

A simple graphical method to explore tail-dependence in stock-return pairs

Klaus Abberger, University of Konstanz, Germany

Abstract:

For a bivariate data set the dependence structure can not only be measured globally, for example with the Bravais-Pearson correlation coefficient, but the dependence structure can also be analyzed locally. In this article the exploration of dependencies in the tails of the bivariate distribution is discussed. For this a graphical method which is called chi-plot and which was introduced by Fisher and Switzer (1985, 2001) is used. Examples with simulated data sets illustrate that the chi-plot is suitable for the exploration of dependencies. This graphical method is then used to examine stock-return pairs. The kind of tail-dependence between returns has consequences, for example, for the calculation of the Value at Risk and should be modelled carefully. The application of the chi-plot to various daily stock-return pairs shows that different dependence structures can be found. This graph can therefore be an interesting aid for the modelling of returns.

Keywords: Association, bivariate distribution, chi-plot, copula, correlation, local dependence, tail-dependence

1 Introduction

In this article bivariate data sets are examined for their dependence structure. The interest lies in the dependences particularly in the tails of bivariate distributions. The Figures 1 and 2 show scatterplots for simulated data sets with similar empirical correlation from two well known distributions. The data in Figure 1 is from a bivariate normal distribution. The normal distribution has well known nice characteristics, such as the equality of zero correlation and independence, or elliptical

t-distribution (standardized), $r=0.7$, $df=3$

VaR	$1 - \alpha$				
a'	0.90	0.95	0.99	0.995	0.999
(1, 1)	1.746	2.508	4.839	6.219	10.871
(2, 1)	2.638	3.787	7.309	9.373	16.203

Standard normal distribution, $r=0.7$

VaR	$1 - \alpha$				
a'	0.90	0.95	0.99	0.995	0.999
(1, 1)	2.365	3.033	4.290	4.747	5.696
(2, 1)	3.582	4.594	6.500	7.186	8.639

Table 1: VaR computations for two different bivariate distributions

lines, of equal density. The form of the point cloud in Figure 1 shows the positive correlation. However, if one looks only at the right upper corner, then no essential local correlation is recognizable in this range. The situation is different in Figure 2. Here the data is from a bivariate t-distribution with two degrees of freedom. In contrast to the first figure there seems to be a positive correlation in the right upper tail. It is the aim of this article to introduce a graphical method that makes judgement of the dependence in the tails easier. This graph is called a chi-plot and was introduced by Fisher and Switzer (1985, 2001). However in their articles the authors do not discuss the problem of tail-dependence considered here. Before we define the graph, we wish to discuss whether the analysis of dependence in the tails is relevant at all.

An important example where the presence of tail dependence is relevant is the calculation of the Value-at-Risk (VaR). Let us consider a portfolio \mathbf{a} with $\mathbf{P}(t)$ the price vector of the assets at time t . The one period VaR with α confidence level is defined by $VaR = \mathbf{F}^{-1}(1 - \alpha)$ with \mathbf{F} the distribution of the random variate $\mathbf{a}'(\mathbf{P}(t) - \mathbf{P}(t-1))$.

Table 1 shows a simple example. The VaR based on the bivariate t-distribution with 3 degrees of freedom and the VaR based on the bivariate standard normal distribution, both with correlation equal to 0.7, are compared. It is obvious that the resulting values are different. However, it can be argued that the comparison of independent t-distributions with independent normal distributions shows a similar pattern because of the different tails, i.e because of the fat tails of the t-distribution. Therefore, for demonstration of the effect of tail dependence, the marginal distri-

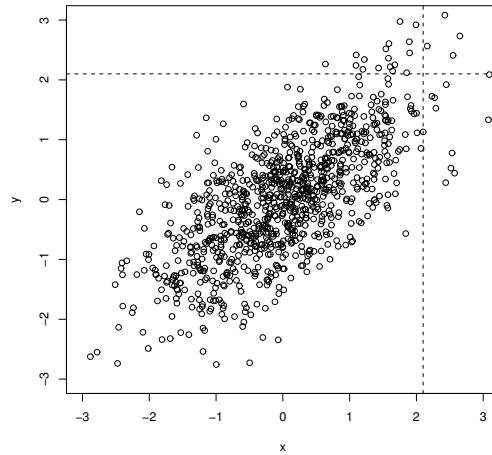


Figure 1: Scatterplot of normal distributed data (empirical correlation=0.7)

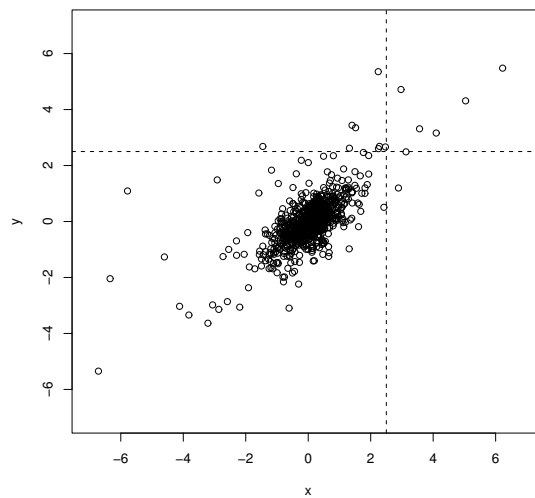


Figure 2: Scatterplot of t-distributed data with 2 degrees of freedom (df) (empirical correlation=0.668)

$$\Sigma = \begin{pmatrix} 1 & 0.82 & 0.44 & 0.36 & 0.33 \\ & 1 & 0.39 & 0.34 & 0.30 \\ & & 1 & 0.37 & 0.31 \\ & & & 1 & 0.31 \\ & & & & 1 \end{pmatrix}$$

(a) Normal marginal distributions with normal copula

	$1 - \alpha$				
a'	0.90	0.95	0.99	0.995	0.999
(1, 1, 1, 1, 1)	7.26	9.33	13.14	14.55	17.45

(b) Normal marginal distributions with Student copula, df=1

	$1 - \alpha$				
a'	0.90	0.95	0.99	0.995	0.999
(1, 1, 1, 1, 1)	5.69	7.95	13.19	15.38	20.06

Source: Bouye et al. (2001).

Table 2: VaR computations for two different five-dimensional distributions

butions and the dependence structure must be separated. This separation can be made possible by the concept of the copula function formally defined in the next section.

Table 2 shows two calculations for five dimensional random variables. In both cases the five marginal distributions are normal distributions and the correlations are equal to the indicated matrix Σ .

In example (a) of Table 2 the five marginal normal distributions are joined together with the dependence structure of the multivariate normal distribution and thus a multivariate normal distribution results. This can be compared with example (b) where the five marginal normal distributions are joined together with the dependence structure of the t-distribution with one degree of freedom. If one compares these calculations, deviations which are based on the different dependence structure in the tails can be found. This example shows how important the consideration of the tail-dependence is for the calculation of VaR.

The concrete application this paper is concerned with is the assessment of tail-dependence between pairs of daily stock returns. For this the chi-plot of GARCH

adjusted returns is used. The next section defines the chi-plot and the concept of tail-dependence, and illustrates them on simulated data sets. Section 3 contains some applications to stock returns.

2 Chi-plot, copula and tail dependence

The chi-plot defined by Fisher and Switzer (1985, 2001) is designed to provide a graph that has characteristic patterns depending on whether the variates are independent, have some degree of monotone relationship, or have more complex dependence structures. The chi-plot depends on the data only through the values of their ranks.

Let $(x_1, y_1), \dots, (x_n, y_n)$ be a random sample from H , the joint (continuous) distribution function for a pair of random variables (X, Y) , and let $1(A)$ be the indicator function of the event A . For each data point (x_i, y_i) , set

$$H_i = \sum_{j \neq i} 1(x_j \leq x_i, y_j \leq y_i) / (n - 1), \quad (1)$$

$$F_i = \sum_{j \neq i} 1(x_j \leq x_i) / (n - 1), \quad (2)$$

$$G_i = \sum_{j \neq i} 1(y_j \leq y_i) / (n - 1), \quad (3)$$

and

$$S_i = \{(F_i - 1/2)(G_i - 1/2)\}. \quad (4)$$

Now calculate

$$\chi_i = (H_i - F_i G_i) / \{F_i(1 - F_i)G_i(1 - G_i)\}^{1/2} \quad (5)$$

and

$$\lambda_i = 4S_i \max\{(F_i - \frac{1}{2})^2, (G_i - \frac{1}{2})^2\}. \quad (6)$$

The chi-plot is a scatterplot of the pairs (λ_i, χ_i) .

At each sample point, χ_i is actually a correlation coefficient between dichotomized X values and dichotomized Y values. Therefore all values of χ_i lie in the interval $[-1, 1]$. If Y is a strictly increasing function of X , then $\chi_i = 1$ for all sample cut points, and when Y is a strictly decreasing function of X , then $\chi_i = -1$ for all sample cut points. The value λ_i is a measure of the distance of the data point (x_i, y_i)

from the center of the data set. All values λ_i must lie in the interval $[-1, 1]$. When the data is a random bivariate sample from independent continuous marginals, then the values of the λ_i are individually uniformly distributed. However, when X and Y are associated, then the values of λ_i may show clustering. In particular, if X and Y are positively correlated, λ_i will tend to be positive and vice versa for negative correlation.

χ_i can be seen also as an empirical measure of the “positive quadrant dependence” (PQD). The PQD is defined as follows (Joe, 1997):

$$P(X \leq a_1, Y \leq a_2) \geq P(X \leq a_1) \cdot P(Y \leq a_2), \text{ for all } a_1, a_2 \in R, \quad (7)$$

i.e. PQD is defined globally for all $a_1, a_2 \in R$. On the other hand χ_i measures this dependence locally and plots it against the distance of the data point to the data center. Joe (1997) defines various dependence concepts and shows the relations between them. Is $(X;Y)$ for example (positively) “associated“, i.e.

$$E[g_1(X)g_2(Y)] \geq E[g_1(X)] \cdot E[g_2(Y)], \quad (8)$$

for all real valued functions g_1, g_2 which are increasing and are such that the expectations exist, then is $(X;Y)$ also PQD. Of course (8) contains the case of positive correlation.

Figure 3 contains four graphs. The upper left graph is a scatter plot of one simulated data set of the size $n = 500$ in which X and Y have a bivariate standard normal distribution with correlation $\rho = 0.5$. The upper right graph is the corresponding chi-plot of the simulated data set. The two lower graphs also show chi-plots but not for the whole data set. Only observations with positive λ_i are included. The λ values are positive for data points which lie in two quadrants, namely in the upper right corner and the lower left corner. The observations into these quadrants are shown separately in the two lower graphics. The left graph is a chi-plot for the observations in the lower left corner of the scatter plot and the right graphic is a chi-plot for the observations in the upper right corner of the scatter plot. By this separation one can recognize asymmetries in the tail-dependences. In the case at hand the tail-dependence is symmetric, what is confirmed by the plots also.

Now we look at the plot for all observations more rigorously. Positive dependence is indicated over almost all the complete range looked at. The chi-plot shows a bent course. The chi-values are near to zero on the edges of the graph. For easier

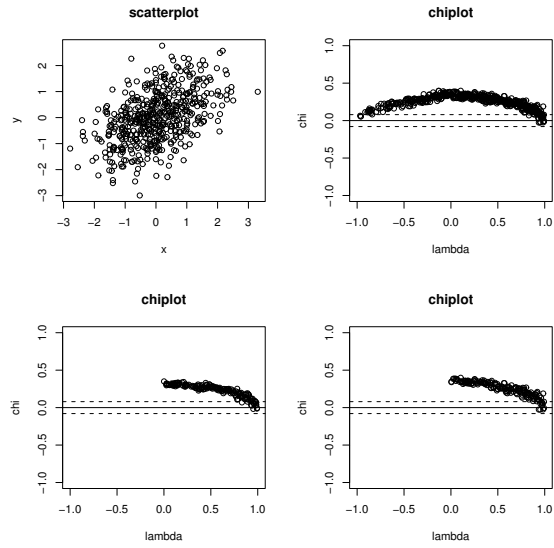


Figure 3: Scatterplot and chi-plots for normal distributed data (correlation=0.5, n=500)
`%beginfigure[p]`

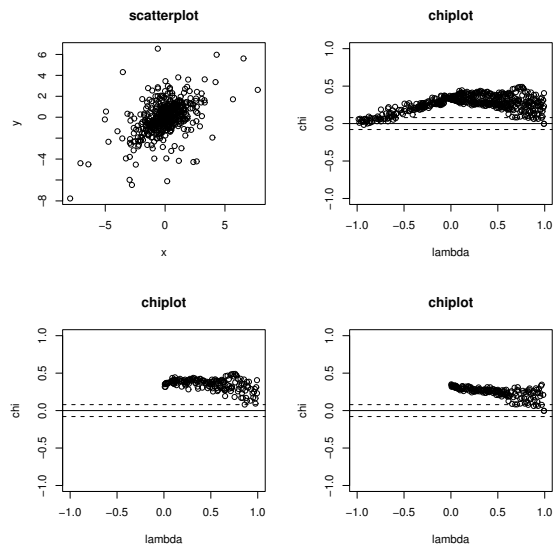


Figure 4: Scatterplot and chi-plots for t-distributed data (df=3, n=500, correlation=0.5)

interpretation the mapping is extended by two horizontal lines at $\pm c/n^{1/2}$, where c is selected so that approximately 95% of the pairs (λ_i, χ_i) lie between the lines. c can be calculated with Monte Carlo methods (Fisher and Switzer, 2001).

The bent course of the chi-plots reports the so-called “tail independence“ characteristic of the bivariate normal distribution, provided that $|\rho| < 1$. Tail dependence is often defined in terms of the copula of a joint distribution (for a complete treatment of copula see Joe (1997). For a bivariate distribution $F(x, y)$ with j -th univariate margin F_j , the copula associated with F is a distribution function $C : [0, 1]^2 \rightarrow [0, 1]$ that satisfies

$$F(x, y) = C(F_1(x), F_2(y)), x, y \in R. \quad (9)$$

So the copula is a distribution function of a random vector, $\mathbf{U} = (U_1, U_2)$, where each $U_j \sim \text{uniform}(0, 1)$. If a bivariate copula C is such that

$$\lim_{u \rightarrow 0} C(u, u)/u = \lambda_L \quad (10)$$

exists, C has lower tail dependence if $\lambda_L \in (0, 1]$, and no lower tail dependence if $\lambda_L = 0$. Similarly the upper tail dependence can be defined (see Joe, 1997).

The theoretical chi-value is zero if $P(X \leq u_1, Y \leq u_2) = P(X \leq u_1) \cdot P(Y \leq u_2)$. This is equal to

$$C(u_1, u_2) = u_1 \cdot u_2 \quad (11)$$

and

$$\frac{C(u_1, u_2)}{u_1} = u_2 \text{ resp. } \frac{C(u_1, u_2)}{u_2} = u_1. \quad (12)$$

If the copula is lower tail independent, we know that $\lim_{u \rightarrow 0} C(u, u)/u = 0$ and so the chi-values are also equal to zero. For the normal distribution it is known that there is no tail dependence (Embrechts et al., 2002) and the chi-plot shows exactly this.

The bivariate t-distribution provides an interesting contrast to the bivariate normal distribution, provided $\rho > -1$ the copula of the bivariate t-distribution is tail dependent (see also Embrechts et al., 2002).

The data underlying Figure 4 are from a t-distribution with three degrees of freedom. For the data generation the following quantities are used. At first bivariate standard normal distributed samples \mathbf{Z} with correlation 0.5 are produced. A χ^2

Copula	$C_\alpha(u, v)$	α
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\theta > 0$
Gumbel	$\exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta})$	$\theta > 0$
Gau	$\Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$	$\rho \in (-1, 1)$
Student	$T_{\nu, \rho}(T_\nu^{-1}(u), T_\nu^{-1}(v))$	$\nu \geq 2, \rho \in (-1, 1)$

Copula	Wertebereich von ρ_S	τ_L	τ_U
Clayton	(0, 1)	$2^{-1/\theta}$	0
Gumbel	(0, 1)	0	$2 - 2^{1/\theta}$
Gau	(-1, 1)	0	0
Student	(-1, 1)	$2(1 - t_{\nu+1}(\frac{\sqrt{\nu+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}))$	

Table 3: Some copulas and their characteristics

distributed variable U is then produced with $df = 3$ degrees of freedom. With $\mathbf{V} = (V_1, V_2)'$ and $V_i = \sqrt{df}Z_i/\sqrt{U}$, $i = 1, 2$, the variable \mathbf{V} is then a sample from the t-distribution with three degrees of freedom and

$$\text{cov}(\mathbf{V}) = \frac{df}{df-2}\Sigma, \quad (13)$$

Σ denoting the variance-covariance matrix of the bivariate normal variables.

In comparison with the normal distributed data, the chi-plot for the t-distributed data shows another course. The chi-plot does not incline on the right side of the graph to the zero line anymore. This shows tail-dependence on the upper right and lower left edge of the data.

The tail-dependence behaves symmetrically at both the normal distribution and the t-distribution. Table 3 shows this analytically. The table consists of a column for the copula ($C_\alpha(u, v)$), a column for the domains of the parameters (α), a column for the possible values of the Spearman correlation (ρ_S) as well as two columns for the values of the upper (τ_U) and lower (τ_L) tail-dependence. In addition to the normal copula and the student copula there are contained two copula in the table which can produce only positive Spearman correlations but which have asymmetric tail-dependence. These asymmetric cases could be relevant for the analysis of stock returns.

3 Application to stock-return pairs

The chi-plot is now applied to various pairs of daily stock returns. The returns r_t are defined by

$$r_t = \log \frac{p_t}{p_{t-1}}, \quad (14)$$

with p_t the price of the stock at time t . Often single return series are modelled with the GARCH model. The univariate GARCH model is

$$r_t = \mu + \varepsilon_t, \quad (15)$$

$$E[\varepsilon_t | \varepsilon_{t-1}] = 0, \quad (16)$$

$$V(\varepsilon_t | \varepsilon_{t-1}) = h_t = c + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}. \quad (17)$$

For details see, for example Gouriéroux (1997). In addition, in this reference the multivariate GARCH model and different variants of this model are represented.

In the following a GARCH(1,1) model with copula is used. Since pairs of returns are analyzed we have $r_{i,t}$, $i = 1, 2$, $t = 1, \dots, T$. Each series follows a GARCH(1,1) model and the bivariate distribution of errors is characterized by both marginal distributions and the copula:

$$y_{i,t} = \mu_i + \sigma_{i,t} \varepsilon_{i,t} \quad (18)$$

$$E[\varepsilon_t | \varepsilon_{t-1}] = 0, \quad (19)$$

$$V(\varepsilon_t | \varepsilon_{t-1}) = 1, \quad (20)$$

$$\sigma_{i,t}^2 = \gamma_i + \alpha_i (y_{i,t-1} - \mu_i)^2 + \beta_i (\sigma_{i,t-1})^2 \quad (21)$$

$$F(\varepsilon_{1,t}, \varepsilon_{2,t}) = C(F_1(\varepsilon_{1,t}), F_2(\varepsilon_{2,t})). \quad (22)$$

This model is a generalization of the GARCH(1,1) model with constant correlation introduced by Bollerslev (1987). In the later model the bivariate distribution is assumed to be normal with constant correlation. In the following a GARCH(1,1) model is estimated for each series separately. Then the GARCH adjusted returns are calculated and further analyzed with the chi-plot.

In Figure 5 the graphical results for the adjusted Bayer and the adjusted BASF¹ returns from June 1, 1993 until October 31, 1995 are represented. The chi-plot

¹These are German stocks. The data are provided by the Deutsche Finanzmarkt Datenbank Karlsruhe.

clearly indicates tail dependence. The tail dependence is present in both the upper right and the lower left tail of the returns distribution.

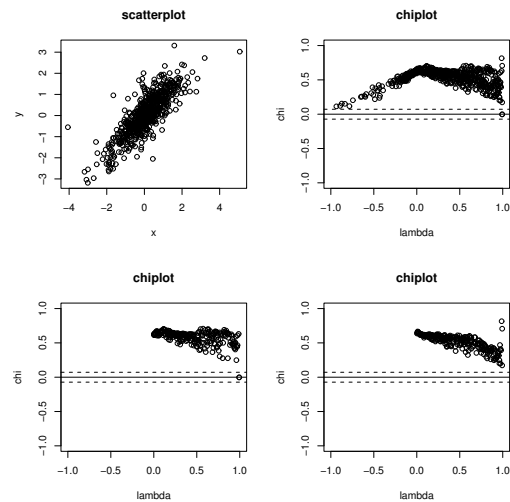


Figure 5: Scatterplot and chi-plots for adjusted Bayer and BASF return pairs

Another dependence structure result for the next return pair is shown in Figure 6. The dependence of Daimler-Benz and BMW returns is examined. The chi-plot shows that between these two returns there is no tail-dependence. Instead the plot explores two outliers. So for this example the conclusion from the plot is that there is no tail-dependence but there are two outliers. However the quality of no dependence is also very interesting for modelling the returns. Figure 7 shows the Q-Q-plot for the quantiles of the BMW returns against the theoretical normal quantiles. The departure of the points from the line indicates a fat tailed return distribution. Therefore the bivariate normal distribution cannot be a suitable assumption. Instead one should use fat tailed marginal distributions and combine them in such a way that no tail-dependence results.

Figure 8 represents the plots for BMW and Allianz. Tail-dependence is present in this pair. There the dependence is asymmetric. In the upper right corner of the scatterplot both are present, tail-dependence and outliers. In contrast to this no tail-dependence is found in the lower left corner.

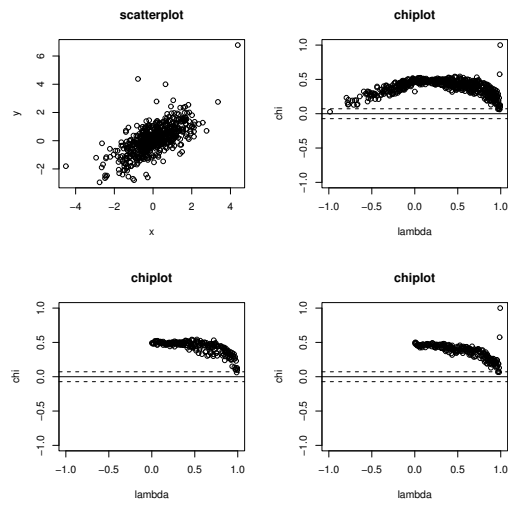


Figure 6: Scatterplot and chi-plots for adjusted Daimler-Benz and BMW return pairs

An asymmetric behaviour also is shown in the last example. Figure 9 contains the results from daily DAX returns and daily FTSE returns from 1991 until 1998. Tail-dependence is detected only in the lower left corner.

4 Conclusions

In this article the chi-plot introduced by Fisher and Switzer (1985, 2001) is used to detect tail-dependence in pairs of daily stock returns. The ability of the chi-plot to detect tail dependence is first demonstrated in some simulated date sets. In this context the concept of copula functions is introduced and the tail-dependence behaviour of the normal and the t-distribution is discussed.

Applied to various stock return pairs, different tail-dependence structures are found. The results range from no tail-dependence to symmetric tail-dependence to asymmetric tail dependence. For the common modelling of returns pairs the following questions have therefore to be answered:

- Are there tail-dependencies?

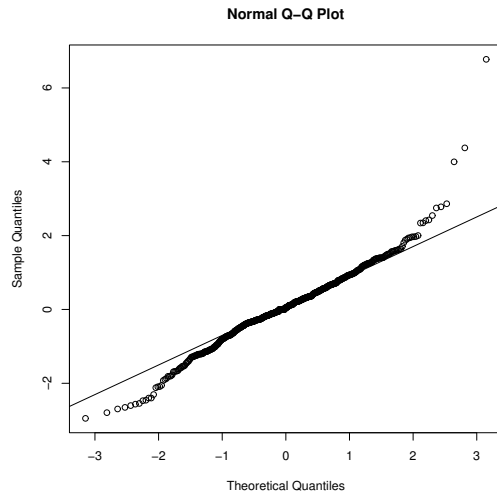


Figure 7: Q-Q plot of GARCH residuals of BMW returns against normal quantiles

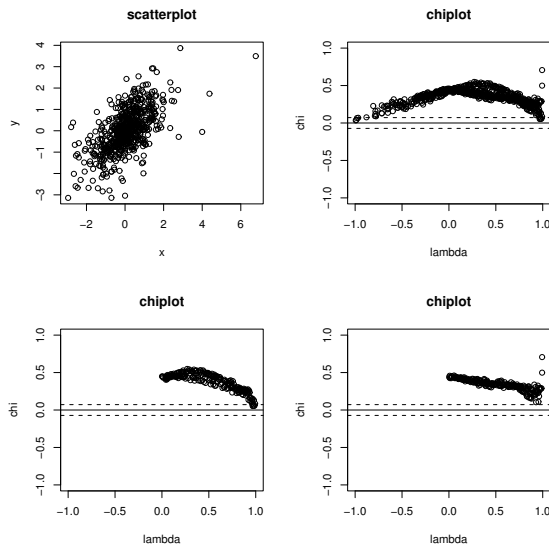


Figure 8: Scatterplot and chi-plots for adjusted BMW and Allianz return pairs

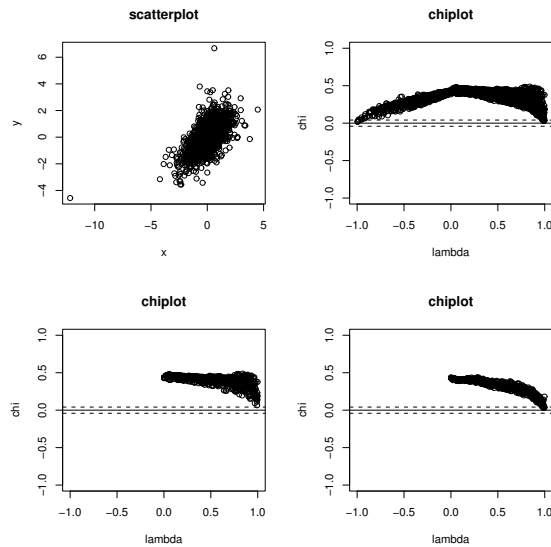


Figure 9: Scatterplot and chi-plots for adjusted DAX and FTSE returns, 1991-1998

- Are the tail-dependencies symmetric or asymmetric?
- Are there no tail-dependencies but perhaps are there outliers?

The chi-plot seems to be a very easy, fast and useful graphical instrument for exploring the dependence structures.

From the analyzed data sets it follows that a flexible approach is needed for the modelling of the data. Fortin and Kuzmics (2002) use for the copula in (22) a convex combination $C(u, v) = \sum_{i=1}^k \lambda_i C_i(u, v)$, $\sum_{i=1}^k \lambda_i = 1$, $\lambda_i \geq 0$, of k copulas. Through this one gets a very flexible approach whose robustness against outliers should be checked, however. Another approach is discussed in Longin and Solnik (2001). They use extreme value theory to model pairs of returns. Also in this case a dependence function is introduced which can have different quantities regarding the dependence structure. The chi-plot therefore makes independently of the modelling strategy, useful insights into the dependence structure possible.

5 Literature

Bollerslev T. (1987): A Multivariate GARCH Model with Constant Conditional Correlations for a Set of Exchange Rates. Diskussion Paper, Northwestern University, Evanston.

Bouye E., Durrleman V., Nikeghbali A., Ribulet G., Roncalli T. (2001): Copulas: an open field for risk management. Discussion paper, Credit Lyonnais.

Embrechts P., Hing A., Juri A. (2003): Using copulae to bound the Value-at Risk for functions of dependent risk. *Finance and Stochastics*, 7, 145-167.

Embrechts P., McNeil A., Strautmann D. (2002): Correlation and dependence in risk management: properties and pitfalls. In: *Risk Management: Value at Risk and Beyond*, ed.: M.A.H. Dempster, Cambridge University Press, Cambridge.

Fisher N.I., Switzer P. (2001): Graphical assessment of dependence: is a picture worth 100 tests? *The American Statistician*, 55, 233-239.

Fisher N.I., Switzer P. (1985): Chi-plots for assessing dependence. *Biometrika*, 72, 253-265.

Fortin I., Kuzmics C. (2002): Tail-Dependence in Stock-Return Pairs. Discussion paper, Institute for Advanced Studies, Vienna.

Gourieroux C. (1997): ARCH Models and Financial Applications. Springer Verlag, Heidelberg.

Gourieroux C., Jasiak J. (2001): Financial Econometrics. Princeton University Press, Princeton.

Hollander P.W., Wang Y.J. (1987): Dependence function for bivariate continuous densities. *Communications in Statistics, Theory and Methods*, 16, 863-876.

Joe H. (1997): Multivariate models and dependence concepts. Chapman and Hall, New York.

Longin F., Solnik B. (2001): Extreme Correlation of International Equity Markets. The Journal of Finance, 56, 649-676.

McNeil A.J., Frey R. (1999): Estimation of Tail-Related Risk Measures of Heteroskedastic Financial Time Series: an Extreme Value Approach. Discussion Paper, ETH Zurich.