15th August, 1956.

My dear Geof,

I was very glad to see your paper which Ted sent me, and especially that he and you are inclined to carry over this business of dispersion on a sphere from the exploratory mathematical phase to the equally exploratory phase of experimental design. I agree with you entirely about Gibbs, who is quite the most original mind produced in the United States and, characteristically, not much appreciated in that country, where they let the journalists for the most part do the judging.

I think the kind of consideration you give of thermodynamic type is sufficient to show that if positions on a sphere have any simple distribution, it must be just that which we have been using. However, in reality it is almost certainly not fully simple, and compounding distributions of a sphere does not, so far as I have seen, lead to any simple result. Probably a compound of the distribution we use, with a ~~whole~~ pole uniformly distributed round a circle, for secular variation, might actually be more realistic, though I should think scarcely worth developing beyond its most obtrusive statistical effects.

About "Scientific Inference", my book on this should be out by now as I have received an advance copy, but not all the other copies I have ordered. You will see that I think that rather simple semantic considerations on the word 'probability' will do a good deal to clear up the mess introduced by the axiomatic approach. In fact, if I were contrasting any distinguished mathematician with Gibbs, in parallel with the statistical situation, I should choose Hilbert rather than Maxwell, but perhaps some East European type would be better still!

I have recently written a note to the Statistical Journal pointing out that one of the tables published by Pearson and Hartley, as a substitute for Behrens' solution of the problem of two normal samples, is rather wildly wrong numerically. I suspect what they have done, as is usual in these cases, is to take the frequency in the wrong reference set, i.e. that although the table is entered with the ratio s_1/s_2, therefore each entry can refer properly only to samples having this ratio, ~~if~~ yet they have equated the level of significance to the proportion found in a much more extensive set of samples in which this ratio is ignored.

This failure to recognize appropriate subsets is, I think, one of the most mischievous fruits of the Neyman-Pearson approach.

Sincerely yours,