MINERVA
Investment Management Society

Stefano Serio
Head of Risk Management
+393385756999

Chiara Iasenzarino
Risk Management Analyst
+393348027877

Steps-Overview:

I. Normality:
- *Variance-Covariance Approach*
- *Montecarlo Simulations Approach*
- *Evidences against normality*
- *Historical Approach*

II. Multivariate t-student Hypothesis
- *Montecarlo Simulations Approach*
- *Montecarlo with Multivariate Skew t-student*

III. Heteroskedasticity Hypothesis
- *Montecarlo Simulations under normality using EWMA for Covariance matrix estimation*
- *Montecarlo Simulations under multivariate t-student using EWMA for Covariance matrix estimation*

Objective:

The main task of this report is to give a sense of the risk embedded in the first portfolio composed by the Investment Division of Minerva Investment Society. The point of view chosen is that of a daily perspective on the potential extreme behavior of a basket of assets, with the future goal of sending signals to the portfolio manager to marginally review its stocks' selection or the weights assigned to each of them.

Main Results:

	MULTIVARIATE NORMAL			EMPIRICAL	MULTIVARIATE T-STUDENT		MULTIVARIATE SKEW T-STUDENT
	VAR-COV	*MONTECARLO*	*MONTECARLO-EWMA*	*HISTORICAL*	*MONTECARLO*	*MONTECARLO - EWMA*	*MONTECARLO*
VAR 95%	€ (1,561.67)	€ (1,349.91)	€ (1,480.84)	€ (1,469.91)	€ (1,551.89)	€ (1,715.27)	€ (1,785.89)
VAR 97.5%	€ (1,860.84)	€ (1,650.96)	€ (1,807.88)	€ (1,844.02)	€ (1,980.96)	€ (2,181.34)	€ (2,208.25)
VAR 99%	€ (2,208.70)	€ (1,990.43)	€ (2,183.96)	€ (2,444.02)	€ (2,558.58)	€ (2,792.78)	€ (2,769.37)
ES 95%		€ (1,745.53)	€ (1,913.05)	€ (2,153.89)	€ (2,184.04)	€ (2,398.27)	€ (2,408.20)
ES 97.5%		€ (2,005.34)	€ (2,198.15)	€ (2,613.86)	€ (2,625.59)	€ (2,875.34)	€ (2,841.85)
ES 99%		€ (2,315.25)	€ (2,536.83)	€ (3,356.77)	€ (3,228.70)	€ (3,530.74)	€ (3,435.68)

HP:

- **100,000.00 € Investment**
- **6 months of log returns**
- **Daily Var**
- **500,000 iterations**
- **λ = 0.94**

I. Normality

We started with the simplest and most optimistic hypothesis possible, that of normality of returns, and we implemented the classical Var-Cov asset normal approach. With this objective, using six months of daily data we estimated the vector of historical standard deviations of the stocks, along with their vector of means and the 38x38 Varcov matrix, assuming the stability of these parameters over time. The VAR for the single stock was computed as: $VARi, daily = Amount(i)\ x\ \sigma(i)\ x\,\alpha$, with Amount(i) equal to $\mathbb{w}(i)\ x\ 100{,}000.00\ €$ and α= Scaling factor, equal to 1.645, 1.96, 2.33 for a 95%, 97.5%, 99% Confidence level respectively, ∀ i=1,...,38. The VAR of the single chosen stocks happen to be very similar one to another, and this is due to the weight assigned to each stock, all similar, but as a decreasing function of the standard deviation registered in the past. Thanks to the normality assumption it was possible to apply the formula: $VARp(daily) = \sqrt{VAR'\ x\ VARCOV\ x\ VAR}$, with VAR equal to the vector of VAR of the stocks and VARCOV equal to the historical variance covariance matrix.

As a second step, we wanted to enrich the normality analysis using Montecarlo Simulations and so fitting the data with a Multivariate Normal, the aim is to increase the number of observations drawn from the desired distribution function. Still thanks to the normality hypothesis, we could still use excel, and in particular the function RAND(), to generate a set of 500,000 random numbers extracted from a uniform distribution constrained between 0 and 1, for each of the stocks. Such a number of iterations was chosen because of convergence issues, and it must be noted that it could be necessary to use R or other computer software in order to fulfil the following step by step because of RAM issues. These random numbers have then been mapped into realizations of a standard normal random variable (through the inversion of its cumulative distribution function). Up to this point we are not considering the correlation among the variables yet. For this reason, we had to use the CHOLESKY variance covariance matrix decomposition to find the triangular matrix **A**, such that **AA'**=VARCOV. As a final step, we applied: $SIMULATED\ DATA\ VECTOR(j) = Xj\ x\ \boldsymbol{A} + \mu$, where the Xjs, with j=1,..,500000, are the row vectors obtained in the previous step and μ is the vector of means. As a final check we compared the Varcov matrix of simulated data with that of original data, looking for divergences. At this point, we simply computed the value of the portfolio in each simulated scenario of returns, using the fixed weights we had, and sorted the P&L from the highest loss to the smallest (highest profit). With the percentile approach we cut the distribution of portfolio loss at the desired level according to the chosen confidence level, ie. 95%, 97.5%, 99%. In this case, we were also able to compute the expected shortfall, equal to the average of losses that exceeded the E(Loss).

Evidences against normality, some tests.

Switching to R, we tested the normality hypothesis, to check if the Multivariate Normal could be a good fit for our data. The results for these tests are shown below.

MultivariateNormality

	Test	Statistic	p value	Result
1	Mardia Skewness	14665.1857449292	2.18645401922033e-194	NO
2	Mardia Kurtosis	19.6005500911968	0	NO

	Test	HZ	p value	MVN
1	Henze-Zirkler	1.000035	0	NO

	Test	H	p value	MVN
1	Royston	689.6822	6.900611e-121	NO

	Test	E	df	p value	MVN
1	Doornik-Hansen	1416.975	76	4.514769e-246	NO

	Test	E	df	p value	MVN
1	Doornik-Hansen	1416.975	76	4.514769e-246	NO

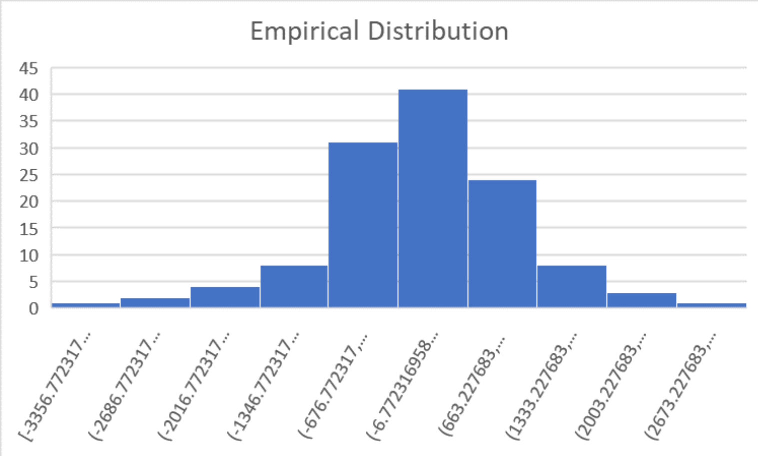
	Test	Statistic	p value	MVN
1	E-statistic	3.693513	0	NO

Analyzing the 38 stocks separately we noticed that the behavior of some of them could be well approximated by a normal distribution, on the other hands, most of them clearly deviate from normality, showing levels of skewness different from zero, and above all, abnormal levels of excess kurtosis, suggesting the presence of fat tails. As far as the multivariate tests are concern, there are no doubts about the exclusion of the gaussian case. Furthermore, as later we will see, trying to fit the data with a Multivariate t-student distribution we noticed that the degree of freedom level is about 7.7, far away from 20. According to the usually implied rule of thumb, in fact, with df<20 we can safely reject normality.
According to these findings, the normal distribution is not able to capture that fatness of the left tail, that is what we are interested in. With this limitation in mind, we tried to implement the historical approach, and so a non-parametric approach, assuming that the historical empirical distribution of the data could be a good stable proxy for the future. In this case, especially for the 99% case, we observe an increase in the forecasted loss if compared to the other models, reflecting the findings of the tests. However, here arise a remarkable problem, the fact that if we increase the number of observation to make the prediction more precise, extending the length of data, we implicitly make less realistic the hypothesis of stability of the distribution. This trade off moves the attention to further model specifications. As far as Expected Shortfall is concerned, in this case it is basically an unreliable estimate, due to the scarcity of data in fact, especially for the 99% case, it is limited to an average of a very limited number of data point.

$univariateNormality

	Test	Variable	Statistic	p value	Normality
1	Shapiro-Wilk	Column1	0.9226	<0.001	NO
2	Shapiro-Wilk	Column2	0.9837	0.1445	YES
3	Shapiro-Wilk	Column3	0.9815	0.0891	YES
4	Shapiro-Wilk	Column4	0.7540	<0.001	NO
5	Shapiro-Wilk	Column5	0.9819	0.0985	YES
6	Shapiro-Wilk	Column6	0.8182	<0.001	NO
7	Shapiro-Wilk	Column7	0.9835	0.1395	YES
8	Shapiro-Wilk	Column8	0.9558	5e-04	NO
9	Shapiro-Wilk	Column9	0.8461	<0.001	NO
10	Shapiro-Wilk	Column10	0.9970	0.9971	YES
11	Shapiro-Wilk	Column11	0.9924	0.7477	YES
12	Shapiro-Wilk	Column12	0.9868	0.2799	YES
13	Shapiro-Wilk	Column13	0.9822	0.1054	YES
14	Shapiro-Wilk	Column14	0.8885	<0.001	NO
15	Shapiro-Wilk	Column15	0.9737	0.0167	NO
16	Shapiro-Wilk	Column16	0.9650	0.0028	NO
17	Shapiro-Wilk	Column17	0.7011	<0.001	NO
18	Shapiro-Wilk	Column18	0.7992	<0.001	NO
19	Shapiro-Wilk	Column19	0.9477	1e-04	NO
20	Shapiro-Wilk	Column20	0.9804	0.0704	YES
21	Shapiro-Wilk	Column21	0.9492	2e-04	NO
22	Shapiro-Wilk	Column22	0.9522	3e-04	NO
23	Shapiro-Wilk	Column23	0.8147	<0.001	NO
24	Shapiro-Wilk	Column24	0.9821	0.1029	YES
25	Shapiro-Wilk	Column25	0.9075	<0.001	NO
26	Shapiro-Wilk	Column26	0.9887	0.4042	YES
27	Shapiro-Wilk	Column27	0.9876	0.3296	YES
28	Shapiro-Wilk	Column28	0.9203	<0.001	NO
29	Shapiro-Wilk	Column29	0.9556	5e-04	NO
30	Shapiro-Wilk	Column30	0.8425	<0.001	NO
31	Shapiro-Wilk	Column31	0.9557	5e-04	NO
32	Shapiro-Wilk	Column32	0.9070	<0.001	NO
33	Shapiro-Wilk	Column33	0.9648	0.0027	NO
34	Shapiro-Wilk	Column34	0.9448	1e-04	NO
35	Shapiro-Wilk	Column35	0.9703	0.0081	NO
36	Shapiro-Wilk	Column36	0.8439	<0.001	NO
37	Shapiro-Wilk	Column37	0.9879	0.3512	YES
38	Shapiro-Wilk	Column38	0.9908	0.5938	YES

	Skew	Kurtosis
1	0.915580981	4.741246179
2	0.050063450	1.047289368
3	0.380638529	0.895407277
4	3.065564214	17.090020669
5	0.195211842	1.022666264
6	-2.563761881	16.922458987
7	-0.337149658	-0.257238005
8	0.826413110	1.909527432
9	2.279544495	13.765240235
10	0.057319747	-0.007204183
11	-0.158561063	-0.249020922
12	0.244523560	0.828764396
13	0.189242869	0.976140196
14	1.646859627	7.041687368
15	-0.482210996	1.982213166
16	0.346185605	1.609216646
17	-2.988010453	26.270843325
18	-1.937224103	16.956031189
19	0.043079622	3.716822494
20	0.201349968	1.278307364
21	0.282151021	2.409461361
22	0.728416041	3.014723141
23	2.372793261	17.203474804
24	0.247370756	1.368398724
25	1.184006513	7.763755806
26	-0.323897684	0.378780300
27	-0.251008814	0.179342704
28	1.280497572	5.872982264
29	-0.504738666	2.419036778
30	2.356334487	14.675523149
31	-0.782189170	2.143277388
32	1.323477444	3.141154376
33	-0.669864954	1.198766297
34	-0.141877780	3.174493751
35	0.058375037	1.150455785
36	-1.579536237	12.086871970
37	0.007157441	0.835122079
38	0.088088829	0.078135404



II. Multivariate t-student hypothesis / Multivariate Skew t-student

Switching to R, we then tried to fit the data with a Multivariate t-student distribution. The fit function of the "LaplaceDemon" package found the parameter for us, and among those, it estimated a df=7.769. Then, inputting these parameters along with the Cholesky triangular matrix into the simulation function, we were able to perform the Montecarlo Simulations. Here again we checked for the equality of the variance covariance matrix of simulated data with that of original data. As in the case of normality, we proceeded computing the P&L of the portfolio for each possible simulated scenario and we used the percentile approach. As can been seen in the table of page 1, especially for the extreme tails, we now got a truly less conservative estimate of the potential losses.
Furthermore, looking at the tests and at the distribution of the historical data, the portfolio returns present some degree of asymmetry in their distribution. For this reason, we tried to fit the data also with a Multivariate Skew t-student. It turned out that we were right, according to the fitting results in fact, we obtained a vector parameter Alpha slightly different from 0, that is instead typical for a symmetric t-student. Df level is now equal to 8.22, still lower than 20. It is interesting to notice that in this case, the results are very similar to those coming out from using EWMA with the symmetric multivariate t-student.

III. Heteroskedastic Hypothesis

Not satisfied by the previous findings, we decided to refine our estimates of the varcov matrix. In particular, we assumed heteroskedasticity, and so we allowed variance and covariances to move over time. In this environment, our previous estimate of the varcov matrix was not reliable anymore and we replaced it using the Exponential Weighted Moving Average Approach (EWMA) with a decay factor, λ, equal to 0.94. Such an estimate of this parameter, close to 1, implies that our estimate is going to be less sensitive to recent observations, since the past observations are assigned with a high level of persistence. However, thanks to this decaying factor, the weights assigned to past observations approach zero very slowly, and so decrease substantially the ECHO effect on the data, ie. big past shocks leaving the data sample will have a small effect on the estimate. In this context, we performed again the Montecarlo Simulations under both the Multivariate Normal HP and the Multivariate t-student HP. In both the cases we got a more conservative estimate of the VAR, projecting higher potential losses, accounting for a time varying variance of stocks that appears to be not that favorable.