[13] In admitting the probability that selection, in this case, has operated primarily upon a visible character, color, I am adopting a viewpoint somewhat different from that expressed by me only recently (Carnegie Institution Yearbook for 1928).

[14] Cf. Sumner, F. B., *Ecology*, **6**, 1925, 352–371, where evidence has been presented for believing that the primary cause of the depigmentation of desert animals has been the climatic factor, though the need for concealment (through selection) has probably accelerated the process in certain cases.

NOTE ON C. S. PEIRCE'S EXPERIMENTAL DISCUSSION OF THE LAW OF ERRORS

By Edwin B. Wilson and Margaret M. Hilferty

Department of Vital Statistics, Harvard School of Public Health

Read before the Academy November 19, 1928

Incident to the preparation of the official biography of C. S. Peirce for this Academy the senior author came across Peirce's experimental discussion of the law of errors[1] wherein the conclusion was that the normal law was, on the whole, verified. This is in accord with the dictum of Poincaré[2] that everybody believes in the law of errors: the mathematicians because they think it empirically demonstrated by experimenters and the experimenters because they think the mathematicians have proved it *a priori*. The series of observations given by Peirce is long, consisting of about 500 records each day for 24 different days of the time elapsed between the making of a sharp sound and the record of reception of the sound by an observer. According to our previous experience such long series of observations generally reveal marked departures from the normal law, and it seems interesting to examine this material of Peirce's from the point of view of the modern theory of frequency functions.[3] Accordingly we have entered in the tables, (1) the median with its standard deviation, (2) the mean with its standard deviation, (3) the semi-interquartile range, (4) two-thirds of the standard deviation, (5) the ratio of probable error as defined by the semi-interquartile range to the probable error as defined by two-thirds or more accurately 0.6745 of the standard deviation, (6) the arithmetic mean error about the median as origin, (7) the number of negative errors greater than 3 σ and the number of positive errors of similar magnitude and their sum, (8) the number of observations within 0.25 σ of the mean and the number expected according to the normal law, (9) the percentage excess of the observed over the expected number, (10) the second moment μ_2 and its standard deviation, (11) the third moment μ_3 and its standard deviation, (12) the Pearsonian constant $\sqrt{\beta_1} = \mu_3/\sigma^3$

(Charlier's measure of skewness), (13) the Pearsonian skewness *sk*, (14) Yule's measure of skewness as the sum of the quartiles less twice the median divided by the semi-interquartile range, (15) the Pearsonian kurtosis $\beta_2 - 3$, (16) the Pearsonian frequency type, with the appended sign + meaning that the sampling error of the fourth moment is infinite and with ++ meaning that the third moment also has an infinite sampling error for samples drawn from the universe defined by the curve.

The following observations may be made on the table:

1. There is on each of the 24 days a greater concentration of small errors in the vicinity of the mean or median than the normal law, fitted with the computed value of the standard deviation, predicts. This may be seen in a number of ways: (*a*) the quotient of the probable error defined in column (3) divided by the probable error in column (4), given in column (5), is uniformly less than 1, varying from a maximum of 0.98 which presumably does not depart significantly from 1 to 0.51 which does greatly depart therefrom, with a mean or median value of about 0.765 which cannot in any way be reconciled with the theoretical value 1. (*b*) The arithmetic mean error in column (6) is never so high as 0.8 σ as required by the normal curve but varies from 0.76 σ to 0.50 σ about a median value of 0.7 σ. (*c*) The number of observations within 0.25 σ of the mean on either side as observed in column (8) is uniformly in excess of the theoretical number by anywhere from 4 to 94%.

2. There is a considerable excess of large errors as compared with the normal law. For example, the normal law would indicate on the basis of 500 observations that one error should exceed 3.1 σ. In column (7) the observed numbers are entered. There is only one day on which the number of errors above 3.1 σ is less than 3 and on five days there are 8. The mean number of negative errors greater than 3.1 σ is 1.7 instead of 0.5 and the mean number of positive errors greater than 3.1 σ is 3.9 instead of 0.5, making the mean value of the total 5.6 instead of 1. When factors of safety of 3 to 8 are observed, can one say that the theoretical distribution is substantiated? When one calculates the ratio of the largest error to the value of σ for each day one finds numbers running up to over 15; and such errors are inconceivably rare on the basis of the normal law.

As a matter of fact, is it not generally true of errors of observation that they exhibit an excess of small and an excess of large errors as compared with the normal law?

3. The curves exhibit positive skewness. Just how to formulate the precise amount or even the sign of this phenomenon is difficult to say. If the ratio $\beta = \mu_3/\sigma^3$ is taken, column (12), we know that according to the books the value of β for samples drawn from any normal universe is $\sqrt{6/n} = 0.11$ when $n = 500$. There is only one value of β lying between -0.11 and $+0.11$, and the values run up to 10.9 with a median of 1. Two

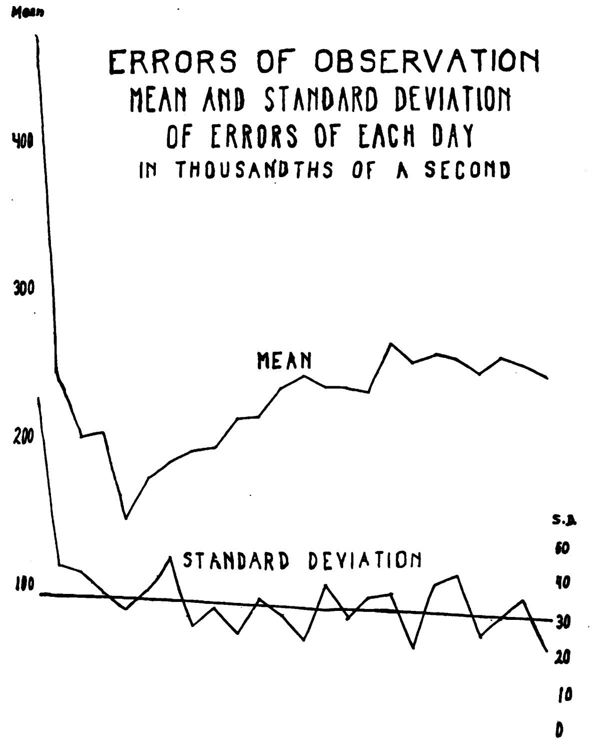
values are negative and 22 are positive. On the basis of Pearson's skewness, column (13), of which the standard deviation is 0.055, there are 4 values less than 0.055 and 20 values greater; there are 10 values less than

	1	2	3	4	5	6	7			8	
			PROBABLE ERROR				ERRORS $>3.1\sigma$			NO. WITHIN 0.25σ OF THE MEAN	
DAY	MEDIAN	MEAN	$\frac{Q_3-Q_1}{2}$	0.6745σ	$\frac{Q_3-Q_1}{2(0.6745\sigma)}$	A.D.	NEG.	POS.	TOTAL	OBS'D.	EXPECTED
1	468.0 ±3.0	475.6 ±4.2	57.9	62.1	0.932	70.1	1	3	4	110	98
2	236.8 ±2.4	241.5 ±2.1	26.6	31.6	0.842	35.7	1	0	1	113	97
3	199.9 ±1.8	203.1 ±2.0	27.5	30.4	0.905	33.7	0	7	7	113	97
4	201.3 ±1.2	205.6 ±1.8	19.4	26.6	0.730	26.9	1	7	8	134	99
5	147.0 ±2.1	148.5 ±1.6	21.4	23.5	0.912	26.7	0	4	4	110	97
6	172.4 ±1.8	175.6 ±1.8	20.0	26.9	0.744	28.0	0	6	6	119	97
7	184.0 ±1.6	186.9 ±2.2	24.6	32.7	0.753	32.3	0	6	6	132	98
8	193.9 ±1.2	194.1 ±1.4	17.2	20.5	0.840	22.0	2	6	8	120	97
9	195.4 ±1.4	195.8 ±1.6	17.9	23.7	0.756	24.0	2	4	6	132	98
10	215.1 ±1.5	215.5 ±1.3	16.4	19.3	0.850	21.0	2	1	3	120	99
11	213.1 ±1.8	216.6 ±1.7	19.7	25.2	0.782	25.9	1	5	6	135	99
12	232.7 ±1.9	235.6 ±1.7	17.0	22.4	0.759	23.4	3	5	8	103	78
13	244.4 ±1.3	244.5 ±1.2	16.3	17.7	0.922	19.9	6	1	7	101	97
14	236.3 ±1.2	236.7 ±1.8	14.7	27.8	0.529	21.6	2	3	5	192	99
15	253.3 ±1.0	236.0 ±1.4	14.3	21.6	0.662	20.8	4	4	8	162	98
16	233.3 ±1.4	233.2 ±1.7	15.6	25.5	0.612	21.7	4	2	6	162	98
17	264.0 ±1.7	265.5 ±1.7	20.6	26.0	0.792	27.6	3	5	8	123	100
18	254.0 ±1.3	253.0 ±1.1	15.8	16.5	0.959	18.6	0	4	4	114	98
19	255.0 ±0.9	258.7 ±1.8	13.9	27.7	0.502	20.4	1	3	4	187	99
20	252.6 ±1.3	255.4 ±2.0	15.2	29.2	0.521	21.6	0	3	3	179	98
21	245.2 ±1.6	245.0 ±1.2	14.6	18.5	0.790	19.2	3	4	7	120	99
22	255.0 ±1.5	255.6 ±1.4	14.8	21.5	0.688	19.0	2	4	6	142	99
23	252.0 ±1.1	251.4 ±1.6	15.0	24.6	0.610	19.8	0	3	3	158	98
24	244.2 ±0.9	243.4 ±1.1	11.6	15.9	0.730	16.8	3	3	6	113	98
							1.7	3.9	5.6		

	9	10	11	12	13	14	15	16
					SKEWNESS			
DAY	EXCESS OBS. OVER EXPECTED	$\mu_2 \pm \sigma_{\mu_2}$	$\mu_3 \pm \sigma_{\mu_3}$	CHARLIER'S μ_3/σ^3	PEARSON'S sk	YULE'S	$\beta_2 - 3$	PEARSONIAN TYPE
	%							
1	12	8475 ±856	918,117 ±312,000	1.18	0.417	+0.107	3.1	IV
2	16	2197 ±168	41,338 ± 14,000	0.43	0.158	+0.192	0.9	IV
3	16	2046 ±219	100,824 ± 37,000	1.09	0.314	−0.029	3.6	IV
4	35	1552 ±237	111,274 ± 64,000	1.82	0.407	+0.108	9.7	IV+
5	13	1210 ± 99	14,511 ± 11,800	0.39	0.126	+0.252	1.3	IV
6	22	1596 ±209	91,972 ± 38,600	1.48	0.375	0	6.4	IV+
7	35	2353 ±547	336,594 ±248,000	2.96	0.592	+0.114	24.9	IV++
8	23	926 ±103	13,856 ± 12,000	0.48	0.100	+0.070	4.1	IV++
9	35	1233 ±220	73,081 ± 59,000	1.71	0.268	−0.056	13.8	IV++
10	21	818 ±120	18,569 ± 23,600	0.84	0.129	−0.012	8.8	IV++
11	36	1394 ±214	83,461 ± 48,900	1.69	0.352	+0.152	9.8	IV++
12	32	1108 ±143	22,133 ± 17,000	0.63	0.118	+0.312	4.7	IV++
13	4	692 ± 67	4,894 ± 5,800	−0.22	−0.046	+0.037	2.6	IV+
14	94	1702 ±616	401,764 ±207,000	5.74	1.595	−0.034	63.6	VI+
15	65	1030 ±252	55,666 ± 25,000	1.68	0.222	−0.042	27.9	IV++
16	65	1426 ±614	345,877 ±328,000	6.39	1.445	+0.006	90.6	VI+
17	23	1486 ±165	14,176 ± 23,700	0.21	0.039	+0.126	4.3	IV++
18	16	596 ± 52	3,879 ± 3,700	0.27	0.080	−0.139	1.8	IV+
19	89	1689 ±910	758,187 ±512,000	10.94	112.0	+0.288	143.9	I
20	81	1884 ±818	632,595 ±427,000	7.71	3.496	+0.006	91.4	VI
21	21	752 ±107	4,724 ± 16,000	0.23	0.030	−0.055	8.2	IV++
22	43	1024 ±383	171,371 ±156,000	5.27	1.08	−0.148	68.1	VI++
23	61	1333 ±343	144,188 ±135,000	2.73	0.420	−0.220	31.1	IV++
24	15	559 ± 68	−342 ± 6,600	−0.02	−0.006	−0.241	5.4	IV++

$3\sigma = 0.165$ but 14 values greater. Still we have two negative and 22 positive values. On the basis of Yule's value, column (14), we find 10 negative, 1 zero and 13 positive values for the skewness; the comparison of the signs and numerical magnitudes in column (14) and in columns (12) and (13) is interesting.

4. The values of the kurtosis are impossible to reconcile with the notion of random sampling from a normal universe. For the value of $\beta_2 - 3$, of column (15), is, as ordinarily given, $0 \pm \sqrt{24/n} = 0 \pm 0.22$, and there is



no value of $\beta_2 - 3$ which is less than 4 times this amount; indeed, the observed values seem to run up almost indefinitely.

The failure of β to lie within the limits $\pm$ 0.11 and of $\beta_2 - 3$ to lie within the limits ± 0.22 justifies the statement that on no single day of the 24 was the distribution of errors such that its values of β and $\beta_2 - 3$ could have arisen by random sampling from a normal universe.

5. On account of the large kurtosis the usual formula for the standard deviation of the standard deviation, $\sigma_\sigma = \sigma\sqrt{2n}$, based on the normal curve cannot be applied; we must use $\sigma_\sigma = \sigma\sqrt{(\beta_2 - 1)/4n}$.

6. The third moments as calculated from the data are not reliable

because the sampling errors as calculated from the data, using sixth moments, are so large—see column (11).

7. The type of Pearsonian frequency curve which fits the data is IV on 19 of the 24 days, is VI on 4 days and I on one day. But it should be observed that on 18 of the 24 days the fitted Pearsonian curve has infinite eighth moments μ_8 and that therefore if the universe of errors of observation were really as thus defined on any one of these days the sampling errors of the values of μ_4 would be infinite. The erratic conduct of β_2-3, observed just above, is possibly fair evidence that in very truth the universe of errors of observation in the case of this whole series is heterotypic and that the values of μ_4 or β_2 are in no way to be depended upon. (In 12 of the 24 days the fitted Pearsonian curve has infinite sixth moments so that the reliability of the third moments must be nearly nil—and judging from the sampling errors as given in column (11) it is not alone in these 12 days that the moments are unreliable.)

8. The distributions cannot be represented by the Charlier A-type expansions using the third and fourth moments (but no higher moments) as computed. The departures from the normal curve are in fact so large that the A-type expansion becomes negative in some ranges of the variable less than 3σ removed from the mean in 20 of the 24 days.

The upshot of this all is that Peirce had observations which could show as completely as one might desire that the departures of the errors from the normal law was for his series uniformly great.

Let us turn from the frequency distribution to the statistical constants. The ordinary statement based on the normal law is that the determination of the median is 25% worse than that of the mean. A comparison of the standard deviations of the median and mean in columns (1) and (2) shows that for these observations the median is better determined than the mean on 13 days, worse determined on 9 days, and equally well determined on 2 days. Roughly speaking this means that mean and median are on the whole about equally well determined.

The changes in the mean from day to day are significant in the statistical sense. The value of the mean gives the correction for the "personal equation" of the observer for this type of observation. For scientific purposes in applying the corrections and in estimating the inexactness of the correction we must include in that estimate of inexactness all variations of the personal equation which are not subject to known correction. As the standard deviation of the mean is $\sigma/\sqrt{n}$ or about 1.7 we may not expect better exactness than this. However, the real (as yet uncontrolled and unexplained) variation of the mean is very much greater. Just how to estimate it from the data is impossible to say because there are obviously progressive changes in the mean. On the first day when the observer is "finding himself" the mean is very large; on the second day it drops to about the terminal value, but continues to drop for three days more; thereafter it rises to a maximum on the 17th day and then falls off.

The actual changes in the mean for the last half of the series are

$$+9,\ -8,\ -1,\ -3,\ +32,\ -12,\ +5,\ -3,\ -10,\ +11,\ -5,\ -8.$$

There is little to suggest anything but statistical fluctuations in these differences. It is evident that whatever may be the proper estimate of the standard deviation of the mean even during the last half of the series when the observer may be considered to have settled down, the value must very greatly exceed the 1.7 obtained by the formula $\sigma/\sqrt{n}$, i.e., a reasonably large Lexian ratio is involved. This illustrates the principle that we must have a plurality of samples if we wish to estimate the variability of some statistical quantity, and that reliance on such formula as $\sigma/\sqrt{n}$ is not scientifically satisfactory in practice, even for estimating unreliability of means.

With respect to the variation of the standard deviation from day to day it is clear that the first day is quite imcomparable with the subsequent days, that beginning with the second day there is a gradual reduction of the value of σ of total extent of perhaps 4 or 5 units in 23 days, superposed upon which there are fluctuations of very considerable size. The standard deviation of σ is indeed about 6.2. The formula $\sigma/\sqrt{2n}$ gives about 0.8 for σ_σ on the basis of a mean value $\sigma = 36$. But, as seen above, the kurtosis of the frequency curves is such as to make necessary the use of $\sigma \sqrt{(\beta_2-1)/4n}$ in place of $\sigma/\sqrt{2n}$. The root mean square of $(\beta_2 - 1)/2$ is 3.7 and if it be appropriate to use this as a factor we should give σ_σ the value 3.0 which still is less than half the observed value 6.2.

It appears, therefore, that the sampling theory is not applicable to the results of the identical experiment repeated on 24 days (when the first day is rejected) even when the analysis is confined to the last half of the series. The mean and standard deviation vary much more than is predicted.

[1] On the "Theory of Errors of Observations" by Assistant C. S. Peirce, *Rep. Super. U. S. Coast Survey* (for the year ending Nov. 1, 1870), Washington, Gov't. Printing Office, 1873, Appendix No. 21, pp. 200–224 and Plate No. 27.

[2] H. Poincaré, *Calcul des Probabilités*, chap. X, p. 171, **1912**.

[3] The only reference to the series that we have found in the literature is one by M. Fréchet, "Sur la loi des erreurs d'observation," *Rec. Soc. Math., Moscow*, **32**, 1924, in which the author states that at his instigation one of his pupils (Samama) had adjusted by the normal law and by the first Laplacian law $e^{-a|x|}$ several of Peirce's curves and had found that, though the two laws present satisfactory adjustments, the first gives the better result.