JMFSP / SPAS

Journal of Mathematical Facts and Short Papers (JMFSP)
Vol. 1 (1), 2018, pages 13 - 17.
DOI: https://dx.doi.org/10.16929/jmfsp/13.002

SPAS - SASP

A Remarkable Identity on Simple and Double summations and Application in Statistics

Gane Samb Lo

[1] LERSTAD, Gaston Berger University, Saint-Louis, Sénégal.
[2] Affiliated to LSTA, Université Pierre et Marie Curie, Paris, France
[3] Associated with African University of Sciences and Technology, AUST, Abuja, Nigeria
1178 Evanston Drive NW, T3P 0J9, Calgary, Alberta, Canada

Received on February 1, 2018; Accepted on February 15, 2018

Abstract. We establish the following identity for $(n-k+1)$ elements of a commutative ring $x_k, ..., x_n$ with $1 \leq k \leq n$:

$$\sum_{i=1}^{k}(x_{n-i+1}-x_{n-k})^2 = 2\sum_{i=1}^{k}\sum_{j=1}^{i} j(1-\delta_{ij}/2)(x_{n-i+1}-x_{n-i})(x_{n-j+1}-x_{n-j}),$$

where $\delta_{ij}=1$ if $i=j$ and 0 elsewhere, is the Kronecker symbol based on the null element and on the unity element of the ring. As an application, we prove that two known statistics in extreme values theory are the same up to a multiplicative coefficient. This fact is an important element of the discovery of thoses estimators.

Key words: remarkable identity; extreme value index estimation; moment Dekkers-Einmahl-de haan's statistic; Lo's Statistics.
AMS 2010 Mathematics Subject Classification : 05A19; 11B65

Corresponding author gane-samb.lo@ugb.edu.sn, gslo@aust.edu.ng, ganesamblo@ganesamblo.net

Résumé (French) Nous établissons l'identité remarquable suivante concernant $(n-k+1)$ éléments $x_k, ..., x_n$ d'un anneau commutatif avec $1 \leq k \leq n$:

$$\sum_{i=1}^{k}(x_{n-i+1}-x_{n-k})^2 = 2\sum_{i=1}^{k}\sum_{j=1}^{i} j(1-\delta_{ij}/2)(x_{n-i+1}-x_{n-i})(x_{n-j+1}-x_{n-j}),$$

où $\delta_{ij} = 1$ if $i = j$ et 0 sinon, est le symbole de Kronecker basé sur l'élement nul et l'élément unité de l'anneau. Une application sur deux statistiques en valeurs extremes est donnée, en montrant que l'une est un multiple de l'autre. Ce fait contintue un élément important sur l'historicité de la découverte de ces estimateurs.

1. Introduction

In this note, we report a remarkable identity and give an historical comparison of two known statistics used in the extreme value index estimation.

2. The lemma

Lemma : A Remarkable identity.

Let $1 \leq k \leq n$ be integers and let $x_k, ..., x_n$ be $(n-k+1)$ elements of a commutative ring. Then we have

$$\sum_{i=1}^{k}(x_{n-i+1}-x_{n-k})^2 = 2\sum_{i=1}^{k}\sum_{j=1}^{i} j(1-\delta_{ij}/2)(x_{n-i+1}-x_{n-i})(x_{n-j+1}-x_{n-j}); \qquad (1)$$

where $\delta_{ij} = 1$ if $i = j$ and 0 elsewhere, is the Kronecker symbol based on the null element the unity elements of the ring.

Proof of the Lemma.

We use the notation $S_r = \sum_{1 \leq i \leq k} x_{n-j+1}^r$, $r = 1, 2$. We also these two formulas :

$$\sum_{i=1}^{h} j(x_{n-j+1}-x_{n-j}) = x_n + ... + x_{n-h+1} - hx_{n-h} \qquad (2)$$

and

$$\sum_{i=1}^{k}\sum_{j=1}^{i} j(x_{n-i+1}-x_{n-i})(x_{n-j+1}-x_{n-j}) = \frac{1}{2}\left\{(\sum_{j=1}^{k} x_j)^2 - \sum_{j=1}^{k} x_j^2\right\}$$
$$= (S_1^2 - S_2)/2. \qquad (3)$$

The first of the formulas is obtained by induction and proved for $h = 2, 3$, etc.. and then, the deduction is easy to get. The second is simply deduced from the developement of the square of the sum of the $n-k+1$ numbers. Now, the second term of Formula (1) is

$$\sum_{i=1}^{k}\sum_{j=1}^{i-1} j(x_{n-i+1}-x_{n-i})(x_{n-j+1}-x_{n-j}) + \frac{1}{2}\sum_{i=1}^{k} i(x_{n-i+1}-x_{n-i})^2 \equiv A+B.$$

Next, using (2), on has

$$\begin{aligned}
A &= \sum_{i=1}^{k}(x_{n-i+1}-x_{n-i})\sum_{j=1}^{i-1} j(x_{n-j+1}-x_{n-j}) \\
&= \sum_{i=1}^{k}(x_{n-i+1}-x_{n-i})(x_n+...+x_{n-i+2}-(i-1)x_{n-i+1}) \\
&= \sum_{i=1}^{k}(x_{n-i+1}-x_{n-i})(x_n+...+x_{n-i+1}-ix_{n-i+1}) \\
&= \sum_{i=1}^{k}\sum_{j=}^{i} x_{n-i+1}x_{n-j+1} \\
&- \sum_{i=1}^{k} ix_{n-i+1}^2 - \sum_{i=1}^{k}\sum_{j=1}^{i} x_{n-i}x_{n-j+1} + \sum_{i=1}^{k} ix_{n-i}x_{n-i+1} \\
&\equiv A_{11}+A_{12}+A_{21}+A_{22}.
\end{aligned}$$

Now by the change of variables $i = h-1$, we have

$$\begin{aligned}
-A_{21} &= \sum_{h=2}^{k+1}\sum_{j=1}^{h-1} x_{n-h+1}x_{n-j+1} = \sum_{h=1}^{k}\sum_{j=1}^{h-1} x_{n-h+1}x_{n-j+1} + \sum_{j=1}^{k} x_{n-j+1}x_{n-k} \\
&= \sum_{h=1}^{k}\sum_{j=1}^{h} x_{n-h+1}x_{n-j+1} - \sum_{h=1}^{k} x_{n-j+1}^2 + \sum_{j=1}^{k} x_{n-j+1}x_{n-k} \\
&= A_{11} - S_2 + x_{n-k}S_1.
\end{aligned}$$

Further

$$2B = \sum_{i=1}^{k} ix_{n-i+1}^2 + \sum_{i=1}^{k} ix_{n-i}^2 - 2\sum_{i=1}^{k} ix_{n-i}x_{n-i+1} = B_1+B_2+B_3$$

with, by change of variable $i = h-1$,

$$\begin{aligned} B_2 &= \sum_{h=2}^{k+1}(h-1)x_{n-h+1}^2 = \sum_{h=1}^{k+1}(h-1)x_{n-h+1}^2 = \sum_{h=1}^{k}(h-1)x_{n-h+1}^2 + kx_{n-k}^2 \\ &= \sum_{h=1}^{k} hx_{n-h+1}^2 - \sum_{h=1}^{k} x_{n-h+1}^2 + kx_{n-k}^2 = \sum_{h=1}^{k} hx_{n-h+1}^2 - S_2 + kx_{n-k}^2. \end{aligned}$$

Finally, we have

$$B = \frac{1}{2}(-A_{12} - A_{21}) - \frac{1}{2}S_2 + \frac{1}{2}kx_{n-k}^2 - A_{22}$$

and the second term of 1 is

$$\begin{aligned} & A_{11} + A_{12} - A_{11} + S_2 - x_{n-k}S_1 + A_{22} - A_{12} - \frac{1}{2}S_2 + \frac{1}{2}kx_{n-k}^2 - A_{22} \\ = & \frac{S_2 - 2x_{n-k}S_1 + kx_{n-k}^2}{2}. \end{aligned}$$

This is nothing that the half of

$$\sum_{i=1}^{k}(x_{n-i+1} - x_{n-k})^2 = \sum_{i=1}^{k} x_{n-i+1}^2 - 2x_{n-k}\sum_{i=1}^{k} x_{n-i+1} + kx_{n-k}^1 = S_2 - 2x_{n-k}S_1 + kx_{n-k^2}.$$

This closes the proof.

3. A consequence in Extreme Value Index Estimation

This identity will show that two known statistics are in reality very closely linked. In the frame extreme value theorem, there exist so many estimators of the extreme value index. We will not enter into the details here. Consider $0 < X_{1,n} \leq X_{2,n} \leq ...... \leq X_{n,n}$ are the order statistics of a sample X_1, X_2, ..., X_n of size n. First Dekkers (1989) introduced the following estimator

$$\gamma_n = H_n + 1 - \frac{1}{2}\left(1 - \frac{H_n^2}{D_n}\right)^{-1}, \tag{4}$$

based of the Hill's statistic Hill (1975)

$$H_n = \frac{1}{k}\sum_{i=1}^{k}(\log X_{n-i+1,n} - \log X_{n-k+1,n}), \tag{5}$$

and the new one they proposed as

$$D_n = \frac{1}{k}\sum_{i=1}^{k}(\log X_{n-i+1,n} - \log X_{n-k+1,n})^2$$

and made a full theory of asymptotic theory. Next, Lo (1992) proposed the following statistic, whose square-root is an estimator of the extreme value index when it is non-negative,

$$L_n = \frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{i} j(1-\delta_{ij}/2)(\log X_{n-i+1,n} - \log X_{n-i,n})(\log X_{n-j+1,n} - \log X_{n-j,n}) \tag{6}$$

where δ_{ij} denoting the Kronecker symbol, and used the couple $(H_n/A_n^{1/2}, H_n)$ to discriminate the domain of attraction with a full theory of asymptotic theory. Clearly, this remarkable identity shows that

$$D_n = 2L_n. \tag{7}$$

Dates of Finding and Publication. The dates used here are those of the publication of the statistics. Due to the delays of publication, we may easily conceive that these statistics have been found years before. The statistic in 7 has been found during the period of 1986-1987 in the preparation of a *Thèse d'Etat*.

References

Dekkers, A.L.M., Einmahl, J.H.J. and De Haan, L.(1989). A moment estimator for the index of an extreme value distribution. *Ann. of Statist.* 17. (4), 1833-1855.

Hill, B.(1975). A simple general approach to the inference about the tail index of a distribution. *Ann Statist.* 3, 1163-1174.

Lo, G.S.(1991). Empirical characterisation of the extremes I, Thèse d'Etat. Available arxiv

Lo, G.S.(1992). Sur la caractérisation empirique des extrêmes. *C.R. Math. Acad. Sci. Canada*; Vol XIV, 2,3, pp.89-94.