NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any right or permission to manufacture, use or sell any patented invention that may in any way be related thereto.
A NOTE ON THE EXACT VARIANCE OF PRODUCTS

by

Leo A. Goodman

Columbia University
Mathematical Statistics

April 27, 1961

This research was sponsored in part by the Office of Naval Research under Contract Number Nonr-266(33), Project Number NR O42-034. Reproduction in whole or part is permitted for any purpose of the United States Government.

This note was written while the author was a Visiting Professor of Mathematical Statistics and Sociology at Columbia University on leave of absence from the University of Chicago. For their very helpful comments, the author is indebted to Gerald J. Glasser, Joseph E. Keilin, and Hilary L. Seal.
A number of readers of [2] have written the author inquiring about the possibility of generalizing the results presented there. It therefore seemed worthwhile to prepare the present brief note indicating how some of the results in [2] can be generalized.

Let \( x_1, x_2, \ldots, x_K \) be \( K \) random variables. Let us denote the expected value of \( x_i \) by \( \mathbb{E}(x_i) = x_i \), the variance of \( x_i \) by \( V_i \), and the square of the coefficient of variation of \( x_i \) by \( V_i/x_i = g_i \). (For the sake of simplicity, we assume that \( x_i \neq 0 \), although some of the results presented do not require this assumption.) We shall make use of the simple identity

\[
(1) \quad \prod_{i=1}^{K} x_i = \prod_{i=1}^{K} x_i \prod_{i=1}^{K} (\delta_i + 1) = \prod_{i=1}^{K} (\triangle_i + x_i) ,
\]

where \( \delta_i = (x_i - x_i)/x_i \) and \( \triangle_i = (x_i - x_i) \). If the \( x_i \) are mutually independent, we find using identity (1) that the variance of \( \prod_{i=1}^{K} x_i \) will be equal to

\[
(2) \quad V(\prod_{i=1}^{K} x_i) = \mathbb{E}\left\{\prod_{i=1}^{K} x_i^2\right\} - \prod_{i=1}^{K} x_i^2 = \prod_{i=1}^{K} x_i^2 \left[\prod_{i=1}^{K} (g_i + 1) - 1\right] ,
\]

which can also be written as

\[
V(\prod_{i=1}^{K} x_i) = \prod_{i=1}^{K} (V_i + x_i^2) - \prod_{i=1}^{K} x_i^2
\]

\[
(3) \quad \prod_{i=1}^{K} x_i^2 = \sum_{i,j} v_{i,j} \prod_{i' 

\[
= \prod_{i=1}^{K} x_i^2 \left[\sum_i g_i + \sum_{i,j} g_i g_j + \sum_{i,j,k} g_i g_j g_k + \ldots + g_1 g_2 \cdots g_K\right] .
\]
where the summation, $\sum_{i_1,i_2,\ldots,i_s}$, is over all values of $i_1 = i_2 = i_3 \ldots = i_s$ ranging over $1,2,\ldots,K$, and where $I_{j\neq i_1,i_2,\ldots,i_s}$ is the product over the $K$-S values different from the $S$ values $i_1,i_2,\ldots,i_s$. Equation (3) here is a generalization of equations (2) and (15) in [2] and equation (a) in [7]; equation (2) here appeared earlier in [3] where it was used to study the case where the distribution of $\prod_{i=1}^{K} x_i$ was (approximately) logarithmic-normal.

We now present an unbiased estimator of $\sqrt[\prod_{i=1}^{K} x_i}$ based on unbiased estimators, $\bar{x}_1$ and $v_1$, of $X_1$ and $V_1$, respectively, where $\bar{x}_1$ is the sample mean and $v_1$ is the sample variance in a sample of $n_1$ observations each having mean $X_1$ and variance $V_1$ $(i=1,2,\ldots,K)$. When the $K$ samples $(i=1,2,\ldots,K)$ are mutually independent, we find that

\begin{equation}
(4) \quad v(\prod_{i=1}^{K} x_i) = \prod_{i=1}^{K} (v_1 + z_i) - \prod_{i=1}^{K} z_i
\end{equation}

\[ = \prod_{i=1}^{K} \left[ \bar{x}_1^2 + v_1 (n_1 - 1)/n_1 \right] - \prod_{i=1}^{K} \left[ \bar{x}_1^2 - v_1 / n_1 \right] \]

is an unbiased estimator of $V(\prod_{i=1}^{K} x_i)$, where $z_i = \bar{x}_1^2 - v_1 / n_1$. 
This follows from the fact that \( E(\bar{x}_i^2) - X^2 = v_1/n_i \). Equation (4) here is a generalization of equation (5) in [2].

The case where the \( x_i \) are not mutually independent is more complicated. From identity (1) we see that the variance of \( \prod_{i=1}^{K} x_i \) is

\[
(5) \quad V(\prod_{i=1}^{K} x_i) = \prod_{i=1}^{K} x_i^2 \left[ E \left\{ \prod_{i=1}^{K} (\delta_i + 1)^2 \right\} - B^2 \right] = E\left\{ \prod_{i=1}^{K} (\Delta_i + x_i)^2 \right\} - M^2,
\]

where \( M = E\left\{ \prod_{i=1}^{K} x_i \right\} \) and \( B = N/\prod_{i=1}^{K} X_i \). The special case of (5) where \( K = 2 \) was studied in [2]. We now consider the case where \( K = 3 \). By straightforward calculation, we find that, when \( K = 3 \), equation (5) can be rewritten as

\[
(6) \quad V(\prod_{i=1}^{K} x_i) = \prod_{i=1}^{3} X_i \left[ \sum_{i=1}^{3} G_i + (B-1)(3-B) + \frac{2}{j,k,l} \right] E \left\{ \delta_j \delta_k \delta_l \right\} h(j,k,l)
\]

where the indices \( j, k, l \) range over the values 0, 1, 2, and where \( h(j,k,l) \) is a symmetric function of \( j, k, l \) having the following values:

\[
h(j,k,l) = \begin{cases} 
0 & \text{for } (j,k,l) = (0,0,0), (0,0,1), (0,1,1) \\
1 & \text{for } (j,k,l) = (0,2,2), (2,2,2) \\
2 & \text{for } (j,k,l) = (0,1,2), (1,2,2) \\
4 & \text{for } (j,k,l) = (1,1,1), (1,1,2)
\end{cases}
\]
Equation (6) here is a generalization of equation (13) in [2], the formula given there for the variance of the product of two random variables (not necessarily independent). In the same way that equation (18) was used in [2] to derive other variance formulas for various product estimators (e.g., equations (20) and (21) in [2]), equation (6) here can also be used to derive other variance formulas for product estimators where, for example, three estimators (rather than two) are multiplied together. We shall not go into these details in this brief note.

References

