



## Application of the Records Method to Identify the Sporadicity of *Percnon gibbesi* Distribution in Greece

Zaher Khraibani<sup>1\*</sup>, Zainab Assaghir<sup>1</sup>, Hussein Khraibani<sup>1</sup>,  
Dimitris Poursanidis<sup>2</sup>, Anis Hoayek<sup>1</sup>, Ghazi Bitar<sup>3</sup>, Hassan Zeineddine<sup>1</sup>,  
Argyro Zenetos<sup>4</sup> and Hussein Badran<sup>1</sup>

<sup>1</sup>Faculty of Sciences Department of Applied Mathematics, Lebanese University, Lebanon.

<sup>2</sup>Foundation for Research and Technology-Hellas (FORTH), Institute of Applied and Computational Mathematics, N. Plastira 100, Vassilika Vouton, 70013, Heraklion, Greece.

<sup>3</sup>Faculty of Sciences, Department of Biology, Lebanese University, Lebanon.

<sup>4</sup>Institute of Marine Biological Resources, Hellenic Centre for Marine Research, Anavyssos 19013, Greece.

### Authors' contributions

This work was carried out in collaboration between all authors. Author ZK designed the statistical study and wrote the protocol and wrote the first draft of the manuscript. Authors HK and HZ managed the literature searches. Author AH wrote the final manuscript. Authors HB and GB managed the experimental process. Authors DP and AZ identified the species of *Percnon gibbesi*. All authors read and approved the final manuscript.

Original Research Article

Received 27<sup>th</sup> December 2013

Accepted 21<sup>st</sup> April 2014

Published 27<sup>th</sup> June 2014

### ABSTRACT

**Aims/objectives:** Mediterranean Sea is one of the most severely affected by alien marine invasions regions worldwide. The reported alien species in the Mediterranean Sea represent approximately 6% of the known biodiversity which is estimated to approximately 17000 species. The number of species increases with a rate of one new record every 1.5 weeks. The Sally Light-foot crab, *Percnon gibbesi* (H. Milne Edwards, 1853) is a primarily algivorous crab of the shallow infra-littoral rocky shores. *Percnon gibbesi* increased rapidly its distribution in the Mediterranean Sea. In this paper, we will use a mathematical framework based on the theory of records to identify if *Percnon gibbesi* is sporadic. A record is a result or some measurement in a given chain of events that exceeds

\*Corresponding author: E-mail: [zaher.khraibani@gmail.com](mailto:zaher.khraibani@gmail.com);

everything that has been encountered previously. Then, we first describe the distribution of *Percnon gibbesi* on site in the Mediterranean Sea, Greece, and after we perform a test to identify if the specie is rare

**Study Design:** Number of records.

**Place and Duration of Study:** Department of Applied Mathematics and Department of Biology, Institute of Applied and Computational Mathematics, between February 2013 and November 2013.

**Methodology:** We apply a new statistical method in marine ecology to study the evolution of some new marine species for which few information are available. We would determine if some species are isolated (sporadic) based on record method. A specie is recorded or have a record if the number of occurrence of this specie exceeds the last registered number. The observations can be presented with a high variability that may come from some random phenomena such as the transfer, the environment or others.

**Results:** We consider the inverse of the waiting time between two observed successive species, *Percnon gibbesi* in the Mediterranean Sea Greece. The first case is founded in Greece in 2001 and the last in 2012. Based on the record method we can easily check the evolution of *Percnon gibbesi* in Greece is a species that spreads with the time.

**Conclusion:** We presented a new method of detection of the beginning of a species marine based on the number of observed records from the first observations. The properties of this method consist first on its convenience for any number  $n$  of observations since the distribution of the number of records is exactly calculated for each  $n$ , and its robustness since this distribution is independent of the cumulative distribution function involved in the distribution of  $\{\Delta S_n\}$ .

**Keywords:** Number of records; Sporadic Mathematics subject classification; 53C25; 83C05; 57N16.

## 1. INTRODUCTION

In recent years, researchers have become more and more interested in records. Many new and interesting applications of the theory of records has been discovered and explored. A record is simply an achievement, a result or some other kind of measurement in a given chain of events that exceeds everything that has been encountered previously. Thus, a new record is always something remarkable which attracts attention nevertheless the occurrence of this record is considered good or not. Records show the boundary of what has been possible so far. This is why the records become more important than other measurements. For example, the famous book Guinness World Records holds its own record as the best-selling copy righted book in history [1]. Records are also of great interest in sports, especially in athletics and in swimming records, Olympic-or world-records, are always something special and observable [2,3].

An area, where records are also important is the context of global warming. The question, how a changing climate affects the number of recorded temperatures we encounter has bothered both the general public and researchers [4]. However records also play an important role in other sciences like physics [5], biology [6] and economy [7]. Moreover, it was recently possible to analyse and model the statistics of records in stock prices [8]. The theory of records in time series of independent and identically distributed (i.i.d.) random variables was already developed many years ago [9].

In this paper, we apply the records to observe the evolution of some new marine species for which few information are available. We would determine if some species are isolated (sporadic) by using the methods of records. More precisely, a species is recorded or have a record if the number of occurrence of this specie exceeds the last registered number. The observations can be presented with a high variability that may come from some random phenomena such as the transfer, the environment or others. These can be one-dimensional, like (the number of new species observed in a period of time in the whole region or an interval of time between the occurrence of the observation of two species), or multi-dimensional (i.e. the number of observations in a period per region).

*Percnon gibbesi* is an opportunistic feeder, which feeds primarily on algae such as filamentous algae, calcareous algae, corticated macrophytes, and folios algae [10] at the shallow (0-8m) infra-littoral rocky shore but also on invertebrates such as pagurid crabs, polychaeta, mollusks (mainly gastropods), other crustacean species and jelly fish [11]. It can be found only in rocky bottoms, especially on boulders with no or no dense vegetation and it has a mainly pre-nocturnal activity. Meanwhile, it creates dense populations which displace other indigenous crabs, like the warty crab, *Eriphia verrucosa*. The success of its invasion and establishment in the Mediterranean sea seems to be based on (a) the flexibility on the food source, (b) the high prolificacy of the spider crab, (c) the large duration of the larval stage [12] which allows the long dispersal of the species, (d) the exceptionally large megalopa from which a robust first crab stage is produced [13], (e) favorable environmental conditions, especially due to the warming of the Mediterranean Sea, in particular after 1998 [14], and (f) absence of substantial competition [15] and predation.

The paper is organized as follows. In section 2, we present a brief of the theory of records. Then, we define the record as a sum of independent random variables. This is important especially in the presence of rare events that are independent and not regular simultaneously. In the section 3, we study the distribution of the number of records when the random variables are independent and identically distributed, which will be used as a statistical test. In section 4, we evaluate the first moment of the records which is approximately in the complexity of  $\ln(n)$ . In section 5, we use a hypothesis test to identify if *Percnon gibbesi* is sporadic or not by studying the case in Greece.

## 2. METHODOLOGY AND HYPOTHESIS

The theory of records and the order statistic are extremely related. In statistics, the  $k$ -th order statistic of a statistical sample is equal to the  $k$ -th smallest value. Important special cases of the order statistics are the minimum and the maximum value of the sample. Formally, let consider a time series  $\{X_1, \dots, X_n\}$  of random variables which can, for example, be a series of temperatures, stock prices, sports results or some other kind of measurement process.

For a sample of size  $n$ , the  $n$ -th order statistic is given by  $M_n = \max(X_1, \dots, X_n)$ . The theory of records handles the exact distributions of some statistic (the number of records values and instants of records) that are not depending on the distribution of the observations. Also this method is able to evaluate the probability distribution for a small sample of observations.

The aim of this work is to develop a mathematical framework for evaluating the frequency of the sally light foot crab *Percnon gibbesi* in the Mediterranean Sea, then the application of a non parametric test on some real world data collected from Greece to identify that *Percnon gibbesi* is sporadic or not.

We consider a renewal process  $\{S_n\}_{n \geq 0}$  on  $\mathbb{R}^+$  that represents the successive occurrence time of a new specie. For example  $S_0$  represents the time of the first occurrence of the species. Let consider  $\{\Delta S_n\}_{n \geq 1}$  be the difference between two successive times of occurrence of the species i.e.  $\Delta S_n = S_n - S_{n-1}$  for  $n=1,2,\dots$ . The set of  $\Delta S_n$  for  $n=1,2,\dots$  form a time series of random variables that are independent and identically distributed.

In the framework of random variables independent identically distributed  $\{\Delta S_n\}$ , the quick evolution of the species corresponds to the smallest values of  $\Delta S_n$ . This is why we find the similar values of the time, i.e.  $S_n$  and  $S_{n-1}$  are close to each other. Hence, we use the inverse  $(\Delta S_n)^{-1}$  in statistical test since we are looking for the highest records. In other words, when  $S_n$  approaches to  $S_{n-1}$  then the difference will be small, thus the inverse that represents the records increases.

For example, the Table 1 shows an example of computing the waiting time  $(\Delta S_n)$  and its inverse  $(\Delta S_n)^{-1}$  between two successive cases of the observation of specie in a region.

**Table 1. Example of calculating the waiting time  $(\Delta S_n)$  and its inverse  $(\Delta S_n)^{-1}$**

Observed time $S_n$	Month/Year	$\Delta S_n = S_n - S_{n-1}$	$(\Delta S_n)^{-1} = 1/\Delta S_n$
$S_1$	06/2001	<i>Trivial record</i>	<i>Trivial record</i>
$S_2$	07/2002	395	0.002
$S_3$	04/2004	640	0.001

### 2.1 Probability Distribution of the Number of Records

Let consider  $M_n = \max(X_1, X_2, \dots, X_n)$  the maximum of the sequence of i.i.d. random Variables  $\{X_k\}_{k \leq n}$ , where  $X_n = (\Delta S_n)^{-1}$ . Let define the number of records R such as

$$R = \sum_{j=1}^n \delta_j \tag{3.1}$$

where  $\delta_j$  is the indicator function of the records defined as follow:

$$\delta_j = \begin{cases} 1 & \text{if } \Delta S_j^{-1} > \max(S_1^{-1}, \dots, \Delta S_{j-1}^{-1}) \\ 0 & \text{elsewhere} \end{cases} \tag{3.2}$$

Returning to the Table 1, we have  $(\Delta S_2)^{-1} = 0.002$ ,  $(\Delta S_3)^{-1} = 0.001$  and  $(\Delta S_4)^{-1} = 0.008$ . The first case of observation gives  $\delta_1 = 1$  (since the specie is found). Since  $(\Delta S_3)^{-1}$  is smaller than  $(\Delta S_2)^{-1}$  then  $\delta_3 = 0$ . Hence,  $R = 1$  and we continue in the same way, let consider  $(\Delta S_4)^{-1} = 0.008$ , this is greater than the maximum of the inverse time observed then  $\delta_4 = 1$  and the number of records becomes  $R = 2$ . Studying the distribution of the number of records helps us to show the independency between the observations. This is given by:

$$P_n(R = r) = \frac{s(n, r)}{n!} \tag{3.3}$$

Where  $\{s(n, r)\}$  is the Stirling number. This can be defined by comparison of the following polynomials [16]:

$$\prod_{j=0}^{n-1} (s - j) = \sum_{r=0}^n s(n, r) s^r \tag{3.4}$$

Moreover, Glick [16] gives a simplified expression for the distribution of records. This expression is obtained by recursion and given by:

$$P_j(R = r) = \left(1 - \frac{1}{j}\right) P_{j-1}(R = r) + \left(\frac{1}{j}\right) P_{j-1}(R = r - 1) \tag{3.5}$$

For all  $r \geq 1$  and  $j \geq 2$  where the initial probabilities are  $P_1[R=0]=0$  and  $P_1[R=1]=1$ . The simulations shown in Fig. 1 illustrate a comparison between the probability distributions of the number of records for different values of  $n$ . We can also remark that the mode of the distribution, that is the highest frequency, increases with  $n$  and the distribution spreads to the left. Table 2 gives the probability to observe at least  $r$  records among  $n$  variables. This table is used to compute the significance level to test the sporadic species.

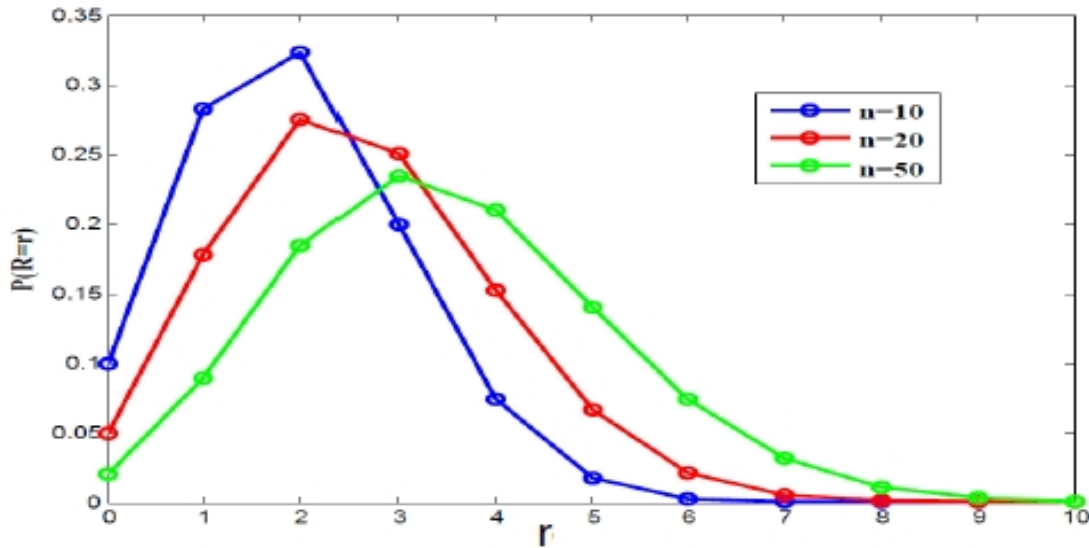


Fig. 1. Comparison of several probability distribution of the number of records

Table 2.  $\{P_n(R \geq r)\}_{r \leq 9}$  derived from(3.3) for various values of  $n$

n\r	1	2	3	4	5	6	7	8	9
10	1	0.9	0.6171	0.2939	0.0945	0.0203	0.0029	0.0003	0
20	1	0.95	0.7726	0.4978	0.2470	0.0944	0.0280	0.0065	0.0012
30	1	0.9667	0.8346	0.5999	0.3475	0.1614	0.0605	0.0185	0.0047
40	1	0.975	0.8687	0.6627	0.4182	0.2163	0.0922	0.0326	0.0097
50	1	0.98	0.8904	0.7060	0.4712	0.2617	0.1212	0.0472	0.0156
100	1	0.99	0.9382	0.8124	0.6194	0.4082	0.2315	0.1133	0.0482

### 2.2 Moments of the Number of Records

Following Renyi [17], the sequence of the indicators of records  $\{\delta_j\}_{j \geq 1}$  is independent. The distribution is then given by:

$$P(\delta_j = 1) = \frac{1}{j} \tag{4.1}$$

The number of records in the interval  $[1, n]$  has approximately a complexity of  $O(\ln n)$ . In other words, when  $n$  increases the appearance of new records becomes improbable. This can be explained by:

$$E(R) = \sum_{j=1}^n \frac{1}{j} \approx \ln n \tag{4.2}$$

where  $E(R)$  is the mean of the number of records. In statistics, we can define the species that are Rare with respect to their time of appearance. Then the instants of records increase as an exponential function  $en$  [9]. Hence, the period of time separating two successive records must be as long as possible to allow the observation of the evolution of the new specie. The time separating records lies with  $n$ . In other words, more the time runs, the more the records become infrequent. Fig. 2 presents the mean of the number of records as function of  $n$ .

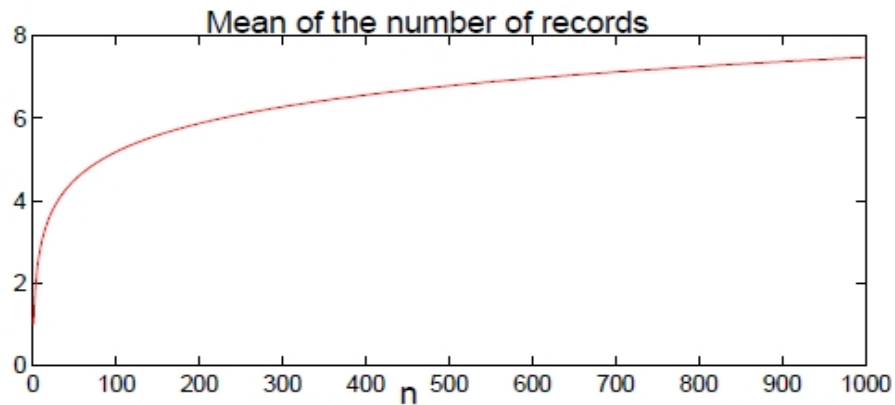


Fig. 2. In a sequence of 1000 observations we expect to observe about 7 records.

### 2.3 A Real World Data Application

In order to study the evolution of *Percnon gibbesi*, we consider the inverse of the waiting time  $\Delta(Sn)^{-1}$  between two observed successive species. *Percnon gibbesi* is observed in several regions in the Mediterranean region as in Italy, Spain, Greece, Turkey [11]. Here in this section, we will illustrate the method explained below on some real world data. We will consider, as an example, a region in the Mediterranean Sea, Greece. We apply the records method to figure out that the observed species of *Percnon gibbesi* are sporadic or not from a small number of observed species. The first case is founded in Greece in 2001 and the last in 2012. Notice that we are using monthly data between 2001 and 2012. Fig. 3 presents the

inverse of the waiting time for which we observe 6 records. We can evaluate the probability of the number of records for the six observed records using the formula (3.3). Then, we use the number of record test at a risk level  $\alpha=5\%$ , where  $\alpha$  is the probability to identify that *Percnon gibbesi* is not sporadic. Fig. 2 presents the mean of the number.

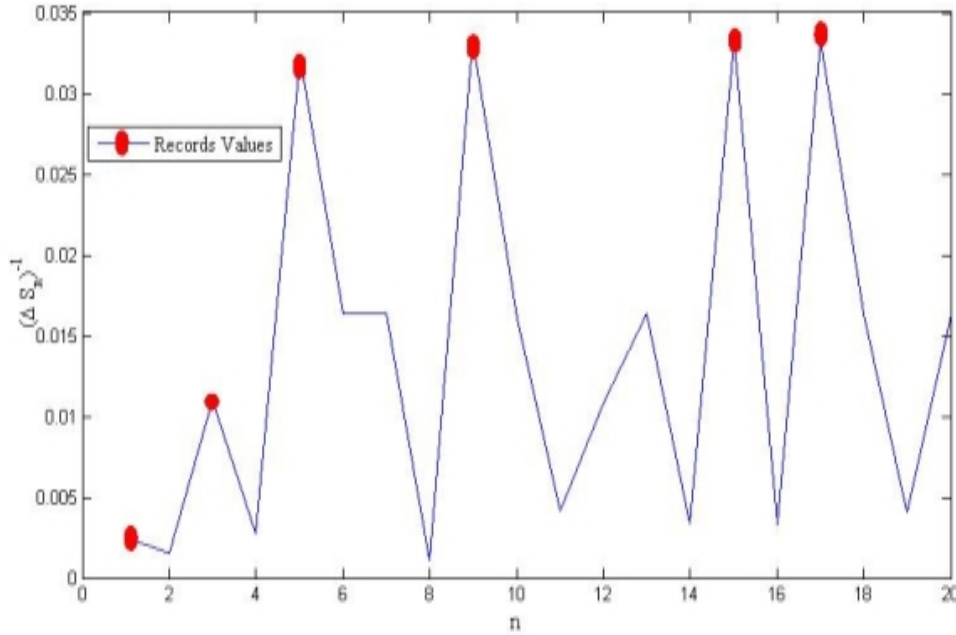


Fig. 3. Successive observed records for the inverse time waiting  $(\Delta S_n)-1$ .

Table 3. Probability distribution and cumulative probability distribution of the number of records in Greece for  $n=15$  and  $1 \leq r \leq 7$

<b>R</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
$P_n(R = r)$	0.0667	0.2168	0.2999	0.2378	0.1221	0.0433	0.011
$P_n(R > r)$	1	0.9333	0.7166	0.4167	0.1789	0.0567	0.0134

In particular according to Table 3, we compute the probability distribution for each value of the records between 1 and 7 and we can deduce  $P_n(R \geq 4) = 0.4167$  for  $n=15$ . Therefore the species are sporadic.

Then we evaluate the cumulative probability  $P_n(R \geq r)$  to determine the critical region in which we consider that *Percnon gibbesi* is not sporadic. To identify that *Percnon gibbesi* is sporadic, we compare the value of the probability with a level risk  $\alpha=0.05$ . Here, we observe that  $P_n(R \geq 6) = 0.0567$  is the nearest value to  $\alpha=0.05$ . Moreover, if we observe more records values, we can easy check the evolution of *Percnon gibbesi* with a probability less than 0.05 (i.e.  $P_n(R \geq 7) = 0.0134$ ). Consequently, we can conclude that *Percnon gibbesi* in Greece is a species that spreads with the time.

### 3. CONCLUSION

In conclusion, we have presented a new method in statistical marine ecology based on the number of records to determine the sporadic species from the first observations. This method has major advantages and is totally robust since the distribution of the statistical test is exactly calculated for each value of  $n$ , and this distribution is independent of the cumulative distribution function involved in the distribution of observations. According to the classical theoretical result which shows that the number of record becomes large enough, then a few records are easily observed in sporadic cases, for small, but as increases, it becomes very difficult to observe additional records. The record approach is particularly suitable for a small number of observations since it does not take into account the smallest data which may be particularly small for a new event. However, when becomes larger, it could be interesting to take into account more observations and not only the record values.

### COMPETING INTERESTS

Authors have declared that no competing interests exist.

### REFERENCES

1. Glenday C. Guinness World records; 2012.
2. Daniel G, John G, Dieter S. Evolution of Athletic Records: Statistical Effects versus Real Improvements. *Journal of Applied Statistics*. 2007;34:529-545.
3. Daniel G, John G, Dieter S. Sports statistics: Trends and random fluctuations in athletics. *Nature*. 2002;417:506-506.
4. Rasmus E, Benestad. Records values, non-stationary tests and extreme value distributions. *Global and Planetary Change*. 2004;44:11-26.
5. Paolo S. Coarse-graining dynamics: Continuous Time Random Walks vs. Record Dynamics, *Europhysics Letters*. 2013;101:30004.
6. Stuart K. The origins of order. *Theoretical Biology*. 1987;128.
7. Balevy G, Nagaraja H. Characterizations in a random record model with a non-identically distributed initial record, Technical report; 2005.
8. Wergen G, Bogner M, Krug J. Record statistics for biased random walks, with an application to financial data. *Physical Review E*. 2011;83:051109.
9. Nevzorov VB. Records: Mathematical Theory, 164. American Mathematical Society; 2001
10. Puccio V, Relini M, Azzuro E, Orsi R. Feeding habits of *Percnon gibbesi* (H. Milne Edwards, 1853) in the Sicily Strait. *Hydrobiologia*. 2006;557:79-84.
11. Katsanevakis S, Poursanidis D, Yokes B, Vesna M, Beqiraj S, Kashta L, Ramzi S, Zakhama-Sraieb R, Benamer I, Bitar G, Bouzaza Z, Magni P, Bianchi CN, Tsiakkios L, Zenetos A. Twelve years after the first report of the crab *Percnon gibbesi* (H. Milne Edwards, 1853) in the Mediterranean: current distribution and invasion rates. *Journal of Biological Research*. 2011;16:224-236.
12. Paula J, Hartnoll RG. The larval and post-larval development of *Percnon Gibbesi* (Crustacea, Brachyura, Grapsidae) and the identity of the larval genus *Pluteocaris*. *Journal of Zoology, London*. 1989;218:17-37.
13. Hartnoll RG. Megalopae and early postlarval stages of East African *Percnon* (Decapoda: Brachyura: Grapsidae). *Journal of Zoology, London*. 1992;228:51-67.



14. Raitzos D, Beaugrand G, Georgopoulos D, Zenetos A, Theochairs A, Papathanassiou E. Global climate change amplifies the entry of tropical species into the eastern Mediterranean Sea. *Limnology and Oceanography*. 2010;55:1478-1484.
15. Pipitone C, Badalamenti F, Sparrow A. Contribution to the knowledge of *Percnon gibbesi* (Decapoda, Grapsidae), an exotic species spreading rapidly in Sicilian waters. *Crustaceana*. 2001;74:1009-1017.
16. Abramowitz M, Stegun Irene A. Hand book of mathematical functions with formulas, graphs, and mathematical tables, New York, Wiley; 1972.
17. Renyi A. On the extreme elements of observations. *Oszt. Kzl.* 1962;12:105-121.

---

© 2014 Khraibani et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*

*The peer review history for this paper can be accessed here:*  
<http://www.sciencedomain.org/review-history.php?iid=578&id=22&aid=5094>