

Marie Hušková

The fourth moment of a general linear statistic

Commentationes Mathematicae Universitatis Carolinae, Vol. 9 (1968), No. 2, 325--327

Persistent URL: <http://dml.cz/dmlcz/105184>

Terms of use:

© Charles University in Prague, Faculty of Mathematics and Physics, 1968

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

THE FOURTH MOMENT OF A GENERAL LINEAR STATISTIC

Marie HUŠKOVÁ, Praha

The purpose of this paper is to derive the fourth moment of the general linear rank statistic. Let R_1, R_2, \dots, R_N be ranks of independent random variables X_1, X_2, \dots, X_n . Let

$$S = \sum_{i=1}^N a(i, R_i)$$

be a general linear rank statistic. For computation it is suitable to put

$$d(i, j) = a(i, j) - a(i, \cdot) - a(\cdot, j) + \bar{a} ,$$

where

$$a(i, \cdot) = \frac{1}{N} \sum_{j=1}^N a(i, j) ,$$

$$a(\cdot, j) = \frac{1}{N} \sum_{i=1}^N a(i, j) ,$$

$$\bar{a} = \frac{1}{N} \sum_{j=1}^N a(\cdot, j) .$$

Evidently

$$(1) \quad \sum_{i=1}^N d(i, j) = \sum_{j=1}^N d(i, j) = 0 .$$

We shall formulate our result in the following theorem:

Theorem: If the distributions of X_1, X_2, \dots, X_N are continuous and equal one another, then

$$\begin{aligned}
E(S-ES)^4 &= \frac{3(N^2-3N+1)}{N(N-1)(N-2)(N-3)} \left(\sum_{i=1}^N \sum_{j=1}^N d^2(i,j) \right)^2 - \\
&- \frac{3}{(N-2)(N-3)} \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N (d^2(i,j)d^2(i,k) + d^2(i,j)d^2(k,j)) + \\
&+ \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^N \sum_{j=1}^N d^4(i,j) + \\
&+ \frac{6}{N(N-1)(N-2)(N-3)} \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N d(i,k)d(i,l)d(j,k)d(j,l).
\end{aligned}$$

Proof: From the definition of the fourth moment we receive

$$\begin{aligned}
(2) \quad E(S-ES)^4 &= E\left(\sum_{i=1}^N d(i, R_i)\right)^4 = \sum_{i=1}^N E d^4(i, R_i) + 4 \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N E d(j, R_j) d^3(i, R_i) + \\
&+ 3 \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N E d^2(i, R_i) d^2(j, R_j) + 6 \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{\substack{k=1 \\ k \neq i \\ k \neq j}}^N E d^2(i, R_i) d(j, R_j) d(k, R_k) + \\
&+ \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{\substack{k=1 \\ k \neq i \\ k \neq j}}^N \sum_{\substack{l=1 \\ l \neq i \\ l \neq j \\ l \neq k}}^N E d(i, R_i) d(j, R_j) d(k, R_k) d(l, R_l).
\end{aligned}$$

Now we shall compute each member of the right side separately. The first member can be arranged as follows

$$\sum_{i=1}^N E d^4(i, R_i) = \frac{1}{N} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N d^4(i, j).$$

If we use (1) and relations

$$\begin{aligned}
P(R_i = k, R_j = l) &= \frac{1}{N(N-1)} \quad \text{if } k \neq l, i \neq j, \\
&= 0 \quad \text{if } k = l, i \neq j,
\end{aligned}$$

then the second member gets the form

$$\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N E d^3(i, R_i) d(j, R_j) = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{\substack{k=1 \\ k \neq i}}^N d^4(i, k).$$

Analogously we obtain the following formulas for the

remaining members

$$\begin{aligned}
 & \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k=1}^N E d^2(i, R_i) d^2(j, R_j) = \frac{1}{N(N-1)} \left(\left(\sum_{i=1}^N \sum_{k=1}^N d^2(i, k) \right)^2 - \right. \\
 & \left. - \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N d^2(i, k) d^2(j, k) - \sum_{i=1}^N \sum_{k=1}^N \sum_{h=1}^N d^2(i, k) d^2(i, h) + \sum_{i=1}^N \sum_{k=1}^N d^4(i, k) \right), \\
 & \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{\substack{k=1 \\ k \neq j}}^N E d^2(i, R_i) d(j, R_j) d(k, R_k) = \frac{1}{N(N-1)(N-2)} \left(\left(\sum_{i=1}^N \sum_{k=1}^N d^2(i, k) \right)^2 - \right. \\
 & \left. - 2 \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N d^2(i, k) d^2(j, k) - 2 \sum_{i=1}^N \sum_{k=1}^N \sum_{h=1}^N d^2(i, k) d^2(i, h) + 4 \sum_{i=1}^N \sum_{k=1}^N d^4(i, k) \right), \\
 & \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{\substack{k=1 \\ k \neq j}}^N \sum_{\substack{h=1 \\ h \neq k}}^N E d(i, R_i) d(j, R_j) d(k, R_k) d(h, R_h) = \\
 & \quad \substack{j \neq i \quad k \neq i \\ k \neq j \quad h \neq j \\ h \neq k} \\
 & = \frac{1}{N(N-1)(N-2)(N-3)} \left(6 \sum_{i=1}^N \sum_{k=1}^N \sum_{j=1}^N \sum_{h=1}^N d(i, k) d(i, h) d(j, k) d(j, h) + \right. \\
 & + 3 \left(\sum_{i=1}^N \sum_{k=1}^N d^2(i, k) \right)^2 - 18 \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N d^2(i, k) d^2(j, k) - \\
 & \left. - 18 \sum_{i=1}^N \sum_{k=1}^N \sum_{h=1}^N d^2(i, k) d^2(i, h) + 36 \sum_{i=1}^N \sum_{k=1}^N d^4(i, k) \right).
 \end{aligned}$$

Now it is sufficient to put these results into (2) and we shall get the requested formula.

Remark: If $a(i, j) = a_i b_j$ then our result coincides with that which appeared in the paper The correlation coefficient test, Suppl. JRSS 4(1937), 225-232, written by E.J.G. Pitman.

(Received May 23, 1968)