

The Zipf's Distribution of the Number of Medals in 2016 Olympic Game in Rio DeJaneiro

Y. Yamada

College of Engineering
Chubu University
Japan

H. H. A. Rêgo

Departamento de Física
Instituto Federal de Educação
Ciência e Tecnológico Maranhão
Brazil

H. E. Stanley

Center for Polymer Studies
Boston University
USA

K Yamamoto, S. Miyazima

Faculty of Science and Engineering
Department of Electrical and Electronic Engineering
Setsunan University Japan

Abstract

In a socio-economical point of view, one can have a common belief that richer countries tend to win more medals in worldwide competitions. The Olympic game is a sport competition that where some country win more prizes overall, challenging the logic of the distribution of medal. In the present study we analyze the distribution of medals by country in 2016 Olympic game held in Rio de Janeiro, Brazil. Despite in a first look one can apparently find it random, after a careful analysis we find that the number of medals achieved by each country indicates Zipf's law distribution. The reason for Zipf's law in the distribution of medals is discussed from the point of view of a relation between the number of participants and GDP.

1. Introduction

The Olympic Games and the Medals

As being one of the most important events in the modern world, the Olympic game features summer and winter sports competitions, originally proposed by Coubertin as a revival of the Ancient Olympic game, which started in B.C. 9th century. The first modern Olympics was held in 1896 in Greece, and the recent one in 2016 in Rio de Janeiro. The Olympic game is a unique event aiming to promote peace across the countries, religions, ideologies, races, and so on. The Olympic game can be "more than an organized set of competitions" that induces a social and economic development of the host cities/countries. More than 11,000 people participated from different countries over 200, and competed on 28 sports in 306 different events, capturing the world's attention in the following question: which country won more medals? The answer to that particular question becomes clear immediately after competition. The country in the 1st, 2nd, or any other place in the rank, varies, reflecting the tradition and the evolution of sports skills. Intuitively, a richer country seems to take more prizes, however, it may not be always true, and countries with low wealth sometimes having strong traditions in sports can overcome richer nations. Looking at a simple statistics of medal number which was achieved by each country, it seems to be random distribution. But when we plot the number of medals against the number of the participants of each country as shown in Fig.1, one may notice that there is a correlation between the total number of achieved medals by each country and the number of participants of each country.

In such plot, the total number achieved in one country is distributed from zero to 46 of Gold Medals, the Silver Medals is from zero to 37, and the Copper Medals is from zero to 38. The number of participant of each country distributes up to 554 from USA.

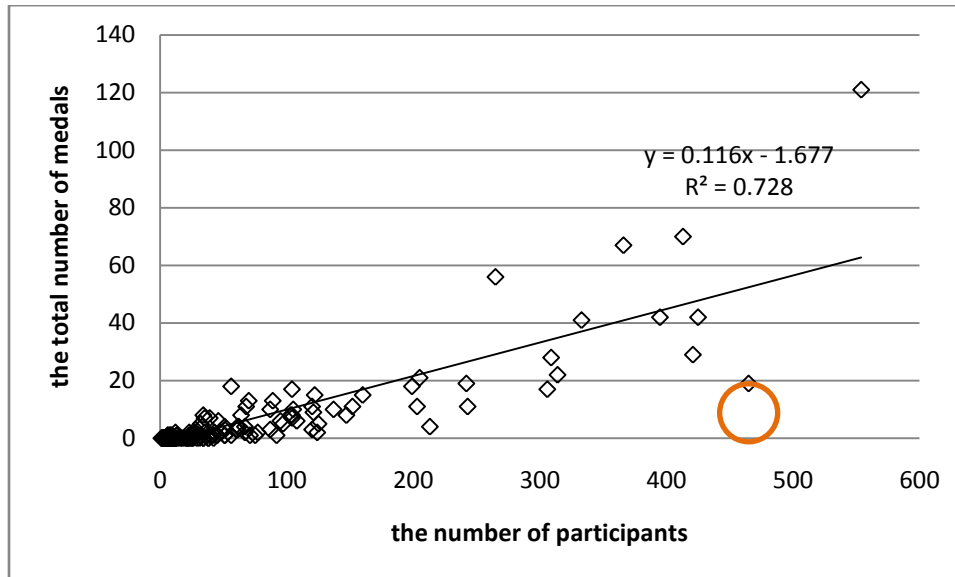


Fig.1 The total number of medals achieved by each country is plotted against the number of participants. A diamond \diamond in a circle is for Brazil, where the host country for Olympic 2016 in Rio de Janeiro. Usually about a half of this year's participants were sent from Brazil.

Fitting by a quadratic equation or higher order, we find that coefficients of higher powers such as $x^2, x^3 \dots$ are very small. The fitting by a line is given in Fig.1 with $R^2 = 0.73$. The correlation coefficient between the total number of medals and the number of participants is 0.85. We may conclude that these two quantities, that is, the number of medals and the number of participants are correlated each other.

2. The Zipf Distribution

Zipf's law is a famous regulation to stand for almost all phenomena even if the phenomena are seemed to be random. In Linguistic studies, Zipf described a very typical example of a power law scaling in data on human behavior¹. By examining the word frequency and its rank in texts, Zipf found that the word frequency distribution is related with its rank, i.e., the frequency of appearance of the k -th element in the rank is proportional to $1/k$.

Similar regulation is found in various fields and a number of researchers have studied about Zipf's law. The Zipf's law can also be applicable to various random and complex phenomena. Gutenberg-Richter's law that the frequency of earthquakes beyond magnitude M is proportional to $(\text{the magnitude})^{-2/3}$ is similar problem to Zipf's law². Recently, the Zipf's law was discussed by many authors, such as many different types of power law in fractal matters^{3,4}, $1/f$ noise⁵ and Pareto theory⁶, frequency of the words which are often used^{7,8}, the family name distribution^{9,10}, the population or scale of cities¹¹⁻¹⁴, scale of firm¹⁵, internet access ranking^{16,17}, the distribution of high income¹⁸⁻²⁴, and economic growth²⁵, etc.

In the next section we will apply Zipf's plotting for the number of medals in the Olympics.

3. Zipf's Law in the number of medals

In Fig. 2, we plot the number of gold medals against the rank which is determined by the number of gold medals represented by the diamond \diamond in a double logarithmic diagram²⁶. The square \square and the triangle Δ are plotted in a similar manner, but with ranking which is decided by the number of silver and copper medals, respectively. We find that the three curves show a pretty straight portion at the initial part with $y = -0.6895x + 3.7642$, that means the number of medals depends as a power law on the rank, i.e., $\sim r^{-0.69}$ up to about $r = 13$. The fitting by a line is given in Fig.2 with $R^2 = 0.94$.

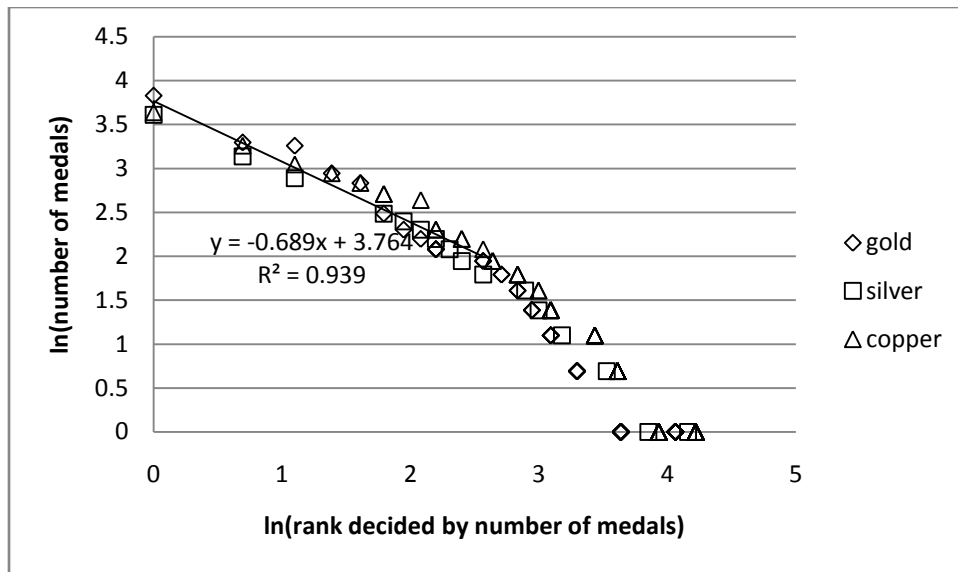


Fig.2. The number of medals is plotted against the rank, where the rank is decided by the number of gold medals, silver medals, and copper medals, respectively.

On the other hand, the data in the lower part of Fig.2 seems not to follow a power law. We can count the number of countries which achieved one gold medal, two gold medals and so on up to about ten gold medals. Also, we can count the number of countries which won silver and copper medals, respectively in the similar way, which is plotted by the square \square and the triangle Δ in Fig. 3a. These points of the diamond \diamond (gold medal), the square \square (silver medal) and the triangle Δ (copper medal) are fitted by lines of solid, dashed and dot-dashed lines, respectively. The fitting lines are $y = -1.2417x + 3.0505$ ($R^2 = 0.86$) for the gold, $y = -1.3799x + 3.2925$ ($R^2 = 0.85$) for the silver and $y = -1.1509x + 3.1122$ ($R^2 = 0.8$) for the copper medals, respectively. The fitting line of the total number of gold, silver and copper medals is $y = -1.2145x + 4.2439$ ($R^2 = 0.93$) as shown in Fig 3b.

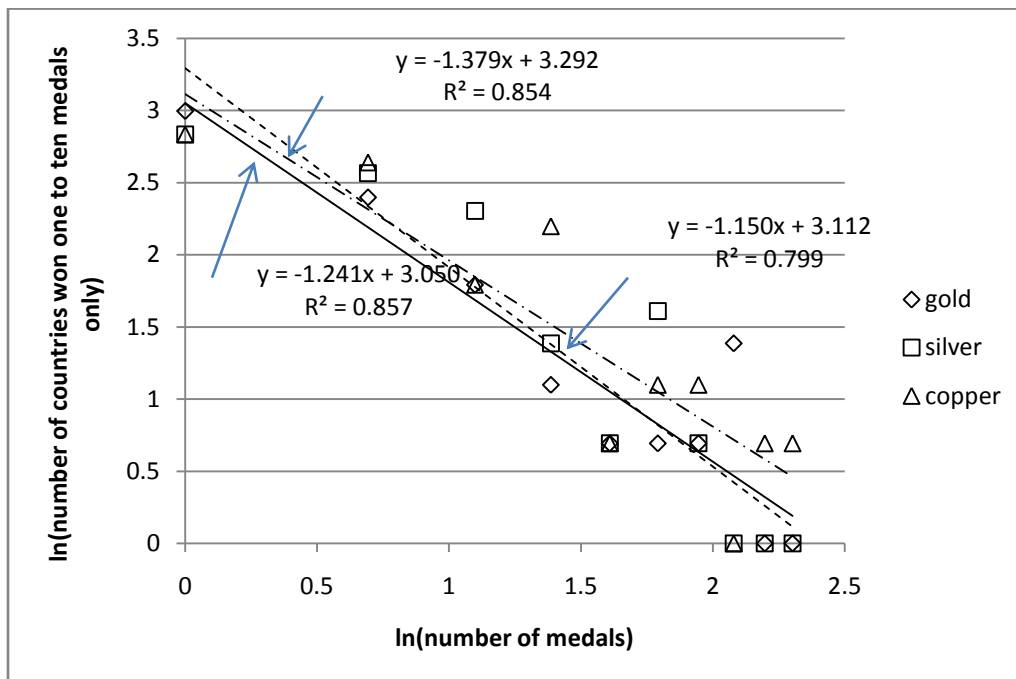


Fig.3a. The log-log plot of the number of countries which achieved only one, two, three up to ten gold medals against number of medals. Values of gold medals are plotted by \diamond , also silver and copper medals are plotted by \square and Δ .

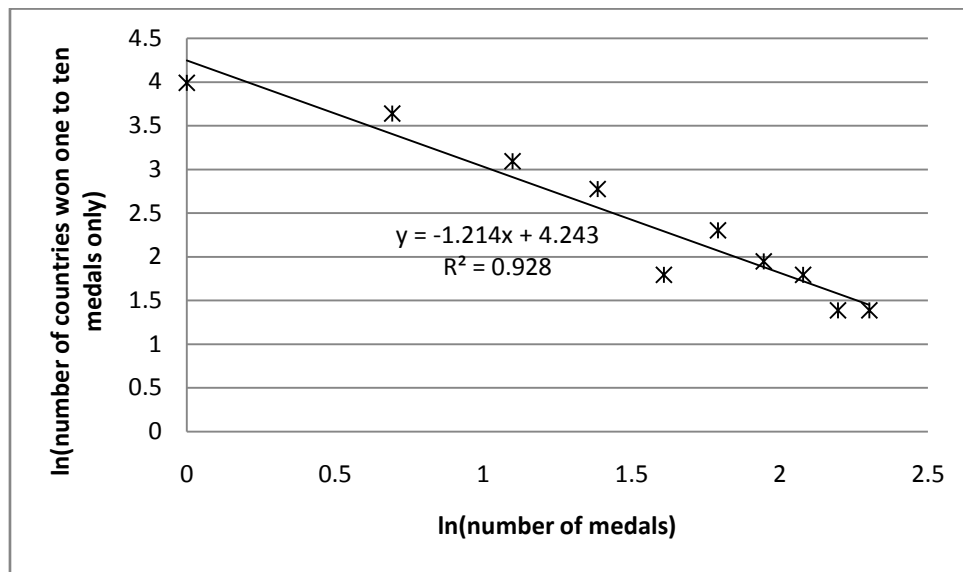


Fig.3b. The number of countries which achieved only one, two, three medals and so on without distinction of the color of medals is plotted against number of medals up to ten gold medals. The line is a fitting line.

The behaviors of competition in Olympic game are similar to those found on complex systems phenomena, like the word distribution, population, earthquakes and so on. From Figs 2 and 3, we may conclude that the medal number also follows to Zipf's law, but it is important to explain why the medal number in Olympic game follows Zipf's law, in the next section.

4. Why the number of medals follows to Zipf's Law?

Now we will consider how the number of participant is decided, because the number of medals is almost proportional to the number of participants as seen in Fig.1. The athletic players have to prepare for a long time in order to grow up their skill, where the most of them need nationwide support in many aspects, especially financial support is most serious one. In addition, most countries seem to believe that the Olympic event is the critical opportunity to appeal their significance towards the world. In order to let players take part in the Olympic game, the nationwide effort to grow superior athletes is critical. The country which is able to send more players to the Olympic is considered to be a more "economically advantageous" country. The more economical capable countries could send the more participants to the Olympic game, in this case, to Brazil. Considering that the GDP index represents a comparative economic allowance of the countries, one can state that the country which has large economic power (GDP) can send a large number of participants.

Thus, by following such logic, we plot the total number of participants against GDP^{27,28} of participants' country in Fig.4, and we find that there is a correlation also between the number of participants and GDP. The fitting lines are drawn by splitting the whole data into two parts, where the first part is of the three economical-big countries (USA, China and Japan) in GDP and the second part is the other countries. The circle \circ shows the top three countries with the fitting line $y = 0.016x + 257.07 (R^2 = 0.98)$. The cross $+$ is for the other countries except USA, China and Japan with the fitting line $y = 0.1474x + 20.981 (R^2 = 0.7)$. The correlation coefficient between the number of participants and GDP of Fig.4 is 0.68, which is small but positive. From Fig.4 we understand that the top three countries send much less participant than expected number which is estimated from the lower fitting line.

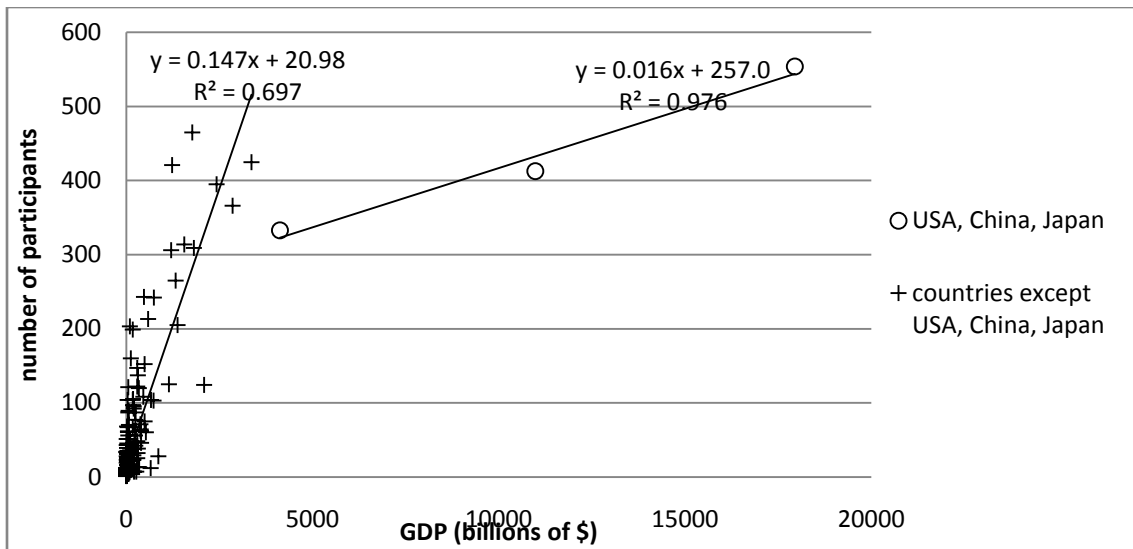


Fig.4. The number of participants against GDP (billions of \$). □ shows the top three countries (USA, China and Japan) and Δ is for the other countries.

We try to give a proof that there is a correlation between the number of participants and the number of medals, also between the number of medals and GDP. By Fig.1 we have already known that the number of achieved medals correlate with the number of participants. In order to check this correlation we plot the two quantities – the number of participants and GDP in Fig. 4.

From Fig.1, it is obvious that countries which send more participants obtain a lot of medals. From Fig.4, we can tell that the country with more GDP can send more participants. Thus, we can conclude that there should be a strong correlation between the number of medals and GDP. Therefore, our results indicate that the number of obtained medals might have been represented by Zipf’s law, because Zipf’s law behavior of GDP itself has been shown by many authors^{25,29-32}.

5. Discussion

In order to understand our result better in a global point of view, we analyze the quantities of participants, number of medals, and GDP distributed by seven regions (North America, Asia, Europe, Latin America, Oceania, Middle East and Africa) in the world³³. First, in Figs5a and 5b, we indicate the number of participants and the number of medals by the world regions such as Asia, Europe, Latin America, Oceania, Middle East, Africa and North America.

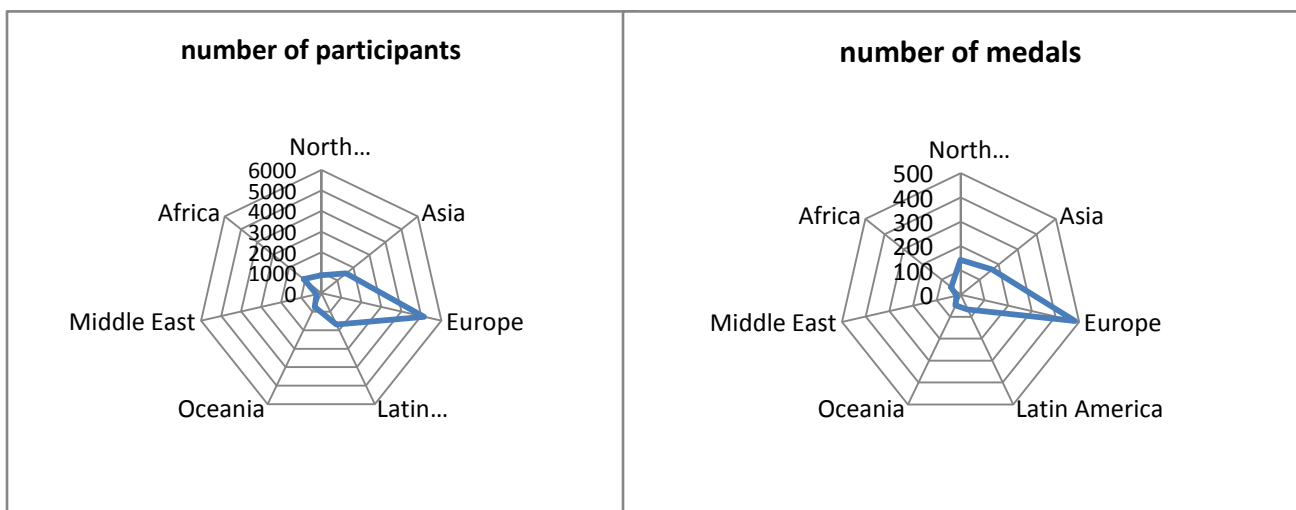


Fig5a. Comparison of the number of participants from seven world regions

Fig5b. Comparison of total number of medals which were achieved by seven world regions

By looking at both Figs. 5a and 5b, it is obvious that most of whole participants were from European countries. In the other hand, Figure 6a shows the total amount of GDP of seven regions. There are two more big peaks of Asia and North America added to the peak of Europe in Fig. 6a which is due to the fact that USA of which GDP is no.1 in the world belongs to North America. Also no.2 and no.3 in the world are China and Japan which belong to Asia. Except these three big countries, the top amount of GDP is shared by European countries as shown in Fig.6b and the radar pattern is similar to Figs 5a and 5b.

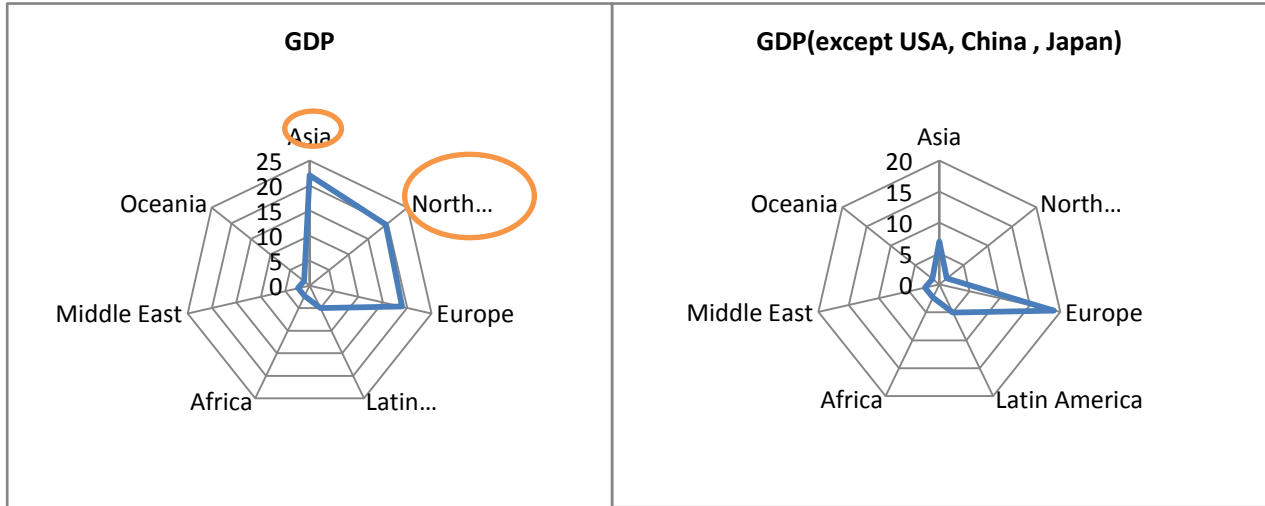


Fig. 6a.The total amount of GDP (the unit is a trillion \$)of 7 world areas. Asia and North America in circles have a large amount of GDP of USA , China and Japan respectively.

Fig.6b.The total amount of GDP (the unit is a trillion \$) of 7 world areas except USA, China and Japan. Then we obtain a very similar pattern as Figs 5a and 5b

The correlation between the number of participants and the number of obtained medals is understood by Fig.1. We also found that there is strong correlation between the number of obtained medals and GDP which shows the economic status of the country from Figs 5 and 6.The Olympic game is considered to be the competition which can symbolize each background of each country as shown in Fig.7.The correlation coefficient is 0.80 which shows rather good correlation.

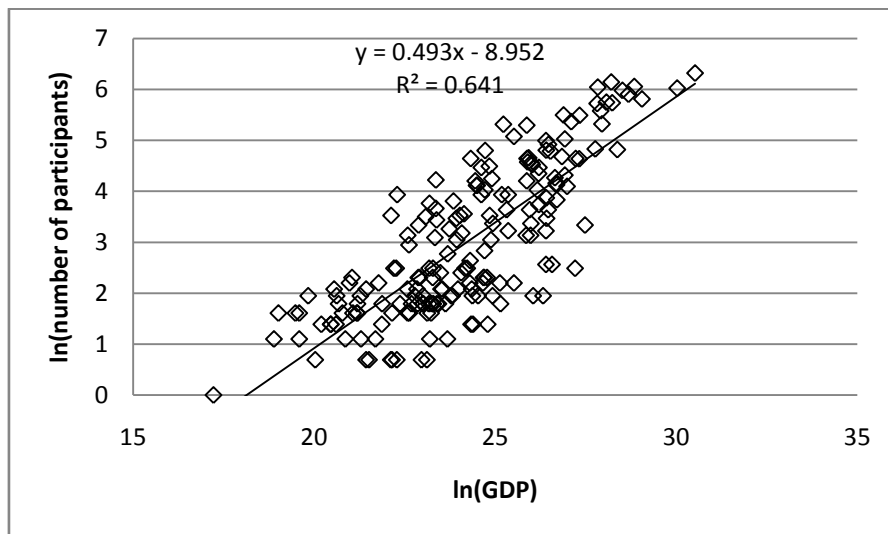


Fig.7. The log-log plot of the number of participants against GDP (billions of \$).

Finally we show that the number of participants satisfies already with Zipf’s law. In Fig.8, we show log-log plot of the number of participants against the rank. The fitting by a line $y = -0.2655x + 6.3633(R^2 = 0.91)$ for top twelve countries (USA, Brazil, Germany, Australia, China, France, UK, Japan, Canada, Italy, Spain and Russia) is given in Fig.8 which shows quite good matching with Zipf’s law.

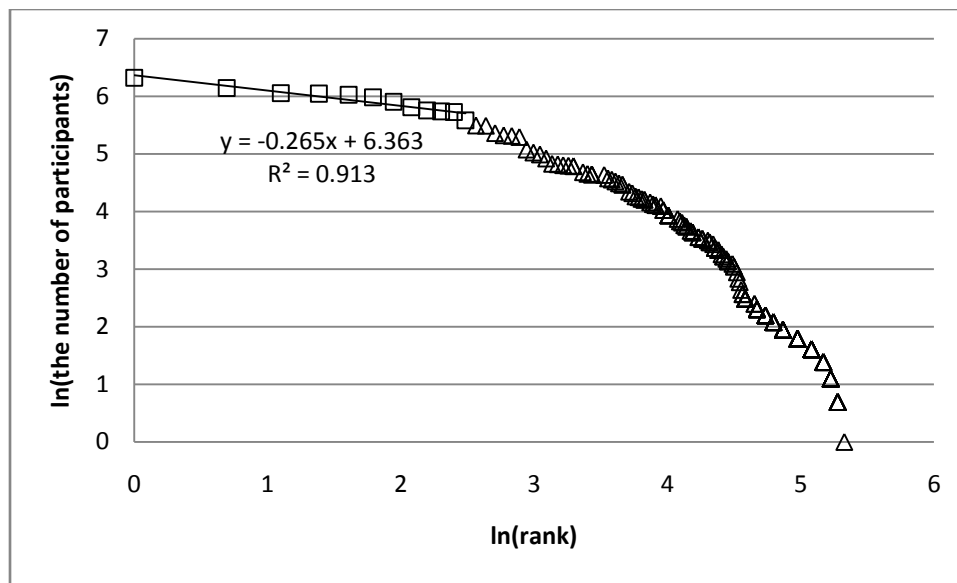


Fig.8. log-log plot of the number of participants against rank

Thus, it is obvious that there is a strong correlation between the number of participants and the economic power.

6. Conclusions

A lot of world records and the number of medals were won by competition of various athletic events in the Rio Olympic game in 2016 after great enormous effort of players. Despite an inspection initially induces that the distribution of the number of medals follows a random distribution, in a more close analysis we found that they indeed follow the very well-known Zipf's distribution. The current analysis has been made by considering the whole GDP as being a key factor to the investments on sports, but of course, this relation is not linear, some countries spend more percentage of its budget in sports than others, so one further step to this work is also considered to be the specific amount of resources spent by each country on those activities. In another point of view, the practice of some sports varies among countries, and the distribution of resources are also clearly non-linear among the different disciplines, and this can naturally affect the final result obtained by a country in Olympics. Such factor should also be considered in our future analysis.

References

- ipf, G. K. Human, "Behavior and the principle of least effort" (Addison-Wesley, Cambridge, MA, 1949)
- B. Gutenberg, C. F. Richter, "Magnitude and energy of earthquakes", *Annali di Geofisica*, Vol.9, n.1, 1956
- B. Mandelbrot, "The fractal Geometry of Nature", 1982. W.H. Freeman and Company
- J. Laherrere, and D. Sornette, "Stretched exponential distributions in Nature and Economy: 'Fat tails' with characteristic scales", *European Physical Journals*, B2, 525-539, 1998
- Y. Chen, "Zipf's law, 1/f noise, and fractal hierarchy", *Chaos, Solitons & Fractals*, Vol.45, Issue 1, 63-73 2012,
- R. Gunther, L. Levitin, B. Shapiro, and P. Wagner, "Zipf's law and the effect of ranking on probability distributions", *International Journal of Theoretical Physics*, **35**(2), 395-417, 1996
- G. K. Zipf, "Selected Studies of the Principle of Relative Frequency in Language". Harvard University Press, Cambridge, 1932.
- S. T. Piantadosi, "Zipf's word frequency law in natural language: a critical review and future directions", *Psychonomic Bulletin & Review*, October, Vol.21, Issue 5, 1112-1130, 2014
- S. Miyazima, Y. Lee, T. Nagamine, H. Miyajima, "Family Name Distribution in Japanese Societies". *J. of Physical Society of Japan*, Vol.68, No.10, Oct.1999
- S. Miyazima, Y. Lee, T. Nagamine, H. Miyajima, "Power-law distribution of family names in Japanese societies". *Physica A* 278, 2000
- P. Bak, "How Nature Works: The science of self-organized criticality", Springer-Verlag, New York, 1996.
- X. Gabaix, "Zipf's law for cities: an explanation", *The Quarterly Journal of Economics*. 114(3), 739-767, 1999.
- X. Gabaix, "Zipf's Law and the Growth of Cities", *American Economic Review*, vol. 89, no. 2, 129-132, 1999

- J. A. O. Aragón, V. d. S. Queiroz, "The Zipf's law and the effects of free trade: The case of Guatemala", *Economía* 15, 82–99, 2014
- M. H. R. Stanley, S. V. Buldyrev, S. Havlin, R. N. Mantegna, M. A. Salinger, & H. E. Stanley, "Zipf plots and the size distribution of firms", *Economics Letters* 49, 453–457, 1995
- R. Albert, H. Jeong, A-L Barabasi, "The Diameter of the World Wide Web", *Nature* **401**:130.
- M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On Power-Law Relationships of the Internet Topology", *SIGCOMM* 251-262, 1999
- G D Champernowne, "A Model of Income Distribution", *Economic J.* 63, 318, 1953
- H F Lydall, *Econometrica*, "The Distribution of Employment Incomes", 27, 110-115, 1959
- K. Yamamoto, S. Miyazima, "Power-law Distributions of High Income Taxes in Japan", *Sociological Theory and Methods* Vol.16, No.1:133-138, 2001
- K. Yamamoto, S. Miyazima, R.K. Koshal, Y. Yamada, "A model for high tax payers", *Physica A* 314, 743-748, 2002
- K. Yamamoto, S. Miyazima, R. Koshal, M. Koshal, Y. Yamada, "A Model for Distribution of High-Tax Payers", *Japan J. Indust. Appl. Math.*, 20, 147-154, 2003
- K. Yamamoto, S. Miyazima, "Power-law behavior in social and economical phenomena", *Physica A* 344, 757-763, 2004
- H. Yamamoto, T. Ohtsuki, A. Fujihara, S. Tanimoto, K. Yamamoto, S. Miyazima, "Asymptotic Analysis of the Model for Distribution of High-Tax Payers", *Japan J. Indust. Appl. Math.*, 24, 211-218, 2007
- X.Gabaix "Power Laws in Economics: An Introduction", *Journal of Economic Perspectives*, Vol.30, No.1, pp185-206, 2016
- <http://www.joc.or.jp/games/olympic/riodejaneiro/medalranking/>
- International Monetary Fund World Economic Outlook (October-2015)
- http://ecodb.net/ranking/imf_ngdpd.html
- D.Furceri, "Zipf's law and world income distribution", *Applied Economics Letters*, Vol.15,-Issue 12, 2008, 921-923
- J. Hinloopen and C.V. Marrewijk, "Comparative Advantage, the Ranksize Rule, and Zipf's Law", Tinbergen Institute Discussion Paper, 2006
- R. K. Skipper, "Zipf's Law and Its Correlation to the GDP of Nations", *The Univ. of Maryland McNair Scholars Undergraduate Research Journal*, Vol.3, 2011, 217-226
- G. Li and C. Fang, "Global mapping and estimation of ecosystem services values and gross domestic product: A spatially explicit integration of national 'green GDP' accounting", *Ecological Indicators*, Vol.46, 2014, 293-314
- IMF - World Economic Outlook Databases