1 May 1931

Professor H. Hotelling,
Dept. of Mathematics,
Stanford University,
Stanford, Calif.,
U.S.A.

Dear Hotelling,

I was glad to see from your letter that I may hope to meet you at Minneapolis towards the end of August. Term there ends on August 29th, and I shall have to leave almost as soon as possible, so that I hope you will be able to pass there shortly before.

It is good to hear that the book is coming on well, and I shall be particularly interested to see the developments you mention in connection with the $z$ distribution. Is it not a curious fact historically that the first appearance of the normal distribution was as an approximation to the binomial series, whereas if Bernoulli and De Moivre had been more advanced analysts and had got the expression for the true sum of a broken binomial series, they would have found the integral of the $x$-distribution?
About semi-invariants I find that Thiele actually defines them as the coefficients of \( \frac{e^{tx}}{x^k} \) in the expansion of

\[
\frac{1}{x^n} \mathcal{S}(e^{tx})
\]

where \( x \) is the value observed in the sample. They are therefore in his usage primarily statistics, and only by the current contemporary confusion between statistics and parameters were used to designate the population values.

If the point is to be governed by historical accuracy, therefore, now that the distribution has been made and is regarded as essential by statisticians, the term "semi-invariants" should be confined to the series of statistics

\[
\mu_1, \mu_2, \mu_3, \mu_4 - 3\mu_2^2, \ldots
\]

while "cumulants" is the first term to be given specifically to the population parameters, of which Thiele's semi-invariants may be regarded as estimates. As such they are consistent, but as Craig's work has shown, introduce serious algebraic difficulties. Still it is all to the good to have a distinctive name for them.

Yours sincerely,