УДК 529.21

Колеснік О. В.

https://doi.org/10.18523/2617-70806202335-38

ЦЕНТРАЛЬНА ГРАНИЧНА ТЕОРЕМА ДЛЯ КІЛЬКОСТІ РЕКОРДІВ У F^α-СХЕМІ

У цій роботі розглянуто твердження, які стосуються виконання центральної граничної теореми (ЦГТ) для кількості рекордів у послідовності незалежних випадкових величин в рамках F^α-схеми рекордів. Наведено методику знаходження точних асимптотичних виразів для математичного сподівання та дисперсії, якими можна замінити справжні характеристики у ЦГТ.

Розглянуто конкретний приклад степеневого зростання експонент F^α-схеми і побудовано ЦГТ лише у термінах моменту спостереження та степені зростання.

У статті є 4 теореми з повним доведенням. Теорема 1 пов'язує математичне сподівання та дисперсію з накопиченою інтенсивністю F^α-схеми. Теорема 2 встановлює виконання ЦГТ у загальному вигляді, а теорема 4 – для конкретного випадку.

Ключові слова: незалежні однаково розподілені випадкові величини, F^α-схема, рекорди, кількість рекордів, центральна гранична теорема.

Вступ

Розглянемо послідовність $\{X_k, k \geq 1\}$ незалежних однаково розподілених випадкових величин, функція розподілу яких є неперервною. Тоді події типу $\{X_i = X_j\}$ мають ймовірність 0, якщо $i \neq j$. Нехай $L(1) = 1$. Для $n \geq 2$ означимо рекурентно випадкові величини

$$L(n) = \inf\{k > L(n-1) : X_k > X_{L(n-1)}\} \tag{1}$$

вважаючи, що $\inf \emptyset := +\infty$. Члени послідовності $L = \{L(n), n \geq 1\}$ називають *моментами рекордів*, побудованими за $\{X_k, k \geq 1\}$. Ми також розглядаємо послідовність випадкових величин $\mu = \{\mu(n), n \geq 1\}$, означену співвідношенням

$$\mu(n) = \#\{k : L(k) \leq n\}, \qquad n \geq 1. \tag{2}$$

Зрозуміло, що $\mu(n)$ – це *кількість рекордів*, що трапились до моменту n включно.

В роботі [10] вперше було розглянуто так звану *F^α-схему*, яка будується за заданої функції розподілу та послідовності додатних чисел $\{\alpha_k\}$. Зрозуміло, що $F^{\alpha_n}(x)$ є функцією розподілу для кожного $n \geq 1$. Сукупність незалежних величин $\{X_n\}$ називають F^α-схемою, якщо функцією розподілу випадкової величини X_n є $F^{\alpha_n}(x)$. Якщо всі α_n є рівними між собою, то F^α-схема – це сукупність незалежних однаково розподілених випадкових величин. Якщо ж не всі α_n є рівними між собою, то F^α-схема – це узагальнення класичного випадку.

При вивченні F^α-схеми корисними є допоміжні випадкові величини

$$I_k = \begin{cases} 1, & \text{якщо } X_k \text{ є рекордом}, \\ 0, & \text{у іншому випадку}. \end{cases}$$

У роботі [7] доведено, що $\{I_k\}$ є незалежними випадковими величинами (див. також [1]) з такими ймовірностями «*успіху*»

$$\mathbf{P}(I_k = 1) = p_k = \frac{\alpha_k}{A_k} = 1 - \frac{A_{k-1}}{A_k},$$

де $A_0 = 0$, $A_1 = \alpha_1$, $A_k = \alpha_1 + \cdots + \alpha_k$, $k \geq 2$. Оскільки I_k – випадкова величина Бернуллі, то

$$\mathbf{E}I_k = p_k, \qquad \mathbf{D}I_k = p_k(1-p_k) \tag{3}$$

Звідси випливає, що

$$\mathbf{E}\mu(n) = \sum_{k=1}^{n} p_k, \tag{4}$$

$$\mathbf{D}\mu(n) = \sum_{k=1}^{n} p_k(1-p_k). \tag{5}$$

Зауваження 1. Якщо $\mathbf{E}\mu(n)$ – обмежене, то за лемою Бореля–Кантеллі для подій $\{I_k = 1\}$ будемо мати майже напевно скінченну кількість рекордів, і тоді не зможемо говорити про асимптотики. Тому надалі $\mathbf{E}\mu(n) \to \infty$, що буде підтверджуватись із контексту.

Асимптотика $\mathbf{E}\mu(n)$ та $\mathbf{D}\mu(n)$

Теорема 1. *Нехай $A_n \to \infty, p_n \to 0, n \to \infty$. Тоді:*

$$\lim_{n\to\infty} \frac{\boldsymbol{E}\mu(n)}{\ln A_n} = 1 \qquad \lim_{n\to\infty} \frac{\boldsymbol{D}\mu(n)}{\ln A_n} = 1$$

Доведення. Представимо $\mathbf{E}(\mu(n) - \mu(s))$ у вигляді наступної суми:

$$\mathbf{E}(\mu(n) - \mu(s)) = \sum_{k=s+1}^{n} p_k = \sum_{k=s+1}^{n} \lambda_k g(\lambda_k),$$

де $g(x) = \frac{1-e^{-x}}{x}, \lambda_k = \ln\left(\frac{1}{1-p_k}\right)$. Запишемо наступну послідовність нерівностей:

$$\mathbf{E}(\mu(n) - \mu(s)) \leq \sup_{s<k\leq n} (g(\lambda_k)) \sum_{k=s+1}^{n} \lambda_k \leq$$

$$\leq \sup_{s<k}(g(\lambda_k)) \sum_{k=s+1}^{n} \lambda_k = \sup_{s<k}(g(\lambda_k)) \left(\ln A_n - \ln A_s\right),$$

де остання рівність випливає з $\lambda_k = \ln\left(\frac{A_k}{A_{k-1}}\right)$. Маємо:

$$\frac{\mathbf{E}\mu(n) - \mathbf{E}\mu(s)}{\ln A_n - \ln A_s} \leq \sup_{s<k} g(\lambda_k) \tag{6}$$

Аналогічно, отримуємо оцінку знизу:

$$\inf_{s<k} g(\lambda_k) \leq \frac{\mathbf{E}\mu(n) - \mathbf{E}\mu(s)}{\ln A_n - \ln A_s} \tag{7}$$

Взявши верхню та нижню границі за n, враховуючи $\ln A_n \to \infty$, з (6) та (7) маємо:

$$\inf_{s<k} g(\lambda_k) \leq \liminf_{n\to\infty} \frac{\mathbf{E}\mu(n)}{\ln A_n} \leq$$

$$\leq \limsup_{n\to\infty} \frac{\mathbf{E}\mu(n)}{\ln A_n} \leq \sup_{s<k} g(\lambda_k)$$

Взявши границю за s будемо мати:

$$\liminf_{k\to\infty} g(\lambda_k) \leq \liminf_{n\to\infty} \frac{\mathbf{E}\mu(n)}{\ln A_n} \leq$$

$$\leq \limsup_{n\to\infty} \frac{\mathbf{E}\mu(n)}{\ln A_n} \leq \limsup_{k\to\infty} g(\lambda_k)$$

При $p_k \to 0$ також $\lambda_k \to 0$. Існує границя

$$\lim_{x\to 0} g(x) = \lim_{x\to 0} \frac{1-e^{-x}}{x} = 1$$

У підсумку, маємо рівність нижньої та верхньої границь для $g(\lambda_k)$ і наступну еквівалентність:

$$\lim_{n\to\infty} \frac{\mathbf{E}\mu(n)}{\ln A_n} = 1$$

Ті ж самі міркування можна провести для $\mathbf{D}\mu(n)$, лише $\widetilde{g}(x) = \frac{1-e^{-x}}{x}e^{-x}$. Будемо мати:

$$\lim_{n\to\infty} \frac{\mathbf{D}\mu(n)}{\ln A_n} = 1$$

Доведення завершено.

ЦГТ для F^{α}-схеми

Теорема 2. *Нехай $\boldsymbol{D}\mu(n) \to \infty, n \to \infty$. Тоді:*

$$\frac{\mu(n) - \boldsymbol{E}\mu(n)}{\sqrt{\boldsymbol{D}\mu(n)}} \Rightarrow N(0,1)$$

Доведення. З'ясуємо виконання ЦГТ для $\mu(n)$, перевіривши достатні умови у формі Ляпунова:

$$\exists\delta > 0 : \lim_{n\to\infty} \frac{\sum\limits_{k=1}^{n} \mathbf{E}\left|I_k - p_k\right|^{2+\delta}}{\left(\sqrt{\mathbf{D}\mu(n)}\right)^{2+\delta}} = 0$$

Розглянемо $\delta = 1$.

$$\frac{\sum\limits_{k=1}^{n} \mathbf{E}\left|I_k - p_k\right|^3}{\left(\mathbf{D}\mu(n)\right)^{3/2}} = \frac{\sum\limits_{k=1}^{n} p_k^3(1-p_k) + (1-p_k)^3 p_k}{\left(\mathbf{D}\mu(n)\right)^{3/2}}$$

$$= \frac{\sum\limits_{k=1}^{n} p_k(1-p_k)(p_k^2 + (1-p_k)^2)}{\left(\mathbf{D}\mu(n)\right)^{3/2}} \leq \frac{\sum\limits_{k=1}^{n} p_k(1-p_k)}{\left(\mathbf{D}\mu(n)\right)^{3/2}}$$

$$= \frac{1}{\sqrt{\mathbf{D}\mu(n)}} \to 0, \text{ оскільки } \mathbf{D}\mu(n) \to \infty, n \to \infty.$$

Доведення завершено.

Далі нас буде цікавити випадок, коли $p_n \to 0$. Саме в ньому спостерігаються цікаві асимптотики. При цьому ми також вважаємо, що $A_n \to \infty$, інакше будемо мати лише скінченну кількість рекордів майже напевно. Подальшою метою буде заміна математичного сподівання та дисперсії на їх асимптотики, та явний підрахунок в окремому випадку.

Розрахунок для степеневого зростання

Розглянемо випадок, коли $\alpha_k = k^s, s > -1$.

Теорема 3.

$$\exists C_s > 0 : |(s+1)\ln n - \boldsymbol{E}\mu(n)| < C_s, \quad n \geq 1$$

Доведення. Спершу оцінимо $A_n = 1^s + \cdots + n^s$ для $s > 0$. Порівняємо зі значенням інтеграла від x^s. Розбивши інтеграл на одиничні інтервали, оцінивши на них мінімальним та максимальним значенням функцію, маємо стандартну оцінку для інтеграла:

$$\sum_{k=1}^{n-1} k^s \leq \int_1^n x^s dx \leq \sum_{k=2}^{n} k^s$$

$$A_n - n^s \leq \frac{n^{s+1}-1}{s+1} \leq A_n - 1$$

Звідки, перетворивши, маємо:

$$\frac{n^{s+1}-1}{s+1} + 1 \leq A_n \leq \frac{n^{s+1}-1}{s+1} + n^s$$

Отримаємо оцінки на $p_k = \frac{\alpha_k}{A_k}$.

$$\frac{(s+1)k^s}{k^{s+1}-1+(s+1)k^s} \leq \frac{\alpha_k}{A_k} \leq \frac{(s+1)k^s}{k^{s+1}-1+s+1}$$

Звідси легко показати, що

$$\frac{s+1}{k+\lceil s\rceil+1} \leq \frac{\alpha_k}{A_k} \leq \frac{s+1}{k}$$

Якщо просумувати за k від 1 до n дані нерівності, то отримуємо зліва та справа часткові суми гармонічного ряду (зліва сума не з одиниці).
І можемо застосувати формулу Ойлера для суми гармонічного ряду. Всі константи внесемо в $O(1)$.

$$(s+1)\ln(n+\lceil s\rceil+1)+O(1) \leq \mathbf{E}\mu(n) \leq$$
$$\leq (s+1)\ln n + O(1)$$

$$(s+1)\ln\frac{n+\lceil s\rceil+1}{n}+O(1) \leq$$
$$\leq \mathbf{E}\mu(n)-(s+1)\ln n \leq O(1)$$

Оскільки $\ln\frac{n+\lceil s\rceil+1}{n}$ прямує до нуля, то він обмежений. Тоді з вказаних нерівностей випливає обмеженість різниці $\mathbf{E}\mu(n)-(s+1)\ln n$.

Випадок $s \in (-1;0)$ розглядається абсолютно аналогічно, лише з тою різницею, що x^s строго спадає, що переставляє верхню та нижню оцінки, але не змінює характеру асимптотики.

Випадок $s=0$ перетворює $\mathbf{E}\mu(n)$ на гармонічний ряд, що очевидно зберігає обмеженість різниці з логарифмом.
Доведення завершено.

Теорема 4. *Нехай* $\alpha_k = k^s, k \geq 1, s > -1$. *Тоді:*

$$\frac{\mu(n)-(s+1)\ln n}{\sqrt{(s+1)\ln n}} \Rightarrow N(0,1)$$

Доведення. Із оцінок для A_n з теореми 3 випливає, що $A_n \sim \frac{n^{s+1}}{s+1}$. Тому $\ln A_n \sim (s+1)\ln n$.
Також $p_n \to 0$, оскільки $\frac{\alpha_n}{A_n} \sim \frac{s+1}{n}$.
З умови $s > -1$ випливає, що $A_n \to \infty$. Тоді за теоремою 1 маємо ланцюг еквівалентностей:

$$(s+1)\ln n \sim \ln A_n \sim \mathbf{E}\mu(n) \sim \mathbf{D}\mu(n) \quad (8)$$

Оскільки $\mathbf{D}\mu(n) \to \infty$ можемо записати результат теореми 2:

$$\frac{\mu(n)-\mathbf{E}\mu(n)}{\sqrt{\mathbf{D}\mu(n)}} \Rightarrow \eta \sim N(0,1) \quad (9)$$

Через еквівалентності (8):

$$\sqrt{\frac{\mathbf{D}\mu(n)}{(s+1)\ln n}} \to 1 \quad (10)$$

Тому за теоремою Слуцького для добутку (9) та (10) маємо:

$$\frac{\mu(n)-\mathbf{E}\mu(n)}{\sqrt{\mathbf{D}\mu(n)}}\sqrt{\frac{\mathbf{D}\mu(n)}{(s+1)\ln n}} \Rightarrow \eta\cdot 1 \quad (11)$$

Тобто:

$$\frac{\mu(n)-\mathbf{E}\mu(n)}{\sqrt{(s+1)\ln n}} \Rightarrow \eta \sim N(0,1) \quad (12)$$

Із теореми 3 випливає:

$$\frac{\mathbf{E}\mu(n)-(s+1)\ln n}{\sqrt{(s+1)\ln n}} \to 0 \quad (13)$$

Тому за теоремою Слуцького для суми (12) та (13) маємо:

$$\frac{\mu(n)-(s+1)\ln n}{\sqrt{(s+1)\ln n}} \Rightarrow \eta \sim N(0,1)$$

Доведення завершено.

Список літератури

1. Borovkov K., Pfeifer D. On record indices and record times. *Journal of Statistical Planning and Inference.* 1995. Vol. 45, No. 1–2,. Pp. 65–79.
2. Buldygin V. V., Indlekofer K.-H., Klesov O. I., Steinebach J. Pseudo-Regularly Varying Functions and Generalized Renewal Processes. Springer Verlag, 2018.
3. Doukhan P., Klesov O. I., Pakes A., Steinebach J. G. Limit theorems for record counts and times in the F?-scheme. *Extremes.* 2013. Vol. 16, No. 2. Pp. 147–171.
4. Doukhan O. I., Klesov J. G., Steinebach J. G. Strong laws of large numbers in an F^α-scheme. *Mathematical Statistics and Limit Theorems. Festschrift in Honour of Paul Deheuvels.* Springer International Publishing, 2015. Pp. 287–303.
5. Gut A. Probability: A Graduate Course. Springer-Verlag, 2005.
6. Klesov O. I. Limit Theorems for Multi-Indexed Sums of Random Variables. Springer-Verlag, 2014.
7. Nevzorov V. B. On record times and inter-record times for sequences of nonidentically distributed random variables.*Zap. Nauchn. Sem. LOMI.* 1985. Vol. 142. Pp. 109–118.
8. Nevzorov V. B. Records: Mathematical Theory. Providence, RI : American Mathematical Society, 2001.
9. R?nyi A. Th?orie des ?l?ments saillants d'une suite d'observations. *Combinatorial Methods in Probability Theory.* Math. Inst., Aarhus Univ., Aarh, 1962. Pp. 10–117.
10. Yang M. On the distribution of the inter-record times in an increasing population. *J. Appl. Prob.* 1975. Vol. 12. Pp. 148–154.

References

1. K. Borovkov and D. Pfeifer, "On record indices and record times," Journal of Statistical Planning and Inference. **45** (1–2), 65–79 (1995).
2. V. V. Buldygin, K.-H. Indlekofer, O. I. Klesov, and J. Steinebach, *Pseudo-Regularly Varying Functions and Generalized Renewal Processes* (Berlin: Springer Verlag, 2018).
3. P. Doukhan, O. I. Klesov, A. Pakes, and J. G. Steinebach, "Limit theorems for record counts and times in the F?-scheme," Extremes. **16** (2), 147–171 (2013).
4. P. Doukhan, O. I. Klesov, and J. G. Steinebach, Strong laws of large numbers in an F^α-scheme, in: *Mathematical Statistics and Limit Theorems. Festschrift in Honour of Paul Deheuvels* (Springer International Publishing, Switzerland, 2015), pp. 287–303.
5. A. Gut,*Probability: A Graduate Course* (Berlin: Springer Verlag, 2005).
6. O. I. Klesov, *Limit Theorems for Multi-Indexed Sums of Random Variables* (Berlin: Springer Verlag, 2014).
7. V. B. Nevzorov, "On record times and inter-record times for sequences of nonidentically distributed random variables," *Zap. Nauchn. Sem. LOMI.* **142**, 109–118 (1985).
8. V. B. Nevzorov, *Records: Mathematical Theory* (Providence, RI: American Mathematical Society, 2001).
9. A. R?nyi, Th?orie des ?l?ments saillants d'une suite d'observations, in: Combinatorial Methods in Probability Theory (Math. Inst., Aarhus Univ., Aarhus, Denmark, 1962), pp. 10–117.
10. M. Yang, "On the distribution of the inter-record times in an increasing population," J. Appl. Prob. **12**, 148–154 (1975).

O. Kolesnik

CENTRAL LIMIT THEOREM FOR THE NUMBER OF RECORDS IN F^α-SCHEME

Consider the sequence $\{X_k, k \geq 1\}$ *of independent identically distributed random variables whose distribution function is continuous. Then events of the type* $\{X_i = X_j\}$ *have probability* 0 *if* $i \neq j$*. Let* $L(1) = 1$*. For* $n \geq 2$*, we define random variables*

$$L(n) = \inf\{k > L(n-1) : X_k > X_{L(n-1)}\}$$

assuming that $\inf \emptyset := +\infty$*. The members of the sequence* $L = \{L(n), n \geq 1\}$ *are called moments of records constructed for* $\{X_k, k \geq 1\}$*. Consider the sequence of random variables* $\mu = \{\mu(n), n \geq 1\}$*, defined by the relation*

$$\mu(n) = \#\{k : L(k) \leq n\}, \qquad n \geq 1.$$

It is clear that $\mu(n)$ *– is the number of records that happened up to the moment* n *inclusive.*

In the work [10], the so-called F^α*-scheme is considered for the first time, which is built using a given distribution function and a sequence of positive numbers* $\{\alpha_k\}$*. It is clear that* $F^{\alpha_n}(x)$ *is the distribution function for each* $n \geq 1$*. The set of independent random variables* $\{X_n\}$ *is called the* F^α *scheme, if the distribution function of the random variable* X_n *is* $F^{\alpha_n}(x)$*. If all* α_n *are equal to each other, then the* F^α *scheme – is a set independent identically distributed random variables. If not all* α_n *are equal to each other, then the* F^α *scheme – is a generalization of the classical case.*

This paper examines the assertions related to the fulfillment of the central limit theorem (CLT) for the number of records in the F^α*-scheme of records. The method of finding exact asymptotic expressions for mathematical expectation and variance, which can be used to replace the real characteristics in CLT, is given.*

A specific example of power-law growth of exponents of the F^α*-scheme was considered, and CLT is constructed only in terms of the moment of observation and the power of growth.*

The article contains 4 theorems with complete proof. Theorem 1 relates the mathematical expectation and variance to the accumulated intensity of the F^α*-scheme. Theorem 2 establishes the implementation of CLT in general, and theorem 4 – for a specific case.*

Keywords: independent identically distributed randoom variables, F^α-scheme, records, number of records, central limit theorem.

Матеріал надійшов 27.12.2023