February 24, 1941

Dear Mr Lyle,

If from each of the observed variances you subtract

$$E = \tfrac{1}{2}(\sigma_x^2 + \sigma_y^2) - \tfrac{1}{2}\sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4r^2\sigma_x^2\sigma_y^2}$$

the three quantities

$$\sigma_x^2 - e, \quad \sigma_x^2\sigma_y^2, \text{ and } \sigma^2 y - e$$

are in geometrical progression, the common ratio of which is the slope of the so-called regression line.

E.g., if σ_x^2 is 7, $r\sigma x\sigma y$ is 6, and σ_y^2 is 12 the e is $\tfrac{1}{2}(19 - \quad 5^2 + 12^2) = 3$, so one has the geometrical progression 4, 6, 9, and y changes by three parts when x changes by 2.

This, of course, is only appropriate when x and y are similar quantities, such as readings before and after a process, having, as often occurs in such cases, errors of the same magnitude, independent of each other, In this case two values lying on the line found altered by independent errors having the same variance will reproduce the sums of squares and products observed.

Yours sincerely,