Sergey Kurenoy Novosibirsk 26 October, 2019

Test problem for Perfect Art interview:
Expected shortfall distribution using Monte-Carlo simulation

To generate the required three-year (N = 750 trading day) time series, I used the standard scheme of Mandelbrot[1] where the daily difference in the natural logarithms of daily prices P_i is modeled by random variables X distributed according to the stable distribution (also known as Lévy alpha-stable distribution, or Pareto-Lévy stable distribution): $X=\ln(P_{i+1})-\ln(P_i)$. The time series of N+1 prices is produced by rearranging this equation for the subsequent price $P_{i+1}=P_i\exp(X)$. (Note that the initial price P_0 is arbitrary and inconsequential because only price ratios are relevant for proportional returns.) The random variables X of the stable distribution with parameters $\alpha, \beta, \gamma, \delta$ (next paragraph) are generated using `SciPy.stats.levy_stable.rvs`. Note that for a symmetric distribution, having skewness parameter $\beta = 0$, the scale parameters $\gamma = c$ and location parameters $\delta = \mu$ coincide in the two competing parametrizations of the stable distribution, eliminating a common source of confusion. The time series of N one-day proportional returns r_i^1 and N–9 overlapping ten-day proportional returns r_i^{10} are then calculated according to the given equations: $r_i^1=(P_{i+1}/P_i)-1$ and $r_i^{10}=(P_{i+10}/P_i)-1$. Finally, the 1% percentile worst-case two week (ten-day) return is found by applying the `NumPy.quantile` function to the generated r_i^{10} time series and $q = 0.01$.

The given stability parameter $\alpha = 1.7$ and skewness parameter $\beta = 0$ of the stable distribution used in this model agree exactly with Mandelbrot's original analysis of cotton prices. These suggested parameter values are also consistent with existing fitting[2] to real-world returns of US-listed stocks, where α generally varies between 1.5 and 1.8, and β is small ($\beta = 0 \pm 0.01$). However, the given scale parameter $\gamma = 1.0$ and location parameter $\delta = 1.0$ are unrealistically large, in the sense that the corresponding expected daily return would be $e^\delta - 1 \approx 170\%$ with a volatility of roughly $\gamma\sqrt{2} \approx 140\%$, which leads to wild price fluctuations and unrestricted exponential growth (final P_i values may exceed the maximum `double` limit). Thus, it is reasonable to restrict the model to much smaller γ and δ parameters, perhaps so that expected daily return and volatility do not exceed around 10% under "normal" conditions. (Note that this and most stochastic financial models presume a "normal market", and therefore cannot possibly account for extreme events such as market crashes.) This means that realistic scale and location parameters are in the following ranges: $0 < \gamma \leq 0.07$ and $-0.10 \leq \delta \leq 0.10$. Due to limited time and computing productivity, in practice I will further restrict these parameters to the "most interesting" ranges: $0.01 \leq \gamma \leq 0.05$ and $-0.02 \leq \delta \leq 0.06$.

---

1 Benoit Mandelbrot, "The Variation of Certain Speculative Prices", *The Journal of Business*, Vol. 36, No. 4 (Oct., 1963), pp. 394-419, The University of Chicago Press. See http://www.jstor.org/stable/2350970

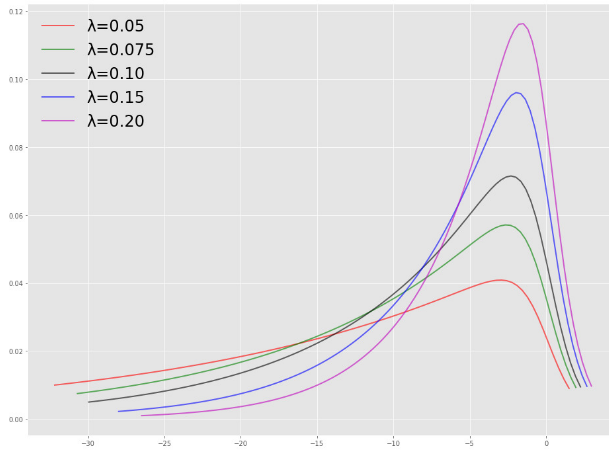
2 Riccardo Donati and Alice Pisani, Redexe "Risk Management and Finance" and University of Parma, presentation "Pareto-Lévy stable distributions in Action!", 23 Nov., 2011. See http://www.redexe.net/docs/redesfull.pdf

The qualitative observations about distributions that emerge when looking at samples of points representing 1% percentile overlapping ten-day returns from independent time series instances with fixed $\{\gamma, \delta\}$ parameters, along with my *proposed "financial interpretations"*, are as follows:

- A continuous non-symmetric or skewed distribution with a prominent peak and a long left/negative tail, but no right/positive tail. That is, *1% percentile returns are expected most frequently at a particular expectation value, although they can sometimes be much worse, but rarely any better than this fixed value.*
- The mode/maximum of the peak is generally centered at a negative value, unless $\gamma >> \delta$. This mode is more negative when γ is small and δ is big. That is, *this aforementioned expectation value will usually be a negative return, unless growth rate dominates volatility. The expected return becomes more negative when the growth rate is lower and the volatility is higher.*
- There is a pileup near return value –1 at the end of the tail, with a higher frequency when γ is big. That is, *the worst possible outcome of nearly –100% return (or default, or near zero final price) is realized more frequently at the 1% percentile when volatility is high.*

Please see the 15-part full page figure on page 4, which more fully illustrates these observations.

Quantitatively, there are several theoretical distribution types that can fit these collections of points by closely corresponding to the shape of the numerically generated distribution and reflect the first two observations. (In fact, Wikipedia's "Expected shortfall" article lists 17 different suggested distributions.) To reflect the third observation about pileups, the theoretical distribution's tail must be cut off when below –1, and lost instances recompensated in the leftmost bin. (Alternatively, cases with large pileups will not fit as well to a long-tailed theoretical distribution.) In any case, it seems unlikely that there is a truly good distribution with strong mathematical motivation (as in Mandelbrot's use of stable distribution) for these types of results. (Otherwise, someone would have invented it, published their results, and received significant attention from the financial analysis community.) Thus, I must pick a "working distribution" in order to attempt a quantitative description of the results.



I used a brute-force readily available method to find the best fit[3] by minimizing the sum of squared errors between a fixed $\{\gamma, \delta\}$ case and every available distribution in `SciPy.stats`, the two winning distributions for cases with small pileups were the Type I generalized logistic (also known as skew-logistic) distribution and Johnson's S_U distribution. I will fit all cases to the skew-logistic distribution because it is simpler (has only one shape parameter λ besides scale and location, while Johnson's S_U has two). The probability density function with positive shape parameter λ results in a left tail. The mode/maximum of the peak in this skew-logistic distribution is easily found explicitly by maximizing its probability density function: $f(y,\lambda)=\lambda e^{-y}/(1+e^{-y})^{\lambda+1}$ with $y = (x - \text{location})/\text{scale}$. This mode occurs at $y = \ln(\lambda)$, or equivalently, at $x = \text{location} + \text{scale} \ln(\lambda)$, and has value $f_{max}=(1+1/\lambda)^{-(\lambda+1)}$.

To test whether the samples of 1% percentile overlapping ten-day returns plausably follow the proposed skew-logistic distribution, I use the chi-square goodness-of-fit test[4]. With $k = 33$ bins, I compute using the `SciPy.stats.chisquare` function (or manually) the chi-square test statistic χ^2:

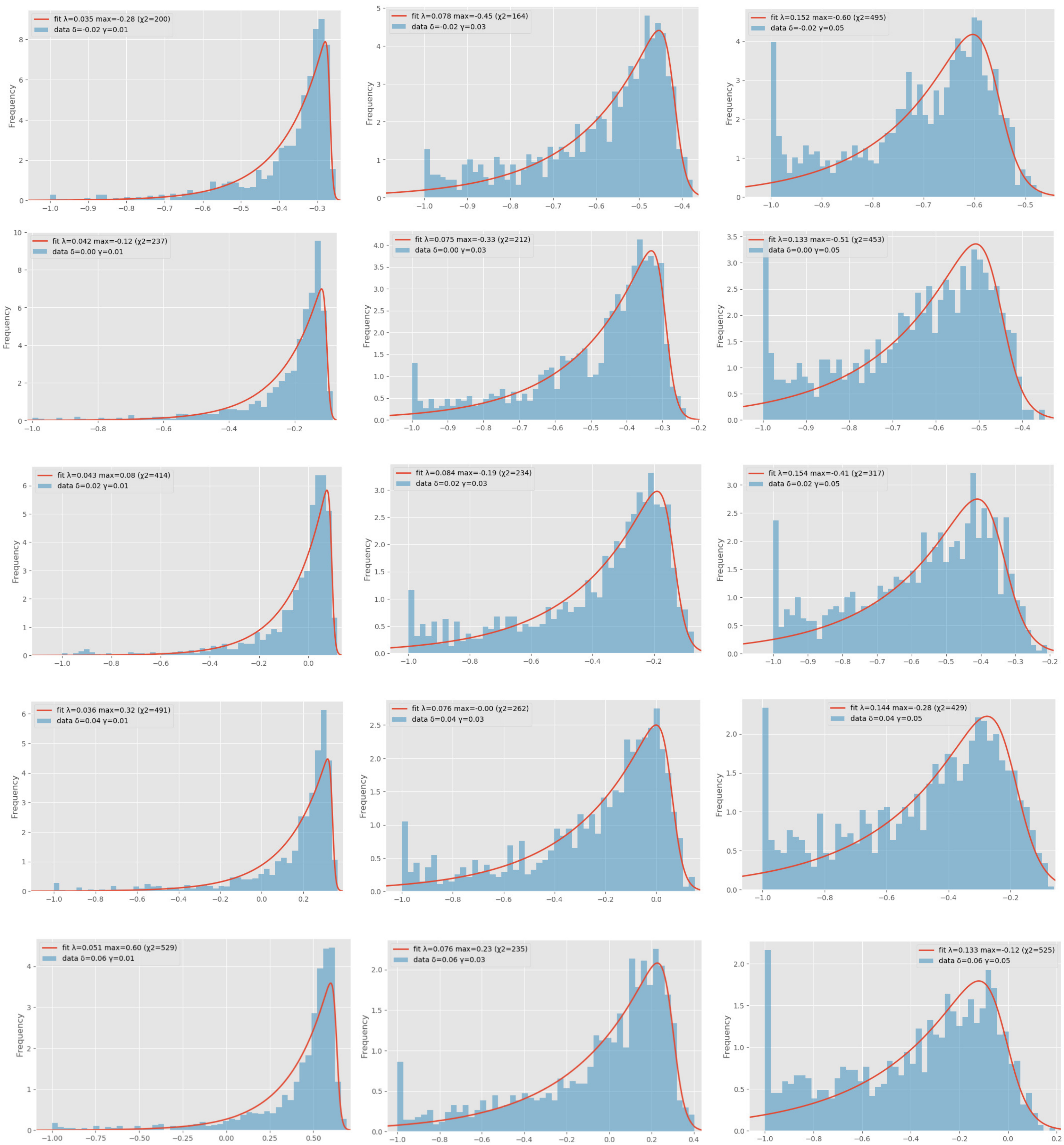
$$\chi^2=\sum_{i=1}^{k}(O_i-E_i)^2/E_i$$

Here, O_i is the observed frequency of the sample in the i-th bin, and $E_i=M(F(x_{upper})-F(x_{lower}))$ is the expected frequency of the proposed distribution, in terms of the sample size M and the difference of the cumulative density function F evaluated at the bin edges x_{upper} and x_{lower}. It is suggested that sample size M should be large enough that every bin's expected frequency is at least 5, so in practice $M > 1000$ is needed for the bins to fill up sufficiently at the tail. (I used $M = 1200$.) For 95% confidence that the sample is distributed according to this theoretical distribution (with 3 parameters: shape, location, and scale), with $k - c = 33 - 3 - 1 = 29$ degrees of freedom, the chi-square test statistic must not exceed the critical value of $\chi^2_{\text{crit}}(0.95, 29) = 44.557$. The resulting χ^2 of my cases range between 164 and 529, and as much as 300 for cases with a large pileup (which indeed produce worse fits). This means, as expected, that it is "highly unlikely" that these samples are really distributed according to the skew-logistic distribution, but some of these fits look convincing (according to the accompanying p-value) in some cases with low χ^2 and minimal pileup.

3 Code borrowed entirely from the top answer by user tmthydvnprt, saved as bestfit.py in my repository. See https://stackoverflow.com/questions/6620471/fitting-empirical-distribution-to-theoretical-ones-with-scipy-python

4 National Institute of Standards and Technology (NIST) *e-Handbook of Statistical Methods* , "1.3.5.15. Chi-Square Goodness-of-Fit Test", See https://www.itl.nist.gov/div898/handbook/eda/section3/eda35f.htm

Figure: full array of results, arranged with γ increasing leftward, δ increasing downward. Note that the frequencies are normalized (using `Density=True`) to correspond to the probability density functions of the theoretical skew-logistic distribution, and independent of sample size $M = 1200$.



With more time, I could find empirically the explicit dependencies of the fit parameters on the γ and δ parameters used to generate different samples. In particular, financial analysts would probably be interested in being able to predict the sample mode value, which is the most likely 1% percentile two-week return at the maximum of the fitting distribution. One could also use the stability property of the stable-distribution (that linear combinations of random variables have the same distribution, up to location and scale) to speculate about how 10-day returns are distributed (location increases by 10, scale increases by $10^{-1/\alpha} \approx 3.87$, but α and β invariant), but they must be strongly correlated because they are overlapping.