Economics Bulletin

Volume 32, Issue 1

Use of the Pearson System of Frequency Curves for the Analysis of Stock Return Distributions: Evidence and Implications for the Italian Market

Fabio Pizzutilo
Dipartimento di studi aziendali e giusprivatistici, University of Bari "Aldo Moro"

Abstract

Pearson's system of continuous probability distributions is used herein to analyse return distributions of the shares in all companies listed on the Italian stock exchange. Results show that when finite time periods are examined, the type IV distribution describes the behaviour of almost all returns on stocks. The occasional exceptions to this rule appear to be linked only with the occurrence of extraordinary events in the life of a company. When an infinite time horizon is assumed, the results do not reject the hypothesis that the distributions are of type VII, which is a special, symmetrical and hyperkurtotical case of type IV distribution that subsumes the Student's t and the Cauchy distributions, and is easier to deal with in practice.

Citation: Fabio Pizzutilo, (2012) "Use of the Pearson System of Frequency Curves for the Analysis of Stock Return Distributions: Evidence and Implications for the Italian Market", *Economics Bulletin*, Vol. 32 No. 1 pp. 272-281.
Contact: Fabio Pizzutilo - fabiopizzutilo@disag.uniba.it.
Submitted: December 15, 2011. **Published:** January 20, 2012.

1. Introduction

An extensive body of empirical evidence has shown that returns on financial assets and equity portfolios are not normally distributed and that moments higher than the second are relevant to the investors' decisions and impact equilibrium prices (Mandelbrot, 1963; Arditti, 1971; Kraus and Litzenberger, 1976; Kon, 1984; Singleton and Wingender, 1986; Loretan and Phillips, 1994; Fang and Lai, 1997; Harvey and Siddique, 2000; Sun and Yan, 2003; to cite only a few). Other distributional forms have been investigated, claiming the inadequacy of the Gaussian for modelling equity returns (Fama, 1965; Praetz, 1972; Officier, 1972; Blattberg and Gonedes, 1974; Fielitz and Rozelle, 1983; Badrinath and Chatterjee, 1991; Mills, 1995; Theodossiou, 1998, Harris and Kucukozmen, 2001, among the others). Despite the robustness of these findings the most common and utilized financial models continue to rely upon the convenient assumption of normality of stock return distributions.

The purpose of the present research is to contribute to the literature on alternatives to the classical mean-variance approach to financial modelling by investigating the distributions of daily returns on Italian listed stocks. The analysis employs the Pearson system of frequency curves. The study firstly aims at verifying: a) if there is in the so-called Pearson system a type that can be reasonably employed to describe Italian stocks daily return distributions or b) the type of distribution varies from equity to equity, or c) the type of distribution depends on the sector, the size of company, the company beta and the trading volume, or d) the type of distribution varies according to the time period analysed.

While the nongaussianity of financial return distributions is nowadays an established fact, there is no agreement in existing literature on the relevance of (co)skewness in explaining financial asset returns (see for instance, Peirò, 1999; Mitton and Vorknik, 2007; Post *et al.*, 2008; Doan *et al.*, 2010). Bearing in mind Fogler and Redcliffe's (1974) notes on the sensitivity of skewness measures, the contradictory results can be, among the others, the outcome of the use of different time periods for the empirical analysis. One possible explanation is that over the short term stock return distributions are significantly skewed while over the long term tends to become less asymmetrical. Thus, we want here to test the hypothesis that assuming an infinite time horizon (or a very long one) the distributions of Italian stocks daily returns tend to be a symmetrical type of the Pearson system.

To our knowledge, an investigation of the distribution of returns using the Pearson system has only been previously undertaken by Hirschberg *et al.* (1992) and Errunza *et al.* (1996). The formers, using monthly returns for a selection of 100 U.S. stocks for the years 1977 to 1989, concluded that the majority of the stocks analysed are of type IV. Errunza *et al.* (1996) classified various national equity market monthly returns according to their Pearson type. They find that developed markets are similar and their distributional forms are of type IV. Emerging markets returns exhibit significant cross-sectional differences even if type IV best describes the majority of the sub-sample.

We adopt an approach that is different from previous literature in that we use daily prices, and cover a sample that is more extensive. We also provide some deeper insights into those cases of divergence from type IV, in order to understand why deviations occur and to try to predict when deviations are most likely. The study also focuses on the behaviour of the distributions of stock returns for a range of different time periods and verifies whether the symmetrical hypothesis can be accepted if an infinite time horizon is assumed.

The contribution of this paper to the existing literature is twofold. (1) with respect to theory, it constitute a possible basis for the development of more realistic financial models in the presence of deviations from normality of the distributions of returns. (2) With respect to practice, its implications are relevant for the design of sound strategies of asset management and for the implementation of more sensible models of Value at Risk for the Italian equity market.

2. Methodology and Data

K. Pearson (1895, 1901 and 1916) designed a system of curves of continuous probability distribution that describes observed data in mathematical terms for all the possible variations of the first four moments. He realized that all continuous probability distributions can be classified into a small number of 'family types', provided that their probability density distributions satisfy the following differential equation:

$$\frac{1}{y}\frac{dy}{dx} = \frac{(a+x)}{b_0 + b_1 x + b_2 x^2} \qquad (1)$$

where a, b_0, b_1, and b_2 are fixed coefficients.

In terms of the parameters of (1), kappa is defined as:

$$k = \frac{b_1^2}{4(b_0 b_2)} \qquad (2)$$

According to kappa and to the roots of the characteristic equation in the denominator of (1) all continuous probability distributions are classified into one of 12 types (labelled I to XII), that constitutes the so-called Pearson system. The majority of empirical distributions fall into three of these types (namely I, IV and VI). The remaining types are special cases or cases of transition (i.e., cases at the limits between types) of the main three. In the Pearson system, the Gaussian distribution is a single case among an infinite set of other possible curves and can be described as a limit function of type I, II, III, IV, V or VII.

The three main types are characterized by the following roots of the characteristic equation of the denominator of (1), and kappa:

Type I: real roots of opposite sign, $k<0$

Type IV: no real roots, $0<k<1$

Type VI: real roots of the same sign but not equal, $k>1$.

If kappa is zero the curves are symmetrical, and are type II (if the curve is leptokutotic), type VII (if the curve is hyperkurtotic), the uniform distribution or the Gaussian. If kappa has a value of 1 and the roots are identical, the distribution is of type V. Where kappa is +/- ∞ and one root is infinite, this implies a distribution of type III.

We estimated the Pearson type for all the stocks of the sample using unconditional daily returns[1] for the whole period of the analysis and for nonoverlapping sub periods of 250 observations (i.e., about a year's worth of daily closing prices) and nonoverlapping sub periods of 500, 750, 1000, 1250 and 1500 observations. If the company was listed, kappa was also calculated for the last 7, 8, 9 and 10 years.

In order to determine kappa, the parameters of (1) and (2) have been expressed in terms of central moments (see Elderton and Johnson, 1969). The confidence levels for the Pearson types estimates have been derived through a bootstrap simulation of 5000 repetition of the same length than the original series. Provided that return time-series cannot be assumed to be independently and identically distributed (i.i.d.), the bootstrap time-series were randomly

[1] $r_t = Ln\frac{P_t}{P_{t-1}}$ where r_t=day t stock return, P_t=last price working day t and P_{t-1}=last price working day t-1.

selected with fixed blocks of 250 observations. Confidence intervals were derived using the percentile method.

The data consisted of the closing prices, corrected for corporate actions and dividends, of all the ordinary shares listed on the Telematic national market of the Italian Stock Exchange on 31.10.2010 that were not classified as foreign. Companies that had first been quoted less than one year earlier, or whose shares were suspended on 31.10.10, or whose shares were suspended for more than 15 banking days as a result of conditions that did not assure the regular course of trading, were excluded from the sample. In total, 252 stocks were analysed. The analysis covered the period 1 January 1999 to 31 October 2010. The source of the data is Bloomberg. The Wilk-Shapiro test rejected the normality hypothesis for all the stocks in the sample with *p*-values less than $1{,}98e^{-10}$.

To assess the hypothesis that over infinite time horizon kappa tends to zero, and thus that daily return distributions tend to be symmetrical types of the Pearson system, a log-logarithmic regression model was employed. For all stocks for which historical data were available for the whole period under analysis, kappa was calculated for 60 incremental observations per period. The first kappa was calculated considering the observations 1 to 300 in order to reduce the bias due to a high standard error associated with the very first values of kappa. The resulting series of kappa (observ. 1-300, observ. 1-360, observ. 1-420 and so on) were regressed using the following log-logarithmic model:

$$Ln(kappa_i) = \beta_0 + \beta_1 \cdot Ln(numb.observ._i) + \mu_i$$

Shares whose Pearson type estimate for the whole period 1 January 1999 to 31 October 2010 is not statistically significant at 95% confidence level were excluded from this analysis. The sub-samples consisted in 95 stocks. We were interested at verifying if the regression coefficient β_1 is significantly negative at usual confidence levels as this supports the inference that Ln(kappa) decreases when the number of observations increases and thus that kappa tends to zero when time tends to infinite.

3. Results

When finite time horizons are analysed, the behaviour of the distributions of Italian stock returns is described by Pearson type IV curves.

When the whole series of daily returns is considered, only five out of 252 (1.98%) stocks may not be classified as type IV. All the distributions that are not of type IV are of type VI. Deviations from type IV seem to occur when an extraordinary event (such as a corporate action, rumours of financial difficulties, mergers and acquisitions) causes a very large positive or negative daily return. In the period that follows the ‘event’, there is in fact a strong tendency for kappa to fall in the range zero-to-one (and for the distribution thus to be type IV), which is statistically appreciable at the 1% confidence level for all companies. Even the distributional form of the five outliers is consistent with type IV when the abnormal return of the ‘event day’ is not taken into account.[2] Result is consistent with Mills (1995) who found

[2] Three cases out of five (an airline, a real estate company, a telecom) are of companies that experienced heavy financial distress and launched a capital increase and a debt restructuring plan to avoid default. During the first few days after the quotation of the warrant linked to the corporate action, the stock generated abnormally high daily returns (in all cases greater than 70%). When these abnormal returns are controlled for, these distributions may also be described as type IV. The other cases relate to: an art services company whose stocks rose +125% the first day after an inverse split of its real estate assets aimed at restructuring its financial position; a foodstuffs

either the Stock Exchange deregulation in October 1986 and the 19 October 1987 market crash quite dramatically alter the return distribution of the FTSE indices.

The estimates of all the no-type IV distributions are not statistically appreciable at usual 95% confidence level. More than 90% of the estimates of the type IV distributions are significant at 99% confidence level. Exceptions to type IV are not statistically attributable to variables such as dimension, sector, beta and trading volume at any of the usual confidence levels.

Type IV distributions fall into what is known to be a heterotypic area, which covers a set of distributions for which it is either doubtful that an adequate description may be provided by the first four moments alone (as assumed in the Pearson family types), or the use of higher moments is required. There are no common statistical distributions that fall into type IV, and the lack of a closed form of its density function makes it difficult to handle in practice. However, Merrington and E. Pearson (1958), shown that the Pearson type IV curve provides a very good approximation to the noncentral t-distribution over a wide range in the skewness-kurtosis plane.

When shorter periods are taken into consideration, type IV remains the main class into which distributions of stock returns fall. For sub periods of 250 observations, type I behaviour is observable, along with type VI. For longer sub periods, deviations from type IV occur, but only to type VI. No other Pearson family type describes the behaviour of the distribution of Italian stock returns over finite time periods. For a total of 6265 sub period tests, only 3.85% of the distributions cannot be described as type IV. Of these, 3.32% are of type VI and 0.53% of type I. Deviations from type IV are related to unusually high or low daily returns that occur in the period under analysis. Table I shows a summary of the main results.

<u>Table I. Pearson family types, Italian listed shares 1999 to 2010</u>

	No. of TEST	% TYPE IV	% OTHERS		% TYPE I	% TYPE VI
FULL PERIOD	252	98.02%*	1.98%**	of whom	0.00%	1.98%
250 OBSERV.	2377	93.61%	6.39%	of whom	1.39%	5.01%
500 OBSERV	1152	97.31%	2.69%	of whom	0.00%	2.69%
750 OBSERV.	783	96.93%	3.07%	of whom	0.00%	3.07%
1000 OBSERV.	557	98.20%	1.80%	of whom	0.00%	1.80%
1250 OBSERV.	381	98.69%	1.31%	of whom	0.00%	1.31%
1500 OBSERV.	353	98.87%	1.13%	of whom	0.00%	1.13%
LAST 7 YEARS	179	97.77%	2.23%	of whom	0.00%	2.23%
LAST 8 YEARS	171	97.66%	2.34%	of whom	0.00%	2.34%
LAST 9 YEARS	165	97.58%	2.42%	of whom	0.00%	2.42%
LAST 10 YEARS	147	97.96%	2.04%	of whom	0.00%	2.04%
TOTAL SUBPERIODS	6265	96.15%	3.85%	of whom	0.53%	3.32%

* statistically significant at 99%: 91.09%, at 95%: 2.83%, not statistically significant: 6.07%

** not statistically significant at 95% confidence level: 100%

company whose shares rose 42,4% when rumors first appeared of its interest in the acquisition of the food branch of Cirio group, which was in default at the time. In all cases, the distributions are type IV if these abnormal returns are not taken into account.

Table II. Log-log regression outputs.

β_1	*p*-value	R^2 adjusted	Number	%
$\beta_1<0$	*p*-value≤0.01	R^2adj≥0.80	20	21.05%
$\beta_1<0$	*p*-value≤0.01	0.50≤R^2adj<0.80	28	29.47%
$\beta_1<0$	*p*-value≤0.01	0.25≤R^2adj<0.50	20	21.05%
$\beta_1<0$	*p*-value≤0.01	R^2adj<0.25	2	2.11%
Total β_1<0; P-value≤1%			70	73.68%
$\beta_1<0$	0.01<*p*-value≤0.05	R^2adj≥0.80	0	0.00%
$\beta_1<0$	0.01<*p*-value≤0.05	0.50≤R^2adj<0.80	0	0.00%
$\beta_1<0$	0.01<*p*-value≤0.05	0.25≤R^2adj<0.50	0	0.00%
$\beta_1<0$	0.01<*p*-value≤0.05	R^2adj<0.25	2	2.11%
Total β_1<0; 1%<P-value≤5%			2	2.11%
$\beta_1<0$	*p*-value>0.05	R^2adj≥0.80	0	0.00%
$\beta_1<0$	*p*-value>0.05	0.50≤R^2adj<0.80	0	0.00%
$\beta_1<0$	*p*-value>0.05	0.25≤R^2adj<0.50	0	0.00%
$\beta_1<0$	*p*-value>0.05	R^2adj<0.25	8	8.42%
Total β_1<0; P-value>5%			8	8.42%
Total β_1<0			**80**	**84.21%**
$\beta_1>0$	*p*-value≤0.01	R^2adj≥0.80	0	0.00%
$\beta_1>0$	*p*-value≤0.01	0.50≤R^2adj<0.80	1	1.05%
$\beta_1>0$	*p*-value≤0.01	0.25≤R^2adj<0.50	4	4.21%
$\beta_1>0$	*p*-value≤0.01	R^2adj<0.25	2	2.11%
Total β_1>0; P-value≤1%			7	7.37%
$\beta_1>0$	0.01<*p*-value≤0.05	R^2adj≥0.80	0	0.00%
$\beta_1>0$	0.01<*p*-value≤0.05	0.50≤R^2adj<0.80	0	0.00%
$\beta_1>0$	0.01<*p*-value≤0.05	0.25≤R^2adj<0.50	0	0.00%
$\beta_1>0$	0.01<*p*-value≤0.05	R^2adj<0.25	1	1.05%
Total β_1>0; 1%<P-value≤5%			1	1.05%
$\beta_1>0$	*p*-value>0.05	R^2adj≥0.80	0	0.00%
$\beta_1>0$	*p*-value>0.05	0.50≤R^2adj<0.80	0	0.00%
$\beta_1>0$	*p*-value>0.05	0.25≤R^2adj<0.50	0	0.00%
$\beta_1>0$	*p*-value>0.05	R^2adj<0.25	7	7.37%
Total β_1>0; P-value>5%			7	7.37%
Total β_1>0			**15**	**15.79%**

The hypothesis that over an infinite time horizon Pearson kappa tends to zero, and thus that cannot be rejected the hypothesis that daily return distributions on Italian listed stocks tend to symmetrical types of the Pearson system, was tested through the log-logarithmic regression model described in Section 2. For the most part of the subsample, the hypothesis is not rejected. More than 75% of the regressed time series show a negative regression coefficient statistically significant at 95% confidence level (more than 73% is statistically significant at 99%).[3] Only 8.42% of the regressions show a positive β_1, statistically significant at usual confidence levels. The results are summarized in Table II.
When kappa is zero, this indicates in the Pearson system of frequency curves a set of symmetrical distributions: the Gaussian, the Uniform, the Pearson type II and the Pearson type VII. Wilk-Shapiro tests (see Section 2) lead to the rejection of the Gaussian as a serious candidate. For obvious reasons, the uniform distribution may also be rejected. Type II and type VII are substantially different in that the former is leptokurtotic and the latter hyperkurtotic. Our dataset show that, at any confidence level, the null hypothesis that distributions of stock returns are leptokurtotic can be rejected. Moreover type VII is a limit case of type IV while type II is a limit case of type I, which according to the results summarized in Table 1 describes the distributions of stock returns only in marginal cases when periods shorter than 2 years are considered. Thus, for the most part of the subsample here analysed, Pearson type VII curve appears to be a reasonable candidate in order to describe daily unconditional returns over an infinite (or very long) time horizon. It is noteworthy, especially for practitioners, that transitional types of the Pearson system, such as type VII, may be employed satisfactorily when the values of kappa approximate the theoretical values, yielding accurate outputs. Clearly, the approximations depend on the standard error of the function in question. Pearson type VII curve has an unlimited range in both directions, is symmetrical, is bell-shaped, and has its origin at the mean (=mode). The Pearson type VII distribution subsumes the Student's t-distribution and hence also the Cauchy distribution. In practice, type VII is easier to handle than type IV.

4. Conclusions

We employed the Pearson system of frequency curves to classify all the distributions of the returns on Italian listed shares into the Pearson system of frequency curves. Our results show that over finite time periods the distribution of almost all stock returns may be described as type IV. The occasional exceptions are linked strictly with extraordinary events that result in abnormal returns. They are more frequent if short time periods are examined, and tend to disappear when the horizon of the analysis is extended. Deviations from type IV are usually to type VI, and occasionally to type I, for periods less than 2 years.
Type IV describes a set of distributions for which is doubtful whether they can be adequately described by only the first four moments, and the use of higher moments may be required. None of the common statistical distributions are of type IV, and they are difficult to handle in practice.
When an infinite time horizon is assumed, the results does not reject the hypothesis that the distribution is symmetrical and that it is of type VII, which is a symmetrical and hyperkurtotical special case of type IV that subsumes Student's t and the Cauchy distribution and is easier to deal with in practice.

[3] These figures rise of about 5% if the analysis is restricted to the company that constitutes the FSTE/MIB All Share Index, that is the micro caps companies are excluded.

The foregoing results constitute a possible basis for the development of more realistic financial models in the presence of deviations from normality of the distributions of returns. With respect to practice, its implications are relevant for the design of sound strategies of asset management and for the implementation of more sensible models of Value at Risk for the Italian equity market.

References

Arditti, F.D. (1971) "Another Look at Mutual Fund Performance" *The Journal of Financial and Quantitative Analysis* **6**, 909- 912.

Badrinath, S.G. and S. Chatterjee (1991) "A Data-Analytic Look at Skewness and Elongation in Common-Stock-Return Distributions" *Journal of Business & Economic Statistics* **9**, 223-233.

Blattberg, R.C. and N.J. Gonedes (1974) "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices" *The Journal of Business*, 47, 244-280.

Doan, P., Lin, C.T. and R. Zurbruegg (2010) "Pricing Assets with higher moments: Evidence from the Australian and US stock markets" *Journal of International Financial Markets*, Institutions and Money **20**, 51-67.

Elderton, W.P. and N.L. Johnson (1969) *Systems of Frequency Curves*, Cambridge University Press: Cambridge.

Errunza, V., Hogan Jr, K. and S.C. Mazumdar (1996) "Behavior of international stock return distributions: a simple test of functional form" *International Review of Economics and Finance,* **5**, 51-61.

Fama, E.F. (1965) "The Behavior of Stock-Market Prices" *The Journal of Business* **38**, 34-105.

Fang, H. and T. Lai (1997) "Co-Kurtosis and Capital Asset Pricing" *The Financial review* **3**, 293-307.

Fielitz, B.D. and J.P. Rozelle (1983) "Stable Distributions and the Mixtures of Distributions Hypotheses for Common Stock Returns" *Journal of the American Statistical Association* **78**, 28-36.

Fogler, R.H. and R.C. Radcliffe (1974) "A Note on Measurement of Skewness" *The Journal of Financial and Quantitative Analysis* **9**, 485- 489.

Harris, R. and C. Kucukozmen (2001) "The Empirical Distribution of UK and US Stock Returns" *Journal of Business Finance & Accounting* **28**, 715–40.

Harvey, C.R. and A. Siddique (2000) "Conditional Skewness in Asset Pricing Tests" *The Journal of Finance* **55**, 1263-1295.

Hirschberg, J., Mazumdar, S., Slottje, D. and G. Zhang (1992) "Analysing functional forms of stock returns" *Applied financial economics* **2**, 221-227.

Kon, S.J. (1984) "Models of Stock Returns-A Comparison" *The Journal of Finance* **39**, 147-165.

Kraus, A. and R.H. Litzenberger (1976) "Skewness Preference and the Valuation of Risk Assets" *The Journal of Finance* **31**, 1085-1100.

Loretan, M. and P.C.B. Phillips (1994) "Testing the covariance stationarity of heavy tailed time series: an overview of the theory with applications to several financial data sets" *Journal of Empirical Finance* **1**, 211-248.

Mandelbrot, B. (1963) "The Variation of Certain Speculative Prices" *The Journal of Business* **36**, 394-419.

Merrington, M. and E.S. Pearson (1958) "An Approximation to the Distribution of Non-Central t" *Biometrika* **45**, 484-491.

Mills, T.C. (1995) "Modelling Skewness and Kurtosis in the London Stock Exchange FT-SE Index Return Distributions" *Journal of the Royal Statistical Society. Series D (The Statistician)* **44**, 323-332.

Mitton, T. and K. Vorkink (2007) "Equilibrium underdiversication and the preference for skewness" *Review of Financial Studies* **20**, 1255-1288.

Officier, R.R. (1972) "The distribution of stock returns" *Journal of the American Statistical Association* **67**, 807-812.

Pearson, K. (1895) "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material" *Philosophical Transactions of the Royal Society of London, Series A* **186**, 343-414.

Pearson, K. (1901) "Mathematical Contributions to the Theory of Evolution. X. Supplement to a Memoir on Skew Variation" *Philosophical Transactions of the Royal Society of London, Series A* **197**, 443-459.

Pearson, K. (1916) "Mathematical Contributions to the Theory of Evolution. XIX. Second Supplement to a Memoir on Skew Variation" *Philosophical Transactions of the Royal Society of London, Series A* **216**, 429-457.

Peirò, A. (1999) "Skewness in financial returns" *The Journal of Banking and Finance* **23**, 847-862.

Post, T., van Vliet, P. and H. Levy (2008) "Risk aversion and skewness preference" *Journal of Banking and Finance* **32**, 1178–1187.

Praetz, P.D. (1972) "The Distribution of Share Price Changes" *The Journal of Business* **45**, 49-55.

Singleton, J.C. and J. Wingender (1986) "Skewness Persistence in Common Stock Returns" *The Journal of Financial and Quantitative Analysis* **21**, 335-341.

Sun, Q. and Y. Yan (2003) "Skewness persistence with optimal portfolio selection" *Journal of Banking and Finance* **27**, 1111-1121.

Theodossiou, P. (1998) "Financial Data and the Skewed Generalized t Distribution" *Management Science* **44**, 1650–61.