Zipf's Law and 7 ± 2 Principle Lead to a Possible Explanation of Daniel's Law

Olga Kosheleva and Vladik Kreinovich
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

In 1961, D. R. Daniel observed that the success of a company is usually determined by three to six major factors. This observation has led to many successful management ideas, but they leave one puzzled: why three to six? why not two or seven? In this paper, we provide a possible explanation to this puzzle; namely, we show that these numbers of factors can be derived from Zipf's Law and from the 7 ± 2 principle.

1 A Problem: An Important Empirical Observation Needs to Be Explained

An important empirical observation. In 1961, D. Ronald Daniel observed that the success of a company is usually determined by three to six major factors [1]. Since then, the concept of critical success factors has been actively (and efficiently) used; see, e.g., [2].

Remaining problem. The success of this approach naturally leads to a question: why three to six? why not seven or eight?

What we do in this paper. We provide a possible explanation of why three to six factors are usually sufficient.

2 Analysis of the Problem and the Resulting Formulas

It is often important to find the exact expression for the dependence. In general, we are interested in controlling the value of a certain quantity y – e.g., of the company's productivity. We cannot control this quantity directly, but we can change the values of some other quantities $x_1, \ldots, x_n$, and we know that the change of these quantities x_i affect the quantity y. To implement such an indirect way of changing the quantity y, we need to know how exactly the change in x_i affects y. In other words, we need to know the expression $y = f(x)$ which determines how y depends on the parameters $x = (x_1, \ldots, x_n)$.

We need an approximate expression for the desired dependence. In general, the space of all possible functions is infinite-dimensional, meaning that we need infinitely many parameters to exactly describe each function. This can be done, e.g., if we select an orthonormal base $e_1(x)$, ..., $e_k(x)$, ...; then, each function $f(x)$ can be represented as a linear combination $f(x) = \sum\limits_{i=1}^{\infty} a_i \cdot e_i(x)$.

In practice, we can only determine and store finitely many parameters $a_1, \ldots, a_k$. Based on the known values $a_1, \ldots, a_k$, we can only determine the approximate expression $f(x) \approx f_k(x) \stackrel{\text{def}}{=} \sum\limits_{i=1}^{k} a_i \cdot e_i(x)$ for the desired dependence.

The combinations $e_1(x)$, ..., $e_k(x)$ serve as *factors* which determine the value of the desired quantity y.

Our objective. We want to determine how many factors we need. For that, we first need to describe what accuracy we need to achieve. Once this is determined, we need to determine how many factors to take into account so that we will be able to achieve the desired accuracy.

What accuracy we need: 7 ± 2 law. It is known that in the first approximation, we divide the objects into 7 ± 2 categories; see, e.g., [3]. In other words, we divide everything into five to nine categories. In particular, instead of the exact value, we have N (five to nine) categories of values.

In general, the interval $[0, V]$ of possible values is divided into N subintervals $(5 \le N \le 9)$ $\left[0, \dfrac{V}{N}\right]$, $\left[\dfrac{V}{N}, \dfrac{2V}{N}\right]$, etc. Within each subinterval, we can select the midpoint as its most accurate representation: the difference between each value from the subinterval and this subinterval's midpoint cannot exceed the subinterval's half-width $\dfrac{V}{2N}$. Thus, for the value to be within the correct category, we need to determine the value with relative accuracy $\dfrac{1}{2N}$.

The value N ranges from $N = 5$ to $N = 9$. For the smallest value $N = 5$, we thus need a relative accuracy $\dfrac{1}{10} = 10\%$, while for the largest value $N = 9$, we need a relative accuracy $\dfrac{1}{18}$.

How approximation accuracy depends on the number of factors k. For the orthonormal basis, we have $\|f\|^2 = \sum\limits_{i=1}^{\infty} |a_i|^2$, where $\|f\| \stackrel{\text{def}}{=}$

$\sqrt{\int (f(x))^2\,dx}$. For the k-term approximation $f_k(x)$, we similarly have $\|f_k\|^2 = \sum\limits_{i=1}^{k} |a_i|^2$.

When k increases, the value $\|f\|_k$ gets closer and close to $\|f\|$. Thus, to get a good approximation, we want to select the number of factors k for which $\|f_k\|$ is close to $\|f\|$.

What accuracy we need. We have mentioned that we need an approximation of relative accuracy $\dfrac{1}{2N}$. Thus, we need to find the number of factors k for which $\|f_k\| \geq \|f\| \cdot \left(1 - \dfrac{1}{2N}\right)$.

Zipf's law. To find out how many factors we need to achieve a given accuracy, we need to know how the parameters a_i depend on i. Since the series $|a_i|^2$ converges, the values a_i tend to 0. In many areas of science and engineering, the decrease of a tending-to-zero quantity a_i can be described by the Zipf's law $a_i = \dfrac{c}{i}$; see, e.g., [4]. It is therefore reasonable to use this law.

Let us now combine all these formulas to get a possible explanation for the puzzling three-to-six factor.

3 Resulting Explanation for Three-to-Six Factors ($k = 3$ through $k = 6$)

The condition for the number of factors. According to the above analysis, k factors are sufficient when $\|f_k\| \geq \|f\| \cdot \left(1 - \dfrac{1}{2N}\right)$, where N can take any value from 5 to 9. Thus, to find the smallest number of factors sufficient to describe the dependence with a given accuracy, we need, for each N from 5 to 9, to find the smallest k for which the above inequality holds.

Simplifying the above condition. Since the values $\|f\|$ and $\|f_k\|$ are defined as square roots, it is convenient to consider not the original inequality, but the equivalent one which is obtained from the original one by squaring both sides:

$$\|f_k\|^2 \geq \|f\|^2 \cdot \left(1 - \frac{1}{2N}\right)^2.$$

Expressing $\|f_k\|^2$ and $\|f\|^2$ in terms of the values a_i, we get the equivalent inequality

$$\sum_{i=1}^{k} a_i^2 \geq \left(\sum_{i=1}^{\infty} a_i^2\right) \cdot \left(1 - \frac{1}{2N}\right)^2.$$

Substituting $a_i = \frac{c}{i}$ into this formula, we get

$$\sum_{i=1}^{k} \frac{c^2}{i^2} \geq \left(\sum_{i=1}^{\infty} \frac{c^2}{i^2}\right) \cdot \left(1 - \frac{1}{2N}\right)^2.$$

Dividing both sides of this inequality by c^2, we get:

$$\sum_{i=1}^{k} \frac{1}{i^2} \geq \left(\sum_{i=1}^{\infty} \frac{1}{i^2}\right) \cdot \left(1 - \frac{1}{2N}\right)^2,$$

It is known that $\sum\limits_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}$. Thus, we conclude that

$$\sum_{i=1}^{k} \frac{1}{i^2} \geq \frac{\pi^2}{6} \cdot \left(1 - \frac{1}{2N}\right)^2.$$

For each N, we need to find the smallest k which satisfies this inequality.

It is sufficient to consider extreme values of N. Let $k(N)$ denote the smallest value k for which the above inequality is satisfied for a given N. We are interested in the range of possible values of $k(N)$ when N takes the values from 5 to 9.

For each N, the above inequality is satisfied for $k = k(N)$. The smaller N, the larger $\frac{1}{N}$ and thus, the smaller the difference $1 - \frac{1}{2N}$ and the smaller the square of this difference. So, for $N' < N$, the above inequality is also satisfied for $k = k(N)$. Therefore, the smallest value $k(N')$ for which the above inequality is satisfied for N' is smaller than or equal to $k(N)$.

In other words, the smaller N, the smaller $k(N)$. Thus, the smallest possible value of $k(N)$ is attained when N is the smallest (i.e., when $N = 5$), and the largest value of N is attained when N is the largest, i.e., when $N = 9$. So, to find the range of possible values of $k(N)$, it is sufficient to consider only two values of N: $N = 5$ and $N = 9$. Let us consider these two values one by one.

Case of $N = 5$. For $N = 5$, we have $1 - \frac{1}{2N} = 0.9$ and thus,

$$\left(1 - \frac{1}{2N}\right)^2 = 0.81.$$

So, we want to find the smallest k for which

$$s_k \stackrel{\text{def}}{=} \sum_{i=1}^{k} \frac{1}{i^2} \geq \frac{\pi^2}{6} \cdot 0.81.$$

Here, $\frac{\pi^2}{6} \approx 1.644$, so $\frac{\pi^2}{6} \cdot 0.81 \approx 1.331$.

For $k = 1$, we have $s_1 = 1.000 < 1.331$. For $k = 2$, we have

$$s_2 = 1 + \frac{1}{4} = 1.250 < 1.331.$$

For $k = 3$, we have

$$s_3 = 1 + \frac{1}{4} + \frac{1}{9} \approx 1.250 + 0.111 = 1.361 > 1.331.$$

Thus, for $N = 5$, the smallest number of factors needed to described the dependence with the corresponding accuracy is $k = 3$.

Case of $N = 9$. For $N = 9$, we have $1 - \frac{1}{2N} = 1 - 118$ and thus,

$$\left(1 - \frac{1}{2N}\right)^2 = \left(1 - \frac{1}{18}\right)^2 = 1 - \frac{1}{9} + \frac{1}{324} \approx 1 - 0.111 + 0.003 = 0.892.$$

So, we have $\frac{\pi^2}{6} \cdot 0.892 \approx 1.644 \cdot 0.892 \approx 1.489$.

Let us find the smallest k for which $s_k \geq 1.489$. We already know that this value $k(9)$ is larger than or equal to $k(5) = 3$, so we start with $k = 3$. For $k = 3$, we have $s_3 \approx 1.361 < 1.489$. For $k = 4$, we have

$$s_4 = s_3 + \frac{1}{16} \approx 1.361 + 0.063 = 1.424 < 1.489.$$

For $k = 5$, we have

$$s_5 = s_4 + \frac{1}{25} \approx 1.424 + 0.040 = 1.464 < 1.489.$$

Finally, for $k = 6$, we have

$$s_6 = s_5 + \frac{1}{36} \approx 1.464 + 0.028 = 1.492 > 1.489.$$

Thus, for $N = 9$, the smallest number of factors needed to described the dependence with the corresponding accuracy is $k = 6$.

Conclusion. When N takes values from 5 to 9 (in accordance with the 7 ± 2 law), the corresponding number of factors ranges from $k = 3$ to $k = 6$. Thus, the above combination of Zipf's law and 7 ± 2 law indeed explains Daniel's law – that we always need three to six factors.

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by a travel grant from Chiang Mai University, Thailand.

The authors are thankful to all the participants of the 7th International Conference of Thailand Econometric Society TES'2014 (Chiang Mai, Thailand, January 8–10, 2014) for valuable discussions.

References

[1] R. D. Daniel, "Management information crisis", *Harvard Business Review*, 1961, Vol. 39, No. 5, pp. 111–121.

[2] J. A. Johnson and M. Friesen, *The Success Paradigm: Creating Organizational Effectiveness through Quality and Strategy*, Quorum Books, New York, 1995.

[3] G. A. Miller, "The magical number seven plus or minus two: some limits on our capacity for processing information", *Psychological Review*, 1956, Vol. 63, pp. 81–97.

[4] G. K. Zipf, *Human Behavior and the Principle of Least Effort*, Addison Wesley, Cambridge, Massachusetts, 1949.