

塔勒布量化开篇之作《肥尾分布的统计效应》（下）

作者：戴国晨

导读：真实世界中的随机性并非无迹可循，在日常经验中，人们发现幂律分布（帕累托分布）可以近似描述大量真实世界的现象，如市场波动，财富分配，流量效应，灾难损失等等。更准确的讲，这些随机事件的尾部都具备幂律特征，在对数坐标系下生存函数以直线形式下降。

这是极为神奇的现象，从标普 500 指数的尾部收益率到大型战争的极端伤亡人数，横跨不同领域的风险事件在肥尾的数学框架中得到了统一。二八定律周而复始的出现，映照出世界背后的底层逻辑：均衡不过是奢望，极端才是常态。

本书的下半部分就从实际肥尾分布出发，引出预测和赔付的关系，最终收尾于期权定价，构建了从现象到规律，思路到工具的闭环。

在标普 500 指数的收益率分布中，可以看到随着周期拉长，收益率峰度逐渐下降，呈现出缓慢的中心极限定理。这也是很多金融产品的共同特点：短期市场总是过度反应，收益率服从幂律尾分布，长期则回归理性，收益率逐渐向正态靠拢。这样的回报特点会给不同策略的鲁棒性带来巨大差异，价值投资聚焦长期，结果随着周期拉长变得愈发稳定，而市场上林林总总的短期策略则极大地暴露在尾部风险中。

在幂律尾分布下，历史极值或条件均值的风险视角有着巨大缺陷，尚未发生的尾部事件可能会极大的影响整体统计性质。因此任何短期策略想要长期存活，必须在极端条件下保护自身。随着市场不断发展，对尾部保护的需求逐渐演变为今天的期权合约，并发展出一系列量化定价方法。本书的期权部分主要从肥尾分布的角度切入，对合约的绝对和相对定价做出了思考。

读塔勒布之前看尾部风险，如同身处黑暗，满是难以量化的混沌与恐惧，读塔勒布之后看尾部风险，朦胧间已然瞥见一抹光亮，尽管依然无法看清风险的全貌，但是已经能够辨别可知与不可知，跳出被随机性愚弄的轮回。以上是一点个人感悟，也希望各位读者可以从中学有所收获。

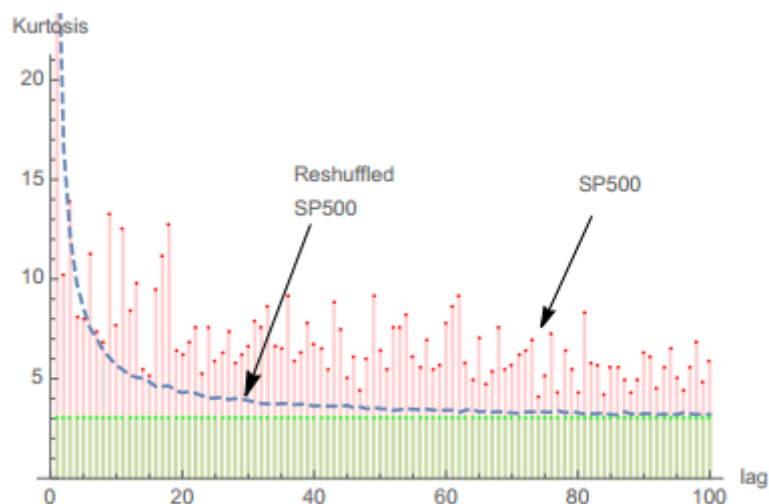
标普 500 分布

通过我们具备的各类工具，可以对美国标普 500 指数进行多角度统计分析。SP500 作为金融市场中最重要指数，其回报率的尾部满足幂律分布（存在一定不对称性），我们通过历史收益率关注如下几点：

- 随着周期增加累计峰度的性质
- 对于不同的 K 其条件期望 $E(X|X>K)$ 的性质
- 各阶矩的 MS 图（极大值贡献 max-to-sum）
- 最大回撤
- 极值分析（极值是否互相独立，波动率集聚程度）

累计峰度

SP500 单日收益率呈现出很高的峰度，但是如果计算不同时间周期对数收益率的分布（日收益率的和分布），根据大数定律该分布应该趋向于高斯分布。但是通过从历史数据得到的峰度结果如下图所示：



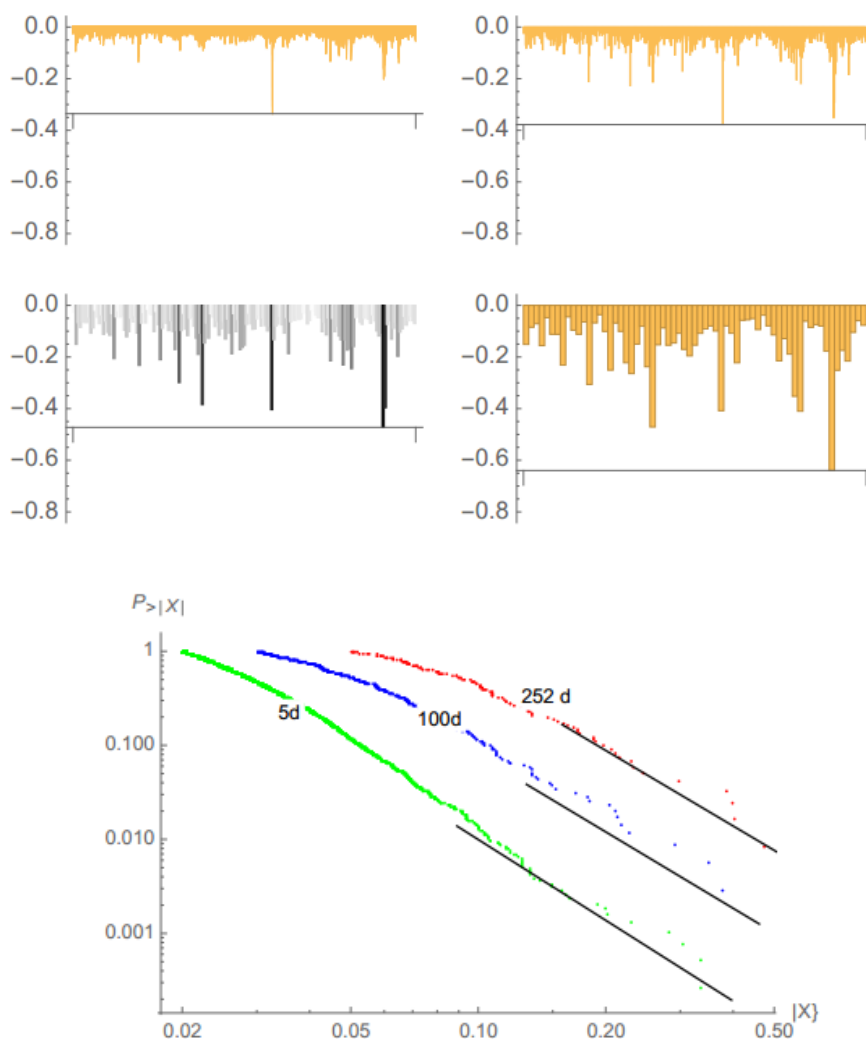
可以看到，拉长收益率周期分布没有明显的迅速收敛到正态，理论上我们应该观察到峰度以 n^{-1} 速度的下降，并收敛于正态分布。但是如果我们将 SP500 收益率随机打乱再分析和分布，其收敛行为就很明显了。因此我们可以认为收益率的肥尾来自数据内部结构，也就是波动率集聚现象。

最大回撤

资产 S 在 t_0 到 $t_0 + n\Delta t$ 之间的最大回撤定义为：

$$\delta(t_0, t, \Delta t) = \text{Min} \left(S_{i\Delta t+t_0} - \left(\text{Min} S_{j\Delta t+t_0} \right)_{j=i+1}^n \right)_{i=0}^n$$

下面是 $n=5, 30, 100$ 和 252 天的回撤，通过 log-log 图可以看到回撤尾部满足帕累托分布

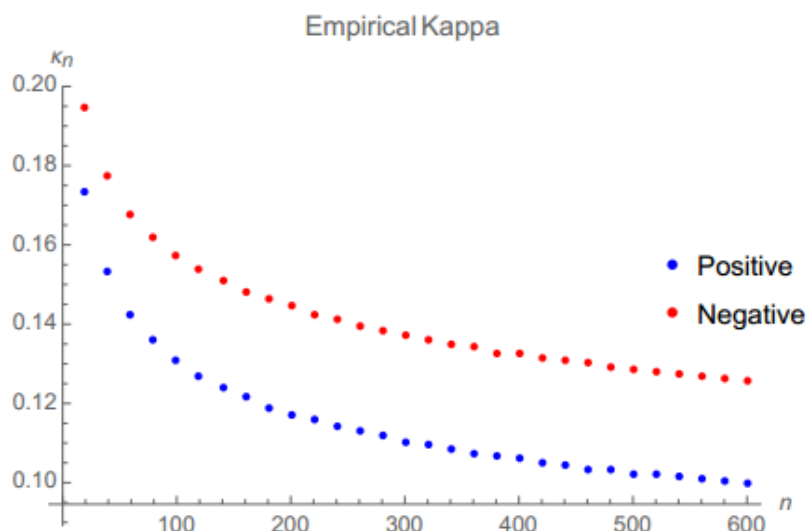


Kappa 值

肥尾指标 κ 由之前的定义：

$$\kappa_{n_0,n} = 2 - \frac{\log(n) - \log(n_0)}{\log\left(\frac{\mathbb{M}(n)}{\mathbb{M}(n_0)}\right)}$$

这里我们令 $\kappa_n = \kappa(1, n)$ ，我们可以看到负收益率的尾部吻合 $\alpha < 3$ 的帕累托分布。



条件期望

我们在这里测试收益率超过给定值 K 的条件期望和 K 的比值，也即：

$$\varphi_K^+ = \frac{E(X)|_{X>K}}{K} \quad \varphi_K^- = \frac{E(-X)|_{X>K}}{K}$$

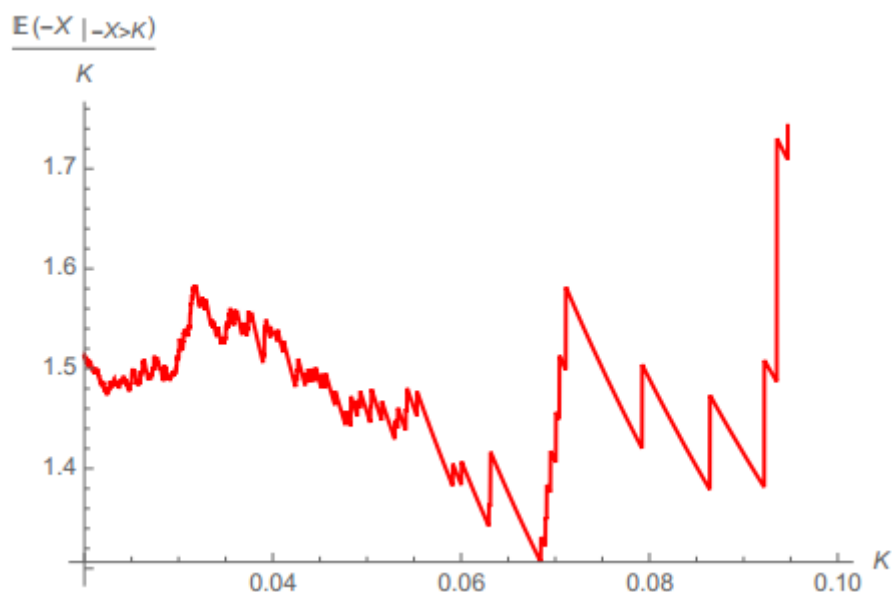
对于非幂律尾部，比如薄尾分布（正态分布）：

$$\lim_{K \rightarrow \infty} \varphi_K = 0$$

对于幂律尾部：

$$\lim_{K \rightarrow \infty} \varphi_K / K = \frac{\alpha}{1 - \alpha}$$

从 SP500 的结果来看，随着 K 增大该比值并没有向 0 收敛，因此可以证明 SP500 的尾部更靠近幂律分布



四阶矩不稳定性

下面的表格中展示了在 SP500 超过 50 年的历史中，单日回报率对峰度的最大贡献高达 79%。这种超大单日极值贡献在其他金融资产中也很常见，比如原油（79%），白银（94%），其他商品和股票指数等。如此依赖极值也说明了金融资产回报率的峰度高度不稳定，甚至很可能并不存在。

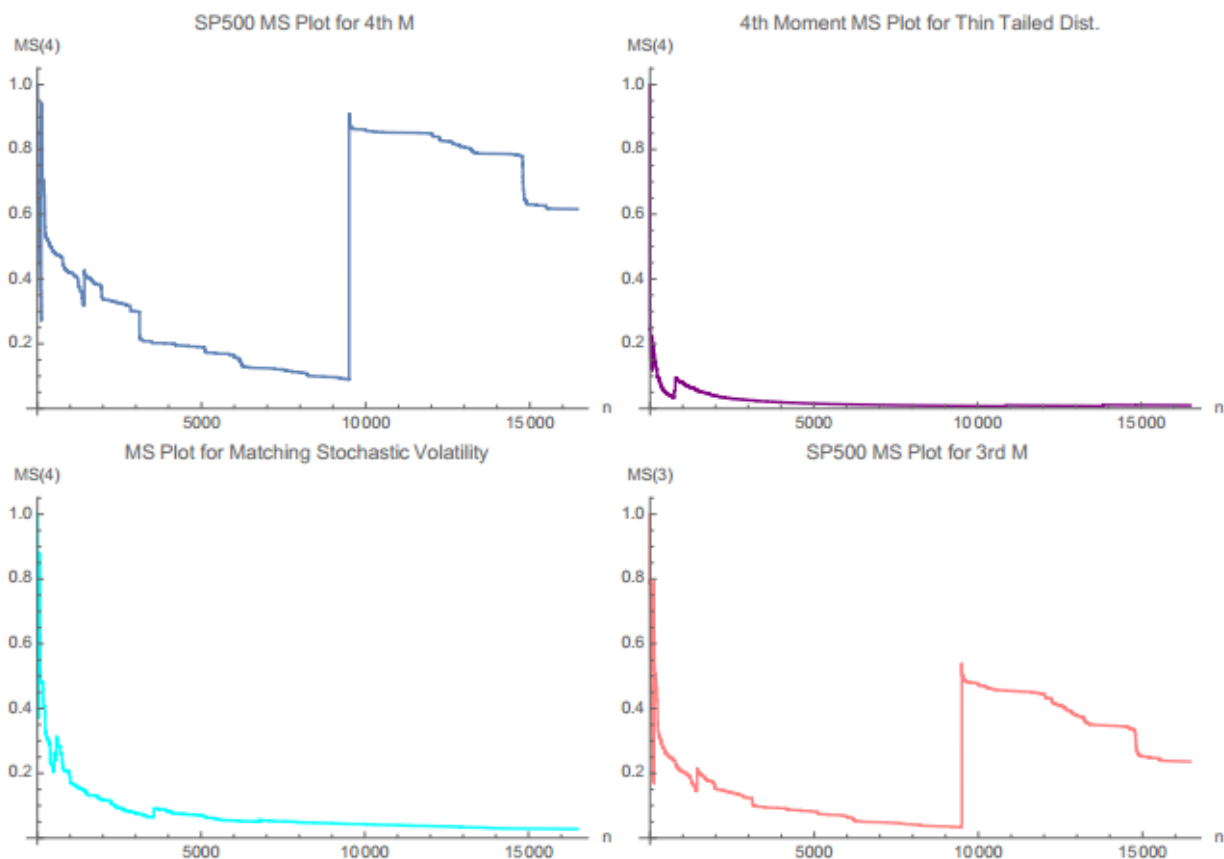
| Security | Max Q | Years. |
|--------------------|-------|--------|
| Silver | 0.94 | 46. |
| SP500 | 0.79 | 56. |
| CrudeOil | 0.79 | 26. |
| Short Sterling | 0.75 | 17. |
| Heating Oil | 0.74 | 31. |
| Nikkei | 0.72 | 23. |
| FTSE | 0.54 | 25. |
| JGB | 0.48 | 24. |
| Eurodollar Depo 1M | 0.31 | 19. |
| Sugar #11 | 0.3 | 48. |
| Yen | 0.27 | 38. |
| Bovespa | 0.27 | 16. |
| Eurodollar Depo 3M | 0.25 | 28. |
| CT | 0.25 | 48. |
| DAX | 0.2 | 18. |

极大值贡献图

另外一种测量 $E(X^p)$ 稳定性的方法是极大值贡献（Max to Sum）图。MS 图依赖大数定律，对于满足独立同分布的 X_1, X_2, \dots, X_n ，如果 $E(X^p) < \infty$ ，那么

$$R_n^p = \frac{M_n^p}{S_n^p} = \frac{\max(X_1^p \dots X_n^p)}{X_1^p + \dots + X_n^p} \rightarrow 0 \quad (n \rightarrow \infty)$$

从图中我们可以看出 SP500 收益率呈现出非常陡峭的幂律分布特征，对于三阶和四阶矩在 50 年（16000 个点）范围的回测上依然显著不为 0，因此不满足大数定律。



极值分析

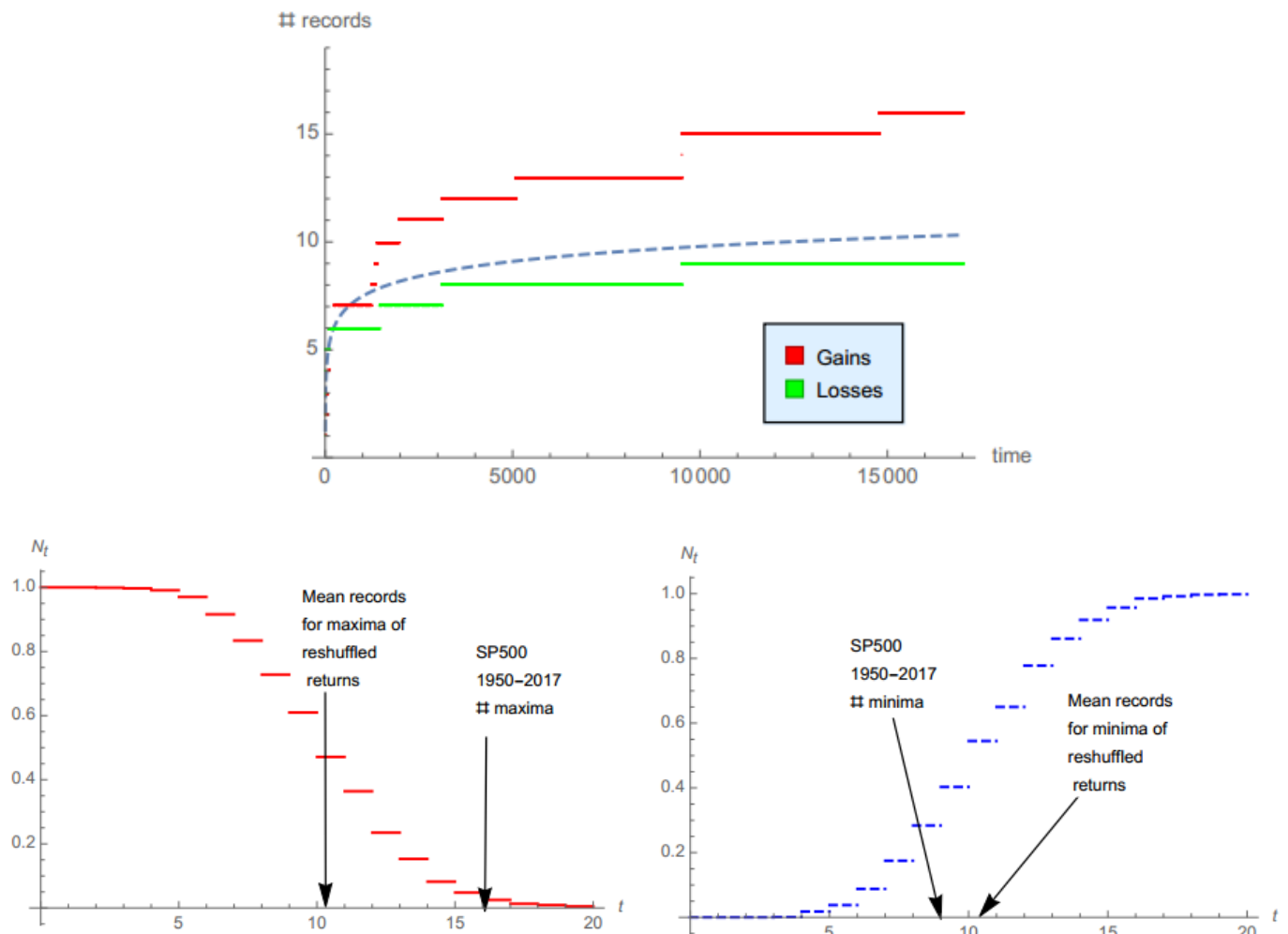
在极值分析中，Gumbel 极值统计方法如下：对于离散时间序列 X_1, X_2, \dots, X_n ，假设 $M_t = \max(X_1, X_2, \dots, X_t)$ ，定义极值数量：

$$N_{1,t} = 1 + \sum_{k=2}^t 1_{X_k > M_{k-1}}$$

也即一段时间内创出收益率极值的次数。该指标呈对数增长，开始时增长较快，随着分布样本量增加逐渐趋于稳定，越到后期需要越来越长的时间来打破历史收益率极值。不管 X 的分布如何，只要分布互相独立， $N_{1,t}$ 的期望都类似调和级数：

$$E(N_{1,t}) = H_t = \sum_{i=1}^t \frac{1}{i}$$

在整个历史中，SP500 的正收益极值突破了 16 次，负收益极值突破了 9 次。如果将收益率打乱重新进行极值分析，正负收益率突破次数的均值都在 10 次左右，此时代表肥尾之间无相关性。因此负收益率的尾部相对更加独立，而正收益率由于比理论值增长更快，说明尾部极值间存在一定的相关性。



预测与不确定性

肥尾分布下的决策

对于不确定条件下的决策，关键在于确定所面临概率与赔付的关系，其中赔付的重要性常常高于概率预测本身。在实际决策过程中，人们往往过度关注预测的正确与否，希望无限提高正确率，但是到头来却在赔付结果上吃了亏，形成决策上的巨大错配。本章中我们对预测和赔付之间的关系进行探讨。

首先我们要认识到：

- 人们的预测会包括各种各样的偏差，外界信息和心理作用都会极大影响预测能力，**如果正确预测是很容易的事情我们就不会看到金融市场中的贪婪和恐惧，自然也不会有周期性。**
- 成为预测大师并不一定能保证有好的表现，主要来自预测和赔付之间的非线性，**一个预测大师可以 99%的时间都正确，然后在 1%的时候赔得底儿掉。**

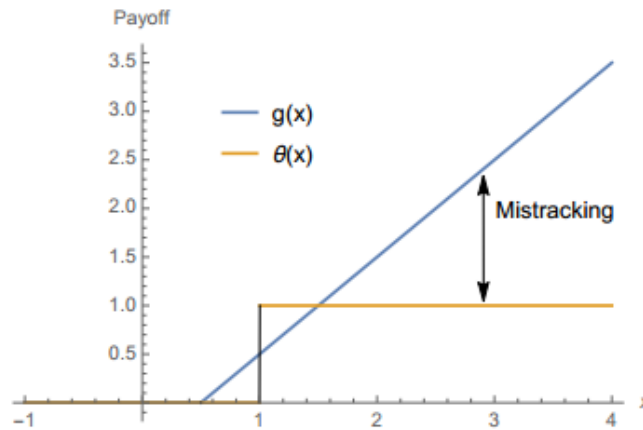
一个例子是老板问手下的交易员：你认为市场会上涨还是下跌？交易员信心满满的说会上涨，然后转头做空。老板非常生气，觉得受到了欺骗，因为他只能接受二元的状态：上涨做多，下跌做空，却无法理解大概率上涨对应“小幅上涨”而小概率下跌对应“大幅下跌”。交易员在这里锚定的是期望而不是预测。

从统计的角度讲，预测本身对应的是概率分布的零阶矩，而赔付往往是概率分布的一阶或高阶矩

可能出现的概率和赔付关系：

- **二元预测和赔付**，如赌博，彩票，选举结果，新药开发等，只有成功和失败两种结果。这时预测或者观点相当于一种投票机制，赔付和预测内在关联，如果两者脱钩则会出现无风险套利机会，如荷兰赌。
- **无边界赔付（包括无上界，无下界）**，如战争伤亡数字，市场崩盘损失，通胀程度，新产品销售和利润率，保险保障等。这样的非线性关系下会出现预测和结果方向相悖的现象，哪怕预测者只有小概率正确，由尾部带来的超大赔付可能依然划算，或者预测者绝大多数时候正确，但是可能会被尾部带来的超大损失击溃。

在金融衍生品中，上述赔付的典型例子分别为二元期权（0 或 1）和普通期权（单向无边界），如下图所示：



特征尺度

为了对冲尾部的风险，有人可能会问：在肥尾分布下，一个典型的或是常规的灾难赔付会有多大？实际上在无特征尺度的肥尾分布中，“典型”的尾部赔付很可能并不存在。为了区分随机变量的薄尾和肥尾，我们定义随机变量 X 分布的特征尺度如下：

$$\lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}(X|_{X>K}) = \lambda$$

- 如果 $\lambda = 1$ ， X 属于薄尾分布类 D_1 ，存在特征尺度
- 如果 $\lambda > 1$ ， X 属于肥尾分布类 D_2 ，不存在特征尺度
- 如果 $\lim_{K \rightarrow \infty} \mathbb{E}(X|_{X>K}) - K = \mu$ ， $\mu > 0$ ，则 X 属于边界指数分布

在肥尾分布中，由于尺度不存在，不论 K 多大都依然会有更大的条件期望，也即风险之上还有风险，想要完全对冲尾部的风险就只有通过无边界赔付的产品，如期权。

上述定义也可以用条件赔付的形式表示，假设 I 为 K 值以上的条件赔付， $g(x)$ 为赔付函数：

$$I_1 = \int_K^{\infty} g(x)f(x)dx$$

$$I_2 = g(K) \int_K^{\infty} f(x)dx$$

这时如果 X 属于薄尾分布类 D_1 ，我们有：

$$\lim_{K \rightarrow \infty} \frac{I_1}{I_2} = 1$$

如果 X 属于肥尾分布类 D_2 ，我们有：

$$\lim_{K \rightarrow \infty} \frac{I_1}{I_2} = \lambda > 1$$

这里的 I_2 代表截断于 K 处的赔付，由于样本值的原因，人们在进行计算时经常会错误的将 I_2 作为尾部的风险，但是实际上风险却在 I_1 的水平。比如某股票过去的 99%VAR 为 15%，投资者可能会简单地认为有 1% 的概率会赔 15%，实际上 1% 的真正风险由 CVAR 表示，在肥尾分布下会远远超出 15%。下表比较了正态分布和帕累托分布下两者的差别，可以看到正态分布（薄尾）下这种错误估计并不会带来太大的问题，条件风险越大误差越小，而帕累托分布下（肥尾）则会带来灾难，导致真实尾部风险被低估一个数量级。

Table 11.1: Gaussian pseudo-overestimation

| p | K_p | $\int_{K_p}^{\infty} x f(x) dx$ | $K_p \int_{K_p}^{\infty} f(x) dx$ | p^* | $\frac{p^*}{p}$ |
|-------------------|-------|---------------------------------|-----------------------------------|-----------------------|-----------------|
| $\frac{1}{10}$ | 1.28 | 1.75×10^{-1} | 1.28×10^{-1} | 1.36×10^{-1} | 1.36 |
| $\frac{1}{100}$ | 2.32 | 2.66×10^{-2} | 2.32×10^{-2} | 1.14×10^{-2} | 1.14 |
| $\frac{1}{1000}$ | 3.09 | 3.36×10^{-3} | 3.09×10^{-3} | 1.08×10^{-3} | 1.08 |
| $\frac{1}{10000}$ | 3.71 | 3.95×10^{-4} | 3.71×10^{-4} | 1.06×10^{-4} | 1.06 |

Table 11.2: Paretan pseudo-overestimation

| p | K_p | $\int_{K_p}^{\infty} x f(x) dx$ | $K_p \int_{K_p}^{\infty} f(x) dx$ | p^* | $\frac{p^*}{p}$ |
|-------------------|-------|---------------------------------|-----------------------------------|-----------|-----------------|
| $\frac{1}{10}$ | 8.1 | 8.92 | 0.811 | 1.1 (sic) | 11. |
| $\frac{1}{100}$ | 65.7 | 7.23 | 0.65 | 0.11 | 11. |
| $\frac{1}{1000}$ | 533 | 5.87 | 0.53 | 0.011 | 11. |
| $\frac{1}{10000}$ | 4328 | 4.76 | 0.43 | 0.0011 | 11. |

总的来说，概率只是积分内部的核函数，真实世界中重要的是赔付，也即概率事件对每个人的实际影响。**金融领域风险管理的本质在于改变赔付关系，而不在于追求正确预测，因为在肥尾分布下你很难进行“正确量级的预测”。**因此只要在赔付关系上有利于自身，哪怕降低预测精度也无妨。反过来说，预测准确率的提高如果对应赔付的大幅恶化，这样的准确并没有意义。如人们所说，同样是犯错误，把熊误认为是石头远远比把石头误认为是熊糟糕的多。

大选预测

上一节中我们谈到了二元预测，比如美国大选，总统在民主党和共和党候选人中二选一产生。在金融衍生品领域，二元期权的定价正是描述这样的过程。当选举不确定性大幅提升的时候，风险中性定价会将对应的期权价格推向 50%，并且越接近到期越趋向 50%。这一点和直觉相悖：当底层资产波动率提高的时候，期权的波动率反而降低了。通过借鉴二元期权的定价方法（鞅随机过程和无套利假设），我们可以更好的对大选结果进行建模预测，比如对于如下问题：

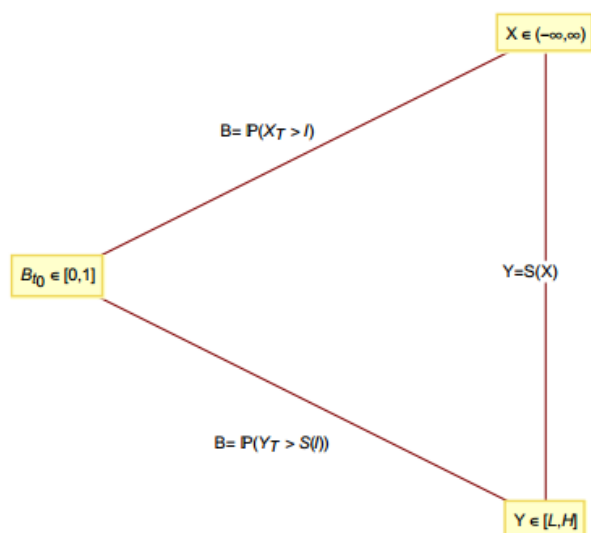
目前民调显示特朗普的支持率为 30%，请问特朗普获胜的概率是多少？

绝大多数人都会预测是 30%，但是其实 30%并不准确。因为 30%只考虑了最新的支持率，遗失了民调的波动率信息。事实上如果我们知道民调的波动率很大，民众忽而支持特朗普，忽而支持希拉里，这样的条件会降低现有民调结果的置信度，特朗普实际获胜的概率会高于 30%。或从另一个角度出发，如果民调显示特朗普支持率为 0%，因为潜在的不确定性，我们无法断定特朗普一定会落选，其获胜概率依然高于 0%。

为了定量求解获胜概率，我们需要对期权定价做一些改进，因为期权标的资产收益率为无界变量，而这里大选投票为有界量，因此我们加入代表投票票数的随机变量 Y ，将满足布朗随机游走的无界变量 X 映射到 Y 上，并使得 Y 为鞅过程。在非线性变换下，此时的 X 不满足鞅过程。

可能有人会问为什么不通过直接假设变量 Y 为有界 Beta 分布的形式来求解。原因在于数学上目前无法通过有界分布逆推随机过程。采用影子随机变量巧妙的解决了这一点，方便以期权定价求解获胜概率，并可以延展到不同的时刻观察获胜概率变化。

由此构建下图所示的结构，其中 B 为二元选举结果， Y 为投票数， X 为定义在 R 上的影子变量（方便计算用）



在风险中性条件下求解结果如下：

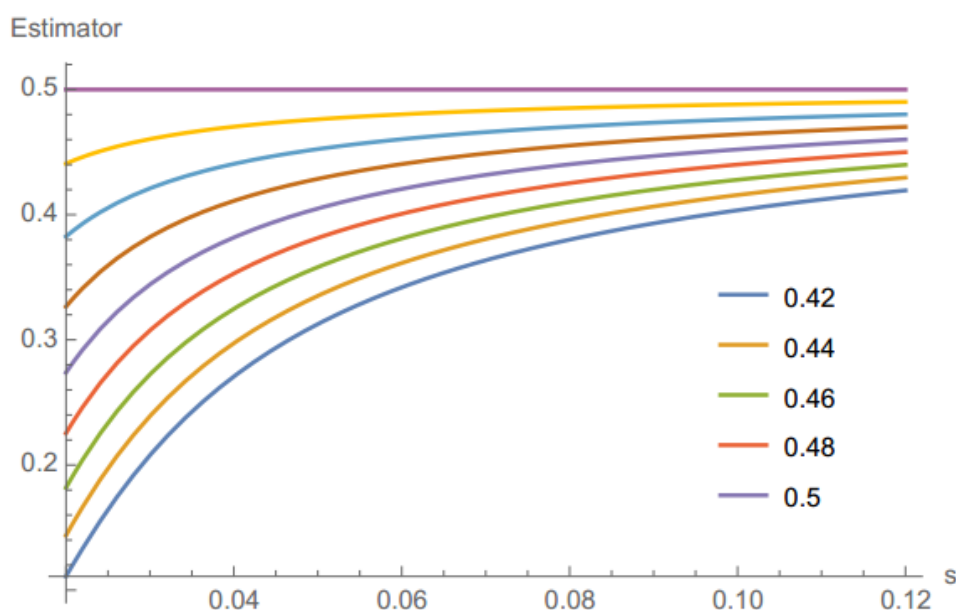
$$B(Y_0, \sigma, t_0, T) = \frac{1}{2} \operatorname{erfc} \left(\frac{l - \operatorname{erf}^{-1}(2Y_0 - 1)e^{\sigma^2(T-t_0)}}{\sqrt{e^{2\sigma^2(T-t_0)} - 1}} \right)$$

其中

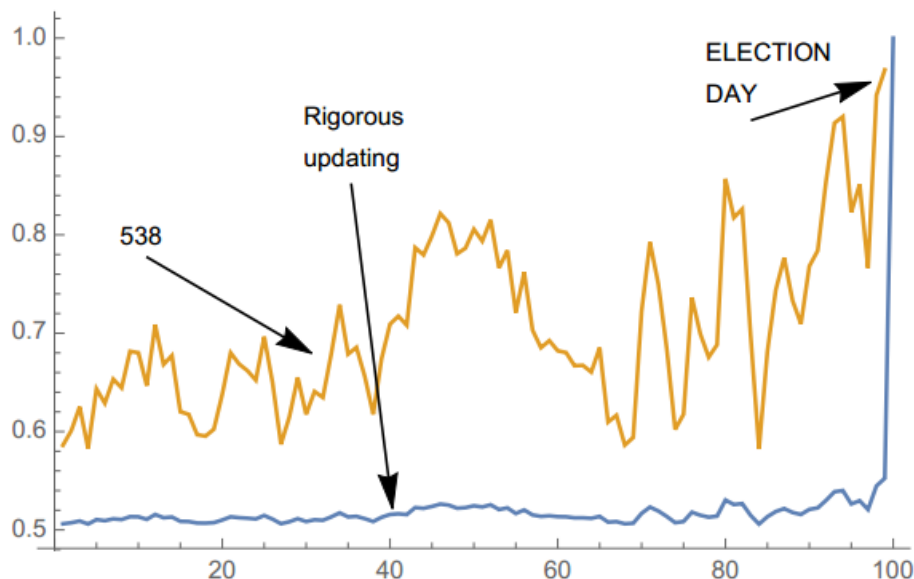
$$\sigma \approx \frac{\sqrt{\log(2\pi s^2 e^{2\operatorname{erf}^{-1}(2Y_0-1)^2} + 1)}}{\sqrt{2(T-t_0)}}$$

$$\operatorname{erfc}(z) = 1 - \operatorname{erf}(z), \quad \operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

Y_0 是初始投票值， T 为到期时间， t_0 是当前时间， s 是 Y 的年化波动率（此处假设波动率恒定）， B 为预测概率，也即无套利假设下的期权价格。 B 和 s 的敏感性如图所示，随着波动率增大，不同初始投票比率对应的潜在获胜概率都趋向于0.5。



有了这样的工具，我们可以通过不同时刻观测到的支持率推出真正的大选获胜概率（考虑获胜概率波动后），其无套利条件下的真实值会相对接近0.5，远远小于我们所看到的支持率波动。不过这里有一个很重要的假设是波动率保持不变，实际上随着大选的进行，有关候选人的信息逐渐披露，支持率的波动也将减小，真实世界的合理定价介于两者之间。



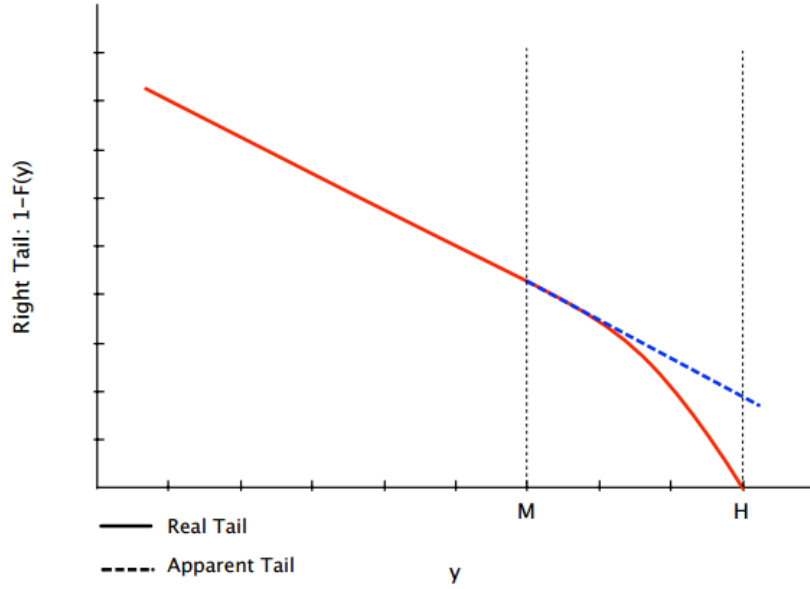
有界肥尾分布

有界帕累托分布

在一部分极度肥尾的数据中，如果光从数据拟合的角度出发，可能会得出一个 $\alpha < 1$ 的分布，也即该分布的均值无限。但是真实世界中的分布一定是有界的，比如战争的伤亡，其值不可能超出全球总人口数。虽然该上界远远大于可能出现的随机变量值，在观察分布时可以忽略，却给定量计算统计量带来了难度，因为有界分布的均值一定有限，而且我们依然希望通过极值理论定量计算条件均值，衡量尾部风险。类似的场景有：

- **操作风险：**公司操作风险的损失上界为全部市值，通过破产的方式截断
- **再保险保单：**再保险的保单存在很大的保额上限
- **战争伤亡：**上界为全球总人口数
- **信用风险：**一笔贷款的损失上界为全部资本金
- **城市规模：**城市人口的分布极度肥尾，上界也为全球（国家）总人口数
- **环境破坏：**破坏的面积上界为地球总面积
- **公司盈利规模：**一个公司的盈利上界为全球（国家）GDP

从数学的角度，假设随机变量 Y 的下界和上界为 $[L, H]$ ，其中 $L \gg 0$ ，一般的观察值 M 满足 $M \ll H < \infty$ 。因此仅仅通过样本拟合得到的尾部如下图虚线，斜率 $\alpha < 1$ 不变，而实际尾部由于上界的原因如实线所示。



在面对这样的情况时，为了拟合并求解分布均值，目前有两种方法：

- 1) 假设帕累托分布的尾部于 H 处截断，将超出部分的概率重新按比例加回到 [L, H] 区域
- 2) 假设 H 处为帕累托尾部的吸收态，超出 H 的概率通过狄拉克函数的形式加到 H 点

这两种方式在计算时由于概率密度的跃变，并不适用于极值理论。这里提出第三种方式——通过无界的中间分布实现概率密度的连续，并求解条件均值。假设随机变量 Z 满足：

$$Z = \varphi(Y) = L - H \log\left(\frac{H - Y}{H - L}\right)$$

$$Y = \varphi^{-1}(Z) = (L - H)e^{\frac{L-Z}{H}} + H$$

因此 Z 为连续分布，且满足 $\varphi^{-1}(L) = \varphi(L) = L$ ， $\varphi^{-1}(\infty) = H$ ，且如果 H 很大时 $Z \approx Y$ 。通过极值理论中 Pickands-Balkema-de Haan 定理，对于 Z 的概率分布函数 G 可以得到 Z 的尾部分布：

$$G_u(\omega) = P(Z - u \leq \omega | Z > u) = \frac{G(u + \omega) - G(u)}{1 - G(u)}$$

当 $u \rightarrow \infty$ 该分布 $G_u(\omega)$ 趋近于广义帕累托分布 $GPD(\omega; \xi, \sigma)$

$$GPD(\omega; \xi, \sigma) = \begin{cases} 1 - \left(1 + \frac{\xi \omega}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - e^{-\frac{\omega}{\sigma}} & \text{if } \xi = 0 \end{cases}$$

其中 ξ 为形状参数，对应 α^{-1} 控制肥尾程度，因此我们从 Z 出发推出随机变量 Y 的概率密度函数，假设 $u = L^* \geq L$ ，使得 $G_u(\omega) \approx GPD(\omega; \xi, \sigma)$

$$\begin{aligned} \int_{L^*}^{\infty} g(z) dz &= \int_{L^*}^{\varphi^{-1}(\infty)} f(y) dy \\ f(y; \alpha, \sigma) &= \frac{H \left(1 + \frac{H(\log(H - L) - \log(H - y))}{\alpha \sigma}\right)^{-\alpha-1}}{\sigma(H - y)} \end{aligned}$$

接下来即可计算超出 L^* 的条件均值：

$$\begin{aligned} e_u(Y) = E(Y - u | Y > u) &= \frac{\int_u^{\infty} (u - y) f(y; \alpha, \sigma) dy}{1 - F(u)} \\ &= (H - L) e^{\frac{\alpha \sigma}{H}} \left(\frac{\alpha \sigma}{H}\right)^{\alpha} \left(\frac{H \log\left(\frac{H - L}{H - u}\right)}{\alpha \sigma} + 1\right)^{\alpha} \Gamma\left(1 - \alpha, \frac{\alpha \sigma}{H} + \log\left(\frac{H - L}{H - u}\right)\right) \end{aligned}$$

由此整体的尾部条件均值可以写成：

$$E(Y|Y > u \geq L^*) = e_u(Y) + u$$

在风险管理领域，VaR 和 ES 预期损失可以通过上述公式结合起来，比如我们可以计算 Y 在 L^* 以上 95% 的预期损失，也即：

$$ES_{0.95} = e_{VaR_{0.95}^Y}(Y) + VaR_{0.95}^Y$$

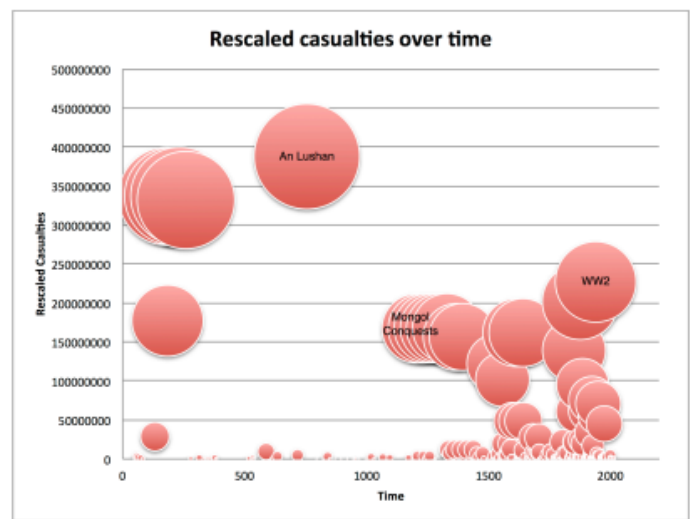
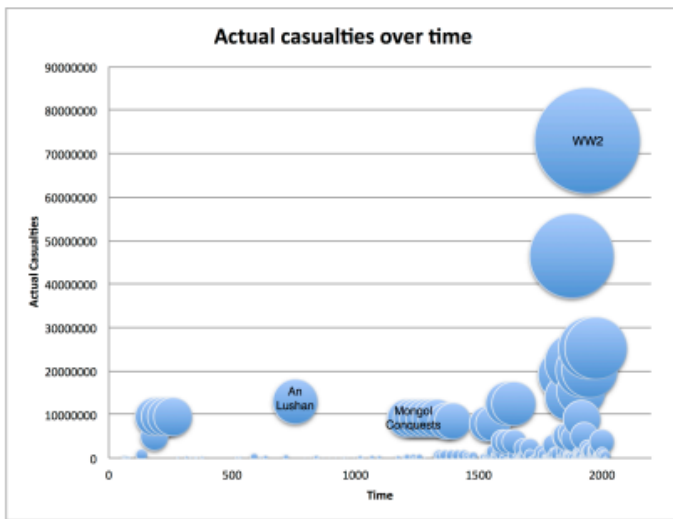
大规模战争和动乱的尾部概率

在人类历史上，战争和动乱是造成大量人员伤亡的主要因素。传统和战争相关的统计分析主要聚焦于优化不完善和不可靠的数据集。这里开辟一个新的视角，通过极值理论观察历史上战争造成人员死亡的分布，并按照有界帕累托分布，通过截断尾部的形式估计伤亡均值，战乱事件发生的周期和相关性。通过该研究尝试回答一个问题：**随着历史的发展，战乱的发生概率或伤亡规模是否有降低的趋势？**

在研究中，我们主要着眼于造成五万人以上死亡的战乱事件（按今天的人口规模估算，在 18 世纪约等于五千人），通过统计分析可以得到以下结论：

- 战乱的风险并没有降低，在伤亡方面以分布估算的均值高于历史均值，我们有可能低估了未来战争的危险性
- 战乱的周期符合无记忆性的到达模型，并没有随着历史发展降低频率

由于全球人口随着时间不断增长，为了保证可比性我们分别对原始数据和尺度重整数据做统计研究，其结论并无很大不同。在战乱事件的定义上，由于“战乱”概念本身比较模糊，不同类型的战乱可能在同一时期重叠，也可能一个战乱横跨几个时期，伴随着饥荒瘟疫等因素。我们这里将超过 25 年的战乱拆解开来，比如蒙古当年在亚欧大陆的入侵，持续了 125 年以上，在多个地方有着不同的记载，因此被拆解为了 12 到 55 个不同的事件。原始数据中最大的死亡人数出现在二战时期（7000-8000 万），而尺度调整后死亡人数最大的是中国唐朝的安史之乱（历史死亡人数 1000-2000 万）。



另外，对已发生战乱的分布进行统计存在生存偏差，我们得以生存的前提是尚未发生过毁灭整个人类的超大型战乱，随着人类掌握核武器等破坏性科技，这一极端风险也日益升高。

1) 伤亡满足帕累托分布，通过 log 变换后拟合的 $0.4 \leq \alpha \leq 0.7$ ，属于极度肥尾分布

通过广义帕累托分布我们可以拟合出满足 $GPD(\omega; \xi, \sigma)$ 的参数，尺度调整后的 $\xi = 1.88 \pm 0.14$ ，因此 $\alpha = \xi^{-1} = 0.53 \pm 0.04$

| Data | Nr. Excesses | ξ | β |
|-----------------|--------------|--------------------|---------------------|
| Raw Data | 307 | 1.5886 (0.1467) | 3.6254 (0.8191) |
| Naive Rescaling | 524 | 1.8718 (0.1259) | 14.3254 (2.1111) |
| Log-rescaling | 524 | 1.8717 (0.1277) | 14.3261 (2.1422) |

2) 战乱发生周期满足齐次泊松过程，并无记忆性，也即无法证明随着历史的发展全球已经变得更加和平
 检验齐次泊松过程可以通过 POT 方法，观察战乱到达时间是否满足指数分布，且发生时间之间不存在自相关性。
 下表为超过 1,2,5,10（百万）死亡人数的战乱的平均发生时间和平均偏差（基于原始数据）

| Threshold | Average | MAD |
|-----------|---------|--------|
| 1 | 26.71 | 31.66 |
| 2 | 42.19 | 47.31 |
| 5 | 57.74 | 68.60 |
| 10 | 101.58 | 144.47 |

下表为基于尺度重整数据的结果：

| Threshold | Average | MAD |
|-----------|---------|-------|
| 1 | 11.27 | 12.59 |
| 2 | 16.84 | 18.13 |
| 5 | 26.31 | 27.29 |
| 10 | 37.39 | 41.30 |
| 20 | 48.47 | 52.14 |
| 50 | 67.88 | 78.57 |

可以看到如果采用原始数据，一千万人以上死亡的大型战争平均周期为 101 年，而偏差可以达到 144 年，因此在二战以来短短不到 100 年的今天，我们并没有办法在统计上说明目前的世界已经进入长期和平的状态。

3) 通过分布拟合，如果未来发生大规模战争，虽然伤亡数字高度不确定，但拟合分布均值约为 3 倍历史均值，历史均值的视角大大低估了可能的伤亡

通过极大似然估计法可以对帕累托分布的条件均值进行更为准确的估计，由于 $\alpha < 1$ ，简单的套用帕累托分布会得到无穷均值。这里我们可以采用处理有界帕累托分布的方法，限制分布上界，再对尺度重整后的数据求解分布均值，并和样本均值进行比较。下表中可以看到极大似然估计下分布的均值是样本均值的三倍左右，且控制死亡人数条件对结论并无显著影响。

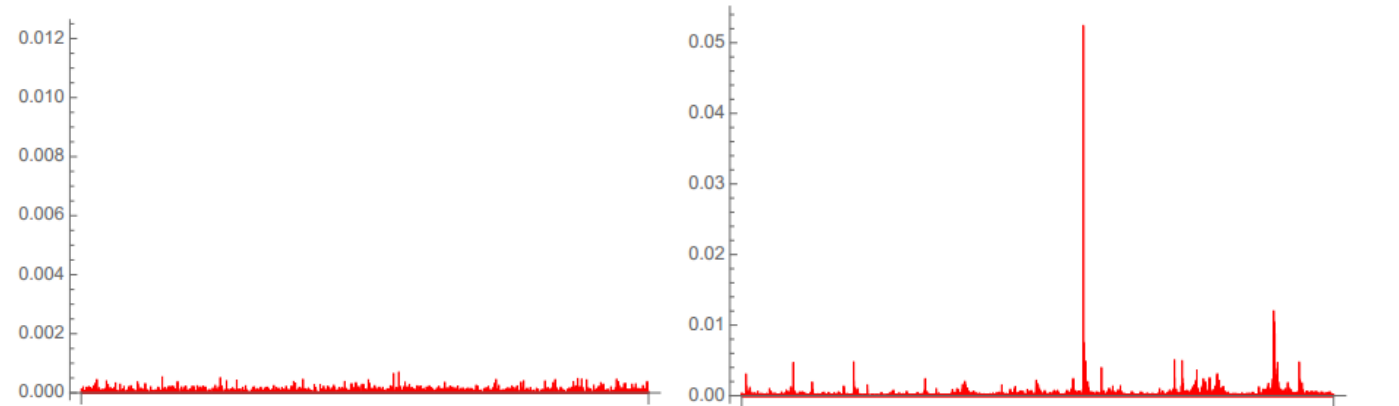
| L | Sample Mean | ML Mean | Ratio |
|------|---------------------|--------------------|-------|
| 10K | 9.079×10^6 | 3.11×10^7 | 3.43 |
| 25K | 9.82×10^6 | 3.62×10^7 | 3.69 |
| 50K | 1.12×10^7 | 4.11×10^7 | 3.67 |
| 100K | 1.34×10^7 | 4.74×10^7 | 3.53 |
| 200K | 1.66×10^7 | 6.31×10^7 | 3.79 |
| 500K | 2.48×10^7 | 8.26×10^7 | 3.31 |

大规模战争是高度不确定的尾部事件，基于上述数千年的历史统计研究，我们虽然无法预知第三次世界大战是否会发生，何时发生，或是会产生多少伤亡，但是至少从统计上可以认为其发生的风险并未降低。如果未来全球能继续延续百年以上的和平，才会在尺度上影响到上述统计结论，证明我们进入了一个更为安全的历史时期。

肥尾条件下的期权定价

期权是非线性的金融衍生产品，早在十七世纪的荷兰就已经出现，当时著名商人和哲学家 Joseph De La Vega 曾描述过为期权定价和管理风险敞口的技巧。之后于 1900 年的法国，Louis Bachelier 以数学模型确定了价格来自于最终赔付的期望，该方法并不限制标的资产的分布（甚至优于 BSM 模型）。后来更为人们所熟知的是 Black-Scholes-Merton 模型（BSM），使得连续时间下期权和标的资产的动态对冲组合满足风险中性条件，并加入了收益率正态分布的假设。期权定价模型的优点在于将不确定的价格变化转化为与资产价格关联的确定性赔付，因此期权不属于单独的资产类别，也不存在 CAPM 理论下的风险溢价。

但是在现实金融市场中，BSM 模型里面几个重要的潜在假设难以实现，比如假定无交易摩擦成本，无市场冲击影响，和最重要的无价格跳跃。虽然学界在后续的衍生定价模型上对于上述假设进行了改进，但还是不能很好的适用真实市场。因为只有连续交易才能满足时刻变化的 Delta 条件，而在收益率呈现出幂律尾分布的市场中，任何想要以动态对冲的方式复制期权收益流的行为都会面临高昂的误差成本。下图中左为 Black-Scholes 世界中的动态对冲误差，而右为真实的动态对冲误差。对于套利组合来说，出现在 1987 年市场崩盘期间的巨大误差很可能是灾难性的。因此在肥尾条件下动态对冲无法有效降低组合的风险。



虽然依赖动态对冲对期权定价的方式有重大缺陷，但是由于看涨看跌期权可以复制出远期合约，期权平价理论的适用范围可以超出 BSM 的假设，不论标的远期价格满足风险中性测度与否，看涨和看跌期权都至少存在一定的对应关系。对于任意的行权价 K 和测度 Q，我们都有：

$$\int_{\Omega} f_C(K)d\mu_Q - \int_{\Omega} f_P(K)d\mu_Q = \int_{\Omega} F_t d\mu_Q - K$$
$$f_C(K) = (S_t - K)^+, f_P(K) = (K - S_t)^+$$

如果对于不同的行权价 K 其平价定价出现偏差，就可以通过构建组合的模式进行多空套利，这种套利不需要面对 BSM 模型中的动态对冲误差问题，因此更为实用。

| | Black-Scholes-Merton 模型 | 看涨看跌期权平价 |
|--------|--|--|
| 套利类型 | 组合连续动态对冲 | 静态对冲 |
| 限制 | 时间维度大数定律 | 行权价维度大数定律 |
| 市场假设 | 1) 市场走势连续无跳跃 2) 所有时间标的资产买入卖空无限制 3) 无交易摩擦成本 | 1) 允许价格跳跃 2) 到期日标的资产无买入卖空限制 3) 期权交易低成本 |
| 概率分布 | 各阶矩有限 | 一阶矩有限 |
| 市场完备性 | 通过动态对冲组合满足完备条件 | 无要求 |
| 假设合理程度 | 低 | 高 |
| 收敛性 | 不确定，随着价格跳跃程度改变 | 表现稳定 |

从 BSM 模型的问题中我们可以看到，期权并不是一个可以很容易进行无风险复制的资产。因此对于期权交易员来说，并不需要通过模型对单一期权准确定价，因为没有人会针对他的报价进行结构性套利。期权交易更多是从标的和期权组合出发进行相对定价。一些常见的规则比如看涨看跌期权平价， $K + \Delta K$ 的看涨期权价格不能低于 K 的看涨期权， $K + 2\Delta K$ 的看涨期权价格不能超过两倍 $K + \Delta K$ 看涨期权，低利率下日历价差不能为负等等。在实际操作中，期权交易员更多的是响应市场供需并进行对冲，而非用尾部概率精确计算期权的价值。

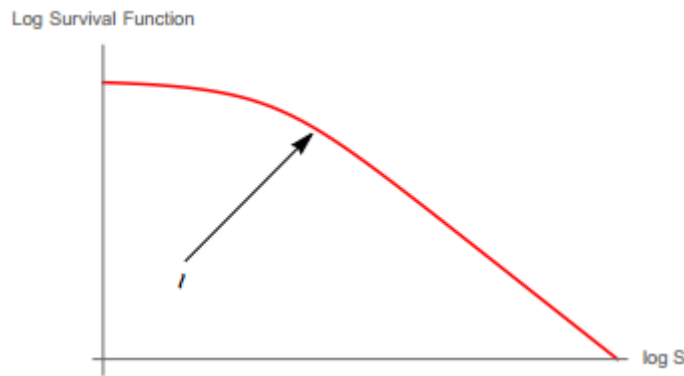
在相对定价方面，如果资产分布尾部满足幂律条件，不同行权价的期权之间会存在一定的合理价格关联。对于幂律分布的尾部来说：

$$P(X > x) \sim L(x) x^{-\alpha}$$

其中对于任意 $k > 0$ ， $L(x)$ 满足：

$$\lim_{x \rightarrow \infty} \frac{L(kx)}{L(x)} = 1$$

在 x 逐渐增大到一定程度的时候， $L(x)$ 将趋近于常数 l ，我们把这里的临界点 l 称为 Karamata 常数。而后再向尾部延伸其生存函数就在 $\log\text{-}\log$ 图中呈现出 $-\alpha$ 的下降斜率，如下图所示。



在测度 P 中，行权价为 K ，标的价格为 S 的看涨期权和看跌期权价格如下：

$$C(K) = E(S - K)^+ = \int_K^{\infty} (S - K) dP$$

$$P(K) = E(K - S)^+ = \int_0^K (K - S) dP$$

假定收益率 $\frac{S-S_0}{S_0}$ 满足幂律分布，我们利用 Karamata 常数 l 可以求得期权价格，并得到不同行权价期权间的相对关系，其最终相对关系仅和分布尾部 α 相关。

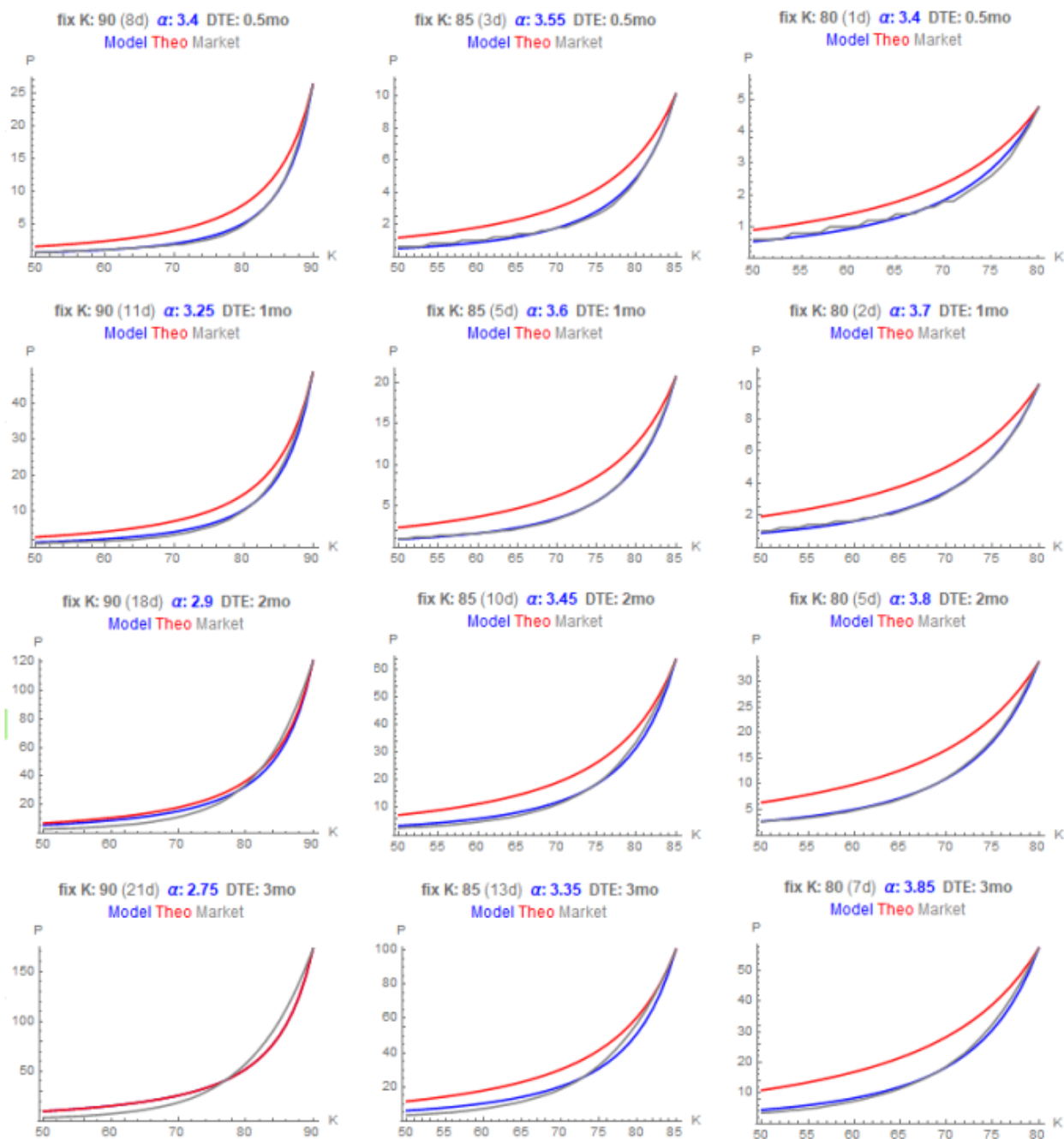
对于 $K_1, K_2 \geq (1 + l)S_0$ ，幂律分布下看涨期权的相对定价有：

$$C(K_2) = \left(\frac{K_2 - S_0}{K_1 - S_0} \right)^{1-\alpha} C(K_1)$$

对于 $K_1, K_2 \leq (1 - l)S_0$ ，幂律分布下看跌期权的相对定价有：

$$P(K_2) = \frac{(-1)^{1-\alpha} S_0^{-\alpha} ((\alpha - 1)K_2 + S_0) - (K_2 - S_0)^{1-\alpha}}{(-1)^{1-\alpha} S_0^{-\alpha} ((\alpha - 1)K_1 + S_0) - (K_1 - S_0)^{1-\alpha}} P(K_1)$$

在这样的相对关系下，我们可以结合市场报价观察期权之间的关系。下图以标普 500 的看跌期权为例，锚定一边的行权价为现价的 90%，85% 和 80%，并计算其余行权价的相对价格曲线。其中红线为市场尾部 $\alpha = 2.75$ 时的相对价格，可以看到此时市场给出价格对应的 α 在 3 到 4，低于长期历史数据显示的尾部风险，因此市场波动率的偏斜并未高估虚值期权价格，尾部保护的成本仍然合理。



对期权的一些认知误区

- **混淆肥尾程度（四阶矩或高阶矩）和波动率（二阶矩）**

很多交易者选用 VIX 指数来押注极端尾部事件，但是 VIX 本身是由平值期权求得的，更贴近波动（二阶矩）而非峰度（四阶矩）。正确押注肥尾的方式是卖出平值期权买入尾部的虚值期权，以二阶矩中性的方式单独做多四阶矩，押注波动率偏斜的增强。

- **计算期权回报时忽略了非线性**

对于押注风险事件的波动率策略，其回报和风险变动呈现出高度非线性关系，只看 VIX 来计算尾部期权的回报是错误的。假设 VIX 为 10% 时买入对应的期权，当 VIX 上涨 4% 时，对应的平值期权价格会上涨 15%，而一个 5 倍标准差的价外期权会上涨 4 倍，10 倍标准差的价外期权则会上涨 144 倍。

- **只看期权成本而忽略保护价值**

单独购买期权如同购买保险，期权的价值只有在和标的资产组合的时候才能进行客观衡量。一个合理的方式是通过计算持有标的+长期购买少量尾部期权的收益与风险。对冲基金 Universa 正是基于这一策略的长期正收益建立。