February 12, 1940

Dear Professor Fréchet,

In my paper of 1922, Phil Trans., which I think you have available, I comment on the use made by Pearson and Filon in 1898 of the formula giving the minimal variance or standard deviation of an estimated parameter.

If I remember right, they arrive at the formula by a very obscure argument involving inverse probability; but probably what they have done is equivalent to proving that $\sqrt{n}$ times the error of an estimate, taken, e.g. approximately at the mode of an inverse probability distribution, will in the limit when $\underline{\underline{n}}$ is large, tend to be normally distributed with variance $\frac{1}{I}$. One of the features which makes their treatment so obscure and indefinite is that they do not notice that different methods of estimation have, in the limit for large samples, different precisions. Consequently they apply the formula, without hesitation, to estimates found by fitting Pearsonian curves by moments, and obtained a number of erroneous formulae, which were, I believe, not corrected till early in the present century when Sheppard showedhow the standard errors of moments could be calculated directly. The erroneous formulae seem thereupon to have been dropped by the Pearsonian school, but there seems to be no hint in Biometrika which would inform the reader that previous misleading formulae were being corrected.

Edgeworth wrote a series of papers about 1908 in the Statistical Society's journal. His attitude seems to be rather over-cautious than over-confident. He refers his readers to the paper by Pearson and Filon in terms which leave little doubt that he regarded it as correct, although in numerous other passages in this series of papers I should have thought he must be taken as recognising that different methods of estimation possessed different precisions, that these precisions have an upper bound, and that an estimate having in the limit the highest possible precision can be found by the procedure which Edgeworth regarded as inverse probability, ~~by~~ but which, as I was concerned to emphasise in the 1930 paper I sent you, can and indeed must be completely dissociated from inverse probability, and which in 1922 I called the method of maximum likelihood.

The confusion of associating this method with Bayes' theorem seems to have been due originally to Gauss, who certainly recognized its merits as a method of estimation, though I do not know whether he proved anything definite about it.

I do not know of any explicit statements of the properties, consistency, efficiency and sufficiency which may characterise estimates prior to my 1922 paper. I had noted the functional peculiarity of sufficiency in an early paper, 1919, in the Monthly Notices of the Astronomical Society.

Yes, I think your definition of probability is consistent with mine, though I should emphasise that the physical magnitude considered is a character of the univers~~al~~e (population sampled)

Yours sincerely,