# ON THE DISTRIBUTION OF AVERAGES OVER THE VARIOUS LAGS OF CERTAIN STATISTICS RELATED TO THE SERIAL CORRELATION COEFFICIENTS*

by

V. K. Murthy

University of North Carolina

This research was supported by the Office of Naval Research under Contract No. Nonr-855(09) for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Institute of Statistics
Mimeograph Series No. 262
August 1960

# ON THE DISTRIBUTION OF AVERAGES OVER THE VARIOUS LAGS OF CERTAIN STATISTICS RELATED TO THE SERIAL CORRELATION COEFFICIENTS*

By

V. K. Murthy

University of North Carolina, Chapel Hill

SUMMARY: The sum of all the different circular serial correlation coefficients defined in the usual manner, with lags of 1, 2, 3, . . . N-1 time units and a sample of $N$ successive observations, turns out to be identically equal to $-1$ while the corresponding sum of the non-circular serial correlation coefficients, defined with the sum of squares of deviations from the mean as common denominator, is identically equal to $-1/2$ . The customary definitions of the circular and non-circular serial correlation coefficients are slightly modified hereby by dropping the correction term due to the sample mean. It is shown in this note that a certain function of the average of these modified circular serial correlation coefficients and another function of the average of modified non-circular serial correlation coefficients based on a random sample of size $N$ from a normal distribution with zero mean and a fixed variance have F-distributions with $N-1$ and $1$ degrees of freedom.

Let $x_1, x_2, \ldots x_N$ be a random sample of size $N$ from a normal distribution with zero mean and unit variance.

---

*This research was supported by the office of Naval Research under Contract No. Nonr-855(09) for Research in probability and statistics. Reproduction in whole or in part for any purpose of the United States Government is permitted.

Define

(1) $$r_k = \sum_{j=1}^{N} x_j x_{j+k} \Big/ \sum_{j=1}^{N} x_j^2$$

where $x_j = x_{N+j}$ , for all j.

Let

(2) $$C = \begin{pmatrix} 0 & 1 & 0 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 0 & 0 & . & . & 0 & 1 \\ 0 & 0 & 0 & 0 & . & . & 0 & 0 \end{pmatrix}$$ , and

(3) $$E_k = \sum_{j=1}^{N} x_j x_{j+k} .$$

Then one can see that

(4) $$E_k = \frac{1}{2} \underset{\sim}{X}' \left( C^K + C^{N-K} + C'^K + C'^{N-K} \right) \underset{\sim}{X} ,$$

where

$$\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \\ . \\ . \\ X_N \end{pmatrix}$$ , and

$C'$ denotes the transpose of the matrix C.

Now

(5) $$\sum_{K=1}^{N-1} E_K = \underset{\sim}{X}' \sum_{K=1}^{N-1} \left( C^K + C'^K \right) \underset{\sim}{X}$$

But

$$(6) \quad \sum_{K=1}^{N-1} \left[ C^K + C'^K \right] = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ 1 & 1 & & \dots & 0 \end{pmatrix} = D(\text{say})$$

Therefore

$$(7) \quad \sum_{K=1}^{N-1} E_K = \underline{X}' D \underline{X}.$$

The characteristic roots of the matrix D are easily seen to be N-1 and -1, the later repeated N-1 times.

But

$$(8) \quad \sum_{K=1}^{N-1} r_k = \frac{\sum_{K=1}^{N-1} E_K}{\sum_{K=1}^{N} x_K^2} = \frac{\underline{X}' D \underline{X}}{\underline{X}' \underline{X}} .$$

If $Y_1, Y_2, \ldots Y_N$ are related to $X_1, X_2, \ldots X_N$ by a certain orthogonal transformation we find from (8) that

$\sum_{K=1}^{N-1} r_K$ is distributed like

$$(9) \quad \frac{(N-1)\, Y_1^2 - Y_2^2 - \ldots - Y_N^2}{Y_1^2 + Y_2^2 + \ldots + Y_N^2} .$$

Now defining $r_0 \equiv 1$, we get

$$\sum_{K=0}^{N-1} r_K \text{ is distributed like}$$

$$\frac{N Y_1^2}{Y_1^2 + Y_2^2 + . + Y_N^2} \quad ,$$

where $Y_1$, $Y_2$, . . $Y_N$ are normally and independently distributed variates with zero mean and a common variance.

Hence

(10) $$\bar{r} = \sum_{K=0}^{N-1} r_K \Big/ N \; ; \quad .$$

is distributed like

(11) $$Y_1^2 \Big/ Y_1^2 + . . . + Y_N^2 \quad .$$

From (10) and (11) one can easily see that

(12) $$F = \frac{1 - \bar{r}}{(N-1)\,\bar{r}} \quad ,$$

has an $F$ distribution with $N-1$ and 1 degrees of freedom.

Now, consider the non-circular serial correlation coefficient of lag K defined by

(13) $$r_K^{*} = \frac{\sum_{j=1}^{N-K} x_j x_{j+k}}{\sum_{j=1}^{N} x_j^2}$$

It is easily seen that

$$(14) \quad E_K^* = \sum_{j=1}^{N-K} x_j x_{j+k} = \frac{1}{2} \; \underset{\sim}{X}' \left( C^K + C'^K \right) \underset{\sim}{X} \; .$$

Therefore,

$$\sum_{K=1}^{N-1} r_K^* = \sum_{K=1}^{N-1} E_K^* \Big/ \sum_{j=1}^{N} x_j^2 \quad ,$$

$$= \frac{1}{2} \; \underset{\sim}{X}' \sum_{K=1}^{N-1} \left( C^K + C'^K \right) \underset{\sim}{X} \quad ,$$

$$= \frac{1}{2} \; \frac{\underset{\sim}{X}' \, D \, \underset{\sim}{X}}{\underset{\sim}{X}' \underset{\sim}{X}} \quad .$$

Hence

$$(16) \quad \bar{r}^* = \sum_{K=0}^{N-1} r_K^* \Big/ N \; ,$$

is distributed like

$$(17) \quad \frac{1}{2N} + \frac{1}{2} \quad \frac{Y_1^2}{Y_1^2 + Y_2^2 + \ldots Y_N^2} \quad ,$$

where $Y_1$ , $Y_2$, . . $Y_N$ are normally independently distributed random variables with zero mean and a common variance.

From (16) and (17) one can easily see that

$$(18) \quad F = \frac{1}{N-1} \left[ \frac{N}{2 \, N \, \bar{r}^* - 1} - 1 \right]$$

has an F - distribution with N-1 and 1 degrees of freedom. The function $F^{-\frac{1}{2}}$ of either (18) or (12) has the student distribution with N-1 degrees of freedom.

ACKNOWLEDGMENT: I thank Prof. Harold Hotelling for his comments and criticisms.