Г.О. Михалін, О.В. Стогній
НПУ імені М.П. Драгоманова

Статистичні ймовірності. Прості випадкові величини. Закон великих чисел.

1. Вступ. У роботах [1], [2] та інших обґрунтовано думку про те, що вивчення елементів стохастики у середній школі доцільно здійснювати не на основі «класичного означення» ймовірності (яке насправді не є означенням, а є тавтологією або зачарованим колом), а на основі поняття «відносної частоти» або «статистичної ймовірності», яке вводиться наступним чином.

Нехай подія $A \subset \Omega$ пов'язана з певним *стохастичним експериментом*, множиною можливих наслідків якого є Ω – *простір елементарних подій* [1, с.6]. Якщо цей експеримент провести n – разів, тобто провести n *випробувань*, і в них подія A відбудеться m разів, то число $P_n^*(A) = \frac{m}{n}$ називають *відносною частотою або статистичною ймовірністю* події A [1, с.19].

При цьому вважають, що сукупність подій утворює так званий *простір подій* S, що визначається умовами [1, с.23]:

1_S. $\Omega \in S$, *тобто вірогідна подія завжди є подією*;

2_S. *Якщо* $A \in S$, *то* $\overline{A} \in S$, *тобто* A *і* $\overline{A}$ *одночасно або є подіями, або ні*;

3_S. *Якщо* $A_i \in S,\ i \in N$, *то* $\sum\limits_i A_i \in S$, *тобто сума скінченної або зчисленної кількості подій завжди є подією*.

Зауважимо, що аналогічні умови задовольняє сукупність фігур, що є частинами, наприклад, одиничного квадрата (куба) і мають площу (об'єм).

З означення статистичної ймовірності випливає, що вона задовольняє властивості, аналогічні властивостям площ та об'ємів [2]:

1_{P^*}. $P_n^*(A) \geq 0 \quad \forall A \in S$ (*невід'ємність*);

2_{P^*}. $P_n^*\left(\sum\limits_i A_i\right) = \sum\limits_i P_n^*(A)$, *коли події* A_i *попарно несумісні, а їх кількість скінченна або зчисленна* (*повна адитивність*);

3_{P^*}. $P_n^*(\Omega) = 1$ (*нормованість*).

Зрозуміло, що $P_n^*(A)$ залежить від n, проте виявляється, що *для досить великих* n *числа* $P_n^*(A)$ *можуть бути у певному розумінні близькими до деякого числа* $P(A)$, яке вже не залежить від n, причому числа $P(A)$, $A \in S$, також задовольняю властивості, аналогічні властивостям площ та об'ємів:

1_P. $P(A) \geq 0\ \forall A \in S$ (*невід'ємність*);

2_P. $P\left(\sum\limits_i A_i\right) = \sum\limits_i P(A)$, *коли події* A_i *попарно несумісні, а їх кількість скінченна або зчисленна* (*повна адитивність*);

3_P. $P(\Omega) = 1$ (*нормованість*).

Числа $P(A)$, $A \in S$ називають *ймовірностями відповідних подій* $A \in S$, а сукупність (Ω, S, P) називають *імовірнісним простором*, що відповідає даному стохастичному експерименту [1, с. 24].

Отже, у зв'язку з наведеними фактами виникає питання (проблема, задача) про те**, як слід пояснювати слова**: *«відносна частота* $P_n^*(A)$ *у певному розумінні є близькою до ймовірності* $P(A)$, *коли кількість випробувань* n *досить велика».*

2. Прості випадкові величини. Відносна частота $P_n^*(A)$ є прикладом так званої випадкової величини [3,с.51,59], множина значень якої скінченна. Виявляється, що за допомогою таких величин можна ефективно розв'язати поставлену проблему.

Тому визначимо спочатку такі величини і дослідимо їхні властивості.

Нехай задано імовірнісний простір (Ω, S, P_n^*).

Простою випадковою величиною називається будь-яка функція $X = X(E)$, $E \in D(X)$, для якої:

1) областю визначення $D(X)$ є простір Ω елементарних подій;
2) множина значень $X(\Omega)$ складається із скінченної кількості попарно різних чисел $x_i \in R$, $i \in \overline{1,k}$ (k – фіксоване число), тобто $X(\Omega) = \{x_i : i \in \overline{1,k}\}$;
3) сукупність елементарних подій, в яких випадкова величина X набуває значень x_i є подією, тобто $A_i := \{E \in \Omega : X(E) = x_i\} = X^{-1}(x_i) \in S$, $i \in \overline{1,k}$.

Приклад 1. Якщо функція $X(E) = c = const$, $E \in \Omega$, то її називають *сталою випадковою величиною*. Множина значень сталої випадкової величини складається з одного елемента c, тобто $X(E) = \{c\}$ для всіх $E \in \Omega$, або $X^{-1}(c) = \Omega$.

Приклад 2. Якщо A – випадкова подія і

$$X_A(E) = \begin{cases} 1, \text{ коли } E \in A, \\ 0, \text{ коли } E \notin A, \end{cases}$$

то таку випадкову величину називають *індикатором події* A. Множина значень цієї простої випадкової величини складається з двох елементів 1 і 0, тобто $X_A(\Omega) = \{0, 1\}$.

Приклад 3. Якщо X – кількість точок на грані грального кубика, якою кубик падає догори після однократного підкидання, а простір подій S складається з будь-яких підмножин простору Ω, то X – проста випадкова величина з можливими значеннями 1, 2, 3, 4, 5, 6, тобто $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$.

Приклад 4. Якщо у прикладі 3 простір подій S не є найширшим, то X вже не є простою випадковою величиною, оскільки для деякого числа $i \in \overline{1,6}$ його прообраз $X^{-1}(i)$ не є подією.

3. На практиці часто є корисним **критерій простої випадкової величини**: *для того, щоб функція* $X = X(E)$, $E \in \Omega$ *була простою випадковою величиною, необхідно і достатньо, щоб існувала скінченна кількість попарно несумісних подій* $A_k \subset \Omega$, $k \in \overline{1,s}$, *для яких* $\sum_{k=1}^{s} A_k = \Omega$ *і* $X(E) = c_k$, $E \in A_k$, $k \in \overline{1,s}$, *але числа* c_k *не обов'язково попарно різні. При цьому сукупність подій* A_k, $k \in \overline{1,s}$ *називають скінченним розкладом простору* Ω *елементарних подій.*

Дійсно, якщо $X = X(E)$, $E \in \Omega$, – проста випадкова величина, то $X(\Omega) = \{x_1, x_2, ..., x_m\}$, де числа x_k попарно різні. Згідно з означенням випадкової величини множини $X^{-1}(x_k) = A_k$ є подією для кожного $k \in \overline{1,m}$. Ці події попарно несумісні, оскільки x_k попарно різні. Зрозуміло, що коли $E \in A_k$, то $X(E) = x_k$, а оскільки $X(\Omega) = \{x_1, x_2, ..., x_m\}$, то $\sum_{k=1}^{m} A_k = \Omega$. Отже, події A_k утворюють скінченний розклад простору Ω, тому необхідність доведено.

Доведемо достатність. Нехай для функції $X = X(E)$, $E \in \Omega$, існує скінченна кількість попарно несумісних подій A_k, $k \in \overline{1,s}$, для яких $X(E) = c_k$, коли $E \in A_k$, і $\sum_{k=1}^{s} A_k = \Omega$, причому числа c_k не обов'язково попарно різні.

Розподілимо усі числа c_k на m груп, в кожну з яких входять рівні числа c_k, а в різні групи входять обов'язково різні числа:

$c_{k_1^{(1)}} = c_{k_2^{(1)}} = ... = c_{k_{i_1}^{(1)}} = x_1$ – числа першої групи;

$c_{k_1^{(2)}} = c_{k_2^{(2)}} = ... = c_{k_{i_2}^{(2)}} = x_2$ – числа другої групи;

……………………………………………………….

$c_{k_1^{(m)}} = c_{k_2^{(m)}} = ... = c_{k_{i_m}^{(m)}} = x_m$ – числа m-ої групи.

Утворимо події $A_j^* = \sum_{i=1}^{i_j} A_{k_i^{(j)}}$, $j \in \overline{1,m}$, і дістанемо, що $X(\Omega) = \{x_1, x_2, ..., x_m\}$, де числа x_j попарно різні, $X(E) = x_j$, коли $E \in A_j^*$, а тому $X^{-1}(x_j) = A_j^*$ є подією для кожного $j \in \overline{1,m}$. Це означає, що функція $X = X(E)$, $E \in \Omega$ є простою випадковою величиною.

4. З доведеного критерія простої випадкової величини легко дістати **твердження про замкненість сукупності простих випадкових величин відносно арифметичних операцій**: *Якщо* X_1 *і* X_2 *– прості випадкові величини, то їх сума* $X_1(E) + X_2(E)$, *різниця* $X_1(E) - X_2(E)$, *добуток* $X_1(E) \cdot X_2(E)$ *і частка* $\dfrac{X_1(E)}{X_2(E)}$, $(X_2(E) \neq 0)$, *є простими випадковими величинами.*

Дійсно, згідно з критерієм простої випадкової величини, існують скінченні розклади $A_k^{(1)}$, $k \in \overline{1,s_1}$ і $A_k^{(2)}$, $k \in \overline{1,s_2}$ простору Ω, для яких $X_1(E) = c_k^{(1)}$, $E \in A_k^{(1)}$, $k \in \overline{1,s_1}$ і $X_2(E) = c_k^{(2)}$, $E \in A_k^{(2)}$, $k \in \overline{1,s_2}$.

Розглянемо події $B_{i,j} = A_i^{(1)} A_j^{(2)}$, $i \in \overline{1,s_1}$, $j \in \overline{1,s_2}$, що утворюють скінченний розклад простору Ω, спільний для випадкових величин X_1 та X_2. При цьому $X_1(E) = c_i^{(1)}$, а $X_2(E) = c_j^{(2)}$, коли

$E \in B_{i,j} = A_i^{(1)} \cdot A_j^{(2)}$, $i \in \overline{1, s_1}$, $j \in \overline{1, s_2}$. Тому $X_1(E) \pm X_2(E) = c_i^{(1)} \pm c_j^{(2)}$, $X_1(E) \cdot X_2(E) = c_i^{(1)} c_j^{(2)}$ і $\dfrac{X_1(E)}{X_2(E)} = \dfrac{c_i^{(1)}}{c_j^{(2)}}$, коли $E \in B_{i,j} = A_i^{(1)} \cdot A_j^{(2)}$, $i \in \overline{1, s_1}$, $j \in \overline{1, s_2}$. Звідси, в силу критерія простої випадкової величини дістанемо, що $X_1 \pm X_2$, $X_1 \cdot X_2$ та $\dfrac{X_1}{X_2}$ (коли $X_2(E) \neq 0$, $E \in \Omega$), є простими випадковими величинами.

Методом математичної індукції легко довести, що сума і добуток довільної скінченної кількості простих випадкових величин $X_k(E)$, $E \in \Omega$, $k \in \overline{1, m}$ також є простою випадковою величиною.

Узагальненням цього є твердження про те, що *коли* $X_k(E)$ *– прості випадкові величини для* $k \in \overline{1, m}$, *то будь-яка функція* $Y = f(X_1(E), X_2(E), ..., X_m(E))$, $E \in \Omega$, *також є простою випадковою величиною за умови, що* $(X_1(E), ..., X_m(E))$ *належить області визначення* $D(f)$ *функції* f, $E \in \Omega$. *Зокрема, якщо* $X(E)$ *– проста випадкова величина, то* $Y(E) = X^n(E)$, $Y(E) = \sqrt[n]{X(E)}$, $Y(E) = e^{X(E)}$, $Y(E) = \ln X(E)$, $Y(E) = \sin X(E)$, $Y(E) = \cos X(E)$ *тощо також є простими випадковими величинами.*

Приклад 5. Нехай X_1 – індикатор події $A = [0; 0{,}5) \subset \Omega = [0;1)$, а X_1 – індикатор події $B = [0{,}25; 1)$. Тоді події $A_1^{(1)} = [0; 0{,}5)$ і $A_2^{(1)} = \overline{A_1^{(1)}} = [0{,}5; 1)$ та $A_1^{(2)} = (0; 0{,}25)$ і $A_2^{(2)} = \overline{A_1^{(2)}} = [0{,}25; 1)$ утворюють скінченні розклади простору Ω, що відповідають випадковим величинам відповідно X_1 та X_2:

$$X_1(E) = 1, \text{ коли } E \in A_1^{(1)} \text{ і } X_1(E) = 0, \text{ коли } E \in A_2^{(1)} = \overline{A_1^{(1)}},$$
$$X_2(E) = 0, \text{ коли} E \in A_1^{(2)} \text{ і } X_2(E) = 1, \text{ коли} E \in A_2^{(2)} = \overline{A_1^{(2)}}.$$

Спільний розклад утворюють події:

$$B_{11} = A_1^{(1)} \cdot A_1^{(2)} = [0;\ 0{,}5) \cap [0; 0{,}25) = [0; 0{,}25),$$
$$B_{12} = A_1^{(1)} \cdot A_2^{(2)} = [0; 0{,}5) \cap [0{,}25; 1) = [0{,}25; 0{,}5),$$
$$B_{21} = A_2^{(1)} \cdot A_1^{(2)} = [0{,}5; 1) \cap [0; 0{,}25) = \varnothing \text{ і}$$
$$B_{22} = A_2^{(1)} A_2^{(2)} = [0{,}5; 1) \cap [0{,}25; 1) = [0{,}5; 1).$$

Тому $X_1(E) + X_2(E) = 1 + 0 = 1$, коли $E \in [0; 0{,}25)$,

$X_1(E) + X_2(E) = 1 + 1 = 2$, коли $E \in [0{,}25; 0{,}5)$, і

$X_1(E) + X_2(E) = 0 + 1 = 1$, коли $E \in [0{,}5;\ 1)$.

Отже,

$$X_1(E) + X_2(E) = \begin{cases} 1, & \text{коли } E \in [0; 0{,}25) \cap [0{,}5; 1), \\ 2, & \text{коли } E \in [0{,}25; 0{,}5). \end{cases}$$

Аналогічно знаходимо $X_1(E) \cdot X_2(E)$ та $\dfrac{X_1(E)}{X_2(E)}$.

Надалі, розглядаючи прості випадкові величини $X = X_k(E)$, $E \in \Omega$, $k \in \overline{1, m}$, вважатимемо, що всім їм відповідає спільний розклад простору Ω, тобто $\Omega = \sum_{i=1}^{s} A_i$, події A_i попарно несумісні і кожна випадкова величина X_k є сталою на множині A_i: $X_k(E) = c_i^{(k)}$, $E \in A_i$, $i \in \overline{1, s}$, $k \in \overline{1, m}$.

5. *Прості випадкові величини* X_1 та X_2 називають *статистично незалежними*, коли для будь-яких чисел $a \in X_1(\Omega)$ та $b \in X_2(\Omega)$ події $X_1^{-1}(a)$ та $X_2^{-1}(b)$ є незалежними відносно міри P_n^* з відповідного ймовірнісного простору (Ω, S, P_n^*), тобто $P_n^*(X_1^{-1}(a) \cdot X_2^{-1}(b)) = P_n^*(X_1^{-1}(a)) \cdot P_n^*(X_2^{-1}(b))$.

В іншому разі X_1 та X_2 називають *статистично залежними випадковими величинами.*

Приклад 6.1. Нехай X_1 та X_2 з прикладу 5. Тоді $X_1^{-1}(1) = [0; 0{,}5)$, $X_1^{-1}(0) = [0{,}5; 1)$, $X_2^{-1}(1) = [0{,}25; 1)$ і $X_2^{-1}(0) = [0; 0{,}25)$. Тому $X_1^{-1}(1) \cdot X_2^{-1}(1) = [0{,}25; 0{,}5)$, $X_1^{-1}(1) \cdot X_2^{-1}(0) = [0; 0{,}25)$, $X_1^{-1}(0) \cdot X_2^{-1}(1) = [0{,}5; 1)$ і $X_1^{-1}(0) \cdot X_2^{-1}(0) = \varnothing$.

Отже, якщо $P_n^*([a, b)) = b - a$, $[a, b) \subset [0,1)$, то

$$P_n^*(X_1^{-1}(1) \cdot X_2^{-1}(1)) = 0{,}5 - 0{,}25 \neq P_n^*(X_1^{-1}(1)) \cdot P_n^*(X_2^{-1}(1)) = \\ = 0{,}5 \cdot 0{,}75 = 0{,}375.$$

Це означає, що випадкові величини X_1 і X_2 залежні відносно такої міри P_n^*.

2. Якщо ж P_n^* визначена так, що $P_n^*([0{,}25;0{,}5))=1$, то $P_n^*([0;0{,}5))=P_n^*([0{,}25;1))=1$ і $P_n^*([0;0{,}25))=P_n^*([0{,}5;1))=0$, а тому

$$P_n^*(X_1^{-1}(1)\cdot X_2^{-1}(1))=1=P_n^*(X_1^{-1}(1))\cdot P_n^*(X_2^{-1}(1))=1\cdot 1,$$
$$P_n^*(X_1^{-1}(1)\cdot X_2^{-1}(0))=0=P_n^*(X_1^{-1}(1))\cdot P_n^*(X_2^{-1}(0))=1\cdot 0,$$
$$P_n^*(X_1^{-1}(0)\cdot X_2^{-1}(1))=0=P_n^*(X_1^{-1}(0))\cdot P_n^*(X_2^{-1}(1))=0\cdot 1,$$
$$P_n^*(X_1^{-1}(0)\cdot X_2^{-1}(0))=0=P_n^*(X_1^{-1}(0))\cdot P_n^*(X_2^{-1}(0))=0\cdot 0.$$

Це означає, що X_1 та X_2 є незалежними випадковими величини відносно останньої міри P_n^*.

Таким чином, залежність чи незалежність випадкових величин суттєва визначається ймовірнісною мірою P_n^*.

Випадкові величини $X_1, X_2, ..., X_m$ називаються попарно незалежними, якщо будь-які дві з них є незалежними.

Приклад 7. Якщо X_1 і X_2 та P_n^* з прикладу 6.2, а $X_3(E)=1$, коли $E\in\Omega$, то легко бачити, що X_1, X_2 та X_3– попарно незалежні прості випадкові величини, оскільки стала випадкова величина з будь-якою випадковою величиною утворюють пару незалежних величин.

Якщо $Y=X+c$, то $Y^{-1}(a)=X^{-1}(a-c)\ \ \forall a\in R$.

Тому *прості випадкові величини* X_1 *та* X_2 *є незалежними тоді тільки тоді, коли незалежні величини* $Y_1=X_1+C_1$ *та* $Y_2=X_2+C_2$.

6. Числові характеристики простих випадкових величин. Нехай задано імовірнісний простір (Ω, S, P_n^*) та просту випадкову величину $X=X(E)$, $E\in\Omega$, множина значень якої $X(\Omega)=\{x_1, x_2, \ldots, x_m\}$. Тоді *статистичне математичне сподівання* $M_n{}^*(X)$ та *статистична дисперсія* $D_n{}^*(X)$ цієї випадкової величини визначається відповідно рівностями

$$M_n^*(X)=\sum_{k=1}^{m} x_k P_n^*(X^{-1}(x_k)) \qquad (1)$$

та

$$D_n^*(X)=M_n^*((X-M_n^*(X))^2)=\sum_{k=1}^{m}(x_k-M_n^*(X))^2 P_n^*(X^{-1}(x_k)) \qquad (2)$$

Зауваження. Суми (1) та (2) називають *інтегралами Лебега* простих випадкових величин $X(E)$ та $(X(E)-M_n^*(X))^2$ і позначають $(L)\int\limits_{\Omega} XdP$ та $(L)\int\limits_{\Omega}(X-M_n^*(X))^2\,dP$.

Приклад 8. 1. Якщо $X=c$ – стала випадкова величина, то $X(\Omega)=\{c\}$, $X^{-1}(c)=\Omega$, $P_n^*(X^{-1}(c))=P_n^*(\Omega)=1$ і тому $M_n^*(c)=c\cdot 1=c$, $D_n^*(c)=M_n^*((c-c)^2)=M_n^*(0)=0$.

2. Якщо $X=X_A(E)$, $E\in\Omega$ – індикатор події A, то $X(\Omega)=\{1{,}0\}$, $X^{-1}(1)=A$, $X^{-1}(0)=\overline{A}$, $P_n^*(X^{-1}(1))=P_n^*(A)=p$, а $P_n^*(X^{-1}(0))=P_n^*(\overline{A})=1-p$. Тому $M_n^*(X_A)=1\cdot p+0\cdot(1-p)=p$, $D_n^*(X_A)=(1-p)^2\cdot p+(0-p)^2\cdot(1-p)=p(1-p)(1-p+p)=p(1-p)$.

Враховуючи, що події $A_k=X^{-1}(x_k)$, $k\in\overline{1,m}$, утворюють скінченний розклад простору Ω, виникає гіпотеза, що замість формул (1) і (2) можна використовувати відповідно формули

$$M_n^*(X)=\sum_{k=1}^{s} c_k P_n^*(A_k) \qquad (1^*)$$

і

$$D_n^*(X)=\sum_{k=1}^{s}(c_k-M^*(X))^2 P_n^*(A_k), \qquad (2^*)$$

де події A_k, $k\in\overline{1,s}$ утворюють довільний скінченний розклад простору Ω і $X(E)=c_k$, коли $E\in A_k$, $k\in\overline{1,s}$. На відміну від чисел x_k числа c_k не обов'язково попарно різні.

Формули (1^*) і (2^*) дійсно правильні, оскільки, як показано при доведенні критерія простої випадкової величини, події A_k визначають події

$$A_j^*=\sum_{i=1}^{i_j} A_{k_i^{(j)}}\ ,\ \ j\in\overline{1,m}$$

на кожній з яких випадкова величина X набуває значення x_j, причому $X(\Omega)=\{x_1,x_2,\ldots,x_m\}$, числа x_j попарно різні і тому $X^{-1}(x_j)=A_j^*$. В силу цього з формули (1*) дістаємо:

$$M_n^*(X)=\sum_{k=1}^{s}c_k P_n^*(A_k)=\sum_{j=1}^{m}\sum_{i=1}^{i_j}c_{k_i^{(j)}}P_n^*(A_{k_i^{(j)}})=\sum_{j=1}^{m}\sum_{i=1}^{i_j}x_j P_n^*(A_{k_i^{(j)}})=$$

$$=\sum_{j=1}^{m}x_j\sum_{i=1}^{i_j}P_n^*(A_{k_i^{(j)}})=\sum_{j=1}^{m}x_j P_n^*(\sum_{i=1}^{i_j}A_{k_i^{(j)}})=\sum_{j=1}^{m}x_j P_n^*(A_{k_i^{(j)}})=$$

$$=\sum_{j=1}^{m}x_j P_n^*(X^{-1}(x_j)).$$

Отже, формули (1) і (1*) рівносильні.

Міркування для формул (2) і (2*) аналогічні.

Зауваження. Суми (1*) та (2*) називають *інтегралами Рімана* простих випадкових величин $X(E)$ та $(X(E)-M_n^*(X))^2$ і позначають $(R)\int_\Omega XdP$ та $(R)\int_\Omega (X-M_n^*(X))^2\,dP$. З доведеного випливає, що для простих випадкових величин їхні інтеграли Рімана і Лебега співпадають.

7. Основні властивості статистичних математичного сподівання і дисперсії:

1. Статистичне математичне сподівання сталої випадкової величини дорівнює цій сталій:

$$M_n{}^*(C)=C.$$

2. Статистична дисперсія сталої випадкової величини дорівнює нулеві:

$$M_n{}^*(C)=0.$$

3. Статистичне математичне сподівання добутку простих незалежних випадкових величин X та Y дорівнює добутку їхніх статистичних математичних сподівань:

$$M_n^*(X\cdot Y)=M_n^*(X)\cdot M_n^*(Y).$$

Зокрема, якщо $Y=c$, то $M_n^*(cX)=cM_n^*(X)$.

Дійсно, якщо $X(\Omega)=\{x_1,x_2,\ldots,x_s\}$ і числа x_i попарно різні, а $Y(\Omega)=\{y_1,y_2,\ldots,y_m\}$ і числа y_j попарно різні, то згідно з формулою (1) а також враховуючи, що випадкові величини X і Y незалежні, одержимо:

$$\begin{aligned}M_n^*(X)\cdot M_n^*(Y)&=\sum_{i=1}^{s}x_iP_n^*(X^{-1}(x_i))\cdot\sum_{j=1}^{m}y_jP_n^*(Y^{-1}(y_j))=\\&=\sum_{i=1}^{s}(\sum_{j=1}^{m}x_i\cdot y_jP_n^*(X^{-1}(x_i)\cdot Y^{-1}(y_j))\end{aligned}\tag{3}$$

Події $B_{ij}=X^{-1}(x_i)\cdot Y^{-1}(y_j)$, $i\in\overline{1,\,s}$, $j\in\overline{1,\,m}$ попарно несумісні і в сумі дають простір Ω. Тому вони утворюють скінченний розклад простору Ω.

Звідси, враховуючи що $X(E)\cdot Y(E)=x_i\cdot y_j$, коли $E\in B_{ij}$, та формулу (1*), дістаємо, що права частина рівності (3) дорівнює $M_n^*(X\cdot Y)$. Властивість 3 доведено.

4. Статистичне математичне сподівання лінійної комбінації простих випадкових величин X та Y дорівнює відповідній лінійній комбінації їхніх математичних сподівань:

$$M_n^*(aX+bY)=aM_n^*(X)+bM_n^*(Y).$$

Дійсно, якщо події A_k, $k\in\overline{1,\,m}$ утворюють скінченний розклад простору Ω, спільний для X та Y, причому $X(E)=x_k$, а $Y(E)=y_k$, коли $E\in A_k$, $k\in\overline{1,\,m}$, то згідно з формулою (1*) маємо:

$$aM_n^*(X)+bM_n^*(Y)=a\sum_{k=1}^{m}x_kP_n^*(A_k)+b\sum_{k=1}^{m}y_kP_n^*(A_k)=$$

$$=\sum_{k=1}^{m}(ax_k+by_k)P_n^*(A_k)=M_n^*(aX+bY).$$

Властивість 4 доведено.

Дана властивість правильна для будь-якої скінченної кількості випадкових величин X_k, $k\in\overline{1,\,m}$:

$$M_n^*(\sum_{k=1}^{m}a_kX_k)=\sum_{k=1}^{m}a_kM_n^*(X_k).$$

5. Якщо $X\ge Y$, то $M_n^*(X)\ge M_n^*(Y)$. Зокрема, якщо $X\ge 0$, то $M_n^*(X)\ge 0$, а тому $D_n^*(X)\ge 0$ для будь-якої простої випадкової величини X.

При цьому вважають, що коли події A_k, $k \in \overline{1, m}$, утворюють скінченний розклад простору Ω, спільний для X і Y, причому $X(E) = x_k$ і $Y(E) = y_k$, коли $E \in A_k$, $k \in \overline{1, m}$, то умова $X \geq Y$ означає, що з нерівності $x_k < y_k$ випливає рівність $P_n^*(A_k) = 0$.

Враховуючи останнє, попередню властивість і формулу (1^*), маємо

$$M_n^*(X) - M_n^*(Y) = M_n(X - Y) = \sum_{k=1}^{m} (x_k - y_k) P_n^*(A_k) \geq 0,$$

оскільки остання сума містить лише невід'ємні доданки. Таким чином

$$M_n(X) - M_n(Y) \geq 0 \Leftrightarrow M_n(X) \geq M_n(Y).$$

Властивість 5 доведено.

6. Якщо X – проста невід'ємна випадкова величина, число $\varepsilon > 0$ і множина $A_\varepsilon = \{E \in \Omega : X(E) \geq \varepsilon\}$, то A_ε є подією, для якої має місце нерівність Чебишова:

$$P_n^*(A_\varepsilon) \leq \frac{1}{\varepsilon} M_n^*(X).$$

Дійсно, якщо $X(\Omega) = \{x_1, x_2, \ldots, x_m\}$, де $0 \leq x_1 < x_2 < \ldots < x_m$, то для фіксованого $\varepsilon > 0$ знайдемо найменший номер n_0, для якого $x_{n_0} \geq \varepsilon$. Тоді нерівність $x_k \geq \varepsilon$ рівносильна нерівності $k \geq n_0$, в силу чого множина $A_\varepsilon = \sum_{n=n_0}^{m} X^{-1}(x_k)$ і тому є подією. Тепер, враховуючи формулу (1), дістаємо

$$M_n^*(X) = \sum_{k=1}^{m} x_k P_n^*(X^{-1}(x_k)) \geq \sum_{k=n_0}^{m} x_k P_n^*(X^{-1}(x_k)) \geq$$

$$\geq \varepsilon \sum_{k=n_0}^{m} P_n^*(X^{-1}(x_k)) = \varepsilon P_n^*(\sum_{k=n_0}^{m} X^{-1}(x_k)) = \varepsilon P_n^*(A_\varepsilon).$$

Отже,

$$M_n^*(X) \geq \varepsilon P_n^*(A_\varepsilon) \Leftrightarrow P_n^*(A_\varepsilon) \leq \frac{1}{\varepsilon} M_n^*(X).$$

Нерівність Чебишова доведено.

7. Статистична дисперсія лінійної комбінації (з коефіцієнтами a і b) незалежних випадкових величин X та Y дорівнює лінійній комбінації їхніх статистичних дисперсій з коефіцієнтами a^2 і b^2:

$$D_n^*(aX + bY) = a^2 D_n^*(X) + b^2 D_n^*(Y).$$

Дійсно,

$$D_n^*(aX + bY) = M_n^*((aX + bY - M_n^*(aX + bY))^2 =$$
$$= M_n^*((a(X - M_n^*(X)) + b(Y - M_n^*(Y)))^2) = M_n^*(a^2(X - M_n^*(X))^2 +$$
$$+ 2ab(X - M_n^*(X))(Y - M_n^*(Y)) + b^2(Y - M_n^*(Y))^2) =$$
$$= a^2 M_n^*((X - M_n^*(X))^2) + 2ab M_n^*((X - M_n^*(X)) \cdot (Y - M_n^*(Y))) +$$
$$+ b^2 M_n^*(Y - M_n^*(Y))^2) = a^2 D_n^*(X) + b^2 D_n^*(Y) +$$
$$+ 2ab M_n^*((X - M_n^*(X)) \cdot (Y - M_n^*(Y))$$

Оскільки величини X і Y є незалежними, то за властивостями 4,3 та 1, маємо:

$$M_n^*((X - M_n^*(X)) \cdot (Y - M_n^*(Y))) = M_n^*(XY - XM_n^*(Y) - YM_n^*(X) +$$
$$+ M_n^*(X) M_n^*(Y)) = M_n^*(XY) - M_n^*(X) M_n^*(Y) - M_n^*(Y) M_n^*(X) +$$
$$+ M_n^*(X) M_n^*(Y) = M_n^*(X) M_n^*(Y) - M_n^*(X) M_n^*(Y) -$$
$$- M_n^*(Y) M_n^*(X) + M_n^*(X) M_n^*(Y) = 0.$$

Тому

$$D_n^*(aX + bY) = a^2 D_n^*(X) + b^2 D_n^*(Y).$$

Дана властивість має місце для будь-якої скінченної кількості попарно незалежних випадкових величин X_j, $j \in \overline{1, r}$:

$$D_n^*\left(\sum_{j=1}^{r} a_j X_j\right) = \sum_{j=1}^{r} a_j^2 D_n^*(X_j).$$

8. Закон великих чисел. Розглянемо послідовність попарно незалежних простих випадкових величин X_k, $k \in N$. Тоді випадкові величини $\sum_{k=1}^{m} X_k$ та $Y_m = \frac{1}{m}\sum_{k=1}^{m} X_k$, $m \in N$ також будуть простими випадковими, причому за властивостями статистичного математичного сподівання і дисперсії

$$M_n^*(\sum_{k=1}^{m} X_k) = \sum_{k=1}^{m} M_n^*(X_k),\ M_n^*(Y_m) = \frac{1}{m}\sum_{k=1}^{m} M_n^*(X_k) \text{ і } D_n^*(Y_m) = D_n^*\left(\frac{1}{m}\sum_{k=1}^{m} X_k\right) = \frac{1}{m^2}\sum_{k=1}^{m} D_n^*(X_k).$$

Згадуючи нерівність Чебишова, дістаємо:

$$P_n^* = \left(\left|\frac{1}{m}\sum_{k=1}^{m} X_k - \frac{1}{m}\sum_{k=1}^{m} M_n^*(X_k)\right|^2 \geq \varepsilon^2\right) = P_n^*\left(\left|\frac{1}{m}\sum_{k=1}^{m} X_k - \frac{1}{m}\sum_{k=1}^{m} M_n^*(X_k)\right| \geq \varepsilon\right) \leq$$

$$\leq \frac{1}{\varepsilon^2} M_n^*\left(\left(\frac{1}{m}\sum_{k=1}^{m} X_k - \frac{1}{m}\sum_{k=1}^{m} M_n^*(X_k)\right)^2\right) = \frac{1}{\varepsilon^2} D_n^*(Y_m) = \frac{1}{m^2\varepsilon^2}\sum_{k=1}^{m} D_n^*(X_k).$$

Отже,

$$P_n^*\left(\left|\frac{1}{m}\sum_{k=1}^{m} X_k - \frac{1}{m}\sum_{k=1}^{m} M_n^*(X_k)\right| \geq \varepsilon\right) \leq \frac{1}{m^2\varepsilon^2}\sum_{k=1}^{m} D_n^*(X_k) \tag{4}$$

а тому ліва частина останньої нерівності є як завгодно малою за умови, що такою є права частина цієї нерівності. У цьому і полягає закон великих чисел для простих випадкових величин.

9. Частинні випадки закону великих чисел. 1. Умова як завгодної малості правої частини нерівності (4) напевно виконується, коли число m досить велике, а

$$D_n^*(X_k) \leq C \quad \forall k \in \overline{1,m}, \tag{5}$$

оскільки тоді $0 \leq \frac{1}{m^2\varepsilon^2}\sum_{k=1}^{m} D_n^*(X_k) \leq \frac{C \cdot m}{m^2 E^2} = \frac{C}{m\,\varepsilon^2}$ і останній дріб є як завгодно малим, коли m досить велике.

Зокрема, якщо за умови (5) виконана ще умова $M_n^*(X_k) = a, \quad \forall\ k \in N$, тобто якщо всі $M_n^*(X_k)$ однакові і дорівнюють a, то нерівність (4) набуває вигляду

$$P_n^*\left(\left|\frac{1}{m}\sum_{k=1}^{m} X_k - a\right| \geq \varepsilon\right) \leq \frac{C}{m\varepsilon^2} \Leftrightarrow P_n^*\left(\left|\frac{1}{m}\sum_{k=1}^{m} X_k - a\right| < \varepsilon\right) \geq 1 - \frac{C}{m\varepsilon^2},$$

а закон великих чисел формулюється так: *статистична ймовірність досить значного відхилення середніх арифметичних значень випадкових величин X_k від їхнього спільного статистичного математичного сподівання є як завгодно малою величиною, коли кількість m випадкових величин досить велика*.

2. Нехай X_k – індикатор події A у k-му випробуванні з послідовності m незалежних випробувань, у кожному з яких подія A відбувається з статистичною ймовірністю $p = P_n^*(A)$ або не відбувається з статистичною ймовірністю $q = 1 - p$, причому $n \geq m$. Тоді (див. приклад 8) $M_n^*(X_k) = p = P_n^*(A)$ і $D_n^*(X_k) = pq \leq \frac{1}{4}\ \forall\ k \in \overline{1,m}$, а $\frac{1}{m}\sum_{k=1}^{m} X_k = P_m^*(A)$, $\frac{1}{m}\sum_{k=1}^{m} M_n^*(X_k) = p = P_n^*(A)$ і $\sum_{k=1}^{m} D_n^*(X_k) = mpq \leq \frac{1}{4}m$. Тому нерівність (4) набуває вигляду

$$P_n^*(|P_m^*(A) - P_n^*(A)| \geq \varepsilon) \leq \frac{pq}{m\varepsilon^2} \leq \frac{1}{4m\varepsilon^2} \Leftrightarrow P_n^*(|P_m^*(A) - P_n^*(A)| < \varepsilon) \geq$$

$$\geq 1 - \frac{pq}{m\varepsilon^2} \geq 1 - \frac{1}{4m\varepsilon^2}.$$

Отже, закон великих чисел для статистичних ймовірностей формулюється так: *якщо у кожному k-му випробуванні подія A відбувається із статистичною ймовірністю $P_n^*(A)$, а кількість m цих випробувань досить велика (а тому досить велике число n), то статистична ймовірність досить значного відхилення $P_m^*(A)$ від $P_n^*(A)$ є як завгодно малою, а статистична ймовірність як завгодної близькості $P_m^*(A)$ до $P_n^*(A)$ є досить великою, тобто як завгодно близькою до* 1.

Таким чином, саме відповідно до закону великих чисел кажуть, що *статистичні ймовірності $P_n^*(A)$ і $P_m^*(A)$ події A для великих серій випробувань скупчуються одна біля одної, а тому й біля певного числа $P(A)$, яке й називають ймовірністю події A.*

10. Висновки. 1. Наведені факти підтверджують думку, що статистичні ймовірності випадкових подій дають добрі наближення ймовірностей цих подій. Це відчував ще три сторіччя тому один з творців з теорії ймовірностей Я. Бернуллі, який писав: "І що не дано нам вивести a priori, то, принаймні, можна дістати a posteriori, тобто з багаточисельних спостережень результатів подібних прикладів. Тому потрібно передбачати, що деяке явище згодом може відбутися у стількох же випадках, у скількох воно раніше було відмічено при подібних же умовах як таке, що відбулося. Цей емпіричний спосіб визначення числа випадків за спостереженнями не новий і не незвичайний…, тому усі дотримуються його у повсякденній практиці"[5].

Зауважимо, що у сучасній практиці цей спосіб використовують незрівнянно частіше і ефективніше, ніж у часи Я.Бернуллі. Причиною цього є надзвичайно високий рівень розвитку сучасної математики у цілому і теорії ймовірностей – одного з найважливіших розділів математики, що має численні практичні застосування, у яких найчастіше використовують статистичні ймовірності.

2. Випадкові величини можуть мати і нескінченну кількість значень, проте виявляється, що кожну з них можна "досить добре наблизити" простими випадковими величинами так, що основні властивості довільних випадкових величин випливатимуть з основних властивостей простих випадкових величин.

3. Частинні випадки закону великих чисел, пов'язані із скупченням статистичних ймовірностей події A біля її ймовірності, можна дістати (див., наприклад, [6 с.43-55]) методами, не пов'язаними з простими випадковими величинами, проте відповідні міркування є значно складнішими, ніж ті, що пов'язані з простими випадковими величинами. Цим ще раз підтверджується принцип "узагальнення для полегшення", згідно з яким загальніши й погляд на проблему часто дозволяє знайти значно ефективніше розв'язання цієї проблеми.

ЛІТЕРАТУРА

1. Жалдак М.І., Михалін Г.О., Елементи стохастики з комп'ютерною підтримкою. – К.:РННЦ "ДІНІТ", 2004–107 с.
2. Жалдак М.І., Михалін Г.О. Про вивчення елементів стохастики в школі // Математика в школі –2004. –№ 9-10, с.7-12.
3. Гихман І.І., Скороход А.В., Ядренко М.І. Теория вероятновтей и математическая статистика.– К.: Вища Школа, 1979-с.408.
4. Феллер В. Введение в теорию вероятностей и ее приложения. – М.: Мир, 1984.–с.528
5. Майстров Л.Е. Теорія ймовірностей. Історичний очерк.– М.: Наука, 1967.–с. 320.
6. Скороход А.В. Вероятность вокруг нас. К.: Наукова думка, 1980.– 196с.