

利用“p100/q100”律对胶东地区金矿资源潜力估计

申维^{1,2}

(1. 中国地质大学 国土资源与高新技术研究中心,北京 100083;
2. 中国科学院 地球化学研究所,贵州 贵阳 550002)

摘要:利用幂函数分布在高端截尾(upper truncation)条件下具有分形性质,即尺度不变的特征,证明了统计量 q 具有分形性质,即它与尺度参数 k 无关。在总体的随机变量 X 服从幂函数分布条件下,从“p100/q100”律的概率形式推出“p100/q100”律的分形形式;“80/20”律是“p100/q100”律的一种特殊分形形式。当 $a = 1.16$, $p = 0.20$ 时, $q = 0.80$;将“p100/q100”律应用于山东胶东地区金矿床储量数据,得到了该地区大中小型金矿床的“p100/q100”律表达式和相应的百分数。比较理论计算的结果与样本观察值计算的结果大小,可预测该地区可能存在1个或2个大型金矿床。

关键词:分形;帕累托(Pareto)律;幂函数分布;金矿床

中图分类号:P628 **文献标识码:**A **文章编号:**1005-2321(2001)04-0463-04

意大利经济和社会学家 Vilfredo Pareto (1848—1923)发现,在一般情况下,占全部人口较小比例的人占据了整个社会财富的大部分。帕累托(Pareto)原理可以应用于其它方面,如大型超大型矿床拥有绝大部分的储量。

帕累托(Pareto)律^[1]是描述一个国家居民收入分布的经验关系。在意大利20%的人口拥有80%的社会财富。从这个原理,可推出启发式“20/80”律。“20/80”律在管理科学中有许多应用实例。例如,20%的消费者占80%的销售额;20%的组成部分占80%的成本或费用。

Crovelli^[2]将“20/80”律推广到一般的“p100/q100”律,在总体的随机变量服从帕累托(Pareto)分布条件下,对“p100/q100”律进行分形研究,并应用于油田资源中。他认为最大几个油田的储量对全部油田的储量有决定性的影响。本文利用幂函数分布在高端截尾(upper truncation)条件下具有分形性质,证明了统计量 q_2 具有分形性质,即它与尺度参数 k 无关。在总体的随机变量 X 服从幂函数分布条件

下从“p100/q100”律的概率形式推出“p100/q100”律的分形形式;通过分析,山东胶东地区金矿床储量总体分布服从帕累托分布,可以得到该地区大中小型金矿床的“p100/q100”律表达式和相应的百分数。比较理论计算的结果与样本观察值计算的结果大小,可预测该地区可能存在1个或2个大型金矿床。

1 “p100/q100”律的统计形式,概率形式和分形形式

1.1 “p100/q100”律的统计形式

设总体随机变量为 X ,它的 N 个子样的随机变量为 X_1, X_2, \dots, X_N , 它们的观察值分别为 x_1, x_2, \dots, x_N , 将这些观察值按由小到大排列:

$x_{(1)} \quad x_{(2)} \quad \dots \quad x_{(N)}$, 其中 $x_{(1)} = \text{Min}\{x_1, x_2, \dots, x_N\}$, $x_{(N)} = \text{Max}\{x_1, x_2, \dots, x_N\}$ 。

它们总和为 $S = \sum_{i=1}^N x_{(i)} = \sum_{i=1}^N x_i$ 。

$$\text{令 } S_1 = \sum_{i=[(1-p_1)N]+1}^N x_{(i)}, S_2 = \sum_{i=1}^{[p_2N]} x_{(i)}$$

其中 $[x]$ 表示 x 的最大整数部分, $p_1 (0 < p_1 < 1)$ 是最大几个观察值的个数在全部观察值的个数 N 中所占的比例, $p_2 (0 < p_2 < 1)$ 是最小几个观察值的个数在全部观察值的个数 N 中所占的比例。 S_1 表示最大 $N - [(1 - p_1)N]$ 个观察值的和, S_2 表示最小 $[p_2N]$

收稿日期:2001-04-29;修订日期:2001-05-17

基金项目:国家自然科学基金项目(49873027;40172099);中国科学院重大项目(K2951-1B-411);国土资源部科技项目(B7-10);国土资源部矿产资源定量预测及勘查评价开放研究实验室基金资助项目

作者简介:申维(1957—),男,博士后,数学地质专业,原长春科技大学教授,主要从事数学地质的科研和教学工作。

N]个观察值的和。

$$\begin{aligned}
 \text{定义统计量 } q_1, q_2 \circ q_1 &= S_1 / S = \frac{\sum_{i=1}^{N-p_1} x_{(i)}}{\sum_{i=1}^N x_{(i)}} \\
 &= \frac{p_1 \sum_{i=1}^{N-p_1} x_{(i)} / p_1 N}{\sum_{i=1}^N x_{(i)} / N} \quad (1)
 \end{aligned}$$

$$q_2 = S_2 / S = \frac{\sum_{i=1}^{p_2} x_{(i)}}{\sum_{i=1}^N x_{(i)}} = \frac{p_2 \sum_{i=1}^{p_2} x_{(i)} / p_2 N}{\sum_{i=1}^N x_{(i)} / N} \quad (2)$$

从上面式子可以看出, q_1 是最大 $N - [(1 - p_1)N]$ 个观察值的和在全部观察值的和(即 S)中所占的比例, q_2 是最小 $[p_2N]$ 个观察值的和在全部观察值的和(即 S)中所占的比例。

1.2 “p100/q100”律的概率形式

设总体随机变量为 X , 它的分布为 $F(x)$, 则统计量 q_1 与 q_2 可以分别写成下列形式:

$$q_1 = \frac{p_1 E(X | X > x_{p_1})}{E(X)} \quad (0 < p_1 < 1) \quad (3)$$

$$q_2 = \frac{p_2 E(X | X < x_{p_2})}{E(X)} \quad (0 < p_2 < 1) \quad (4)$$

其中 $E(X)$ 是随机变量 X 的数学期望, $E(X | X > x_{p_1})$ 是在 $X > x_{p_1}$ 条件下随机变量 X 的条件数学期望, $E(X | X < x_{p_2})$ 是在 $X < x_{p_2}$ 条件下随机变量 X 的条件数学期望, $p_1 = P(X > x_{p_1})$, x_{p_1} 称为上 p_1 分位数, $p_2 = P(X < x_{p_2})$, x_{p_2} 称为下 p_2 分位数。如果随机变量是连续型的, 则 $P(X < x) = P(X \leq x)$ 。

1.3 “p100/q100”律的分形形式

如果总体随机变量 X 服从幂函数分布^[3](power-function distribution), 则

$$F(x) = P(X \leq x) = (x/k)^a \quad a > 0 \quad 0 < x < k \quad (5)$$

其中 k 称为尺度参数, a 称为形状参数或指数, 则它的数学期望为

$$E(X) = ak(a+1)^{-1} \quad a > 0 \quad (6)$$

我们证明了幂函数分布的高端截尾(upper truncation)条件下具有分形性质^[4], 即

$$P(X \leq x | X \leq k) = (x/k)^a \quad a > 0 \quad 0 < x < k \quad (7)$$

$$\text{取 } k' = x_{p_2}, \text{ 可得 } E(X | X < x_{p_2}) = \frac{ax_{p_2}}{a+1} \quad a > 0 \quad (8)$$

由(4), (6)和(8)式, 可知

$$q_2 = \frac{p_2 E(X | X < x_{p_2})}{E(X)} = \frac{p_2 ax_{p_2} / (a+1)}{ak / (a+1)} = \frac{p_2 x_{p_2}}{k} \quad (9)$$

由(5)式, 可知

$$p_2 = P(X < x_{p_2}) = \left(\frac{x_{p_2}}{k}\right)^a \quad (10)$$

$$x_{p_2} = kp_2^{1/a} \quad (11)$$

$$\text{由(9)和(11)式, 可知 } q_2 = p_2^{1+1/a} \quad (0 < p_2 < 1) \quad (12)$$

从(12)式中可以看出, 统计量 q_2 与尺度参数 k 无关, 即具有分形性质。 q_2 与不仅具有幂指数关系, 而且是分数。因此, 统计量 q_1 是分形。

$$\text{令 } u = E(X), \text{ 由(6)可得 } a = \frac{u}{k-u} \quad (13)$$

$$\text{因此, (12)式可改写为 } q_2 = p_2^{k/u} \quad (0 < p_2 < 1) \quad (14)$$

Crovelli^[2]认为: 如果总体随机变量 X 服从帕累托分布, 则统计量 q_1 与尺度参数 k 无关, 即具有分形性质。

设帕累托分布^[3](Pareto distribution)为

$$G(x) = P(X \leq x) = 1 - (x/k)^{-a} \quad a > 0 \quad x > k > 0 \quad (15)$$

其中 k 称为尺度参数, a 称为形状参数或指数。则它的数学期望为

$$E(X) = ak(a-1)^{-1} \quad a > 1 \quad (17)$$

帕累托分布在低端截尾(lower truncation)条件下具有分形性质^[5,6], 即

$$P(X \leq x | X \leq k) = 1 - (x/k)^{-a} \quad a > 0 \quad x > k > 0 \quad (18)$$

$$\text{取 } k' = x_{p_1}, \text{ 可得 } E(X | X > x_{p_1}) = \frac{ax_{p_1}}{a-1} \quad a > 0 \quad (19)$$

由(3), (17)和(19)式, 可知

$$q_1 = \frac{p_1 E(X | X > x_{p_1})}{E(X)} = \frac{p_1 ax_{p_1} / (a-1)}{ak / (a-1)} = \frac{p_1 x_{p_1}}{k} \quad (20)$$

由(16)式, 可知

$$p_1 = P(X > x_{p_1}) = \left(\frac{x_{p_1}}{k}\right)^{-a} \quad (21)$$

$$x_{p_1} = kp_1^{-1/a} \quad (22)$$

$$\text{由(20)和(22)式,可知 } q_1 = p_1^{1-1/a} \quad (0 < p_1 < 1) \quad (23)$$

从(23)式可以看出,统计量 q_1 与尺度参数 k 无关,即具有分形性质。 q_1 与不仅具有幂指数关系,而且是分数。因此,统计量 q_1 是分形。

$$\text{令 } u = E(X), \text{ 由(17),可得 } a = \frac{u}{u-k} \quad (24)$$

$$\text{因此,(23)式可改写为 } q_1 = p_1^{k/u} \quad (0 < p_1 < 1) \quad (25)$$

1.4 “20/80”律、“80/20”律及其相应的百分数

在(23)式中,取 $a = 1.16$, 从统计量 q_1 可推出“20/80”律和相应的百分数。 $q_1 = p_1^{1-1/1.16} = p_1^{0.138}$ ($0 < p_1 < 1$)。当 $p_1 = 0.20$ 时, $q_1 = (0.20)^{0.138} = 0.80$ 。此结果表明:占全部人口 20% 高收入的人拥有社会财富的 80%。当 $p_1 = 0.10$ 时, $q_1 = (0.10)^{0.138} = 0.73$ 。当 $p_1 = 0.30$ 时, $q_1 = (0.30)^{0.138} = 0.85$ 。

在(12)式中,取 $a = 0.161$, 从统计量 q_2 可推出“80/20”律和相应的百分数。 $q_2 = p_2^{1+1/0.161} = p_2^{7.211}$ ($0 < p_2 < 1$)。当 $p_2 = 0.80$ 时, $q_2 = (0.80)^{7.211} = 0.20$ 。此结果表明:占全部人口 80% 中低收入的人只拥有社会财富的 20%。当 $p_2 = 0.70$ 时, $q_2 = (0.70)^{7.211} = 0.076$ 。当 $p_2 = 0.90$ 时, $q_2 = (0.90)^{7.211} = 0.478$ 。

2 应用实例

山东胶东地区是中国著名的金矿化集中区和重要的黄金产地,其黄金储量占全国黄金储量的 20%,占山东全省黄金储量 80%。至 1996 年底已探明金矿 81 处,提交 D 级以上储量共 942 391 kg。其中超大型金矿 5 处,储量分别为 84 853, 71 066, 66 720, 60 196, 55 601 kg; 大型金矿 8 处,储量分别为 48 643, 44 000, 43 567, 40 447, 34 496, 29 514, 28 415, 20 440 kg; 中型金矿 20 处,储量分别为 18 543, 17 456, 17 365, 15 656, 15 621, 14 156, 10 284, 9 440, 8 506, 8 187, 7 681, 7 357, 7 052, 6 276, 5 639, 5 616, 5 587, 5 489, 5 402, 5 280 kg; 小型金矿 48 处,它们的储量总和为 117 840 kg(数据来源:山东省胶东地区大一中型金矿成矿规律及找矿

方向的研究研究报告,山东省地质矿产勘查局,1997 年)。

令总体随机变量 X 为金矿床的储量。考虑 81 个大、中、小型金矿床(即子样与观察值),其储量总和 $S = \sum_{i=1}^{81} x_{(i)} = \sum_{i=1}^{81} x_i = 942\,391$ kg, 此时 $N = 81$ 。如果取 $p_1 = 0.16$, 则 $[(1 - p_1)N] + 1 = [0.84 \times 81] + 1 = 69$, 最大 13 个金矿床(即大型超大型金矿床)储量的和 $S_1 = \sum_{i=69}^{81} x_{(i)} = 627\,958$ kg, 统计量 $q_1 = S_1/S = 627\,958/942\,391 = 0.67$ 。这表明大型、超大型金矿床储量占整个储量的 67% 左右,而大型、超大型金矿床个数占整个矿床个数的 16% 左右。如果取 $p_1 = 0.06$, 则 $[(1 - p_1)N] + 1 = [0.94 \times 81] + 1 = 77$, 最大 5 个金矿床(即超大型金矿床)储量的和 $S_1 = \sum_{i=77}^{81} x_{(i)} = 338\,436$ kg, 统计量 $q_1 = S_1/S = 338\,436/942\,391 = 0.36$ 。这表明超大型金矿床储量占整个储量的 36% 左右,而超大型金矿床个数占整个矿床个数的 6% 左右。如果取 $p_1 = 0.10$, 则 $[(1 - p_1)N] + 1 = [0.90 \times 81] + 1 = 73$, 最大 9 个金矿床(即超大型金矿床和部分大型金矿床)储量的和 $S_1 = \sum_{i=73}^{81} x_{(i)} = 515\,093$ kg, 统计量 $q_1 = S_1/S = 515\,093/942\,391 = 0.54$ 。如果取 $p_1 = 0.20$, 则 $[(1 - p_1)N] + 1 = [0.80 \times 81] + 1 = 65$, 最大 17 个金矿床(即超大型金矿床和部分大型金矿床)储量的和 $S_1 = \sum_{i=65}^{81} x_{(i)} = 696\,978$ kg, 统计量 $q_1 = S_1/S = 696\,978/942\,391 = 0.74$ 。

我们根据 81 个大、中、小型金矿床及其储量值,可得到样本的平均值 $u = 942\,391/81 = 11\,648$ 和尺度参数 $k = 2\,803$ 。由(24)式,可得 $a = \frac{u}{u-k} = 11\,648/(11\,648 - 2\,803) = 1.317$ 。

由(23)式,可推出山东胶东地区大、中、小型金矿床的“p100/q100”律表达式和相应的百分数: $q_1 = p_1^{1-1/1.317} = p_1^{0.241}$ ($0 < p_1 < 1$)。当 $p_1 = 0.06$ 时, $q_1 = (0.06)^{0.241} = 0.51$ 。此结果表明:占全部金矿床 6% 超大型金矿床占整个金矿床储量的 51%。当 $p_1 = 0.10$ 时, $q_1 = (0.10)^{0.241} = 0.57$ 。当 $p_1 = 0.16$ 时, $q_1 = (0.16)^{0.241} = 0.64$ 。当 $p_1 = 0.20$ 时, $q_1 = (0.20)^{0.241} = 0.68$ 。

通过分析,山东胶东地区金矿床储量总体分布服从参数 $a = 1.317$ 的帕累托分布,比较上面理论计算的结果与样本观察值计算的结果大小,如当 $p_1 = 0.06$ 时, 0.51 (理论数) 大于 0.36 (经验数),可预测该地区可能存在 1 个或 2 个大型金矿床。在数理统计中,金矿床储量可以认为是随机变量,金矿床储量数据实际上是观察值,这些数据在整体上基本服从幂函数分布(参见文献[7])。

3 结论

(1) 幂函数分布在高端截尾(upper truncation)条件下具有分形性质,即尺度不变的特征,证明了统计量 q_2 具有分形性质,即它与尺度参数 k 无关。如果总体的随机变量 X 服从幂函数分布,则从“p100/q100”律的概率形式推出“p100/q100”律的分形式,此时参数 q 仅仅是参数 p 和形状参数 a 的乘幂函数。

(2) “80/20”律是“p100/q100”律的一种特殊分形式。当 $a = 1.16$, $p = 0.20$ 时, $q = 0.80$ 。

(3) 山东胶东地区金矿床储量总体分布服从帕累托分布,通过分析可以得到该地区大、中、小型金矿床的“p100/q100”律表达式和相应的百分数。比较理论计算结果与样本观察值计算结果的大小,可预测该地区可能存在 1 个或 2 个大型金矿床。

“p100/q100”律产生的原因在于:地球作为一个耗散形态,在漫长的岁月中经历了多次非均匀非线性的地质作用,地下的岩性物性均显现出很强的非均匀非线性特性。地质环境中非线性过程的相互作用是造成地壳元素与矿化不均匀分布的根本原因,并导致了元素含量与大小矿床分布的分形结构。

(4) “p100/q100”律不仅应用于一些地质数据(如矿床的数量与储量)的检查和预测,而且可应用于其他方面,例如资源、环境和灾害的评价。

参考文献:

- [1] TURLA P A, HAWKINS K L. *Time Management Made Easy* [M]. New York: E. P. Dutton, 1938.
- [2] CROVELLI R A. The Generalized 20/80 law using probabilistic fractals applied to petroleum field size [J]. *Nonrenewable Resources*, 1995, 4(3): 233-241.
- [3] JOHN N L, KOTZ S. *Continuous Univariate Distributions-I* [M]. New York: John Wiley & Sons Inc, 1970.
- [4] 申维,赵鹏大. 分形统计模型的理论研究及其在地质学中的应用 [J]. *地质科学*, 1998, 33(2): 234-243.
- [5] CHRISTOPHER C B, la PAUL R P. *Fractals in Petroleum Geology and Earth Processes* [M]. New York: Plenum Press, 1995.
- [6] MANDELBROT B B. Forecasts of future prices, unbiased markets, and “martingale” models [J]. *Journal of Business*, University of Chicago, 1966, 39: 242-255.
- [7] 申维. 分形模型与大型、超大型矿床 [J]. *中国数学地质*, 1998, 9: 52-58.

ASSESSMENT OF GOLD ORE RESOURCES POTENTIAL IN EASTERN SHANDONG BY THE “p100/q100” LAW

SHEN Wei^{1,2}

(1. Research Centre of Applied New Techniques to Land and Resources, China University of Geosciences, Beijing 100083, China;

2. Institute of Geochemistry, Chinese Academy of Sciences, Guiyang 550002, China)

Abstract: This paper shows that the statistic q possesses the fractal property, which does not depend upon the original scale parameter k , from the power-function distribution of the fractal property of scaling invariance under upper truncation. The fractal form of the “p100/q100” law can be obtained from the probabilistic form of the “p100/q100” law under the random variables X being of the power-function distribution. The “80/20” law is a special case of the “p100/q100” law. When $a = 1.16$, $p = 0.20$, the statistic q is 0.80. The expression and the corresponding percentage of “p100/q100” law are got by the gold ore reserve data in eastern Shandong. By comparing theoretical data with empirical results, it shows that there could be one or two large-sized gold mineral deposit in eastern Shandong.

Key words: fractal; Pareto law; power-function distribution; gold deposit