
Large Deviations of Factor Models with Regularly-Varying Tails: Asymptotics and Efficient Estimation

Farzad Pourbabaee
(farzad@berkeley.edu)

Department of Economics, UC Berkeley

April 4, 2017

Overview

- *Heavy tails* in Economics and Finance: firm and city size, CEO compensations, income and wealth distribution (Gabaix 2016).
 - Majority of financial assets, in particular stocks have heavy tails; that precludes Gaussian/log-Gaussian modeling (for e.g the kurtosis of S&P 500 return's empirical distribution is somewhere around 10).
 - The large deviation analysis of heavy tails is less studied in the literature; in particular simulating extreme events and their likelihood estimation.
-

Problem Statement

- Suppose X_1, \dots, X_N are independent, \mathbb{R} -valued random variables (N is fixed). What are the asymptotics of $\mathbb{P}[X_1 + \dots + X_N > x]$ as $x \rightarrow \infty$? and how can we estimate that?
 - Where do these types of likelihood appear?
 - Extreme losses or profits of a portfolio exposed to multiple independent risk factor.
 - Frequency of large economic downturns, and significant GDP departures from equilibrium trend, caused by the heavy-tail nature of microeconomic shocks (Acemoglu, Ozdaglar & Tahbaz-Salehi 2017).
 - What is particularly important for this estimation? As $x \rightarrow \infty$ the deviation probability becomes excessively small, and finding *nontrivial* confidence intervals as a measure of estimator's efficiency becomes more difficult.
-

-
- Why simulation?
 - Because, there may be no closed form solution for large deviation likelihood.
 - We may not know the form of individual distribution, and we have to rely on the empirical observations which look like heavy tails.
 - Finally, the estimation approach is different for light tailed distributions and heavy tails.
 - In general there are two broad techniques for extreme event simulation: **Importance sampling and Monte-Carlo methods.**
-

Example of Market Portfolio (Gaussian Environment)

- Assume there are M assets available in the market whose returns are driven by k latent factors $\phi = (\phi_1, \dots, \phi_k)$. The return to asset i is

$$\eta_i = \langle \beta_i, \phi \rangle + \varepsilon_i, \quad (1)$$

where $\phi \sim \mathcal{N}(0, I_k)$ and $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$.

- Let the market portfolio be the uniformly weighted average return of individual assets: $\xi = \frac{1}{M} \sum_{i=1}^M \eta_i = \langle \bar{\beta}, \phi \rangle + \bar{\varepsilon}$.
- For large x how to calculate $\mathbb{P}[\xi > x]$? **Importance sampling with an appropriate choice of sampling measure.**

$$\xi \sim \mathcal{N} \left(0, \|\bar{\beta}\|_2^2 + \sum_{i=1}^M \sigma_i^2 / M^2 \right) \quad (2)$$

- The cumulative generating function $\psi(\theta)$ exists for Gaussian distribution for all $\theta \in \mathbb{R}$,

$$\psi(\theta) = \log \mathbb{E} \left[e^{\theta \xi} \right] = \frac{\theta^2}{2} \left(\|\bar{\beta}\|_2^2 + \frac{1}{M^2} \sum_{i=1}^M \sigma_i^2 \right) \quad (3)$$

- One candidate for the importance sampling distribution is \mathbb{P}_θ , where

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = e^{\theta \xi - \psi(\theta)} \quad (4)$$

- Now we can generate n samples from \mathbb{P}_θ , and form the following sample average, which represents the unbiased estimator under the new measure \mathbb{P}_θ :

$$\frac{1}{n} \sum_{i=1}^n 1_{[\xi_i > x]} \frac{d\mathbb{P}}{d\mathbb{P}_\theta}(\xi_i) \quad (5)$$

- Denote the per-sample estimator by $Z(x) = 1_{[\xi > x]} \frac{d\mathbb{P}}{d\mathbb{P}_\theta}(\xi)$. The next definition spells out two notions that characterize relative error optimality.

Notions of Estimator's Efficiency

Definition 1. The estimator $Z(x)$ has *bounded relative error* if

$$\limsup_{x \rightarrow \infty} \frac{\text{Var}(Z(x))}{\mathbb{E}[Z(x)]^2} < \infty, \quad (6)$$

and is *logarithmically efficient* (a weaker notion) if for some $\varepsilon > 0$

$$\limsup_{x \rightarrow \infty} \frac{\text{Var}(Z(x))}{\mathbb{E}[Z(x)]^{2-\varepsilon}} = 0. \quad (7)$$

Importance Sampling Efficiency (Asmussen 2008)

Theorem 2. The exponential change of measure in (4) is logarithmically efficient for the unique parameter θ that solves $x = \psi'(\theta)$.

Example of Market Portfolio (Gaussian Environment)

- As a result of twisting the sampling distribution, the relative error scales as $\mathbb{P}[\xi > x]^{-\varepsilon/2}$, compared to $\mathbb{P}[\xi > x]^{-1/2}$ without any measure transformation.
 - But, what if the factors or the individual asset risk terms are not Gaussian? **The moment generating function may not exist, then whole measure change methods breaks down!**
 - Therefore, to find the optimal measure change we have to appeal to heuristic methods, or use **Monte-Carlo methods**.
-

Regularly-Varying Environment

RV Distribution

Definition 3. A distribution function F has a *regularly varying* tail, if $\bar{F}(x) \sim L(x)/x^\alpha$ as $x \rightarrow \infty$, where $\alpha > 0$ and $L(\cdot)$ varies *slowly* at infinity, i.e.

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1 \text{ for all } t > 0. \quad (8)$$

A Regularity Condition

Condition 4. F satisfies the *h-condition*, if there exists an eventually increasing function $h(x)$ such that $\lim_{x \rightarrow \infty} h(x) = \infty$ and

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x + h(x))}{\bar{F}(x)} = 1. \quad (9)$$

Sum Asymptotics

Theorem 5. Suppose X_1, \dots, X_N are independent random variables in \mathbb{R} , such that:

- (i) An RV distribution F exists, such that $\bar{F}_i(x) \sim c_i \bar{F}(x)$ for all i 's and at least one $c_i \neq 0$,
- (ii) F satisfies the *h-condition*,

then the following asymptotic result holds:

$$\mathbb{P}[X_1 + \dots + X_N > x] \sim \sum_{i=1}^N \mathbb{P}[X_i > x] \sim \left(\sum_{i=1}^N c_i \right) \bar{F}(x) \quad (10)$$

Maximum Factor Asymptotics (Almost Catastrophe Principle)

Theorem 6. Suppose X_1, \dots, X_N are independently drawn from F_1, \dots, F_N , and take values in \mathbb{R} . Then, under the same conditions **i** and **ii** of theorem 5, the following asymptotic result holds: then the following asymptotic result holds:

$$\begin{aligned} \sum_{i=1}^N \mathbb{P}[X_i > x] + o(\bar{F}(x)) &\leq \mathbb{P}\left[\max_{1 \leq i \leq N} X_i > x\right] \\ &\leq (1 - e^{-1})^{-1} \sum_{i=1}^N \mathbb{P}[X_i > x] + o(\bar{F}(x)) \end{aligned} \tag{11}$$

Estimation of the Large Deviation Probability

- One candidate is to take $\sum_{i=1}^N \mathbb{P}[X_i > x]$ as an estimator, that becomes more precise as $x \rightarrow \infty$. However, it turns out that it performs very weakly.
 - A conditional Monte-Carlo (CMC) algorithm is developed in (Asmussen, Kroese et al. 2006) to cope with the tail probability of sum of i.i.d heavy tails. That idea is incorporated here to obtain an estimator for the sum of independent but non-identical factors.
 - CMC steps:
 - (i) Take a sample draw X_i from its corresponding distribution F_i .
 - (ii) Let $M_N = \max\{X_i : i \in [N]\}$.
 - (iii) $Z(x) = \sum_{i=1}^N \mathbb{P}[S_N > x, M_N = X_i | X_{-i}]$ gives an unbiased estimator for $\mathbb{P}[S_N > x]$.
-

CMC Efficiency Analysis

- The proposed CMC estimator has *bounded relative error*, i.e. $\limsup_{x \rightarrow \infty} \frac{\text{Var}(Z(x))}{\mathbb{E}[Z(x)]^2} < N^{2\alpha} < \infty$.
- Let $\mu(x) = \mathbb{P}[X_1 + \dots + X_N > x]$, and $\sigma(x)^2 = \text{Var}(Z(x))$; repeat the CMC n times, and take the sample average $\bar{Z}_n(x)$.
- **CLT confidence interval**: CLT implies $\frac{\bar{Z}_n(x) - \mu(x)}{\sigma(x)/\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$, and

$$\bar{Z}_n(x) \in (\mu(x)(1 - \kappa), \mu(x)(1 + \kappa)) \quad \text{w.p.} \left(2\Phi(\kappa\sqrt{n}N^{-\alpha}) - 1 \right) \quad (12)$$

- **Markov confidence bound**:

$$\mathbb{P} \left[|\bar{Z}_n(x) - \mu(x)| > \kappa\mu(x) \right] \leq \frac{\mathbb{E} \left[(\bar{Z}_n(x) - \mu(x))^2 \right]}{\kappa^2 \mu(x)^2} \leq \frac{N^{2\alpha}}{\kappa^2 n} + o_x(1), \quad (13)$$

CMC vs. Crude Monte-Carlo

Sample mean estimator after n repetition is

$$\hat{\mu}_n(x) := \frac{1}{n} \sum_{k=1}^n 1_{\left[X_1^{(k)} + \dots + X_N^{(k)} > x \right]} \quad (14)$$

where $X_i^{(k)}$ is the k th independent draw from F_i .

Comparison of Estimators' Concentration Speed

Proposition 7. For any precision level $0 < \kappa < 1$, the CMC estimator $\bar{Z}_n(x)$ is exponentially more efficient than $\hat{\mu}_n(x)$. Namely, for any $0 < r < 2\kappa^2 N^{-2\alpha}$

$$\limsup_{x \rightarrow \infty} \lim_{n \rightarrow \infty} \left\{ rn + \log \left(\frac{\mathbb{P} \left[\left| \bar{Z}_n(x) - \mu(x) \right| > \kappa \mu(x) \right]}{\mathbb{P} \left[\left| \hat{\mu}_n(x) - \mu(x) \right| > \kappa \mu(x) \right]} \right) \right\} = 0. \quad (15)$$

Market Portfolio Large Deviation

- One of the main motivations of studying RV distributions was to capture the large deviations of asset returns, as initially laid out for the Gaussian case.
- Remember the factor model: $\eta_i = \langle \beta_i, \phi \rangle + \varepsilon_i$ and the *market portfolio*: $\bar{\eta} = \langle \bar{\beta}, \phi \rangle + \bar{\varepsilon}$.
- Assume all factors (ϕ_1, \dots, ϕ_k) and individual idiosyncratic noises $(\varepsilon_1, \dots, \varepsilon_M)$ have RV tails. Then, what can we say about $\mathbb{P}[\bar{\eta} > x]$?

$$\begin{aligned}
 \mathbb{P} \left[\langle \bar{\beta}, \phi \rangle + \frac{1}{M} \sum_{i=1}^M \varepsilon_i > x \right] &\sim \mathbb{P} [\langle \bar{\beta}, \phi \rangle > x] + \mathbb{P} \left[\frac{1}{M} \sum_{i=1}^M \varepsilon_i > x \right] \\
 &\stackrel{(*)}{\sim} \mathbb{P} [\langle \beta, \phi \rangle > x]
 \end{aligned}
 \tag{16}$$

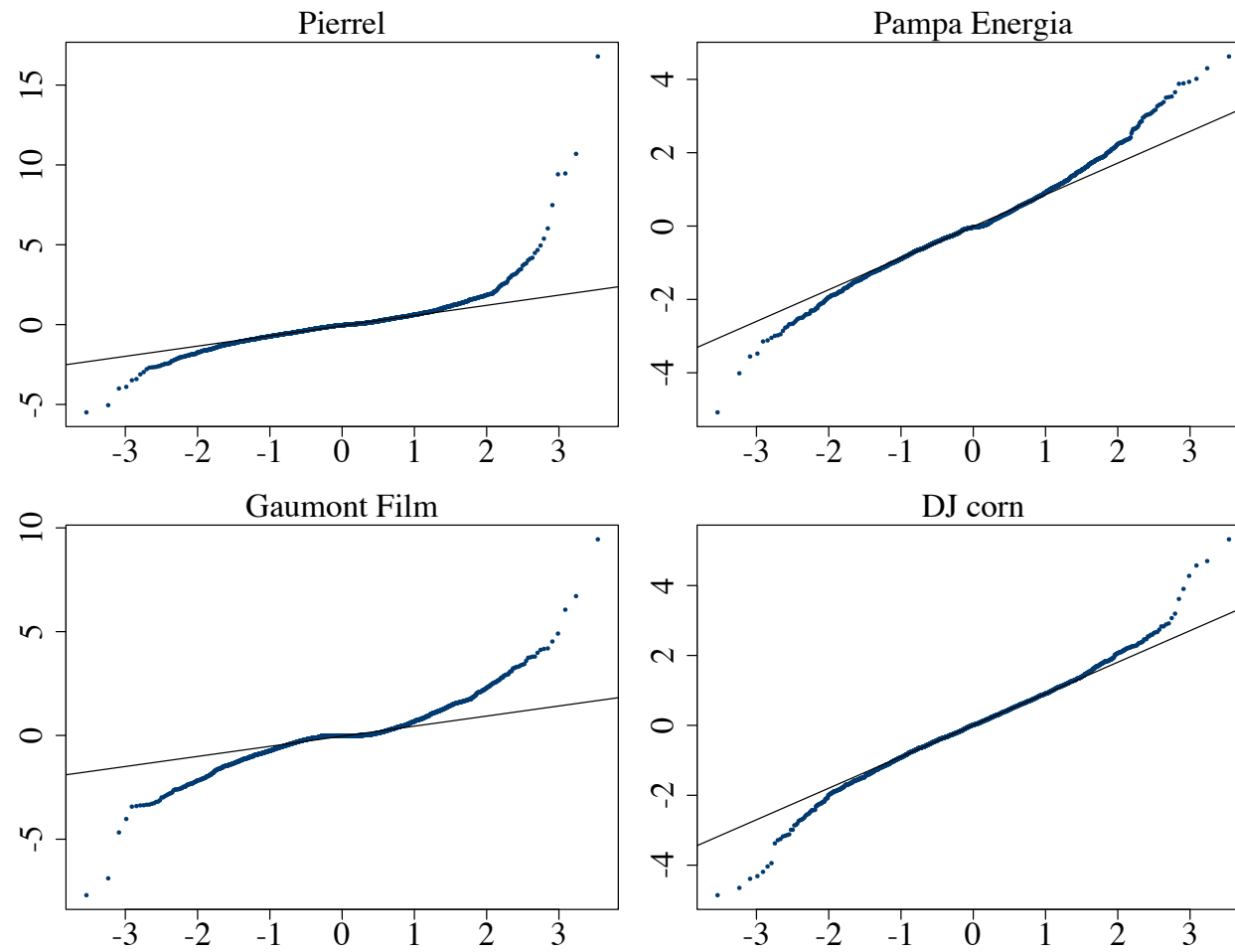
Empirical Implications

- Historical daily closed prices of Pierrel (the Italian pharmaceutical company), Pampa Energia (the electricity company in Argentina), Gaumont Film (the oldest film production company located in France) and Dow Jones commodity corn index over the period of 12/2006-12/2016 are pulled out from compustat.
- AR(1)+GARCH(1,1) is then fit to all four stock returns:

$$\begin{aligned}
 r_t &= \mu_t + \underbrace{\eta_t}_{\sigma_t \varepsilon_t} := \mathbb{E} [r_t | \mathcal{F}_{t-1}] + (r_t - \mathbb{E} [r_t | \mathcal{F}_{t-1}]) \\
 \sigma_t^2 &= \omega + \gamma \eta_{t-1}^2 + \theta \sigma_{t-1}^2 \\
 \mu_t &= \lambda + \varphi \mu_{t-1}
 \end{aligned} \tag{17}$$

- QQ-plot of *standardized returns* v.s $\mathcal{N}(0, 1)$:

Figure 1: QQ-plot versus standard Normal



-
- Gaussianity over both left and right tails are violated! So we can fit a Generalized Pareto distribution to both tails:

$$f(x; \sigma, \alpha) = \frac{1}{\sigma} \left(1 + \xi \frac{x}{\sigma} \right)^{-(1+1/\xi)} \quad \text{for } x > 0. \quad (18)$$

- The largest 10% of both positive and negative residuals in absolute value are taken, then the ML estimates of the shape and scale parameters are found.
-

Table 1: GPD ML estimates for positive and negative tails

| | Pierrel | Pampa Energia | Gaumont Film | DJ corn |
|------------|------------------|-------------------|------------------|------------------|
| ξ_+ | 0.527 (0.141) | -0.006 (0.126) | 0.086 (0.098) | 0.050 (0.082) |
| σ_+ | 0.600 (0.097) | 0.631 (0.099) | 0.903 (0.126) | 0.554 (0.067) |
| ξ_- | 0.101 (0.092) | 0.009 (0.084) | 0.059 (0.073) | 0.111 (0.114) |
| σ_- | 0.541 (0.068) | 0.536 (0.064) | 0.727 (0.079) | 0.563 (0.081) |

Pierrel's QQ-plot together with the first row of table 1 confirm that it has the heaviest positive tail among four, and roughly second to DJ corn for the negative tail

- This observation on heavy tailed nature of residual series backs up the need for the efficient methods of deviation probability estimation, as many events are directly or indirectly relied on such likelihoods.

References

- Acemoglu, Daron, Asuman Ozdaglar & Alireza Tahbaz-Salehi (2017), ‘Microeconomic origins of macroeconomic tail risks’, *American Economic Review* **107**(1), 54–108.
- Asmussen, Søren (2008), *Applied probability and queues*, Vol. 51, Springer Science & Business Media.
- Asmussen, Søren, Dirk P Kroese et al. (2006), ‘Improved algorithms for rare event simulation with heavy tails’, *Advances in Applied Probability* **38**(2), 545–558.
- Gabaix, Xavier (2016), ‘Power laws in economics: An introduction’, *The Journal of Economic Perspectives* **30**(1), 185–205.
-

....

Thanks
