knife-edge $u_{1}$ comes in front of the vertical plane through $n_{1}$; and, since the lower point of attachment of the compensating spring $t_{1}$ is far below $n_{1}$, a couple is introduced which compensates for the greater upward force. The same is the case in the reverse order, when the lever is deflected upwards. Hence if the pull exerted by $t_{1}$ and the other conditions mentioned below be properly adjusted, the horizontal lever may be made to have any desired period of free oscillation. In actual practice some positive stability must be given to the lever in order that its position of equilibrium may be definite ; but its period may be made so great that, even if oscillations of considerable amplitude in its own period are set up, they will be so slow compared with those of the earthquake, that the undulating line so drawn will still be practically straight, so far as the earthquake record is concerned. In order to insure good compensation, the condition must be fulfilled that the rate of variation of the compensating couple is always the same as that of the supporting couple. If this be not the case, the pendulum must either be left with excessive positive stability for small deflections, or it will be continually liable to become unstable by the compensating couple becoming too great when the deflection exceeds a certain limit. In the present instance, let the modulus of the supporting spring be M , the arm at which it acts $a$; let the modulus of the compensating spring be $\mathrm{M}_{1}$, and the distance between $n_{1}$ and $u_{1}$ be $a_{1}$. Then for a deflection of the lever equal to $\theta$ we have, on the supposition that the length of the supporting spring and link is great compared with $a_{1}$, for the return couple $\mathrm{M} a^{2} \cos \theta \sin \theta-\mathrm{M}_{1} a_{1}{ }^{2} \cos \theta \sin \theta-\mathrm{M}_{1} \beta \sin \theta$, where $\beta+a_{1}$ is the total elongation of the spring for the horizontal position of the lever. Now our condition necessitates $\beta$ being either zero or negative; and in order to keep within this condition the length of the unstretched spring and link are made to reach a little above $n_{1}$, and the height of $u_{1}$ is made adjustable, so that $M_{1} a_{1}{ }^{2}$ can be adjusted to be as near $M a^{2}$ as may be desired.

## XLI. On Discordant Observations. By F. Y. Edgeworth, M.A., Lecturer at King's College, London*.

DISCORDANT observations may be defined as those which present the appearance of differing in respect of their law of frequency from other observations with which they are combined. In the treatment of such observations there is great diversity between authorities ; but this discordance of

[^0]methods may be reduced by the following reflection. Different methods are adapted to different hypotheses abont the cause of a discordant observation ; and different hypotheses are true, or appropriate, according as the subject-matter, or the degree of accuracy required, is different.

To fix the ideas, I shall specify three hypotheses : not pretending to be exhaustive, and leaving it to the practical reader to estimate the $\grave{a}$ priori probability of each hypothesis.
(a) According to the first hypothesis there are only two species of erroneous observations-errors of observation proper, and mistakes. The frequency of the former is approximately represented by the curve $y=\frac{h}{\sqrt{\pi}} e^{-h^{2} \mathbf{x}^{2}}$; where the constant $h$ is the same for all the observations. But the mathematical law* only holds for a certain range of error. Beyond certain limits we may be certain that an error of the first category does not occur. On the other hand, errors of the second category do not occur within those limits. The smallest mistake is greater than the largest error of observation proper. The following example is a type of this hypothesis. Suppose we have a group of numbers, formed each by the addition of ten digits taken at random from Mathematical Tables. And suppose that the only possible mistake is the addition or subtraction of 100 from any one of these sums. Here the errors proper approximately conform to a probability curve (whose $\dagger$ modulus is $\sqrt{\overline{165}}$ ), and the mistakes $\ddagger$ are quite distinct from the errors proper.

Here are seven such numbers: each of the first six was formed by the addition of ten random digits, and the seventh by prefixing a one to a number similarly formed-

$$
45, \quad 23, \quad 31, \quad 50, \quad 42, \quad 45, \quad 136 .
$$

[^1]The hypothesis entitles us to assert that 23 is an error-proper -an accidental deviation from 45 ; though the odds against such an event before its occurrence are considerable, about 100 to 1 . On the other hand, we may know for certain that 136 is a mistake.
$(\beta)$ According to the second hypothesis, the type of error is still the probability-curve with unvarying constant. But the range of its applicability is not so accurately known beforehand. We cannot at sight distinguish errors proper from mistakes. We only know that mistakes may be very large, and that the large mistakes are so infrequent as not to be likely to compensate each other in a not unusually numerous group of observations. This hypothesis may thus be exemplified :As before, we have a series of numbers, each purporting to be the sum of ten random digits. But occasionally, by mistake, the sum (or difference) of two such numbers is recorded. The mistake might be large, but it would not always exceed the limits of accidental deviation ( 100 and 0 ); which need not be supposed known beforehand. Here is a sequence of seven such numbers, which was actually obtained by me (in the course of 280 decades) -

$$
50, \quad 54, \quad 41, \quad 73, \quad 46, \quad 38, \quad 49 .
$$

The hypothesis leaves it doubtful whether 73 may not be a mistake ; the odds against it being an ordinary accidental deviation being, before the event, about 250 to 1 .
( $\gamma$ ) According to the third hypothesis all errors are of the type $y=\frac{h}{\sqrt{\pi}} e^{-h^{2} x^{2}}$. But the $h$ is not the same for different observations. Mistakes may be regarded as emanating from a source of error whose $h$ is very small. This hypothesis may be thus illustrated. Take at random any number $n$ between certain limits, say 1 and 100. Then take at random (from Mathematical Tables) $n$ digits, add them together and form their Mean (the sum $\div n$ ), and multiply this Mean by ten. The series of Means so formed may be regarded as measurements of varying precision ; the real value of the object measured being 45 . The weight, the $h^{2}$, being proportionate to $n$, one weight is $\grave{a}$ priori as likely as another. In order that the different degrees of precision, the equicrescent values of $h$, should be à priori equiprobable, it would be proper, having formed our $n$ as above, to take the mean of (and then multiply by 10 ), not $n$, but $n^{2}$ digits. Here is a series formed in this latter fashion :-

| $n \ldots \ldots \ldots \ldots \ldots \ldots$. | 5 | 7 | 6 | 1 | 10 | 8 | 1 |
| :---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| $n^{2} \ldots \ldots \ldots \ldots .$. | 25 | 49 | 36 | 1 | 100 | 64 | 1 |
| $\left.\begin{array}{c}10 \times \text { Mean of } n^{2} \\ \text { random digits }\end{array}\right\}$ | 31 | 45 | 43 | 100 | 43 | $47 \cdot 5$ | 100 |

In this table the firstrow is obtained by taking at random ten digits from a page of Statistics, 0 counting for ten. The second row consists of the squares of these numbers. The third row was thus formed from the second:-I took 25 random digits, and divided their sum by 25 ; then multiplied this mean by 10 . I similarly proceeded with 49 (fresh) digits, and so on. It will be noticed how the defective precision of the fourth and seventh observations makes itself felt. It was, however, a chance that they both erred as far as they conld, and in the same direction.

In the light of these distinctions I propose now to examine the different methods of treating discordant observations. For this purpose the methods may be arranged in the following groups :-
I. The first sort of method is based upon the principle that the calculus of probabilities supplies no criterion for the correction of discordance. All that we can do is to reject certain hage errors by common sense or simple induction as distinguished from the calculation of $a$ posteriori probability.
II. Or, secondly, we may reject observations upon the ground that they are proved by the Calculus of Probability to belong to a much worse category than the observations retained.
III. Or, thirdly, we may retain all the observations, affecting them respectively with weights which are determined by à posteriori probability.
IV. In a separate category may be placed a method which, as compared with* the simple Arithmetical Mean, reduces the effect (upon the Mean) of discordant observations-- the method which consists in taking the Median $\dagger$ or "Centralwerth" $\ddagger$ of the observations.

I propose now to test these methods by applying them in turn to all the hypotheses above specified.
I. (a) The first method-which is none other than Airy's, as I understand his contribution§ to this controversy-is adapted to the first hypothesis. Upon the second hypothesis $(\beta)$ the first method is liable to error, which, as will be shown under the next heading, is avoidable. ( $\gamma$ ) Upon the third hypothesis the first method is not theoretically the most precise; but it may be practically very good.
II. Under the second class I am acquainted with three

[^2]species : the criteria of Prof. Stone*, Prof. Chauvenet $\dagger$, and Prof. Peirce $\ddagger$.
II. (1) Prof. Stone's method is to reject an observation when it is more likely to have been a mistake than an error of observation of the same type as the others. In determining this probability he takes account of the $\grave{a}$ priori probability of a mistake. He puts for that probability $\frac{1}{n}$, admitting that $n$ cannot be determined precisely. The use of undetermined constants like this is, I think, quite legitimate§, and, indeed, indispensable in the calculation of probabilities. This being recognized, Prof. Stone's method may be justified upon almost any hypothesis. Hypothesis (a) presents two cases : where the discordant observation exceeds that limit of errors proper which is known beforehand, and where that limit is not exceeded. For example, in the instance $\|$ given above-where $\leq 5$ is the Mean, and the Modulus $\Phi$ is about 13the discordant observation might be either above 100 (e.g.110). or below io (e.g. 84). Now let us suppose that the a priori probability of a mistake is not infinitesimal, but say of the order 1000 . Since the deviation of 110 from the Mean is about five times the Modulus, the probability of this deviation occurring under the typical law of error is nearly a millionth. This observation is therefore rejected by Method II. (1), which so far agrees with Method I. Again, the probability of 84 being an accidental deviation is less than a forty-thousandth; 84-45 being about three times the Modulus. Therefore 84 also is rejected by the criterion. And we thus lose an observation which is by hypothesis (a) a good one. But this loss occurs very rarely. And the observation thrown away is, to say the least, not** a particularly good one, though doubtless it may happen that it is particularly wanted-as in the case of Gen. Colby, adduced $\dagger \dagger$ by Sir G. Airy.
II. (1) ( $\beta$ ) The second hypothesis is that to which Prof. Stone's criterion is specially adapted. Upon this hypothesis, 84 may be a mistake. In rejecting such discordant observations, we may indeed lose some good observations, especially if

[^3]we have exaggerated the $\grave{a}$ priori probability of a mistake. But it may be worth while paying this price for the sake of getting rid of serious mistakes. Especially is this position tenable according to the definition of the quesitum in the Theory of Errors*, which Laplace countenances. According to this view, the desideratum in a method of reduction is not so much that it should be most frequently right, as that it should be most advantageous; account being taken, not only of the frequency, but also of the seriousness, of the errors which it incurs. Prof. Stone's method might diminish our chance of being right (in the sense of being within a certain very small distance from the true mark $\dagger$ ); and yet it might be better than Method I., if it considerably reduced the frequency of large and detrimental mistakes.
II. (1) ( $\gamma$ ) Prof. Stone's method is less applicable to the third hypothesis. Though even in this case, if the smaller weights are à priori comparatively rare, it may be safe enough to regard ( $m-1$ ) of the $m$ observations as of one and the same type ; and to reject the $n$th if violently discordant with that supposed type.

The only misgiving which I should venture to express about this method relates, not to its essence and philosophy, but to a technical detail. Prof. Stone says:-" If we find that value which makes $\frac{2}{\sqrt{ }} \frac{\pi}{\pi} \int_{\frac{p}{a}}^{\infty} e^{-y^{2}} d y=\frac{1}{n}[$ where $p$ is the deviation of a discordant observation, and $a$ is the modulus of the probability-curve under which the other observations range, and $\frac{1}{n}$ is the $\grave{a}$ priori probability of a mistake], all larger values of $p$ are with greater probability to be attributed to mistakes." But ought we not rather to equate to $\frac{1}{n}$, not the left-hand member of the equation just written, which may be called $\theta\left(\frac{p}{a}\right)$, but $\theta^{m}\left(\frac{p}{a}\right)$, where $m$ is the number of observations. I am aware that the point is delicate, and that high authority could be cited on the other side. There is something paradoxical in Cournot's $\ddagger$ proposition that a certain

[^4]Phil. Mag. S. 5. Vol. 23. No. 143. April 1887. 2 C
deviation from the Mean in the case of Departmental returns of the proportion between male and female births is significant and indicative of a difference in kind, provided that we select at random a single French Department; bat that the same deviation may be accidental if it is the maximum of the respective returns for several Departments. There is something plausible in De Morgan's* implied assertion that the deficiency of seven in the first 608 digits of the constant $\pi$ is theoretically not accidental; because the deviation from the Mean 61 amounts to twice $\dagger$ the Modulus of that probabilitycurve which represents the frequency of deviation for any assigned digit. I submit, however, that Cournot is right, and that De Morgan, if he is serious in the passage referred to, has committed a slight inadvertence. When we select out of the ten digits the one whose deviation from the Mean is greatest, we ought to estimate the improbability of this deviation occurring by accident, not with De Morgan as $1-\theta(1 \cdot 63)$, corresponding to odds of about 45 to 1 against the observed event having occurred by accident ; but as $1-\theta^{10}(1 \cdot 63)$, corresponding to odds of about 5 to 1 against an accidental origination.
II. (2) Prof. Chauvenet's criterion differs from Prof. Stone's in that he makes the $\grave{a}$ priori probability of a mistake -instead of being small and undetermined-definite and considerable. In effect he assumes that a mistake is as likely as not to occur in the course of $m$ observations, where $m$ is the number of the set which is under treatment. It is not within the scope of this paper to consider whether this assumption is justified in the case of astronomical or of any other observations. It suffices here to remark that this assumption coupled with hypothesis ( $\alpha$ ) commits us to the supposition that huge mistakes occur on an average once in the course of $2 m$ observations. Upon this supposition no doubt Method II. (2), is a good one. Hypothesis $(\beta)$ expressly $\ddagger$ excludes this supposition ; the mistakes which, according to II. (2), are as likely as not, must, according to this second hypothesis, be of moderate extent. Thus, in the case above put of sums of ten digits, suppose that the number of such sums under observation is ten. According to Prof. Chauvenet's criterion we must reject any sum which lies outside $45 \pm k$, where

$$
\theta\left(\frac{k}{13}\right)=\frac{2 n-1}{2 n}=\frac{19}{20}=\cdot 95 .
$$

[^5]This gives for the required limit about 15. According, then, to II. (1) ( $\beta$ ), any observation greater than 60 , or less than 30 , is more likely than not to be a mistake in the sense of not belonging to the same law of frequency as the observations within those limits. But why on that ground should the discordant observation be rejected? Suppose there were not merely a bare preponderance of probability, but an actual certainty, that the suspected observation belonged to a different category in respect of precision from its neighbours, the best course certainly would be if possible (as Mr. Glaisher in his paper "On the Rejection of Discordant Observations" suggests) to retain the observation affected with an inferior weight. But if we have only the alternative of rejecting or retaining whole, it is a very delicate question whether retention or rejection would be in the long run better. There is not here the presumption against retention which arises when, as in II. (1), the discordant observation is large and rare ; so that, if it is a mistake, it is likely to be a serious and an uncompensated one. However, Prof. Chauvenet's method may quite possibly be better than the No-method of Sir G. Airy. Much would turu upon the purpose of the calculator-whether he aimed at being most frequently right* or least seriously wrong. The same may be said with reference to hypothesis $(\gamma)$.

There is a further difficulty attaching particularly to this species of Method II. In its precise determination of a limil, it takes for granted that the probability-curve to which we refer the discordant observation is accurately determined. But, when the number of observations is small, this is far from being the case. Neither of the parameters of the curve, neither the Mean, nor the Modulus, can be safely regarded as accurate. The "probable error" of the Mean is $\cdot 477 \frac{c}{\sqrt{ } n}$, where $c$ is the Modulus. The probable error of the Modulus is conjectured to be not inconsiderable from the fact that, if we took $m$ observations at random, squared each of them and formed the Mean-square-of-error, the " probable error" of that Mean-square-of-error would be • $477 \frac{c^{2}}{\sqrt{n}} \dagger$. This, however, is not the most accurate expression for the probable error of the Modulus-squared as inferred $\ddagger$ from any given $n$ observations.

[^6]To appreciate the order of error which may arise from these inaccuracies, we may proceed as in my paper of last October*. First, let us confine our attention to the Mean, supposing for a moment the Modulus accurate. Let $k$ have been determined according to Prof. Chauvenet's method, so that

$$
\theta\left(\frac{a}{c}\right)=\frac{2 m-1}{2 m} .
$$

To determine more accurately the probability of an observation not exceeding $a$ we must put for $a, a+z$, where $z$ is the error of the Mean subject to the law of frequency

$$
y=\frac{\sqrt{m}}{\sqrt{ } \pi c} e^{-\frac{m_{2} 2}{c^{2}}} .
$$

The proper course is therefore to evaluate the expression

$$
\int_{-\infty}^{\infty} \theta\left(\frac{a+z}{c}\right) \frac{\sqrt{m}}{\sqrt{\pi c}} e^{-\frac{m z^{2}}{c^{2}}} d z .
$$

Expanding $\theta$, and neglecting the higher powers of $z \dagger$, we have for the correction of $\theta\left(\frac{a}{c}\right)$ the subtrahend $\frac{2 \beta}{\sqrt{\prime}_{\pi n}^{\prime}} e^{-\beta^{2}}$, where $\beta$ is put for $\frac{a}{c}$. Call this modification of $\theta, \partial \theta$. To see how the prim $\hat{a}$ facie limit $\beta$ is affected by this modification, let us put
whence

$$
\begin{gathered}
{[\theta+\partial \theta](\beta+\Delta \beta)=\frac{2 n-1}{2 n} ;} \\
\partial \theta(\beta)+\Delta \beta \times \theta^{\prime}(\beta)=0 .
\end{gathered}
$$

Whence

$$
\Delta \beta=\frac{2 \beta}{\sqrt{\pi} n} e^{-\beta^{2}} \div \frac{2}{\sqrt{ } \pi} e^{-\beta^{2}}=\frac{\beta}{n} ;
$$

an extension of the limit which may be sensible when $n$ is small.

In the example given by Prof. Chauvenet the uncorrected limit as found by him is $1 \cdot 22$. This divided by the Modulns [which $=\sqrt{2} e=\cdot 8$ ] is $1 \cdot 5$. This result, our $\beta$, divided by 15 the number of observations, gives 1 as the correction of $\beta$; $\cdot 08$ as the correction of the limit $a$. The limit must be advanced to $1 \cdot 30$. This does not come up to the discordant observation $1 \cdot 40$. But we have still to take into account that we have been employing only the apparent Modulus (and Mean Error), not the real one. In virtue of this consideration I find-by an analysis analogous to that given in the paper

[^7]just referred to-that the limit must be pushed forward as much again ; so that the suspected observation falls within the corrected limit. I have similarly treated the example given by Prof. Merriman in his Textlook on The Method of Least Squares (131). The limit found by him is $4 \cdot 30$, and he therefore rejects the observation $4 \cdot 61$. But I find that this observation is well within the corrected limit *.
II. (3) Prof. Peirce's criterion is open to the same objections as that of Prof. Chauvenet. Indeed it presents additional difficulties. If by $y$ the author designates that quantity which Prof. Stone calls $\frac{1}{n}$, and which I have termed the "a priori" probability of a mistake, I am unable to follow the reasoning by which he obtains a definite value for this $y$. But I am aware how easy it is on such subjects to misunderstand an original writer.
III. We come now to the third class of method, of which I am acquainted with three species. (1) There is the procedure indicated by De Morgan and developed $\dagger$ ly Mr. Glaisher ; which consists in approximating to the weights waich are to be assigned to the observations respectively, after the analogy of the Reversion of Series and similar processes. (2) Another method, due to Prof. Stone $\ddagger$, is to put
$$
\mathrm{P}=h_{1} h_{2} \ldots e^{-h_{1}\left(x-x_{1}\right)^{2}-h_{2}^{2}\left(x-x_{2}\right)^{2}} \cdots \times d h_{1} d h_{2} \ldots
$$
as the $\grave{\alpha}$ posteriori probability of the given observations having resulted from a particular system of weights $h_{1}{ }^{2} h_{2}^{2} \& c$., and a particular Mean $x$; and to determine that system so that P should be a maximum. (3) Another variety is due to Prof. Newcomb §.
III. (1) \& (2) Neither of the first two Methods are well adapted to the first two hypotheses. Both indeed may successfully treat mistakes by weighting them so lightly as virtually to reject them. But both, I venture to think, are liable to err in underweighting observations, which, upon the first two hypotheses, have the same law of frequency as the others. Both, in fact, are avowedly adapted to the case where the observations

[^8]are not presumed beforehand to emanate from the same source of error. The particular supposition concerning the $\grave{a}$ priori distribution of sources which is contemplated by the De-Morgan-Glaisher Method, has not perhaps been stated by its distinguished advocates. The particular assumption made by the other Method is that one value of each $h$ is as likely as another over a certain range of values-not necessarily between infinite limits. I have elsewhere* discussed the validity of this assumption. I have also attempted to reduce the intolerable labour involved by this method. Forming the equation in $x$ of $(n-1)$ degrees,
$$
n x^{n-1}-(n-1) \mathrm{S} x_{1} x^{n-2}+(n-2) \mathrm{S} x_{1} x_{2} x^{n-3}-\& \mathrm{c} .=0,
$$

I assume that the penultimate (or antepenultimate) limiting function or derived equation will give a better value than the last-derived equation $n x-\mid n-1 \mathrm{~S} x_{1}$, which gives the simple Arithmetic Mean. Take the observations above instanced under hypothesis ( $\gamma$ ),

$$
31,45,43,100,43,47 \cdot 5,100 .
$$

For convenience take as origin the Arithmetical Mean of these observations 58.5 , say 58 . Then we have the new series

$$
-27,-13,-15,+42,-15,-11,+42
$$

Here $\mathbb{S} x_{1} x_{2}=-2494$. And the penultimate limiting equation is

$$
7 \times 6 \times 5 \times 4 \times 3 x^{2}+5 \times 4 \times 3 \times 2 \times 1 \times-2494=0 .
$$

Whence $x^{2}=119$. And $x= \pm 11$ nearly. To determine which of these corrections we ought to adopt, the rule is to take the one which makes P greatest; which is $\dagger$ the one which makes $\left(x-x_{1}\right)\left(x-x_{2}\right)\left(x-x_{3}\right) \ldots\left(x-x_{7}\right)$ smallest ; each of the differences being taken positively.

The positive value, +11 , gives the differances

$$
38,24,26,21,36,22,21 .
$$

For the negative value, -11 , the differences are

$$
16,2,4,53,4,0,53
$$

(where 0 of course stands for a fraction). The continued product of the second series is the smaller. Hence -11 is

[^9]the correction to be adopted. Deducting it from 58, or rather $58 \cdot 5$, we have $47 \cdot 5$, which is a very respectable approximation to the real value, as it may be called, viz. 45.
III. (3) Prof. Newcomb* soars high above the others, in that he alone ascends to the philosophical, the utilitarian, principles on which depends the whole art of reducing observations. Here are whole pages devoted to estimating and minimizing the Evil incident to malobservation. With Gauss, Prof. Newcomb assumes $\dagger$ "that the evil of an error is proportional to the square of its magnitude." He would doubtless admit, with Gauss, that there is something arbitrary in this assumption. Another somewhat hypothetical datum is what he $\ddagger$ describes as the "distribution of precisions." In view of this looseness in the data, it becomes a nice question whether it is worth while expending much labour upon the calculation. The answer to this question depends upon an estimate of probability and utility, concerning which no one is competent to express an opinion who has not, on the one hand, a philosophical conception of the Theory of Errors, and, on the other hand, a practical acquaintance with the art of Astronomy. The double qualification is probably possessed by none in a higher degree than by the distinguished astronomer to whom we owe this method.
IV. It remains to consider the fourth Method. But the length and importance of this discussion will require another paper.

## XLII. On the Action of Heat on Potassic Chlorate and Perchlorate. By Еdmund J. Mills, D.Se., F.R.S.§

TT has been pointed out by Teed $\|$, and subsequently by P. Frankland and Dingwallा, that potassic chlorate and perchlorate may be decomposed by heat in such a manner as to lead in each case to various relations among the products of decomposition.

It has occurred to me that both of these chemical changes are instances of Cumulative Resolution**, from which point of view they admit of very simple, and at the same time perfectly adequate, representation.

* American Journal of Mathematics, vol. viii. No. 4.
$\dagger$ § 3, p. 348.
$\ddagger$ § 9, p. 359.
§ Communicated by the Author.
II Proc. Chem. Soc. xii. p. 105 ; xvi. p. 141 ; xxxiii. pp. 24 \& 25.
बI Ibid. xvi. p. 141; xxxii. p. 14; and Trans. Chem. Soc. 1887, p. 274.
** Phil. Mag. [5] iii. p. 492 (1877).


[^0]:    * Communicated by the Author.

[^1]:    * This follows from the supposition that an error of observation is the joint result of a considerable, but finite, number of small sources of error. The law of facility is in such a case what Mr. Galton calls a Binomial, or rather a Multinomial. (See his paper in Phil. Mag. Jan. 1875, and the remarks of the present writer in Camb. Phil. Trans. 1886, p. 145, and Phil. Mag. April 1886.)
    $\dagger$ I may remind the reader that I follow Laplace in taking as the constant or parameter of probability-curves the reciprocal of the coefficient of $x$ : that is $\frac{1}{\vec{h}}$, according to the notation used above. It is $\sqrt{2}$ times the "Mean Error" in the sense in which that term is used by the Germans, beginning with Gauss, and many recent English writers (e. g. Chauvenet) ; and it is $\sqrt{\pi}$ times the Mean Error in the (surely more natural) sense in which Airy, after Laplace, employs the term Mean Error (Chauvenet's Mean of the Errors).
    $\ddagger$ In physical observations the limit of errors proper must, I suppose, be more empirical than in this artificial example.

[^2]:    * This is pointed out by Mr. Wilson in the Monthly Notices of the Astronomical Society, vol, xxxviii., and by Mr. Galton, Fechner, and others.
    $\dagger$ Cournot, Galton, \&c.
    $\ddagger$ Fechner, in Abhandl. Sax. Ges. vol. [xvi.].
    § Gould's Astronomical Journal, vol. iv. pp. 145-147.

[^3]:    * Month. Not. Astronom. Soc. Lond. vol. xxviii, pp. 165-168.
    $\dagger$ 'Astronomy,' Appendix, Art. 60.
    $\ddagger$ Ibid. Art. 57.
    § See my paper on à priori Probabilities, in Phil. Mag. Sept. 1884 ; also "Philosophy of Chance," Mind, 1884, and Camb. Phil. Trans. 1885, pp. 148 et seq.
    || Page 365.
    II $\sqrt{\overline{1} 6 \overline{5}}$, exactly. As determined empirically by me from the mean-square-of-error of 280 observations (i.e. sums of 10 digits), the Modulus was $\sqrt{160}$.
    ** See the remark made under II. (2) ( $\beta$ ).
    $\dagger \dagger$ Gould's Astronom. Journ. vol. iv. p. 188.

[^4]:    * See my paper on the " Method of Least Squares," Phil. Mag. 1883, vol. xvi. p. 363 ; also that on "Observations and Statistics," Camb. Phil. Tr. 1885; and a little work called 'Metretike' (London : Temple Co., 1887).
    $\dagger$ The sense defined by Mr. Glaisher, 'Memoirs of the Astronomical Society,' vol. xl. p. 101.
    $\ddagger$ Erposition de la theorie des Chances, Arts. 102, 114. "Nous ne nous dissimulons pas ce qu'il y a de delicat dans toute cette discussion," I may say with Cournot.

[^5]:    * 'Budget of Paradoxes,' p. 291.
    $\dagger$ If we take many batches of random digits, each batch numbering 608, the number of sevens per batch ought to oscillate about the Mean 61, according to a probability-curve whose Modulus is $\sqrt{\frac{18}{100} 608}=10 \cdot 4$.
    $\quad \ddagger$ Above, p. 366 .

[^6]:    * See the remarks above, p. 369.
    + Todhunter, art. 1003 (where there is no necessity to take the origin at one of the extremities of the curve).
    $\ddagger$ I allude here to delicate distinctions between genuine Inverse Probability and other processes, which I have elsewhere endeavoured to draw, Camb. Phil. Trans. 1885.

[^7]:    * Phil. Mag. 1886, vol. xxii. p. $371 . \quad \dagger$ See the paper referred to.

[^8]:    * These corrections may be compensated by another correction to which the method is open. In determining whether the suspected observation belongs to the same type as the others, would it not be more correct to deduce the characters of that type from those others, exclusive of the suspected observation? The effect both on the Mean and the Modulus would be such as to contract the limit.
    $\dagger$ Memoirs of the Astronomical Society.
    $\ddagger$ Monthly Notices of the Astronomical Society, 1874. This Method was proposed by the present writer in this Journal, 1883 (vol. xvi. p. 360), in ignorance of Prof. Stone's pricrity.
    § American Journal of Mathematics, vol. viii. No. 4.

[^9]:    * Camb. Phil. Trans. 1885, p. 151.
    $\dagger$ See Phil. Mag. 1883, vol, xvi. p. 371.

