Copulas, Tail dependence and Value at Risk

Rajeeva L. Karandikar
Director
Chennai Mathematical Institute
rlk@cmi.ac.in, rkarandikar@gmail.com

2 Multivariate Normal Distribution

Multivariate Normal Distribution occupies a central place in statistical theory- so much so that a traditional course on Multivariate Statistics ends up being a course on Multivariate Normal Distribution. Indeed, other than some bivariate distributions, Wishart distribution and Dirichlet distribution we do not come across other examples of multivariate distributions.

Beyond Normal distribution

Normal distribution seems to be appropriate for modeling quantities arising from natural phenomenon- such as heights of individuals, monthly rainfall in a city, transmission error in a radio signal,...

However in the first course on Econometrics we learn that for describing income distribution, Normal (rather log-normal) distribution underestimates tail of the distribution.

Beyond Normal distribution..

In other words, if we fit a log-normal distribution to Income data, it underestimates the proportion of ultra rich people. The same phenomenon is observed if we try to fit normal distribution to daily stock return data. This has been observed for stocks of various companies over different time periods and across different countries.

Beyond Normal distribution..

Let S_n be the price of stock of a company on n^{th} day. The return on n^{th} day is

$$R_n' = \frac{S_n - S_{n-1}}{S_{n-1}}$$

Note that $R_n' = \frac{S_n}{S_{n-1}} - 1 \approx -(1 - \frac{S_n}{S_{n-1}}) \approx \log(\frac{S_n}{S_{n-1}})$ if $\frac{S_n}{S_{n-1}} - 1$ is close to zero. Hence the return on log scale defined as

$$R_n = \log(S_n) - \log(S_{n-1})$$

is approximately close to R_n' as long as day to day price variations are not huge.

Beyond Normal distribution..

The model used by Black-Scholes assumes that R_n has Normal distribution. Data from various markets strongly suggests that the Normal model underestimates the tails of the true distribution- one sees deviations of over 4 σ often - which has extremely low probability under the Normal model.

Distribution with fatter tails (than normal) seem to fit the data much better- double exponential, stable distributions among others.

Beyond Normal distribution..

Let us look at some data from Mexican stock market. We look at two companies, Telmex and Cemex . The daily log-returns data from for Telmex fro 1991 and for Cemex from 1999 has been taken and in each case and percentiles, mean, median, variance etc have been shown for the empirical distribution as well as best-fit Normal, Double exponential and logistic distribution

Log returns for Telmex..

	Empirical	DoubExp	logistic	Normal
Mean	0.0006	0.0000	0.0005	0.0007
Median	0.0000	-0.0001	0.0005	0.0006
Std. dev	0.0213	0.0213	0.0212	0.0213
1st Percentile	-0.0572	-0.0592	-0.0537	-0.0483
5th Percentile	-0.0323	-0.0347	-0.0341	-0.0340
95th Percentile	0.0349	0.0348	0.0351	0.0357
99th Percentile	0.0616	0.0588	0.0549	0.0505
1percentTailVaR	-0.0731	-0.0736	-0.0646	-0.0559

Log returns for Cemex..

	Empirical	DoubExp	logistic	Normal
Mean	-0.0002	0.0002	0.0001	0.0001
Median	0.0000	0.0001	0.0001	0.0002
Std. dev	0.0284	0.0266	0.0286	0.0287
1st Percentile	-0.0806	-0.0728	-0.0725	-0.0662
5th Percentile	-0.0434	-0.0433	-0.0464	-0.0469
95th Percentile	0.0399	0.0437	0.0469	0.0474
99th Percentile	0.0758	0.0747	0.0722	0.0670
1percentTailVaR	-0.1190	-0.0905	-0.0890	-0.0764

Beyond Normal distribution..

In both the cases, we see that Normal distribution underestimates the tail probabilities by a wide margin - the 1 percentile and 99 percentiles for the best fit normal distribution are wide off the mark. The double exponential seems a better fit.

Beyond Normal distribution....

For a r.v. Z with normal distribution (variance 1),

$$P(\mid Z \mid > t) \sim C \exp(-t^2).$$

For a r.v. X with double exponential distribution,

$$P(\mid X \mid > t) \sim C \exp(-t)$$

while for a r.v. Y with Cauchy distribution,

$$P(\mid Y \mid > t) \sim C \frac{1}{t}.$$

The last is an example of a fat-tailed distribution. Other examples are, t-distribution with small degrees of freedom, Pareto distribution, stable distributions.

Beyond Normal distribution....

Since risk depends on the tail of the distribution, while the normal and observed distributions of stock prices differ the most in the tails, we may see massive underestimation of risk. So it is crucial to choose the correct distribution while computing risk.

Value-at-Risk

Traditionally, risk associated with a asset (which is stochastic) was measured by its standard deviation. In the Gaussian world, the mean and standard deviation of course capture the full picture.
Once people started examining distributions other than gaussian, need was felt for a new *risk measure*. The new measure that has been widely accepted is called Value-at-Risk.

Value-at-Risk

Consider an asset whose value at time T is modeled as S_T. If S_T has Normal distribution with mean μ and variance σ then

$$P(S_T \leq \mu - 2.33\sigma) = 0.01.$$

Then we can be assured that the loss will not be more than $V = -(\mu - 2.33\sigma)$ with 99% probability and V gives us a measure or Risk associated with the underlying asset.

Value-at-Risk

Thus in general, we look for a value V such that

$$P(S_T \leq -V) = 0.01$$

then our loss is less than or equal to V with 99% probability, or loss exceeds V with only 1% probability.

V is called the 1% Value-at-Risk or 1% VaR.

$-V$ is the lower 1^{st} percentile of the asset distribution, or V is the 99^{th} percentile of the loss distribution $L = -S_T$.

We can similarly talk about 1% VaR or α% VaR.

Value-at-Risk

VaR is not additive: VaR of sum of random variables is not equal to sum of the VaR's- it is bounded above by the sum. Thus, to obtain VaR for a portfolio we must look at the joint distribution of the underlying assets.

Value-at-Risk

According to Basel II accord, the capital adequacy for a Bank is to be determined based on VaR of its assets and so accurate measurement of VaR for a basket of assets is an important step. Using Gaussian distribution when the assets have a fatter tail results in under estimation of the VaR.

Beyond Normal distribution..

When considering a portfolio of stocks m which are possibly correlated, we need to model

$$(R_n^1, \ldots R_n^m),\ n = 1, 2, \ldots$$

We need a multivariate distribution whose marginals are say double exponential and with a given correlation matrix.

Beyond Pearson's Correlation coefficient

- Perfectly positively dependent random variables do not necessarily have a correlation of 1; perfectly negatively dependent random variables do not necessarily have a correlation of -1.
- Possible values of correlation depend on the marginal distribution of the random variables. All values between -1 and 1 are not necessarily attainable.

Beyond Pearson's Correlation coefficient

- Correlation is not invariant under transformations of the random variables. For example, $\log(X)$ and $\log(Y)$ generally do not have the same correlation as X and Y.
- Correlation is only defined when the variances of the random variables are finite. It is not an appropriate dependence measure for very heavy-tailed random variables where variances may be infinite.

Spearman's Rank Correlation coefficient

The Spearman's Rank Correlation coefficient between two continuous random variables X, Y is defined as

$$\rho(X,Y) = r(F(X), G(Y))$$

where r denotes the Pearson's Correlation coefficient and F, G are distribution functions of X, Y respectively.

Spearman's Rank Correlation coefficient.....

Usually, Spearman's Rank Correlation coefficient is defined only for a finite sample. Given random variable (X, Y), if (X_1, Y_1), (X_2, Y_2),..., (X_n, Y_n) is a random sample from (X, Y), and if

$$\rho_n = \text{Rank Correlation}\big((X_1, X_2, \ldots, X_n), (Y_1, Y_2, \ldots, Y_n)\big)$$

then it can be shown that

$$\rho_n \to \rho(X, Y) \;\; a.s.$$

Spearman's Rank Correlation coefficient.....

It is easy to see that if u, v are strictly increasing functions then for any random variables X, Y

$$\rho(u(X), v(Y)) = \rho(X, Y).$$

24 Multivariate Distribution with double exponential marginals

Question: How do we construct a joint distribution in m-dimension where all marginals are double exponential and all rank correlations are say 0.75. Of course, these conditions do not determine the joint distribution uniquely. But can we construct one such?

25 Multivariate Distribution with double exponential marginals

One answer: Let $X = (X_1, X_2, \ldots, X_m)$ have multivariate normal distribution with mean vector 0, all variances equal to one and all (Pearson) correlations equal to 0.75. Let Φ denote the distribution function of standard normal distribution. Let $U_i = \Phi(X_i)$, $1 \leq i \leq m$. Now all marginal distributions of $U = (U_1, U_2, \ldots, U_m)$ are Uniform. Let G denote the distribution function of double exponential distribution and let

$$Y_i = G^{-1}(U_i),\ 1 \leq i \leq m.$$

Then each Y_i has the desired distribution G.

26 Multivariate Distribution with double exponential marginals

Clearly the components are not independent. Explicit computation of correlation between Y_i and Y_j seems difficult. It can be shown that Spearman correlation ρ_{ij} between Y_i and Y_j is 0.73414. If we had started with $corr(X_i, X_j) = 0.765362$ for $i \neq j$, then we would have got $\rho_{ij} = 0.75$.

Multivariate Distribution with given marginals

The construction is flexible enough- all component marginal distributions could have different parameters or could come from different families (continuous distributions). Also, the correlations need not be same for all pairs as we will see shortly

Copula

A Copula (in m dimensions) is a joint distribution in m dimensions such that each of its marginals is Uniform $(0,1)$.

Example: $U = (U_1, U_2, \ldots, U_m)$ where $U_i = \Phi(X_i)$, $1 \leq i \leq m$ and $X = (X_1, X_2, \ldots, X_m)$ have multivariate normal distribution with mean vector 0, all variances equal to one and correlations r_{ij}. This copula (distribution of U) is called the Gaussian copula. The Spearman rank correlations ρ_{ij} between U_i, U_j for $i \neq j$ are given by

$$\rho_{ij} = \frac{6}{\pi} \arcsin\left(\frac{r_{ij}}{2}\right). \tag{1}$$

Multivariate Distribution with given marginals..

Now if we want to construct a random vector

$Z = (Z_1, Z_2, \ldots, Z_m)$

in m dimension such that

Z_i has marginal distribution G_i (assumed to be continuous)

and

spearman correlation between Z_i and Z_j is ρ_{ij} for $i \neq j$,

we can proceed as follows.

Multivariate Distribution with given marginals..

Let $X = (X_1, X_2, \ldots, X_m)$ have multivariate normal distribution with mean vector 0 and variance -covariance matrix $\Sigma = (\sigma_{ij})$ with $\sigma_{ii} = 1$ and

$$\sigma_{ij} = 2 \sin\left(\frac{\pi \rho_{ij}}{6}\right).$$

Then let $U_i = \Phi(X_i)$, $1 \le i \le m$ and $Z_i = G^{-1}(U_i)$.
Now $Z = (Z_1, Z_2, \ldots, Z_m)$ is such that Z_i has marginal distribution G_i and spearman correlation between Z_i and Z_j is ρ_{ij} for $i \ne j$. Here we need to use the fact that the spearman correlation does not change under monotone transformations of components.

Multivariate Distribution with given marginals..

The joint distribution of Z is completely determined by the marginal distributions G_i and the Gaussian copula with spearman correlation matrix (ρ_{ij}) (with $\rho_{ii} = 1$.)

Multivariate Distribution with given marginals..

This use of Gaussian copula was made popular by David X. Li in the finance literature. This was seen as a way of dealing with correlated assets whose distribution was not Gaussian and was widely used.

A search for gaussian copula wall street in google will yield several links to articles titled Recipe for Disaster: The Formula That Killed Wall Street. The reference to the formula here is the Gaussian Copula.

Financial Times also has a page titled The formula that felled Wall St again referring to the Gaussian copula.

Copula examples

What is wrong with the Gaussian copula. Well before we explore this, let us have examples of other copulas.

An important family of copulas is the class of Archimedean copula. A continuous, strictly decreasing, convex function $\phi : [0,1] \to [0,\infty]$ satisfying $\phi(1) = 0$ and $\phi(0) = \infty$ is known as an Archimedean copula generator. Given such a function ϕ, let

$$C(u_1, u_2, \ldots, u_m) = \phi^{-1}(\phi(u_1) + \phi(u_2) + \ldots + \phi(u_m)).$$

Copula examples...

If $\psi = \phi^{-1}$ is completely monotone i.e. if for all k

$$(-1)^k \frac{d^k}{(dt)^k} \psi(t) \geq 0,$$

Then

$$C(u_1, u_2, \ldots, u_m) = \phi^{-1}(\phi(u_1) + \phi(u_2) + \ldots + \phi(u_m))$$

defines a copula.

Copula examples...

Example:

Gumbel copula: $\phi(t) = (-\log(t))^{\theta}$ for $\theta \geq 1$

Clayton copula: $\phi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$ for $\theta > 0$

Frank copula: $\phi(t) = -\log(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1})$ for $\theta \neq 0$

In each of these copulas the rank correlation ρ_{ij} between the i^{th} and j^{th} component is a constant for $i \neq j$. This puts a limitation on the use of Archimedean copula beyond dimension two.

t-copula

Let $X = (X_1, X_2, \ldots, X_m)$ have multivariate normal distribution with mean vector 0 and variance -covariance matrix $\Sigma = (\sigma_{ij})$. Let S have a χ^2 distribution with n-degrees of freedom. Let

$$Y_i = \frac{X_i}{(\frac{\sqrt{S}}{n})}.$$

Note that the same χ^2 random variable in the denominator for all i.

So if H_n is distribution function of t distribution with n degrees of freedom, the marginal distribution of each Y_i is H_m and hence $U_i = H_m(Y_i)$ has uniform distribution.

t-copula....

The joint distribution of $(U_1, U_2, \ldots, U_m)$ so constructed is called the t- copula with n degrees of freedom. Let $\Gamma = (\rho_{ij})$ be the rank correlation matrix of $(U_1, U_2, \ldots, U_m)$. An explicit relation between Σ (the pearson correlation matrix of $(X_1, X_2, \ldots, X_m)$) and Γ is not available to the best of my knowledge. The same can be obtained via a massive simulation effort.

Then one can obtain t copula with n degrees of freedom and a given rank correlation matrix Γ.

Copula

To recap, a joint distribution function can be viewed as combination of marginal distributions and the underlying copula, with copula measuring the dependence structure. Also, if the marginals are given and are such that one can simulate from the distribution (easily), then for Gaussian or t - copula with a given correlation structure, it is easy to simulate from the required multivariate distribution.

Tail dependence.

Here is an interesting fact. Let (X, Y) have bivariate normal distribution with mean $(0,0)$, variances 1 and correlation ρ with $|\rho| < 1$. Let ξ_p denote the p^{th} percentile of Normal distribution. Then

$$\lim_{p \downarrow 0} P(Y < \xi_p \mid X < \xi_p) = 0$$

and

$$\lim_{p \uparrow 1} P(Y > \xi_p \mid X > \xi_p) = 0.$$

This holds even when $\rho = 0.99$!

Tail dependence

For a univariate random variable X, let us denote by $\gamma(X,p)$ the p^{th} percentile of the distribution of X. For random variable X,Y let

$$\lambda_L(X,Y) = \lim_{p\downarrow 0} P(Y < \gamma(Y,p) \mid X < \gamma(X,p))$$

$$\lambda_U(X,Y) = \lim_{p\uparrow 1} P(Y > \gamma(Y,p) \mid X > \gamma(X,p)).$$

$\lambda_L(X,Y)$ measures the lower tail dependence and $\lambda_U(X,Y)$ measures upper tail dependence.

Tail dependence

If f and g are increasing functions, it is easy to see that

$$\lambda_L(X,Y) = \lambda_L(f(X), g(Y))$$

$$\lambda_U(X,Y) = \lambda_U(f(X), g(Y))$$

and thus λ_L and λ_U are functions of the underlying copula and not the marginal distribution functions. Thus if X, Y have marginal distributions F, G and the underlying copula is the Gaussian copula, then we have

$$\lambda_L(X,Y) = \lambda_U(X,Y) = 0.$$

Tail dependence

Consider a portfolio of stocks m which are same sector companies and R_n^j denote the returns on log scale on j^{th} stock. Then it is empirically observed that when one of them goes down by a big margin - so do the others. In other words, empirical data suggests that there is lower tail dependence, while using a Gaussian copula to model

$$(R_n^1, \ldots R_n^m),\ n = 1, 2, \ldots$$

would imply no lower tail dependence! So Gaussian copula is not appropriate to model the log returns.

Tail dependence

A better choice seems to be t copula with small degrees of freedom. For a t copula with ν degrees of freedom with Pearson correlation r for the underlying normals, one has

$$\lambda_L = \lambda_U = 2H_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-r)}{1+r}}\right)$$

where H_ν is the distribution function of the t distribution. Even for $r = 0$ there is some tail dependence.

Tail dependence

While t with 1 or 2 degrees of freedom would be better than Gaussian copula, for a t copula $\lambda_L = \lambda_U$ whereas one would like $\lambda_L > \lambda_U$.

Copula and VaR

We had seen that using Gaussian distribution when the assets have a fatter tail results in under estimation of the VaR. Likewise, using Gaussian copula in the context of assets which have tail dependence too underestimates the tail probabilities for the portfolio. The next two slides give an idea of tail probabilities for Gaussian and t-copulas.

Table 5.2. Joint quantile exceedance probabilities for bivariate Gauss and t copulas with correlation parameter values of 0.5 and 0.7. For Gauss copulas the probability of joint quantile exceedance is given; for the t copulas the factors by which the Gaussian probability must be multiplied are given.

			Quantile			
ρ	Copula	ν	0.05	0.01	0.005	0.001
0.5	Gauss		1.21×10^{-2}	1.29×10^{-3}	4.96×10^{-4}	5.42×10^{-5}
0.5	t	8	1.20	1.65	1.94	3.01
0.5	t	4	1.39	2.22	2.79	4.86
0.5	t	3	1.50	2.55	3.26	5.83
0.7	Gauss		1.95×10^{-2}	2.67×10^{-3}	1.14×10^{-3}	1.60×10^{-4}
0.7	t	8	1.11	1.33	1.46	1.86
0.7	t	4	1.21	1.60	1.82	2.52
0.7	t	3	1.27	1.74	2.01	2.83

Table 5.3. Joint 1% quantile exceedance probabilities for multivariate Gaussian and t equicorrelation copulas with correlation parameter values of 0.5 and 0.7. For Gauss copulas the probability of joint quantile exceedance is given; for the t copulas the factors by which the Gaussian probability must be multiplied are given.

			Dimension d			
ρ	Copula	ν	2	3	4	5
0.5	Gauss		1.29×10^{-3}	3.66×10^{-4}	1.49×10^{-4}	7.48×10^{-5}
0.5	t	8	1.65	2.36	3.09	3.82
0.5	t	4	2.22	3.82	5.66	7.68
0.5	t	3	2.55	4.72	7.35	10.34
0.7	Gauss		2.67×10^{-3}	1.28×10^{-3}	7.77×10^{-4}	5.35×10^{-4}
0.7	t	8	1.33	1.58	1.78	1.95
0.7	t	4	1.60	2.10	2.53	2.91
0.7	t	3	1.74	2.39	2.97	3.45