

# 分布の特性の表現に対するパレート図の利用について

東京理科大学 牧 野 都 治

(1978年8月 受付)

## On the Use of Pareto Diagrams for Representation of Characteristics of Some Distributions\*

Toji Makino

(Science University of Tokyo)

In this paper, we investigate characteristics of Pareto Diagram for the Erlang and the Weibull Distributions and examine the relationship between the Weibull and the Normal Distributions by means of Pareto Diagram.

\* The technical term Pareto Diagram is conventionally used in Operations Research but essentially the same as the so-called Lorenz' Diagram.

### 1. はじめに

オペレーションズ・リサーチの分野では、よくABC分析が用いられる[8]。たとえば在庫管理の場で、累積品目数百分率を横軸、累積金額百分率を縦軸にとり、図1のように、A, B, Cの3種類に分類し、能率的に重点管理を行なう。

これが、ABC分析とよばれる手法であるが、ORでは図1の累積曲線(ないしは折れ線)のことを、ふつうパレート曲線、パレート図などよんでいる。そこで、本稿でもそのような呼び方に従うが、この曲線は、所得格差の議論の中にみられるローレンツ曲線と、本質的には何ら異なることがない。ただ、われわれはパレート曲線を書くにあたって、金額のかさむ方から累積していくので、曲線が対角線(均等線)の上側にできるのに対し、従来ふつうに用いられているローレンツ曲線は、金額の低い方から累積するので、下側にふくらむ。

同一のデータにもとづき、金額の高い方から累積して書いた曲線と、低い方から累積した曲線とをくらべてみると、それらは正方形の中心に関して点対称になっていることはいうまでもない。

われわれは本稿において、まずパレート曲線の性質について考察し、その性質を用いて、ORの分野でしばしば用いられるアーラン分布、ワイブル分布などの特性を調べ、あわせてワイブル分布と正規分布の関係を議論したい。

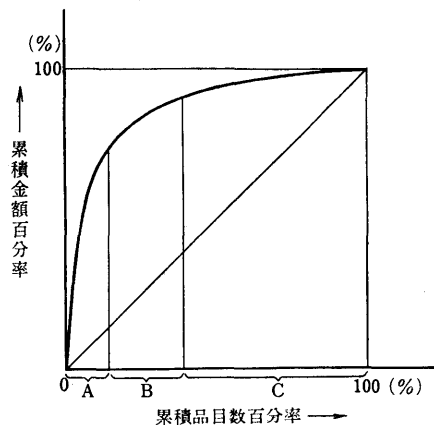


図 1. ABC 分析

## 2. パレート曲線の性質

いま、在庫管理を例にとることにして、出庫金額  $T$  の分布の確率密度関数を  $f(t)$  と書くことにする。

注 本稿では、確率変数はすべて連続的かつ平均、分散の存在するもののみを扱う。また、とくに断わり書きのない場合には、非負の値をとる確率変数を考えているものとする。

このような分布に対応するパレート曲線において、累積百分率での 100% を 1 と書くことにすると（以下、つねにこのように扱う）、横座標  $x$ 、縦座標  $y$  の関係を、媒介変数  $t$  を用いて次のように表わすことができる。

$$x = \int_t^{\infty} f(t) dt \quad (1)$$

$$y = \frac{1}{\nu} \int_t^{\infty} t \cdot f(t) dt \quad (2)$$

ただし、 $\nu = E(T)$

これらは、 $T$  の累積分布関数  $F(t)$  を用いて、次のように変形することができる。

$$x = 1 - F(t) \quad (3)$$

$$y = \frac{1}{\nu} \left[ t[1 - F(t)] + \int_t^{\infty} \{1 - F(t)\} dt \right] \quad (4)$$

したがって、パレート曲線上で、 $x$  に対応する  $\nu y$  の値を、図 2 における斜線部分の面積によって表わすことができる。

もちろん、平均  $\nu$  は

$$\nu = \int_0^{\infty} \{1 - F(t)\} dt$$

であるから、これは曲線  $F(t)$  と縦軸 1 を通る横軸への平行線とによって囲まれる部分の面積に他ならない。

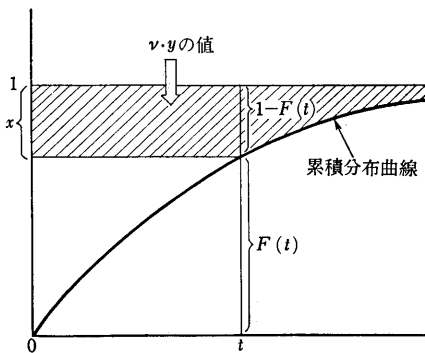


図 2. パレート曲線での  $x, y$  の関係

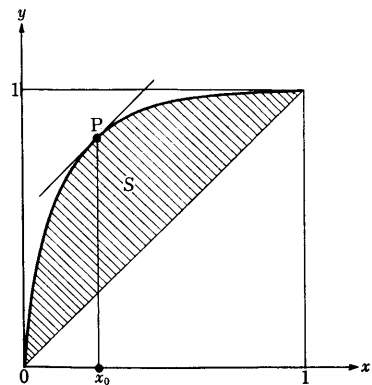


図 3. ふくらみ最大の点と弓形の面積

つぎに、パレート曲線でのふくらみ最大の点、つまり均等線との距離が最大になる点の  $x$  座標を、図 3 のように  $x_0$  と書くことにすると、 $x_0$  は

$$x_0 = \int_{\nu}^{\infty} f(t) dt$$

となることが、よく知られている。

また、不平等度をあらわす弓形の面積  $S$  は

$$\begin{aligned} S &= \int_0^1 y dx - \frac{1}{2} \\ &= \frac{1}{\nu} \int_0^{\infty} t f(t) F(t) dt - \frac{1}{2} \end{aligned}$$

によって求められ、これは尺度母数に関して不変であることも知られている。

この性質が、いくつかの分布の類似性を調べたりするのに役立つ。

### 3. アーラン分布, ガンマ分布のパレート図

待ち行列でよく用いられる、位相  $k$  のアーラン分布 ( $E_k$  分布と略記) のパレート曲線について考える。

$T$  が平均  $\frac{1}{\alpha}$  の  $E_k$  分布にしたがうとき、その確率密度関数は次のように表わされる。

$$f(t) = \begin{cases} \frac{(k\alpha)^k t^{k-1}}{\Gamma(k)} e^{-k\alpha t} & (t \geq 0) \\ 0 & (t < 0) \end{cases} \quad (5)$$

よって、パレート曲線の  $x$  座標,  $y$  座標を次のように表わすことができる。

$$x = \int_i^{\infty} \frac{(k\alpha)^k t^{k-1}}{\Gamma(k)} e^{-k\alpha t} dt \quad (6)$$

$$y = \alpha \cdot \int_i^{\infty} \frac{(k\alpha)^k t^k}{\Gamma(k)} e^{-k\alpha t} dt \quad (7)$$

さて、よく知られているように、アーラン分布はガンマ分布の特別な場合であって、(5) をガンマ分布の確率密度関数と考えてよい。ただ、アーラン分布ではパラメータ  $k$  は自然数のみをとるのに対し、ガンマ分布ではそのように限定しないことだけが異なる。

そこで、 $k$  がいろいろな場合のときのガンマ分布に対するパレート曲線を書き、ふくらみ最大の点をプロットしてみると、図4のようになる。

すなわち、この点は  $k=0$  に対応する点  $(0, 1)$  からスタートして、 $k$  の値が大きくなるに従い、点線のような動きを示し、 $k \rightarrow \infty$  のとき  $(\frac{1}{2}, \frac{1}{2})$  に到達する。

$k \rightarrow \infty$  のとき、この分布は単位分布になるので、そのパレート曲線は本来  $(0, 0)$ ,  $(1, 1)$  を結ぶ対角線になるはずであるが、この曲線でのふくらみ最大の点は対角線の中心  $(\frac{1}{2}, \frac{1}{2})$  であると考えられることになる。

つぎに、 $E_k$  分布の場合の弓形の面積について調べてみる。

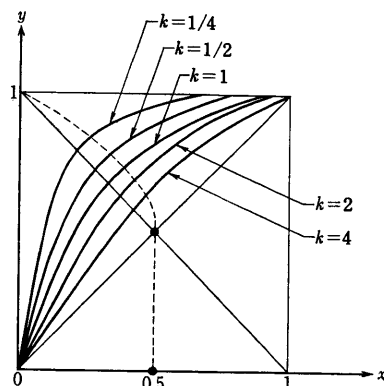


図4. ガンマ分布の場合のふくらみ最大の点

(7) 式は、部分積分により

$$y = \frac{1}{\Gamma(k)} e^{-k\alpha t} (\alpha t)^k \cdot k^{k-1} + \int_t^{\infty} \frac{(k\alpha)^k t^{k-1}}{\Gamma(k)} e^{-k\alpha t} dt$$

のように表わされるので、これと(6)式とにより、

$$y - x = \frac{1}{\Gamma(k)} e^{-k\alpha t} \cdot (\alpha t)^k k^{k-1}$$

を得る。したがって弓形の面積  $S$  は

$$\begin{aligned} S &= \int_0^1 (y - x) dx \\ &= \int_0^{\infty} \left\{ \frac{1}{(k-1)!} e^{-k\alpha t} (\alpha t)^k k^{k-1} \right\} \cdot \left\{ \frac{(k\alpha)^k}{(k-1)!} t^{k-1} e^{-k\alpha t} \right\} dt \\ &= \frac{1}{2^{2k}} \cdot \binom{2k-1}{k} \end{aligned} \quad (8)$$

になる。

とくに  $k=1$ 、つまり指数分布の場合には、弓形の面積が

$$S = \frac{1}{2^2} \times \binom{1}{1} = \frac{1}{4}$$

という、きれいな値になることは、よく知られている[6].

#### 4. ワイブル分布と正規分布

##### 4.1 ワイブル分布におけるふくらみ最大の点 $x_0$

つぎに、信頼性理論などによく用いられるワイブル分布の場合について調べてみる。ワイブル分布は、よく知られているように、位置のパラメータ、尺度パラメータ、形のパラメータをもっているが、ここでの考察においては、位置のパラメータは0として、確率密度関数が

$$f(t) = \frac{m \cdot t^{m-1}}{t_0} e^{-\frac{t^m}{t_0}} \quad (9)$$

であるとしておく。ただし、 $m$  は形のパラメータ、 $t_0$  は尺度パラメータである。この場合

$$\text{平均 } E(T) = t_0^{1/m} \Gamma\left(1 + \frac{1}{m}\right)$$

であるから、パレート曲線でふくらみ最大の点の  $x$  座標は

$$x_0 = \int_{t_0^{1/m} \Gamma\left(1 + \frac{1}{m}\right)}^{\infty} \frac{m t^{m-1}}{t_0} e^{-\frac{t^m}{t_0}} dt = e^{-\left\{\Gamma\left(1 + \frac{1}{m}\right)\right\}^m} \quad (10)$$

ここで、 $m \rightarrow 0$  のとき  $x_0 \rightarrow 0$  となることは、すぐわかるが、 $m \rightarrow \infty$  のときに  $x_0$  がどのような値になるかを調べてみるとおもしろい。前節で調べたガンマ分布の場合には、それが  $\frac{1}{2}$  に近づいたわけであるが、ワイブル分布の場合には  $0.57037\cdots$  という値に近づくことが、次の計算からわかる。

いま

$$\left\{ \Gamma \left( 1 + \frac{1}{m} \right) \right\}^m = \eta_m$$

とおくと、

$$\begin{aligned} \log \eta_m &= m \log \Gamma \left( 1 + \frac{1}{m} \right) \\ &= \frac{\log \Gamma \left( 1 + \frac{1}{m} \right)}{\frac{1}{m}} \end{aligned}$$

したがって

$$\begin{aligned} \lim_{m \rightarrow \infty} (\log \eta_m) &= \lim_{m \rightarrow \infty} \frac{\log \Gamma \left( 1 + \frac{1}{m} \right)}{\frac{1}{m}} \\ &= -\gamma \end{aligned}$$

を得る。ただし、 $\gamma$  は Euler の定数で

$$\gamma = 0.57721$$

である。

このことから

$$\lim_{m \rightarrow \infty} e^{-\left\{ \Gamma \left( 1 + \frac{1}{m} \right) \right\}^m} = e^{-e^{-\gamma}} = 0.57037 \quad (11)$$

となることがわかる。

図5は、このような様子を説明している。

#### 4.2 ワイブル分布と正規分布のふくらみ最大の点 $x_0$ による比較

つぎに、ワイブル分布と正規分布の関係について調べてみる。

ワイブル分布は、形のパラメータ  $m$  が1のときに指数分布になるが、 $m$  がある値になると正規分布に近い分布になるといわれている。このことについて、従来いくつかの試算結果が発表されている。たとえば、Kao [1] は  $m=3.25$  のとき、ワイブル分布の確率密度関数が正規分布のそれと類似することを示し、Weibull [2] 自身は、 $m$  を適当に選ぶならば、正規分布の1~99%の範囲内で、ワイブル分布からのよい近似が得られることを示唆している。また Plait [3] は

$$\begin{aligned} \text{位置パラメータ} &= 0 \\ \text{尺度パラメータ} &= 1 \\ \text{形のパラメータ} &= 3.25 \end{aligned}$$

のワイブル分布の確率密度関数を計算し、これが  $N(0.8963, 0.303^2)$  の確率密度関数によくあうことから、大体  $m=3.25$  がよいと主張している。これに対し、Dubey [4] は分布の裾の方に関心をもつ場合には、 $m=3.60232$  という値がいちばんうまく適合するといっている。

一方、われわれはかつて、瞬間平均故障率 [10] というものを定義した。それは、器材の寿命を  $T$ 、その分布関数を  $F(t)$ 、確率密度関数を  $f(t)$  として、

$$H[\lambda(t)] = \int_0^{\infty} \left\{ \frac{f(t)}{1-F(t)} \right\} f(t) dt \quad (12)$$

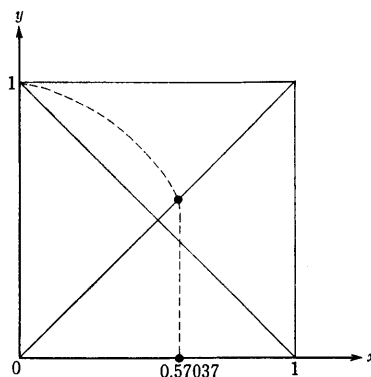


図5. ワイブル分布の場合のふくらみ最大の点

によって定められる量のことである。これはいろいろ興味ある性質をもつのであるが、この量を用いて、ワイブル分布と正規分布の関係を調べたところ、 $m=3.43927$  のとき、正規分布にいちばんよく似てくることがわかった[11]。

以上の経緯のもとに、こんどはパレート図の上で、2つの分布をくらべてみたならばどうなるかを調べてみる。ただしここでは、ワイブルと正規分布との形の類似性を調べるのがねらいなので、ワイブル分布の位置のパラメータは0として、その確率密度関数を(9)式のようにしておく。そして、はじめにパレート曲線上でのふくらみが最大になる点の  $x$  座標である  $x_0$  が  $\frac{1}{2}$  になるような  $m$  を求めてみる。このことは

$$\text{平均} = \text{メジアン}$$

となる  $m$  を求めることに他ならない。このような  $m$  は、(10)式を用いて

$$e^{-\left\{r\left(1+\frac{1}{m}\right)\right\}^m} = \frac{1}{2}$$

から求められて、

$$m = 3.4395$$

となる。これは、われわれが瞬間平均故障率をそろえることによって得た値

$$m = 3.43927$$

と、かなり似通ったものになっている。

それでは、弓形の面積によって比べてみるとどうなるかを、次に考察してみよう。

#### 4.3 弓形の面積による比較

正規分布は非負の値をとるように限定できないので、

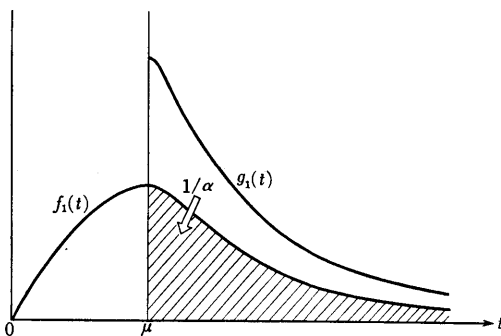


図 6.  $f_1(t)$  と  $g_1(t)$  の関係

パレート図の上でワイブルと比較するのに、そのままの形では困る。そこで、両者の右半分とか左半分どうして比べてみることにする。

はじめに、平均からの右半分について比較してみよう。そのために、まずワイブル分布での、平均から右半分をもとにして作られるパレート図の弓形の面積を計算しておく。これは、本来の確率密度関数が(9)式の  $f(t)$  (ここでは、とくに  $f_1(t)$  と書く)であるのに対し、図6に示すように、新たに  $g_1(t)$  という確率密度関数をもつ分布を考えようというのである。

つまり、(9)式で表わされる確率変数  $T$  の平均は

$$E(T) = t_0^{1/m} \Gamma\left(1 + \frac{1}{m}\right)$$

となるのであるが、これを  $\mu$  と書くことにして

$$P(T > \mu)$$

を計算すると

$$P(T > \mu) = e^{-\left\{r\left(1+\frac{1}{m}\right)\right\}^m}$$

になる。これを便宜上  $\frac{1}{\alpha}$  とおくことにする。

かくして、ワイブルで平均からの右半分に対応する分布の確率密度関数として

$$g_1(t) = \alpha \cdot f_1(t + \mu) = \alpha \cdot \frac{m}{t_0} (t + \mu)^{m-1} \cdot e^{-\frac{(t+\mu)^m}{t_0}} \quad (t > 0)$$

を得る。この分布に対する新しい平均は

$$\nu = \int_0^{\infty} t \cdot g_1(t) dt$$

であり、新しい分布関数は

$$G_1(t) = \int_0^t g_1(t) dt = \alpha \left\{ e^{-\frac{\mu^m}{t_0}} - e^{-\frac{(t+\mu)^m}{t_0}} \right\}$$

になるが、

$$\alpha \cdot e^{-\frac{\mu^m}{t_0}} = 1$$

であるから、 $G_1(t)$  はまた

$$G_1(t) = 1 - \alpha \cdot e^{-\frac{(t+\mu)^m}{t_0}}$$

とも書ける。

これらを用いて、新しい分布に対する弓形の面積  $S$  を（少し煩雑になるが）計算することができて、次のようになる。

$$S_1 = \frac{1}{2} - \left(\frac{1}{2}\right)^{1+\frac{1}{m}} \cdot e^{\left\{ \Gamma\left(1+\frac{1}{m}\right) \right\}^m} \cdot \frac{\int_2^{\infty} \left\{ \Gamma\left(1+\frac{1}{m}\right) \right\}^m x^{\frac{1}{m}-1} e^{-x} dx}{\int_1^{\infty} \left\{ \Gamma\left(1+\frac{1}{m}\right) \right\}^m x^{\frac{1}{m}-1} e^{-x} dx} \quad (13)$$

一方、正規分布の右半分の分布に対するパレート曲線の弓形の面積  $S_2$  は、次のようにして容易に計算することができる。すなわち新しい（右半分の）分布の確率密度関数は

$$g_2(t) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{t^2}{2\sigma^2}} \times 2$$

また、平均は

$$\begin{aligned} E(T) &= \int_0^{\infty} t \cdot \left\{ \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{t^2}{2\sigma^2}} \times 2 \right\} dt \\ &= \sqrt{\frac{2}{\pi}} \cdot \sigma \end{aligned}$$

となる。したがってパレート曲線の  $x$  座標および  $y$  座標は

$$x = \int_1^{\infty} g_2(t) dt, \quad y = \sqrt{\frac{\pi}{2}} \cdot \frac{1}{\sigma} \int_1^{\infty} t \cdot g_2(t) dt$$

となり、弓形の面積はこれらの  $x, y$  を用いて

$$S_2 = \int_0^1 y dx - \frac{1}{2}$$

を計算することにより、結局

$$S_2 = \frac{\sqrt{2}-1}{2} = 0.2071 \quad (14)$$

を得る.

(注) さきに述べたように, 弓形の面積は尺度パラメータ  $\sigma$  に関して無関係である.

ここで, ワイブルと正規分布をくらべるために, (13) 式と (14) 式を等置して  $m$  を求めてみると, ふくらみ最大の点に注目したときとは少し異なり

$$m = 2.9291$$

を得る. すなわち, 弓形の面積に注目して, 平均の右半分でくらべてみると, ワイブルの形のパラメータ  $m$  が 2.9291 のときに正規分布に近い分布になることがわかった.

それでは, 左半分どうしでくらべたらどうか. このことについて, いまと同様の方法で  $m$  を求めてみたところ,

$$m = 4.309$$

のときに弓形の面積が 0.2071 になることがわかった.

図7は, 上のことに関連して, 形のパラメータがいくつときの右半分または左半分の弓形の面積がいくらになるかを示したものである.

これらの  $m$  の間には, かなりのくい違いがみられるので, むしろ, 左半分にもとづく弓形の面積と右半分によるそれとが等しくなるような  $m$  を求めるのがよいかもしれない. このようにして求められた  $m$  の値は

$$m = 3.5962$$

である.

つまり, 左右対称という面から眺めた場合ワイブルでの形のパラメータ  $m$  が 3.5962 のとき, 正規分布に類似した分布になるといってよいであろう.

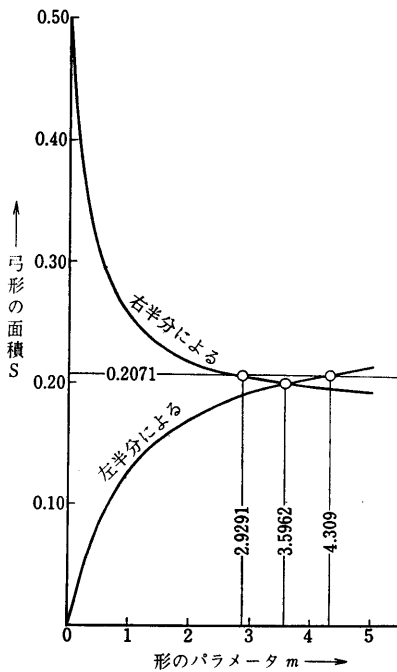


図7. 形のパラメータと弓形の面積の関係

**謝辞** 本稿を草するにあたって, 茨城大学工学部森雅夫氏から有益な御意見をいただいたこと, また数値計算の面で同学部遠藤令子さんにいろいろお手伝いしていただいたことを記して, 感謝の意を表します.

なお, 査読者の方々から貴重な助言をいただき, 本稿を修正することができました. あつくお礼申し上げます.

#### 参 考 文 献

- [1] Kao, J.H.K. (1960) A summary of some new techniques in failure analysis, *Proceedings Sixth National Symposium on Reliability and Quality Control*.
- [2] Weibull, W. (1961) *Fatigue Testing and Analysis of Results*, The Macmillan Company,



New York.

- [3] Plait, A. (1962) The Weibull distribution — with the tables, *Industrial Quality Control*, No. 19.
- [4] Dubey, S.D. (1967) Normal and Weibull distributions, *Naval Research*, **14**.
- [5] Taguchi, T. (1968) Concentration-curve methods and structures of skew populations, *Ann. Inst. Statist. Math.*, **20**, 107-141.
- [6] 田口時夫 (1960) 日本における資本の諸集中形態と構造分析の方法としてのローレンツ・カーブについて, 統計数理研究所彙報第8巻第1号.
- [7] 田口時夫 (1964) パレート分布とパレート曲線の分析, 統計数理研究所彙報第12巻第1号.
- [8] 日本工業規格 (1967) オペレーションズリサーチ用語, 日本規格協会発行.
- [9] 田口時夫 (1974) 経済統計における分布と分配構造の解析法の研究, 統計数理研究所彙報第22巻第1号.
- [10] 津村善郎・牧野都治 (1968) 瞬間平均故障率について, 日本 OR 学会研究発表予稿集 (昭 43.5).
- [11] 佐々木昭・牧野都治 (1968) 正規分布とワイブル分布, 日本 OR 学会研究発表予稿集 (昭 43.10).
- [12] 牧野都治 (1976) ワイブル分布, ガンマ分布と正規分布, 日本数学会研究発表予稿集 (昭 51.10).
- [13] 牧野都治 (1976) 待ち行列タイプの問題に対するパレート分析, 統計数理研究所彙報第23巻第2号.