ACCURACY OF MUTATION RATES

By W. L. STEVENS

Statistical Department, Rothamsted Experimental Station, Harpenden, Herts

THE device of attaching a standard error to an estimate, in order to indicate the limits within which the true value may, at any chosen level of significance, be considered to lie, is now so well known that many writers are using it in cases where, owing to non-normality of the distribution, it is very misleading and even quite meaningless.

Estimates of mutation rates are commonly based on only a few observed mutations n in a large number N of cultures. The estimated rate is then $p=n/N$, and a standard error has commonly been calculated from the formula

$$\sqrt{(p/N)} \quad \text{or more exactly} \quad \sqrt{\{p(1-p)/N\}}.$$

Amongst innumerable examples, we may take the first line of data in the second table of a paper by Timoféeff-Ressovsky *et al.* (1935):

Cultures	Mutations	Percentage mutation rate
3708	7	$0{\cdot}19 \pm 0{\cdot}07$ (more exactly $0{\cdot}1888 \pm 0{\cdot}0713$)

These figures purport to show that, taking say a level of significance of 1% corresponding to 2·576 standard deviations, then the result contradicts mutation rates lower than

$$0{\cdot}1888 - 2{\cdot}576 \times 0{\cdot}0713 = 0{\cdot}0051\,\%$$

or higher than 0·3725%, or, more precisely, that if the mutation rate were as low as 0·0051% there would be only one chance in two hundred of getting seven or more mutations, and if as high as 0·3725% only one chance in two hundred of getting seven or fewer mutations.

An exact computation shows, however, how false these statements are. If the true mutation rate is p, the probabilities of finding 0, 1, 2, ..., n, ... mutations in N cultures, are given by the terms of the binomial series

$$(1-p)^N, \quad Np(1-p)^{N-1}, \quad \frac{N(N-1)}{2!}p^2(1-p)^{N-2}, \ldots.$$

Taking $N=3708$, $p=0{\cdot}000051$ we find the probability of seven or more mutations to be only one in 700 million instead of one in two hundred.

Taking $p = 0{\cdot}003725$ we find the probability of seven or fewer mutations to be one in twenty-nine instead of one in two hundred.

In the next line of data in the same paper, the use of the standard error is more obviously absurd.

Cultures	Mutations	Percentage mutation rate
1431	3	$0{\cdot}21 \pm 0{\cdot}12$

Here the estimated mutation rate is only one and three-quarter times its own standard error, so if we are to use the standard error we should conclude that the data do not very strongly contradict (4% significance) even a negative mutation rate!

Lest it should be thought that these are only theoretical criticisms, let us turn to an example where the application of the standard error has led to a false conclusion.

In a paper on the process of structural change in chromosomes of *Drosophila*, Muller (1940) investigates the relation between X-ray dosage and number of translocations. The second line of data on p. 20 reads:

Dose	Cultures	Number of translocations	Percentage of translocations
400	1803	2	$0{\cdot}11 \pm 0{\cdot}08$

Muller is discussing whether the percentage translocations is proportional to the 3/2 power of the dosage, in which case the expectation would be 0·31%, and concludes that 'the value...does disagree significantly from the expectation based on the 3/2 power rule'.

One assumes that Muller finds that the difference 0·20% between the observed 0·11% and the theoretical 0·31% is two and a half times the standard error and that he deduces that the probability of so large a discrepancy arising by chance is only about one in eighty. If this were true then indeed Muller would be right in saying that the observed value differs significantly from that predicted by the 3/2 power rule.

But, in fact, the chance of getting 0, 1, or 2 translocations, if the true probability were 0·31%, is about one in twelve. Since we would also be prepared to reject the rule if the result were equally unlikely because there were too many translocations, the significance to be attached to the result is about one in six. This is a very different matter from one in eighty, and is not high enough to justify Muller's conclusion.

A NEW TABLE FOR TESTING SIGNIFICANCE

It is possible that many who are aware of the extreme inaccuracy of the standard error test have been deterred from making the exact test by the labour of computation. We have accordingly prepared a table which enables limits to be determined or tests of significance made,

Binomial and Poisson distributions: fiducial limits of the expectation

		Probability of n or more					Probability of n or fewer				
n	$p=n/N$	0·005	0·01	0·025	0·05	0·1	0·1	0·05	0·025	0·01	0·005
0	10 %	$=1/N$	Lower limits are all zero				2·06	2·59	3·09	3·69	4·11
	5 %						2·17	2·78	3·37	4·11	4·65
	0						2·30	3·00	3·69	4·61	5·30
1	10 %	0·0050	0·0100	0·0253	0·512	0·105	3·37	3·94	4·45	5·04	5·44
	5 %	0·0050	0·0100	0·0253	0·512	0·105	3·62	4·32	4·97	5·78	6·34
	0	0·0050	0·0101	0·0253	0·513	0·105	3·89	4·74	5·57	6·64	7·43
2	10 %	0·106	0·152	0·247	0·361	0·538	4·90	5·65	6·34	7·17	7·74
	5 %	0·105	0·150	0·245	0·358	0·535	5·11	5·96	6·77	7·73	8·44
	0	0·103	0·149	0·242	0·355	0·532	5·32	6·30	7·22	8·41	9·27
3	10 %	0·348	0·448	0·634	0·834	1·12	6·28	7·16	7·96	8·93	9·61
	0	0·338	0·436	0·619	0·818	1·10	6·68	7·75	8·77	10·04	10·98
4	10 %	0·694	0·847	1·12	1·40	1·77	7·60	8·58	9·47	10·54	11·31
	0	0·672	0·823	1·09	1·37	1·74	7·99	9·15	10·24	11·60	12·59
5	10 %	1·11	1·32	1·66	2·01	2·47	8·88	9·94	10·91	12·08	12·90
	0	1·08	1·28	1·62	1·97	2·43	9·27	10·51	11·67	13·11	14·15
6	10 %	1·58	1·84	2·26	2·67	3·20	10·14	11·27	12·30	13·56	14·44
	0	1·54	1·79	2·20	2·61	3·15	10·53	11·84	13·06	14·57	15·66
7	10 %	2·10	2·39	2·88	3·35	3·96	11·37	12·58	13·67	14·99	15·92
	0	2·04	2·33	2·81	3·29	3·89	11·77	13·15	14·42	16·00	17·13
8	10 %	2·65	2·98	3·53	4·06	4·73	12·59	13·86	15·01	16·40	17·37
	0	2·57	2·91	3·45	3·98	4·66	12·99	14·43	15·76	17·40	18·58
9	10 %	3·22	3·60	4·21	4·78	5·51	13·79	15·12	16·32	17·78	18·80
	0	3·13	3·51	4·12	4·70	5·43	14·21	15·70	17·08	18·78	20·00
10	10 %	3·82	4·23	4·90	5·53	6·31	14·99	16·37	17·62	19·13	20·20
	0	3·72	4·13	4·80	5·43	6·22	15·41	16·96	18·39	20·14	21·40
11	10 %	4·44	4·89	5·61	6·28	7·12	16·17	17·61	18·91	20·47	21·57
	0	4·32	4·77	5·49	6·17	7·02	16·60	18·21	19·68	21·49	22·78
12	10 %	5·07	5·56	6·33	7·05	7·94	17·35	18·83	20·18	21·80	22·93
	0	4·94	5·43	6·20	6·92	7·83	17·78	19·44	20·96	22·82	24·14
13	10 %	5·72	6·24	7·06	7·82	8·76	18·52	20·05	21·44	23·11	24·28
	0	5·58	6·10	6·92	7·69	8·65	18·96	20·67	22·23	24·14	25·50
14	10 %	6·39	6·94	7·81	8·61	9·59	19·68	21·25	22·69	24·41	25·61
	0	6·23	6·78	7·65	8·46	9·47	20·13	21·89	23·49	25·45	26·84
		Correction to be added to n									
15 and over	10 %	1·53	1·20	0·80	0·45	0·14	1·13	1·41	1·73	2·15	2·47
	0	1·87	1·49	0·99	0·62	0·26	1·33	1·72	2·15	2·73	3·19

very easily and rapidly.[1] The method of use of the table will be demonstrated by examples.

Example 1. *Mutation rate.* Taking the data first quoted we have

$$n=7, \quad N=3708, \quad p=n/N=0{\cdot}189\,\%.$$

Wishing to find the limits corresponding to significance of one in a hundred, we enter the table for $n=7$, under the first and last columns (notice that these are headed 0·005 because following the usual convention, the *sum* of the probabilities of rejection at either limit is to be 0·01). Since 0·189 is very near zero, we read off the entries against $p=0$ (which is exactly equivalent to regarding the distribution as of the Poisson type)

$$2{\cdot}04; \quad 17{\cdot}13.$$

Dividing each of these by 3708 we find limits to the mutation rate of

$$0{\cdot}055\,\% \quad \text{and} \quad 0{\cdot}462\,\%.$$

Hence the result decisively contradicts mutation rates outside these limits. They should be compared with the limits implied by the standard error, as found above.

In practice, when we wish to give plausible limits to our estimate, one would prefer the closer limits given by the centre two columns of the table.

Example 2. Let us now turn to Muller's data. We wish to test whether two translocations in 1803 cultures disprove a rate of 0·31 %. Multiplying N, the number of cultures, by the theoretical rate gives an expectation of 5·6 translocations

$$1803\times 0{\cdot}0031=5{\cdot}6.$$

Reference to the table against $n=2$, $p=0$ shows that 5·6 lies between the entries under 0·1 and 0·05

0·1	0·05	0·025
5·32	6·30	7·22

We therefore conclude that the data do not show a very strong disagreement with the hypothesis. A rough interpolation gives 0·08 for the significance, or one in twelve, as we saw before. Again let it be emphasized that if we are using the conventional 0·05 level of significance (equivalent to two standard deviations) and are testing for deviations in *either* direction, then the predicted expectation would have to lie

[1] An extended form of this table, covering all possible samples from binomial distributions (i.e. $p=0$–$50\,\%$), is shortly to be published (Fisher & Yates, 1941).

outside the range 0·242–7·22, tabulated under 0·025 ($=\frac{1}{2}\times 0{\cdot}05$) to indicate significance.

Example 3. *Testing a Mendelian segregation.* Data expected to show a 15:1 segregation actually yielded only 12 recessives in 355 offspring. Is the observed proportion too low?

The expected number of recessives is

$$355 \div 16 = 22{\cdot}19.$$

Reference to the table shows, for $n=12$, under the upper 0·01 column

10 %	21·80
0	22·82

The expected number lies between these limits, and so the chance of getting as few as 12 or fewer is about one in a hundred. However, if one is striving after accuracy, one may interpolate. The observed proportion of recessives is

$$p = n/N = 12/355 = 3{\cdot}4\,\%.$$

The difference between the tabular entries is 1·02. Multiplying by 0·34 we get 0·35 which subtracted from 22·82 gives 22·47. The theoretical expectation is below this limit, and so the result just fails to reach the one in a hundred level of significance in the direction of being too small. (Notice here that since the direction is specified by the question, and by the answer, we have not doubled the significance level read from the table.)

Example 4. *Recombination fraction.* In eighty offspring from a selfed heterozygote in coupling, the two recombination classes were unrepresented. Find an upper limit to the recombination fraction.

Since when $n=0$ the observed proportion is necessarily zero, expectations have been tabulated instead for three values of $1/N$. To find the one in twenty limit for the expected number we therefore take

$$1/80 = 1{\cdot}25\,\%,$$

and interpolate between the entries for $n=0$, $1/N = 5\,\%$ and 0. Thus for the one in twenty limit we read

5 %	2·78	
0	3·00	difference = 0·22.

Since 1·25 % is a quarter of 5 % we find the upper limit to be

$$3{\cdot}00 - \tfrac{1}{4}\,(0{\cdot}22) = 2{\cdot}95.$$

But the combined probability of the two recombination classes is $\frac{1}{2}\rho\,(2-\rho)$, where ρ is the recombination fraction. Equating the expected number to the upper limit just determined, we have

$$80\times\tfrac{1}{2}\rho\,(2-\rho)=2{\cdot}95,$$
$$\rho^2-2\rho+0{\cdot}0738=0,$$
$$\rho=0{\cdot}0376.$$

Hence the recombination fraction is not greater than 0·0376 unless the result is less likely than one in twenty.

Example 5. In a backcross experiment, 20 cross-overs were found in a total of 512 offspring. Find the limits of the recombination fraction, corresponding to a level of significance of one in twenty.

When n is greater than fourteen and therefore falls outside the range of the main table, the bottom two lines provide corrections to the limits found by the usual standard error method. It will be noticed that one can still not afford to neglect these corrections, for they can amount to as much as a difference between one in twenty and one in a hundred significance.

The estimated recombination fraction is $20/512=0{\cdot}0391$. The standard error of the expectation is

$$\sqrt{\{Np\,(1-p)\}}=\sqrt{\{n\,(1-p)\}}=\sqrt{(20\times0{\cdot}961)}=\sqrt{19{\cdot}62}=4{\cdot}43.$$

In the usual way, we add and subtract 1·96 times the standard error from 20, to give the limits

11·32 28·68.

The corrections given by the table are

10 %	0·80	1·73
0	0·99	2·15.

Interpolating for $p=3{\cdot}91$ % we find

0·91 1·98.

Adding these corrections to the above limits gives the true limits

12·23 30·66.

Dividing by 512 gives the limits of the recombination fraction of 0·0239 to 0·0599.

Summary

The use of the standard error of estimates of mutation rates, and other small proportions, is shown to be highly erroneous. A table is provided for rapidly making the exact tests of significance.

REFERENCES

FISHER, R. A. & YATES, F. (1941). *Statistical Tables for Biological, Agricultural and Medical Research*, 2nd ed. (in the Press). Edinburgh: Oliver and Boyd.

MULLER, H. J. (1940). An analysis of the process of structural change in chromosomes in *Drosophila*. *J. Genet.* **40**, 1–66.

TIMOFÉEFF-RESSOVSKY, N. W., ZIMMER, K. G. & DELBRUCK, M. (1935). Über die Natur der Genmutation und der Genstruktur. *Nach. Ges. Wiss. Göttingen* F.G. **6**.