中心极限定理：
——场跨越两百年的传奇

郭　旭

中心极限定理（Central limit theorems, CLT）是概率论和统计学中的一个核心定理。在统计学的实践中，中心极限定理无处不在。中心极限定理表明，在一定条件下，若干随机变量的算术平均值的分布可以用正态分布近似。中心极限定理简洁优美且有普遍适用性。它并不需要明确知道参与求和的随机变量的精确分布，而只要满足一些较弱的条件即可。然而和大数定律 (Law of large numbers) 不同，中心极限定理并不是那么自然。我们很容易理解硬币正面朝上的频率应该逐渐接近于朝上的概率，而不太容易理解为什么即使每个变量是从偏态分布比如指数分布、卡方分布中抽取的，它们的样本均值都会近似服从正态分布。从 1733 年棣莫弗（Abraham De Moivre）首次得出了一个特殊版本的中心极限定理到 1935 年费勒和莱维给出中心极限定理的充要条件，无数伟大的数学家和统计学家做出了巨大的贡献，共同谱写了一场跨越两百年的传奇。在开启这段传奇之旅前，我们先通过两个例子简单认识下中心极限定理。

假设我们抛掷一枚硬币一万次，正面朝上的次数是九千次，现在问这枚硬币是均匀的吗？我们用反证法进行思考。如果这枚硬币是均匀的，那么根据中心极限定理正面朝上的频率将近似服从期望为 1/2，方差为 $0.25/10^4$ 的正态分布。也就是说频率应围绕 1/2 波动，而波动的范围以近似 99.7% 的概率在 $0.5-3*0.5/100$ 到 $0.5+3*0.5/100$ 之间。然而我们实际观测到的频率取值却是 0.9，并不在上述范围中。这就表明这枚硬币不应该是均匀的。上面的例子体现了中心极限定理在假设检验问题中的应用。

此外如何直观理解中心极限定理成立呢？对此我们考虑这样一个例子。令 X_1 和 X_2 分别表示掷两枚骰子的点数，那么容易知道 $P(X_k=i)=1/6$, $k=1, 2$; $i=1,\ldots, 6$。显然 X_1 和 X_2 本身并不呈现正态分布所具有的钟形特征。如果把它们加起来会如何呢？我们发现 $P(X_1+X_2=12)=P(X_1+X_2=2)=1/36$，但 $P(X_1+X_2=7)=1/6$，因为有 6 种可能得到 7。这表明 X_1 和 X_2 的和取中间值的概率在增大，而取边缘值的概率在减小。将更多类似的变量加在一起这种效应将不断加强，最终会呈现正态分布所具有的钟形特征。

在有了以上准备之后，我们正式开始这段传奇之旅。

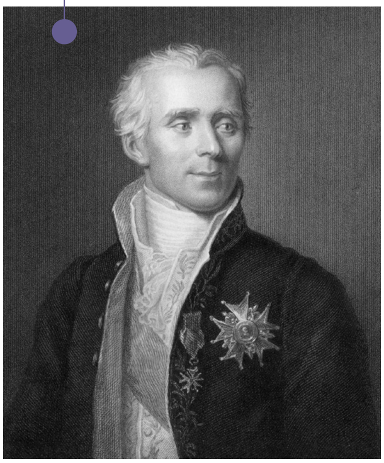
1. 法国数学家的贡献

棣莫弗（1667–1754）

这一伟大传奇是从法国数学家棣莫弗开启的。他在 1733 年给出了中心极限定理的雏形。这也是正态分布的第一次出现。他利用由他和斯特林（James Stirling）共同发展的公式（现今被称为斯特林公式）给出了二项分布的正态近似。他的这一发现也被称为棣莫弗 - 拉普拉斯定理。他的工作是雅各布 • 伯努利（Jacob Bernoulli）大数定律的进一步发展。尽管大数定律表明频率和概率很接近，但却无法给出均匀硬币在 n 次抛掷中 $n/2$ 次正面朝上的概率的更加精细的刻画。然而棣莫弗仅仅将这一超越时代的成果看作是二项分布的近似，而并没有意识到中心极限定理的普遍存在。他在概率论方面的贡献还包括撰写了概率论早期的重要专著之一《机会论》（*The Doctrine of Chances*）。除了在概率论方面的贡献之外，棣莫弗还发现了如下简洁优美的数 学公式：对于两个复数 $Z_1 = r_1(\cos(\theta_1) + i\sin(\theta_1))$, $Z_2 = r_2(\cos(\theta_2) + i\sin(\theta_2))$, 那么

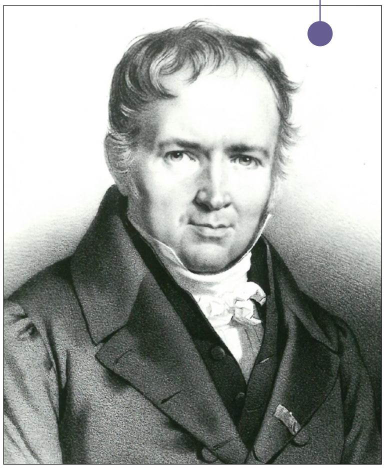
$$Z_1 Z_2 = r_1 r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)].$$

这一结果被称为棣莫弗定理。棣莫弗于 1697 年当选为英国皇家学会会员，并被柏林科学院和巴黎科学院接纳为院士。



拉普拉斯（1749–1827）

下一个重大的突破是由拉普拉斯（Pierre Laplace）给出的。1812 年他发表了重要的《概率分析论》（*Théorie Analytique des Probabilités*）一书。他使用由他在 1785 年引入的特征函数进行了论证。这一证明思路甚至影响到了 100 年之后的李亚普诺夫（Aleksandr Lyapunov）。尽管在 1785 年拉普拉斯已得出了一些初步的结果，但非常奇怪的是直到近 40 年后他才得到一个较一般化的中心极限定理。拉普拉斯所得出的中心极限定理是第一个一般化的结果。他的结果可应用于有界的随机变量。



泊松（1781-1840）

泊松（Simeon Poisson）在 1824 年和 1829 年发表了两篇文章来讨论中心极限定理。他希望能够对拉普拉斯的结果给出更加严格的论证。他的贡献主要是两个方面：1. 他创造了 "choses" 这一概念，可以看作是现代观点下的“随机变量”的开端；2. 他给出了一些反例比如柯西分布来说明中心极限定理有时并不总是成立。

这一时期的研究主要将中心极限定理看作是个工具而不是数学对象本身。在早期，概率论并不被看作是严格的数学理论而更多的是一种常识。到了 19 世纪后期，很多数学家希望给出中心极限定理更加严格的证明。这其中包括德国数学家狄利克雷（Dirichlet），贝塞尔（Friedrich Bessel）和法国数学家柯西（Augustin Cauchy）等。

狄利克雷和贝塞尔在他们的证明中引入了 " 不连续因子 "，进一步发展和完善了泊松的证明。狄利克雷还尝试给出正态近似的误差。尽管这一尝试并不是很成功，但这是第一次对近似误差的较深入的研究。柯西是第一批认真地将概率论看作是纯粹数学的大数学家之一。在与比内梅（Irénée Bienaymé）关于最小二乘的讨论中，柯西建立了正态近似误差的一个上界。

关于中心极限定理的早期证明往往不够严谨，定理所需的条件并没有清晰地给出，同时一般限定所讨论的变量有界。从 1870 年到 1910 年，俄国数学家开始在历史舞台上展露他们的才华。以切比雪夫（Pafnuty Chebyshev），马尔可夫（Andrey Markov）和李亚普诺夫为代表的“圣彼得学派”在中心极限定理的发展史上写下了浓墨重彩的篇章。

欧洲大陆上的数学发展对俄国的数学产生了巨大影响。在这其中，积极向西欧学习先进技术的彼得大帝起到了重要作用。在 1724 年 1 月，彼得大帝批准了按照他的旨意拟订的要在圣彼得堡建立科学院的草案，并作了若干修改和补充。在科学院设立之初，当时数学领域的院士全是外国人，包括欧拉和丹尼尔•伯努利（Daniel Bernoulli）。通过聘请国外知名的大数学家，彼得科学院的设立也极大地促进了俄国数学的发展。切比雪夫就是俄国数学发展的典型代表。切比雪夫不仅自己在数论、力学、概率论和统计学等领域取得巨大成就，更培养了包括马尔可夫和李亚普诺夫在内的多位数学家。其中马尔可夫在随机过程领域做出重要工作，他所提出和深入研究的马尔可夫链和马尔可夫过程现已进入本科教学课程，并对当前的强化学习领域起到重要作用；而李亚普诺夫则以他在动态系统的稳定性方面做出的贡献而闻名。在切比雪夫的组织和领导之下，圣彼得堡数学学派逐渐形成，让俄国数学逐渐走到了世界前列。