

稳定分布与广义中心极限定理

Jake @集智俱乐部 2011-6-16

一、真的是幂律吗？

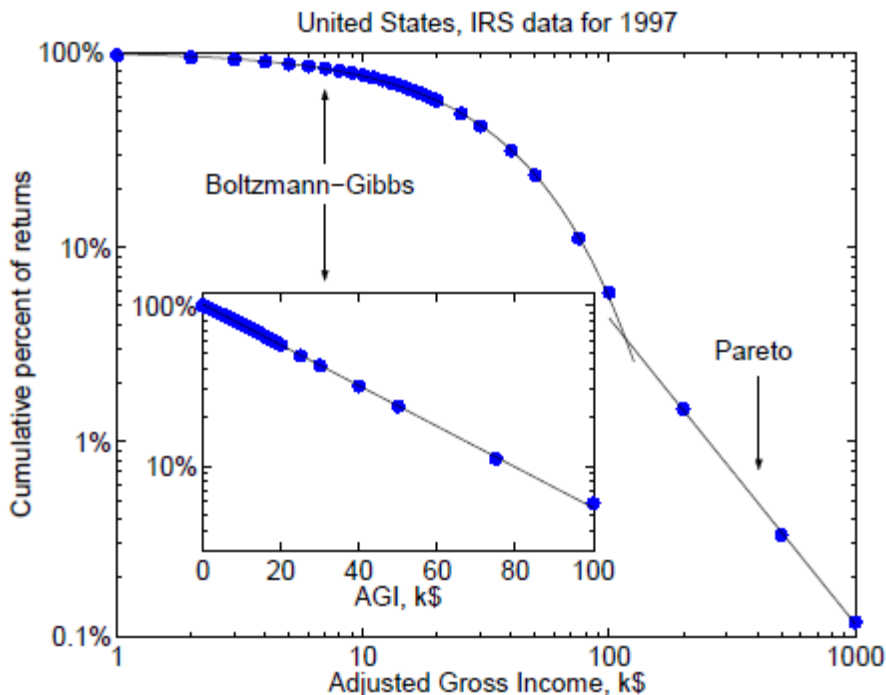
幂律分布是复杂系统中一种非常常见的现象，例如收入的分布、地震等级的分布、单词使用频率的分布等等，几乎横跨了从自然到社会的多种学科(关于幂律分布的实证分析，请参看 [Mark Newmann 的文章](#)^[1]，有关幂律分布成因的讨论，请参看[这里](#)^[2]。我们集智也曾对幂律分布进行过讨论，参见 [1](#)、[2](#))。

然而，幂律分布真的那么普遍吗？随着对大量数据的收集，我们发现，所谓的幂律仅仅适用于分布曲线的高端部分（也就是 x 很大的部分），这一点从分析幂律分布的数学表达式也能看出来。我们知道，如果 $f(x)$ 是幂律的概率密度函数，那么：

$$f(x) = cx^{-\alpha} \quad (1)$$

假如 x 表示的是社会上人们的收入，那么显然 x 有可能取 0 （社会上存在着收入为 0 的人）。但是，这在幂律分布公式的角度来看是不可能的，这是因为对于 $\alpha > 0$ ，即 x 上面的指数是一个负数，这样就意味着，当 x 趋近于 0 的时候， $f(x)$ 的值会很快趋近于无穷大。所以，在幂律分布中， x 必须大于等于一个常数 m ，这就意味着，幂律分布实际上只适用于尾端部分： $x > m$ 。那么人们自然要问，收入低端的分布是什么样的呢？难道，收入低于 m 值的概率是 0 ？低收入的人们都死绝了吗？

事实不是这样的，Yakovenko 等人^[3]用美国真实税收情况估计出的收入分布曲线如下图：



美国收入分布的累积概率图，图片来源[3]

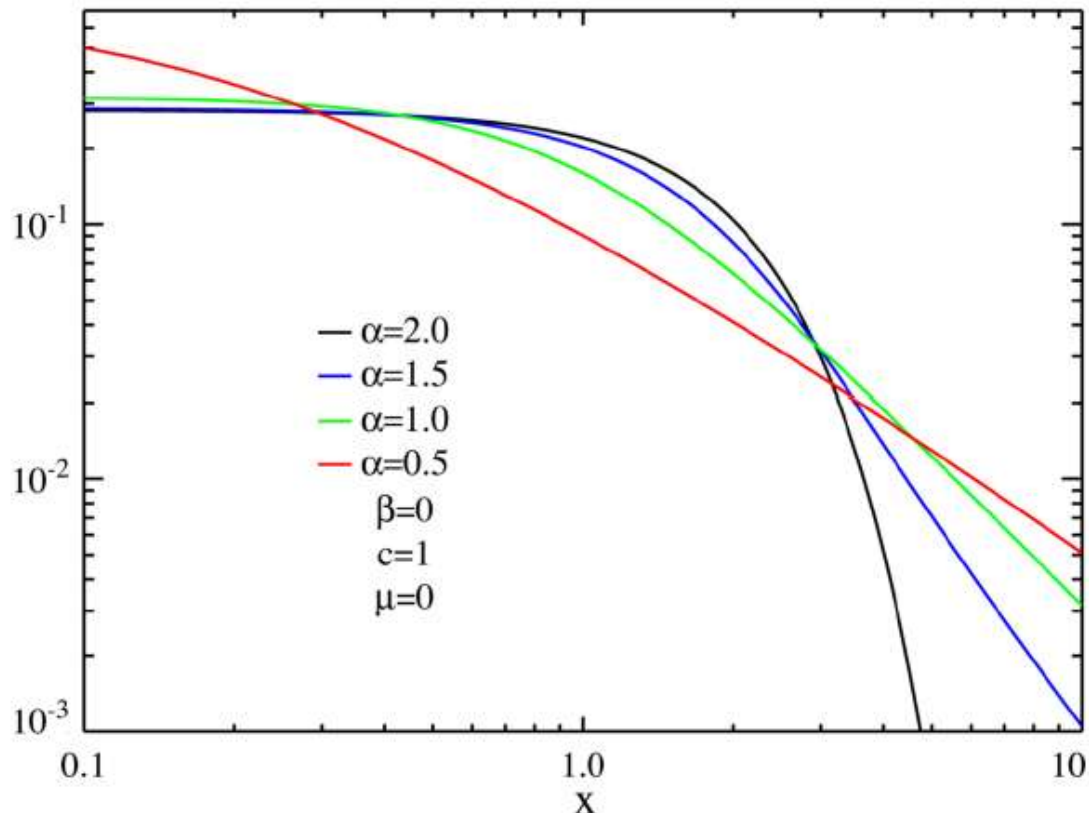
这是 1997 年 Yakovenko 等从美国税收数据推出的整个美国收入分布的曲线。其中，横坐标是收入的数据值，纵坐标是累积概率分布，也就是收入大于横坐标 x 的人口比例。我们看到，收入分布曲线分成了明显的两段，第一段是左边的那段弧形曲线部分，第二段是右下角的直线。由于图中的坐标是双对数的，所以直线段是幂律的，也就是作者标出的 Pareto 部分，而第一段是曲线的，因而不是幂律的，作者称这部分为 Boltzmann-Gibbs 分布，也就是指数分布。为了看清楚低端是指数分布，作者用小图表示出了在横坐标轴不取对数，纵坐标轴取对数的图，我们知道指数分布在这样的半对数坐标下就变成了直线，因而收入的低端应该是指数分布。

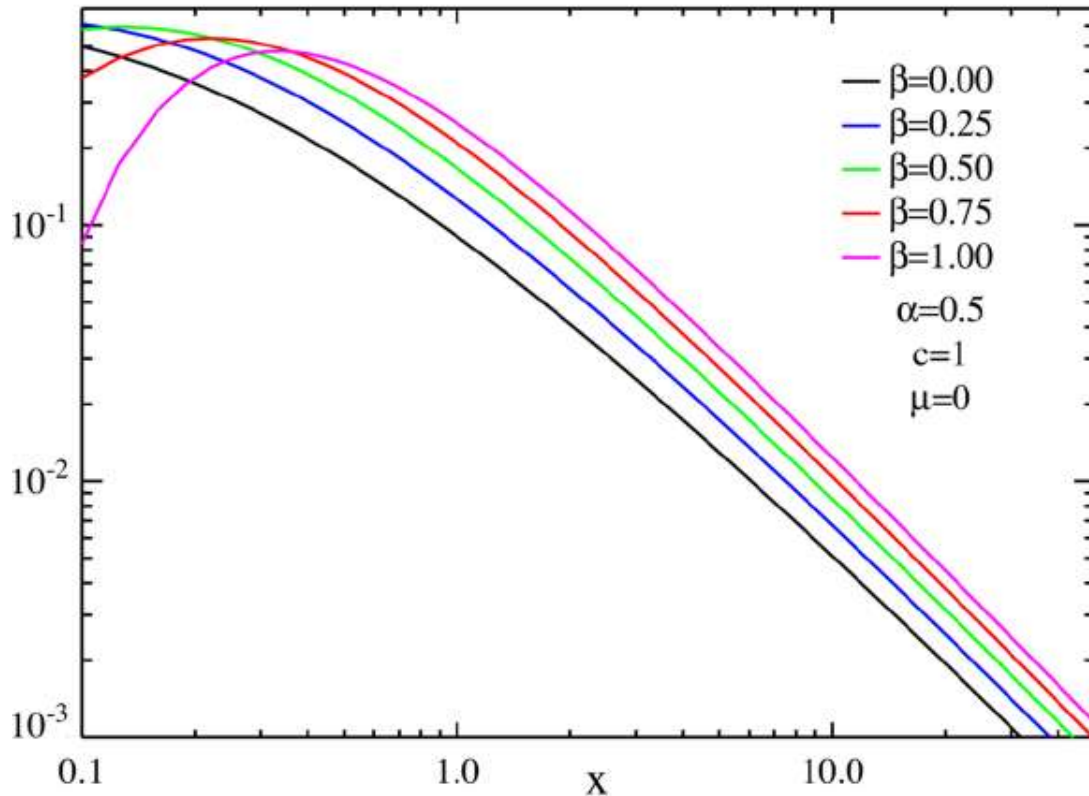
不仅仅是收入分布，很多以前被看作是幂律分布的系统，例如股票收益率的分布、复杂网络上节点的度的分布等等都存在着这种尾端趋近于幂律，头端则明显偏离幂律的现象。人们不禁要问，难道这些实证上的分布真的是幂律的，还是有其它分布？

二、稳定分布 (Stable Distribution)

当你与真正的统计物理学家或者数学家交流的时候，会发现，他们更多地提及一种称之为 [Levy Stable Distribution](#) 或者叫 α -Stable Distribution 的模型^[4]而非幂律分布。例如，大名鼎鼎的 Mandelbrot 在 1967 年的时候就曾经用这种 Stable 分布来拟合棉花期货价格的波动分布^[5]。经济物理学之父 H.E. Stanley 和 Mantegna 曾用结尾 Stable 分布来拟合股指波动的分布^[6]。

为什么这些数学、物理学家都更倾向于稳定分布而非幂律分布呢？原因是，这个稳定分布的概率密度函数恰恰具备尾端趋近于幂律分布，而在头端 ($x \rightarrow 0$) 偏离幂律，趋向于指数分布的性质。





图片来源：http://en.wikipedia.org/wiki/Stable_distribution

上两张图都是通过变换 Stable Distribution 的不同参数而得到的在双对数坐标下面的概率密度函数图。我们看到这些曲线都具备类似收入分布的性质：尾端呈现出幂律的形式（双对数下是直线）而左面则呈现出类似指数函数的性质。

你一定会产生这样的疑问：既然 Stable Distribution 这套分布函数能够那么好地逼近真实数据，那为什么还有那么多人使用幂律，而不是稳定分布呢？答案就在于 Stable Distribution 实际上不能给出密度函数的解析表达式，而仅仅能给出它的特征函数（也就是密度函数的 Fourier 变换）的表达式。因此，你要得到概率密度函数，你就要做 Fourier 逆变换，但是对于大多数的稳定分布来说，这个逆 Fourier 变换是不存在解析解的。上面的两张图，则是相应的数值解^[7]。

有关 Fourier 变换和特征函数，请参看下面的框：

我们知道，对于任意一个实数域上的函数 $f(x)$ 来说，只要它满足在整个实数域上面平方的积分有限，即：

$$\int_{-\infty}^{+\infty} f^2(x) dx < \infty$$

我们就可以对它进行 Fourier 变换，也就是一个全新的定义在实数域 s 上面的函数 $\tilde{f}(s)$ ：

$$\tilde{f}(s) = \int_{-\infty}^{+\infty} f(x) e^{-2\pi x i} dx$$

其中， i 为 -1 的平方根。反过来，如果已经知道了函数 $\tilde{f}(s)$ ，我们也可以通过求 Fourier 逆变换把它的原函数 $f(x)$ 求出来。可以证明：

$$f(x) = \int_{-\infty}^{+\infty} \tilde{f}(s) e^{2\pi x i} ds$$

其中， $\tilde{f}(s)$ 是 $f(x)$ 的 Fourier 变换，而 $f(x)$ 是 $\tilde{f}(s)$ 的 Fourier 逆变换。虽然对于任何的函数 $f(x)$ ，都可以定义它的 Fourier 变换及其逆变换，但是这并不意味着积分式都可以计算出来，有很多函数对应的 Fourier 变换的积分是写不出具体表达式的。

下面，我们来考虑一类特殊的函数，也就是说 $f(x)$ 表示一个概率分布的密度函数(density function)，那么，显然 $f(x)$ 满足平方积分有限的条件，于是，我们便可以对 $f(x)$ 做 Fourier 变换得到 $\tilde{f}(s)$ 。在概率论中， $\tilde{f}(s)$ 有个特殊的名字，它叫做 $f(x)$ 的特征函数。例如，如果 $f(x)$ 为标准正态分布，即：

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

它的特征函数就是：

$$\tilde{f}(s) = e^{-2\pi^2 s^2}$$

这个特征函数有一些很好的性质。例如， $f(x)$ 为概率密度函数，因此 $f(x)$ 在整个数轴上的积分就必须是 1，这就意味着其 Fourier 变换函数的 0 点值为 1，因为：

$$\tilde{f}(0) = \int_{-\infty}^{+\infty} f(x) e^{-2\pi \cdot 0 \cdot i} dx = \int_{-\infty}^{+\infty} f(x) dx = 1$$

另外， $\tilde{f}(s)$ 在 0 点的 n 阶导数正比于原随机变量 X 的 n 阶矩，这可以写为：

$$(-2\pi i)^n \tilde{f}^{(n)}(0) = \int_{-\infty}^{+\infty} x^n f(x) dx$$

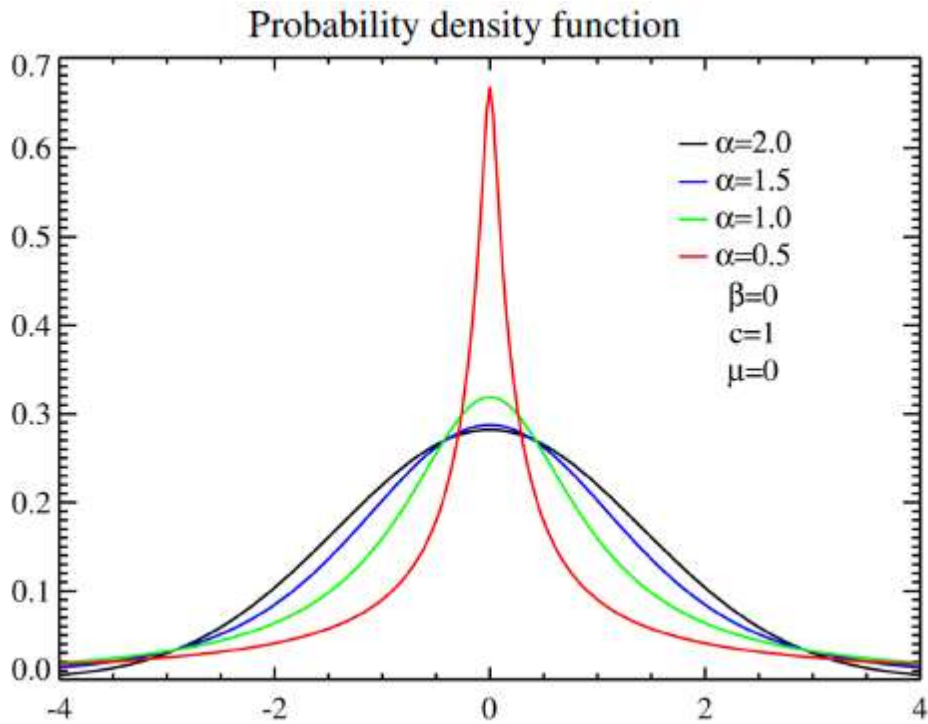
Stable Distribution 的特征函数的表达式如下^[4]：

$$\tilde{f}(s) = \begin{cases} \exp\left\{i\mu s - |cs|^\alpha \left[1 - i\beta \tan\left(\frac{\alpha\pi}{2}\right) \operatorname{sgn}(s)\right]\right\}, & \alpha \neq 1 \\ \exp\left\{i\mu s - |cs|^\alpha \left[1 - i\beta \frac{2}{\pi} \log(s) \operatorname{sgn}(s)\right]\right\}, & \alpha = 1 \end{cases} \quad (1)$$

注意，在这个函数中，自变量是 s ， i 是虚数单位， sgn 表示符号函数，也就是当 $s > 0$ 的时候为 +1， $s < 0$ 的时候为 -1， $s = 0$ 的时候为 0。 α, β, μ, c 都是参数，它们有取值范围：

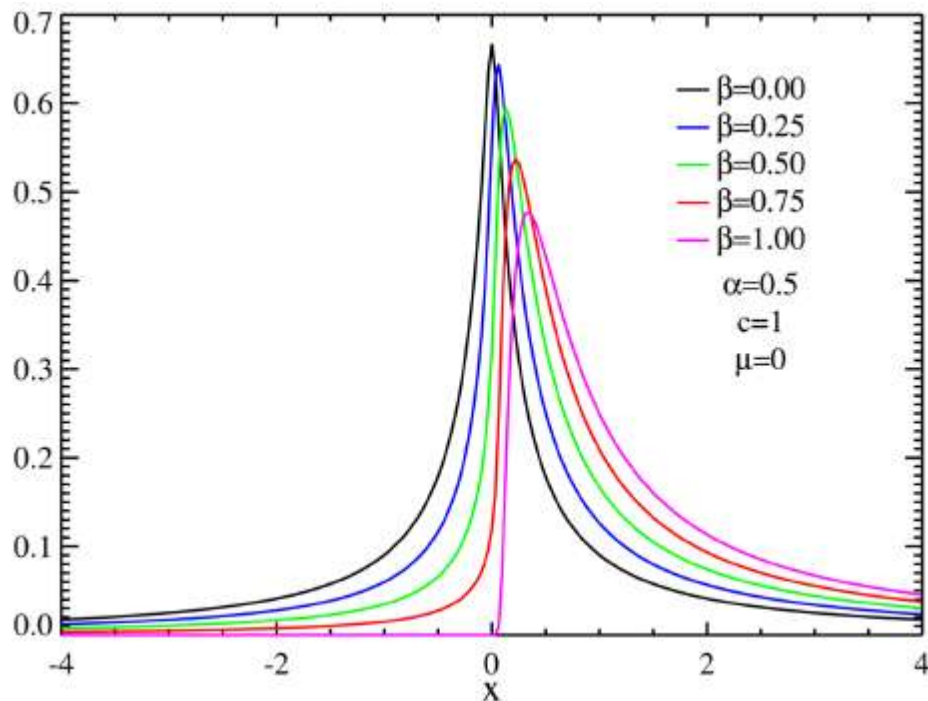
$$\alpha \in (0, 2], \beta \in [-1, 1], \mu \in (-\infty, +\infty), c \in (0, +\infty)$$

这些参数影响了密度曲线 $f(x)$ ($\tilde{f}(s)$ 的 Fourier 逆变换) 的形状和位置。其中 μ 影响了曲线中心的位置，是一个平移参数， c 则是曲线在水平方向上的缩放参数。而 α, β 则分别影响了概率密度曲线的陡峭程度和偏斜程度，如下两张图所示：



图片来源: http://en.wikipedia.org/wiki/Stable_distribution

当 $\beta = 0$ 的时候, 概率密度函数都是关于 $x=0$ 轴对称的(图 1、2 两图仅仅画出了 $x>0$ 的曲线)。我们看到这些曲线都呈现出了尖峰胖尾的形状(也就是中间有一个凸起, 而两侧的尾端非常扁平, 而且比正态分布 ($\alpha = 2$) 的黑色曲线要厚一些, 也就是胖尾的特性。有趣的是, 股票收益率分布曲线也是尖峰胖尾的^[6])



图片来源: http://en.wikipedia.org/wiki/Stable_distribution

这些曲线则是固定 α , 变换 β 得到的, 我们看到 β 越大, 则曲线越不对称, 所以, 我们称 β 为偏斜程度。可以证明稳定分布对应的概率密度函数在 x 很大的时候具有如下的渐进形

式^[4]:

$$f(x) \xrightarrow{x \rightarrow \infty} \mu^\alpha c(1 + \beta)x^{-1-\alpha}$$

也就是稳定分布的尾端趋近于幂律分布，并且幂指数就是 α -Stable 分布的指数 α 。

实际上，当我们让 α, β 取不同的特殊值的时候，我们能够得到常见的曲线。例如，如

果让 $\alpha = 2, \beta = 0$ ，则（1）式变为：

$$\tilde{f}(s) = \exp\{i\mu s - (cs)^2\}$$

这个函数的 Fourier 逆变换是可以求解出来的，它是：

$$f(x) = \frac{\sqrt{\pi}}{|c|} \exp\left\{-\frac{(x + \mu)^2}{4c^2}\right\}$$

这恰恰是一个均值为 $-\mu$ ，方差为 $|c|$ 的正态分布。所以，实际上正态分布仅仅是 Stable Distribution 的一个特例。

接下来，你一定会有这样的疑问：为什么一定要用那么怪怪的特征函数来表达这些稳定分布呢？这个怪异的数学式子（1）是怎么来的呢？下面，我会告诉你，原来这个稳定分布的来头是很大的：它恰恰来源于大名鼎鼎的中心极限定理。

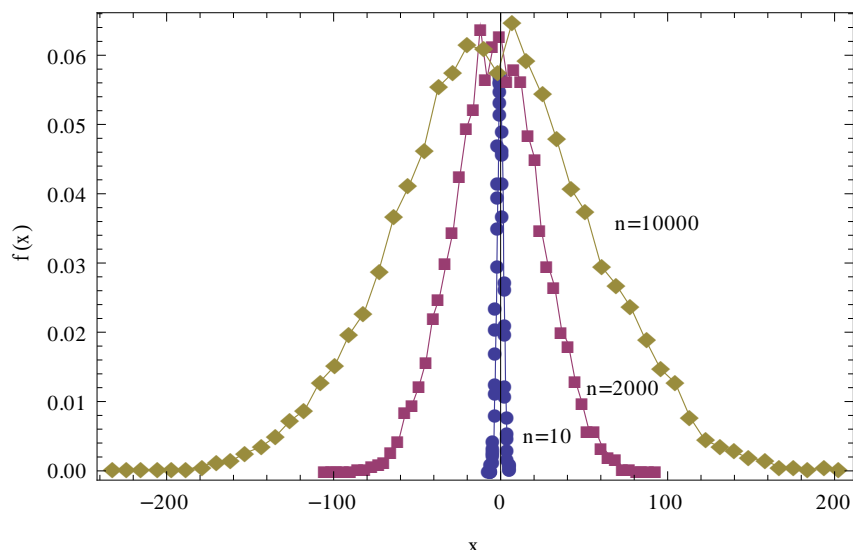
三、中心极限定理

学过概率论课程的人一定都还记得中心极限定理这个整个概率论中非常核心的数学定理。中心极限定理说：对于 n 个（ n 趋于无穷大）独立同分布的、方差有限的随机变量： X_1, X_2, \dots, X_n ，它们的和会趋向于一个稳定的分布：正态分布。具体来讲：

$$\sum_i^n X_i \rightarrow N(n\mu, \sqrt{n}\sigma) \quad (2)$$

上式表明， n 个独立同分布的随机变量和会趋近于一个正态分布，其中这个正态分布的均值是 $n\mu$ ，这里 μ 表示任意一个随机变量 X_i 的均值，方差的平方根是 $\sqrt{n}\sigma$ ，这里 σ 是任意一个随机变量 X_i 的方差的平方根。

下图用图形表示出了计算机模拟的 n 个独立同分布的二项分布的和在不同 n 值时候的分布图：



可以看到,随着 n 逐渐增大,分布曲线逐渐接近于钟型曲线,而且曲线变得越来越扁平,即方差越来越大。

这个定理是怎么推出来的呢?这就牵扯到了分布的特征函数,也就是原密度函数的 Fourier 变换。

首先,我们知道,如果两个独立的随机变量分别是 X 和 Y , 它们的密度函数分别是 $f(x)$ 和 $g(y)$, 那么这两个随机变量的和 $X+Y=Z$ 的概率密度函数是 $h(z)$, 我们知道:

$$\int_{-\infty}^z h(t)dt = \Pr\{Z < z\} = \Pr\{X + Y < z\} = \iint_{x+y < z} f(x)g(y)dxdy = \int_{-\infty}^z \int_{-\infty}^{z-x} f(x)g(y-x)dxdy$$

对上式得两边取对 z 的导数, 得到:

$$h(z) = \int_{-\infty}^{+\infty} f(x)g(z-x)dx = f \otimes g$$

也就是说 Z 的概率密度函数等于 f 和 g 这两个密度函数的卷积(请参看[计算士写的一篇非常有趣的科普文章](#))。如果对上式两侧同时作 Fourier 变换, 根据 Fourier 变换的性质^[7], f 和 g 卷积的 Fourier 变换等于 f 的福氏变换与 g 的福氏变换的乘积, 也就是:

$$\tilde{h} = \tilde{f} \cdot \tilde{g}$$

也就是两个独立的随机变量和的特征函数为这两个随机变量特征函数的乘积。

这个结论不难扩展到我们讨论的中心极限定理的情况, 那里有 n 个独立且同分布的随机变量的和, 所以最后的极限分布就是每一个随机变量特征函数的 n 次幂, 于是经过如下方框中的推理步骤, 我们不难证明中心极限定理。

中心极限定理的完整证明:

首先, 我们要先对每一个随机变量进行归一化, 也就是说, 定义随机变量 Y_i , 使得:

$$Y_i = \frac{X_i - \mu}{\sqrt{n\sigma}}$$

我们假设 x_i 的特征函数是 $\tilde{g}(s)$ 。根据特征函数的性质, 那么 Y_i 的特征函数就可以写成:

$$\tilde{h}(s) = e^{2\pi i \frac{\mu s}{\sqrt{n\sigma}}} \tilde{g}\left(\frac{s}{\sqrt{n\sigma}}\right) \stackrel{\text{记为}}{=} \tilde{f}\left(\frac{s}{\sqrt{n\sigma}}\right)$$

当 n 很大的时候 $\frac{s}{\sqrt{n\sigma}}$ 会趋近于 0，所以，我们可以把函数 $\tilde{f}\left(\frac{s}{\sqrt{n\sigma}}\right)$ 在 0 点附近做关于

$\frac{s}{\sqrt{n\sigma}}$ 的 Tylor 展开：

$$\tilde{f}\left(\frac{s}{\sqrt{n\sigma}}\right) = \tilde{f}(0) + \tilde{f}'(0) \frac{s}{\sqrt{n\sigma}} + \frac{1}{2} \tilde{f}''(0) \left(\frac{s}{\sqrt{n\sigma}}\right)^2 + o\left(\left(\frac{s}{\sqrt{n\sigma}}\right)^2\right)$$

我们知道：

$$\tilde{f}(0) = \tilde{g}(0) = \int_{-\infty}^{+\infty} g(y) e^{-2\pi i y s} \Big|_{s=0} dy = \int_{-\infty}^{+\infty} g(y) dy = 1$$

其中 $g(y)$ 是随机变量 Y_i 的概率密度函数，上面的式子=1 是概率密度函数的归一化条件。

$$\begin{aligned} \hat{f}'(0) &= \left(e^{2\pi i \mu s} \tilde{g}(s) \right) \Big|_{s=0} = 2\pi i \mu \tilde{g}(0) + \tilde{g}'(0) = 2\pi i \mu \int_{-\infty}^{+\infty} g(y) dy + (-2\pi i) \int_{-\infty}^{+\infty} y g(y) dy \\ &= 2\pi i \mu + (-2\pi i) \mu = 0 \end{aligned}$$

$$\begin{aligned} \hat{f}''(0) &= \left(e^{2\pi i \mu s} \tilde{g}(s) \right) \Big|_{s=0} = -4\mu^2 \pi^2 \tilde{g}(0) + 4i\mu\pi \tilde{g}'(0) + \tilde{g}''(0) = \\ &= 4\pi^2 \mu^2 - 4\pi^2 \int_{-\infty}^{+\infty} y^2 g(y) dy = -4\pi^2 \left(\int_{-\infty}^{+\infty} y^2 g(y) dy - \mu^2 \right) \\ &= -4\pi^2 \int_{-\infty}^{+\infty} (y - \mu)^2 g(y) dy = -4\pi^2 \sigma^2 \end{aligned}$$

其中用到了 Fourier 的特殊性质：Fourier 变换函数在 0 点的 n 阶导数正比于原函数的 n 阶矩。把这些代入展开式得：

$$\tilde{f}\left(\frac{s}{\sqrt{n\sigma}}\right) = 1 - 2\pi^2 \sigma^2 \left(\frac{s}{\sqrt{n\sigma}}\right)^2 + o\left(\left(\frac{s}{\sqrt{n\sigma}}\right)^2\right) = 1 - \frac{2(\pi\sigma)^2}{n} + o\left(\left(\frac{s}{\sqrt{n\sigma}}\right)^2\right)$$

于是：

$$\left(\tilde{f}\left(\frac{s}{\sqrt{n\sigma}}\right) \right)^n = \left(1 - \frac{2(\pi\sigma)^2}{n} + o\left(\left(\frac{s}{\sqrt{n\sigma}}\right)^2\right) \right)^n \approx \left(1 - \frac{2(\pi\sigma)^2}{n} \right)^n \xrightarrow{n \rightarrow \infty} \exp(-2\pi^2 s^2)$$

也就是说 n 个随机变量的和的特征函数为 $\exp(-2\pi^2 s^2)$ ，它恰好就是标准正态分布的特征函数，这一点只要通过对它进行 Fourier 逆变换就知道了：

$$\int_{-\infty}^{+\infty} \exp(-2\pi^2 s^2) \exp(-2\pi i s y) ds = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)$$

因此，我们有结论：

$$\sum_{i=1}^n Y_i = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \rightarrow N(0,1)$$

也就是：

$$\sum_{i=1}^n X_i \rightarrow N(n\mu, \sqrt{n}\sigma)$$

从而证明了中心极限定理（2）式。

注意，在经典的中心极限定理中，我们并没有对每个独立同分布的随机变量 X_i 的分布形式有任何要求，只要满足 X_i 的方差有限，就能保证这 n 个随机变量的和会趋近于一个正态分布。所以，正态分布是一种非常鲁棒的分布，它相当于是密度函数空间中的一个吸引域，无论你从什么分布开始，最终都会很容易地掉入正态分布这个吸引域。这也解释了为什么正态分布那么常见。

四、广义中心极限定理

虽然经典的中心极限定理非常强大，已被应用于很多实际问题中，但是这个定理存在着一个缺陷，这就是：定理要求每个独立的随机变量 X_i 都必须方差有限。但是，在很多实际的随机变量中，这并不会一定会被满足。例如，我们通常所说的幂律分布，当幂指数小于 2 的时候，该随机变量的方差就是发散的。这是因为它的二截距：

$$\int_m^{+\infty} x^{-1-\alpha} x^2 dx = \int_m^{+\infty} x^{1-\alpha} dx = \frac{1}{2-\alpha} x^{2-\alpha} \Big|_m^{+\infty}$$

可以看出，当 $\alpha < 2$ 的时候，上面的式子就是无穷大了。所以，针对这种幂律分布，我们就不能应用经典的中心极限定理了。

在 20 世纪初的时候，[P.P. Levy](#), [A.Kolmogrov](#) 等人对经典的中心极限定理进行了扩展，这被称为广义的中心极限定理，正是在这个扩展中，他们发现了 **Stable Distribution** 这种新型的概率分布形式。

所谓的广义中心极限定理^[8]，也就是说对于 n 个方差可以无限（当然也可以有限，从而涵盖经典中心极限定理的情形）的独立同分布的随机变量 Y_1, Y_2, \dots, Y_n ，它们的和会收敛到一个稳定分布（**stable distribution**） Y ，也就是：

$$\sum_{i=1}^n Y_i = Y_1 + Y_2 + \dots + Y_n \xrightarrow{n \rightarrow \infty} Y \quad (3)$$

其中 Y 就是一个稳定分布，我们只能定义它的特征函数满足（1）式，而在一般情况下不能给出 Y 的概率密度函数的具体形式。

五、稳定分布的简单推导

然而，到此为止，我们只是说极限分布是稳定分布，稳定分布的特征函数满足(1)式，但是仍然没有给出究竟(1)式是怎么得来的。下面，我们将展示如何得出（1）式的思路说明。

首先，我们可以对经典的中心极限定理（2）进行一个小的变形，实际上经典的中心极

限定理也可以表述成：对于 n 个方差有限的独立同分布的随机变量 X_1, X_2, \dots, X_n 的和，有：

$$\sum_{i=1}^n X_i = X_1 + X_2 + \dots + X_n \xrightarrow{n \rightarrow \infty} a_n X + b_n = (\sigma n)^{1/2} X + n\mu \quad (4)$$

其中， X 是一个服从标准正态分布 $N(0,1)$ 的随机变量。同样，也可以对 (3) 式进行这个变形：

$$\sum_{i=1}^n Y_i = Y_1 + Y_2 + \dots + Y_n \xrightarrow{n \rightarrow \infty} a_n Y + b_n = (cn)^{1/\alpha} Y + b_n \quad (5)$$

其中， Y 是一个 $\mu=0, c=1$ 的稳定分布（参见 (1) 式）。注意到 (4) 式与 (5) 式的最主要区别就是 (4) 中 a_n 正比于 n 的 $1/2$ 次幂，而在 (5) 式中 a_n 正比于 n 的 $1/\alpha$ 次幂。

进一步，因为 n 非常大，因此，我们总可以把上面的和式 (5) 分成两半，前面有 m 项，后面有 $n-m$ 项，也就是：

$$\sum_{i=1}^n Y_i = (Y_1 + Y_2 + \dots + Y_m) + (Y_{m+1} + Y_{m+2} + \dots + Y_n)$$

由于 n 趋近于无穷大，因而 m 也是无穷大， $n-m$ 也是无穷大，这样：

$$\sum_{i=1}^m Y_i = Y_1 + Y_2 + \dots + Y_m \xrightarrow{m \rightarrow \infty} (cm)^{1/\alpha} Y + b_m,$$

以及：

$$\sum_{i=1}^{n-m} Y_{m+i} = Y_{m+1} + Y_{m+2} + \dots + Y_n \xrightarrow{m-n \rightarrow \infty} (c(m-n))^{1/\alpha} Y + b_{n-m}$$

这样，原始的和式就可以表述成：

$$\sum_{i=1}^n Y_i \xrightarrow{n \rightarrow \infty} (cm)^{1/\alpha} Y + b_m + (c(n-m))^{1/\alpha} Y + b_{n-m} \quad (6)$$

比较 (5) 和 (6) 式，也就是说对于稳定分布的随机变量 Y 它应该满足下式：

$$(cm)^{1/\alpha} Y + (cn)^{1/\alpha} Y = (c(m+n))^{1/\alpha} Y + b \quad (7)$$

其中， $b = b_{m+n} - b_m - b_n$ ，我们将 (6) 中的 $n-m$ 换成了 n ，而把 n 换成了 $n+m$ ，就得到了 (7) 式。事实上，(7) 式恰恰就是稳定分布随机变量的定义。也就是说，如果随机变量 Y 是稳定分布的，那么它就应该满足 (7) 式，反过来任何满足 (7) 式的随机变量 Y 都是稳定分布的。

(7) 式表达的是随机变量的关系，其中 $=$ 号表示的是方程两边的随机变量具有相同的分布。如果我们假设 Y 的概率密度函数为 $f(y)$ ，那么将 (7) 式改写成关于概率密度函数 $f(y)$ 的表达式：

$$\int_{-\infty}^{+\infty} \frac{f\left(\frac{y}{(cm)^{1/\alpha}}\right)}{(cm)^{1/\alpha}} \frac{f\left(\frac{z-y}{(cn)^{1/\alpha}}\right)}{(cn)^{1/\alpha}} dy = \frac{f\left(\frac{z-b}{(c(m+n))^{1/\alpha}}\right)}{(c(m+n))^{1/\alpha}}$$

我们不妨对上式进行 Fourier 变换，这样左边的卷积就可以变成函数相乘：

$$\tilde{f}((cm)^{1/\alpha} s) \tilde{f}((cn)^{1/\alpha} s) = \tilde{f}((c(m+n))^{1/\alpha} s) e^{2\pi i b s}$$

对上式两边取 log，乘积就变成了求和：

$$\log \tilde{f}((cm)^{1/\alpha} s) + \log \tilde{f}((cn)^{1/\alpha} s) = \log \tilde{f}((c(m+n))^{1/\alpha} s) + 2\pi i b s$$

我们不妨把函数 $\log \tilde{f}$ 简记成 g ，于是就得到：

$$g((cm)^{1/\alpha} s) + g((cn)^{1/\alpha} s) = g((c(m+n))^{1/\alpha} s) + 2\pi i b s \quad (8)$$

我们知道 Y 遵从稳定分布的充分必要条件就是使得方程(8)成立，因此只需要求解 (8) 式所代表的函数方程就可以求解出稳定分布的具体表达式。观察 (8) 式发现，函数 g 是一个未知的函数，只要得到了 g 的一般形式，也就知道了特征函数 \tilde{f} 。通过下面框中的计算，可以证明，满足 (8) 式的 g 的一般形式如下：

$$g(s) = \begin{cases} i\mu s - |cs|^\alpha \left[1 - i\beta \tan\left(\frac{\alpha\pi}{2}\right) \text{sgn}(s) \right], & \alpha \neq 1 \\ i\mu s - |cs|^\alpha \left[1 - i\beta \frac{2}{\pi} \log(s) \text{sgn}(s) \right], & \alpha = 1 \end{cases}$$

因此，将其代入 $\tilde{f}(s) = \exp(g(s))$ 就能得到了 (1) 式。

(1) 式的导出 (参照[9])

首先，为了简化，我们不妨让 $b=0, c=1$ ，这样 (8) 就变成了简化的方程：

$$g(m^{1/\alpha} s) + g(n^{1/\alpha} s) = g((m+n)^{1/\alpha} s) \quad (8')$$

注意，在(8')中， m 和 n 都可以是任意的自然数。这样我们实际上可以让 $m=n=1$ ，就得到了：

$$2g(s) = g(2^{1/\alpha} s)$$

再让 $m=2, n=1$ ，得到：

$$g(2^{1/\alpha} s) + g(s) = 3g(s) = g(3^{1/\alpha} s)$$

所以，不断运用上面的过程，不难验证对于任意的自然数，甚至是任意正实数 p 都有：

$$pg(s) = g(p^{1/\alpha} s)$$

我们令 $t = p^{1/\alpha}, s = 1$ ，就有：

$$g(t) = t^\alpha g(1)$$

其中 $g(1)$ 是一个常数， $t > 0$ 。由于 $g(t)$ 这个函数是 $f(x)$ 这个函数的 Fourier 变换之后取对数得到的，所以，一般情况下 $g(t)$ 是一个复数，因而 $g(1)$ 也是一个复数。

又因为特征函数：

$$|\tilde{f}(s)| = \left| \int_{-\infty}^{+\infty} f(x) e^{-2\pi i x s} dx \right| \leq \int_{-\infty}^{+\infty} |f(x)| \cdot |\exp(-2\pi i x s)| dx = \int_{-\infty}^{+\infty} f(x) \cdot dx = 1$$

于是：

$$g(s) = \log \tilde{f}(s) = \log(re^{i\theta}) = \log(r(s)) + i\theta(s)$$

这里 $r(s), \theta(s)$ 分别是 $\tilde{f}(s)$ 的模和相角，而我们知道 $r(s) \leq 1$ ，所以：

$$\operatorname{Re}(g(s)) \leq 0$$

也就是说 $g(s)$ 的实部是小于 0 的，这样函数 $g(t)$ 就可以写成：

$$g(t) = -|t|^\alpha (c_1 - ic_2)$$

其中 c_1 为大于 0 的实数， c_2 为任意实数。这里 t 加上了绝对值，于是 t 就可以取任意的实数了。进一步，把 c_1 提取出来这个式子又可以改写成：

$$g(t) = -c_1 |t|^\alpha (1 - ic_2/c_1)$$

首先， c_1 是正数，于是它可以融合进 $|t|^\alpha$ 写成： $|c_0 t|^\alpha$ ，其中 $c_0 = c_1^{1/\alpha}$ 。而第二部分：

c_2/c_1 可以写成： $\beta \tan(\alpha\pi/2) \operatorname{sgn}(t)$ 当 $\alpha \neq 1$ ，其中 $\beta \in [-1, 1]$ 为引入的任意参数。我们知道当 $\tan(\alpha\pi/2)$ 的取值为任意实数，所以， c_2/c_1 也可以取遍所有实数。对于 $\alpha = 1$ 的情况也可以进行类似地处理，这样， $g(t)$ 就可以写成：

$$g(t) = \begin{cases} -|c_0 t|^\alpha \left[1 - i\beta \tan\left(\frac{\alpha\pi}{2}\right) \operatorname{sgn}(t) \right], & \alpha \neq 1 \\ -|c_0 t| \left[1 + i\beta \frac{2}{\pi} \log(|t|) \operatorname{sgn}(t) \right], & \alpha = 1 \end{cases}$$

下面，我们再考虑 c 不为 1， b 不为 0 的一般情况。这种情况相当于把 $g(t)$ 函数变成了 $g(t/c) + 2\pi i \mu s$ ，其中缩放因子 c 可以吸收到 c_0 中，而 μ 相当于密度函数 $f(x)$ 的平移参数。所以，一般情况下的 $g(t)$ 表达式就是：

$$g(t) = \begin{cases} i\mu t - |c_0 t|^\alpha \left[1 - i\beta \tan\left(\frac{\alpha\pi}{2}\right) \operatorname{sgn}(t) \right], & \alpha \neq 1 \\ i\mu t - |c_0 t| \left[1 + i\beta \frac{2}{\pi} \log(|t|) \operatorname{sgn}(t) \right], & \alpha = 1 \end{cases}$$

从而，原始的 $\tilde{f}(s)$ 就具有一般的函数形式：

$$\tilde{f}(s) = \begin{cases} \exp\left\{ i\mu s - |cs|^\alpha \left[1 - i\beta \tan\left(\frac{\alpha\pi}{2}\right) \operatorname{sgn}(s) \right] \right\}, & \alpha \neq 1 \\ \exp\left\{ i\mu s - |cs| \left[1 + i\beta \frac{2}{\pi} \log(|s|) \operatorname{sgn}(s) \right] \right\}, & \alpha = 1 \end{cases}$$

这就得到了稳定分布的特征函数表达式。

下面我们对参数 α 的取值范围进行一定的讨论：

首先，我们知道：

$$g''(t) = -(c_1 - ic_2)\alpha(\alpha - 1) |t|^{\alpha-2}$$

而：

$$g''(0) = (\log \hat{f}(0))'' = \frac{\hat{f}''(t)\hat{f}'(t) - (\hat{f}'(t))^2}{\hat{f}^2(t)} \Big|_{t=0} = \sigma^2$$

也就是原函数的方差。当 $\alpha = 2$ 的时候，我们知道方差是一个常数，这恰恰对应了正态分布的情形。而当 $\alpha < 2$ 并且 t 趋近于 0 的时候，方差是无穷大的。如果 $\alpha > 2$ ，那么当 t 趋近于 0 的时候，方差为 0，也就是：

$$\sigma = \int_{-\infty}^{+\infty} (x - E(x))^2 f(x) dx = 0$$

这个等式只有当 $f(x)$ 处处为 0 的时候才能成立，而此时 $f(x)$ 已经不再是一个概率密度函数了。另外，当 $\alpha \leq 0$ ，而且 t 趋近于 0 的时候， $g(0)$ 趋于无穷大，因此 $\tilde{f}(0) \rightarrow \infty$ ，但是这与 $\tilde{f}(0) = 1$ 的要求矛盾，所以 α 必须 > 0 。因而， α 必须处于区间 $(0, 2]$ 。

当然，这里的推导仅仅是一种示意性的，并不严格，严格的证明请参照这本书：[\[8\]](#)。

六、总结

本文指出了有可能自然界普遍存在的幂律分布仅仅适用于很多实际密度函数的尾端，而在头端，则会出现类似指数函数的形式。这种现象使得我们怀疑幂律分布函数有可能并不是普遍形式。

数学家 Levy 早在 1920 年左右提出来的 Levy α -Stable 分布则具有类似于现实实证分布的前端指数，后端幂律的形式。但是，很可惜的是，稳定分布仅仅能给出其特征函数的解析表达式，而不能给出一般的概率密度表达式。

进一步，这种稳定分布实际上是中心极限定理的推广，对于 n 多个方差可能无限的独立同分布变量来说，它们就会收敛到稳定分布。而且，从无穷多个独立同分布的随机变量和的形式可以推出稳定分布的定义式。

正态分布之所以非常普遍是因为它是一切独立同分布变量和的极限分布，因此正态分布对于方差有限的随机变量形成了一个吸引子。同样的道理，之所以在各类复杂系统中，我们能够观察到如此众多的尾端是幂律形式的分布，也是因为稳定分布构成了更一般的情况：方差可能无限的随机变量和的吸引域。也就是说，恰恰是因为稳定分布的鲁棒性，才导致了我们在实际系统中很容易看到具有幂律尾特征的分布形式。

因此，可以说为什么幂律分布如此普遍这个问题早在已经被数学家 Paul Levy 于 1920 年代就解决了！

参考文献

- [1] M. E. J. Newman: Power laws, Pareto distributions and Zipf's law, arXiv:cond-mat/0412004
- [2] M. Mitzenmacher: A Brief History of Generative Models for Power Law and Lognormal Distributions, Internet Mathematics Vol. 1, No. 2: 226-251
- [3] A. Drăgulescu, V. M. Yakovenko: Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States, cond-mat/0103544, v.2 March 28, 2001
- [4] Nolan, John P. (2009). "[Stable Distributions: Models for Heavy Tailed Data](http://academic2.american.edu/~jpnolan/stable/chap1.pdf)" (PDF). <http://academic2.american.edu/~jpnolan/stable/chap1.pdf>.
- [5] Mandelbrot, B. (1967). The variation of some other speculative prices. Journal of Business 40, 393{413.
- [6] R.N. Magtegn, H.E. Stanley: [An Introduction to Econophysics: Correlations and Complexity in Finance](#), Cambridge University Press, 1999
- [7] Ronald Bracewell: The Fourier Transform and Its Applications, McGraw-Hill Science, 1999
- [8] B.V. Gnedenko, A.N., Kolmogorov: [Limit distributions for sums of independent random variables](#), Addison-Wesley, 1954
- [9] A. A. Dubkov, B. Spagnolo, V. V. Uchaikin: Levy Flight Superdiffusion: An Introduction, arXiv:0810.1492, 2008