From the Geometry of Extreme Value Distributions to Improved Tail Fitting in Financial Market Data

William F. Shadwick
Omega Analysis Limited
IMPA June 2018

Joint work with Ana Cascon and William H. Shadwick

The goal of Extreme Value Theory (EVT) is to make statistical estimates of the likelihood and severity of 'random' events which may never, or only rarely, have been observed, based on observed data–for example:

What is the probability of the VIX Index rising by more than 50% in a day or in a week and what is the average increase if that happens?

How likely was a repeat of the worst 20th Century flood in Manitoba Canada (1955) and what level should have be expected if that record flood were exceeded? (As it was in 2011.)

This is a very ambitious goal but two approaches to this sort of problem are very tractable due to remarkable ‘Extreme Value’ limit theorems analogous to the Central Limit Theorem.

These results can be unified, explained and extended in terms of geometric invariants which are precisely analogous to the curvature of a surface.

Just as there are only three types of constant curvature surfaces, the ‘three types’ of distributions at the heart of Extreme Value Theory are the exceptional ones with constant invariants.

Omega Analysis

Our new invariants provide simple and easy to use characterisations of domains of attraction in EVT and explain the relationship between the EVT and Generalised Pareto Distributions.

The invariants also provide an intrinsic measure of the rate of convergence to the limiting distributions.

This isn't just of mathematical interest.

It has led us to highly efficient tail models which produce excellent results in financial market data.

Omega Analysis

First Approach to Extremes: Sample Maxima

Let $X_1, \ldots, X_N$ be a sample of N independent, identically distributed random variables with distribution function F. Let X_{Max} be the sample maximum.

If $X_{Max} < r$ then all of the sample draws must be less than r and the probability of this is $F^N(r)$.

Thus the distribution of X_{Max} is F^N.

In 1928 Fisher and Tippett addressed the question:

Does there exist a sequence of 'location-scale' transformations $x \to a_N x + b_N$ and a distribution G such that

$$F^N(a_N x + b_N) \to G(x) \tag{1}$$

as N tends to ∞?

Any such attractor G *must* be the distribution of its own extremes. This is the so-called Stability Property.

Omega Analysis

Fisher and Tippett proved that there are only three 'types' of distributions with the Stability Property.

$$\Phi(x,\alpha) = e^{-(-x)^{\alpha}},\ x \in (-\infty, 0],\ \alpha > 0 \tag{2}$$

$$\Psi(x,\alpha) = e^{\frac{-1}{x^{\alpha}}},\ x \in [0, \infty),\ \alpha > 0 \tag{3}$$

$$\Lambda(x) = e^{-e^{-x}},\ x \in (-\infty, \infty). \tag{4}$$

(Weibull, Fréchet and Gumbel distributions respectively)

Omega Analysis

The three types of distributions Fisher and Tippett discovered are really a one-parameter family as Richard von Mises showed in 1936.

They are more conveniently denoted on variable domains depending on α as follows:

$$E_\alpha(x) = exp(\frac{-1}{(1+\frac{x}{\alpha})^\alpha}), \quad \alpha \neq 0 \tag{5}$$

This is the Weibull type, defined on $[-\infty, -\alpha]$ when $\alpha < 0$ and is the Fréchet type defined on $[-\alpha, \infty,]$ when $\alpha > 0$.

As $|\alpha| \to \infty$, both types have the Gumbel distribution $E_\infty = e^{-e^{-x}}$ as their limit.

Omega Analysis

Fisher and Tippett showed that the sample maxima limit for the Normal distribution was Gumbel type.

They gave no method for determining if a given distribution had a limit, or if it did, what the limit was.

It took 15 years to fill this gap.

In 1943 Gnedenko gave simple criteria for convergence to the Fréchet and Weibull types in terms of asymptotic scaling behaviour.

He produced a variety of necessary and sufficient conditions for F to be in the domain of attraction of the Gumbel distribution but was not satisfied that any of them were either definitive or practical.

The key to the Fréchet and Weibul distributions is invariance under *scaling transformations* but in the Gumbel case we have shown that it is *translation invariance*.

Second Approach to Extremes: 'Peaks over threshold'

Let X be have distribution F defined on $[\alpha(F), \omega(F)]$.

If the distribution of X, conditional on X exceeding a threshold T, tends to a limit as T tends to $\omega(F)$ (up to location scale transformations) then we say F has a PoT limit.

In 1975 Picklands (and independently Balkema and de Haan) showed that there was a remarable connection between PoT limits and Domains of Attraction of Extreme Value distributions.

F is in the domain of attraction of E_α if and only if the PoT limit of F is, up to a location-scale transformation, a Generalised Pareto distribution G_α as $x \to \omega(F)$ where

$$G_\alpha(x) = 1 - \frac{1}{(1 + \frac{x}{\alpha})^\alpha}, \quad \alpha \neq 0 \tag{6}$$

and

$$G_\infty(x) = 1 - e^{-x} \tag{7}$$

G_α is defined on $[0, -\alpha]$ when $\alpha < 0$ and on $[0, \infty)$ when $\alpha > 0$. As $|\alpha| \to \infty$, both types have the exponential distribution $G_\infty = 1 - e^{-x}$ as their limit.

But now we have a real mystery.

Everything in the Domain of Attraction of E_α is converging to everything else in that Domain of Attraction. For example, it's easy to check that for each ν, the Student t distribution $S(x, \nu)$ is in the domain of attraction of E_ν.

So what is special about Generalised Pareto distributions and what is behind the connection between Extreme Value limits and PoT limits?

Geometry answers these questions.

The Geometry of Extreme Value Distributions is the information invariant under what statisticians call the 'location scale' transformations and mathematicians call the proper affine group on the line $\mathcal{A}$.

These are the transformations of the form

$$x \rightarrow ax + b. \tag{8}$$

where $a > 0$.

The geometry is all of the information that is invariant under the group $\mathcal{A}$.

The most powerful method for discovering this geometry was produced by Elie Cartan extending the 19th Century results of Sophus Lie.

Cartan's *Method of Equivalence* allows us to construct a collection of *differential invariants* (like the curvature of a surface) which completely characterise the geometry.

The geometry always identifies *exceptional cases* such as constant curvature surfaces.

Omega Analysis

Let $I = log(F)$ and $J = I_{xx}/I_x^2$.

It turns out that all of the geometric information about F under location scale transformations is determined by the relation between I and J.

Since we only have one independent variable, J must be functionally dependent on I.

And all of the geometry is encoded in the functional relation $J = H(I)$. (See Appendix 2 for the solution of the equivalence problem).

Omega Analysis

The cases where J is a constant are exceptional (like surfaces with constant curvature).

For example, the Uniform distribution is completely characterised by $J = -1$.

So every distribution F on an interval $[A, B]$ for which $J = -1$ can be translated and re-scaled to the standard uniform distribution $U(x) = x$ on $[0, 1]$.

We'll come back to the rest of the exceptional cases.

The Stability Property is the condition that a distribution F and F^N be in the same equivalence class with respect to $\mathcal{A}$.

But there's no need to restrict this question to integer powers.

It turns out that if we ask what distributions F are in the same $\mathcal{A}$ equivalence class as F^λ for all positive values of λ the answer is still the Extreme Value distributions.

Suppose that $[F^\lambda] = [F]$ for all $\lambda > 0$.

Because the relation $J = H(I)$ determines equivalence classes we must have $J_{F^\lambda} = H(I_{F^\lambda})$

From the definitions of I_F and J_F we have $I_{F^\lambda} = \lambda I_F$ and $J_{F^\lambda} = \frac{1}{\lambda} J_F$ so $\frac{1}{\lambda} H(I) = H(\lambda I)$ for all $\lambda > 0$.

Differentiating with respect to λ and evaluating at $\lambda = 1$ shows that $[F^\lambda] = [F]$ if and only if there is a constant c such that

$$H(I) = \frac{c}{I} \tag{9}$$

Omega Analysis

Thus $[F^\lambda] = [F]$ if and only if $IJ = c$ for constant c.

Each value of c determines a distinct equivalence class.

It is easy to see that the Extreme Value distributions provide normal forms for these equivalence classes.

The equivalence class of E_α is given by $c = 1 + \frac{1}{\alpha}$ for $\alpha \neq 0$ and E_∞ is given by $c = 1$.

We have a new invariant for the 1-parameter family of equivalence classes $[F^\lambda]$ because

$$I_{F^\lambda} J_{F^\lambda} = I_F J_F. \tag{10}$$

for all $\lambda \ \epsilon \ (0, \infty)$.

So K is independent of λ and the *exceptional distributions* for which $K = IJ$ is constant are precisely the Extreme Value distributions.

Omega Analysis

The Geometry of Domains of Attraction Theorem (Cascon and Shadwick)

Let F be a distribution defined on $[\alpha(F), \omega(F)]$.

F is in the domain of attraction of an Extreme Value distribution if and only if the limit of K_F as x approaches $\omega(F)$ is the constant invariant of the Extreme Value distribution.

It is easy to use our result for any of the standard probability distributions to determine which of the Extreme Value distributions E_α and E_∞ they have as their limits.

Unlike Gnedenko's theorem, there's only one test and it's just as simple for the Gumbel attractor as it is for Weibull or Fréchet cases.

It is also easy to verify that for each α the Generalised Pareto distribution G_α is in the Domain of Attraction of E_α and that G_∞ is in the Domain of Attraction of E_∞.

Omega Analysis

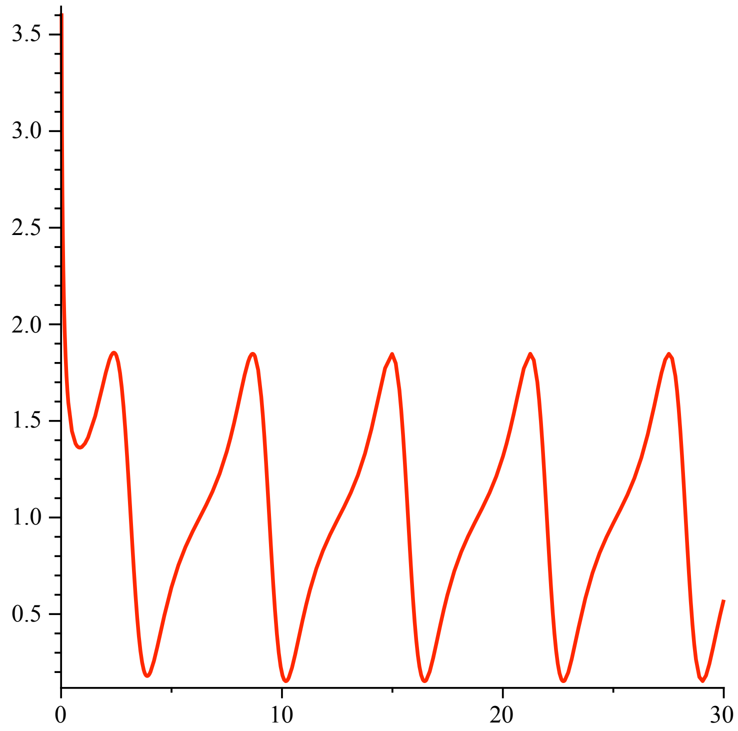
The following example, due to Richard Von Mises, shows that smooth distributions need not have an Extreme Value limit.

If F is defined on $[0, \infty]$ by

$$F(x) = 1 - exp(-x - \frac{sin(x)}{2}). \tag{11}$$

then $I_F J_F$ has no limit as $x \to \infty$.

Our invariant makes this apparent.



$I_F J_F$ has no limit as $x \to \infty$

The invariant quickly becomes periodic and clearly has no limit. By $x = 20$, the difference between F and 1 is less than one part in 10^9.

Omega Analysis

The Picklands Mystery Again

The utility of the Generalised Pareto distributions is that they converge rapidly to their Extreme Value limits. This means that less data is required to make a reasonable fit.

But why the Generalised Pareto distributions and not others in the same Domain of Convergence?

Omega Analysis

Another Approach to Extremes: 'Peaks under threshold'

If F is defined on $[\alpha(F), \omega(F)]$ and $T \in (\alpha(F), \omega(F)]$ the distribution conditional on $x < T$ is $F_T = \frac{F}{F(T)}$. If $[F_T]$ tends to a limiting distribution as $T \to \alpha(F)_-$ then F is said to have a PuT limit.

Such distributions *must* be their own PuT limits so we have the question of PuT stability:

When is $[F_T] = [F]$?

It turns out that once again, geometry answers this question.

The PuT stable distributions are the *exceptional ones* for which J is constant.

Each constant determines an equivalence class of distributions and all constants are possible.

Omega Analysis

It is easy to integrate $J = c$ to produce normal forms, where $\alpha = \frac{1}{c}$:

$$\widehat{G}_{\alpha}(x) = \frac{1}{(1 - \frac{x}{\alpha})^{\alpha}}, \quad \alpha \neq 0 \tag{12}$$

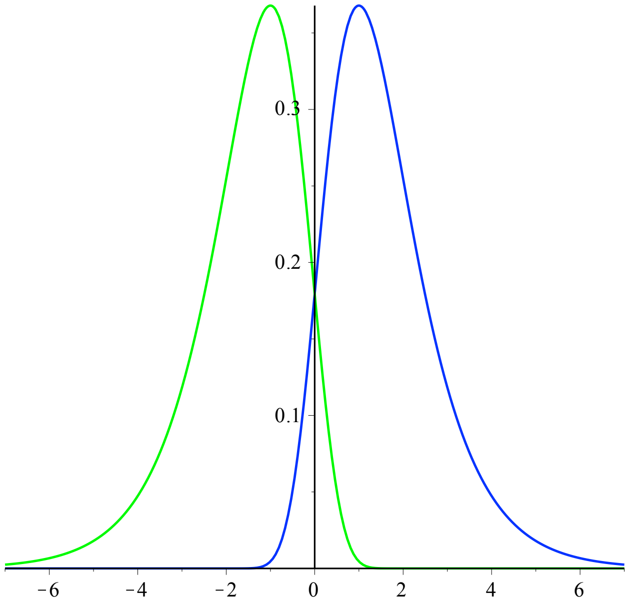
and

$$\widehat{G}_{\infty}(x) = e^{x} \tag{13}$$

$\widehat{G}_{\alpha}$ is defined on $[\alpha, 0]$ when $\alpha < 0$ and on $(-\infty, 0]$ when $\alpha > 0$.

As $|\alpha| \to \infty$, both types have the exponential distribution $\widehat{G}_{\infty} = e^{x}$, on $(-\infty, 0]$ as their limit.

For any probability density function f defined on $[A, B]$ there's a 'mirror image' probability density $\widehat{f}$ on $[-B, -A]$ defined by $\widehat{f}(x) = f(-x)$.



A Gumbel density (blue) and its mirror image

Omega Analysis

If F is the distribution with density f then we will refer to the distribution $\widehat{F}$ whose density is $\widehat{f}$ as the 'mirror image' of F.

It's easy to check (just differentiate) that $\widehat{F}$ is given on $[-B, -A]$ by

$$\widehat{F}(x) = 1 - F(-x). \tag{14}$$

The Generalised Pareto distributions introduced by Picklands are precisely the mirror images of the exceptional distributions corresponding to constant differential invariant J.

Nature only makes so many exceptional, 'constant curvature' objects. The Extreme Value distributions and the (mirror image) Generalised Pareto distributions both have this property, but for different 'curvatures'.

Here's the relationship between them.

Omega Analysis

Duality of Domains of Attraction Theorem (Cascon and Shadwick)

If F is a distribution on $[\alpha(F), \omega(F)]$ and $\widehat{F}$ is the mirror image distribution on $[-\omega(F), -\alpha(F)]$ then

$$\lim_{x \to \omega(F)_-} I_F J_F = 1 + c. \tag{15}$$

if and only if

$$\lim_{x \to -\omega(F)_+} J_{\widehat{F}} = c. \tag{16}$$

The geometry we have uncovered unifies and explains 70 years worth of discoveries in Extreme Value Theory.

But it does much more than that.

It provides an *intrinsic scale* on which we can measure the rate of convergence of a distribution to its EVT or PuT attractor.

The values of the invariants J and K at quantiles are also invariants.

The difference between one of these values and the EVT constant is an *intrinsic measure of convergence*.

The more rapidly a distribution converges to its EVT limit, the less data is necessary to discover that limit.

So being able to compare rates of convergence has a very important statistical application.

Rapid convergence is a key reason for the utility of Generalised Pareto distributions in fitting tails.

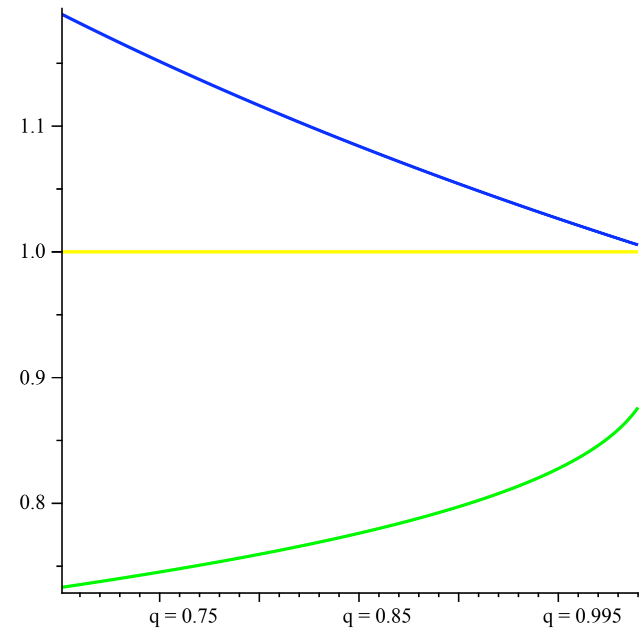
Generalised Pareto distributions converge to their EVT limits incredibly quickly.

But this is a result that's asymptotic and is no guarantee that there's an advantage over all quantiles as we'll also see by using these new invariants.

The EVT limit for the Normal distribution is the Gumbel distribution. Its PoT limit is the exponential distribution (which is the Generalised Pareto distribution G_∞) so both of these distributions are converging to the Gumbel distribution. And the Gumbel distribution has $K = 1$.

The invariant $K = IJ$ provides us with an intrinsic means of comparison the rates at which the other distributinos approach this limit.

We simply compare values of K at the same quantiles. This shows that there's no contest.



K_{Normal} in green, K_{G_∞} in blue and the Gumbel constant

The graph shows that G_∞ converges to E_∞ much more rapidly than the Normal distribution does

Omega Analysis

Fisher and Tippett developed their ‘Penultimate Approximation’ to deal with this slow convergence.

They observed that there was always a distribution in the Weibull family $E_{-\alpha}$ which was a better approximation to the Normal tail that the ‘ultimate’ Gumbel distribution limit $E_{-\infty}$ below any given quantile. (See Appendix 3).

It's easy to check using our invariant that in 'head to head' competitions, i.e. in quantile to quantile comparisons, the convergence of the Generalised Pareto distributions to their EVT limits is faster than that of any of the textbook distributions.

For example, both the $Student(x, 3)$ distribution and $G(x, 3)$ have the same EVT limit E_3.

If you were to ask about convergence near a quantile of interest such as the 99% level where you could want to use either distribution to estimate VaR and Expected Shortfall from financial data, you would see that $G(x,3)$ was much closer to its limit than $Student(x,3)$.

But tail fitting with a Generalised Pareto distribution is *not* going to put its 99% level up against the 99% level of the $Student(x,3)$ distribution.

That's because you don't model the entire distribution by the Generalised Pareto–you only use it to model the *tail*.

You may only be fitting, for example, the top 5% of the data using the Generalised Pareto distribution.

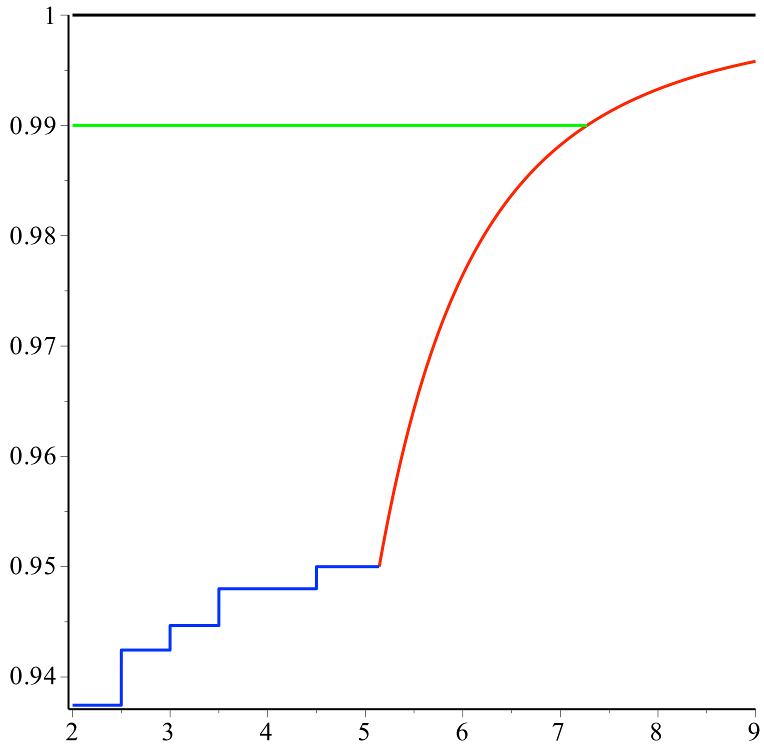
In this case you're putting the convergence at only its $q = 0.8$ level up against the $q = 0.99$ for the $Student(x, 3)$ distribution. And in that contest the $Student(x, 3)$ distribution wins, hands down.

To see this, suppose we want to find the VaR at the 0.99% quantile. The model for the distribution is

$$D = \frac{19}{20} Empirical + \frac{1}{20} G(x-u, 3) \tag{17}$$

where u is the point at which the Empirical distribution reaches $q = 0.95$.

Omega Analysis



$\frac{1}{20}GP(x-u,3)$ attached at $q = .95$.

The $q = 0.99$ level for the tail plus empirical distribution is the solution to $\frac{19}{20} + \frac{1}{20}G(x-u,3) = \frac{99}{100}$.

For $x > u$ the distribution is $D = \frac{19}{20} + \frac{1}{20}G(x - u, 3)$ so the answer to the to the question "When is $D(x) = 0.99$?" is the answer to the question

"When is $\frac{19}{20} + \frac{1}{20}G(x - u, 3) = \frac{99}{100}$?".

The answer to that is the value at which $\frac{1}{20}G(x - u, 3) = \frac{4}{100}$ or $G(x - u, 3) = 0.8$.

If instead, the model for the distribution is

$$D = Student(x, 3), \tag{18}$$

the question is "When is $Student(x, 3) = 0.99$?".

So we're comparing the efficiency of $G(x, 3)$ at $q = 0.8$ with $Student(x, 3)$ at $q = 0.99$.

Note that the value of the invariant K at the $q = 0.8$ level is the same for every one of the family of Generalised Pareto distributions $GP(\frac{x-b}{a}, 3)$. (In the example, I've just set the scale parameter a equal to 1 for convenience.)

So we can make the comparison with the value of the invariant for the distribution $Student(x, 3)$ at $q = 0.99$ using the standard Generalised Pareto distribution $GP(x, 3)$.

At those values the Generalised Pareto distribution's K is twice as far from its EVT limit of $4/3$ as the Student distribution's K is.

This is a handicap that not even the Generalised Pareto distribution can overcome.

(But there's no way to know that without our invariant measure of convergence.)

This is *not* a recommendation to fit tails with $Student$ distributions rather than Generalised Paretos.

It's just an illustration that in spite of their remarkable convergence properties, it's easy to find examples of distributions that are closer to their EVT limits *over the quantile range that matters in practice*.

Such distributions are more efficient in use with short data sets.

This is a tremendous advantage in financial market data.

The ability to make good estimates of VaR and ES using short data windows allows you to observe and respond to changes in risk while there's still time to take advantage of the information.

Omega Analysis has developed proprietary tail fits that converge more rapidly to their attractors than the Generalized Pareto Distributions do, over a range of quantiles of practical significance–even though the latter eventually converge more rapidly.

Our distributions are very efficient models of tails in financial market returns and in other fields where decisions must be based on short data sets.

One example was provided by the dramatic slump in U.S. and European bank share prices following the U.K. referendum.

Our tail models allow us to make very good VaR and ES estimates for 5-day returns at the 99% level. (As judged by comparing the number of VaR breaches over long histories with the number that should have been observed.)

Here's what we predicted the prospects for drawdowns were *prior to the UK referendum* compared with what happened by 6 July 2016.

Omega Analysis

Instrument	Value at Risk (VaR) 99% 5-day	Expected Shortfall (ES) 99% 5-day	Worst 5-day Loss (since 23 June 2016)
KBW Nasdaq Bank Index	-10.9%	-16.8%	-9.3%
Stoxx® Europe 600 Banks	-14.3%	-21.8%	-16.8%
Banca Monte dei Paschi	-31.8%	-47.4%	-32.5%
Barclays	-14.2%	-21.3%	-27.1%
Deutsche Bank	-19.4%	-28.4%	-21.5%
HSBC	-10.8%	-15.7%	N/A
JPMorgan	-10.9%	-17.8%	-7.6%
UniCredit	-21.5%	-31.3%	-27.7%

As of 6 July 2016

With the right technology, the losses were entirely predictable as our measured risk levels for banks had *doubled* in the previous year.

This also showed the information gap between market prices and CDS spreads. Our risk measure showed that the tails of the distribution of JP Morgan returns were significantly fatter than those of HSBC's returns.

But the CDS rate for JP Morgan was only about two thirds that of HSBC at the time.

Low volatility (standard deviation of returns) does not mean low risk.

Imagine taking a sample of 250 points (a year's worth of data) from a Cauchy distribution. The sample mean and variance are finite and by re-scaling you can make the volatility as low as you like.

But the tails are still fat. The tail parameter is (by definition) invariant under re-scaling.

This is completely invisible if you only look at the volatility. A more efficient statistic is essential.

The volatility of daily returns in the S&P 500 Index (and in the VIX Index which is used to measure and speculate on changes in the Volatility of the S&P 500 was extremely low through 2017.

This led many investors to hold Short VIX positions which they regarded as safe.

But although *volatility* was low, the tails of the distributions were been fattening dramatically.

In June 2017 we our tail model showed that a jump in the VIX large enough to be catastrophic for the Short VIX positions should be expected at least once in 14 months.

By August 2017 our estimate had changed to once in less than 3 months.

By January 2018 the VIX index was at a historic low (and many thought, safer than ever to be short). But the tails were just as fat as in August.

In February 2018 record increases in the VIX wiped out the Short positions.

Flash crashes and Jamie Dimon's statistics

In April 2015 Jamie Dimon's shareholder letter was headline news but not for the right reason. He was trying to make important points about liquidity in the U.S. Treasury market and the Swiss National Bank's impact on the Euro Swiss Franc exchange rate.

But his message was swamped by the reaction to his ridiculous observation that the October 2014 'flash crash' in U.S. Treasuries was "...supposed to happen once in every 3 billion years or so..."

Of course you can only generate a claim like that by using a Normal distribution to turn a number of standard deviations into a probability estimate. Not a very smart thing to do.

Our tail model showed that the 40 basis point move was, in fact, a daily high-low that should be expected every two to three years.

The really important point was that *the sort of move which an entire day's trading should produce only a few times per decade occurred in less than 15 minutes*

I sent our analysis of this to Jamie Dimon but he hasn't gotten back to me yet.

The other important point he hinted at without being explicit was pretty obviously aimed at what some market participants would call the Swiss National Bank's market vandalism.

When they pulled the plug in January 2015, there was a 38 standard deviation move in the Euro Swiss Franc exchange rate.

Omega Analysis

Central bankers I have talked to seem to think this all worked out just fine.

But there were a lot of losers. If you weren't big enough for threats of legal action to be effective you probably had your guaranteed stops blown out.

There were fund managers who were just on the right side of that line while other market participants, for example IG Index, took major losses.

To see just how outrageous that 38 sigma move was, we can ask a really ridiculous question.

What was the 1 day in 10 year VaR and ES before the Swiss National Bank's action?

Our model provides an answer that's perfectly reasonable. The VaR was 4.6% and the ES conditional on a VaR breach was 7.9%. In the 15 year history of the Euro there was in fact one prior move of almost 8% (in the opposite direction to January 2015)

This shows why some people think the term market vandalism is appropriate.

Even if you had the best current risk technology and even if you were willing to believe and act on a prediction at the 1 day in 10 year level, *you could not possibly have been prepared for the* 14.4% *move the SNB precipitated.*

Imagine what the UK regulator's reaction would have been if a non-Central Bank market participant in London had pulled this off!

Omega Analysis

Important Notice

LEGAL NOTICE
Please read this notice carefully:

The contents of this document are for illustrative and informational purposes only. No information in this document should be considered a solicitation or offer to buy or sell any financial instrument or to offer any investment advice or opinion as to the suitability of any security in any jurisdiction. All information is subject to change and correction due to market conditions and other factors. This document has been created without any regard to the specific investment needs and objectives of any party in any jurisdiction. Specific instruments are mentioned in this document but this should not be construed in any way as a recommendation to invest in them or in funds or other instruments based on them. They are used for informational purposes only.

Omega Analysis Limited provides statistical analysis services. Omega Analysis Limited does not provide investment advice. Investors need to seek advice regarding suitability of investing in any securities or investment strategies. Any decisions made on the basis of information contained herein are at your sole discretion and should be made with your independent investment advisor.

References

1) *Limiting forms of the frequency distribution of the largest or smallest member of a sample*,R.A. Fisher and L.C.H. Tippett, Proceedings of the Cambridge Philosophical Society **XXIV**, part II, pp 180-190 (1928)

2) *Sur La Distribution Limite Du Terme Maximum D'Une Série Aléatoire*, B. Gnedenko, Annals of Mathematics **44**, No. 3, pp 423-453 (1943)

3) *Statistical Inference Using Extreme Order Statistics*, J. Picklands III, Annals of Statistics, **3**, No.1, pp. 119-131 (1975)

Omega Analysis

References Cont'd

4) *Residual Life Time at Great Age*, A.A. Balkema and L. de Haan, Annals of Probability, **2**, No.5, pp. 792-804 (1974)

5) *What Just Happened?*Bank risk is rocketing up on both sides of the Atlantic. 6 July 2016. OmegaAnalysis.com.

6) *JP Morgan Needs Better Statistics* W.F. Shadwick. 10 April 2015. OmegaAnalysis.com

Appendix 1. Gnedenko's Conditions for Fréchet and Weibull Attractors

Let F be defined on $[\alpha(F), \omega(F)]$

A necessary and sufficient condition for F to be in the domain of attraction of the Fréchet distribution $F(x, \alpha)$ is

$$\lim_{x \to \infty} \frac{1 - F(x)}{1 - F(tx)} = t^{\alpha} \tag{19}$$

for all $t > 0$.

Gnedenko's Conditions for Fréchet and Weibull Attractors Cont'd

For the Weibull distribution $W(x, \alpha)$ $\omega(F) = B$ must be finite and

$$\lim_{x \to 0_-} \frac{1 - F(tx + B)}{1 - F(x + B)} = t^{\alpha} \qquad (20)$$

for all $t > 0$.

Note that both conditions depend only on the asymptotic scaling behaviour of F (and it is not at all clear what this should have to do with F^N).

Appendix 2. The Equivalence Problem

Coframe for the affine group

$$\theta_1 = ydx \tag{21}$$

$$\theta_2 = \frac{dy}{y} \tag{22}$$

Coframe adapted to the (invariant) $I = log(F)$

$$\omega_1 = dI = \frac{F_x}{F}dx \tag{23}$$

$$\omega_2 = \frac{dy}{y} \tag{24}$$

We have

$$\omega_1 = L\theta_1 \tag{25}$$

where

$$L = \frac{F_x}{yF} \tag{26}$$

Any diffeomorphism that preserves both co-frames also preserves L so we have a second functionally independent invariant.

Now

$$dI = \omega_1 \tag{27}$$

and

$$\frac{dL}{L} = J\omega_1 - \omega_2, \tag{28}$$

where

$$J = \frac{F_{xx}F}{F_x^2} - 1 = \frac{I_{xx}}{I_x^2}. \tag{29}$$

But J depends only on x so it must be a function of I.

It follows that the remaining information is in the functional dependence of J on I.

This dependence relation determines equivalence classes of F.

The cases where J is a constant are exceptional (like surfaces with constant curvature).

For example, the Uniform distribution is completely characterised by $J = -1$.

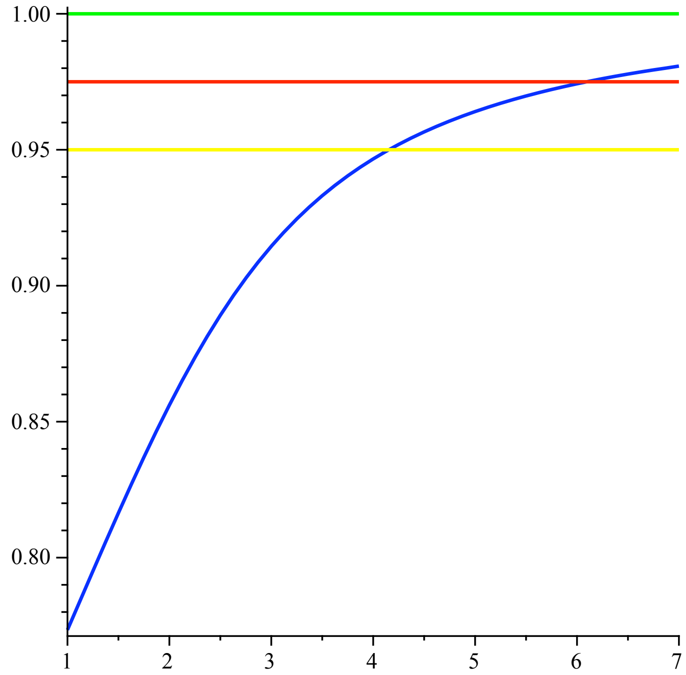
So every distribution F on an interval $[A, B]$ for which $J = -1$ can be translated and re-scaled to the standard uniform distribution $U(x) = x$ on $[0, 1]$.

Appendix 3. Fisher and Tippett's 'Penultimate Approximation'

The very slow convergence of the Normal distribution to its EVT limit was noted by Fisher and Tippett.

In terms of our invariants we can see that the Normal distribution remains closer to E_{-20} for which $K = 0.95$ than it is to E_{∞} for all quantiles that are realistic for actual applications.

You have to care about values with a probability of less than 1 in 1 billion for the actual limit to be a better approximation than E_{-20}, for example.



E_{-20} is a better approximation than E_{∞} for all $x < 6.09$

For $x > 6.09$ the Normal distribution differs from 1 by only one part in 10^9.