U.S. DEPARTMENT OF THE INTERIOR

U.S. GEOLOGICAL SURVEY

# Fractal Lognormal Percentage Assessment of Petroleum Field Sizes in a Play—Application of a Generalized 20/80 Law

by

Robert A. Crovelli[1]

Open-File Report 95-646

This report is preliminary and has not been reviewed for conformity with U.S. Geological Survey editorial standards and stratigraphic nomenclature.

[1]U.S. Geological Survey, Box 25046, MS 971
Denver Federal Center, Denver, Colorado 80225 U.S.A.

1995

## ABSTRACT

The 20/80 law is a heuristic law that has evolved over the years into the following rule of thumb for many populations: 20% of the population accounts for 80% of the total value. This law states quantitatively that often a relatively small portion of a population accounts for most of the total value of the population. This principle is certainly applicable in the case of petroleum field sizes in plays where invariably the few largest fields have an overwhelming amount of the total resources.

The general $p100/q100$ law in statistical form is defined with the statistic $q$ as a function of $p$ where $p$ is the population proportion and $q$ is the proportion of total value. The general $p100/q100$ law in probabilistic form is defined with the parameter $q$ as a function of $p$ for any probability distribution that models the population distribution. Using the lognormal distribution, the $p100/q100$ law in lognormal form is derived with the lognormal $q$ being a fractal, where $q$ possesses the scale invariance property.

The $p100/q100$ law in lognormal form was applied to data on petroleum field sizes in a test play to obtain: $p100$% of the oil fields account for $q100$% of the total oil resources of the fields. The theoretical percentages of total oil resources using the lognormal $q$ are extremely close to the empirical percentages from the petroleum field size data using the statistic $q$. For example, 20% of the 175 oil fields account for 45.87% of the total oil resources of the fields if we use the lognormal $q$, or for 48.06% if we use the statistic $q$.

## INTRODUCTION

Vilfredo Pareto (1848-1923), an Italian engineer turned economist and sociologist, was a pioneer in econometrics, applying the statistical methods of the physical sciences to the social sciences. Pareto discovered that, generally, a relatively small portion of a population accounts for most of the total value of the population. That is, in any set of elements, the critical elements usually constitute a minority of the set (Douglass and Douglass, 1993). Pareto's principle is certainly applicable in the case of petroleum field sizes in a play where the few largest fields have an overwhelming amount of the total resources.

As originally proposed, Pareto's law (Pareto, 1897; Turla and Hawkins, 1983) was an empirical relation describing the distribution of income among the population of a country; viz.,

20 percent of the people in Italy owned 80 percent of the wealth.

From this principle, the heuristic "20/80 law" has evolved over the years into the following simple rule of thumb, which is applicable to many populations:

20% of the population accounts for 80% of the total value.

Numerous applications of this law exist in the management sciences. For example,

20% of the customers account for 80% of the sales;

20% of the employees account for 80% of the work; and,

20% of the components account for 80% of the cost.

The principle is observed in other disciplines as well. The objective of the present paper is to generalize "the 20/80 law" using the lognormal distribution and apply the generalization to data on petroleum field sizes in a test play to obtain: $p100\%$ of the oil fields account for $q100\%$ ($q$ is a function of $p$) of the total oil resources of the fields. Fractal lognormal percentage theory is developed first, then the theory is applied to an example of petroleum field size data.

## FRACTAL LOGNORMAL PERCENTAGE THEORY

### THE *p*100/*q*100 LAW IN STATISTICAL FORM

Consider a population of $N$ values for a random variable $X$ arranged in increasing order of magnitude:

$$x_1, x_2, \ldots, x_N.$$

The total value of the population is:

$$\sum_{i=1}^{N} x_i .$$

Given a proportion $p$ $(0 \leq p \leq 1)$ of largest values of the population, the partial sum of the $p100\%$ of largest values of the population is:

$$\sum_{i=[(1-p)N]+1}^{N} x_i \, ,$$

where $[\cdot]$ is the greatest integer value function.

Let the proportion of the total value of the population accounted for by $p100\%$ of the population be given by $q$ $(0 \leq q \leq 1)$, where the statistic $q$ is defined as:

$$q = \sum_{i=[(1-p)N]+1}^{N} x_i \bigg/ \sum_{i=1}^{N} x_i .$$

An alternative form is:

$$q = \frac{p \sum_{i=[(1-p)N]+1}^{N} x_i \bigg/ pN}{\sum_{i=1}^{N} x_i \bigg/ N} .$$

Summarizing, $p$ is a proportion (or fraction) of the population, and $p100\%$ is a percentage of the population; whereas, $q$ is a proportion (or fraction) of the total value, $q100\%$ is a percentage of the total value, and $p100\%$ of the population accounts for $q100\%$ of the total value.

## THE *p*100/*q*100 LAW IN PROBABILISTIC FORM

Given a population of values of a random variable $X$ having a probability distribution that models the population distribution, then $p100\%$ of the population values account for $q100\%$ of the total value, where the parameter $q$ is called the proportion of total value and defined in terms of conditional expectation, $E(X|\cdot)$, as follows.

$$q = \frac{pE(X \mid X > x_p)}{E(X)}, \quad 0 \leq p \leq 1,$$

with $p = P(X > x_p)$ and $x_p$ is called the $p100$th fractile.

Various probability distributions could be used as models for the population distribution. I will consider the lognormal probability distribution. The percentage theory based on the Pareto distribution has been developed and applied to petroleum field size data (Crovelli, 1995).

## THE LOGNORMAL PROBABILITY DISTRIBUTION

The lognormal distribution is discussed in detail in Aitchison and Brown (1957), Johnson and others (1994), and Crow and Shimizu (1988). A nonnegative random variable $X$ has a lognormal distribution if the random variable $Y = \ln X$ has a normal distribution.

The lognormal probability density function of $X$ is

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma\,x} \exp \frac{-(\ln x - \mu)^2}{2\sigma^2}, \qquad x > 0$$

where $\sigma > 0$ is a shape parameter and $\mu$ is a scale parameter (Law and Kelton, 1991, p. 337).

The expected value or mean of $X$ is

$$E(X) = e^{\mu + \sigma^2/2}$$

The variance of $X$ is

$$V(X) = e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right)$$

The scale parameter $\mu$ and the shape parameter $\sigma$ of the lognormal distribution are computed from the following formulas:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} \ln x_i$$

and

$$\sigma = \left[ \frac{N\sum_{i=1}^{N} (\ln x_i)^2 - \left( \sum_{i=1}^{N} \ln x_i \right)^2}{N^2} \right]^{1/2}.$$

The following notation will be used throughout the remainder of the paper:

$X \sim$ lognormal (shape parameter $\sigma$, scale parameter $\mu$),

where the symbol ~ is read "is distributed as."

## THE LOGNORMAL DISTRIBUTION HAS A FRACTAL PROPERTY

If $X$ ~ lognormal (shape parameter $\sigma$, scale parameter $\mu$), let

$$X' = cX, \qquad c > 0.$$

The fractal property of $X$ under a positive multiplicative constant results in:

$X'$ ~ lognormal (shape parameter $\sigma$, scale parameter $\mu + \ln c$ ).

Note that the distributions of $X$ and $X'$ are identical (with the same shape parameter) except for scale. A probabilistic fractal is a probability distribution that is invariant except for scale (scale invariant or self-similar) under a given transformation (Mandelbrot, 1983, p. 343). Under a positive multiplicative constant, the lognormal distribution remains a lognormal distribution with fixed $\sigma$ and changed $\mu$.

## THE *p*100/*q*100 LAW IN LOGNORMAL FORM

Consider the $p100/q100$ law in probabilistic form when the population distribution is modeled as a lognormal distribution; i.e.,

$$q = \frac{pE(X \mid X > x_p)}{E(X)}, \quad 0 \le p \le 1$$

when $X$ ~ lognormal (shape parameter $\sigma$, scale parameter $\mu$).

Recall that the expected value of $X$ is

$$E(X) = e^{\mu + \sigma^2/2}.$$

It can be shown (Johnson and others, 1994) that the conditional expectation is given by

$$E(X \mid X > x_p) = e^{\mu+\sigma^2/2} \frac{R((\ln x_p - \mu - \sigma^2)/\sigma)}{R((\ln x_p - \mu)/\sigma)}$$

where $R$ denotes the complementary cumulative distribution function of a standard normal random variable $Z$, i.e., $R(z) \equiv P(Z > z)$.

We define the following notation:

$$R(z_p) = P(Z > z_p) = p$$

where $z_p$ is called the $p100$th fractile of $Z$.

Because

$$R((\ln x_p - \mu)/\sigma) = R(z_p) = p$$

and

$$R((\ln x_p - \mu - \sigma^2)/\sigma) = R(z_p - \sigma)$$

we have

$$E(X \mid X > x_p) = e^{\mu+\sigma^2/2} \frac{R(z_p - \sigma)}{p}$$

Substituting into the general formula for $q$ and simplifying, we get

$$q = R(z_p - \sigma), \quad 0 < p < 1$$

or

$$q = 1 - \Phi(z_p - \sigma), \quad 0 < p < 1$$

where $\Phi$ denotes the standard normal cumulative distribution function.

Observe that if $p \to 0$, then $q \to 0$, and if $p \to 1$, then $q \to 1$.Thus we define $q = 0$ if $p = 0$, and $q = 1$ if $p = 1$.The special case when $p = 0.5$ and $z_p = 0$ yields

$$q = R(-\sigma) = 1 - \Phi\,(-\sigma)$$

The parameter $q$ has the very remarkable property of being scale-free; i.e., $q$ does not depend upon the scale parameter $\mu$. Because the parameter $q$ is scale-free, $q$ stays the same for all lognormal distributions with the same shape parameter $\sigma$, as in the case of the fractal property of the lognormal distribution under a positive multiplicative constant. Because it is scale invariant, the proportion $q$ can be considered to be a parametric fractal.

The proportion $q$ of the $p100/q100$ law in statistical form is called the statistic $q$, in probabilistic form it is called the parameter $q$, and in lognormal form it is called the lognormal $q$.

## APPLICATION OF THEORY TO PETROLEUM FIELD SIZES

The fractal lognormal percentage theory is applied in this section to petroleum field size data. An illustrative example consists of 175 fields producing either oil, or oil and gas, in million barrels of oil equivalent (MMBOE) known recovery from the northern Michigan Silurian reef play. The source of data is the Significant Oil and Gas Fields of the United States Data Base, a product of NRG Associates, Inc. (1992). The version used includes discoveries from the years up to and including 1990. Known recovery refers to the sum of cumulative production plus reserves. Included in the NRG files are those fields with at least 1 MMBOE of known recovery and also those smaller, but expected to be revised to at least 1 MMBOE. Let $X$ be the size of the oil or oil and gas field in MMBOE.

This new probabilistic methodology for percentage assessment lends itself as an ideal spreadsheet software application. A frequency distribution of the 175 petroleum field sizes (in MMBOE) is displayed in Figure 1. Because the plotted random variable is the logarithm of MMBOE, Ln (MMBOE), the bell-shaped distribution suggests that the petroleum resources are approximately distributed as a lognormal distribution.

Spreadsheet software was used to make various calculations involving the petroleum field size data. Table 1 is a spreadsheet with 175 rows corresponding to the 175 petroleum fields. The data are listed and several calculations are made, including those necessary for the statistic $q$. The lognormal parameters were computed: scale parameter $\mu = 0.3080$ and shape parameter $\sigma = 0.7380$.

Table 2 is a spreadsheet of the calculations from the petroleum field size data that are necessary for the lognormal $q$ as a function of various values of $p$. The formula for the lognormal $q$ in the case of the petroleum field size data is the following:

$$q = R(z_p - 0.7380) = 1 - \Phi(z_p - 0.7380), \quad 0 < p < 1$$

Therefore, for any specified value of $p$, we can obtain: $p100\%$ of the oil fields account for $q100\%$ of the total oil resources of the fields. The fractal lognormal percentage assessment of oil resources is summarized in Table 3. The corresponding graph of the summary is given in Figure 2.

Note that in Table 3 and Figure 2, the theoretical percentages of total oil resources using the lognormal $q$ are extremely close to the empirical percentages from the petroleum field size data using the statistic $q$. For example, 20% of the 175 oil fields account for 45.87% of the total oil resources of the fields if we use the lognormal $q$, or for 48.06% if we use the statistic $q$. Plays presumably vary in their degree of heterogeneity, and this can be quantitatively captured by this method. The method also forms a basis for comparing plays of different geologic styles, and similar geologic styles.

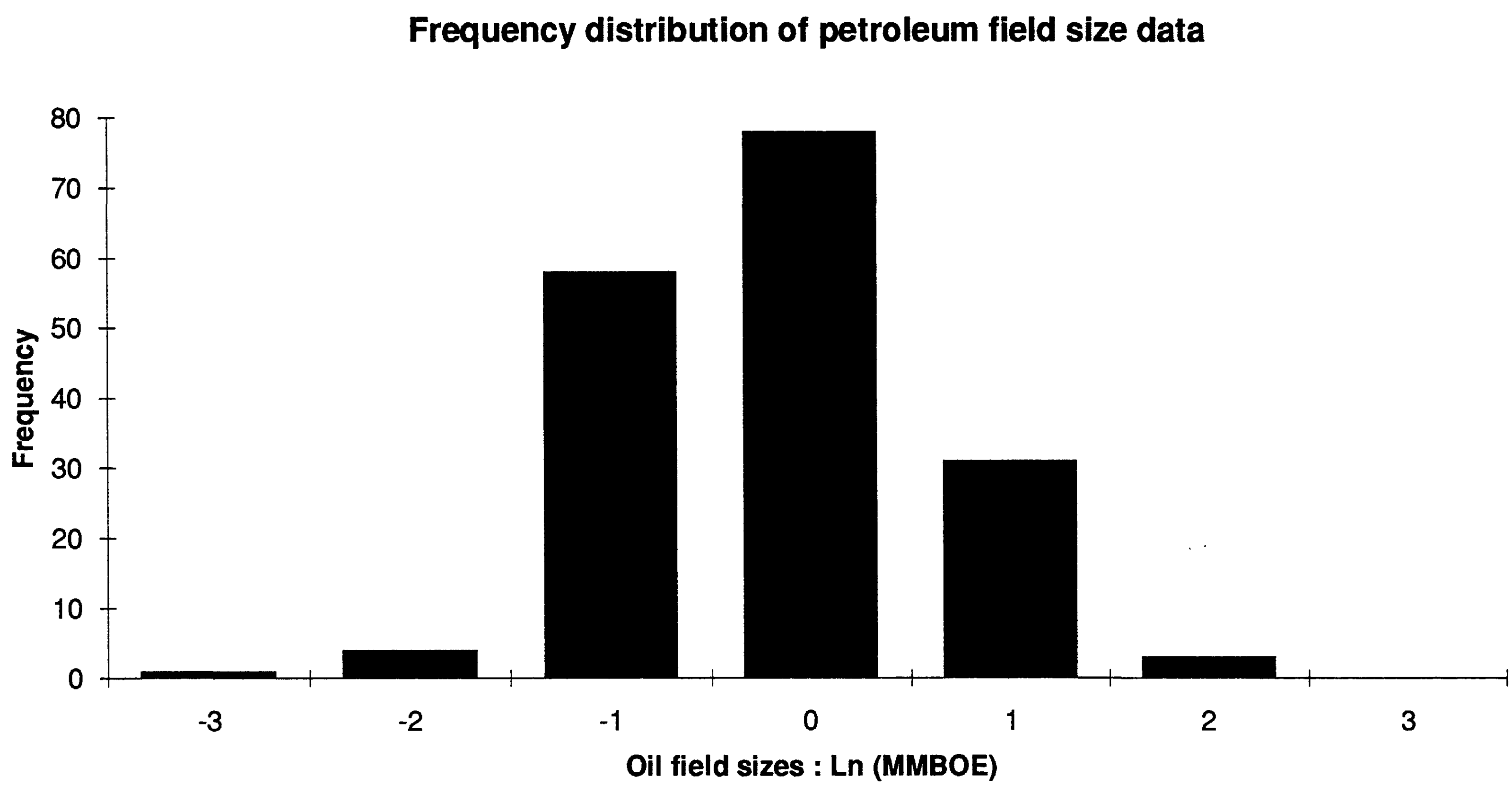


Figure 1. Frequency distribution of 175 petroleum field sizes (in million barrels of oil equivalent). Fields are in the northern Michigan Silurian reef play. Data and bin frequencies are included with Table 1.

Table 1. Spreadsheet of petroleum field size data of Figure 1 and calculations for the statistic $q$.

| $n$ | Field size (MMBOE) | Ln of Field size | Partial sum of field sizes | Statistic $q$ Partial/Total |
|---|---|---|---|---|
| 1 | 14.25 | 2.6568 | 14.25 | 0.0449 |
| 2 | 12.00 | 2.4849 | 26.25 | 0.0827 |
| 3 | 7.80 | 2.0541 | 34.05 | 0.1072 |
| 4 | 5.40 | 1.6864 | 39.45 | 0.1242 |
| 5 | 5.20 | 1.6487 | 44.65 | 0.1406 |
| 6 | 5.15 | 1.6390 | 49.80 | 0.1568 |
| 7 | 5.10 | 1.6292 | 54.90 | 0.1729 |
| 8 | 4.60 | 1.5261 | 59.50 | 0.1873 |
| 9 | 4.60 | 1.5261 | 64.10 | 0.2018 |
| 10 | 4.50 | 1.5041 | 68.60 | 0.2160 |
| 11 | 4.50 | 1.5041 | 73.10 | 0.2302 |
| 12 | 4.20 | 1.4351 | 77.30 | 0.2434 |
| 13 | 3.95 | 1.3737 | 81.25 | 0.2558 |
| 14 | 3.90 | 1.3610 | 85.15 | 0.2681 |
| 15 | 3.80 | 1.3350 | 88.95 | 0.2801 |
| 16 | 3.70 | 1.3083 | 92.65 | 0.2917 |
| 17 | 3.65 | 1.2947 | 96.30 | 0.3032 |
| 18 | 3.60 | 1.2809 | 99.90 | 0.3145 |
| 19 | 3.60 | 1.2809 | 103.50 | 0.3259 |
| 20 | 3.45 | 1.2384 | 106.95 | 0.3367 |
| 21 | 3.40 | 1.2238 | 110.35 | 0.3474 |
| 22 | 3.40 | 1.2238 | 113.75 | 0.3582 |
| 23 | 3.30 | 1.1939 | 117.05 | 0.3685 |
| 24 | 3.30 | 1.1939 | 120.35 | 0.3789 |
| 25 | 3.30 | 1.1939 | 123.65 | 0.3893 |
| 26 | 3.15 | 1.1474 | 126.80 | 0.3992 |
| 27 | 3.00 | 1.0986 | 129.80 | 0.4087 |
| 28 | 3.00 | 1.0986 | 132.80 | 0.4181 |
| 29 | 3.00 | 1.0986 | 135.80 | 0.4276 |
| 30 | 2.95 | 1.0818 | 138.75 | 0.4369 |
| 31 | 2.90 | 1.0647 | 141.65 | 0.4460 |
| 32 | 2.80 | 1.0296 | 144.45 | 0.4548 |
| 33 | 2.75 | 1.0116 | 147.20 | 0.4635 |
| 34 | 2.75 | 1.0116 | 149.95 | 0.4721 |
| 35 | 2.70 | 0.9933 | 152.65 | 0.4806 |
| 36 | 2.60 | 0.9555 | 155.25 | 0.4888 |
| 37 | 2.60 | 0.9555 | 157.85 | 0.4970 |
| 38 | 2.60 | 0.9555 | 160.45 | 0.5052 |
| 39 | 2.45 | 0.8961 | 162.90 | 0.5129 |
| 40 | 2.40 | 0.8755 | 165.30 | 0.5205 |
| 41 | 2.40 | 0.8755 | 167.70 | 0.5280 |

| *Bin* | *Frequency* |
|---|---|
| –3 | 1 |
| –2 | 4 |
| –1 | 58 |
| 0 | 78 |
| 1 | 31 |
| 2 | 3 |
| 3 | 0 |

Table 1. (cont.)

| *n* | **Field size (MMBOE)** | **Ln of Field size** | **Partial sum of Field sizes** | **Statistic *q* Partial/Total** |
|---|---|---|---|---|
| 42 | 2.40 | 0.8755 | 170.10 | 0.5356 |
| 43 | 2.38 | 0.8671 | 172.48 | 0.5431 |
| 44 | 2.35 | 0.8544 | 174.83 | 0.5505 |
| 45 | 2.35 | 0.8544 | 177.18 | 0.5579 |
| 46 | 2.25 | 0.8109 | 179.43 | 0.5650 |
| 47 | 2.10 | 0.7419 | 181.53 | 0.5716 |
| 48 | 2.08 | 0.7324 | 183.61 | 0.5781 |
| 49 | 2.00 | 0.6931 | 185.61 | 0.5844 |
| 50 | 2.00 | 0.6931 | 187.61 | 0.5907 |
| 51 | 1.95 | 0.6678 | 189.56 | 0.5969 |
| 52 | 1.90 | 0.6419 | 191.46 | 0.6028 |
| 53 | 1.90 | 0.6419 | 193.36 | 0.6088 |
| 54 | 1.85 | 0.6152 | 195.21 | 0.6146 |
| 55 | 1.82 | 0.5988 | 197.03 | 0.6204 |
| 56 | 1.80 | 0.5878 | 198.83 | 0.6260 |
| 57 | 1.75 | 0.5596 | 200.58 | 0.6315 |
| 58 | 1.75 | 0.5596 | 202.33 | 0.6371 |
| 59 | 1.71 | 0.5365 | 204.04 | 0.6424 |
| 60 | 1.70 | 0.5306 | 205.74 | 0.6478 |
| 61 | 1.68 | 0.5188 | 207.42 | 0.6531 |
| 62 | 1.68 | 0.5188 | 209.10 | 0.6584 |
| 63 | 1.66 | 0.5068 | 210.76 | 0.6636 |
| 64 | 1.65 | 0.5008 | 212.41 | 0.6688 |
| 65 | 1.65 | 0.5008 | 214.06 | 0.6740 |
| 66 | 1.60 | 0.4700 | 215.66 | 0.6790 |
| 67 | 1.60 | 0.4700 | 217.26 | 0.6841 |
| 68 | 1.60 | 0.4700 | 218.86 | 0.6891 |
| 69 | 1.60 | 0.4700 | 220.46 | 0.6941 |
| 70 | 1.55 | 0.4383 | 222.01 | 0.6990 |
| 71 | 1.50 | 0.4055 | 223.51 | 0.7037 |
| 72 | 1.45 | 0.3716 | 224.96 | 0.7083 |
| 73 | 1.42 | 0.3507 | 226.38 | 0.7128 |
| 74 | 1.40 | 0.3365 | 227.78 | 0.7172 |
| 75 | 1.40 | 0.3365 | 229.18 | 0.7216 |
| 76 | 1.40 | 0.3365 | 230.58 | 0.7260 |
| 77 | 1.35 | 0.3001 | 231.93 | 0.7303 |
| 78 | 1.35 | 0.3001 | 233.28 | 0.7345 |
| 79 | 1.30 | 0.2624 | 234.58 | 0.7386 |
| 80 | 1.30 | 0.2624 | 235.88 | 0.7427 |
| 81 | 1.30 | 0.2624 | 237.18 | 0.7468 |
| 82 | 1.30 | 0.2624 | 238.48 | 0.7509 |

Table 1. (cont.)

| *n* | Field size (MMBOE) | Ln of Field size | Partial sum of Field sizes | Statistic *q* Partial/Total |
|---|---|---|---|---|
| 83 | 1.30 | 0.2624 | 239.78 | 0.7550 |
| 84 | 1.28 | 0.2469 | 241.06 | 0.7590 |
| 85 | 1.27 | 0.2390 | 242.33 | 0.7630 |
| 86 | 1.27 | 0.2390 | 243.60 | 0.7670 |
| 87 | 1.26 | 0.2311 | 244.86 | 0.7710 |
| 88 | 1.22 | 0.1989 | 246.08 | 0.7748 |
| 89 | 1.22 | 0.1989 | 247.30 | 0.7787 |
| 90 | 1.21 | 0.1906 | 248.51 | 0.7825 |
| 91 | 1.20 | 0.1823 | 249.71 | 0.7862 |
| 92 | 1.20 | 0.1823 | 250.91 | 0.7900 |
| 93 | 1.20 | 0.1823 | 252.11 | 0.7938 |
| 94 | 1.18 | 0.1655 | 253.29 | 0.7975 |
| 95 | 1.18 | 0.1655 | 254.47 | 0.8012 |
| 96 | 1.17 | 0.1570 | 255.64 | 0.8049 |
| 97 | 1.16 | 0.1484 | 256.80 | 0.8086 |
| 98 | 1.16 | 0.1484 | 257.96 | 0.8122 |
| 99 | 1.15 | 0.1398 | 259.11 | 0.8158 |
| 100 | 1.15 | 0.1398 | 260.26 | 0.8195 |
| 101 | 1.15 | 0.1398 | 261.41 | 0.8231 |
| 102 | 1.15 | 0.1398 | 262.56 | 0.8267 |
| 103 | 1.14 | 0.1310 | 263.70 | 0.8303 |
| 104 | 1.13 | 0.1222 | 264.83 | 0.8338 |
| 105 | 1.10 | 0.0953 | 265.93 | 0.8373 |
| 106 | 1.08 | 0.0770 | 267.01 | 0.8407 |
| 107 | 1.06 | 0.0583 | 268.07 | 0.8440 |
| 108 | 1.02 | 0.0198 | 269.09 | 0.8473 |
| 109 | 1.02 | 0.0198 | 270.11 | 0.8505 |
| 110 | 1.00 | 0.0000 | 271.11 | 0.8536 |
| 111 | 1.00 | 0.0000 | 272.11 | 0.8568 |
| 112 | 1.00 | 0.0000 | 273.11 | 0.8599 |
| 113 | 0.99 | –0.0101 | 274.10 | 0.8630 |
| 114 | 0.96 | –0.0408 | 275.06 | 0.8661 |
| 115 | 0.96 | –0.0408 | 276.02 | 0.8691 |
| 116 | 0.96 | –0.0408 | 276.98 | 0.8721 |
| 117 | 0.95 | –0.0513 | 277.93 | 0.8751 |
| 118 | 0.95 | –0.0513 | 278.88 | 0.8781 |
| 119 | 0.93 | –0.0726 | 279.81 | 0.8810 |
| 120 | 0.93 | –0.0726 | 280.74 | 0.8839 |
| 121 | 0.93 | –0.0726 | 281.67 | 0.8869 |
| 122 | 0.91 | –0.0943 | 282.58 | 0.8897 |
| 123 | 0.91 | –0.0943 | 283.49 | 0.8926 |

Table 1. (cont.)

| *n* | Field size (MMBOE) | Ln of Field size | Partial sum of Field sizes | Statistic *q* Partial/Total |
|---|---|---|---|---|
| 124 | 0.90 | −0.1054 | 284.39 | 0.8954 |
| 125 | 0.90 | −0.1054 | 285.29 | 0.8983 |
| 126 | 0.90 | −0.1054 | 286.19 | 0.9011 |
| 127 | 0.88 | −0.1278 | 287.07 | 0.9039 |
| 128 | 0.87 | −0.1393 | 287.94 | 0.9066 |
| 129 | 0.86 | −0.1508 | 288.80 | 0.9093 |
| 130 | 0.85 | −0.1625 | 289.65 | 0.9120 |
| 131 | 0.85 | −0.1625 | 290.50 | 0.9147 |
| 132 | 0.84 | −0.1744 | 291.34 | 0.9173 |
| 133 | 0.83 | −0.1863 | 292.17 | 0.9199 |
| 134 | 0.83 | −0.1863 | 293.00 | 0.9225 |
| 135 | 0.82 | −0.1985 | 293.82 | 0.9251 |
| 136 | 0.80 | −0.2231 | 294.62 | 0.9276 |
| 137 | 0.80 | −0.2231 | 295.42 | 0.9302 |
| 138 | 0.80 | −0.2231 | 296.22 | 0.9327 |
| 139 | 0.78 | −0.2485 | 297.00 | 0.9351 |
| 140 | 0.78 | −0.2485 | 297.78 | 0.9376 |
| 141 | 0.77 | −0.2614 | 298.55 | 0.9400 |
| 142 | 0.75 | −0.2877 | 299.30 | 0.9424 |
| 143 | 0.75 | −0.2877 | 300.05 | 0.9447 |
| 144 | 0.75 | −0.2877 | 300.80 | 0.9471 |
| 145 | 0.74 | −0.3011 | 301.54 | 0.9494 |
| 146 | 0.73 | −0.3147 | 302.27 | 0.9517 |
| 147 | 0.71 | −0.3425 | 302.98 | 0.9540 |
| 148 | 0.70 | −0.3567 | 303.68 | 0.9562 |
| 149 | 0.68 | −0.3857 | 304.36 | 0.9583 |
| 150 | 0.68 | −0.3857 | 305.04 | 0.9605 |
| 151 | 0.68 | −0.3857 | 305.72 | 0.9626 |
| 152 | 0.66 | −0.4155 | 306.38 | 0.9647 |
| 153 | 0.66 | −0.4155 | 307.04 | 0.9668 |
| 154 | 0.65 | −0.4308 | 307.69 | 0.9688 |
| 155 | 0.65 | −0.4308 | 308.34 | 0.9708 |
| 156 | 0.65 | −0.4308 | 308.99 | 0.9729 |
| 157 | 0.63 | −0.4620 | 309.62 | 0.9749 |
| 158 | 0.63 | −0.4620 | 310.25 | 0.9769 |
| 159 | 0.62 | −0.4780 | 310.87 | 0.9788 |
| 160 | 0.60 | −0.5108 | 311.47 | 0.9807 |
| 161 | 0.60 | −0.5108 | 312.07 | 0.9826 |
| 162 | 0.58 | −0.5447 | 312.65 | 0.9844 |
| 163 | 0.58 | −0.5447 | 313.23 | 0.9862 |
| 164 | 0.55 | −0.5978 | 313.78 | 0.9880 |

Table 1. (cont.)

| $n$ | Field size (MMBOE) | Ln of Field size | Partial sum of Field sizes | Statistic $q$ Partial/Total |
|---|---|---|---|---|
| 165 | 0.50 | −0.6931 | 314.28 | 0.9895 |
| 166 | 0.49 | −0.7133 | 314.77 | 0.9911 |
| 167 | 0.45 | −0.7985 | 315.22 | 0.9925 |
| 168 | 0.40 | −0.9163 | 315.62 | 0.9938 |
| 169 | 0.40 | −0.9163 | 316.02 | 0.9950 |
| 170 | 0.38 | −0.9676 | 316.40 | 0.9962 |
| 171 | 0.34 | −1.0788 | 316.74 | 0.9973 |
| 172 | 0.33 | −1.1087 | 317.07 | 0.9983 |
| 173 | 0.24 | −1.4271 | 317.31 | 0.9991 |
| 174 | 0.18 | −1.7148 | 317.49 | 0.9997 |
| 175 | 0.11 | −2.2073 | 317.60 | 1.0000 |

| | | |
|---|---|---|
| 0.11 | −2.2073 | <== Min |
| 14.25 | 2.6568 | <== Max |
| 317.60 | 53.8958 | <== Total |
| 1.8149 | 0.3080 | <== Mean (μ) |
| 1.7391 | 0.7380 | <== Standard Deviation (σ) |

Table 2. Spreadsheet of calculations from petroleum field size data of Figure 1 for the lognormal $q$ as a function of various values of $p$. NORMSINV is the spreadsheet function that computes $z$; NORMSDIST is the function that computes $1 - q$.

| $p$<br>Pop. Prop. | $(p)(N)$<br>$(p)(175)$ | $1 - p$ | $z$<br>NORMSINV | $z - \sigma$<br>$(z - 0.7380)$ | $1 - q$<br>NORMSDIST | lognormal $q$<br>(theoretical) | statistic $q$<br>(empirical) |
|---|---|---|---|---|---|---|---|
| 0.01 | 1.75 | 0.99 | 2.3263 | 1.5884 | 0.9439 | 0.0561 | 0.0732 |
| 0.025 | 4.375 | 0.975 | 1.9600 | 1.2220 | 0.8891 | 0.1109 | 0.1304 |
| 0.05 | 8.75 | 0.95 | 1.6449 | 0.9069 | 0.8178 | 0.1822 | 0.1982 |
| 0.1 | 17.5 | 0.9 | 1.2816 | 0.5436 | 0.7066 | 0.2934 | 0.3089 |
| 0.2 | 35 | 0.8 | 0.8416 | 0.1036 | 0.5413 | 0.4587 | 0.4806 |
| 0.3 | 52.5 | 0.7 | 0.5244 | −0.2136 | 0.4154 | 0.5846 | 0.6058 |
| 0.4 | 70 | 0.6 | 0.2533 | −0.4846 | 0.3140 | 0.6860 | 0.6990 |
| 0.5 | 87.5 | 0.5 | 0.0000 | −0.7380 | 0.2303 | 0.7697 | 0.7729 |
| 0.6 | 105 | 0.4 | −0.2533 | −0.9913 | 0.1608 | 0.8392 | 0.8373 |
| 0.7 | 122.5 | 0.3 | −0.5244 | −1.2624 | 0.1034 | 0.8966 | 0.8912 |
| 0.8 | 140 | 0.2 | −0.8416 | −1.5796 | 0.0571 | 0.9429 | 0.9376 |
| 0.9 | 157.5 | 0.1 | −1.2816 | −2.0195 | 0.0217 | 0.9783 | 0.9759 |

Table 3. Summary of fractal lognormal percentage assessment of oil resources in the northern Michigan Silurian reef play

| Percentage of oil fields | Percentage of total oil resources $q100\%$ | |
|---|---|---|
| $p100\%$ | Theoretical (lognormal $q$) | Empirical (statistic $q$) |
| 0 | 0 | 0 |
| 1 | 5.61 | 7.32 |
| 2.5 | 11.09 | 13.04 |
| 5 | 18.22 | 19.82 |
| 10 | 29.34 | 30.89 |
| 20 | 45.87 | 48.06 |
| 30 | 58.46 | 60.58 |
| 40 | 68.60 | 69.90 |
| 50 | 76.97 | 77.29 |
| 60 | 83.92 | 83.73 |
| 70 | 89.66 | 89.12 |
| 80 | 94.29 | 93.76 |
| 90 | 97.83 | 97.59 |
| 100 | 100 | 100 |

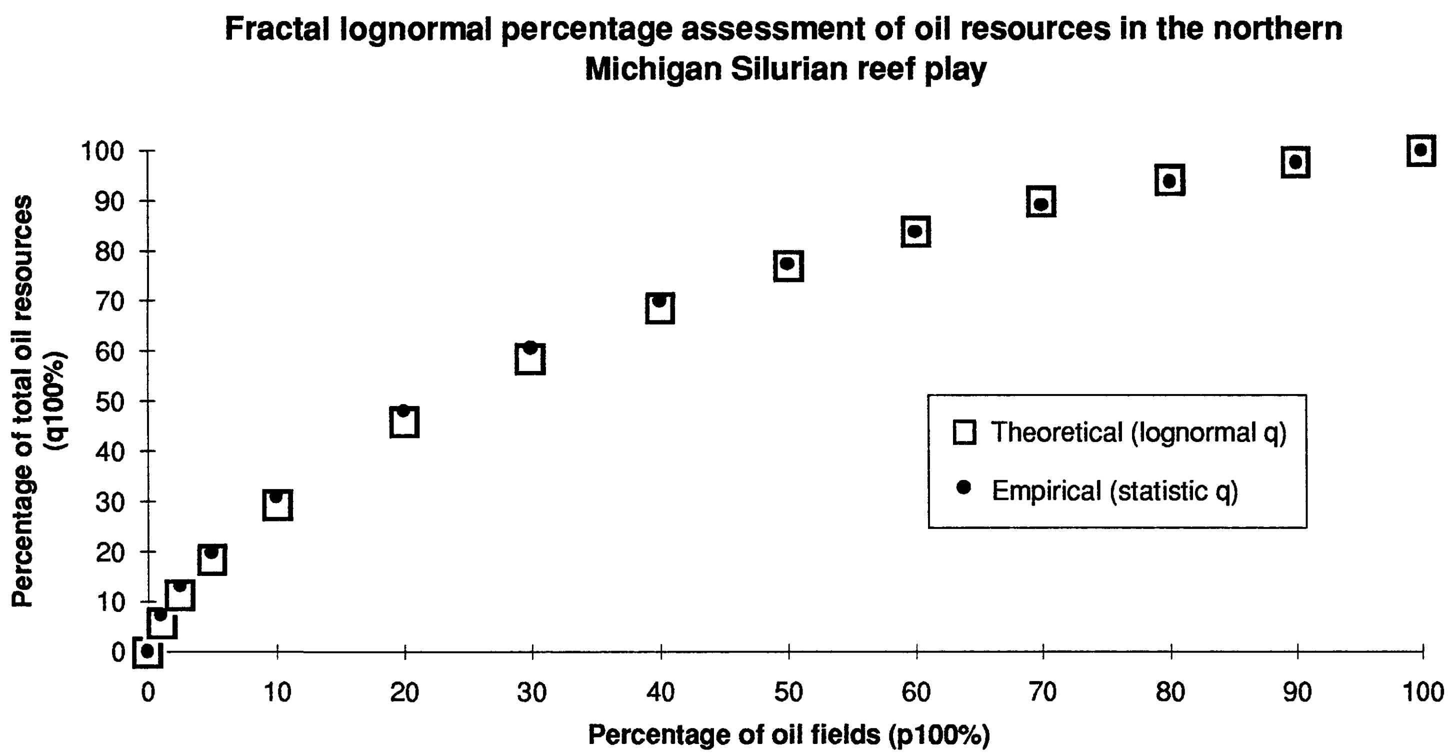


Figure 2. Theoretical percentage of total oil resources using the lognormal $q$ as a function of the percentage of oil fields, along with the corresponding empirical values using the statistic $q$. Data are from Table 3.

## SUMMARY AND CONCLUSIONS

1. "The 20/80 law" is generalized as the $p100/q100$ law in probabilistic form for any probability distribution that models the population distribution. The parameter $q$ is called the proportion of total value and defined in terms of the population proportion $p$ and conditional expectation.

2. When the population distribution is modeled with the lognormal distribution, the $p100/q100$ law in probabilistic form produces the $p100/q100$ law in lognormal form. The proportion of total value is derived, and the parameter $q$ is a function only of $p$ and the shape parameter $\sigma$. The parameter $q$ is scale-free; i.e., it does not depend upon the lognormal scale parameter $\mu$. Because it is scale invariant, the proportion $q$ is a parametric fractal.

3. The $p100/q100$ law in lognormal form was applied to petroleum field sizes in a test play. The lognormal $q$ is a function of $p$ such that $p100\%$ of the oil fields account for $q100\%$ of the total oil resources of the fields. The theoretical percentages of total oil resources using the lognormal $q$ are extremely close to the empirical percentages from the data for the test play using the statistic $q$.

4. This new probabilistic methodology for percentage assessment of resources lends itself as an ideal application for spreadsheet software.

## ACKNOWLEDGMENT

The author wishes to gratefully acknowledge Richard H. Balay at Metropolitan State College of Denver for assisting in the spreadsheet software application.

## REFERENCES

Aitchison, J. and Brown, J.A.C., 1957, The lognormal distribution: Cambridge, Cambridge University Press, 176 p.

Crovelli, R.A., 1995, The generalized 20/80 law using probabilistic fractals applied to petroleum field size: Nonrenewable Resources, v. 4, no. 3, p. 233-241.

Crow, E.L. and Shimizu, Kunio, eds., 1988, Lognormal distributions—Theory and applications: New York and Basel, Marcel Dekker, Inc., 387 p.

Douglass, M.E. and Douglass, D.N., 1993, Manage your time your work yourself: New York, AMACOM, 194 p.

Johnson, N.L., Kotz, S., and Balakrishnan, N., 1994, Continuous univariate distributions (2d ed.), v. 1: New York, John Wiley & Sons, Inc., 756 p.

Law, A.M. and Kelton, W.D., 1991, Simulation modeling and analysis (2d ed.): New York, McGraw-Hill, 759 p.

Mandelbrot, B.B., 1983, The fractal geometry of nature: San Francisco, W.H. Freeman and Co., 460 p.

NRG Associates, Inc., 1992, The significant oil and gas fields of the United States [through December 31, 1990]: Available from Nehring Associates, Inc., P.O. Box 1655, Colorado Springs, Colorado 80901.

Pareto, V., 1897, Cours d'Economie Politique: Lausanne and Paris, Rouge and Cie.

Turla, P.A. and Hawkins, K.L., 1983, Time management made easy: New York, E.P. Dutton, 206 p.