Жалдак М.І. , Михалін Г.О., Біляй І.М.
НПУ імені М.П. Драгоманова

Використання між предметних зв'язків та аналогій у процесі навчання теорії ймовірностей майбутніх учителів математики

1. Вступ. В Україні шкільний курс математики збагатився новою змістовою лінією – стохастичною майже 20 років тому, проте, на жаль, досі опанування стохастичного матеріалу і в загальноосвітніх школах, і в багатьох педагогічних університетах здійснюється на досить низькому рівні, який відомий американський математик У. Феллер влучно охарактеризував як напівмістичний [1, с. 11]. Основною причиною цього є те, що за основу побудови навчального курсу з теорії ймовірностей беруть так зване «класичне означення» ймовірності, запропоноване П. Лапласом майже 200 років тому. У найкращому випадку так побудовані навчальні курси відповідають рівню розвитку теорії ймовірностей двохсотрічної давнини. Разом з тим за останні 200 років теорія ймовірностей не тільки збагатилася величезною кількістю важливих понять і тверджень, а й перетворилася із «*сукупності результатів напівмістичних міркувань і маніпуляцій*» у струнку математичну теорію, що має численні важливі практичні застосування у багатьох галузях. І зараз теорія ймовірностей продовжує бурхливо розвиватися, збагачуючись новими математичними фактами та новими сферами застосувань.

Врахування сучасного стану розвитку математичної науки у навчальних курсах з теорії ймовірностей і для школярів, і для майбутніх учителів дозволяє не тільки строго математично подати навчальний матеріал, а й зробити це досить прозоро, спираючись на доступні школярам і студентам аналогії, широко використовуючи міжпредметні зв'язки.

2. Поняття випадкової величини. Побудова навчального курсу з теорії ймовірностей на основі «класичного означення» ймовірності, як правило, поєднується з намаганням ввести поняття випадкової величини безвідносно до поняття ймовірнісного простору. Таке намагання У. Феллер назвав *«демонстрацією мистецтва вводити в оману»* [1, с. 14]. Досить часто при цьому навіть не розкривають зв'язок між поняттями «випадкова величина» та «відповідність», «відображення», «функція», ігноруючи *теоретико-множинний підхід* до вивчення цих понять. Думку про недоцільність такого підходу до навчання шкільного курсу математики часто обумовлюють тим, що вчителі не володіють цим підходом, а майбутніх вчителів не навчають такого підходу, оскільки він у шкільних підручниках не використовується.

Разом з тим при використанні *теоретико-множинного підходу* з'являються можливості досить прозоро ввести найабстрактніші математичні поняття і продемонструвати їх практичні застосування [2].

Так, серед означень поняття функції, які зустрічаються у шкільних підручниках, найвдалішим можна вважати таке: *функція – це така залежність між двома змінними, при якій кожному значенню незалежної змінної відповідає єдине значення залежної змінної* [3]. Проте відповідь на запитання: «Що таке змінна та залежність між змінними?» жоден шкільний підручник не надає. Разом з тим за допомогою теоретико-множинного підходу одержати таку відповідь досить просто.

2.1. Відповідність, відображення, функція. Перш ніж починати вивчення поняття випадкової величини, доцільно перевірити, наскільки добре володіють майбутні учителі поняттями відповідності, відображення, функції та пов'язаними з ними поняттями образа, прообраза, оберненої відповідності тощо. Якщо за результатами перевірки знання деяких студентів виявляються невисокими, доцільно, щоб вони опрацювали необхідний матеріал самостійно або під керівництвом викладача за звичайним підручником або за відповідним дистанційним курсом. В результаті такого опрацювання студенти повинні чітко уявляти, що:

- *змінну* можна тлумачити як довільний (неконкретизований) елемент x даної множини A, а *значенням цієї змінної* може бути кожен (конкретизований) елемент $a \in A$, і тоді пишуть $x = a$; щоб задати змінну, треба задати множину A її значень;
- якщо задано непорожні множини A і B (не обов'язково різні), то з елементів цих множин $x \in A$ та $y \in B$ можна утворити *впорядковані пари* (x, y) та ввести *співвідношення рівності цих пар*: $(x_1, y_1) = (x_2, y_2)$ тоді і тільки тоді, коли $x_1 = x_2$ і $y_1 = y_2$;
- сукупність усіх впорядкованих пар (x, y), де $x \in A$, $y \in B$, називають *декартовим добутком множин* A і B і позначають $A \times B$;
- будь-яку підмножину F декартового добутку $A \times B$ множин A і B називають *відповідністю між множинами* A і B або *залежністю між змінними* $x \in A$ та $y \in B$; при цьому x називають незалежною змінною а y – *залежною змінною*; якщо пара $(x, y) \in F$, то кажуть, що *елемент* $y \in B$ *відповідає елементу* $x \in A$ при відповідності F між множинами A і B;

- для довільної відповідності (залежності) F між множинами A і B кожному елементу $x \in A$ може відповідати або один, або кілька, або жоден з елементів множини B; сукупність усіх елементів $y \in B$, що відповідають елементу $x \in A$, називають *образом* цього *елемента* при відповідності (залежності) F, а об'єднання образів усіх елементів множини $U \subset A$ називають *образом цієї множини* при відповідності (залежності) F і позначають $F(U)$; зокрема $F(\{x\})$ – це образ елемента x, який позначають також $F(x)$;
- для довільної відповідності F між множинами A і B кожен елемент $y \in B$ може відповідати одному, або кільком, або жодному елементові із множини A; сукупність усіх елементів $x \in A$, яким поставлено у відповідність елемент $y \in B$, називається прообразом елемента y при відповідності F, а об'єднання прообразів усіх елементів множини $V \subset B$ називають прообразом цієї множини при відповідності F і позначають $F^{-1}(V)$; зокрема $F^{-1}(\{y\})$ – це прообраз елемента y, який позначають також як $F^{-1}(y)$. При цьому $F^{-1}(F(U)) = U$ для будь-якої множини $U \subset A$, $A \neq \varnothing$, $F(U) \neq \varnothing$, а також $F^{-1}(F(V)) = V$ для будь-якої множини $V \subset B$, $V \neq 0$, $F^{-1}(V) \neq \varnothing$;
- для будь-якої відповідності F між непорожніми множинами A і B прообрази мають наступні важливі властивості: Якщо $F(A) = B$, $V \subset B$, $V_i \subset B$, $V \neq \varnothing$, $V_i \neq \varnothing$, $B \setminus V \neq \varnothing$, $F^{-1}(V) \neq 0$, $F^{-1}(V_i) \neq \varnothing$, $F^{-1}(B \setminus V) \neq \varnothing$, то $F^{-1}(B) = A$, $F^{-1}(B \setminus V) = A \setminus F^{-1}(V)$ і $F^{-1}\left(\bigcup_i V_i\right) = \bigcup_i F^{-1}(V_i)$.

  Звідси випливає, що коли S_A – сукупність деяких підмножин множини A, що задовольняє умови 1_s-3_s простору подій, тоді сукупність S_B підмножин V множини B, для яких $F^{-1}(V) \in S$, також задовольняє умови 1_s-3_s;
- *відображенням* множини A у множину B називають таку відповідність (залежність) F між цими множинами, коли кожному елементу (кожному значенню змінної) $x \in A$ відповідає один елемент $y \in B$, який називають *образом елемента x*; при цьому позначають F: $A \to B$, множину A називають *множиною (областю) визначення відображення* F, а множину $F(A) \subset B$ називають *множиною (областю) значень відображення F*;
- *відображення F*: $A \to B$ називають також *функцією* F з множини A у множину B і позначають $y = F(x)$, $x \in A$, $y \in B$; при цьому елемент y називають *значенням функції F* у точці x, множину A називають *множиною (областю) визначення функції F*, а B – *множиною (областю) значень* цієї *функції* і кажуть також, що за функцією F здійснюється відображення множини A у множину B;
- *відображення F* множини A у множину B називають *взаємно однозначним* і позначають F: $A \leftrightarrow B$, коли кожен елемент $y \in B$ є образом єдиного елемента $x \in A$. При цьому існує *обернене відображення* F^{-1}: $B \leftrightarrow A$ (обернена функція $x = F^{-1}(y)$, $y \in B$, $x \in A$), коли кожному значенню $y \in B$ відповідає таке значення $x \in A$, для якого $F(x)=y$. *Множини A* і *B* називають *еквівалентними* (і кажуть також, що вони *мають однакову кількість елементів*), якщо існує відображення F: $A \leftrightarrow B$.

2.2. Випадкові величини, як вимірні функції. Намагання досліджувати і тлумачити випадковість, не звертаючись до відповідних математичних моделей, майже завжди вводить в оману. Про це свідчать відомі парадокси теорії ймовірностей. Саме прагнення позбутися таких парадоксів призвело до побудови аксіоматичних математичних теорій, включаючи і теорію ймовірностей.

Слід зауважити, що доступність навчального матеріалу досягається не за рахунок свідомого ігнорування сучасного рівня розвитку математики і апелювання до першоджерел сто, двісті чи навіть тисячолітньої давнини, а за рахунок суттєвого збільшення індуктивної складової процесу навчання у поєднанні з постановкою проблем (питань) та розв'язуванні їх разом із студентами.

Наприклад, перш ніж формулювати аксіоми 1_s-3_s:

1_s. $\Omega \in S$.

2_s. Якщо $A \in S$, то і $\overline{A} \in S$.

3_s. Якщо $A_i \in S$, $i = 1, 2, \ldots$, то і $\bigcup_i A_i \in S$.

якими визначають поняття простору S випадкових подій, доцільно за допомогою системи доступних прикладів і питань підвести студентів до висновку, що саме лише означення якоїсь операції над математичними об'єктами не гарантує існування результату цієї операції і треба ще задати умови такого існування. Саме тому випадковими подіями можна вважати лише такі підмножини простору Ω елементарних подій, які утворюють сукупність *S*, замкнену відносно операцій над подіями. Постановка питання: «Чи можна забезпечити цю замкненість якомога меншою кількістю умов?» приводить до висновку, що достатньо виконання лише трьох умов 1_s-3_s.

Так само індуктивно слід вводити поняття випадкової величини для майбутніх вчителів математики, пам'ятаючи, що у майбутньому вони можуть використати цей підхід для ознайомлення своїх учнів з поняттям випадкової величини.

Вивчення поняття випадкової величини можна розпочати з простого, але важливого прикладу функції, заданої на просторі Ω елементарних подій рівністю

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

де *A* – фіксована підмножина простору Ω елементарних подій.

Виникає питання: «*Коли і як можна знайти ймовірність того, що дана функція набуде значення 1?*» Зрозуміло, що умова «*X*(*E*) набуває значення 1» виконується тоді й тільки тоді, коли $E \in A$, тобто ця умова рівносильна тому, що відбувається подія *A*, тому підмножина $A \subset \Omega$ повинна бути подією із заданого простору подій *S*. Отже, якщо крім простору Ω елементарних подій задано простір подій *S* та ймовірність *P* на просторі *S*, то ймовірність того, що функція *X* набуває значення 1, можна знайти тоді й тільки тоді, коли $A \in S$. Тоді «*X*(*E*) набуває значення 1» означає відбування події $X^{-1}(\{1\}) = \{E \in \Omega : X(E) = 1\} = A$. При цьому $X^{-1}(\{0\}) = \{E \in \Omega: \ X(E) = 0\} = \overline{A}$ також є подією простору *S* і $P(X^{-1}(\{0\}) = P(\overline{A}) = 1 - P(A)$. Зрозуміло, що коли число $a \neq 0$ і $a \neq 1$, то множина $X^{-1}(\{a\}) = \{E \in \Omega : X(E) = a\} = \varnothing$ і тому є подією простору *S*, оскільки $\varnothing = \overline{\Omega} = \Omega \setminus \Omega \in S$, причому $P(X^{-1}(\{a\})) = P(\varnothing) = 0$.

Таким чином, для функції (1) і для будь-якого фіксованого числа *a* можна знайти ймовірність того, що $X = a$, тоді й тільки тоді, коли задано ймовірнісний простір (Ω, S, P) і множина $X^{-1}(\{a\}) = \{E \in \Omega : X(E) = a\}$ є подією з простору *S*.

Природним завершенням розглянутого прикладу є формулювання означення простої випадкової величини: *функція X(E),* $E \in \Omega$, *називається простою випадковою величиною стосовно простору* (Ω, S, P), *коли множина* Ω_X *значень цієї функції є скінченною, тобто* $\Omega_X = \{x_1, x_2, ..., x_k\}$, *і для будь-якого* $x_i \in \Omega_X$, $X^{-1}(\{x_i\}) \in S$. Зокрема це означає, що для будь-якого числа *a* множина розв'язків рівняння *X*(*E*)=*a* є подією з простору *S*, бо коли $a \overline{\in} \Omega_X$, $X^{-1}(a) = \varnothing \in S$.

Шляхом розв'язування достатньої кількості задач, пов'язаних з поняттям простої випадкової величини бажано підвести майбутніх учителів до наступного важливого питання: *якщо X – проста випадкова величина стосовно* (Ω, S, P), *то множина розв'язків рівняння X(E)=a є подією з простору S для будь-якого числа a. А чи буде тоді подією з простору S множина розв'язків нерівності* $X(E) < a$
($X(e) \leq a$, $X(e) > a$, $X(e) \geq a$)?

Разом із студентами можна досить просто знайти позитивну відповідь на поставлені питання. Логічним наслідком цієї відповіді буде цілком природне узагальнення поняття простої випадкової величини: *функція X(E),* $E \in \Omega$, *називається випадковою величиною стосовно простору* (Ω, S, P), *коли для будь-якого числа a множина розв'язків нерівності* $X(E) < a$, *є подією з простору S.*

Таким чином, задача визначення, чи є задана функція випадковою величиною стосовно ймовірнісного простору (Ω, S, P), зводиться до звичних студентам задач:

- розв'язування рівнянь, коли множина значень функції скінченна;
- розв'язування нерівностей, коли множина значень функції довільна (скінченна або нескінченна).

Після опанування наведених означень, можна узагальнити поняття випадкової величини на випадок S/S_X-вимірних функцій: *функція X*(*E*), $E \in \Omega$, *називається випадковою величиною стосовно простору* (Ω, S, P) *або* S/S_X*-вимірною функцією, якщо для довільного* $V \in S_X$ $X^{-1}(V) \in S$.

Згідно з наведеними раніше важливими властивостями прообразів сукупність S_X задовольняє визначальні властивості простору подій 1_s-3_s.

Таким чином, в означенні S/S_X-вимірної функції сукупність S_X, визначається за простором (Ω, S, P) та функцією $X(E)$, $E \in \Omega$: серед усіх підмножин множини Ω_X вибирають такі підмножини V, для яких $X^{-1}(V) \in S$.

Слід зауважити разом з тим, що сукупність S_X не обов'язково містить всі підмножини $V \subset \Omega_X$, для яких $X^{-1}(V) \in S$.

Якщо серед підмножин сукупності S_X міститься кожен проміжок $(-\infty; x)$, $x \in (-\infty; +\infty)$, то такі випадкові величини називають *S-вимірними функціями*. Насправді тут маються на увазі також S/S_X-вимірні функції, але сукупність S_X явно не вказується.

Зауважимо, що S-вимірні функції є частковим випадком S/S_X-вимірних залежностей між множинами Ω та Ω_X. Щоб функція $X(E)$, $E \in \Omega$, була S-вимірною, необхідно, щоб Ω_X була множиною в деякому координатному просторі. Для S/S_X-вимірних залежностей належність до деякого координатного простору як множини Ω_X, так і множини Ω необов'язкова, і крім того така залежність необов'язково має бути функціональною.

Таким чином, при вивченні поняття випадкової величини студенти насамперед повинні зрозуміти, що:

- без задання ймовірнісного простору (Ω, S, P) математично коректно ввести поняття випадкової величини неможливо;
- не кожна функція, задана на просторі Ω елементарних подій, є випадковою величиною стосовно будь-якого простору (Ω, S, P);
- для того, щоб кожна функція, задана на просторі Ω, була випадковою величиною стосовно простору (Ω, S, P), необхідно й достатньо, щоб простір подій S був найширшим з можливих;
- в означенні випадкової величини не є важливим спосіб задання ймовірнісної міри P на просторі подій S; зокрема *ця ймовірність може бути й статистичною, проте це ані спрощує, ані ускладнює відповідну теорію*.

3. Розподіл ймовірностей на множині значень випадкової величини. Важливою характеристикою дійсної функції $X(E)$, $E \in \Omega$, є її графік, за яким можна знайти будь-яке значення цієї функції. Дві дійсні функції із спільною областю визначення вважають рівними тоді й тільки тоді, коли вони мають однакові графіки, тобто коли у довільній точці області визначення ці функції набувають однакових значень.

Для випадкових величин їх значення в окремих точках часто вже не є такими важливими, як для звичайних функцій. Значно важливішим є те, з якою ймовірністю ця випадкова величина набуває тих чи інших значень.

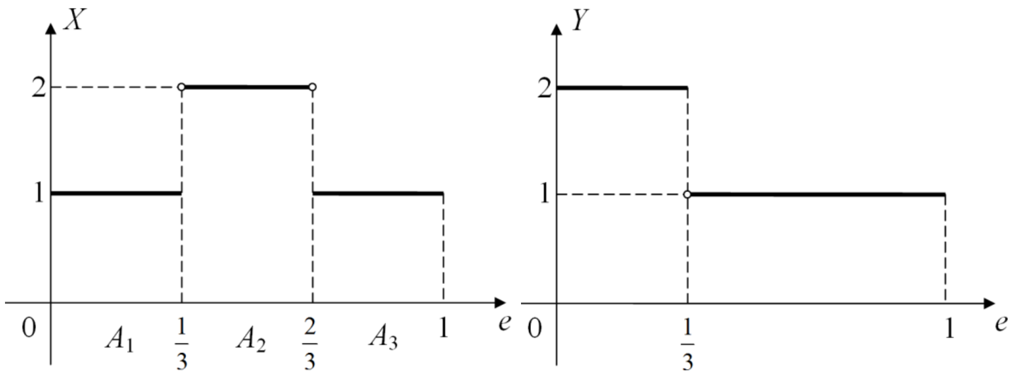

Рис. 1

Наприклад, функції $X(E)$ та $Y(E)$, $E \in [0;1]$, графіки яких зображено на рис. 1, є суттєво різними, як звичайні функції, і є по суті однаковими, як випадкові величини, оскільки вони мають однакові множини значень $\Omega_X = \{1, 2\}$, $\Omega_Y = \{1, 2\}$ і набувають однакових значень з однаковими ймовірностями: $P(X^{-1}(\{2\})) = P(Y^{-1}(\{2\})) = \dfrac{1}{3}$, а $P(X^{-1}(\{1\})) = P(Y^{-1}(\{1\})) = \dfrac{2}{3}$ за умови, що $\Omega = [0;1]$, S – сукупність підмножин Ω, що мають довжину (міру Лебега), і P – міра Лебега (довжина) множин (подій) $A \in S$.

Справді, за цих умов $P(X^{-1}(\{2\})) = P\left(\left(\frac{1}{3};\frac{2}{3}\right)\right) = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$, $P(Y^{-1}(\{2\})) = P\left(\left[0;\frac{1}{3}\right]\right) = \frac{1}{3}$,

$P(X^{-1}(\{1\})) = P\left(\left[0;\frac{1}{3}\right] \cup \left[\frac{2}{3};1\right]\right) = P\left(\left[0;\frac{1}{3}\right]\right) + P\left(\left[\frac{2}{3};1\right]\right) = \frac{1}{3} + 1 - \frac{2}{3} = \frac{2}{3}$, $P(Y^{-1}(\{1\})) = P\left(\left[\frac{1}{3};1\right]\right) = 1 - \frac{1}{3} = \frac{2}{3}$.

За допомогою означення випадкової величини (S/S_X-вимірної функції) $X(E)$, $E \in \Omega$, стосовно простору (Ω, S, P) можна побудувати новий ймовірнісний простір (Ω_X, S_X, P_X), для якого $\Omega_X = (-\infty;+\infty)$, S_X – найвужчий простір подій, що містить кожен проміжок $(-\infty;x)$ (цей *простір подій* S_X називають *борелівським* і кожну подію (множину) з S_X також називають *борелівською*), а ймовірність P_X задається умовою $P_X((-\infty;x)) = P(X^{-1}((-\infty;x)))$ для кожного $x \in (-\infty;+\infty)$. При цьому відповідність між подіями простору S_X та ймовірностями цих подій називають *розподілом ймовірностей* P_X *на множині* Ω_X *за подіями простору* S_X.

У випадку, коли виділено кілька подій $A_i \in S_X$, i=1, 2, …, які попарно несумісні, причому $\sum_i P_X(A_i) = 1$, говорять про *розподіл ймовірностей* P_X *на множині* Ω_X *за подіями* A_i, i=1, 2, …. Такий розподіл можна задати у вигляді таблиці 1:

Табл. 1

A_i	A_1	A_2	…	A_k	…
$p_i=P_X(A_i)$	p_1	p_2	…	p_k	…

У випадку, коли існують одноелементні події $A_i = \{x_i\}$, i=1, 2, …, простору S_X, для яких $p_i = P_X(\{x_i\}) > 0$ і $\sum_i p_i = 1$, одержується *дискретний розподіл ймовірностей* P_X *на множині* Ω_X *за елементарними подіями* x_i, i=1, 2, ….

Якщо $A_i = [a_{i-1};a_i) \in S_X$, i=1, 2, …, k, де $a_0 < a_1 < ... < a_k$, причому $P_X([a_{i-1};a_i)) = p_i$ і $\sum_{i=1}^{k} p_i = 1$, то говорять про *інтервальний розподіл ймовірностей* P_X *на множині* Ω_X *за проміжками* $[a_{i-1};a_i)$, i=1, 2, ….

Так, на множинах Ω_X та Ω_Y значень випадкових величин X та Y, графіки яких зображено на Рис. 1, однакові дискретні розподіли ймовірностей за елементарними подіями $x_1 = 1$ та $x_2 = 2$. Таблиця 1 для цих розподілів матиме вигляд:

Табл. 2

$x_i = y_i$	1	2
$P_X(\{x_i\}) = P_Y(\{y_i\})$	$\frac{2}{3}$	$\frac{1}{3}$

Отже, зовсім різні, за виглядом графіка, випадкові величини X та Y можуть мати однакові розподіли ймовірностей, а тому й однакові розподіли ймовірностей та числові характеристики цих розподілів.

При вивченні розподілів ймовірностей цілком природно і безумовно корисно звертатися до фізичних аналогій: розподілів маси за окремими точками (тобто дискретних) та за частинами суцільних фізичних тіл з певною густиною. Окрім цього корисно згадати деякі відомості, які студенти вивчали у курсі математичного аналізу, наприклад:

- кожна диференційовна функція є неперервною, проте не навпаки (існують функції, що є неперервними на $(-\infty;+\infty)$, проте не мають похідної у жодній точці);
- якщо функція $F(x)$ монотонна і неперервна на $(-\infty;+\infty)$, то вона обов'язково має похідну $F'(x)=f(x)$ майже у кожній точці x, проте не завжди за цією похідною можна відновити функцію $F(x)$ шляхом інтегрування, тобто не обов'язково $F(x) = \int_{-\infty}^{x} F'(t)dt$ навіть коли функція $F'(x) = f(x)$ є інтегровною на кожному проміжку $(-\infty;x)$. Це означає, що *теоретично поняття щільності розподілу ймовірностей можна ввести для будь-якого розподілу, проте ця щільність буде корисною лише для абсолютно неперервних розподілів.* На практиці часто про це забувають і обґрунтування того, що функція $F(x)$ розподілу ймовірностей є абсолютно неперервною, зводять до відшукання похідної $F'(x) = f(x)$,

забуваючи про перевірку умови $F(x)=\int\limits_{-\infty}^{x} f(t)dt$.

4. Математичне сподівання, як середнє значення випадкової величини. Вивчення однієї з найважливіших числових характеристик розподілу ймовірностей на множині Ω_X значень випадкової величини X – математичного сподівання $M[X]$ природно розпочати з нагадування фізичної аналогії розподілу ймовірностей P_X на множині Ω_X значень випадкової величини X: це є розподіл одиничної маси. А з курсів фізики та математичного аналізу відомо, що такий розподіл може мати центр мас, а також відомо, як можна знайти цей центр мас.

4.1. Математичне сподівання простої випадкової величини. Найпростіше координата центра мас x_c (яку часто позначають також $\bar{x}$) визначається, коли маса (ймовірність) P_X розподілена дискретно за точками x_i, i=1, 2, …, k, для яких $P_X(\{x_i\})=p_i>0$ і $\sum\limits_{i=1}^{k} p_i=1$. Тоді $x_c=\sum\limits_{i=1}^{k} x_i p_i$, випадкову величину X з множиною значень $\Omega_X=\{x_1, x_2, ..., x_k\}$ можна вважати простою стосовно певного ймовірнісного простору (Ω, S, P), причому ймовірність P_X пов'язана з ймовірністю P рівністю $P_X(\{x_i\})=P(X^{-1}(\{x_i\}))$, i=1, 2, …, k.

Природним завершенням сказаного є формулювання означення математичного сподівання простої випадкової величини: *число* $x_c=\sum\limits_{i=1}^{k} x_i p_i$ *називають математичним сподіванням або середнім значенням простої випадкової величини* X (*стосовно простору* (Ω, S, P)), *якщо* $p_i=P_X(\{x_i\})=P(X^{-1}(x_i))$ *і* $\sum\limits_{i=1}^{k} p_i=1$. При цьому позначають $x_c=M[X]$ і тому

$$M[X]=\sum_{i=1}^{k} x_i p_i=\sum_{i=1}^{k} x_i P_X(\{x_i\})=\sum_{i=1}^{k} x_i P(X^{-1}(x_i)) \tag{2}$$

для будь-якої простої випадкової величини X з множиною значень $\Omega_X=\{x_1, x_2, ..., x_k\}$.

Останню суму називають *інтегралом Лебега від функції* $X(E)$, $E\in\Omega$, *стосовно міри* P і позначають $\int\limits_{\Omega} X(E)dP$, або $(L)\int\limits_{\Omega} X(E)dP$, або $\int\limits_{\Omega} X(E)P(dE)$.

Оскільки $P(X^{-1}(\{x_i\}))=P_X(\{x_i\})$, то

$$M[X]=\sum_{i=1}^{k} x_i P_X(\{x_i\})=\int\limits_{\Omega_X} xP_X(dX)=\int\limits_{\Omega} X(E)P(dE).$$

Враховуючи, що міра $P_X(\{x_i\})$ цілком визначається за функцією $F_X(x)$ розподілу ймовірностей P_X, тобто $P_X(\{x_i\})=F_X(x_i+0)-F_X(x_i)$, маємо:

$$M[X]=\sum_{i=1}^{k} x_i (F_X(x_i+0)-F_X(x_i)).$$

Останню суму називають інтегралом Стілтьєса від функції $\psi(x)=x$, $x\in(-\infty;+\infty)$ стосовно функції F_X розподілу ймовірностей P_X на множині значень простої випадкової величини X і позначають $\int\limits_{-\infty}^{+\infty} xdF_X(x)$.

Таким чином, з математичним сподіванням $M[X]$ (середнім значенням $x_c=M[X]$) простої випадкової величини X пов'язані рівності

$$x_c=M[X]=\sum_{i=1}^{k} x_i P(X^{-1}(\{x_i\}))=\int\limits_{\Omega} X(E)P(dE)=\int\limits_{\Omega_X} xP_X(dx)=\int\limits_{-\infty}^{+\infty} xdF_X(x). \tag{3}$$

Окрім цього з математичним сподіванням пов'язане поняття інтеграла Рімана. Це обумовлене тим, що *функція X(E),* $E\in\Omega$, *є простою випадковою величиною стосовно простору* (Ω, S, P) *тоді й тільки тоді, коли існує скінченна кількість попарно несумісних подій* $A_i\in S$ *і чисел* a_i, *i=1, 2, ..., n, таких, що* $\sum\limits_{i=1}^{n} A_i=\Omega$ *і* $X(E)=a_i$, $E\in A_i$, *i=1, 2, ..., n*. Враховуючи цей критерій та розподільний, переставний і сполучний закони для операцій над дійсними числами, легко переконатися, що

$$M[X]=\sum_{i=1}^{n}a_i P(A_i)\,. \qquad (4)$$

Останню суму називають *інтегралом Рімана* функції $X(E)$, $E\in\Omega$, стосовно міри P і позначають $(R)\int\limits_{\Omega}X(E)P(dE)$.

Відмінність між інтегралами Лебега та Рімана полягає у тому, що доданки у відповідних сумах (3) і (4) визначаються:

- для інтеграла Лебега – значеннями x_i випадкової величини X, і за цими значеннями знаходять множини $B_i=X^{-1}(\{x_i\})$, на кожній з яких функція $X(E)$ стала і набуває значення x_i;
- для інтеграла Рімана – множинами A_i, і за цими множинами знаходять числа a_i.

На практиці найчастіше точки E кожної фіксованої множини A_i є близькими між собою і в цих близьких точках E відповідні значення функції $X(E)$ також повинні бути близькими – це характерне для інтеграла Рімана. А для інтеграла Лебега характерним є те, що за близькими до x_i значеннями функції $X(E)$ знаходять множини $X^{-1}(\{x_i\})$ точок E, які між собою не обов'язково близькі, а значення функції $X(E)$ у цих точках є близькими між собою і до числа x_i.

Наприклад, для функції $X(E)$, $E\in[0;1]$, графік якої зображено на рис. 1, маємо

$A_1=\left[0;\frac{1}{3}\right]$; $X(E)=1=a_1$, коли $E\in A_1$;

$A_2=\left(\frac{1}{3};\frac{2}{3}\right)$; $X(E)=2=a_2$, коли $E\in A_2$;

$A_3=\left[\frac{2}{3};1\right]$; $X(E)=1=a_3$, коли $E\in A_3$;

$X^{-1}(1)=A_1\cup A_3=\left[0;\frac{1}{3}\right]\cup\left[\frac{2}{3};1\right]$; $X^{-1}(2)=\left(\frac{1}{3};\frac{2}{3}\right)$.

Тому інтеграл Лебега

$$\int\limits_{[0;1]}X(E)P(dE)=1\cdot P\left(\left[0;\frac{1}{3}\right]\cup\left[\frac{2}{3};1\right]\right)+2\cdot P\left(\left(\frac{1}{3};\frac{2}{3}\right)\right)=1\cdot\left(\frac{1}{3}+1-\frac{2}{3}\right)+2\cdot\left(\frac{2}{3}-\frac{1}{3}\right)=\frac{4}{3},$$

а інтеграл Рімана

$$(R)\int\limits_{[0;1]}X(E)dP=1\cdot P\left(\left[0;\frac{1}{3}\right]\right)+2\cdot P\left(\left(\frac{1}{3};\frac{2}{3}\right)\right)+1\cdot P\left(\left[\frac{2}{3};1\right]\right)=\frac{1}{3}+\frac{2}{3}+\frac{1}{3}=\frac{4}{3}=(L)\int\limits_{[0;1]}X(E)P(dE)\,.$$

Отже, *інтеграли Рімана і Лебега* для розглянутої простої випадкової величини *однакові* і це має місце *для довільної простої випадкової величини* $X(E)$, $E\in\Omega$, тобто

$$(R)\int\limits_{\Omega}X(E)P(dE)=(L)\int\limits_{\Omega}X(E)P(dE)=M[X]\,.$$

Для інших функцій існування інтеграла Рімана завжди гарантує існування інтеграла Лебега та рівність цих інтегралів. Навпаки не так: множина функцій, інтегровних за Лебегом, є значно ширшою, ніж інтегровних за Ріманом. Про це слід нагадати студентам і надалі мати справу лише з інтегралом Лебега.

4.2. Математичне сподівання дискретної випадкової величини. За означенням *випадкова величина* $X(E)$, $E\in\Omega$, стосовно простору (Ω,S,P) є *дискретною*, якщо існує скінченна або зчисленна кількість значень x_i, i=1, 2, …, для яких $p_i=P_X(\{x_i\})=P(X^{-1}(\{x_i\}))>0$ і $\sum\limits_i p_i=1$.

Якщо позначити $A_i=X^{-1}(\{x_i\})$, i=1, 2, …, $A_0=\overline{\bigcup\limits_i A_i}=\bigcap\limits_i\overline{A_i}$, $x_0=0$, $X_i(E)=x_iI_{A_i}(E)$, де $I_{A_i}(E)$ – індикатор події A_i, то легко переконатися, що майже напевно (стосовно ймовірності P) $X(E)=\sum\limits_i X_i(E)$ і кожна функція $X_i(E)$ є простою випадковою величиною (стосовно (Ω,S,P)), яка набуває значення x_i на множині A_i і нуль – на множині $\overline{A_i}$.

Оскільки для простої випадкової величини X_i її математичне сподівання $M[X_i]=\int\limits_{\Omega}X_i(E)P(dE)$, то природно покласти

$$M[X] = M[\sum_i X_i] = \int_\Omega X(E)P(dE) = \int_\Omega \sum_i X_i(E)P(dE) = \sum_i \int_\Omega X_i(E)P(dE) = \sum_i x_i P(X^{-1}(x_i)) = \sum_i x_i P_X(x_i).$$

Проте одразу виникає питання: «*Чи існує інтеграл від суми функціонального ряду* $\sum_i X_i(E)$ *і чи можна цей ряд інтегрувати почленно*?»

Звертаючись до курсу математичного аналізу (див., наприклад, [4, с. 384]), дістаємо відповідь на поставлене питання: *для того, щоб функціональний ряд* $\sum_i X_i(E)$ *майже напевно збігався до функції X*(*E*) *і була правильною рівність* $\int_\Omega X(E)P(dE) = \sum_i \int_\Omega X_i(E)P(dE)$, *необхідно й достатньо, щоб* $\sum_i \int_\Omega |X_i(E)|P(dE) < +\infty$.

Оскільки

$$\sum_i \int_\Omega |X_i(E)|P(dE) = \sum_i \int_\Omega |x_i| I_{A_i}(E)P(dE) = \sum_i |x_i| P(A_i) = \sum_i |x_i| P(X^{-1}(x_i)) = \sum_i |x_i| P_X(x_i),$$

то за умови $\sum_i |x_i| P_X(x_i) < +\infty$ відповідь на поставлене запитання позитивна і тому природно, що *математичне сподівання* $M[X]$ *або середнє значення* $\bar{x} = M[X]$ *дискретної випадкової величини* $X(E)$, $E \in \Omega$, *визначається рівністю*

$$\bar{x} = M[X] = \sum_i x_i P(X^{-1}(\{x_i\})) = \sum_i x_i P_X(\{x_i\})$$

за умови $\sum_i |x_i| P_X(\{x_i\}) < +\infty$. При цьому *функцію X*(*E*) вважають *інтегровною на множині* Ω *за Лебегом стосовно міри P*, а функцію $\psi(x) = x$ – інтегровною на множині $\Omega_X = (-\infty; +\infty)$ стосовно міри P_X і стосовно функції $F_X(x) = P_X(-\infty; +\infty)$, причому

$$\bar{x} = M[X] = \sum_i x_i P_X(x_i) = \int_\Omega X(E)P(dE) = \int_{\Omega_X} x P_X(dx) = \int_{-\infty}^{+\infty} x dF(x). \tag{5}$$

Оскільки для простої випадкової величини умова $\sum_i |x_i| P(X^{-1}(x_i)) < \infty$ очевидно виконується, то наведене означення математичного сподівання дискретної випадкової величини є природним узагальненням означення математичного сподівання простої випадкової величини. При цьому міжпредметні зв’язки з курсом математичного аналізу допомогли обґрунтувати необхідність умови $\sum_i |x_i| P_X(x_i) < +\infty$. Без цього обґрунтування було не зрозуміло, чому математичне сподівання дискретної випадкової величини *X* не можна визначити рівністю

$$M[X] = \sum_i x_i P(X^{-1}(x_i))$$

лише за умови, що остання сума існує.

4.3. Математичне сподівання довільної випадкової величини. Розглянуті означення математичного сподівання $M[X]$ або середнього значення $\bar{x} = M[X]$ простої і дискретної випадкової величини та одержані при цьому рівності (4) і (5) підказують природне узагальнення цих означень на довільну випадкову величину за допомогою рівності

$$\bar{x} = M[X] = \int_\Omega X(E)P(dE) = \int_{\Omega_X} x P_X(dx) = \int_{-\infty}^{+\infty} x dF_X(x), \tag{6}$$

де міра P_X на борелівському просторі S_X задається рівністю $P_X((-\infty; x)) = P(X^{-1}((-\infty; x)))$, а $F_X(x) = P_X((-\infty; x))$.

Але одразу виникає питання: «*Що розуміти під кожним з інтегралів рівності (6) і яка умова забезпечує існування та рівність цих інтегралів*?»

Відповідь на це питання випливає із загальної теорії інтеграла Лебега: *означення (6) є коректним для будь-якої випадкової величини* $X(E)$, $E \in \Omega$ (*стосовно* (Ω, S, P)), *що є інтегровною за Лебегом на множині* Ω *стосовно міри P*. При цьому умова інтегровності функції *X*(*E*) гарантує інтегровність за Лебегом на множині $\Omega_X = (-\infty; +\infty)$ функції $\psi(x) = x$ стосовно міри P_X і стосовно функції $F_X(x) = P_X((-\infty; x)) = P(X^{-1}((-\infty; x)))$, а також виконання умови

$$\int_{\Omega} |X(E)|P(dE) = \int_{\Omega_X} |x| P_X(dx) = \int_{-\infty}^{+\infty} |x| dF_X < +\infty .$$

Якщо студенти – майбутні учителі математики не мають уявлення про теорію інтеграла Лебега (а таке, на жаль, має місце досить часто), доцільно і необхідно сформувати у них таке уявлення, використовуючи для цього, можливо, дистанційні навчальні курси.

Загальне означення інтеграла Лебега можна ввести і на лекції. Здійснити це можна, наприклад, так.

Насамперед доцільно сформулювати **основну ідею**: *будь-яку випадкову величину X(E) можна як завгодно добре наблизити дискретними (і навіть простими) випадковими величинами $X_h(E)$, математичні сподівання яких $M[X_h]$ можуть як завгодно мало відрізнятися від певного числа. Тоді це число і називають математичним сподіванням $M[X]$ випадкової величини X(E), $E \in \Omega$, або інтегралом Лебега $\int_{\Omega} X(E)P(dE)$.*

Оскільки разом з існуванням інтеграла $\int_{\Omega} X(E)P(dE)$ необхідно забезпечити ще умову $\int_{\Omega} |X(E)|P(dE) < +\infty$, спочатку можна розглянути довільну невід'ємну випадкову величину $X(E)$, для якої $|X(E)| = X(E)$, $E \in \Omega$. Тоді множину $\Omega_X = [0; +\infty)$ можливих значень цієї випадкової величини можна подати як об'єднання проміжків $[(i-1)h; ih)$, i=1, 2, …, де $h > 0$ – довільне фіксоване число: $\Omega_X = \bigcup_{i=1}^{\infty} [(i-1)h; ih)$. З кожним з таких проміжків пов'язується подія $A_i = \{E \in \Omega : (i-1)h \le X(E) < ih\} \in S$, а за сукупністю цих подій можна ввести дискретну випадкову величину $X_h(E) = (i-1)h$, коли $E \in A_i$, i=1, 2, ….

Зрозуміло, що $|X(E) - X_h(E)| \le h$ для будь-якого $E \in \Omega$, тобто зменшуючи $h > 0$, можна досягти того, що $X_h(E)$ як завгодно добре наближає $X(E)$.

Для дискретної величини $X_h(E)$ її математичне сподівання

$$M[X_h] = \sum_{i=1}^{\infty} (i-1)hP(A_i) = \sum_{i=1}^{\infty} (i-1)hP_X([(i-1)h; ih)) = \sum_{i=1}^{\infty} (i-1)h(F_X(ih) - F_X((i-1)h) .$$

за умови збіжності останніх рядів.

Виявляється, що, коли ця умова виконується для деякого $h > 0$, тоді вона виконується і для будь-якого $h > 0$ [4, с. 417-418]. При цьому обов'язково існує скінченна границя $\lim_{0<h\to 0} M[X_h]$, яку позначають $M[X]$ і називають *математичним сподіванням, або середнім значенням невід'ємної випадкової величини X*, або *інтегралом Лебега*:

- $\int_{\Omega} X(E)P(dE)$ – від функції $X(E)$ на множині Ω стосовно міри P;
- $\int_{\Omega_X} xP_X(dx)$ – від функції $\psi(x) = x$ на множині Ω_X стосовно міри P_X;
- $\int_{0}^{+\infty} xdF_X(x)$ – функції $\psi(x) = x$ на множині Ω_X стосовно функції $F_X = P_X((-\infty; x))$.

Саму невід'ємну випадкову величину X називають при цьому *інтегровною на множині Ω стосовно міри (ймовірності) P*, а функцію $\psi(x) = x$ – інтегровною на множині Ω_X стосовно міри (ймовірності) P_X або стосовно функції F_X .

Зауважимо, що, коли випадкова величина X є абсолютно неперервною, тобто $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$, де f_X – щільність розподілу ймовірностей P_X , тоді

$$M[X_h] = \sum_{i=1}^{\infty} (i-1)hP(A_i) = \sum_{i=1}^{\infty} (i-1)hP_X([(i-1)h; ih)) = \sum_{i=1}^{\infty} (i-1)h \int_{(i-1)h}^{ih} f_X(t)dt =$$

$$= \sum_{i=1}^{\infty} (\int_{(i-1)h}^{ih} tf_X(t)dt + \int_{(i-1)h}^{ih} ((i-1)h - t) f_X(t)dt) = \int_{0}^{+\infty} tf_X(t)dt + \sum_{i-1}^{\infty} \int_{(i-1)h}^{ih} ((i-1)h - t) f_X(t)dt , \qquad (7)$$

причому

$$\left|\sum_{i=1}^{\infty}\int_{(i-1)h}^{ih}((i-1)h-t)f_X(t)dt\right| \le h\sum_{i=1}^{\infty}\int_{(i-1)h}^{ih}f_X(t)dt = h\int_0^{+\infty}f_X(t)dt = h.$$

Тому з рівності (7) випливає, що $M[X_h] \to \int_0^{+\infty} tf_X(t)dt$, коли $h \to 0$. Отже, для абсолютно неперервної інтегровної невід'ємної випадкової величини X її математичне сподівання (середнє значення) визначається за формулою

$$\bar{x} = M[X] = \int_{-\infty}^{+\infty} xf_X(x)dx. \qquad (8)$$

Випадок довільної випадкової величини X, що може набувати як додатних, так і від'ємних значень, зводиться до випадку невід'ємної випадкової величини за допомогою двох функцій:

$$X^+(E) = \begin{cases} X(E), & \text{коли } X(E) > 0, \\ 0, & \text{коли } X(E) \le 0, \end{cases} \text{ та } X^-(E) = \begin{cases} -X(E), & \text{коли } X(E) < 0, \\ 0, & \text{коли } X(E) \ge 0. \end{cases}$$

Кожна з цих функцій невід'ємна і є випадковою величиною стосовно (Ω, S, P), коли X є такою величиною.

Справді, нехай a – довільне фіксоване число. Тоді можливі два випадки:

1) $a \le 0$ і тоді $(X^+ < a) = \{E \in \Omega : X^+(E) < a\} = \varnothing \in S$;
2) $a > 0$ і тоді $(X^+ < a) = \{E \in \Omega : X^+(E) = 0\} \bigcup \{E \in \Omega : 0 < X^+(E) < a\} =$
$= \{E \in \Omega : X(E) \le 0\} \bigcup \{E \in \Omega : 0 < X(E) < a\} = \{E : X(E) < a\} \in S$.

Звідси за означенням X^+ є випадковою величиною. Міркування для X^- аналогічні.

Враховуючи, що $X(E) = X^+(E) - X^-(E)$, а $|X(E)| = X^+(E) + X^-(E)$, $E \in \Omega$, приходимо до висновку, що *випадкову величину X доцільно назвати інтегровною за Лебегом на множині* Ω *стосовно міри P, якщо такими є випадкові величини* X^+ *та* X^-. При цьому за означенням $\int_\Omega X(E)P(dE) = \int_\Omega X^+(E)P(dE) - \int_\Omega X^-(E)P(dE)$ називають *математичним сподіванням* або *середнім значенням випадкової величини X* і позначають $M[X]$. В іншому разі *випадкову величину називають неінтегровною за Лебегом на множині* Ω *стосовно міри P і вона не має математичного сподівання*.

Таким чином, для довільної інтегровної випадкової величини X обґрунтовано рівності (6), які для простих випадкових величин перетворюються у рівності (4), для дискретних – у рівності (5), а для абсолютно неперервних – у рівність (8).

5. Математичне сподівання складеної функції випадкового аргументу. Однією з операцій, за допомогою яких утворюються різноманітні класи функцій (включаючи і випадкові величини), є операція суперпозиції функцій, тобто утворення складених функцій.

Ознайомлення із суперпозиціями випадкових величин доцільно розпочати з постановки питання: «*Коли є випадковою величиною складена функція* $Y(E) = \psi(X(E))$, $E \in \Omega$, *за умови, що X є випадковою величиною стосовно простору* (Ω, S, P), *і як за відомим розподілом ймовірностей* P_X *на множині* Ω_X *значень випадкової величини X знайти розподіл ймовірностей* P_Y *на множині* Ω_Y *значень випадкової величини Y*?».

Щоб знайти відповідь на поставлене питання, потрібно для довільного числа a знайти прообраз $Y^{-1}((-\infty, a))$ множини $(-\infty; a)$, тобто множину $\{E : Y(E) < a\} = Y^{-1}((-\infty; a))$ розв'язків нерівності $Y(E) < a$. Для цього можна спочатку розв'язати нерівність $\psi(x) < a$ і дістати множину $\psi^{-1}((-\infty; a)) = \{x \in (-\infty; +\infty) : \psi(x) < a\}$, а потім знайти множину $\{E \in \Omega : X(E) \in \psi^{-1}((-\infty; a))\} = X^{-1}(\psi^{-1}((-\infty; a))) = Y^{-1}((-\infty; a))$.

З останньої рівності і з властивостей випадкових величин випливає, що $Y^{-1}((-\infty; a)) = X^{-1}(\psi^{-1}((-\infty; a))) \in S$, коли $\psi^{-1}((-\infty; a))$ є борелівською множиною для будь-якого дійсного числа a. У цьому випадку *функцію* ψ також називають *борелівською* і клас таких функцій є досить широким (зокрема, він містить кусково сталі, монотонні, неперервні функції та багато інших).

Отже, складена функція $Y(E) = \psi(X(E))$, $E \in \Omega$, напевно буде випадковою величиною стосовно простору (Ω, S, P), коли $X(E)$, $E \in \Omega$, є випадковою величиною стосовно цього простору, а $\psi(x)$, $x \in (-\infty; +\infty)$, є борелівською функцією. За цих умов маємо, що ймовірність P_Y на борелівському просторі подій задається рівністю $P_Y((-\infty; y)) = P_X(Y^{-1}((-\infty; y))) =$

$= P(X^{-1}(\psi^{-1}((-\infty; y))))$, $y \in (-\infty; +\infty)$. Тому функція F_Y розподілу ймовірностей P_Y визначатиметься рівністю

$$F_Y(y) = P_Y((-\infty; y)) = P_X(Y^{-1}((-\infty; y))) = P(X^{-1}(\psi^{-1}((-\infty; y)))),\ y \in (-\infty; +\infty).$$

В залежності від того, чи є ця функція дискретною, неперервною чи абсолютно неперервною, матимемо і відповідний розподіл ймовірностей P_Y на множині значень випадкової величини *Y*.

З наведеного видно, як можна замість формального дедуктивного подання навчального матеріалу застосувати проблемне навчання: сформулювати проблему і знайти її вирішення (бажано все це робити за активною участю студентів).

У зв'язку з розглядуваною проблемою можна сформулювати низку питань, відповіді на які студенти можуть знайти самостійно:

- Чи є умова «функція ψ – борелівська» необхідною для того, щоб складена функція $Y(E) = \psi(X(E))$, $E \in \Omega$, була випадковою величиною стосовно простору (Ω, S, P)? А навпаки?
- Чи обов'язково $Y(E) = \psi(X(E))$ є дискретною (простою) випадковою величиною, коли *X*(*E*) є такою величиною, а ψ довільна функція? А навпаки?
- Чи обов'язково $Y(E) = \psi(X(E))$ є неперервною (абсолютно неперервною) випадковою величиною, коли *X*(*E*) є такою величиною, а ψ – довільна борелівська функція? А навпаки?

Підкреслимо, що розв'язування задач, пов'язаних з питанням, коли складена функція $Y(E) = \psi(X(E))$ є випадковою величиною, є особливо корисними для майбутніх учителів математики, оскільки допомагає їм ширше та глибше поглянути на задачі розв'язування рівнянь і нерівностей, які у середній школі відіграють важливу роль. Розглянемо кілька таких задач.

6. Деякі зв'язки стохастики та інших розділів шкільного курсу математики. У пункті 4.3 обґрунтовано рівності (6) для обчислення математичного сподівання $M[X]$ довільної інтегровної випадкової величини *X*(*E*), $E \in \Omega$, стосовно заданого ймовірнісного простору (Ω, S, P).

Застосувавши ці рівності до складеної випадкової величини $Y(E) = \psi(X(E))$, $E \in \Omega$, дістанемо:

$$M[Y] = \int_{\Omega} \psi(X(E)) P(dE) = \int_{\Omega_Y} y P_Y(dy) = \int_{\Omega_X} \psi(x) P_X(\psi^{-1}(dy)) = \int_{-\infty}^{+\infty} \psi(x) dF_X(x) \qquad (9)$$

за умови, що функція $x = X(E)$, $E \in \Omega$, є випадковою величиною стосовно (Ω, S, P), а $\psi(x)$ – обмеженою борелівською функцією і тому інтегровною випадковою величиною стосовно (Ω_X, S_X, P_X).

Зокрема, якщо функція $F_X(x)$ (а отже і випадкова величина *X*) є абсолютно неперервною з щільністю $f_X(x)$ розподілу ймовірностей на множині Ω_X значень випадкової величини *X*, з рівності (9) і з означення $M[X]$ дістаємо рівність

$$M[Y] = \int_{-\infty}^{+\infty} \psi(x) f_X(x) dx. \qquad (10)$$

У випадку, коли ймовірність P_X розподілена рівномірно на деякій множині $A \subset \Omega_X$, маємо:

$$f_X(x) = \begin{cases} \dfrac{1}{l(A)}, & \text{коли } x \in A, \\ 0, & \text{коли } x \notin A, \end{cases}$$

де *l*(*A*) – міра Лебега (довжина) множини $A \subset \Omega_X$.

Тоді рівність (10) набуває вигляду

$$M[Y] = \frac{1}{l(A)} \int_A \psi(x) dx, \qquad (11)$$

а, якщо $\psi(x) \ge 0$, $x \in A$, то:

$$M[Y] = \frac{m(\Phi)}{l(A)} \text{ або } m(\Phi) = M[Y] \cdot l(A), \qquad (12)$$

де *m*(*Φ*) – міра Лебега (площа) узагальненої криволінійної трапеції *Φ*, що обмежена знизу множиною *A* на осі абсцис, а зверху – графіком функції $y = \psi(x)$, $x \in A$.

Рівності (11) і (12) розкривають деякі зв'язки стохастики та шкільного курсу математики.

6.1. Математичне сподівання та площа трикутника. Припустимо, що графік функції $y = \psi(x)$, $x \in [0; a]$, співпадає з двома сторонами трикутника, а третьою стороною цього трикутника є відрізок $[0;a]$ на осі абсцис, причому висота трикутника дорівнює h (Рис. 2). Вважатимемо також, що відрізок $[0;a]=\Omega_X$ є множиною значень деякої випадкової величини $X(E)$, $E\in\Omega$, стосовно деякого простору (Ω, S, P), причому ймовірність P_X рівномірно розподілено на відрізку $[0;a]$.

За вказаних умов можна сформулювати низку завдань, професійно значущих для вчителів математики:

- Перевірити, чи задають однозначно наведені умови аналітичний вираз функції $\psi(x)$.
- З'ясувати, чи є складена функція $Y(E)=\psi(X(E))$, $E\in\Omega$, випадковою величиною стосовно (Ω, S, P), і якщо є, то якою саме: дискретною, неперервною чи абсолютно неперервною.
- Обчислити різними способами математичне сподівання $M[Y]$ випадкової величини $Y(E)=\psi(X(E))$ і вказати найпростіший із цих способів.
- З'ясувати, чи залежить $M[Y]$ від заданих a, h і аналітичного виразу функції ψ.

6.1.1. Аналітичний вигляд функції $\psi(x)$ легко знайти за допомогою рисунка 2.

Використовуючи загальне рівняння прямої, що проходить через дану точку із заданим кутовим коефіцієнтом, маємо

$$\psi(x)=\begin{cases}\dfrac{h}{b}x, & \text{коли } 0\le x\le b,\\ -\dfrac{h(x-a)}{a-b}, & \text{коли } b<x\le a.\end{cases} \tag{13}$$

За даних умов про число b можна лише стверджувати, що $b\in(0;a)$. Тому *наведені умови не задають однозначно аналітичний вираз функції* ψ.

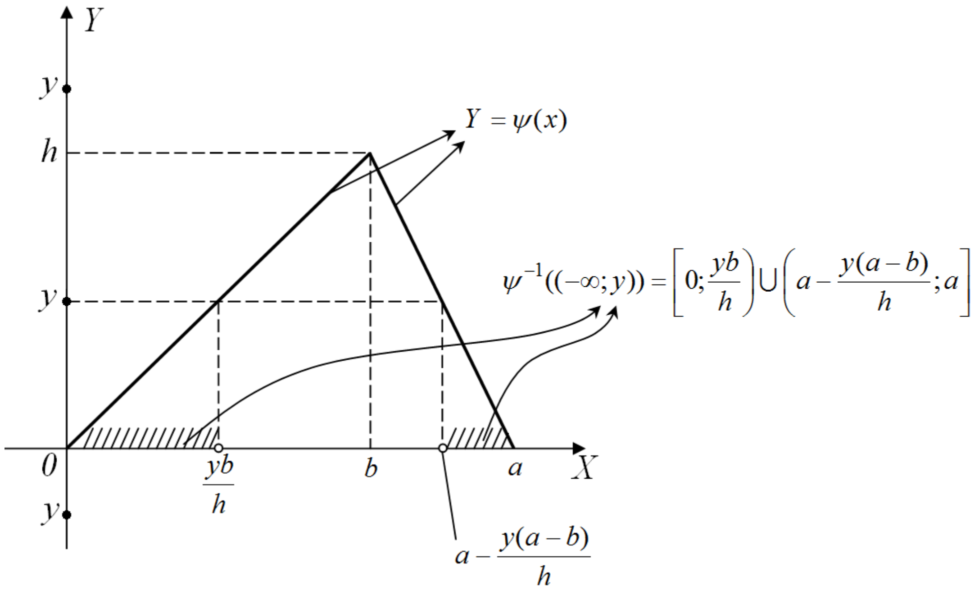

Рис. 2

6.1.2. Оскільки функція ψ неперервна на відрізку $[0;a]$, то вона є борелівською, а тому *складена функція* $Y(E)=\psi(X(E))$, $E\in\Omega$, *є випадковою величиною стосовно* (Ω, S, P), *коли такою є функція X(E),* $E\in\Omega$. Тому ймовірність P_Y, розподілена на множині Ω_Y значень випадкової величини $Y(E)=\psi(X(E))$, $E\in\Omega$, визначається за рівністю

$$P_Y((-\infty;y))=P_X(\psi^{-1}((-\infty;y)))=\int\limits_{\psi^{-1}((-\infty;y))} f_X(x)dx,\ \ y\in(-\infty;+\infty).$$

Оскільки за умовою ймовірність P_X рівномірно розподілена на відрізку $[0;a]$, то щільність такого розподілу має вигляд

$$f_X(x)=\begin{cases}0, & \text{коли } x\le 0 \text{ або } x\ge a,\\ \dfrac{1}{a}, & \text{коли } x\in(0;a).\end{cases} \tag{14}$$

За допомогою рис. 2 для різних $y\in(-\infty,+\infty)$ знаходимо $\psi^{-1}((-\infty;y))=\{x\in[0;a]:\psi(x)<y\}$:

- якщо $y\le 0$, то $\psi^{-1}((-\infty;y))=\varnothing\in S_X$ і тому $P_X(\psi^{-1}((-\infty;y)))=P_X(\varnothing)=0$;

- якщо $0 < y \le h$, то $\psi^{-1}((-\infty; y)) = \left[0; \frac{yb}{h}\right) \cup \left(a - \frac{y}{h}(a-b); a\right]$ і тому $P_X(\psi^{-1}((-\infty; y))) =$
  $= \frac{1}{a}\left(\frac{yb}{h} - 0 + a - a + \frac{y}{h}(a-b)\right) = \frac{y}{h}$;
- якщо $y > h$, то $\psi^{-1}((-\infty; y)) = [0; a]$ і тому $P_X(\psi^{-1}((-\infty; y))) = 1$.

Враховуючи знайдене, маємо

$$F_Y(y) = P_Y((-\infty; y)) = P_X(\psi^{-1}((-\infty; y))) = \begin{cases} 0, \text{ коли } y \le 0, \\ \dfrac{y}{h}, \text{ коли } 0 < y \le h, \\ 1, \text{ коли } y > h. \end{cases}$$

Тепер легко переконатися, що функція $F_Y(y)$ є абсолютно неперервною функцією розподілу ймовірностей P_Y на множині значень випадкової величини Y з щільністю

$$f_Y(y) = \begin{cases} 0, \text{ коли } y \le 0 \text{ або } y \ge h, \\ \dfrac{1}{h}, \text{ коли } 0 < y < h. \end{cases} \tag{15}$$

Отже, складена випадкова величина $Y(E) = \psi(X(E))$, $E \in \Omega$, є абсолютно неперервною, причому ймовірності P_Y рівномірно розподілені на відрізку $[0; h]$ з щільністю (15), яка залежить лише від h і не залежить ані від a, ані від b, тобто не залежить від аналітичного виразу функції ψ.

6.1.3. Згідно з рівностями (9) та (12), (14) і (15), математичне сподівання $M[Y]$ можна обчислити трьома способами.

Перший спосіб: $M[Y] = \int\limits_{\Omega_Y} y P_Y(dy) = \int\limits_0^h y dF_Y(y) = \int\limits_0^h y f_Y(y) dy = \int\limits_0^h y \cdot \frac{1}{h} dy = \left.\frac{y^2}{2h}\right|_0^h = \frac{h}{2}$.

Другий спосіб: $M[Y] = \int\limits_{\Omega_X} \psi(x) P_X(\psi^{-1}(dy)) = \int\limits_0^a \psi(x) dF_X(x) = \int\limits_0^a \psi(x) f_X(x) dx =$

$$= \int\limits_0^b \frac{h}{b} x \cdot \frac{1}{a} dx + \int\limits_b^a -\frac{h(x-a)}{a-b} \cdot \frac{1}{a} dx = \frac{h}{ab} \left.\frac{x^2}{2}\right|_0^b - \frac{h}{a(a-b)} \left.\frac{(x-a)^2}{2}\right|_b^a = \frac{hb}{2a} + \frac{h(a-b)}{2a} = \frac{h}{2}.$$

Третій спосіб: $M[Y] = \frac{m(\Phi)}{l(A)}$, де $m(\Phi)$ – це площа трикутника, основою якого є відрізок $A = [0; a]$ з довжиною $l(A)=a$ та з висотою h. Тому

$$M[Y] = \frac{\frac{1}{2} ah}{a} = \frac{h}{2}.$$

6.1.4. Оскільки $M[Y] = \frac{h}{2}$, то це математичне сподівання не залежить ані від a, ані від аналітичного вигляду виразу $\psi(x)$ (аби тільки графік цієї функції утворював разом з відрізком $[0; a]$ трикутник з висотою h і вершина трикутника проектувалася на його основу).

Аналогічно можна дослідити зв'язки математичного сподівання функцій від випадкових аргументів з площами трапецій, півкругів, півеліпсів та багатьох інших фігур, які можна розглядати як криволінійні трапеції, об'ємами конусів, пірамід, в тому числі зрізаних, еліпсоїдів, параболоїдів і ін.

7. Висновки. Міжпредметні зв'язки притаманні будь-якому навчальному курсу, включаючи і теорію ймовірностей, а використання цих зв'язків слід розуміти як акцентування на них уваги з метою підвищення мотивації навчання і рівня опанування навчальним матеріалом. У процесі навчання майбутніх учителів теорії ймовірностей (і будь-якої іншої математичної дисципліни) важливо використовувати міжпредметні зв'язки не тільки з іншими математичними дисциплінами, а й зі шкільним курсом математики, з методикою навчання математики. Саме у процесі навчання суто математичних дисциплін можна розкрити майбутнім учителям математики справжнє призначення методики навчання математики: знаходити такі методи подання навчального матеріалу, щоб навчати математики на сучасному і водночас доступному для учнів рівні. Завдяки сучасному рівню розвитку математики і методики її навчання, а також сучасних комп'ютеризованих інформаційних технологій навчання задачі, розв'язування яких раніше вимагало значних знань і зусиль, зараз можуть успішно розв'язувати учні середньої школи.

Література

1. Феллер В. Введення в теорию вероятностей и ее приложения. Том 1. – М.: Мир, 1984, 528 с.
2. Жалдак М.І., Михалін Г.О. Елементарні факти теорії множин у шкільному курсі математики // Математика в школі. – 2011, №3. – С. 12-23.
3. Алгебра: 7 клас / Ю.Н. Макарычев и др. / Под ред. С.А. Теляковского. – М.: Просвещение, 1993. – 240 с.
4. Жалдак М.І., Михалін Г.О., Деканов С.Я. Математичний аналіз. Функції багатьох змінних. – К.: НПУ імені М.П. Драгоманова, 2007. – 430 с.