010 1020 1030

Millibars.

The second example will show that the classification of Table I. is not intended for general use. The terms as defined in this table form an extremely simple and convenient classification of mean weekly values, a classification which is easily applied and easily understood.

The probability values given in Table I. are subject to chance fluctuations from one period of years to another. Take the case

[3] Weekly Weather Report.

[4] An account of the method of and reasons for the selection of these limits is given by R. G. K. Lempfert in the Journal of the Board of Agriculture, Vol. xiv., p. 1.

[5] Yule, Theory of Statistics, p. 88.

[6] „ „ „ „ p. 96.

of the moderate values which occur on the average four times out of 12 possible occasions. For such values we have $p=\frac{1}{3}$, $p=\frac{2}{3}$. Suppose that this value of p is obtained from a period of 18 years. The standard deviation σ would be given by $\sigma = \sqrt{\frac{1}{3} \cdot \frac{2}{3} \cdot 18} = 2.$ If other periods of 18 years were available the range which would cover practically all the numbers of times the moderate values would occur is $6 \pm 3 \times 2$,

$i.e.$, 0 to 12.

Thus a value of $p=\frac{1}{3}$ in one period of 18 years might become equal to anything from 0 to $\frac{2}{3}$ in another period of 18 years from fluctuations due to chance alone.

The difficulty in fixing the value of p lies in the fact that 18 years is too short a period. In experiments such as coin-tossing we have a theoretical value for p, but in most meteorological work a theoretical value for p cannot be obtained. All we can do is to use the observed value, and this only becomes reliable when the number of observations is very large.

Let $p = \frac{1}{3}$, $q = \frac{2}{3}$ again, but let the period from which p is determined be 72 instead of 18 years (18 was selected in order to get a simple numerical value for σ). The standard deviation for $n = 72$ would be 4. If other periods of 72 years were taken the range of the numbers of times the value would occur would be $\frac{1}{3} \times 72 \pm 3 \times 4$,

$i.e.$, 24 ± 12,

$i.e.$, 12 to 36.

Thus p would vary from $\frac{1}{6}$ to $\frac{1}{2}$ for other periods of 72 years. Increasing the number of years from 18 to 72 has decreased the probable range for p from 0 to $\frac{2}{3}$

to $\frac{1}{6}$ to $\frac{1}{2}$.

It will be seen, therefore, that a table such as Table I., even if based on as long a period as 72 years would by chance fluctuations be liable to alteration for another period of 72 years.

In the general case where p is the chance of a success in n events, and q the chance of a failure, $p + q = 1$, the frequencies of 0, 1, 2, ——, n successes in N trials each consisting of n events are given by the successive terms of the binomial expansion $N(q + p)^n$ Thus when p is known not only can we find the mean pn, and the standard deviation $\sqrt{pqn}$, but we can give a theoretical expression for the frequency distribution of the chance fluctuations.

The binomial expansion $N(q + p)^n$ gives $n + 1$ ordinates corresponding to 0, 1, 2, ——, n successes in the n events. If we join the tops of these ordinates we get a frequency polygon or frequency distribution. Now for all values of p and q smooth continuous curves can be fitted to these frequency distributions. Such curves are called frequency curves.

In the symmetrical case[7] for which $p = q = \frac{1}{2}$ the frequency curve is the very important curve, known as the normal curve of errors, or law of errors.

[7] See Yule, Theory of Statistics, p. 301

Up to the present we have dealt entirely with cases of simple two-fold classification. In Meteorology, however, the various elements are more frequently given according to a numerical scale. When this is the case we can calculate the arithmetic mean of the element for the period of time to which the observations refer. We can then calculate the standard deviation as the root-mean-square deviation instead of from the formula $\sqrt{pqn}$. We cannot obtain directly a theoretical frequency distribution corresponding to the binomial expansion $N(q+p)^n$, but we can group the observations and form the actual frequency distribution. To this frequency distribution we can fit a curve. When the variations in the observations of the element are due to causes similar to those which produce chance fluctuations, this frequency curve will take the form of the normal curve of errors. M. Angot,[8] in his "Études sur le climat de France," has shown that variations in mean temperature follow the normal law. W G. Reed,[9] in a paper entitled "Frost in the United States," has shown that variations in the dates of last killing frosts in spring and first killing frosts in autumn follow the normal law. No doubt variations of other meteorological elements follow the same law. We are, therefore, justified in adopting the normal curve of errors as a basis for the classification of observations expressed on a numerical scale.

In a normal distribution, if m be the mean and σ the standard deviation, 68 per cent. of the observations lie within the range $m \pm \sigma$. Outside the range $m \pm 2\sigma$ we have 4·6 per cent. of the observations, or 2·3 per cent., on each side. Outside the range $m \pm 3\sigma$ we should have in all 0·3 per cent. of the observations, or more accurately ·135 per cent. on each side. All this is expressed in the form of a diagram in Figure I., the curve being the normal curve of errors.

It appears from this that a most convenient classification of observations for general use would be to name those *not unusual* which lie within the range $m \pm \sigma$. Those lying outside this range could be named *unusual*. Observations outside the range $m \pm 2\sigma$ may be termed *very unusual*, while those lying outside the range $m \pm 3\sigma$ may be termed *exceptional*. Accurate decimal fractions for this system of classification would be:—

Exceptionally deficient ...	·00135	of the observations.
Very unusually deficient ...	·02140	
Unusally deficient	13591	
Not unusual	·68268	
Unusually excessive	13591	
Very unusually excessive ...	·02140	
Exceptionally excessive ...	·00135	

These figures are only strictly true for a normal distribution, but for other forms of distribution they would hold more often than not provided the number of observations were large.

[8] See M.O. 223, V., pp. 6–7.
[9] Proc. Second Pan-Amer., Sci. Cong. Wash., 1917.

Up to the present we have dealt entirely with cases of simple two-fold classification. In Meteorology, however, the various elements are more frequently given according to a numerical scale. When this is the case we can calculate the arithmetic mean of the element for the period of time to which the observations refer. We can then calculate the standard deviation as the root-mean-square deviation instead of from the formula $\sqrt{pqn}$. We cannot obtain directly a theoretical frequency distribution corresponding to the binomial expansion $N(q+p)^n$, but we can group the observations and form the actual frequency distribution. To this frequency distribution we can fit a curve. When the variations in the observations of the element are due to causes similar to those which produce chance fluctuations, this frequency curve will take the form of the normal curve of errors. M. Angot,[8] in his "Études sur le climat de France," has shown that variations in mean temperature follow the normal law. W. G. Reed,[9] in a paper entitled "Frost in the United States," has shown that variations in the dates of last killing frosts in spring and first killing frosts in autumn follow the normal law. No doubt variations of other meteorological elements follow the same law. We are, therefore, justified in adopting the normal curve of errors as a basis for the classification of observations expressed on a numerical scale.

In a normal distribution, if m be the mean and σ the standard deviation, 68 per cent. of the observations lie within the range $m \pm \sigma$. Outside the range $m \pm 2\sigma$ we have 4·6 per cent. of the observations, or 2·3 per cent., on each side. Outside the range $m \pm 3\sigma$ we should have in all 0·3 per cent. of the observations, or more accurately ·135 per cent. on each side. All this is expressed in the form of a diagram in Figure I., the curve being the normal curve of errors.

It appears from this that a most convenient classification of observations for general use would be to name those *not unusual* which lie within the range $m \pm \sigma$. Those lying outside this range could be named *unusual*. Observations outside the range $m \pm 2\sigma$ may be termed *very unusual*, while those lying outside the range $m \pm 3\sigma$ may be termed *exceptional*. Accurate decimal fractions for this system of classification would be:—

Exceptionally deficient ...	·00135	of the observations.
Very unusually deficient ...	·02140	
Unusally deficient	·13591	
Not unusual	·68268	
Unusually excessive	·13591	
Very unusually excessive ...	·02140	
Exceptionally excessive ...	·00135	

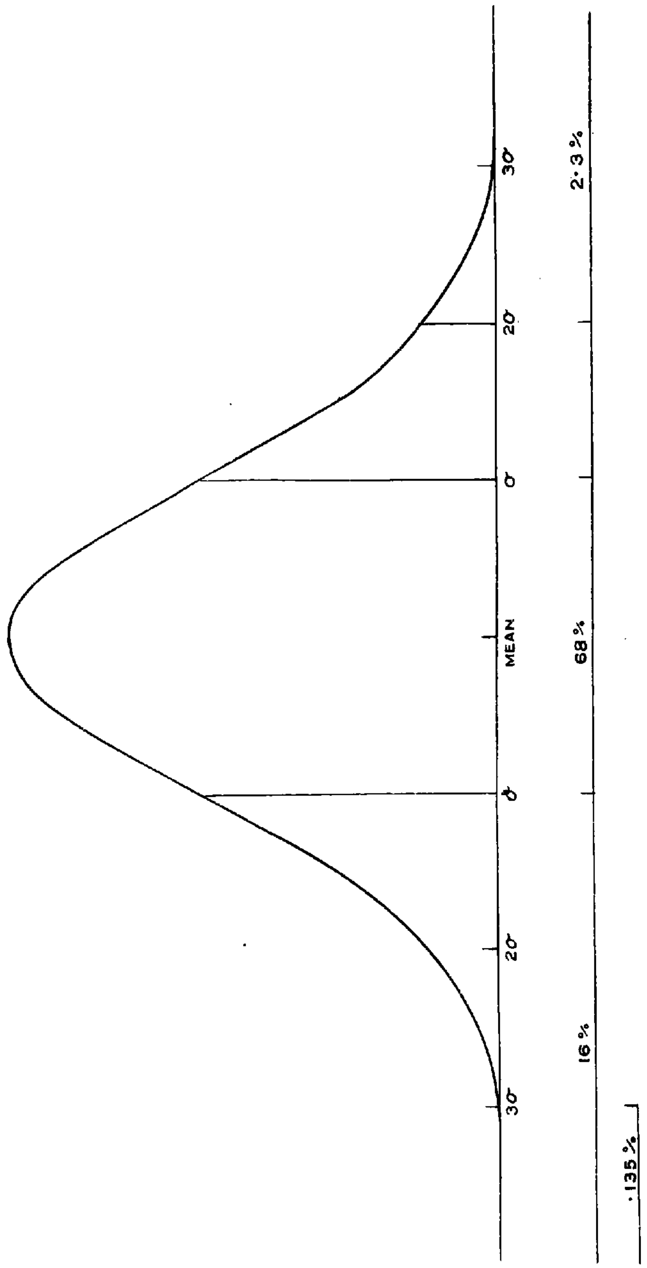
These figures are only strictly true for a normal distribution, but for other forms of distribution they would hold more often than not provided the number of observations were large.

---

[8] See M.O. 223, V., pp. 6–7.
[9] Proc. Second Pan-Amer., Sci. Cong. Wash., 1917.

THE NORMAL CURVE OF ERRORS.

THE NORMAL CURVE OF ERRORS.



WOOLACOMBE.

SEASON	BAROMETER 21h PREVIOUS TO RAIN-FALL. INS.	NUMBER OF OBSER-VATIONS.	RAINFALL 24 HOURS 9h.—9h. ACCORDING TO HEIGHT OF BAROMETER AT 21h. PREVIOUS EVENING. MEAN INS.	σ INS.
YEAR	30·00 OR ABOVE	563	·04	·11
YEAR APRIL -SEPT.	29·75 29·99 BELOW 29·75	612	·10	·16
OCT. -MARCH	BELOW 29·75	286	·15	·21

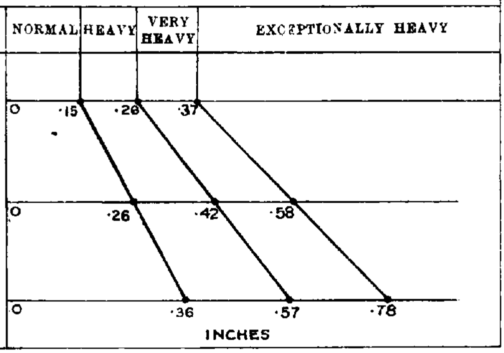


Figure II.

WOOLACOMBE.

WIND	NUMBER OF OBSER-VATIONS.	RAINFALL 24 HOURS 9h.-9h. GROUPED ACCORDING TO WIND DIRECTION NIGHT BEFORE 21h.-9h. MEAN INS.	σ INS.
N.W. & N.E.	192	·03	·08
S.W. & S.E.	228	·09	·11

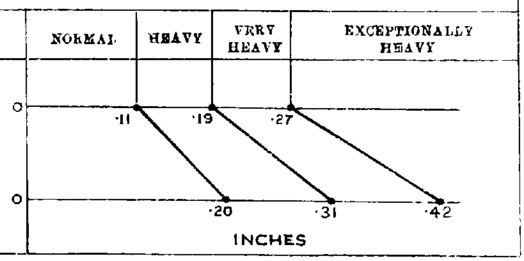


Figure III.

Wyman & Sons, Ltd., M.O. Press, London, S.W. 7. Ps. 4410. 40659. S.O.P. 521. 600 4/19

As chance figures we should have on the average for this method of classification—

UNUSUALLY DEFICIENT 1 in 6		NOT UNUSUAL 2 in 3		UNUSUALLY EXCESSIVE 1 in 6
VERY UNUSUALLY DEFICIENT 1 in 44				VERY UNUSUALLY EXCESSIVE 1 in 44
EXCEPTIONALLY DEFICIENT 1 in 741				EXCEPTIONALLY EXCESSIVE 1 in 741

Our suggestion is that where numerical values of the element are given the mean and standard deviation should be calculated, and the classification determined by $m \pm \sigma$, $m \pm 2\sigma$, $m \pm 3\sigma$. When numerical values of the element are not given, or when it is not desired to work from them, the suggested classification can be got sufficiently accurately from Table II.

TABLE II.

Descriptive Adjectives.	Observation occurs on the average.
Exceptionally deficient ...	1 out of 750 times.
Very unusually deficient ...	1 „ 50 „
Unusually deficient	1 „ 6 „
Not unusual	2 „ 3 „
Unusually excessive	1 „ 6 „
Very unusually excessive ...	1 „ 50 „
Exceptionally excessive ...	1 „ 750 „

It may be pointed out that if Table II. were in use exceptional values would not occur for many meteorological elements such as mean monthly temperatures for a particular calendar month, since observations are not available for 750 years. When the arithmetic mean and standard deviation (root-mean-square deviation) are known, however, exceptional values are possible even if the number of observations is small. Take, for example, September rainfall for London[10] for the 25 years 1888–1912. The mean is 1·68 ins., the standard deviation 1·04 ins. An exceptionally heavy rainfall would be one greater than

$$1{\cdot}68 + 3 \times 1{\cdot}04 \text{ ins.},$$

i.e., greater than 4·80 ins. Such a value occurred in 1896, the September rainfall for that year being 5·43 ins. It will be seen from this example that there is an advantage in working from the numerical values of the element rather than from a table like Table II.

We shall now see what results can be obtained by applying the classification $m \pm \sigma$, $m \pm 2\sigma$, $m \pm 3\sigma$ to a few selected examples.

[10] M.O. 223 V., p. 20.

As a first example we can refer to a previous paper[11] by the present writer in which the rainfall at Woolacombe, North Devon, was under consideration. The observations for four years were divided into two groups (i) summer months April to September, (ii) winter months October to March. The rainfall considered was the amount for the 24 hours 9 h. to 9 h. The two seasonal groups were subdivided into three groups according to the height of the barometer at 21 h. the night preceding the 24 hours' rainfall. The sub-groups in which mean and standard deviation are approximately equal are combined, and the ranges of rainfall calculated from the values $m + \sigma$, $m + 2\sigma$, $m + 3\sigma$ are shown in Figure II. It is to be noted that in this and the next figure the standard deviation in every case was greater than the mean so that no rainfall was *not unusual*.

The observations for Woolacombe for the two years 1904–5 were also grouped according to wind direction during the 12 hours 21h. to 9h. previous to the 24 hours 9h. to 9h. for which the rainfall was given. From the means and standard deviations the observations resolved themselves into two wind groups (i) S.W. and S.E., (ii) N.W. and N.E. The classification of the rainfall is given in Figure III.

As a second example consider the monthly totals of rainfall at Kew and Valencia, means and standard deviations for which have been calculated by the present writer[12] from data for the 47 years 1869–1915. The classification based on $m \pm \sigma$, $m \pm 2\sigma$, $m \pm 3\sigma$ is given in Figure IV for Kew, and Figure V. for Valencia. The complete difference in type between the monthly rainfall totals at Kew and Valencia is well shown by these two diagrams. There are many striking features about Figures IV. and V., but most of them are well known to meteorologists. Some of the irregularities in the diagrams are no doubt due to the unequal lengths of the calendar months. It is not proposed to discuss here the details of these diagrams. The suggestion may be made that rainfall maps showing for each month exceptionally heavy values of rainfall for the British Isles would be interesting and instructive. Such maps would help in the study of rainfall.

For the sake of comparison the monthly rainfall totals for the year 1916 are indicated on the diagrams of Figures IV and V.

There is one extremely valuable use to which the standard deviation σ can be put when the distribution is normal. Suppose that we have a particular deviation x from the mean. Then from the value of $\frac{x}{\sigma}$ we can find the probabililty that a deviation as large as, or larger than, x should occur. Such probability values can be obtained from the tables[13] of the probability integral $F = \frac{1}{2}(1+a)$.

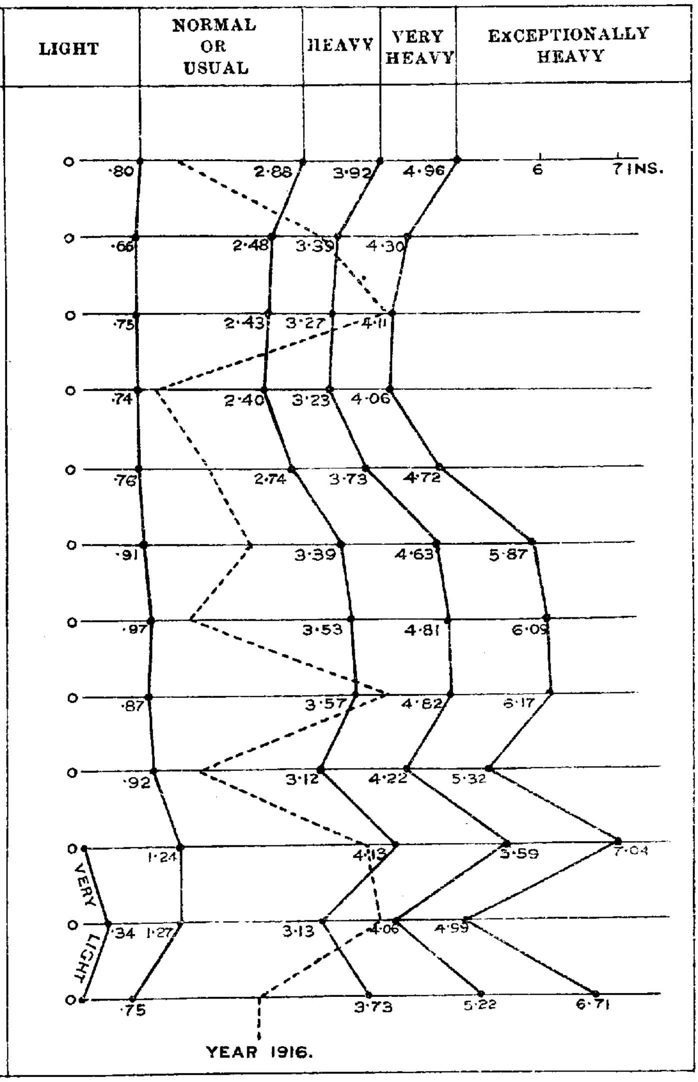
A most interesting use of such probability values is to be found in the paper by W. G. Reed already mentioned[9] In this paper the author gives the mean and standard deviation of the date

---

[11] Barometric Changes and Rainfall. Journal Roy. Met. Soc., October 1915.
[12] Atmospheric Pressure and Rainfall. Journal Roy. Met. Soc., October, 1916.
[13] Tables for Statisticians and Biometricians, Camb. Univ. Press, p. xvii and p. 2.

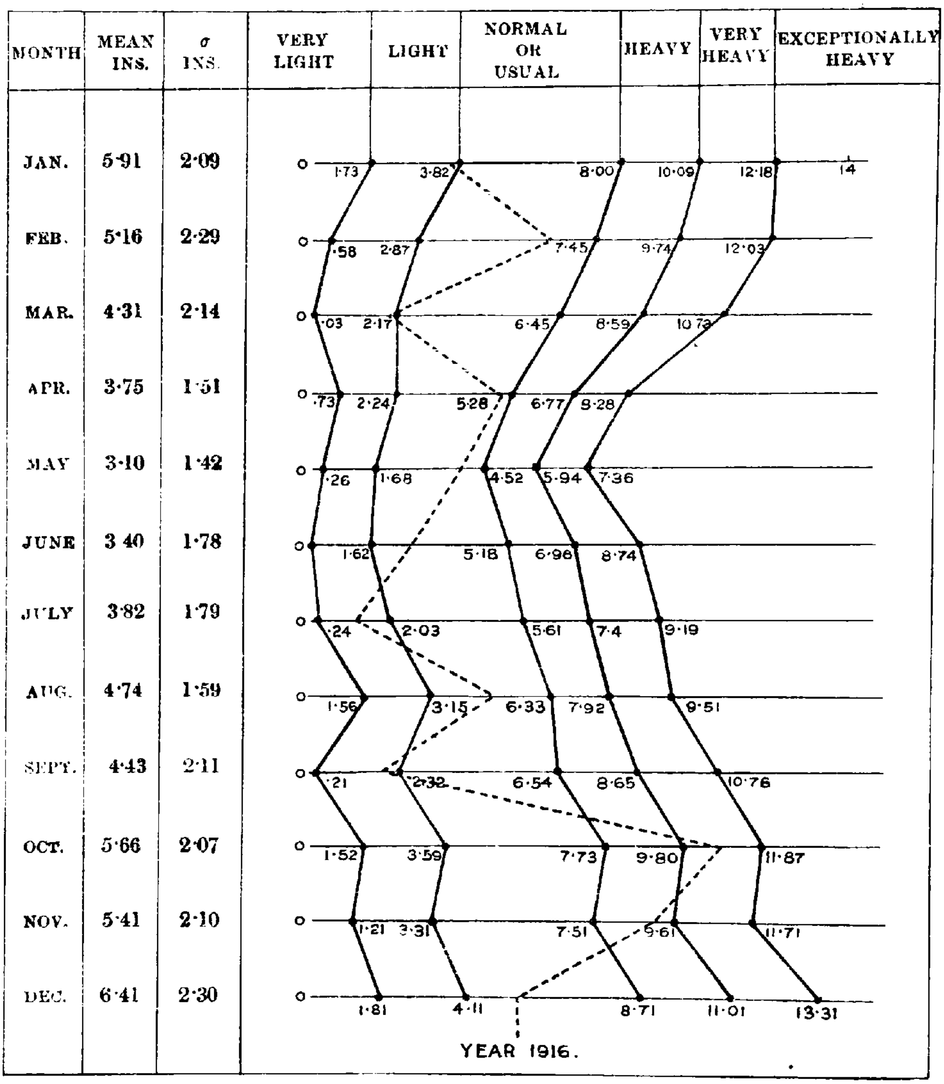
KEW MONTHLY RAINFALL. INCHES.

MONTH	MEAN INS.	σ INS.
JAN.	1·84	1·04
FEB.	1·57	0·91
MAR.	1·59	0·84
APR.	1·57	0·83
MAY	1·75	0·99
JUNE	2·15	1·24
JULY	2·25	1·28
AUG.	2·22	1·35
SEP.	2·02	1 10
OCT.	2·69	1·45
NOV.	2·20	0·93
DEC.	2·24	1·49



Wyman & Sons, Ltd., M.O. Press, London, S.W. 7. Fs. 4419 40C59. S.O.P. 521. 600. 4/19

VALENCIA MONTHLY RAINFALL. INCHES.



of (i) the last killing frost in spring, and (ii) the earliest killing frost in autumn for 569 stations in the United States. The author assumes a normal distribution for dates of killing frosts. On this assumption he tabulates certain particular values of $\frac{x}{\sigma}$ and the corresponding probability values. For example, if A_s is the average date of the last killing frost in spring, and σ_s the standard deviation of that date, the chances are 1 in 15 against a killing frost occurring after the date $A_s + 1{\cdot}5\sigma$.

The author gives the actual dates of the last killing frost in spring, and the earliest killing frost in autumn for 40 years for a station called Bismarck (Burleigh County, N. Dakota). Mean dates and standard deviations are given. Applying the classification $m \pm \sigma$, $m \pm 2\sigma$, $m \pm 3\sigma$, suggested in the present paper to these data for Bismarck we get the following results:—

LAST KILLING FROST IN SPRING.

Occurred in 40 years 0	0	3	31	4	2	0
Exceptionally early	Very early	Early	Not unusual	Late	Very late	Exceptionally late
...........10	11...20	21......1	2......21	22......1	2......10	11..............
	April.		May.		June.	

FIRST KILLING FROST IN AUTUMN.

Occurred in 40 years 0	1	2	32	3	2	0
Exceptionally early	Very early	Early	Not unusual	Late	Very late	Exceptionally late
...........16	17...27	28......8	9......30	1......11	12.. 22	23..............
	August.		September.		October.	

It would be interesting to see similar results obtained from the phenological reports of the Royal Meteorological Society.

The classification suggested in this paper depends on the arithmetic mean m, and the standard deviation σ. The usual formula for the probable error of the arithmetic mean is $\frac{\cdot 6745\sigma}{\sqrt{n}}$, where n is the number of observations. It is assumed in this formula that the n observations are uncorrelated amongst themselves. For a normal distribution the probable error of the standard deviation σ is $\frac{\cdot 6745\sigma}{\sqrt{2n}}$. For other distributions the probable error of σ takes a more complicated form. For a normal distribution we can write the probable error of m and σ respectively $f_1\sigma$ and $f_2\sigma$ where $f_1 = \frac{\cdot 6745}{\sqrt{n}}$ and $f_2 = \frac{\cdot 6745}{\sqrt{2n}}$.

The calculation of these probable errors is facilitated by the use of Table V., pp. 12–18 of Tables for Statisticians and Biometricians. In this table the values of f_1 and f_2 are given from $n = 1$ to $n = 1000$.

In the system of classification determined by $m \pm \sigma$, $m \pm 2\sigma$, $m \pm 3\sigma$ it is easy to see that a fluctuation in the arithmetic mean m, from one set of observations to another, would move the whole scale of classification, numerically speaking, to the right or left by an amount equal to the fluctuation. A fluctuation in the standard deviation σ would alter the separate parts of the classification scale. The *not unusual* part would be altered by the fluctuation of σ by an amount twice as large as the amount any other portion of the scale would be altered. Fluctuations in the values of m and σ from one set of observations to another cannot be avoided. The only thing that can be done to minimise them is to make the number of observations as large as possible.