Ratio Mathematica 11
(1996), 47 - 57

HIGHER ORDER MOMENTS OF A SUM OF RANDOM VARIABLES: REMARKS AND APPLICATIONS

Luisa Tibiletti*, Ernesto Volpe di Prignano**

SUNTO - I momenti di una somma di variabili casuali dipende sia dai momenti puri che da quelli misti. In questa nota si introduce un indice relativo dell'importanza di questi ultimi. Fissati i momenti puri, si individua la relazione funzionale tra gli addendi che conduce ai valori estremi. Sono, infine, suggerite delle applicazioni in Finanza, Teoria delle Decisioni e Scienze Attuariali.

ABSTRACT - The moments of a sum of random variables depend on both the pure moments of each random addendum and on the addendum mixed moments. In this note we introduce a simple measure to evaluate the relative importance to attach to the latter. Once the pure moments are fixed, the functional relation between the random addenda leading to the extreme values is also provided. Applications to Finance, Decision Theory and Actuarial Sciences are also suggested.

KEYWORDS - Higher Order Moments; Cauchy-Schwarz Inequality; Pearson Correlation Coefficient.

---

* Dipartimento di Statistica e Matematica applicata alle Scienze Umane "Diego de Castro", Università di Torino, Piazza Arbarello 8, I-10122 Torino (Italy), Tel. +39-11-670.6229, Fax +39-11-670.6238.
e-mail TIBILETTI@CSIVMS.CSI.IT
** Dipartimento di Matematica per le Decisioni Economiche Finanziarie e Assicurative, Università "La Sapienza", Via Castro Laureziano 9, I-00161 Roma (Italy), Tel. +39-6-4959361, Fax +39-6-49766765

INTRODUCTION[1]

The n-th order moment $(n \geq 2)$ of a sum of random variables depend on both the n-th moments of each random addendum and on the mixed moments of the addenda. The aim of this note is to go a step further in the investigation of the latter influence. A new index which is a relative measure of the importance of the mixed moments in the determination of the sum moments is introduced. As the Pearson coefficient is a measure of the *linear dependence* strength between the two random addenda, so the new index can be interpreted as a measure of strength of a particular *non-linear dependence.*

This index can also be a useful tool for applications. Some examples of these concerning Finance, Decision Theory and Actuarial Sciences are provided.

The plan of the paper is as follows. In Section 1 the problem is explained. In Section 2 the higher order index is introduced and in Section 3 some special cases are discussed. Further remarks involving the sum of m, $m \geq 2$, random variables are pointed out in Section 4. Section 5 is devoted to applications.

1. THE PROBLEM

Let X and Y be two non-degenere random variables (r.v.'s) endowed with finite moments $\mathrm{E}\left(X^{n-k}Y^k\right)$ for $n \geq 2$ and $k = 0,..,n$. Then, the n-th order moment of the sum $S = X + Y$ is given by

$$\mathrm{E}(X+Y)^n = \mathrm{E}\left(X^n\right) + \mathrm{E}\left(Y^n\right) + \sum_{k=1}^{n-1}\binom{n}{k}\mathrm{E}\left(X^{n-k}Y^k\right). \qquad (1)$$

Above displays the fact that the n-th order moment of S depends on both the n-th order pure moments of X and Y and also on a linear combination of their mixed moments. We are aimed to explore the rôle played by the dependence structure between X and Y in the determination of the latter.

[1] Although the paper is co-authored we specify that Sections 2, 3, 4, 5 are to be attributed to Luisa Tibiletti, and Section 1 to E. Volpe di Prignano.

2. A RELATIVE HIGHER ORDER INDEX

It is well-known that the second moment of a sum is given by

$$\mathrm{E}(X+Y)^2 = \mathrm{E}\left(X^2\right)+\mathrm{E}\left(Y^2\right)+2\mathrm{E}(XY) \tag{2}$$

Denoting by

$$r^{(2)} = \frac{\mathrm{E}(XY)}{\sqrt{\mathrm{E}\left(X^2\right)}\sqrt{\mathrm{E}\left(Y^2\right)}}$$

we can write (2) as

$$\mathrm{E}(X+Y)^2 = \mathrm{E}\left(X^2\right)+\mathrm{E}\left(Y^2\right)+2r^{(2)}\sqrt{\mathrm{E}\left(X^2\right)}\sqrt{\mathrm{E}\left(Y^2\right)}.$$

By the Cauchy-Schwarz inequality, $r^{(2)}$ is bounded $(-1 \le r^{(2)} \le 1)$. Consequently, $r^{(2)}$ provides a simple *relative measure* of the influence of the mixed moment over the *sum variance*, whenever the second moments of X and Y are fixed. This result suggests a way for defining an analogous index for $n > 2$. In fact, by means of the Cauchy-Schwarz inequality we have

$$\left|\sum_{k=1}^{n-1}\binom{n}{k}\mathrm{E}\left(X^{n-k}Y^k\right)\right| \le \sum_{k=1}^{n-1}\binom{n}{k}\mathrm{E}\left|X^{n-k}Y^k\right| \le \sum_{k=1}^{n-1}\binom{n}{k}\sqrt{\mathrm{E}\left(X^{2(n-k)}\right)}\sqrt{\mathrm{E}\left(Y^{2k}\right)}$$

It is worth noting that the right-hand-side of the above inequality depends only on the moments of X and Y and no mixed moments appear. Let us define the index

$$r^{(n)} = r^{(n)}(X,Y) = \frac{\sum_{k=1}^{n-1}\binom{n}{k}\mathrm{E}\left(X^{n-k}Y^k\right)}{\sum_{k=1}^{n-1}\binom{n}{k}\sqrt{\mathrm{E}\left(X^{2(n-k)}\right)}\sqrt{\mathrm{E}\left(Y^{2k}\right)}},\ n \ge 2\ .$$

Clearly, it results $-1 \le r^{(n)} \le 1$. Formula (1) can be rewritten

$$\mathrm{E}(X+Y)^n = \mathrm{E}\left(X^n\right)+\mathrm{E}\left(Y^n\right)+r^{(n)}\left\{\sum_{k=1}^{n-1}\binom{n}{k}\sqrt{\mathrm{E}\left(X^{2(n-k)}\right)}\sqrt{\mathrm{E}\left(Y^{2k}\right)}\right\} \tag{3}$$

So, once the moments until the $2(n-1)$-th order of X and Y are fixed, $r^{(n)}$ can be interpreted as a *relative measure* of the importance which is to be attached to the dependence structure between X and Y in the determination of their n-th order sum.

Remark 1. Cases $n=3$ and $n=4$ are relevant from the application view point. In fact once the pure moments of X and Y are fixed, $r^{(3)}$ and $r^{(4)}$ provide two *relative measures* of the influence of the dependence structure between X and Y on the *asymmetry* and the *kurtosis* of their sum.

Remark 2. Once the moments until the $2(n-1)$-th order of X and Y are fixed, a lower and an upper bound for $\mathrm{E}(X+Y)^n$ is given by

$$E(X+Y)^n \geq E\left(X^n\right)+E\left(Y^n\right)-\Sigma_{k=1}^{n-1}\binom{n}{k}\sqrt{E\left(X^{2(n-k)}\right)}\sqrt{E\left(Y^{2k}\right)}$$

and

$$E(X+Y)^n \leq E\left(X^n\right)+E\left(Y^n\right)+\Sigma_{k=1}^{n-1}\binom{n}{k}\sqrt{E\left(X^{2(n-k)}\right)}\sqrt{E\left(Y^{2k}\right)}$$

3. SPECIAL VALUES FOR $r^{(n)}$

It is known that the index $r^{(2)}$ is a measure of the strength of X and Y along a linear function, *i.e.*, the more X and Y tend to cluster around the graph of a linear function the more $r^{(2)}$ tends to the extreme values[2].

A spontaneous question which arises is what is the function type clustered by X and Y when the n-th order index $r^{(n)}$ tends to the extreme values. In other terms, *once the marginal distributions of X and Y are fixed,* what is the functional relation between X and Y which induces a greater influence over $\mathrm{E}(X+Y)^n$?

---

[2] That derives from the fact that $\left|r^{(2)}\right|=1$ if $|\mathrm{E}(XY)|=\sqrt{\mathrm{E}\left(X^2\right)}\sqrt{\mathrm{E}\left(Y^2\right)}$. That occurs if and only if there exists a real number α such that $\mathrm{P}\{Y=\alpha X\}=1$, see for example ROHATGI [8].

The answer can easily be derived from rewriting the following linear combination of mixed moments

$$\sum_{k=1}^{n-1}\binom{n}{k}\mathrm{E}\left(X^{n-k}Y^{k}\right)=\mathrm{E}\left(W.Z_{(n)}\right)$$

where $W = XY$ and $Z_{(n)}=\sum_{k=1}^{n-1}\binom{n}{k}\mathrm{E}\left(X^{n-k-1}Y^{k-1}\right)$. Applying the Cauchy-Schwarz inequality, we have

$$\left|\mathrm{E}\left(W.Z_{(n)}\right)\right|\leq\sqrt{\mathrm{E}\left(W^{2}\right)}\sqrt{\mathrm{E}\left(Z_{(n)}\right)^{2}}$$

with equality if and only if there exists a real number α such that $\mathrm{P}\left\{W=\alpha Z_{(n)}\right\}=1$. Therefore, the more the r.v.'s tend to be tied by the following functional relation

$$xy=\alpha\sum_{k=1}^{n-1}\binom{n}{k}x^{n-k-1}y^{k-1}, \qquad \alpha\in\Re \tag{4}$$

the more $r^{(n)}$ tends to the extreme values.
The functional relation between X and Y for the special cases $n=3$ and $n=4$ will be set out, by an analogous procedure that for $n\geq 5$ can be easily achieved.

3.1.a THIRD ORDER INDEX

Note that $Z_{(3)}=3(X+Y)$. The linear dependence between W and $Z_{(3)}$ implies that $XY=3\alpha(X+Y)$, thus

$$Y=\frac{3\alpha X}{X-3\alpha},\ \text{ where } \alpha\in\Re .$$

If $\alpha<0$, $r^{(3)}$ tends to the minimum; vice versa, it tends to the maximum if $\alpha>0$.
In conclusion, the more the r.v.'s are clustered around a hyperbole the more $r^{(3)}$ tends to the extreme values.

3.1.b FOURTH ORDER INDEX

Let $Z_{(4)} = 2\left[2(X^2+Y^2)+3XY\right]$. If a linear dependence between the two non-degenere r.v.'s W and $Z_{(4)}$ holds, there exists $\alpha \in \Re$, $\alpha \neq 0$ such that

$$XY = 2\alpha\left[2(X^2+Y^2)+3XY\right],$$

or equivalently,

$$X^2+Y^2+\beta XY = 0\text{, where } \beta = \frac{6\alpha-1}{4\alpha},\ \alpha \neq 0. \tag{5}$$

Equation (5) is solved whenever a non-zero number γ such that

$$Y = \gamma X,$$

exists[3]. If $\gamma < 0$; $r^{(4)}$ tends to -1; vice versa[4], if $\gamma > 0$ it tends to +1. Consequently, when X and Y are clustered around a linear function $r^{(2)}$ and $r^{(4)}$ tend to the extreme values.

3.2 A NECESSARY CONDITION FOR THE STOCHASTIC INDEPENDENCE

Whenever X and Y are zero-mean[5] r.v.'s, $r^{(2)}$ coincides with the Pearson correlation coefficient. The vanishing of this last, is a necessary but a non-sufficient condition for declaring X and Y independently distributed. A spontaneous question which could arise concerns the connections between the stochastic independence and the vanishing of the higher order indexes $r^{(n)}$. Simple calculations prove that the independence between two non-degenere null-mean r.v.'s implies null-indexes until the third order, nevertheless for higher orders that could *not* be still true. For example,

---

[3]Imposing $\beta = \dfrac{-\gamma^2-1}{\gamma}$, then $Y = \gamma X$ is a solution of equation (5).

[4]If $\gamma < 0$, then $\beta \geq 2$ and $\alpha < 0$. Vice versa, if $\gamma > 0$, then $\beta \leq -2$ and $\alpha > 0$.

[5]If X and Y are not zero-mean r.v.'s, we can replace them by the zero-mean r.v.'s $X^* = X - \mu_x$ and $Y^* = X - \mu_y$ where $\mathrm{E}(X) = \mu_x$, $\mathrm{E}(Y) = \mu_y$.

$r^{(4)} > 0$ (see the Appendix). That shows that the mixed moments could play an important rôle even when X and Y are independently distributed. A necessary condition for independence is stated below.

Property 3. *Let X and Y be non-degenere and null-mean r.v.'s. If the stochastic independence holds, then*

$$r^{(n)} = 0 \qquad \text{if } n \le 3 \qquad (6)$$

$$r^{(n)} = \frac{\sum_{k=2}^{n-2}\binom{n}{k}\mathrm{E}\left(X^{n-k}\right)\mathrm{E}\left(Y^{k}\right)}{\sum_{k=1}^{n-1}\binom{n}{k}\sqrt{\mathrm{E}\left(X^{2(n-k)}\right)}\sqrt{\mathrm{E}\left(Y^{2k}\right)}} \qquad \text{if } n \ge 4 \qquad (7)$$

The proof follows from the fact that $\mathrm{E}\left(X^{n-k}Y^{k}\right) = \mathrm{E}\left(X^{n-k}\right)\mathrm{E}\left(Y^{k}\right)$ for $k = 1,..,n-1$ and $\mathrm{E}(X) = \mathrm{E}(Y) = 0$.
Clearly, (6)-(7) for all $n \ge 2$ are not sufficient conditions for independence, since they only imply

$$\mathrm{E}\left(g(X).h(Y)\right) = \mathrm{E}\left(g(X)\right).\mathrm{E}\left(h(Y)\right),$$

for all power functions g and h (see the Appendix for a necessary and sufficient condition for independence).
Now, a generalisation of the Pearson coefficient can be introduced

$$\rho^{(n)}(X,Y) = r^{(n)}(X,Y) - \frac{\sum_{k=1}^{n-1}\binom{n}{k}\mathrm{E}\left(X^{n-k}\right)\mathrm{E}\left(Y^{k}\right)}{\sum_{k=1}^{n-1}\binom{n}{k}\sqrt{\mathrm{E}\left(X^{2(n-k)}\right)}\sqrt{\mathrm{E}\left(Y^{2k}\right)}}, \qquad \text{for } n \ge 2,$$

and Proposition 1 can be restated as follows.

Property 4. *Let X and Y be non-degenere and zero-mean r.v.'s. If the stochastic independence holds, then $\rho^{(n)} = 0$ for $n \ge 2$.*

4. AN EXTENSION

Above results can be easily extended to the sum of m $(m \geq 2)$ r.v.'s. $X_1,...,X_m$:

$$E(X_1+...+X_m)^n = E\left(X_1^n\right)+...+E\left(X_m^n\right)+\sum \frac{n!}{n_1!...n_m!} E\left(X_1^{n_1}....X_1^{n_m}\right), \quad n \geq 2$$

the sum being taken over each solution $(n_1,....,n_m)$ of the equation $n_1+....+n_m = n$, where $n_i \in N$ and $n_i \neq n$, for all $i = 1,..,m$. The index

$$r^{(n)}(X_1,...,X_m) = \frac{\sum \frac{n!}{n_1!...n_m!} E\left(X_1^{n_1}....X_1^{n_m}\right)}{\sum \frac{n!}{n_1!...n_m!} \sqrt{E\left(X_1^{2n_1}\right)}.....\sqrt{E\left(X_m^{2n_m}\right)}}, \qquad n \geq 2$$

can be introduced. Since $-1 \leq r^{(n)}(X_1,..,X_m) \leq 1$, it results a relative measure of the dependence structure influence.

5. APPLICATIONS

In this section we collect some topics where the above notion can be used.

5.1 ASYMMETRY OF STOCK RETURN DISTRIBUTION

A relevant issue under discussion in Finance concerns the asymmetry of the distribution of stock-returns. This topic has long been investigated in the literature and massive empirical studies have been developed (see for example LAU *et al.* [6]). As suggested by PECCATI and TIBILETTI [7], a possible explanation of this fact can be found in the presence of positive dependence between the r.v.'s involved, inducing positive values for $r^{(3)}$. Empirical exploration over more than thirty stock-returns covering three years seems to confirm this suggestion (DEANTONI and TIBILETTI [2]). This result tends to discredit a common opinion in Finance about the diversification effect on skewness (see, for example, SINGLETON and WINGENDER [12], p. 338). In fact, if $r^{(3)} > 0$, the portfolio skewness may *not* diminish as the number of different stocks increases. Moreover, $r^{(3)}$ can be used for selecting stock-portfolios with the desired asymmetry.

5.2 MOMENT PREFERENCE

A significant number of articles studies the preference of a decision maker towards the moments of a random amount. The investigation has been focused on two questions: can the direction of preference for each moment be determined by some a-priori grounds and if so, what is the preference direction for each moment? For the third moment the first question has received considerable attention (see ARDITTI [1]; and KRAUS and LITZENBERGER [5] and SIMKOWITZ and BEEDLES [11]). With reference to the second question, ARDITTI [1] and KRAUS and LITZENBERGER [5] have implied a positive preference direction for the third moment (see ROSSI and TIBILETTI [9]). For higher moments, SCOTT and HORVARTH [10] proved that under appropriate assumptions, rational risk-averse investors prefer the odd moments and dislike the even moments of the distribution. In line with these results, multidimensional asset pricing models have been set out (see, for example, HOMAIFAR and GRADDY [4]). Clearly, whenever the random amount under consideration is given by the sum of two uncertain quantities, then their proper choice must take into account the values of $r^{(n)}$.

5.3 STOCHASTIC INEQUALITIES

A large number of commonly used stochastic inequalities involves the higher order moments of the distribution (see, for example, FELLER [3], p. 152 for a fairly general method for deriving non-trivial inequalities depending *only* on the distribution moments). If the r.v. under discussion is given by a r.v.-sum then the above mentioned inequalities can be sharpened by varying the dependence structure among the r.v.-addenda.

5.4 RISK PREMIUM FOR AGGREGATE CLAIMS

Let $S = X + Y$ be a portfolio of individual random claims X and Y provided with finite mixed moments $\mathrm{E}\left(X^{n-k}Y^k\right)$ for $n \geq 2$ and $k = 0,..,n$. Denoted by u the decision maker 's utility function, the risk premium P for S can be defined via

$$\mathrm{E}\left[u(W+S)\right] = u(W-P)$$

where W is the initial reserve. If u is a sufficiently smooth function (see TIBILETTI [13]), then the risk premium admits the following expression

$$P = W - u^{-1}\left\{u(W) + u'(W)\mu_1 + \frac{u''(W)}{2}\mu_2 + \frac{u'''(W)}{3!}\mu_3 + \ldots + \frac{u^{(n)}(W)}{n!}\mu_n + o(\mu_n)\right\}$$

where μ_i stands for the i-th order moment of S. Then, using formula (3) it is possible to rewrite P in order to emphasise the contribution of the moments of X and Y from that of their dependence structure.

APPENDIX

There exists a large literature on the search for conditions guaranteeing stochastic independence. A theorem containing various results states the following. The r.v.'s X and Y are stochastically independent iff

$$\mathrm{E}\big(g(X).h(Y)\big) = \mathrm{E}\big(g(X)\big).\mathrm{E}\big(h(Y)\big)$$

for all Borel-measurable functions g and h, provided that the expectations involved exist.
This statement will be useful in the following.
Property A1. *Let X and Y be non-degenere and zero-mean r.v.'s . If stochastic independence holds, then* $r^{(4)} > 0$.
Proof. Consider

$$\mathrm{E}(X+Y)^4 = \mathrm{E}\left(X^4\right) + \mathrm{E}\left(Y^4\right) + 4\mathrm{E}\left(XY^3\right) + 4\mathrm{E}\left(X^3Y\right) + 6\mathrm{E}\left(X^2Y^2\right) =$$

$$= \mathrm{E}\left(X^4\right) + \mathrm{E}\left(Y^4\right) + 6\mathrm{E}\left(X^2\right)\mathrm{E}\left(Y^2\right)$$

therefore, it results $r^{(4)} > 0$.

Acknowledgments

The Authors are grateful to Luigi Montrucchio and Lorenzo Peccati for some helpful comments and suggestions. However, they remain the only responsible for any remaining errors.

REFERENCES

1. F., ARDITTI, *Risk and required return on equity*, Journal of Finance, March (1967), 19-36.
2. L. DEANTONI and L. TIBILETTI, *Skewness of stock-return distribution: a theoretical and empirical investigation*, presented at the 17-th EURO Working Group on Financial Modelling, Bergamo, June 1-3, 1995.
3. W. FELLER, *An Introduction to Probability Theory and Its Applications*, vol. II, New York: John Wiley & Sons, 1971.
4. G. HOMAIFAR and D.B. GRADDY, *Equity yields in models considering higher moments of the return distribution*, Applied Economics, 20 (1988), 325-334.
5. A. KRAUS and R. LITZENBERGER, *Skewness preference and the valuation of risk assets*, Journal of Finance, September (1988), 1085-1100.
6. H. LAU, J.R. WINGENDER and A.H. LAU, *On Estimating Skewness in Stock Returns*, Management Science, 35, n.9 (1989), 1139-1142.
7. L. PECCATI and L. TIBILETTI, *On the asymmetry of stock-return distribution*, presented at the 14-th EURO Working Group on Financial Modelling, Mantova, Nov. 25-27, 1993.
8. V.K. ROHATGI, *An Introduction to Probability Theory and Mathematical Statistics*, New York: John Wiley & Sons, 1976.
9. G.A. ROSSI and L. TIBILETTI, *Higher Order Polynomial Utility Functions: Advantages in their Use*, Quaderno del Dipartimento di Statistica e Matematica Applicata alle Scienze Umane "Diego de Castro", Sezione Matematica, Università di Torino, n. 1 (1996).
10. R. SCOTT and P. HORVARTH, *On The Direction of Preference for Moments of Higher Order Than The Variance*, The Journal of Finance, 35, n. 4 (1980), 915-919.
11. M.A SIMKOWITZ and W.L. BEEDLES, *Diversification in a Three Moment World*, Journal of Financial and Quantitative Analysis, 33, n. 1 (1978), 927-941.
12. J.C SINGLETON and J. WINGENDER, *Skewness Persistence in Common Stock Returns*, Journal of Financial and Quantitative Analysis, 21, n. 3 (1986), 335-341.
13. L. TIBILETTI, *Risk Premium for Higher Order Moments*, Atlantic Economic Journal, Anthology, vol. 22, n. 3, sept. (1994), 82.