The minimum coefficient of variation has been considered as a criterion for determining the transformation in the family \((X+C)^k\) which comes the closest to achieving a normal distribution. If \(y_1, \ldots, y_n\) are positive, however, the coefficient of variation of \(y_1^k, \ldots, y_n^k\) achieves its minimum value at \(k=0\) and always increases as \(k\) departs from zero in either direction.
The coefficient of variation has been considered by Rao [1] as a criterion for measuring the closeness to normality of the transformed chance variable \( Z = (X+C)^k \). On the basis of empirical sampling from a normal distribution for \( X \), Rao observed that the sample coefficient of variation as a function of the integer \( k \neq 0 \) attained its minimum value at either \( k = -1 \) or \( k = +1 \). We shall show below that this property of the coefficient of variation of a simple sample is unrelated to the fact that the sample came from a normal population and is merely a consequence of the fact that the \( n \) sample values are positive numbers.

We shall prove that if \( Y_1 = X_1 + C, \ldots, Y_n = X_n + C \) are positive numbers then the ratio

\[
f(k; Y_1, \ldots, Y_n) = \frac{\sum Y_i^k}{\left( \sum Y_i^2 \right)^{\frac{3}{2}}}\]

as a function of \( k \) attains its minimum value at \( k = 0 \) and increased as \( k \) departs from 0 in either direction. The same property can then be ascribed to the coefficient of variation,

\[
\text{C.V.} = \sqrt{\frac{1}{n-1} \left[ \sum Y_i^k - \frac{1}{n} (\sum Y_i^2)^{\frac{3}{2}} \right]} = \sqrt{\frac{n^3}{n-1} \left[ \frac{\sum Y_i^k}{(\sum Y_i^2)^{\frac{3}{2}}} - \frac{1}{n} \right]}
\]

with a minimum value of 0 at \( k = 0 \).

This result is obtained by showing that for fixed positive \( Y_1, \ldots, Y_n \) the derivative of \( f(k; Y_1, \ldots, Y_n) \) with respect to \( k \) is positive when \( k > 0 \) and negative when \( k < 0 \).

\[
\frac{df}{dk} = \frac{2(\sum Y_i^2)^2 \sum Y_i^k \log Y_i - 2\sum Y_i^k \sum Y_i^{k-2} \sum Y_i^k \log Y_i}{(\sum Y_i^k)^4}
\]

\[
= \frac{2}{(\sum Y_i^2)^3} \left[ \sum Y_i^{k-2} \sum Y_i^k \log Y_i - \sum Y_i^k \sum Y_i^k \log Y_i \right]
\]
\[
\begin{align*}
\frac{2}{(\Sigma Y_i^k)^3} \left\{ \sum_{i \neq j} Y_i^k Y_j^k \log Y_j - \sum_{i \neq j} Y_i^k Y_j^k \log Y_i \right\} \\
= \frac{2}{(\Sigma Y_i^k)^3} \left\{ \sum_{i < j} Y_i^k Y_j^k \log Y_j + Y_i^k Y_j^k \log Y_i - Y_i^k Y_j^k \log Y_i - Y_j^k Y_i^k \log Y_j \right\} \\
= \frac{2}{(\Sigma Y_i^k)^3} \left\{ \sum_{i < j} Y_i^k Y_j^k \log Y_j + Y_i^k Y_j^k \log Y_i - Y_i^k Y_j^k \log Y_i - Y_j^k Y_i^k \log Y_j \right\} \\
= \frac{2}{(\Sigma Y_i^k)^3} \left\{ \sum_{i < j} \left( Y_i^k - Y_j^k \right) \left( \log Y_i - \log Y_j \right) \right\}
\end{align*}
\]

Now if \( k > 0 \) then \( \frac{df}{dk} > 0 \) since for every pair \( i, j \) (\( i < j \)),
\[
\sgn(Y_i^k - Y_j^k) = \sgn \log \frac{Y_i}{Y_j},
\]

and if \( k < 0 \) then \( \frac{df}{dk} < 0 \) since in this case
\[
\sgn(Y_i^k - Y_j^k) = -\sgn \log \frac{Y_i}{Y_j},
\]

Reference