王立新《关于李德毅院士抄袭本人研究成果的情况说明》

一、时间表：

1. 2010 年 5 月 19 日：本人应李德毅院士的邀请，参加在北京京通宾馆召开的“不确定表示的数学基础研讨会”，会上第一次听到“云模型”（早年由李院士提出）。由于会期紧，没有学习“云模型”的细节。
2. 会后至 2011 年 7 月：参加由李院士主导、主要由参加研讨会的成员组成的网上讨论。
3. 2010 年 11 月 8 日下午 15 时 33 分：收到由李院士的学生发来的“云模型”的算法细节，开始对其进行研究。
4. 2010 年 11 月 9 日下午 20 时 51 分：将一天来的研究心得写成“高斯云的定义与数学特性”一文（见下面附件 1），发给全体讨论组成员。
5. 2011 年 4 月：李德毅院士作为第一作者，在《中国工程科学》（2011, 13(4): 20-23）发表文章“正态云模型的重尾性质证明”（见下面附件 2），其核心结果（定理 1 的证明）完全抄袭本人“高斯云的定义与数学特性”一文中的定理 4（请读者比较下面附件 1 定理 4 的证明与附件 2 定理 1 的证明）。

二、想要说的话：

1. 将研究成果首先发给同行学者，这在国际学术界是通行的做法。收到成果的人将成果作为自己的成果去发表，只字不提原创者，这是极其严重的抄袭行为。
2. 李德毅院士是国内信息科学界德高望重的学者，李院士是中国人工智能学会理事长，国家自然科学基金委员会信息科学部主任，软件工程国家重点实验室学术委员会主任，973 首席科学家，……。
3. 本人于 92 年获美国南加州大学电机工程系博士学位，93 至 07 年任教于香港科技大学电机与电子工程系。07 年从工作了 14 年、具有终身合同（到 65 岁）的香港科大辞职，移居北京，大隐于繁华，以“独立研究者”之身份向科学的高峰攀登。本人的研究成果在国际学术界的影响可上 google scholar 或 SCI 搜 lx wang。

王立新

2012 年 4 月 1 日（愚人节）于北京

附件 1：

高斯云的定义及数学特性

李老师说过云模型是在概率的基础上通过计算机算法实现的，是基于泛高斯分布形成的认知模型。所以我想高斯云的生成算法是李老师云模型理论的出发点与核心。下面我从高斯云的生成算法出发，按顺序做以下几件事：

1. 反推出高斯云的数学定义；
2. 证明高斯云主要的几个数学特性；
3. 给出从采样数据求得高斯云三个参数的具体算法与详细公式；
4. 简单说明云综合的一些方法；
5. 最后以一个笑话结束。

先引用李老师的高斯云生成算法，如下：

算法 1：正态云模型

输入 *表示定性概念 A 的 3 个数字特征值 Ex, En, He，云滴数 N。*

输出 *N 个云滴的定量值，以及每个云滴代表概念 A 的确定度。*

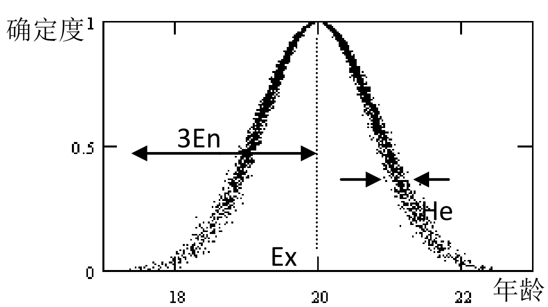
(1) 以 En 为期望值，He 为标准差的一个正态随机数 En′；

(2) 生成以 Ex 为期望值，abs (En′)为标准差的正态随机数 x；

(3) 令 x 为定性概念 $\tilde{A}$ 的一次具体量化值，称为云滴；

(4) 计算 y = exp [−(x−Ex)2/2(En′)2]，令 y 为 x 能够代表定性概念 A 的确定度；

(5) 重复 Step 1 至 Step 4，直到产生 N 个云滴。



通过三个数字特征生成的云图

根据以上算法 1，我们反推出高斯云的数学定义，如下：

<u>定义 1</u>：（高斯云的数学定义）考虑两个论域 U 和 V。在 U 上定义均值为 Ex、标准差为 y 的高斯随机变量 X，即 X 的密度函数为

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

在 V 上定义均值为 En、标准差为 He 的高斯随机变量 Y，即 Y 的密度函数为

$$f(y)=\frac{1}{He\sqrt{2\pi}}e^{-\frac{(y-En)^2}{2He^2}} \tag{2}$$

通过函数 g(x,y)=x（可以理解为 g(x,y)=1*x+0*y）定义二维论域$U\times V$上的随机变量

$$Z=g(X,Y)=X \tag{3}$$

设 X 和 Y 相互独立，则 Z 的密度函数为

$$f(z)=\int_{-\infty}^{\infty}\frac{1}{y\sqrt{2\pi}}e^{-\frac{(z-Ex)^2}{2y^2}}\frac{1}{He\sqrt{2\pi}}e^{-\frac{(y-En)^2}{2He^2}}dy \tag{4}$$

这样定义的 Z 称作高斯云。

从定义 1 可以看出高斯云 Z 是定义在二维论域$U\times V$上的一个随机变量，如果要获得 Z 的一个实现（采样）需要分别在 V 和 U 上各进行一次独立的随机数产生过程，这正是算法 1 所完成的。具体地讲，先根据密度函数（2）在论域 V 上产生随机变量 Y 的一个实现 y=En’，然后根据密度函数（1）在论域 U 上产生随机变量 X 的一个实现 x（暂且忽略算法 1 中 abs(En’)的影响），这样的 x 就是算法 1 第 3 步产生的云滴。

下面给出高斯云 Z 的几个基本数学特性（定理 1 到定理 4）。

定理 1：高斯云 Z 的均值等于 Ex，即

$$E\{Z\}=Ex \tag{5}$$

证明：根据定义

$$E\{Z\}=\int_{-\infty}^{\infty}zf(z)dz \tag{6}$$

将式（4）的 f(z)代入式（6），并进行积分次序的互换，我们得到

$$\begin{aligned}E\{Z\}&=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}z\frac{1}{y\sqrt{2\pi}}e^{-\frac{(z-Ex)^2}{2y^2}}dz\right]\frac{1}{He\sqrt{2\pi}}e^{-\frac{(y-En)^2}{2He^2}}dy\\&=\int_{-\infty}^{\infty}Ex\frac{1}{He\sqrt{2\pi}}e^{-\frac{(x-En)^2}{2He^2}}dy\\&=Ex\end{aligned} \tag{7}$$

■

<u>定理 2</u>：高斯云 Z 的方差

$$\mathrm{Var}\{Z\} = E\{[Z - E(Z)]^2\} = En^2 + He^2 \tag{8}$$

证明：类似于定理 1 的证明，并利用定理 1 的结果 E{Z}=Ex，我们有

$$\begin{aligned}\mathrm{Var}\{Z\} &= \int_{-\infty}^{\infty}[z - E(Z)]^2 f(z)dz \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}(z - Ex)^2 \frac{1}{y\sqrt{2\pi}} e^{-\frac{(z-Ex)^2}{2y^2}} dz\right] \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy \\ &= \int_{-\infty}^{\infty} y^2 \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy \\ &= E\{Y^2\} = He^2 + En^2 \end{aligned} \tag{9}$$

■

<u>定理 3</u>：高斯云 Z 的三阶中心矩(central moment)等于零，即

$$E\{[Z - E(Z)]^3\} = 0 \tag{10}$$

证明：用前面同样的方法，我们有

$$\begin{aligned}E\{[Z - E(Z)]^3\} &= \int_{-\infty}^{\infty}[z - E(Z)]^3 f(z)dz \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}(z - Ex)^3 \frac{1}{y\sqrt{2\pi}} e^{-\frac{(z-Ex)^2}{2y^2}} dz\right] \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy \\ &= \int_{-\infty}^{\infty} 0 \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy = 0 \end{aligned} \tag{11}$$

■

<u>定理 4</u>：高斯云 Z 的四阶中心矩

$$E\{[Z - E(Z)]^4\} = 9He^4 + 18He^2En^2 + 3En^4 \tag{12}$$

证明：运用同样的积分互换方法，并注意到对均值为零、方差为σ^2的高斯随机变量 X 我们有$E\{X^4\} = 3\sigma^4$, 得

$$\begin{aligned}E\{[Z - E(Z)]^4\} &= \int_{-\infty}^{\infty}[z - Ex]^4 f(z)dz \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}(z - Ex)^4 \frac{1}{y\sqrt{2\pi}} e^{-\frac{(z-Ex)^2}{2y^2}} dz\right] \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy\end{aligned}$$

$$= \int_{-\infty}^{\infty} 3y^4 \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy$$

$$= 3\int_{-\infty}^{\infty} [(y-En)+En]^4 \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy$$

$$= 3\int_{-\infty}^{\infty} [(y-En)^4 + 4(y-En)^3 En + 6(y-En)^2 En^2 + 4(y-En)En^3 + En^4] \frac{1}{He\sqrt{2\pi}} e^{-\frac{(y-En)^2}{2He^2}} dy$$

$$= 3(3He^4 + 6He^2 En^2 + En^4) \tag{13}$$

■

根据定理 1－4，我们可以进行如下讨论。常说 He 用来衡量高斯云偏离正态分布的程度，现在有了精确的公式定理 1－4，我们看看这种说法是否正确。为此我们来构建一个与高斯云 Z 尽可能接近的高斯随机变量 X’，这里“尽可能接近”的含义是 Z 和 X’ 的各阶中心矩要尽可能地相等，这个构建的高斯随机变量 X’ 的密度函数为

$$f(x') = \frac{1}{\sqrt{2\pi(En^2+He^2)}} e^{-\frac{(x'-Ex)^2}{2(En^2+He^2)}} \tag{14}$$

因为 X’ 的均值、方差及三阶中心矩分别为 Ex，En^2+He^2 和 0，这与高斯云 Z 的完全相同，所不同的是 X’ 的四阶中心矩为

$$E\{(X'-Ex)^4\} = 3(En^2+He^2)^2 = 3He^4 + 6He^2En^2 + 3En^4 \tag{15}$$

比较式（15）和（12），我们发现高斯云 Z 的四阶中心矩要比 X’ 的四阶中心矩大 $6He^4+12He^2En^2$，据此我们得出结论：定性地讲 He 确实可以用来衡量高斯云 Z 偏离正态分布的程度，而这个偏离程度的更准确的定量指标是 $6He^4+12He^2En^2$ 的大小。

下面我们看看如何从采样数据求得高斯云 Z 的三个参数 Ex,En 和 He。由于有了定理 1－4 的具体公式，我们将会到这些参数的求取是非常方便的。给定高斯云 Z 的 N 个采样值 $z_1, z_2, \cdots, z_N$，先算出其样本均值

$$\bar{m} = \frac{1}{N}\sum_{i=1}^{N} z_i \tag{16}$$

样本方差

$$\bar{v} = \frac{1}{N-1}\sum_{i=1}^{N-1}(z_i - \bar{m})^2 \tag{17}$$

及样本四阶中心矩

$$\overline{\mu_4} = \frac{1}{N-1}\sum_{i=1}^{N-1}(z_i - \bar{m})^4 \tag{18}$$

然后根据定理 1、2、4（式（5）、（8）、（12）），我们有

$$\mathrm{Ex} = \overline{\mathrm{m}} \tag{19}$$

$$\mathrm{En}^2 + \mathrm{He}^2 = \overline{\mathrm{v}} \tag{20}$$

$$9\mathrm{He}^4 + 18\mathrm{He}^2\mathrm{En}^2 + 3\mathrm{En}^4 = \overline{\mu_4} \tag{21}$$

联立求解（20）和（21），我们得到

$$\mathrm{En} = \sqrt[4]{\frac{9\overline{\mathrm{v}}^2 - \overline{\mu_4}}{6}} \tag{22}$$

$$\mathrm{He} = \sqrt{\overline{\mathrm{v}} - \sqrt{\frac{9\overline{\mathrm{v}}^2 - \overline{\mu_4}}{6}}} \tag{23}$$

于是我们有了下面的算法：

算法 2：逆向正态云算法

输入 *高斯云 Z 的 N 个采样值* $z_1, z_2, \cdots, z_N$ *。*

输出 *高斯云 Z 的三个参数 Ex, En 和 He 的样本估计值。*

(1) *根据式（16），（17）和（18）算出样本均值，样本方差及样本四阶中心矩；*

(2) *根据式（19），（22）和（23）求得高斯云 Z 的三个参数 Ex，En 和 He 的样本估计值。*

由于式（22）和（23）中涉及求四阶及二阶平方根，大家会问如果$9\overline{\mathrm{v}}^2 - \overline{\mu_4} < 0$或者 $\overline{\mathrm{v}} < \sqrt{\frac{9\overline{\mathrm{v}}^2 - \overline{\mu_4}}{6}}$怎么办？一种简单（且粗暴）的解决方法是如果这种情况发生，就说该组数据不适合用高斯云模型来描述。或者哲学点地说就是该组数据不能形成一个云概念。

下面简要地说说云综合。设 Z1、Z2 是二维论域$\mathrm{U} \times \mathrm{V}$上两朵美丽的高斯云，如何用这两朵云合成出更美丽的云朵呢？由于 Z1、Z2 说到底是随机变量，所以最基本的合成无非是加、减、乘、除，再加上取最大、最小及圆周化，等等。具体地讲，合成出的云 Z 可以是：

$$\mathrm{Z} = \mathrm{Z1} \pm \mathrm{Z2}$$

$$\mathrm{Z} = \mathrm{Z1} \times \mathrm{Z2}$$

$$\mathrm{Z} = \mathrm{Z1}/\mathrm{Z2}$$

$$\mathrm{Z} = \max(\mathrm{Z1}, \mathrm{Z2})$$

$$\mathrm{Z} = \min(\mathrm{Z1}, \mathrm{Z2})$$

$$\mathrm{Z} = \sqrt{\mathrm{Z1}^2 + \mathrm{Z2}^2}$$

等等。由于 Z1、Z2 的密度函数已知（见式（4）），以上合成后的 Z 的密度函数可以用教科书上的方法求得，这里不再多述。

写了这么多数学公式，最后应该轻松一下。都说云模型是认知模型，我没有研究过认知理论，感觉好像是介于科学和哲学之间的一个学科吧？由此想到一个关于科学家（工程师）、经济学家和哲学家的经典故事，这里说来和大家一起乐一乐。

一天，一个科学家（工程师）、一个经济学家、和一个哲学家来到一间黑屋子的门口，看见一只黑猫跑进了黑屋子。科学家二话没说一头扎进黑屋子里去寻找黑猫。经济学家想了想，觉得在黑屋子里抓黑猫太难，于是假设黑猫没有在黑屋子里，写了一篇论文，提出一套在阳光下抓黑猫的理论。哲学家环顾四周，发现旁边有一间没有黑猫的黑屋子，于是走进去，然后没过多久就大喊道：“我抓住了！我抓住了！”

如果这时又来了一位认知学家，会发生什么呢？会不会捧过经济学家的论文，然后向哲学家的喊声奔去？

（初稿没有细查，有错请朋友们更正。）

附件 2：

正态云模型的重尾性质证明

李德毅 [1]，刘常昱 [2]，淦文燕 [2]

（1 中国电子系统工程研究所，北京，100840

2 解放军理工大学指挥自动化学院，南京，210007）

摘要：正态分布和重尾分布在概率研究中具有非常重要的地位，二者具有完全不同的数学形式和物理意义。正态分布的密度函数以指数函数衰减至 0，服从正态分布的随机变量，其绝大多数取值在其期望附近，偏离期望很大的取值很少。而服从重尾分布的随机变量，其尾分布函数具有重尾特性，密度函数以幂指数衰减至 0。本文证明了正态云模型是具有均值的重尾分布，是介于正态分布与重尾分布之间的中间状态，正态云模型的参数超熵 He 是可以实现正态分布向重尾分布转换的桥梁。

关键词：正态分布，重尾分布，正态云模型，峰度

中图分类号：TP182

1、 引言

在概率论与随机过程的研究中，正态分布的地位举足轻重。正态分布的密度函数和分布函数具有比较简单的数学形式和一些很好的数学性质。正态分布是许多重要概率分布的极限分布，许多非正态的随机变量是正态随机变量的函数，这些都使得正态分布在理论和实际中应用非常广泛。

中心极限定理从理论上阐述了产生正态分布的条件，中心极限定理简单直观的阐述是：如果决定某一随机变量结果的是大量微小的、独立的随机因素之和，并且每一随机因素的单独作用相对均匀的小，没有一种因素可起到压倒一切的主导作用，那么这个随机变量一般近似服从于正态分布。正态分布广泛存在于自然现象、社会现象、科学技术以及生产活动中，在实际中遇到的许多随机现象都服从或者近似服从正态分布。例如正常生产条件下的产品质量指标，随机测量误差，同一生物群体的某种特征，某地的年平均气温等等。

通常在讨论有关概率问题时，分布函数 $F(x)$起着非常重要的作用，其尾分布函数 $1-F(x)$在实际中的应用尤为重要。例如，在有关可靠性的分布中，可靠性与失效率等概念都同尾分布函数有关。自上世纪 60 年代以来，国外出现了大量重尾分布的研究文献[1][2][3]。但是究竟什么是重尾分布？至今仍未有一个确切统一的定义来描述，甚至名称也没有完全统一。不同文献中，重尾分布(heavy-tailed distributions)

也被称为胖尾(fat-tailed)，厚尾(thick-tailed)或者长尾(long-tailed)分布。在本文中，把这些名词当作同义词，在不引起混淆的情况下，统称为重尾分布。

重尾分布在分支过程、排队论和可靠性的研究中，特别是近年来在金融工程，数量经济和保险精算等研究领域都有广泛应用。特别是 1998 年 BA 模型的提出[4]，使全世界不同领域的众多学者对幂律分布产生了极大的兴趣和热情。而幂律分布就是一类重尾分布。由于优胜劣汰的竞争机制在生物界、人类社会等诸多领域中都或多或少的存在，因此，越来越多的研究者倾向于认为，重尾分布具有不亚于正态分布的意义。

无论从产生条件、数学形式还是物理意义而言，重尾分布都和正态分布截然不同。那么，二者之间到底有无内在联系？有无转换桥梁？这将是本文所要论述的中心问题。

2、 正态分布与重尾分布

定义 1 若随机变量 X 的概率分布函数形式为：

$$F\left(x,\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}\exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right)\mathrm{d}u \text{。}$$

或者概率密度函数为：

$$f(x;\mu,\sigma^2)=(2\pi)^{-1/2}\sigma^{-1}\exp\left[-(x-\mu)^2/2\sigma^2\right]$$

则称 X 为正态分布，记为 X~N(μ,σ^2)。其中，μ 和 σ^2 分别是正态分布的期望和方差，分别表征随机变量的最可能取值以及一切可能取值的离散程度。

重尾分布则有很多等价的数学定义，其中 Embrechts 的定义应用最为广泛[1]。

定义 2 称随机变量 X 是重尾分布，如果它不存在指数阶矩，即对任意 λ>0，有 $Ee^{\lambda X}=\int_0^{+\infty}e^{\lambda x}dF(x)=\infty$。其中 F(x)是 X 的分布函数。

还有一种定义是相对于正态分布而言，以四阶中心矩为基础的。四阶中心矩具有峰度（kurtosis）的含义，峰度是统计中描述分布状态的一个重要特征值，用以判断概率密度函数曲线相比于正态分布的尖平程度。如果将正态分布视为常峰态，密度函数曲线的形状比正态分布更高更瘦的称为高峰态，否则称为低峰态。

定义 3[5] 随机变量 X 称为是重尾的，如果 $E[\frac{(X-\mu)^4}{\sigma^4}]>3$，其中 μ，σ 分别为 X 的期望和标准差。

正态分布的峰度为 3，因此该性质被称为超过或大于峰度。但是，该定义只适用于四阶矩存在的情况。

另一种是判断重尾分布较为直观的定义。

定义 4[1] 如果密度函数是以幂指数衰减至 0 的，则该分布函数为重尾的；如果密度函数是以指数函数衰减至 0 的，称该分布函数为轻尾的。

形象而言，重尾分布就是密度函数”尾巴”比较长的分布。在应用领域，它的物理意义可以解释为极端事件的概率不为 0。如保险业的大额索赔问题，在 N 次索赔中有一次索赔的额度非常大，以至于其它 N-1 次索赔相对于这次索赔而言是微不足道的，这类大额索赔问题就需要用重尾分布来处理。这类问题就是所谓的极端事件问题，如地震、洪水、股灾等。这类问题研究在现实中有很强的应用价值，尤其是 911 事变后，大量保险公司破产，极端事件的研究更成为新的前沿研究热点。因此，重尾分布的研究也受到越来越多学者的关注。

3、 云模型——超熵与重尾分布

云模型是利用三个数字特征，结合特定算法实现的定性概念及其定量表示之间的转换模型。其三个数字特征及具体算法定义如下[6]。

期望 Ex：云滴在论域空间分布的期望。即最能够代表定性概念的点。

熵 En：代表定性概念的粒度，熵越大，概念越宏观。

超熵 He：熵的不确定性度量，即熵的熵。

定义 5：设 U 是一论域，C 是 U 上的定性概念，（Ex, En, He）为 C 的数字特征。若定量值 $x\in U$ 是定性概念 C 的一次随机实现，x 满足：

$x \sim N(Ex, En'^2)$， 其中，$En' \sim N(En, He^2)$， $N(En, He^2)$表示期望为 En，方差为 He^2 的正态分布。

x 对 C 的确定度 y 根据下式计算： $y = e^{-\frac{(x-Ex)^2}{2(En')^2}}$

那么 x 在论域 U 上的分布 X 称为正态云。

容易计算，正态云 X 的概率密度函数为 $f_X(x)=\int_{-\infty}^{\infty}\frac{1}{2\pi He|y|}e^{-\frac{(x-Ex)^2}{2y^2}-\frac{(y-En)^2}{2He^2}}dy$。

文献[6]已经证明，正态云模型的期望和方差分别为 Ex 和 En^2+He^2。其概率密度函数以 Ex 为中心左右对称。接下来，我们将证明，正态云 X 还具有重尾特性。

定理 1：当 He>0 时，正态云模型为重尾分布。

证明：正态云模型 X 的四阶中心矩为

$$E\{[X-E(X)]^4\}=\int_{-\infty}^{+\infty}[x-Ex]^4 f(x)dx$$

$$=\int_{-\infty}^{+\infty}[\int_{-\infty}^{+\infty}(x-Ex)^4\frac{1}{y\sqrt{2\pi}}e^{-\frac{(x-Ex)^2}{2y^2}}dx]\frac{1}{\sqrt{2\pi}He}e^{-\frac{(y-En)^2}{2He^2}}dy$$

$$=\int_{-\infty}^{+\infty}3y^4\frac{1}{\sqrt{2\pi}He}e^{-\frac{(y-En)^2}{2He^2}}dy$$

$$=3\int_{-\infty}^{+\infty}[(y-En)+En]^4\frac{1}{\sqrt{2\pi}He}e^{-\frac{(y-En)^2}{2He^2}}dy$$

$$=3\int_{-\infty}^{+\infty}[(y-En)^4+4(y-En)^3En+6(y-En)^2En^2+4(y-En)En^3+En^4]\frac{1}{\sqrt{2\pi}He}e^{-\frac{(y-En)^2}{2He^2}}dy$$

$$=3(3He^4+6He^2En^2+En^4)$$

由于正态云 X 的方差 En^2+He^2，故正态云模型的峰度为

$$K(X)=\frac{E[X-E(X)]^4}{(En^2+He^2)^2}$$

$$=\frac{3(3He^4+6He^2En^2+En^4)}{(En^2+He^2)^2}$$

$$=9-\frac{6}{(1+\frac{He^2}{En^2})^2}>3$$

故由定义 3，当 He>0 时，正态云模型为重尾分布。

根据定理 1，我们可以很容易看出超熵 He 的物理意义。当 He=0 时，正态云模型退化为正态分布。随着 He 的增大，云滴分布将由正态分布向重尾分布转换。可以说，正态云模型的数字特征超熵 He 是正态分布与重尾分布之间的桥梁。

4、结束语

大量的随机因素作用会导致正态分布，在自然界，完全由大量随机因素主导的现象会比较多，因此，在处理自然现象时，正态分布可能会发挥比较好的作用。而在人类社会或者有生命存在的地方，通常会存在不同程度的竞争、适者生存等因素，

优先依附的准则在很多情况下适用，因此幂律分布这类重尾分布也受到了不同领域研究者的青睐。但是，人类社会中的诸多现象或者人类行为，既不会完全由随机因素主宰，因为人类会理智思考，选择最利己的行动。也不会任由富者越富，大小通吃，因为会有类似国家政府的干预等因素。因此，真实的社会中，更多的现象是既存在极端事件，也有大量中间成分，是介于正态分布（平均主义）与幂律这样的重尾分布（完全不平衡）之间的中间状态：有期望的重尾分布。可以说云模型，就是介于正态与重尾之间的中间分布。

通过本文的研究，证明了正态云模型是一类有期望值的重尾分布。这类有平均值的重尾分布具有何种数学性质？其在数学领域的诸多性质证明，与其它重尾分布的比较研究，其内在形成机制的阐述证明，以及其在极端事件预测等领域的应用研究等，都将是下一步的重点关注问题。

参考文献

[1]Embrechts P, Klupperberg C, Mikosch T.Modelling Extremal Events for Insurance and Finance[M].Berlin:Springer-Verlag,1997

[2]Mandjes M. Overflow behavior in Queues with Many Long-tailed Inputs[J].J Appl Probab,2002,37:1150-1167

[3] Baltrunas A. Some Asymptotic Results for Transient Random Walks with Applications to Insurance Risk[J].J Appl Probab, 2001,38:108-121

[4] Albert R, Barabasi Al,Emergence of scaling in random networks[J].Science 286(5439):12-19.1999

[5] Thomas Werner, Christian Upper. Time Variation in the Tail Behavior of Bund Futures Returns. http://www.bundesbank.de/download/volkswirtschaft/dkp/2002/200225dkp.pdf

http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp199.pdf

[6]李德毅，刘常昱。论正态云模型的普适性。中国工程科学，2004，6（8）：28-34

Proof of the Heavy-tailed Property of Normal Cloud Model

Deyu Li[1], Changyu Liu[2], Wenyan Gan[2]

(1 *Institute of Electronic System Engineering,Beijing ,China,100039*

2 *Institute of Command Automation,PLA University of Science and Technology*

Nanjing, China 210007)

Abstract：Normal distributions and heavy-tailed distributions are very important in probability theories. They have very different mathematical forms and physical meanings. The probability density function of a normal distribution decay exponentially to 0. And the majority of a normal random variable values around the mathematical expectation. And obey the heavy-tailed random variable, it means a heavy-tailed characteristics, the probability density function decay a power exponentially to 0.In this paper, we proved that the normal cloud model is a heavy-tailed distribution and its mathematical expectation exists.It is a intermediate between notmal distributions and heavy-tailed distributions. The parameter He(hyper-entropy) of the normal cloud model is the bridge from normal distributions to heavy-tailed distributions.

Keyword：Normal distribution, Heavy-tailed distribution, Normal cloud model, Kurtosis