# Лекція

# Тема 3. ВИПАДКОВІ ВЕЛИЧИНИ

**І ЗАКОНИ РОЗПОДІЛУ ЯК ОСНОВА БІЗНЕС-АНАЛІТИКИ**

* 1. Випадкові величини в бізнес-аналізі
	2. Закони розподілу випадкових величин: сутність і види
	3. [Нормальний розподіл](#_TOC_250035)
	4. [Біноміальний розподіл](#_TOC_250034)
	5. [Апроксимація біноміального розподілу нормальним](#_TOC_250033)

[Питання для самоконтролю](#_TOC_250030)

* 1. **Випадкові величини в бізнес-аналізі Випадковою величиною** називають величину, яка внаслідок

проведення дослідження може прийняти те чи інше значення (лише

одне), причому до проведення дослідження невідомо яким саме воно буде. Випадкові величини позначають літерами *X*, *Y*…, а їх окремі значення відповідно *x, y* ….

У бізнес-аналітиці доволі часто виникають ситуації, які потребують роботи з випадковими величинами, наприклад, під час визначення ефективності портфеля інвестицій або в рамках проведення маркетингового дослідження з метою виявлення ступеня задоволення продукцією підприємства. Кожного разу, коли частиною випадкового експерименту є одне єдине число, ми маємо справу з випадковими величинами або так званими **випадковими змінними**.

Економістам важливо вміти розраховувати та аналізувати узагальнюючі характеристики об’єктів, які є випадковими величинами, а також виявляти ймовірність подій, які залежать від результатів спостережень, наприклад, імовірність того, що 90% клієнтів залишаться задоволеними послугами або продукцією підприємства.

Випадкові величини також можна розглядати у вигляді джерел даних. Багато з тих наборів даних, які ми обговорювали у ***Розділі 2***, були отримані в результаті спостережень та фіксації значень деяких випадкових величин. У такому випадку самі випадкові величини є генеральною сукупністю, у той час як отримані значення випадкових величин є результатом вибірки.

Наведемо декілька *прикладів* випадкових величин. Варто зауважити, що кожна з цих величин є випадковою лише до того моменту, поки не буде зафіксований конкретний результат спостереження.

**Випадок перший.** Підприємство доставило в магазин на продаж товарів загальною вартістю 1 млн. грн. Таким чином, обсяг продажу цих товарів покупцям протягом періоду буде охарактеризовано деяким числом в інтервалі від нуля до мільйону.

**Випадок другий.** За 30 днів підприємством було виготовлено 2500 одиниць продукції типу А. Кількість бракованих одиниць продукції буде знаходитися в межах від 0 до 2500.

**Випадок третій.** Кількість кваліфікованих фахівців, які претендують на певну посаду. Якщо загальна кількість осіб, які подали документи HR-менеджеру, буде дорівнювати 20 особам, то випадкова величина може мати значення від 0 до 20.

**Випадок четвертий.** Кількість викликів, які поступають на телефонну лінію технічного обслуговування підприємства, може варіювати від нуля до певного цілого значення числа.

**Випадок п’ятий.** Заробітна плата працівника підприємства, який буде учасником соціологічного опитування. Якщо рівень заробітної плати на підприємстві варіює від 4 тис. грн. до 18 тис. грн., то і випадкова величина заробітної плати даного працівника буде знаходитися в цьому діапазоні.

Випадкову величину також можна визначити як описання численного результату певного експерименту. Саме значення називають **значенням, що спостерігають*.*** Наприклад, розмір прибутку підприємства в наступному кварталі – це випадкова величина, яка детермінує й описує число, що буде отримане в результаті експерименту, який полягає в тому, щоб дочекатися кінця кварталу та розрахувати отриманий розмір прибутку. Реальне значення, яке буде досягнуто в майбутньому, наприклад 30 млн. грн., є значенням випадкової величини, що спостерігають.

Варто підкреслити, що існує відмінність між випадковою величиною (яка відноситься до будь-якого випадкового процесу) і значенням, що спостерігають (яке є фіксованим числом, що реєструється в ході спостереження).

Правило визначення ймовірностей значень випадкових величин називають **правилом розподілу ймовірностей*.***

Серед випадкових величин можна виділити дискретні та неперервні.

**Дискретною випадковою величиною** називається така величина, кількістю можливих значень якої є або скінчена, або нескінченно злічена множина.

**Неперервною випадковою величиною** називається така величина, можливі значення якої заповнюють певний інтервал (скінчений або нескінченний) числової осі.

У деяких випадках знайти відмінність між дискретною та неперервною випадковими величинами є досить складно. Скажімо, величина прибутку підприємства в наступному кварталі може скласти 320457 грн. або 345004,89 грн., або будь-яке інше число, яке буде меншим установленої межі, наприклад 350000 грн. У суворому розумінні, ця величина прибутку є дискретною величиною, оскільки всі ці значення можна перелічити, однак з практичної точки зору, оскільки інтервал між сусідніми значеннями є дуже малим (1 копійка), то з цією випадковою величиною можна працювати як з неперервною.

Якщо для дискретної випадкової величини є відомим перелік усіх можливих значень із ймовірностями їх отримання, то про відповідний бізнес-процес відомо абсолютно все. Із використанням такого переліку можна розрахувати будь-яку характеристику, що є цікавою для дослідника (наприклад, імовірність утрат у результаті певної виробничої операції та їх можливе значення в гривнях). Розглянемо декілька прикладів дискретних випадкових величин.

1. Обсяг виробництва деталі А на підприємстві протягом наступного тижня. Можливими значеннями будуть 0,1,2,3 ….
2. Кількість працівників, які подали заявку на проводження відпустки у будинку відпочинку підприємства. Лист можливостей у цьому випадку має такі самі значення: 0,1,2,3 ….
3. Розмір бюджету грантової пропозиції на проведення екологічного фестивалю за чотирьох можливих варіантів: 25 тис. дол., 50 тис. дол., 75 тис. дол. та 100 тис. дол. США. Отже, перелік можливих значень буде складатися саме з цих чисел: 25000, 50000, 75000, 100000.

Такий перелік значень разом із ймовірністю їх виникнення є розподілом імовірності для дискретної випадкової величини.

Розглянемо дискретну випадкову величину *Х*, значення якої *x1, x2,…, xn* нам відомі. Знання про можливі значення випадкової величини ще не дають нам повного описання самої випадкової величини, тому що ми не можемо сказати, як часто можуть з’явитися ті чи інші її значення. Для цього необхідно знати закон розподілу ймовірностей випадкової величини.

Внаслідок проведення досліду випадкова величина *Х* набуде одного із своїх можливих значень, тобто відбудеться одна подія із повної групи несумісних подій: *X* = *x1*, *X* = *x2*,…, *X* = *xn*.

Позначимо ймовірності цих подій літерами *p* з відповідними індексами: *P(X=x1) = p1, P(X=x2) = p2,…, P(X=xn) = pn.*

Імовірності повинні приймати значення тільки додатних чисел або нуля, а сума ймовірностей усіх можливих значень випадкової величини *Х* дорівнює одиниці:

*n n*

 *P*( *X*  *x* )   *p*  1

*i*

*i*

(3.1)

Ця сумарна ймовірність якимось чином розподілена між окремими значеннями випадкової величини. Дискретна випадкова величина буде повністю описаною з точки зору ймовірності, якщо буде вказано, яку ймовірність має кожна з подій. Так ми встановимо закон розподілу випадкової величини.

Заданий таким чином розподіл дозволяє знайти середнє значення, середньоквадратичне відхилення та ймовірність будь-якої події, що відповідає цій випадковій величині.

Розглянемо це за допомогою прикладів. Припустимо, підприємство вирішило розширити асортимент виробництва продукції і готує в наступному кварталі випустити на ринок новий інноваційний товар, аналогів якому на даний час майже не має. Залежно від того, як буде сприйнятий товар потенційними покупцями, маркетологи підприємства розробили декілька сценаріїв фінансового результату наприкінці періоду (табл. 3.1).

*Таблиця 3.1*

## Можливі сценарії фінансового результату підприємства

|  |  |  |
| --- | --- | --- |
| Сценарій | Фінансовий результат,млн. грн. | Імовірність |
| Дуже гарний | 20,0 | 0,200 |
| Добрий | 10,0 | 0,400 |
| Нормальний | 5,0 | 0,160 |
| Поганий | -1,0 | 0,140 |
| Дуже поганий | -5,0 | 0,100 |

Цей розподіл імовірностей дуже легко використовувати для розрахунку ймовірностей усіх подій, пов’язаних з отриманням фінансового результату. Наприклад, імовірність отримання фінансового результату в розмірі 20,0 млн. грн. складає 0,200 або 20,0%. Імовірність отримання фінансового результату в розмірі

10,0 млн. грн. і більше можна визначити за допомогою підсумовування ймовірностей: 0,200+0,400 = 0,600.

**Середнє** або **очікуване значення (математичне сподівання)** дискретної випадкової величини – це певне число, що буде характеризувати типове значення цієї величини, подібно тому, як певний набір даних характеризується середнім значенням. Середнє значення для випадкової величини позначається як  або *Е(Х)* і розраховується за наступною формулою:

  *Е*( *Х* )   *ХР*(*х*)

Таким чином, середнє або очікуване значення випадкової величини визначається як сума значень помножених на їх імовірності.

Якщо б імовірності отримання різних варіантів фінансового результату були однаковими, ми б отримали просте осереднення існуючих значень. Однак, оскільки ймовірності різні, ми повинні використовувати зважування всіх значень, у якому в якості вагів виступають відповідні ймовірності.

Для розглянутого вище прикладу середній очікуваний фінансовий результат буде розрахований наступним чином: *Фінансовий результат = 20,0\*0,200 + 10,0\*0,400 + 5,0\*0,160 + (-1,0)\*0,140 + (-5,0)\*0,100 = 8,160 млн. грн.*

Отже, очікуваний фінансовий результат складе 8,16 млн. грн. Це значення характеризує всі можливі результати (20,0; 10,0; 5,0; -1,0; - 5,0) одним числом, яке враховує ймовірність виникнення кожного з них.

**Середньоквадратичне відхилення** дискретної випадкової величини приблизно вказує на скільки реальні значення цієї випадкової величини можуть відрізнятися від середньої. У багатьох випадках господарської діяльності підприємств цей показник характеризує ризик, що вказує якою невизначеною є ситуація. Формула розрахунку середньоквадратичного відхилення наведена нижче: 

( *Х*   )2 *Р*(*х*)

Розрахуємо середньоквадратичне відхилення для нашого прикладу (табл. 3.2).

У цьому випадку середньоквадратичне відхилення буде

дорівнювати наступній величині:

  7,749 *млн*. *грн*.,

60,054

що свідчить про значний рівень ризику випуску нового виду продукції. Очікуваний фінансовий результат може виявитися вищим або нижчим за 8,160 млн. грн. на 7,749 млн. грн.

*Таблиця 3.2*

## Можливі сценарії фінансового результату підприємства

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Сценарій | Фінансовий результат, млн. грн.*Х* | Імовірність*Р(х)* | *Х*   | ( *Х*  )2 | ( *Х*  )2 *Р*(*х*) |
| Дуже гарний | 20,0 | 0,200 | 11,840 | 140,186 | 28,037 |
| Добрий | 10,0 | 0,400 | 1,840 | 3,386 | 1,354 |
| Нормальний | 5,0 | 0,160 | -3,160 | 9,986 | 1,598 |
| Поганий | -1,0 | 0,140 | -9,160 | 83,906 | 11,747 |
| Дуже поганий | -5,0 | 0,100 | -13,160 | 173,186 | 17,319 |
| Усього | х | 1,000 | х | х | 60,054 |

Аналіз випадкових величин допоможе і в інших ситуаціях, як то в інвестиційному аналізі. Розглянемо це на наступному прикладі.

Допустимо, підприємство має у своєму розпорядженні 100,0 тис. у.о, які воно бажає інвестувати у певний проект.

Менеджери підприємства повинні обрати одну з трьох проектних пропозицій, яка буде видаватися найбільш привабливою як з точки зору суми повернених коштів, так і з точки зору ймовірності їх отримання (ризику). Вихідні дані проектів наведені в табл. 3.3.

*Таблиця 3.3*

## Інвестиційні проекти підприємства

|  |  |  |
| --- | --- | --- |
| Проект | Сума повернених коштів, млн. грн. | Імовірність |
| A | 110,0 | 1,000 |
| B | 120,0 | 0,500 |
| 130,0 | 0,500 |
| С | 250000,0 | 0,001 |
| 0,0 | 0,990 |

Знайдемо середні значення суми повернених коштів для кожного з проектів за допомогою формули 3.2. Так, для *проекту А* це буде 110,0 тис. у.о, для *проекту В* – 125,0 *(120,0\*0,500 + 130,0\*0,500)* тис. у.о, для *проекту С*– 250,0*(250000,0\*0,001 + 0,0\*0,990)* тис. у.о.

 *(продовження нижче)*

Отже:

*Е*( *Х* )  *А*  110,0 *тис*. *у*. *о*.

*Е*( *Х* )  *В*  125,0 *тис*. *у*. *о*. *Е*( *Х* )  *С*  250,0 *тис*. *у*. *о*.

Якщо аналізувати лише розміри середніх величин, то можна зробити висновок, що найбільш привабливим для підприємства є *проект С*, однак він є найбільш ризикованим, адже ймовірність отримання коштів складає лише 1 %. Продовжуючи наш аналіз розрахуємо середньоквадратичне відхилення для аналізу ризиків кожного з проектів:

(110,0 110,0)2 \*1,000

 *А* 

 0,0 *тис*. *у*. *о*.

*В* 

*С* 

 5,0 *тис*. *у*. *о*.

 7897,8 *тис*. *у*. *о*.

(120,0 125,0)2 \* 0,500  (130,0 125,0)2 \* 0,500

(250000,0  250,0)2 \* 0,001 (0,0  250,0)2 \* 0,990

Розрахунки середньоквадратичного відхилення підтверджують, що *проект С* є найбільш ризикованим серед інших проектів, адже у випадку його реалізації ризик дорівнює 7997,8 тис. у.о. У той же час неможливо однозначно відповісти, який проект буде більш привабливим для підприємства, оскільки ставлення до ризику у підприємців буває різним. Деякі з них оберуть можливість отримання меншого прибутку з більшою ймовірністю, інші – більші прибутки за меншої ймовірності їх отримання.

## Закони розподілу випадкових величин:

**сутність і види**

**Законом розподілу** випадкової величини називають будь-яке співвідношення, що встановлює зв'язок між можливими значеннями випадкової величини та відповідними ймовірностями.

Найпростішою формою задання закону розподілу випадкової величини *X* є таблиця, в якій наведені можливі значення випадкової величини з відповідними ймовірностями (табл. 3.4).

Така таблиця також носить назву *ряду розподілу випадкової величини Х.* ЇЇ можна представити й у графічному вигляді. У цьому випадку можливі значення випадкової величини відкладають по осі абсцис, а по осі ординат – відповідні ймовірності. Вершини з’єднують відрізками прямих, а таку фігуру називають

багатокутником розподілу. Сума ординат багатокутника розподілу завжди дорівнює *одиниці*.

*Таблиця 3.4*

## Форма завдання закону розподілу

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *хі* | *х1* | *х2* | *…* | *хn* |
| *pi* | *p1* | *p2* | *…* | *pn* |

Розглянутий ряд розподілу є дуже зручною формою представлення закону розподілу для дискретної випадкової величини із скінченим числом можливих значень. Для неперервної випадкової величини ряд розподілу взагалі не можна побудувати. Тобто необхідно мати такий розподіл імовірності, щоб її можна було застосовувати для найрізноманітніших випадкових величин.

Найзагальнішою формою закону розподілу випадкової величини *Х* є так звана ***функція розподілу***. Кожен закон розподілу – це деяка функція, що повністю описує випадкову величину з імовірнісної точки зору.

**Функцією розподілу**, або **інтегральним законом розподілу** випадкової величини *Х* називається задання ймовірності виконання нерівності *X < x*, що розглядається як функція від аргументу *х*:

*F(x) = P(X<x)* (3.4)

Функція розподілу повністю характеризує випадкову величину з точки зору ймовірності, тому вона є однією з форм закону розподілу.

Для дискретної випадкової величини функція розподілу матиме вигляд:

*F* (*x*)   *P*( *X*

*xi*  *x*

 *xi* )

(3.5)

З цього виразу випливає, що функція розподілу дискретної випадкової величини стрибкоподібно зростає.

Функція розподілу неперервної випадкової величини має один недолік – за нею досить важко робити висновки про характер розподілу випадкової величини у невеликому колі тієї чи іншої точки числової осі. Для цього використовують функцію, що носить назву ***щільності розподілу ймовірності*** або ***диференціального закону розподілу випадкової величини****.*

Крива, що зображує щільність розподілу *f(x)* випадкової величини, називається **кривою розподілу**.

Випадкова величина *Х* називається **неперервною**, якщо її функція розподілу *F(x)* неперервна на всій осі *Ох*, а щільність розподілу *f(x)* існує скрізь, окрім, можливо, скінченої кількості точок. На практиці висновки щодо розподілу ймовірностей випадкової величини *Х* часто доводиться робити лише за результатами випробувань. Повторюючи випробування, ми повинні кожного разу

реєструвати, чи стала нам цікавою подія *А*, чи ні.

**Відносною частотою** (або **часткою**) випадкової події *А* називається відношення числа *nA* появ цієї події до загального числа *n* проведених випробувань. При цьому ми приймаємо, що значення відносних частот випадкових подій близькі до їх можливостей. Це є тим більш вірним, чим більшим є число проведених досліджень. При цьому частоти, як і ймовірності, необхідно відносити не до окремих значень випадкової величини, а до інтервалів значень. Це означає, що весь діапазон можливих значень випадкової величини *Х* потрібно розбити на інтервали. Проводячи серії випробувань, які дають емпіричні значення величини *Х*, потрібно фіксувати числа *nx* влучень результатів у кожен інтервал. При великій кількості випробувань *n* співвідношення *nx/n* (частоти потрапляння в інтервали) повинні бути близькими до можливості попадання в ці інтервали. Залежність частот *nx/n* від інтервалів визначає емпіричний розподіл імовірностей випадкової величини *Х*, графічне представлення якої називається **гістограмою**.

Закон розподілу повністю характеризує випадкову величину з точки зору ймовірності. Однак, при вирішенні багатьох практичних задач немає необхідності характеризувати випадкову величину повністю, а достатньо мати тільки деяке загальне уявлення про неї. У теорії ймовірностей для характеристики випадкової величини використовують величини, які носять назву ***числових характеристик випадкової величини****.* Їх призначення – у стислій формі показати найбільш суттєві особливості того чи іншого розподілу.

Про кожну випадкову величину необхідно знати, перш за все, її середнє значення, біля якого групуються можливі значення випадкової величини, а також число, що характеризує ступінь розкиду цих значень навколо середнього. Крім цих характеристик використовують й інші, як то: *дисперсія, коефіцієнт асиметрії, коефіцієнт ексцесу, мода, медіана та моменти розподілу*.

Основними законами розподілу випадкових величин є:

1. Нормальний розподіл.
2. Біноміальний розподіл.
3. Розподіл Пуасона.
4. Експоненціальний розподіл.
5. Рівномірний розподіл.
6. Геометричний розподіл.
7. Логнормальний розподіл.
8. Розподіл Вейбула тощо.

Розглянемо ці види розподілу більш детально.

## Нормальний розподіл

***Нормальний закон розподілу*** або закон Гауса, який ми частково розглянули в другому розділі, грає виключно важливу роль у бізнес- статистиці та займає серед інших законів розподілу особливе місце. Головна характеристика, що виділяє нормальний закон серед інших, полягає в тому, що він є граничним законом, до якого наближаються інші закони розподілу при типових умовах.

**Нормальний розподіл** – це непереривний розподіл, який має графічне зображення у вигляді симетричної дзвіноподібної кривої, наведеної на рис. 3.1.

### Рис. 3.1. Крива Гауса. Нормальний розподіл випадкової величини

Нормальний закон розподілу характеризується щільністю ймовірності у вигляді:

*Р*(*x*) 

( *x* )2

*e* 2

2

1 2

, (3.6)

де π - співвідношення довжини кола та його діаметра, що дорівнює приблизно 3,142;

*е* – основа натурального логарифму, що дорівнює приблизно 2,718;

*Р* – імовірність;

µ – математичне сподівання;

 2 – дисперсія;

*х* – випадкова величина.

Отже, конкретна форма нормального розподілу залежить від 2-х параметрів: математичного сподівання та дисперсії. Параметр µ визначає центр розподілу, якому відповідає максимальна висота

графіка. Дисперсія  2 характеризує розмах варіації, тобто

«розмазаність» даних. Таким чином, математичне сподівання зміщує центр розподілу вправо або вліво, не впливаючи на саму форму кривої щільності; дисперсія визначає загостреність кривої. Коли дані мають малий розкид, то вся їх маса сконцентрована в центрі. Якщо ж у даних великий розкид, то вони будуть розкидані по широкому діапазону.

Дзвіноподібна крива дозволяє більш наочно представити ймовірність для випадку нормального розподілу. Більш вірогідними видаються спостереження значень, які розташовані поблизу центру кривої, біля її самої високої точки (рис. 3.2, зелена область). Поблизу країв кривої, де вона йде нижче, спостереження відповідних значень стають менш правдоподібними (рис. 3.2, синя область). Імовірність того, що значення попаде в деякий інтервал на числовій прямій, дорівнює площині відповідної області під кривою, як це зображено на рис. 3.2.

**Стандартний нормальний розподіл** – це нормальний розподіл із значенням математичного сподівання µ = 0 та середньо-

квадратичним відхиленням   1 . Для визначення випадкової

величини вчені зазвичай використовують літеру *Z*.

Один із методів розрахунку ймовірності для нормального розподілу полягає у використанні таблиць імовірності стандартного розподілу (*Додаток А*), оскільки взагалі неможливо розробити таблиці для кожної комбінації середньоквадратичного відхилення та

математичного сподівання. За допомогою стандартного нормального розподілу можна подати будь-який нормальний розподіл у тому випадку, коли аналізовані величини виступають не в реальних одиницях вимірювання (наприклад гривнях, метрах, кілограмах тощо), а в кількості середньоквадратичних відхилень у більший або менший бік від середнього значення.



### Рис. 3.2. Імовірність того, що випадкова величина приймає значення, які лежать у певному інтервалі

*Додаток А* ймовірностей для стандартного нормального розподілу містить у собі ймовірність того, що випадкова величина *Z*, яка має стандартне нормальне відхилення, приймає значення менше деякого заданого числа *z*. Так, імовірність того, що величина *Z* менша 1,51, складає 0,9345.

Розглянемо приклад вирішення задач на визначення ймовірності за умови нормального розподілу. Припустимо, фінансовий відділ підприємства спрогнозував прибуток у третьому кварталі поточного року в розмірі 21 000 000 грн. Фактичні значення були дещо більшими і наприкінці періоду підприємство отримало прибуток у розмірі 23 300 000 грн. Прогнозований розмір прибутку на четвертий квартал складає 22 000 000 грн., середньоквадратичне відхилення, визначене з минулого досвіду, дорівнює 2 000 000 грн. З урахуванням того, що прибуток підприємства відповідає закону нормального розподілу, необхідно визначити ймовірність того, що наступний

квартал буде охарактеризований спадом ділової активності, а прибуток підприємства буде менший за 17 000 000 грн.

Проаналізувавши задану інформацію, ми можемо зробити висновок, що прибуток у розмірі 23 300 000 грн. і прогнозоване значення на третій квартал 21 000 000 грн. описують минулі події, а отже не мають ніякого впливу на вирішення поточної задачі.

У той же час варто приділити увагу наступним фактам:

* у розглянутій задачі присутній нормальний розподіл, отже його математичне сподівання згідно умови складає µ = 22 000 000 грн.;
* стандартне відхилення складає  = 2 000 000 грн.

Для того, щоб вирішити дану задачу нам необхідно провести нормування цих значень (крім математичного сподівання та середньоквадратичного відхилення), що дозволить використовувати для пошуку відповіді таблицю ймовірностей нормального розподілу із додатку *А*. Нормальне значення (*Z*) являє собою кількість середньоквадратичних відхилень від середнього значення в більший чи менший бік. Процес нормування значень можна виконати за допомогою спеціальної функції програми STATISTICS або з використанням наступної формули:

*z*  *Нормоване*

*значення*   *значення*  



(3.7)

У нашому прикладі розрахунок буде мати наступний вигляд:

*z*  17000000 22000000  2,50

2000000

Це означає, що в даному випадку величина 17 000 000 грн. відповідає величині *z* = - 2,50. Таким чином, ми трансформували вихідну задачу в наступну задачу визначення ймовірності для стандартного нормального розподілу. Згідно з даними додатку *А*, ймовірність того, що квартал буде охарактеризований спадом ділової активності, складає 0,0062 або 0,62%. Тобто негативний результат для діяльності компанії є майже неймовірним, що обіцяє гарні перспективи для подальшої її діяльності.

Крім наведеного прикладу, існують ще три інших варіанти вирішення задач на пошук імовірностей. Тобто, в цілому ці задачі можна розділити на чотири типи, кожному з яких буде властивий певний спосіб вирішення (табл. 3.5). Значення *z, z1, z2* – нормовані значення з постановок задач.

*Таблиця 3.5*

## Варіанти вирішення задач на знаходження ймовірностей

|  |  |
| --- | --- |
| Умови завдання | Спосіб вирішення |
| Необхідно знайти ймовірність того, що величина Z: | знайти в таблиці ймовірностей ймовірність для z |
| є меншою за z |
| є більшою за z | відняти від 1 попередній результат |
| знаходиться між z1 та z2 | знайти в таблиці ймовірності для z1 та z2, а потім відняти від більшої меншу |
| знаходиться за межами інтервалу між z1 та z2 | відняти від 1 попередній результат |

## Біноміальний розподіл

**Біноміальний закон розподілу** описує ймовірність настання події *А m* раз в *n* незалежних випробуваннях, за умови, що ймовірність *p* настання події *А* в кожному випробуванні постійна.

Наприклад, у бізнес-статистиці достатньо часто використовують при розрахунках відсотки. Якщо кількість настання події виражається як відсоток до загальної кількості можливостей, то кількість разів настання події повинно мати біноміальний розподіл. У якості прикладу величин, які мають біноміальний розподіл, можна навести наступні:

1. Кількість бракованих одиниць продукції з партії товару в 100 одиниць.
2. Чисельність чоловіків у колективі відділу маркетингу з 20

осіб.

1. Кількість голосів «за» (або «проти») під час захисту

дисертаційної роботи.

1. Чисельність осіб серед опитаних 200, які виказали бажання поїхати на екскурсію до музею.
2. Кількість осіб із тест-групи, які виказали бажання придбати продукцію, що вони спробували.

Графічно розподіл випадкових величин, що підкоряється біноміальному закону, можна подати в наступному вигляді (рис. 3.3).



### Рис. 3.3. Схематичне зображені біноміального розподілу

Розглянемо деяку конкретну подію, яка під час експерименту трапляється або не трапляється. Наявність тільки двох варіантів результату пояснює приставку «бі» у назві розглянутого закону. Випадкова величина *Х*, яка являє собою число настання певної події в результаті *n* випробувань, має біноміальний розподіл тільки в наступних випадках:

* якщо всі спроби під час експерименту є незалежними одна від одної; ця вимога означає неможливість побачити результат заздалегідь. Наприклад, під час голосування людина голосує «за», оскільки бачить, що більшість інших осіб голосує так само, а вона не хоче виділятися. Для того, щоб розподіл був біноміальним, вибір повинен бути незалежним від будь-яких факторів;
* якщо в кожної з *n* спроб імовірність настання події Ω одна й таж сама.

Біноміальна пропорція *Р* – це подання випадкової змінної *Х*, що має біноміальний розподіл, у вигляді частки кількості спроб *n*:

*P*  *Х*

*п*

 *Кількість випадків настання події*

*Кількість спроб*

(3.8)

Необхідно звернути увагу на те, що *Ω* – це фіксована величина, яка визначає ймовірність настання події, а *P* – випадкова величина, що залежить від даних, які спостерігають.

Розглянемо біноміальний закон на конкретному прикладі. Припустимо, маркетинговий відділ підприємства почергово

презентує керівництву три нових типи продукції. З попереднього досвіду відомо, що близько 70,0 % (тобто Ω = 0,700) пропозицій відділу приймаються дирекцією для виконання. Необхідно визначити ймовірність того, що всі три пропозиції будуть прийняті на підприємстві, а також імовірність того, що не буде прийнята жодна з них. Для того, щоб вирішити цю задачу існує два способи, перший з яких являє собою древо ймовірностей. Дерево ймовірностей (рис. 3.4) дає повне описання ситуації та ілюструє результат для кожного з трьох проектів відділу.



### Рис. 3.4. Дерево ймовірностей для послідовності трьох проектів

Таким чином, наведений рисунок свідчіть, що ймовірність того, що керівництво підприємства схвалить усі три види нової продукції, складає 34%, а ймовірність того, що не буде прийнятий жоден із проектів, дорівнює 3%.

Варто звернути увагу на те, що показані впродовж гілок ймовірності завжди дорівнюють 0,7 та 0,3, оскільки проекти розглядаються незалежно один від одного. Також відмітимо, що існує

три способи ухвалення двох видів нової продукції і, відповідно, три гілки дерева, які це показують.

Для того, щоб отримати розподіл імовірності кількості схвалених проектів, потрібно підсумувати всі ймовірності різних способів реалізації даного проекту (табл. 3.6).

*Таблиця 3.6*

## Ймовірність отримання певної кількості замовлень

|  |  |  |
| --- | --- | --- |
| Кількість схвалених нових видів товару | Відсоток схвалених нових видів товару | Імовірність |
| 0 | 0,0 | 0,03 |
| 1 | 33,3 | 0,18 (0,06+0,06+0,06) |
| 2 | 66,7 | 0,45 (0,15+0,15+0,15) |
| 3 | 100,0 | 0,34 |

Виходячи із розподілу ймовірностей можна знайти будь-яку необхідну ймовірність за допомогою складання відповідних значень. Так, ймовірність схвалення не менш ніж 2 проектів складає 0,45 + 0,34 = 0,79 або 79 %.

У даному випадку ми мали лише три проекти, що зробило наше дерево ймовірностей не дуже громіздким. Однак, бувають ситуації, коли величин у рази більше, тоді використання дерева ймовірностей є недоцільним, що підводить нас до другого варіанту розрахунків, пов’язаних із використанням математичного сподівання і стандартного відхилення.

Середня кількість випадків настання події (математичне сподівання) у випадку біноміального розподілу виражається за допомогою формули *Е(Х) = пΩ,* тобто воно розраховується як добуток числа ймовірностей реалізації події на ймовірність її реалізації. Математичне сподівання для частки дорівнює ймовірності настання події:

*Е*( )  *Е*( *р*)  

*Х*

*п*

(3.9)

Існують також формули для визначення середньоквадратичного відхилення для величини, якій властивий біноміальний розподіл:

 

(1 )

*п*

 

(3.10)

(3.11)

*п*(1 )

Формула 3.10 використовується для розрахунку середньоквадратичного відхилення для кількості випадків настання

події *Х*, а формула 3.11 – для частки або відсотку

*P*  *Х* .

*п*

У розглянутому раніше прикладі з новою продукцією відділу маркетингу *п* = 3, а Ω = 0,7. Якщо ми використаємо наведені формули, то математичне сподівання та середньоквадратичне відхилення можна буде знайти наступним чином (табл. 3.7).

*Таблиця 3.7*

## Розрахунок показників математичного сподівання та середньоквадратичного відхилення для кількості нових товарів відділу маркетингу

|  |  |  |
| --- | --- | --- |
| Показник | Кількість випадків настання події *Х* | Частка або відсоток,*P*  *Х**п* |
| Математичне сподівання | *Е*( *Х* )  3\* 0,7  2,1 новихтоварів | *Е*( *р*)    0,7 або 70 % |
| Середньоквадратичне відхилення |   3\*0,7(1 0,7)  0,794 нових товарів |   0,7 \*(0,7 1)  0,2653або 26,5 % |

Отже, варто очікувати, що 2,1 одиниць загальної кількості нових товарів будуть схвалені керівництвом; у деяких випадках їх буде менше (3) або більше (2). Величина цієї невизначеності розраховується за допомогою середньоквадратичного відхилення від середньої, в даному випадку це 0,794 нових запропонованих товарів. Останнє число, 26,5 %, що характеризує стандартне відхилення для відсотка, трактується як відсоток від одиниці замість відсотка від деякої кількості. Це означає, що в ситуації, коли ми очікуємо 70 % імовірності настання події (ймовірності схвалення нового товару), то в реальності ми можемо отримати на 26,5 % одиниць більше цього результату (70,0 + 26,5 = 96,5 %) або на 26,5 % менше цього

результату (70,0 – 26,5 = 43,5 %).

Розглянемо інший приклад, коли величини *п* та Ω є відомими, а нам необхідно знайти ймовірність того, що *Х* буде дорівнювати певному значенню *х*. Для цього існує спеціальна формула, яку використовують у біноміальному розподілі:

*Р*( *Х*  *х*) 

*п*! *х*!(*п*  *х*)!

 *х* (1 )*п**х* 

1\* 2 \*3\*...\* *п* (1\* 2\*3\*...\* *х*)\* (1\* 2\*3...\*(*п*  *х*)!

*х* (1 )*п**х*

(3.12)

Якщо застосувати цю формулу для кожного зі значень *х* від 0 до

*п*, можна повністю розрахувати розподіл імовірностей.

Припустимо, маркетинговий відділ презентував п’ять нових проектів (*п* = 5) для розгляду. Імовірність успіху кожного з них дорівнює 60 % (Ω = 0,6). Розрахуємо ймовірність того, що керівництвом буде схвалено 2 проекти, тобто *х* = 2. Розрахунок буде мати наступний вигляд:

*Р*( *Х*  2)  5! 0,62 \*0,43  1\* 2\*3\* 4\*5 0,36\*0,0384  0,138.

2!\*(5  2)!

1\* 2\*1\* 2\*3

Таким чином, ми можемо зробити висновок, що з імовірністю 13,8 % керівництво схвалить одночасно два нових проекти. Якщо існує необхідність визначити ймовірність успіху для 2 та більше разів, потрібно використати формулу ще три рази (для *х* = 3, *х* = 4, *х* = 5). Необхідна ймовірність буде розрахована як сума отриманих значень.

Ці розрахунки також можна проводити в редакторі *Exсel* за допомогою функції *БИНОМРАСП*.

## Апроксимація біноміального розподілу нормальним

Між біноміальним і нормальним розподілами випадкових величин існують дві великі відмінності: по-перше, в результаті будь- якого нормального розподілу можна отримати дробові числа, наприклад, 7,8; 5,563 тощо, у той час як результатом біноміального розподілу завжди виступає тільки ціле число; по-друге, нормальний розподіл завжди є симетричним, тоді, коли біноміальний – асиметричний (окрім випадків, коли Ω = 0,5).

Однак для зручності розрахунків біноміальний розподіл можна апроксимувати за допомогою нормального розподілу якщо *n* є достатньо великим, а ймовірність Ω не дуже близька до 1 чи 0. Краще за все, якщо *п* більше 5, а Ω знаходиться в районі 0,5 [23, c. 61]. Апроксимація допоможе визначити ймовірності для біноміального розподілу складних і громіздких розрахунків на більш прості.

Використовуючи нормальний розподіл у якості заміни біноміальному, ми робимо припущення, що значення дискретної випадкової величини 2 знаходиться в проміжку між 1,5 і 2,5, що називається **поправкою на неперервність**.

Розглянемо можливість апроксимації на конкретному прикладі. Так, підприємство кожен день виробляє велику кількість шоколадних

батончиків, 40 % із яких є бракованими. Для перевірки якості працівник обирає 10 батончиків, виготовлених протягом дня. Яка ймовірність того, що 7 чи більше з них є бракованими?

Для вирішення цієї задачі, перш за все, використаємо більш громіздкий метод і скористаємось формулою 3.12. Нам буде необхідно використати форму 4 рази для *x* = 7, *Х* = 8, *x* = 9 та *x* = 10.

У результаті розрахунків отримуємо:

*Р*(*x*  7) 

10!

7!(10  7)!

0,47 (1  0,4)3  1\* 2 \* 3 \* 4 \* 5 \* 6 \* 7 \*8 \* 9 \*10 0,47 \* 0,63  0,042 ,

1\* 2 \* 3\* 4 \* 5 \* 6 \* 7 \*1\* 2 \* 3

*Р*(*x*  8) 

*Р*(*x*  9) 

10!

8!(10  8)!

10!

0,48(1  0,4)2  1\* 2 \* 3 \* 4 \* 5 \* 6 \* 7 \* 8 \* 9 \*10 0,48 \* 0,62  0,011,

1\* 2 \* 3\* 4 \* 5 \* 6 \* 7 \* 8 \*1\* 2

0,49 (1  0,4)1  1\* 2 \* 3 \* 4 \* 5 \* 6 \* 7 \* 8 \* 9 \*10 0,49 \* 0,61  0,002 ,

*Р*(*x*  10) 

9!(10  9)!

10!

10!(10 10)!

0,410 (1 0,4)0 

1\* 2 \* 3 \* 4 \* 5 \* 6 \* 7 \* 8 \* 9 \*1

1\* 2 \* 3\* 4 \*5\* 6 \* 7 \*8\*9 \*10

1\* 2 \*3\* 4 \*5\* 6 \* 7 \*8\* 9 \*10\* 0

0,410 \* 0,60  0,000 .

Підсумувавши всі отримані ймовірності (0,042 + 0,011 + 0,002 + 0,000 = 0,055), ми можемо зробити висновок, що з імовірністю 5,5 % у вибірці з 10 шоколадних батончиків ми знайдемо 7 і більше бракованих одиниць.

Розрахунки із заміною біноміального розподілу нормальним достатньо прості. Дискретне значення, що дорівнює 7, необхідно представити неперервною випадковою величиною в межах від 6,5 до 7,5. Замість того, щоб знаходити ймовірність дискретної величини 7 та більшої кількості бракованих одиниць, знайдемо ймовірність значення неперервної випадкової величини більшої ніж 6,5 бракованих одиниць.

За формулою 3.2 ми можемо визначити, що математичне сподівання дорівнює 4 бракованим одиницям (   10\* 0,4  4 ). Середньоквадратичне відхилення знаходиться за наступною формулою:

*п*(1 )

 

(3.13)

Воно складає 1,549 одиниць (



 1,549 ).

За формулою 3.7 розрахуємо значення *z* для 6,5:

10  0,4(1 0,4)

*z*  6,5  4  1,613 .

1,549

У таблиці ймовірностей нормального розподілу знаходимо, що ймовірність того, що серед 10 обраних шоколадних батончиків буде знаходитися 7 або більше бракованих одиниць, дорівнює 1 – 0,9463 = 0,0537 або 5,4%. Отже, отриманий результат майже не відрізняється від результатів розрахунків при біноміальному розподілі.

Розглянемо ще один приклад заміни біноміального розподілу нормальним для ситуації з відсотковим співвідношенням. З минулих спостережень орнітологів відомо, що із 1000 лебедів, що прилітають зимувати на озеро, 40 є чорними. Яка ймовірність того, що цієї зими на озеро прилетить більше 50 чорних лебедів?

Математичне сподівання, тобто середня величина, буде

дорівнювати

  40  0,040 . Стандартне відхилення в цьому

1000

випадку буде розраховуватися за формулою:

(1 )

*п*

 

та буде дорівнювати  

(3.14)

 0,006 . Для подальшого

0,040(1 0,040)

1000

вирішення здійснимо нормування величини z за вже відомою формулою 3.7:

*z*  0,050  0,040  1,667

.

0,006

Таким чином, згідно з додатком *А*, ймовірність того, що на озеро прилетить більше 50 чорних лебедів, складає 1 – 0,9525 = 0,0475 або

4,8 %.

## Питання для самоконтролю

1. Що таке випадкова величина?
2. Яка відмінність між випадковою величиною та числом?
3. Наведіть приклади випадкових величин.
4. Що таке дискретна випадкова величина?
5. Що таке неперервна випадкова величина?
6. Що таке правило розподілу ймовірностей?
7. Дайте визначення закону розподілу випадкових величин.
8. Назвіть закони розподілу, які ви знаєте.
9. Що таке функція розподілу?
10. Охарактеризуйте нормальний закон розподілу.
11. Охарактеризуйте біноміальний закон розподілу.
12. Назвіть основний недолік використання дерева ймовірностей при розрахунках біноміального закону розподілу.
13. У чому сутність апроксимації біноміального закону нормальним?