February 16, 1937

Dear Fraser Roberts,

I have your letter of February 14th with its interesting puzzle. If $\bar{x}$ stands for the mean I.B. of a sample of 192, and $\bar{y}$ for the mean I.Q., and further $\bar{x}'$ is the mean of the whole 2553 children, I think that what you want the sampling variance of is

$$\bar{y} - c (\bar{x}' - \bar{x})$$

If $\bar{x}'$ were known with certainty this would be

$$\sqrt{(\bar{y})(\bar{x}' - \bar{x})^2}$$

i.e., the variance of the estimate $\bar{y}$ at the given value $\bar{x}'$

Since $\bar{x}'$ is itself liable to error, I suppose one must add a term, I hope only a small one, representing variance due to this cause. I think this would be met by adding the term $c^2 \sqrt{(\bar{x})}$, ignoring the fact that your sample is part of the larger group, for, even if the group were identical, the variance of $\bar{y}$ in the formula is calculated from $s^2$, the estimated variance in arrays, and not from the whole variance of $y$. The second term also is derived from the variance within arrays, but the variance of $\bar{x}$ is, of course

$$\frac{1}{2553} (34.985)^2.$$
For the standard deviation I think you want the variance of \( b \cdot \sigma_x \)

and if the two factors were determined independently this would be

\[
\sigma_x^2 \cdot \text{Var}(b) + b^2 \cdot \text{Var}(\sigma_x')
\]

I think this is all right.

Yours sincerely,