the red as water-vapour bands. While this identification is not certain, it at least suggests that the hydrogen in the upper atmosphere is present as part of the water molecule. While this may be reasonable for the light of the night-sky, one would expect water-vapour to be dissociated during auroral displays. Furthermore, no water-vapour bands have been observed in auroral spectra.

Although the present evidence is not conclusive, it was thought worth while to direct attention to this remarkable effect of hydrogen on the nitrogen afterglow and to its possible relationship to the question which was put at the beginning of this note.

JOSEPH KAPLAN.

University of California at Los Angeles.
Aug. 30.

Statistical Tests

PROF. FISHER[1] is an apt controversialist, but he knows as well as I do that what I understand by *graduation* is not confined to *curves*; that I should term graduation the fitting of a binomial to a series of observations, or the determining whether a system of correlation coefficients could be reasonably supposed to have arisen from samples of material drawn from a population with a given correlation coefficient. The difference between Prof. Fisher and myself lies in the use (and abuse) of the acceptance and rejection of 'hypotheses'. There is only one case in which an hypothesis can be definitely rejected, namely, when its probability is zero. He cites a case which I criticised in the paper he refers to, in which two recessives (say) had produced a dominant, and theory was absolutely contradicted. It did not require an *application* of the (P, χ^2) test to assert that either theory or observations must be rejected! I merely showed that the (P, χ^2) test did not fail in this case. But let us look into what actually happens, and I cannot do better than illustrate it on some statistics provided by Sir James Jeans in NATURE of September 14, 1935 (p. 432). He is comparing the eccentricities of visual binaries, 116 in number, against a theory of equipartition (not a *curve*, but frequencies are considered). His data expressed by a frequency series run as follows :—

Eccentricity.	Observed.	Theory for 116 stars, $e<1$	Theory for 83 stars, $e \leqslant 0{\cdot}06$.
0·00–0·01	0	4·5	9
0·01–0·02	11		
0·02–0·03	9	6	12
0·03–0·04	14	8	16
0·04–0·05	24	10·5	21
0·05–0·06	25	13	25
0·06–0·07	6	15	—
0·07–0·08	13	17	—
0·08–0·09	7	20	—
0·09–1·00	7	22	—

If the P, χ^2 test be applied to the total 116 binaries, we have $P < 0{\cdot}000{,}0005$. On the other hand, if it be applied to the 83 stars of lowest eccentricity, $P = 0{\cdot}79$. In neither case can you say the hypothesis is true or false. You reject it in the former case because it is a poor graduation, you say in the latter case that it is a reasonably good graduation because 79 per cent of random samples would, were the 'hypothesis' true, give a worse result than the observations do. But in accepting it as a working graduation, you do not assert its truth any more than you assert the falsity of the hypothesis applied to the whole 116 stars; you merely say the latter is a bad graduation, and try for a better. Had Sir James Jeans taken all stars with eccentricity $\leqslant 0{\cdot}07$, instead of $\leqslant 0{\cdot}06$, he would have found $P = 0{\cdot}105$, and if he had proceeded to $e \leqslant 0{\cdot}08$, the result would have been $P = 0{\cdot}00001$, that is, he might have got a worse sample in 100,000 trials. Actually he gives his reasons for cutting off the higher eccentricities. With them I am not concerned, although the exact cutting off at $e = 0{\cdot}06$ is not discussed; the difficulty of detecting high eccentricity binaries and of then determining their orbits may account for the irregularity of the last four frequency entries, as he holds, or there may be other reasons why the falling-off occurs at $e = 0{\cdot}06$. *Hypotheses non fingo!*

Now Prof. Fisher refers to rejecting hypotheses as a function of the P, χ^2 method, and of accepting them as a logical fallacy. I have in my letter of August 24 stated that the tests are there to ascertain whether a reasonable *graduation* has been reached; not to assert whether one or other hypothesis is true or false. We should accept Sir James Jeans's equipartition as a reasonable graduation for the observed binaries $e \leqslant 0{\cdot}06$ ($P = 0{\cdot}79$) and reject it as a graduation for the observed binaries $e \leqslant 0{\cdot}08$ ($P = 0{\cdot}000{,}01$). It is not for statisticians to say whether an hypothesis is false except when $P = 0$. All that they can legitimately say is that it gives a poor *graduation*. In particular, it is very unwise in *my* opinion to form tables which provide only the values of $P = 0{\cdot}01$ and $P = 0{\cdot}05$, and consider 'hypotheses' which give a value of $P < 0{\cdot}01$ as 'false', and those with a value between 0·01 and 0·05 as 'doubtful', and for the rest of the scale of P have no descriptive category, for you must not say that such values prove hypotheses to be true. Hence I repeat my assertion, in the face of all the authority of Prof. Fisher and his followers, that all the P, χ^2 test ascertains is goodness of graduation, and I hold that 'goodness' of graduation is relative to the nature of the material investigated, our experience of similar material and the purpose to which we intend to put our graduation. The value of P at which we consider goodness or badness of graduation starts cannot be fixed without regard to the special problem under consideration.

There seems somewhere a logical fallacy in the position of both Prof. Fisher and Mr. Buchanan Wollaston. They both apparently assert that the P, χ^2 test enables one to say an hypothesis is false, yet never to say that an hypothesis is true, but if an hypothesis be *false*, its reverse must be true. If you assert that the hypothesis that a sample is drawn from a normal curve is false, the reverse hypothesis that it is *not* drawn from a normal curve must be true. As a matter of fact, the P, χ^2 test has only measured its 'goodness of fit' by a probability coefficient, and it is as idle to say as a result of it, that the hypothesis is 'false', as that the reverse of it is 'true'. The only exception to this rule is when the observations show the existence of individuals in a frequency class which the hypothesis asserts cannot exist.

The 'laws of Nature' are only constructs of our minds; none of them can be asserted to be true or to be false, they are good in so far as they give good fits to our observations of Nature, and are liable at any time to be replaced by a better 'fit', that is, by a construct giving a better graduation.

KARL PEARSON.

Biometrika Office,
University College,
London, W.C.1.

[1] NATURE, **136**, 474, Sept. 21, 1935.