29 March 1934.

Dr. H.Jeffreys,
St. John's College,
Cambridge.

Dear Jeffreys,

Thanks for your letter. We seem to have a deal of ground to cover in our ten pages apiece. I think your draft shows very well what your line of approach is; at least, I find it much more reasonable in this form than I did in the paper in the Proceedings. I want rather to modify the tone of my note, without much altering its substance, e.g. instead of saying that in your definition notions analogous to probability and those analogous to amount of information are "confused", I have altered the word to "admissable".

From the point of view of interesting the general scientific public, which really ought to be much more interested than it is in the problem of inductive inference, probably the most useful thing we could do would be to take one or more specific puzzles and show what our respective methods made of them. This might mean more work than you want to undertake and will certainly require that we should agree quite exactly as to what the data were and what questions we want to elucidate from them. I had this in mind in mentioning the plant problem, as it was the first case I had come across in which the notions of probability and likelihood appeared to be harnessed tandum, instead of only one or the other of them.

In case you feel it is a point worth discussing I will state it again in a form which, apart from complications, which I think are logically unessential, is that in which it might arise in practice.

A man collects a number of plants of the same species from a wild population, grows them in experimental cultures, and after some years of much labour, has succeeded in classifying a small number. Three of those tested he finds to be identical in respect to his method of classification and these he calls Type A. Two more differ from A, but agree with each other. He calls these B. Two more differ from A and B, but being alike, constitute a new type C. The remaining three which he has tested differ from A,B and C, and from each other, and so constitute solitary representatives of Types D,E and F. He has then tested ten plants and found them to belong to six different types with frequencies of occurrence represented by the partition $(3 2^2 1^3)$ of the number 10.

He knows, or thinks he knows, that the frequencies with which the different sorts of plants distinguishable by his tests occur in the wild population constitute a geometric progression like the frequencies of Planck's oscillators containing 0,1,2,3 quantum of energy. He is not concerned to test this hypothesis, but using it as part of his data to estimate the probable outcome in respect of the discovery of nww types of continuing his labours by testing more specimens from the wild population, the kind of question he might properly ask is, "What is the probability that the eleventh plant testede will turn out to belong to one of the six types already found, or alternatively to add a new type to the list?"

If he knew the value of r, the constant ratio of his geometric progression, this probability would, I think, be determinate and for different values of r from 0 to 1, the probability of getting a new type next time must also, I suppose, run from Oto 1. He does not,in fact, know r, but his observed partition gives him some information about r, e.g. if the observed partition has been $(1)^{10}$ he wouldthink r was higher than he would if the observed partition had been(10).   He can in fact, asign to each value of r a certain likelihood, which will be higher if his partition is made up of a lot of little parts and lower if it has a few big parts, and for any observed partition he can specify the particular value of r which is most likely and the likelihood relative to it of any other value of r.

But this takes me no further a nd to tell the botanist that I know which of thepossible values of the probability he is seeking has the highest likelihood, and I know how much proportionately the likelihood is lower for any other value of the probability. The real question is whether he has any right to expect from me more definite information than this and this may be considered under the heads a) are any further deductions to be drawn vigorously from the bad data as stated b) can we <u>properly</u> suppliment these data by axiomatic truths revalent to our problem, which will enable further conclusions to be drawn.

So much for the case that puzzles me. Without suggesting that you should modify your draft, it may be worthwhile for me to make a few comments. You say I object to the introduction of an *a priori* element, as I should like to get this notion a little more precise. What I object to is the assertion that we must, in considering the possible values of an unknown perameter, used to specify the population sampled, introduce a frequency distribution or a probability distribution of this perameter, supposingly known *a priori*. If we really have knowledge of this kind, as in some problems, which can be reconstructed with dice or urns, I do not deny that it should be introduced, but, I say that, we often do not possess this knowledge, i.e. either that we are in absolute ignorance or that our knowledge is so vague or unsatisfactory that we may properly prefer not to introduce it into the basis of a mathematical discussion, but rather prefer to keep it in reserve and see ~~if~~ what the observations can prove without it and whether it, for what it is worth is confirmed or contradicted by the observations. I claim, in fact, that it is at least a legitimate question to ask: "What can the observations tell us when we know the form of our population, but know nothing of the *a priori* probability of the differentvalues that its perameters may have. This is the situation which I treat as the typical one in the Theory of Estimation, but it would be quite legitimate to say that our assumed knowledge of the form of the distribution, which is needed before there can be anything to estimate , is in the nature of *a priori* knowledge. Infact, such knowledge seems to me essential in what I call problems of specification, but out of place in the next stage, when problems of estimation arise.

I think this may explain your difficulty about fitting a smooth curve to a number of discrepant observations. If the form of the curve is given, e.g. a polynomial ofthe fourth degree, thenyou will probably agree that the best curve of this form will be one specified by maximum likelihood. This would not prevent one from concluding, if after much experience, it were found that some simplification of the formula were possible, e.g. that the coefficient of the fourth power seldom differed significantly fromthe difference between the coefficients of the second and third powersthat ~~our~~ specification could be revised with advantage, but while any specification is in use the problem of estimation can be made perfectly definite and it is primarilyfrom the advantage of separating all the complications of the theory of Estimation from the logical diffulties of specification, that I prefer to think of the thing in these two stages. In practice, as we know, simplicity is sought in specification, but whether iti is to be justified primarily by convenience, subject to the satisfaction of tests of goodness of fit or whether, as you have proposed, it should more properly be justified in terms of probability, I have not any strong opinion. If our object were to explain the habits of a group of organisms, known as scientific men, I do not see that convenience is a less effective explanation, though certainly a less distinguished one than probability.

This is enough and more for one letter. I am returning the drafts.

Yours sincerely,