浅析正态分布在计量学中的现实意义

□李东岳[1]　王宁[2]

(1. 辽宁省计量科学研究院，沈阳 110004　2. 中国合格评定国家认可中心，北京 100062）

【摘　要】正态分布是一个在数学、物理及工程等领域都非常重要的概率分布，有极其广泛的实际背景，在统计学的许多方面有着重大的影响力。生产与科学实验中很多随机变量的概率分布都可以近似地用正态分布来描述。在计量学中，正态分布经常用于表述测得数据在数值区间内的分布情况。本文通过分析生活实际中的概率学和统计学的实例，简要概述正态分布在计量学中的现实意义。

【关键词】正态分布；概率论；统计学

文献标识码：B　　　　**文章编号：**1003-1870（2023）10-0012-02

引言

正态分布是统计学中的一个最重要的数学模型，也被称为“高斯分布”（Gaussian distribution）。高斯在1795 年发现了质数分布定理和最小二乘法，在这些基础之上，得到高斯钟形曲线(正态分布曲线), 其函数被命名为标准正态分布(或高斯分布), 并在概率计算中大量使用。按照国际计量局（BIPM），国际标准化组织（ISO）与国际法制计量组织（OIML）以及国际临床化学联合会（IFCC）、国际理论和应用化学联合会（IUPAC）和国际理论与应用物理学联合会（IUPAP）等七个国际组织联合制订的《国际通用计量学基本术语》（1993 年版），计量学被定义为“测量学科”，并在注解中说明：“计量学包括涉及测量理论和实用的各个方面，不论其不确定度如何，也不论其用于什么测量技术领域（Metrology is the science of measurement, made at a known level of uncertainty, in any field of human activity.）。”其中，在不确定度（Uncertainty）的定义和计算中，即应用了正态分布的数学理论[3][5][6]。

1　正态分布的来源

正态分布理论来源于一个著名的物理实验：高尔顿钉板实验（Galton knocked boards）见图1。即从上方入口处放进一个直径略小于两颗钉子之间的距离的小圆玻璃球，当小圆球向下降落过程中，碰到密集的钉子层后皆以1/2 的概率向左或向右滚下，于是又碰到下一层钉子。如此继续下去，直到滚到底板的一个格子内为止。把许许多多同样大小的小球不断从入口处放下，只要球的数目相当大，它们在底板将堆成近似于正态曲线的密度函数图形，即是正态分布曲线（即: 中间高，两头低，呈左右对称的古钟型）。

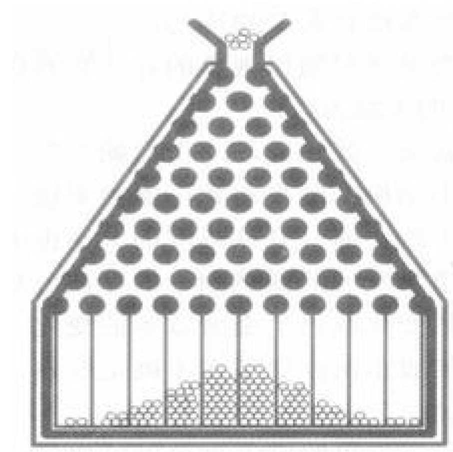

图1　高尔顿钉板实验

其实，只要一个问题或事件中涉及到很多个随机量，这些随机量的分布叠加起来就非常接近正态分布。日常中大多事件在正常情况下都符合正态分布。例如，一个学校里，某次班级的考试成绩，中等成绩的人占大多数，特别优秀或特别差的人占少数。或者某一个地区的男性人群中，中等身高的人

占多数，身高特别高或特别矮的人占少数。

2　正态分布的特征

正态分布曲线见图2，其中 $p(x)$ 纵坐标为概率密度（也有概率的意义），横坐标 x 为测量统计数据，μ 是数学期望（也就是数据平均值的意思）。σ 是标准偏差（也称标准差），在区间 $(\mu-\sigma,\mu+\sigma)$ 阴影中的面积 p 就是测得统计数据出现的概率大小。

经过科学计算得到统计数据 x 的一些基本结论，即

$\mu-\sigma\le x\le\mu+\sigma$ 的概率为68.3%，即包含因子 $k=1$；

$\mu-2\sigma\le x\le\mu+2\sigma$ 的概率为95.5%，即包含因子 $k=2$；

$\mu-3\sigma\le x\le\mu+3\sigma$ 的概率为99.7%，即包含因子 $k=3$。

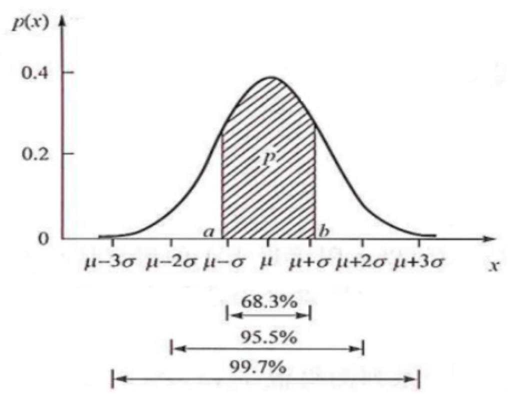

图2　正态分布曲线图

假如说将所有测量统计数据统计出来，利用正态分布这种方法，就可以最终知道下一次测量数据在一定范围内的概率了。在这里我们举个实例：某地区市场监督部门在某超市（流通领域）随机抽样，专项检查商品过度包装，对于某同一品牌的商品空隙率进行检测。该商品随机抽样的同一种商品的10个样品包装的空隙率经过检测分别为38.4%，39.6%，40.3%，42.3%，39.9%，40.4%，42.2%，45.3%，43.3%，40.3%；根据GB23350–2021《限制商品过度包装要求食品和化妆品》的国家标准，包装空隙率不得大于60%，通过计算得出该品牌商品在包装空隙率方面不合格的概率是多少。

首先，求出该品牌同一种商品的10个样品包装空隙率平均值 μ；

$$\mu=\frac{1}{n}\sum_{i=1}x_i=\frac{1}{10}(38.4\%+39.6\%+....+43.3\%+40.3\%)=41.2\%$$

利用贝塞尔公式计算出该组数据的标准差 σ

$$\sigma=\sqrt{\frac{1}{n}\sum_{i=1}(x_i-\mu)^2}=2.1\%$$

该商品的包装空隙率与标准规定的要求之间差了 $\Delta x=|41.2\%-60\%|=18.8\%$，这个数据约等于 9.0σ。这时通过图1的正态分布曲线图来看，不合格的概率基本上落在了3σ之外，通过计算大概得出该品牌商品包装空隙率不合格的可能性只有0.1%，从检测出的数据概率上来说，该品牌基本上不会有过度包装的相关商品。

3　结语

正态分布是许多统计方法的理论基础。检验、方差分析、相关和回归分析等多种统计方法均要求分析的指标服从正态分布。在计量学中广泛利用正态分布来归纳统计数据，计算不确定度，有着非常重要的现实意义。许多统计方法虽然不要求分析指标服从正态分布，但相应的统计量在大样本时近似正态分布，因而大样本时这些统计推断方法也是以正态分布为理论基础的。

参考文献

[1] 中国计量测试学会. 注册计量师基础知识及专业实务[M]. 北京：中国质检出版社, 2017.3.

[2] 孙洪祥 柳金浦主编. 概率论与数理统计（二）[M]. 北京：辽宁大学出版社, 2012.3.

[3] 颜素容 崔红新主编. 概率统计基础[M]. 北京：国防工业出版社, 2011.1.

[4] 全国法制计量管理计量技术委员会. JJF 1070—2005 定量包装商品净含量计量检验规则[S]. 北京：中国标准出版社，2006.1.

[5] 蔡高玉 正态分布及其应用[J]. 北京：环球市场信息导报, 2016.4.

[6] 周富臣 正态分布及其应用（一）[J]. 上海：上海计量测试, 2001.5.

作者简介

李东岳（1983−）, 男，汉族，辽宁沈阳人，高级工程师，主要从事力学计量研究。