



THE MODERN ACTUARY - CHALLENGE • INFLUENCE • LEAD  
ASTIN • IAAHS • IAALS • IACA • PBSS

2 - 5 April  
**2019**  
Cape Town  
South Africa  
**CTICC**

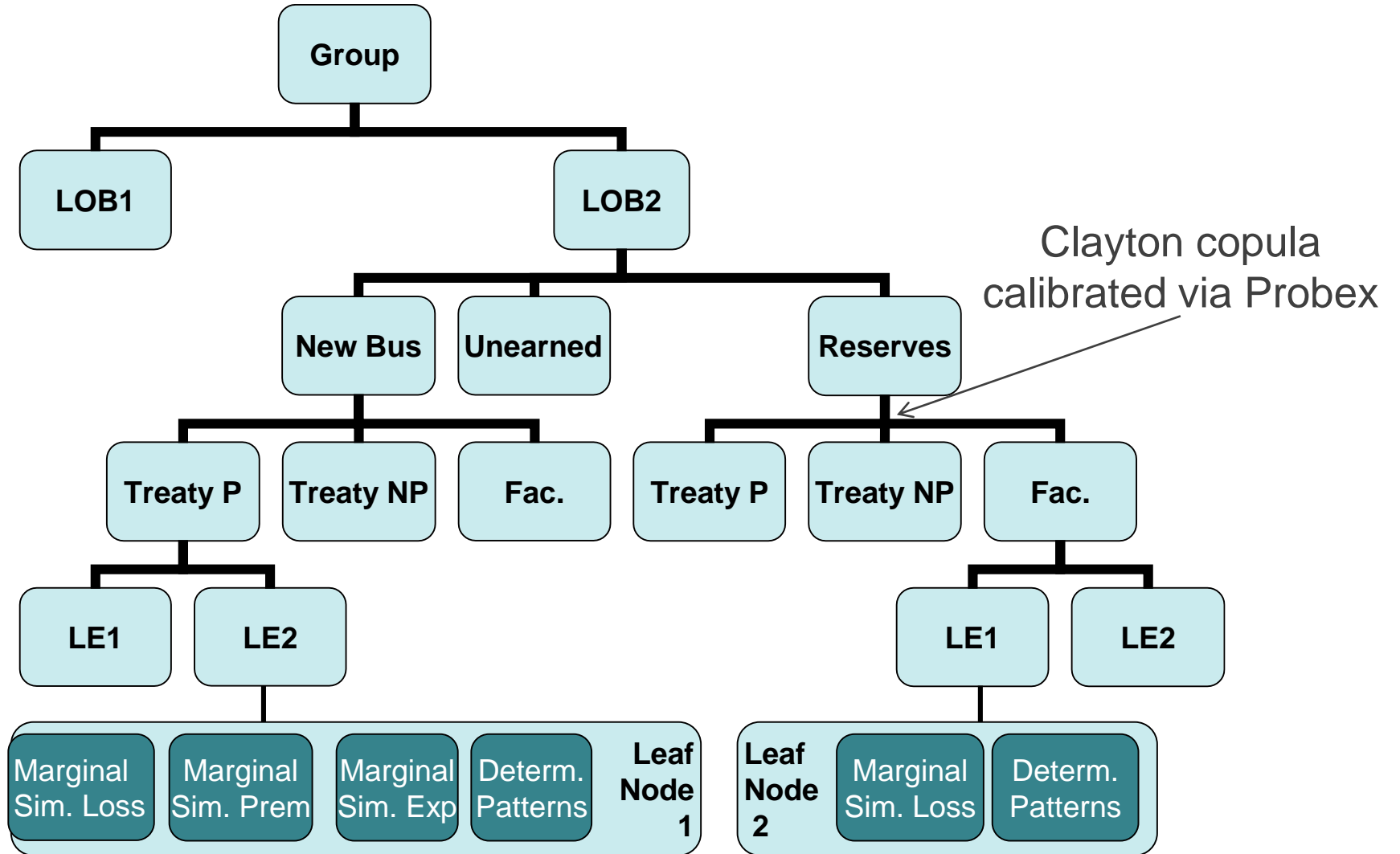
# Impact of an aggregation tree over the skewness of the final risk distribution

## Eric Dal Moro

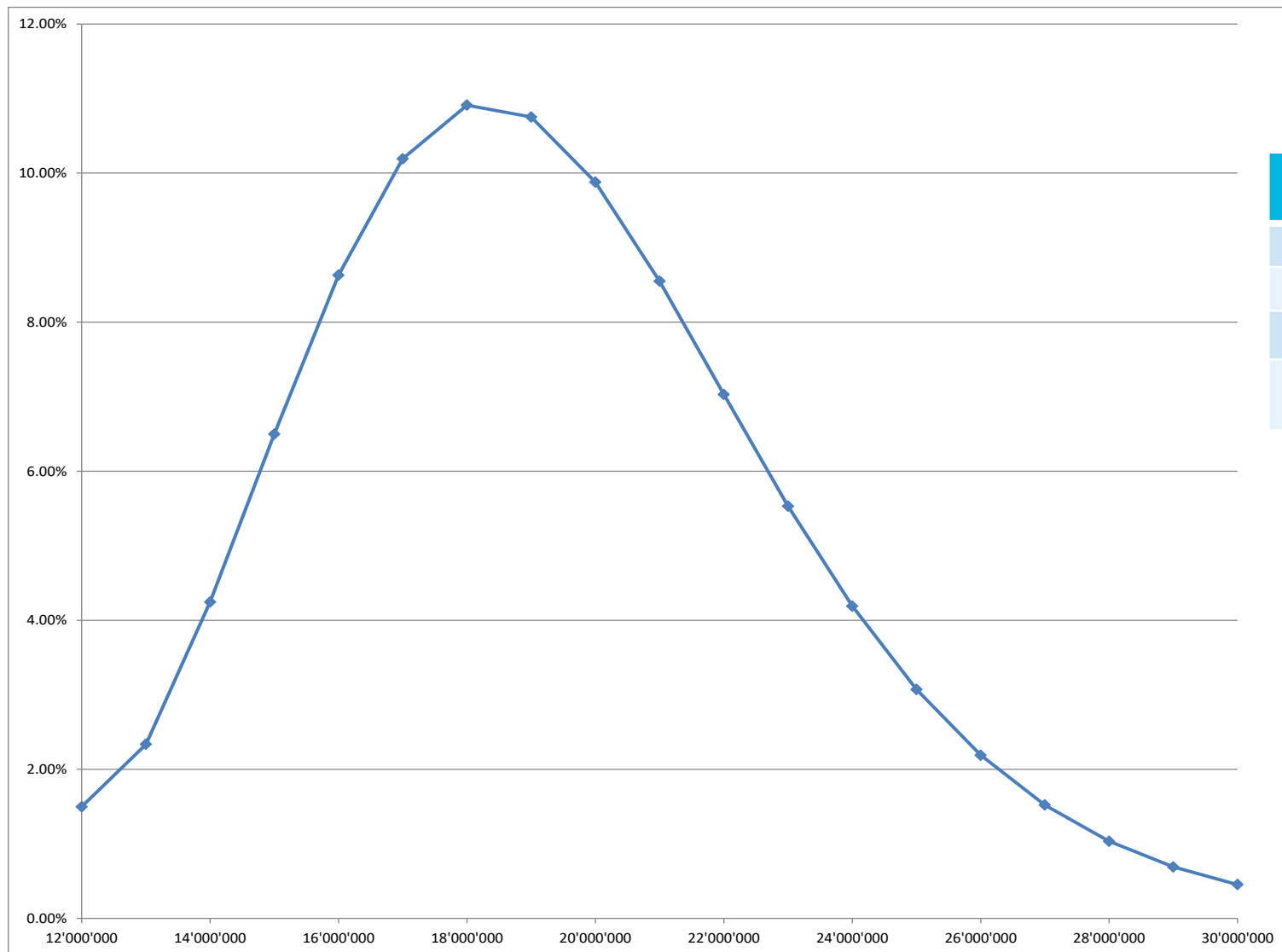
[www.colloquium2019.org.za](http://www.colloquium2019.org.za)

Hosted by  
**ACTUARIAL**  
**SOCIETY**  
OF SOUTH AFRICA





# Internal Model – Results - Example



	Internal Model	LogN	Pareto
Best Estimate	100	100	100
CoV	6.1%	6.1%	6.1%
Skewness	1.5	0.183	2.4
VaR 99.5% (Cornish-fisher)	124.29	116.75	129.47

- Skewness of lognormal distribution  
 $Skewness = (3 + CoV^2)CoV$
- Skewness of a pareto distribution





$$\text{Solve } k^2 - 2k - \frac{1}{CoV^2} = 0$$

$$Skewness = \frac{2(1+k)}{k-3} \sqrt{\frac{k-2}{k}}$$



# Agenda

---

-  **Skewness of two aggregated variables**
-  **Skewness vs diversification**
-  **Comparison of aggregation trees**
-  **Conclusion**

## Skewness of two aggregated variables

Let's assume two random variables  $X_1$  and  $X_2$  for which we know their joint copula distribution. We will focus on the third moment of their aggregated distribution i.e. on

$$SK(X_1 + X_2) = E \left[ \left( (X_1 + X_2) - E(X_1 + X_2) \right)^3 \right]$$

We define the third moment of  $X_1$  as:

$$SK(X_1) = E \left[ \left( X_1 - E(X_1) \right)^3 \right] = E(X_1^3) - 3E(X_1)E(X_1^2) + 2E(X_1)^3$$

It is easy to prove that:

$$SK(X_1 + X_2) = SK(X_1) + SK(X_2) + 3Cov(X_1^2, X_2) + 3Cov(X_1, X_2^2) - 6[E(X_1) + E(X_2)]Cov(X_1, X_2)$$

We will now use the following general property of copulas:

If  $C_{X,Y}$  is the copula of  $(X,Y)$ , any increasing transformation of  $(X,Y)$  has the same copula i.e.

If  $f$  is increasing and  $g$  is increasing,  $C_{f(X),g(Y)}(u,v) = C_{X,Y}(u,v)$

With this property, we know that the dependence between  $X_1$  and  $X_2$  and  $X_1^2$  and  $X_2$  is the same i.e.:

$$Cov(X_1, X_2) = \rho \sigma_{X_1} \sigma_{X_2}$$

$$Cov(X_1^2, X_2) = \rho \sigma_{X_1^2} \sigma_{X_2}$$

## Skewness of two aggregated variables

As a result, we are interested in the finding the parameter  $\alpha_1$  of the following equation:

$$\text{Cov}(X_1^2, X_2) = \alpha_1 \text{Cov}(X_1, X_2)$$

$$\alpha_1 = \frac{\sigma_{X_1^2}}{\sigma_{X_1}}$$

If we denote:

$$KT(X) = E[(X - E(X))^4]$$

$$SK(X) = E[(X - E(X))^3]$$

$$\text{Var}(X) = E[(X - E(X))^2]$$

we have:

$$\alpha_1 = \frac{\sigma_{X_1^2}}{\sigma_{X_1}} = \sqrt{\frac{KT(X_1) + 4SK(X_1)E(X_1) + 4E(X_1)^2 \text{Var}(X_1) - \text{Var}(X_1)^2}{\text{Var}(X_1)}}$$

And:

$$SK(X_1 + X_2) = SK(X_1) + SK(X_2) + 3[\alpha_1 + \alpha_2 - 2E(X_1) - 2E(X_2)]\text{Cov}(X_1, X_2)$$

## Skewness of two aggregated variables

Comments on the equation below:

$$SK(X_1 + X_2) = SK(X_1) + SK(X_2) + 3[\alpha_1 + \alpha_2 - 2E(X_1) - 2E(X_2)]Cov(X_1, X_2)$$

$$\text{As } \alpha_1 - 2E(X_1) = \sqrt{\frac{KT(X_1) + 4SK(X_1)E(X_1) + 4E(X_1)^2 Var(X_1) - Var(X_1)^2}{Var(X_1)}} - 2E(X_1) \geq 0$$

and, most of the time,  $Cov(X_1, X_2) \geq 0$  as risks are usually positively correlated,

**the aggregated skewness is higher than the sum of both skewness.**

Estimation of  $Cov(X_1, X_2)$

Elliptical copulas: Coefficient given directly by the correlation matrix.

Archimedian copula: Let's assume we know the Kendall Tau  $\tau$  of our dependence structure, Embrechts et al. (2015) propose the following transformation:

$$\rho = \sin\left(\frac{\pi}{2}\tau\right)$$

From the above, it is possible to estimate the Spearman Rho which will be used as proxy:

$$\rho_S = \frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right)$$

# Skewness of two aggregated variables – Numerical applications

Gaussian copula:

	Param. 1	Param. 2	Correlation	0.174	0.309	0.415	0.5	0.568	0.623	0.669	0.707	0.997
Normal	20.000	20.000	Simulations	-0.004	-0.006	0.001	-0.006	0.003	-0.003	0.001	0.000	-0.002
	10.000	10.000	Proxy	-0.004	-0.001	-0.001	-0.004	0.002	-0.001	-0.004	-0.001	-0.002
LogNormal	2.991	0.100	Simulations	0.927	0.931	0.917	0.924	0.938	0.931	0.931	0.942	0.932
	2.228	0.385	Proxy	0.929	0.934	0.913	0.928	0.936	0.929	0.933	0.940	0.932
Pareto	8.540	25.860	Simulations	3.155	2.968	3.104	3.068	3.270	3.191	3.158	3.080	3.273
	5.770	32.910	Proxy	3.155	2.981	3.110	3.058	3.279	3.183	3.158	3.079	3.285
Weibull	8.933	30.940	Simulations	-0.273	-0.275	-0.271	-0.272	-0.273	-0.264	-0.269	-0.275	-0.272
	5.381	43.171	Proxy	-0.273	-0.275	-0.272	-0.270	-0.271	-0.268	-0.271	-0.271	-0.270

Skewness comparison between simulated distribution and formula (Proxy)

Mirrored Clayton copula:

	Param. 1	Param. 2	Theta	0.25	0.5	0.75	1	1.25	1.5	1.75	2	40
Normal	20.000	20.000	Simulations	0.642	0.721	0.748	0.762	0.763	0.756	0.743	0.732	0.585
	10.000	10.000	Proxy	0.566	0.619	0.645	0.669	0.687	0.696	0.700	0.703	0.731
LogNormal	2.991	0.100	Simulations	1.108	1.200	1.214	1.213	1.199	1.204	1.183	1.160	0.930
	2.228	0.385	Proxy	1.017	1.067	1.085	1.102	1.104	1.141	1.138	1.130	1.131
Pareto	8.540	25.860	Simulations	3.474	3.619	3.827	3.990	3.711	3.940	3.757	3.669	3.478
	5.770	32.910	Proxy	3.622	3.820	4.175	4.570	4.112	4.701	4.434	4.296	4.376
Weibull	8.933	30.940	Simulations	-0.148	-0.092	-0.065	-0.056	-0.061	-0.065	-0.072	-0.082	-0.355
	5.381	43.171	Proxy	-0.288	-0.299	-0.306	-0.307	-0.310	-0.309	-0.310	-0.312	-0.317

Skewness comparison between simulated distribution and formula (Proxy)

Gumbel copula:

	Param. 1	Param. 2	Theta	1.125	1.25	1.375	1.5	1.625	1.75	1.875	2	21
LogNormal	2.991	0.100	Simulations	1.089	1.121	1.146	1.110	1.114	1.097	1.078	1.072	0.928
	2.228	0.385	Proxy	1.028	1.060	1.112	1.096	1.126	1.135	1.129	1.140	1.145
Pareto	8.540	25.860	Simulations	3.540	3.575	3.760	3.662	3.736	3.831	3.994	3.751	3.567
	5.770	32.910	Proxy	3.623	3.748	4.248	4.147	4.345	4.547	5.069	4.541	4.624

Skewness comparison between simulated distribution and formula (Proxy)

## Skewness vs diversification

Let's denote:  ${}_xVaR_\alpha$  the value at risk above the mean of the reserve at the quantile  $\alpha$  (Cornish-Fisher expansion):

$${}_xVaR_\alpha = \sigma \left( z_\alpha + \frac{1}{6}(z_\alpha^2 - 1)\beta_1 \right)$$

For two random variables, we have:

$${}_xVaR_\alpha = z_\alpha(1 - D) (\sigma_{X_1} + \sigma_{X_2}) + \frac{1}{6}(z_\alpha^2 - 1) \frac{SK(X_1) + SK(X_2) + 3f\rho\sigma_{X_1}\sigma_{X_2}}{[(1 - D) (\sigma_{X_1} + \sigma_{X_2})]^2}$$

where:  $f = [\alpha_1 + \alpha_2 - 2E(X_1) - 2E(X_2)]$

Reduction of capital due to the diversification against the increase of capital requirement due to the increased skewness:

$$D z_\alpha (\sigma_{X_1} + \sigma_{X_2}) \quad \text{against} \quad \frac{1}{6}(z_\alpha^2 - 1) \frac{3f\rho\sigma_{X_1}\sigma_{X_2}}{(1-D)^2(\sigma_{X_1} + \sigma_{X_2})^2}$$

Simplifications:

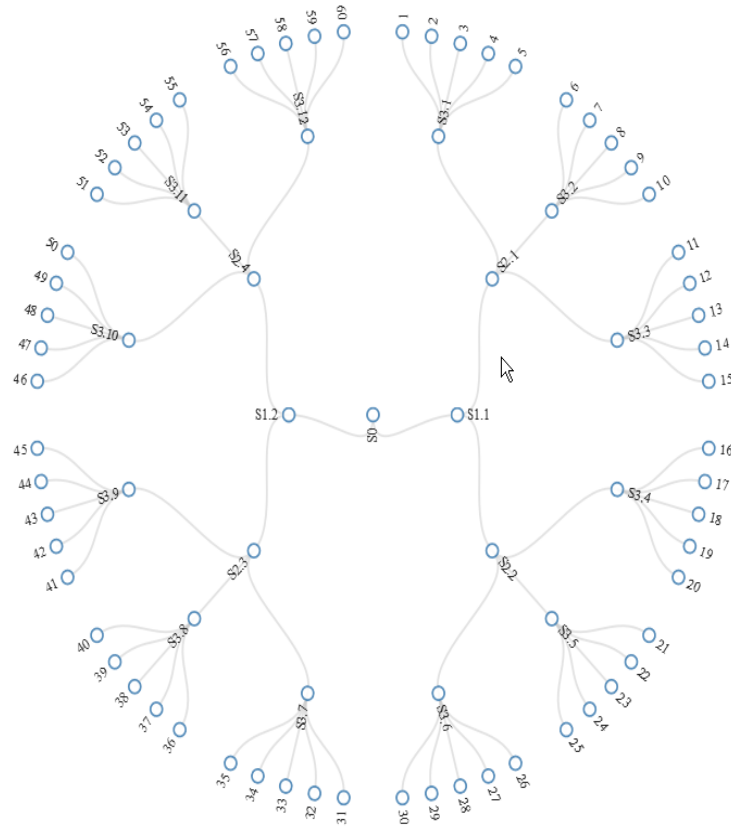
- We estimate the 99.5% quantile:  $z_\alpha = 2.5758$
- $\sigma_{X_1} = \sigma_{X_2}$ .
- $KT(X) = 3\sigma_X^4$
- Skewness of X is nil: This corresponds to a normal distribution.

⇒ As long as D, the diversification benefit, is not smaller than  $\frac{2.1876}{16} CoV_X$  (in the case where the CoV is 30%, it would roughly be 4%), the diversification benefit is bigger than the skewness impact.

# Comparison of aggregation trees

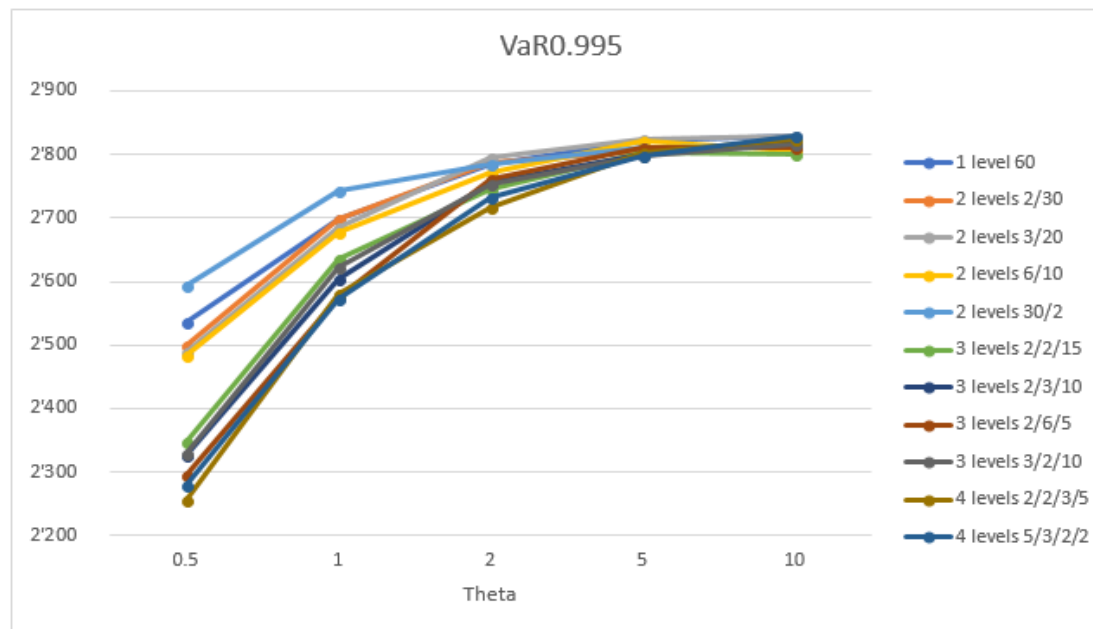
60 lognormal risks are aggregated with different aggregation trees.  
 The different aggregation trees have different layers and, for each layer, the number of aggregated risks vary.  
 The 60 lognormal risks are the same with parameters  $\mu$  equal to 3.34082 and  $\sigma$  equal to 0.19804.

Aggregation 4 levels 2/2/3/5:



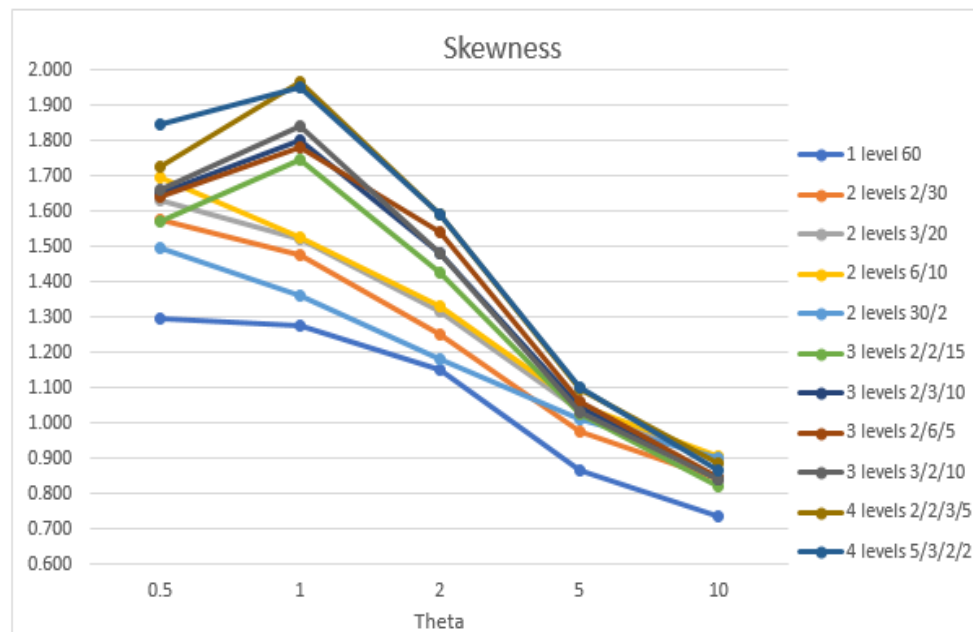
# Comparison of aggregation trees – VaR 99.5%

VaR0.995	Theta				
	0.5	1	2	5	10
<b>1 level 60</b>	2'537	2'700	2'786	2'820	2'828
<b>2 levels 2/30</b>	2'498	2'698	2'787	2'806	2'813
<b>2 levels 3/20</b>	2'489	2'684	2'795	2'825	2'829
<b>2 levels 6/10</b>	2'483	2'677	2'771	2'820	2'806
<b>2 levels 30/2</b>	2'594	2'744	2'784	2'812	2'813
<b>3 levels 2/2/15</b>	2'347	2'635	2'747	2'803	2'802
<b>3 levels 2/3/10</b>	2'327	2'604	2'757	2'801	2'819
<b>3 levels 2/6/5</b>	2'295	2'575	2'760	2'811	2'812
<b>3 levels 3/2/10</b>	2'329	2'622	2'753	2'798	2'816
<b>4 levels 2/2/3/5</b>	2'256	2'581	2'717	2'802	2'823
<b>4 levels 5/3/2/2</b>	2'280	2'572	2'733	2'798	2'829



# Comparison of aggregation trees – Skewness

Skewness	Theta				
	0.5	1	2	5	10
<b>1 level 60</b>	1.297	1.276	1.148	0.864	0.734
<b>2 levels 2/30</b>	1.573	1.475	1.249	0.978	0.839
<b>2 levels 3/20</b>	1.633	1.520	1.317	1.033	0.899
<b>2 levels 6/10</b>	1.697	1.526	1.330	1.052	0.904
<b>2 levels 30/2</b>	1.498	1.358	1.182	1.010	0.898
<b>3 levels 2/2/15</b>	1.571	1.746	1.428	1.026	0.819
<b>3 levels 2/3/10</b>	1.653	1.798	1.480	1.045	0.841
<b>3 levels 2/6/5</b>	1.640	1.781	1.543	1.062	0.844
<b>3 levels 3/2/10</b>	1.660	1.839	1.480	1.032	0.839
<b>4 levels 2/2/3/5</b>	1.724	1.965	1.591	1.096	0.884
<b>4 levels 5/3/2/2</b>	1.844	1.948	1.592	1.102	0.867



## Conclusion

The structure of the aggregation tree e.g. the number of layers play a major role in the determination of the overall coefficient of variation and skewness and, as a consequence, of the Value at Risk 99.5%.

The higher the number of layers is, the higher the skewness is but the lower the coefficient of variation. This is as expected.

The coefficient of variation and the related diversification benefit play a bigger role than the skewness in the determination of the Value at Risk.

	Internal Model	LogN	Pareto
Best Estimate	100	100	100
CoV	6.1%	6.1%	6.1%
Skewness	1.5	0.183	2.4
VaR 99.5% (Cornish-fisher)	124.29	116.75	129.47

Remark: The R code developed to estimate the presented formulae is available on the URL: <https://drive.google.com/open?id=1tX324aD6HqQI1ZA8ghVcOyR2SV-IdH9A>