British Journal of Mathematics & Computer Science
9(6): 473-482, 2015, Article no.BJMCS.2015.218
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org

The Law of Large Numbers: A Multicriterial Analysis

Ana-Maria Rîtea[1*]

[1]*Department of Mathematical Analysis and Probability, Transylvania University of Braşov, Str. Iuliu Maniu Nr.50, Braşov–500091, România.*

Article Information

DOI: 10.9734/BJMCS/2015/17980
Editor(s):
(1) Vyacheslav Pivovarchik, Department of Applied Mathematics and Computer Science, South-Ukrainian National Pedagogical University, Ukraine.
(2) Dijana Mosi´c, Department of Mathematics, University of Niˇs, Serbia.
(3) Paul Bracken, Department of Mathematics, The University of Texas-Pan American Edinburg, TX 78539, USA.
Reviewers:
(1) Anonymous, University of North Carolina at Charlotte, USA.
(2) Anonymous, Egerton University, Kenya.
(3) Anonymous, Marie Curie-Skłodowska University, Poland.
Complete Peer review History: http://sciencedomain.org/review-history/9771

Original Research Article

Received: 31 March 2015
Accepted: 21 May 2015
Published: 15 June 2015

Abstract

The paper presents the law of large numbers followed by a multicriterial analysis of the six theorems that compose the weak form. This group of theorems work on issues of sequences limit laws of sums of random variables in the sense of convergence almost surely of such sequences.

Keywords: Bernoulli; Chebyshev; the law of large numbers; Markov; multicriterial analysis; Poisson.

1 Introduction

The fundamental limit theorems of the probability theory can be classified into two groups. A group dealt with the limit laws problems of the sums of sequences of random variables in the sense of convergence almost sure of such sequences. The Central Limit Problem development for independent terms can be divided into three periods. The first refers to the case of Bernoulli with corresponding to limit theorems of Bernoulli, DeMoivre and Poisson. The first two theorems gave birth to concepts, from which the classic problem of the Central Limit Law of Large Numbers and normal convergence.

The development is to simplify the existing demonstrations, modeling and consolidation; determining the general concepts underlying their results and expanding their fields. Analysis of this growth will put into relief the role and the interconnections of fundamental limit theorems.

*Corresponding author: ana_ritea_maria@yahoo.com;

First, the italian mathematician G. Cardano (1501-1576) affirmed, without demonstrations, that the precision of an empirical statistics tends to improve with the sample sizes. This was formalized then as a law of large numbers. Later on, another great mathematician, J. Bernoulli (1654-1705) established in his work "Ars conjectandi" published in 1713, that the new mathematical theory it would be the foundation of the study of mass phenomenon. It took him more than 20 years to develop a rigorous demonstration of this, which was published in opera that we cited earlier. By "theorem of large numbers" or "law of large numbers" this is a direct connection between frequency and probability, after a large number of samples. This theorem is the foundation of mathematical statistics. Obviously, probability theory and statistics are in a continuous development. Thus, it is justified the application of probability theory in so many areas.

The purpose of this paper is to expose the six theorems that make up the weak form of the law of large numbers. Here, we detail the present theorems demonstrations, then we make a multicriterial analysis thereof. From the discussions with researchers in other domains than statistics or probability, the author noticed that they encountered difficulties in applying the theorems in their fields of research, so using multicriterial analysis, we make a ranking to urge the one who applied, from which point of view to take one or the other. As we have mentioned, the applicability of probability theory being so ample in so many domains; we come to be helpful to a nonmatematician to help him making a decision as best and fastest. We emphasize that this multicriterial analysis applied in this paper is an own contribution.

2. The Law of Large Numbers (The Weak Form)

Theorem 1. [1,2] *Let $(Y_n)_{n\in\mathbb{N}^*}$ be a string of random variables and mean values and finite variances*

$$E[Y_n] = E_n < \infty, \qquad Var[Y_n] = \sigma_n^2 < \infty.$$

If there exist a constant M such that

$$\lim_{n\to\infty} E_n = M$$

and

$$lim_{n\to\infty}\,\sigma_n^2 = 0 \tag{1}$$

Then $\quad \{Y_n\} \xrightarrow{p} M.$

Demonstration

We observe that for any $\varepsilon > 0$

$$\{|Y_n - M| > \varepsilon\} \subseteq \left\{|Y_n - M_n| > \tfrac{\varepsilon}{2}\right\} \cup \left\{|M_n - M| > \tfrac{\varepsilon}{2}\right\}.$$

we can write

$$P(\{|Y_n - M| > \varepsilon\}) \le P\left(\left\{|Y_n - M_n| > \tfrac{\varepsilon}{2}\right\}\right) + P\left(\left\{|M_n - M| > \tfrac{\varepsilon}{2}\right\}\right). \tag{2}$$

Applying random variable Y_n on Chebyshev's inequality we obtain

$$P(\{|Y_n - M| > \varepsilon\}) \le \tfrac{Var[Y_n]}{\varepsilon^2} = \tfrac{\sigma_n^2}{\varepsilon^2}. \tag{3}$$

Since

$$\lim_{n\to\infty} E_n = M$$

It results that $\forall \varepsilon > 0, \exists N(\varepsilon) \in \mathbb{N}$ such that $\forall n \geq N(\varepsilon)$ the inequality $|E_n - M| > \frac{\varepsilon}{2}$ becomes imposible, ie

$$\forall n \geq N(\varepsilon): P(\{|E_n - M| > \varepsilon\}) = 0.$$

Passing to the limit in (2), taking into account of (3) and the ipothesis (1) it results that

$$\lim_{n\to\infty} P(\{|Y_n - M| > \varepsilon\}) = 0.$$

Theorem 2. (**the weak form of the Jacob Bernoulli theorem**)

The first formulation [2,3]: *If A is an event caracterized by $P(A) = p \neq 0$ and $\frac{S_n}{n}$ of the relative frequency achievement of the event A in the first n samples of an infinite number of independent samples, then*

$$\forall \varepsilon > 0 : \lim_{n\to\infty} P\left(\left\{\left|\frac{S_n}{n} - p\right| > \varepsilon\right\}\right) = 0$$

or, equivalent,

$$\forall \varepsilon > 0 : \lim_{n\to\infty} P\left(\left\{\left|\frac{S_n}{n} - p\right| \leq \varepsilon\right\}\right) = 1.$$

The second formulation: [4,5] *If $\{X_n\}_{n\in\mathbb{N}}$ is a set of independent random variables such that $X_n: B_i(n, p), n \in \mathbb{N}$, then the string of random variables converges in probabiliy to p.*

$$\left\{\frac{1}{n}\sum_{k=1}^{n} X_k\right\}_{n\in\mathbb{N}}$$

Demonstration

We demonstrate Bernoulli's theorem in the second form. We note

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

and considering that

$$E[Y_n] = \frac{1}{n}\sum_{k=1}^{n} E[X_k] = \frac{1}{n} np = p,$$

$$Var[Y_n] = \frac{1}{n^2}\sum_{k=1}^{n} Var[X_k] = \frac{npq}{n^2} = \frac{pq}{n}.$$

So

$$\lim_{n\to\infty} E[Y_n] = p$$

and

$$\lim_{n\to\infty} Var[Y_n] = 0$$

which indicates that the hypotheses of the theorem 1 are satisfied, then

$$\left\{\frac{1}{n}\sum_{k=1}^{n} X_k\right\} \xrightarrow{p} p.$$

Remark 1. The first statement of the Bernoulli theorem approach is theoretical justification made between the concepts of relative frequency and probability of an event. It expresses the fact that for very large strings evidence any significant difference between the relative frequency and probability of an event is very improbable. So convergence of relative frequencies to probability of the event it does not to be understood as a common numerical convergence. It is not apparent from the above theorem and is not accepted the intuitive idea that with increasing number of relative samples frequency values are necessarily closer to probability.

Remark 2. The two formulations result one of the other if we consider the absolute frequency of realization of A in n samples. Let X_n be a random variable distributed $B_i(n,p)$. If we note

$$S_n = \sum_{k=1}^{n} X_k$$

then

$$\frac{S_n}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

and

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{S_n}{n} - p\right| > \varepsilon\right\}\right) = 0$$

which means that the

$$\left\{\frac{1}{n}\sum_{k=1}^{n} X_k\right\} \xrightarrow{p} p$$

Remark 3. The theorem result is the theoretical justification of the definition of a parameter estimation of a distribution.

Theorem 3. [6,7] (**Chebyshev**) *If $(X_n)_{n\in\mathbb{N}^*}$ is a set of independent random variables two by two having dispersions bounded by the same constant*

$$Var[X_n] \le c^2, \qquad n \in \mathbb{N}$$

then

$$\forall\varepsilon > 0 : \lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E[X_k]\right| > \varepsilon\right\}\right) = 0$$

or, equivalent,

$$\forall \varepsilon > 0: \lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E[X_k]\right| \le \varepsilon\right\}\right) = 1,$$

i.e. occurs the convergence in probability string of random variables,

$$\left\{\frac{1}{n}\sum_{k=1}^{n}(X_k - E[X_k])\right\} \xrightarrow{p} 0$$

Demonstration

If

$$Y_n = \frac{1}{n}\sum_{k=1}^{n}(X_k - E[X_k])$$

then

$$E[Y_n] = E[\frac{1}{n}\sum_{k=1}^{n}(X_k - E[X_k])] = \frac{1}{n}\sum_{k=1}^{n} E[X_k] - \frac{1}{n}\sum_{k=1}^{n} E[X_k] = 0$$

and because the random variables X_n are independent two by two, it results that $(X_k - E[X_k])$ are independent two by two and therefore

$$Var[Y_n] = Var\left[\frac{1}{n}\sum_{k=1}^{n}(X_k - E[X_k])\right] = \frac{1}{n^2}\sum_{k=1}^{n} Var(X_k - E[X_k]) = \frac{1}{n}\sum_{k=1}^{n} Var[X_k].$$

Applying to random variable Y_n the Chebyshev's inequality and taking into account the assumptions of the problem, we can write

$$P\left(\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E[X_k]\right| > \varepsilon\right\}\right) \le \frac{1}{n^2\varepsilon^2}\sum_{k=1}^{n} Var[X_k] < \frac{nc^2}{n^2\varepsilon^2} = \frac{c^2}{n\varepsilon^2}$$

and so

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E[X_k]\right| > \varepsilon\right\}\right) \le \lim_{n\to\infty}\frac{c^2}{n\varepsilon^2} = 0$$

Theorem 4. [8,2] **(Markov)** *If the random variables $X_1, X_2, \dots, X_n, \dots$ have the property that:*

$$\frac{Var\left[\sum_{i=1}^{n} X_i\right]}{n^2} \to 0, \qquad n \to \infty$$

then we have the law of large numbers.

Demonstration

It must be demonstrated that the string $\{X_n\}_{n\in\mathbb{N}}$ verifies the relation:

$$\forall \varepsilon > 0 : \lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E[X_k]\right| \le \varepsilon\right\}\right) = 1.$$

Applying Chebyshev's inequality for the random variables:

$$Y = \sum_{i=1}^{n} \frac{X_i}{n}, Var[Y] = Var\left[\sum_{i=1}^{n} \frac{X_i}{n}\right] = \frac{1}{n^2} Var\left[\sum_{i=1}^{n} X_i\right] \to 0, \;\; n \to \infty$$

Therefore $Var[Y] < \infty$ and then Chebyshev's inequality holds:

$$P(|Y - E[Y]| \le \varepsilon) \ge 1 - \frac{1}{\varepsilon^2}\frac{1}{n^2} Var\left[\sum_{i=1}^{n} X_i\right].$$

But

$$\frac{1}{n^2} Var\left[\sum_{i=1}^{n} X_i\right] \to 0, n \to \infty$$

Passing to the limit in the previous inequality, we obtain the assertion of the theorem.

Remark 4. [9] In the particular case when the random variables are independent, provided the statement is written:

$$\frac{Var\left[\sum_{i=1}^{n} X_i\right]}{n} \to 0, n \to \infty$$

Theorem 5. [4,2] (**Khincin**) *If* $\{X_n\}_{n\in\mathbb{N}}$ *is a set of independent random variables with the same distribution two by two and having finite dispersions, then*

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - M\right| > \varepsilon\right\}\right) = 0$$

where $M = E[X_n], n \in \mathbb{N}$.

Demonstration

If the random variable $\{X_n\}_{n\in\mathbb{N}}$ follow the same distribution, it results that $Var[X_n] = \sigma^2$, $n \in \mathbb{N}$, which implies satisfying the assumptions of Chebyshev's theorem.

Remark 5. This result substantiates the usual theory of approximation of the average of statistical characteristic of a population studied by the arithmetic mean of the observed values. Indeed, if $X_1, X_2, \ldots, X_n, \ldots$ are seen as potential values of a random variable X obtained by a number of independent samples, the theorem shows that their average converges to $M = E[X]$.

As a special case of Chebyshev's theorem we obtain Poisson's theorem.

Theorem 6. [4,9] (**Poisson**) *Let A an event of which the probability to output varies during a series of independent samples so that the sample of rank k,* $P(A) = p_k, k = 1,2, \dots$ *and let* $\frac{S_n}{n}$ *the frequency of execution of the event A in the first n samples. Then*

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{S_n}{n} - \frac{p_1 + p_2 + \cdots + p_n}{n}\right| > \varepsilon\right\}\right) = 0$$

Demonstration

We define for each $k = 1,2, \dots, n, \dots$ the random variable X_k takes the values 1 or 0 as A is realized or not in the sample with the rank k. Hereupon, X_k is definited

$$X_k: \begin{pmatrix} 0 & 1 \\ 1 - p_k & p_k \end{pmatrix}$$

Noting with S_n the number of realisations of A in the first n samples, we denote that

$$f_n(A) = \frac{S_n}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

Because $X_1, X_2, \dots, X_n, \dots$ are independent random variables and because

$$E[X_k] = p_k, Var[X_k] = E[X_k^2] - E[X_k]^2 = p_k - p_k^2 = p_k(1 - p_k) \le \frac{1}{4}$$

it results that the string $\{X_n\}_{n\in\mathbb{N}}$ verify the hypotheses of the Chebyshev's theorem and so

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{S_n}{n} - \frac{p_1 + p_2 + \cdots + p_n}{n}\right| > \varepsilon\right\}\right) = \lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E[X_k]\right| > \varepsilon\right\}\right) = 0.$$

3. The Multicriterial Analysis of the Dynamics That Make up the Application Form Weak Theorems of Law of Large Numbers

Multicriterial analysis can serve very well in obtaining all kinds of rankings with subjects in the same area or in different areas of activity, contemporaries or not, in that the subjectivism is greatly removed. It is very important that this kind of analysis is in relation to the criteria chosen, an analysis that gives – in a large proportion – an objective character of its results.

We have the following variants:

- Variant (a): The extent of which one applies more frequently theorems
- Variant (b): The difficulty to apply the 6th theorems
- Variant (c): The reliable of application of the 6th theorems

6 criteria have been chosen (Table 1):

Table 1. The criteria table

Current no.	The theorem
1.	T. Bernoulli 1
2.	T. Bernoulli 2
3.	T. Chebyshev
4.	T. Markov
5.	T. Khincin
6.	T. Poisson

Based bet on score, weighting of the criteria resulted as follows (Table 2):

Table 2. The weights of the criteria

	T. Bernoulli 1	T. Bernoulli 2	T. Chebyshev	T. Markov	T. Khincin	T. Poisson	Points	Level	γ_i
T. Bernoulli 1	1/2	1	0	0	0	1/2	2	6	0,5
T. Bernoulli 2	0	1/2	0	1	1/2	1/2	2,5	3,5	1
T. Chebyshev	1	1	1/2	1	0	0	3,5	2,5	2,4
T. Markov	1	0	0	1/2	0	1	2,5	3,5	1
T. Khincin	1	1/2	1	1	1/2	0	4	1	3,8
T. Poisson	1/2	1/2	1	0	1	1/2	3,5	2,5	2,4

The γ_i weighting coefficients can be calculated with different formulas. We chose to use FRISCO practice formula (empirical formula given by a renowned creative group from San Francisco – US) that has been recognized worldwide as being the most performance and is long used.

Therefore, with,

$$\boldsymbol{\gamma_i = \frac{p + \Delta p + m + 0,5}{-\Delta p' + \frac{N_{crt}}{2}}}$$

where,

- $\boldsymbol{p}$ is the sum of points obtained (on line) of the considered element
- $\boldsymbol{\Delta p}$ the difference between the score of the considered element and the score at the top level element; if the element taken into account is the one located on the top floor, results $\boldsymbol{\Delta p}$ with the value 0
- $\boldsymbol{m}$ the number of outclassed criteria (exceeded from terms of score) the by the criteria taken into account
- $\boldsymbol{N_{crt}}$ the number of considered criterion
- $\boldsymbol{\Delta p'}$ the difference between the score of the considered element and the score of the first element (resulting with a negative value); taken into account if the item is located on the first level, $\boldsymbol{\Delta p'}$ results with the value 0

We obtain

$$\gamma_{Bernoulli\ 1} = \frac{2 + (2-2) + 0 + 0,5}{-(2-4) + \frac{6}{2}} = \frac{2,5}{5} = 0,5$$

$$\gamma_{Bernoulli\ 2} = \frac{2,5 + (2,5-2) + 1 + 0,5}{-(2,5-4) + \frac{6}{2}} = \frac{4,5}{4,5} = 1 = \gamma_{Markov}$$

$$\gamma_{Poisson} = \gamma_{Cebâşev} = \frac{3{,}5 + (3{,}5 - 2) + 3 + 0{,}5}{-(3{,}5 - 4) + \frac{6}{2}} = \frac{8{,}5}{3{,}5} = 2{,}4$$

$$\gamma_{Hincin} = \frac{4 + (4 - 2) + 5 + 0{,}5}{-(4 - 4) + \frac{6}{2}} = \frac{11{,}5}{3} = 3{,}8$$

It is noted that the main diagonal of the array contains only quadratic criteria for scoring 1/2 values because no criteria may be more important or less important than the criteria itself.

According to the criteria there were given the following notes for each variant N_i (Table 3).

Table 3. The scores of variants

Criteria	Variant (a)	Variant (b)	Variant (c)
	$\mathbf{N_i}$	$\mathbf{N_i}$	$\mathbf{N_i}$
T. Bernoulli 1	7	6	8
T. Bernoulli 2	9	9	9
T. Chebyshev	10	10	7
T. Markov	6	5	4
T. Khincin	9	4	5
T. Poisson	8	7	7

It may take into account different weight now and consequence of each criterion, complementing and enhancing the Table 4 above notes (lines) with the coefficient of importance:

Table 4. The products between notes and weighting coefficients

Criteria	γ_i	Variant (a)		Variant (b)		Variant (c)	
		N_i	$N_i \times \gamma_i$	N_i	$N_i \times \gamma_i$	N_i	$N_i \times \gamma_i$
T. Bernoulli 1	0,5	7	3,5	6	3	8	4
T. Bernoulli 2	1	9	9	9	9	9	9
T. Chebyshev	2,4	10	24	10	24	7	16,8
T. Markov	1	6	6	5	5	4	4
T. Khincin	3,8	9	34,2	4	15,2	5	19
T. Poisson	2,4	8	19,2	7	16,8	7	16,8
Final clasament			95,9		73		69,6

4. Conclusions

Multicriterial analysis technique is useful in the composition of an ranking, while qualitatively and quantitatively, of product variants, objects, methods, models, equipment, structures, creations, etc. A first valence would be that the result of such analysis in order not only put options, but it quantifies in value terms.

Rankings, to a large extent, have a high degree of subjectivity and aims the most of the time only the qualitative aspect. Multicriterial analysis technique gives, from the viewpoint of its user, results found to a great extent objectives (ie, this technique objectifies in an important measure the results).

It is noted that after the ranking did, the theorems of the weak large numbers law are preferred to be taken in the variant (a).

We want to emphasize that point III is our own creation. Efforts have been made to develop this multicriterial analysis applied to the law of large numbers, hoping that we will develop it in the future. It is useful specialy for nonmathematicians, because they can decide earlier which of the theorems they have to apply in genetics, management, mechanics etc.

Competing Interests

Author has declared that no competing interests exist.

References

[1] Klebaner C. Fima. Introduction to stochastic calculus with applications. Imperial College Press, London; 2004.

[2] Pletea A, Popa L. Teoria probabilităţilor. Ed. Universităţii Tehnice Gh. Asachi, Iaşi; 1999.

[3] Orman GV. Handbook of limit theorems and stochastic approximation. Transilvania University Press, Braşov; 2003. ISBN 973-635-227-7.

[4] Cenuşa G, Burlacu V. Săcuiu I. Teoria probabilităţilor şi statistică matematică. Editura ASE; 1999.

[5] Geiss C, Geiss S. An introduction to probability theory. University of Jyvaskyla; 2004.

[6] Feller W. An introduction to probability theory and its applications. Third Edition. John Wiley & Sons, Inc., New York. 1968;1. ISBN 0-471-25711-7.

[7] Loéve M. Fundamental limit theorems of probability theory. The Annals of Mathematical Statistics. 1950;21(3):321-338.

[8] Capinski M. Kopp E. Measure, integral and probability. Springer-Verlag; 2003.

[9] Øksendal B. Stochastic differential equations – An introduction with applications. Springer, Berlin; 2007;113-p.137. ISBN-10:3-540-04758-1, ISBN-13:978-3-540-04758-2.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/9771