ПРАКТИЧНЕ ЗАНЯТТЯ № 3. (*Додаток*).
Граничні теореми в схемі Бернуллі.

1. Розподіл Пуассона. 2. Схема випробувань Бернуллі з «малою» кількістю випробувань. 3. Схема випробувань Бернуллі в випадку «масових» явищ. 4. Гранична теорема Пуассона. 5. Застосування граничної теореми Пуассона.

1. *Розподіл Пуассона.*

Випадкова величина π має розподіл *Пуассона з параметром* λ, якщо вона може приймати всі цілі невід'ємні значення $\pi \in \{0, 1, 2 \ldots \}$ і при цьому:

$$p_k = P(\pi = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!},\ k = 0, 1, 2, \ldots :.$$

Гранична теорема Пуассона: Нехай випадкова величина ξ_n визначає число «*Успіхів*» в серії з n випробувань Бернуллі.
Якщо ймовірність «*Успіху*» (p_n) залежить від кількості (n) випробувань,

$$p_n \to 0, \text{ при } n \to \infty,$$

а також

$$\lim_{n\to\infty} (p_n \cdot n) = \lambda,$$

то для довільного значення $k = 0, 1, 2, \ldots$

$$\lim_{n\to\infty} P(\xi_n = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

На практиці це означає, що для біноміального розподілу з параметрами (n, p), у випадку, коли p «*достатньо мале*», n – «*достатньо велике*», а $\lambda = n \cdot p$, то для довільного k:

$$P(\xi_n = k) = C_n^k \cdot p^k \cdot (1-p)^{n-k} \approx e^{-\lambda} \cdot \frac{\lambda^k}{k!},\ k = 0, 1, 2, \ldots .$$

2. *Схема випробувань Бернуллі з «малою» кількістю випробувань.*

Приклад 1. Ймовірність виграти в лотереї дорівнює **0,01**. Маємо **5** білетів лотереї. Яка ймовірність того, що ***принаймні один*** з них виграє?

Розв'язок. Введемо схему випробувань Бернуллі, яка описує процедуру перевірки окремих білетів лотереї. «УСПІХОМ» будемо називати ситуацію, коли білет є виграшним. Згідно з умовою:

$$p = P\{«УСПІХУ»\} = 0{,}01.$$

Оскільки маємо 5 білетів і необхідно перевірити кожен з них незалежно від інших, то кількість випробувань:

$$n = 5.$$

Введемо випадкову величину ξ, яка визначає кількість виграшних білетів серед 5 білетів, що маємо в розпорядженні. Тоді приймаючи до уваги приклад 5 (ПРАКТИЧНЕ ЗАНЯТТЯ № 1), можемо стверджувати, що випадкова величина ξ має *біноміальний розподіл з параметрами* (5, 0,01).

$$\xi \Leftrightarrow B(5, 0{,}01).$$

Випадкова подія: A = {«Принаймні один білет є виграшним»}
означає, що $A = \{\xi \geq 1\}$. Оскільки згідно з визначенням:

$$Q(k, n, p) = P\{B(n, p) \geq k\}$$

то

$$P\{«Принаймні один білет є виграшним»\} = P\{A\} = P\{\xi \geq 1\} =$$
$$= P\{B(5, 0{,}01) \geq 1\} = Q(1, 5, 0{,}01) = 0{,}049.$$

(**Відповідь: *P* = 0,049.**)

Приклад 2. Брошура має **100** сторінок. Відомо, що в тексті цієї брошури є ***вісім*** помилок. Нехай випадкова величина ξ означає кількість помилок на *випадково вибраній* сторінці.
1) Знайти розподіл випадкової величини ξ.
2). Використовуючи таблиці ймовірностей $Q(k, n, p)$ для біноміального розподілу підрахувати ймовірність того, що на *випадково вибраній* сторінці буде ***принаймні одна*** помилка.

Розв'язок.
1) Введемо схему випробувань Бернуллі, яка описує те, яким чином помилки можуть потрапити на цю сторінку. Оскільки в умові немає жодних припущень щодо процесу розташування помилок на сторінках брошури, то логічно припустити, що ймовірність потрапити кожної *конкретної* помилки на кожну *конкретну* сторінку є *такою самою* для всіх сторінок, тобто дорівнює 0,01. Припустимо, що сторінку ми вже вибрали. «УСПІХОМ» будемо називати ситуацію, коли якась серед *восьми* помилок, що є в тексті, виявиться саме на цій вибраній сторінці. Згідно з умовою:

$$p = P\{\text{«УСПІХУ»}\} = 0{,}01.$$

Оскільки маємо 8 помилок і немає жодної інформації, що між ними існує якийсь зв'язок, то необхідно вважати, що кожна з них може потрапити на кожну сторінку незалежно від інших. Тому необхідно провести $n = 8$ незалежних «*випробувань*», щоб перевірити кожну з них. Отже приймаючи до уваги приклад 5 (ПРАКТИЧНЕ ЗАНЯТТЯ № 1), можемо стверджувати, що випадкова величина ξ має *біноміальний розподіл з параметрами* (8, 0,01).

$$\xi \Leftrightarrow B(8, 0{,}01).$$

$$C = \{\text{«На } \textit{випадково вибраній} \text{ сторінці буде принаймні одна помилка»}\}.$$

$$C = \{\xi \geq 1\}.$$

$$P\{C\} = P\{\xi \geq 1\} = P\{B(8, 0{,}01) \geq 1\} = Q(1, 8, 0{,}01) = 0{,}077.$$

(**Відповідь:** $\boldsymbol{P = 0{,}077}$.)

Стохастичні експерименти, що розглядаються в цих прикладах, формальним чином описує схема випробувань Бернуллі, а випадкові величини мають *біноміальний розподіл*. Якщо випробувань (n) не дуже багато ($0 < n < 20$), то кожне додаткове випробування може мати суттєвий вплив на остаточний результат. Це добре видно на приведених прикладах. Тому з метою полегшення виконання «*технічних*» обчислень побудовані різноманітні таблиці для окремих значень n та p. Зокрема таблиця ймовірностей $Q(k, n, p)$ для біноміального розподілу:

$$Q(k, n, p) = P\{B(n, p) \geq k\} = P\{B(n, p) = k\} + \ldots + P\{B(n, p) = 0\}.$$

3. *Схема випробувань Бернуллі в випадку «масових» явищ*.

Приклад 3.. Ймовірність виграти в лотереї дорівнює **0,01**. В лотереї беруть участь **200** осіб кожна з яких має 1 білет. Яка ймовірність того, що принаймні ***троє*** з них виграють?

Розв'язок. Міркуючи цілком аналогічно, як і в прикладі 1, приходимо до висновку, що цей стохастичний експеримент можна описати за допомогою схеми випробувань Бернуллі, для якої

$$p = P\{\text{«УСПІХУ»}\} = 0{,}01; \; n = 200.$$

Отже випадкова величина ξ, яка визначає кількість осіб, що виграли в лотереї, має *біноміальний розподіл з параметрами* (200, 0,01).

$$\xi \Leftrightarrow B(200,\, 0{,}01).$$

Тому для події:

D = {«Принаймні ***троє*** з учасників лотереї виграють»}

отримаємо наступну рівність:

$$P\{D\} = P\{\xi \geq 3\} = P\{B(200,\, 0{,}01) \geq 3\} = \sum_{k=3}^{200} C_{200}^{k} \cdot (0{,}01)^{k} \cdot (0{,}99)^{200-k}.$$

Можна дещо спростити отриману відповідь, використовуючи наступну властивість ймовірності (див. Лек.№4):

Ймовірність протилежної події $\overline{A} = \Omega \setminus A$ *обчислюємо за формулою*:

$$P(\overline{A}) = P(\Omega \setminus A) = 1 - P(A).$$

Отже:

$$\mathrm{P}\{\xi \geq 3\} = 1 - \mathrm{P}\{\xi < 3\} = 1 - \sum_{k=0}^{2} C_{200}^{k} \cdot (0{,}01)^{k} \cdot (0{,}99)^{200-k}.$$

Але це не розв’язує проблеми, оскільки відповідних таблиць не існує, а безпосереднє обчислення (хоча й можливе) але позбавлене практичного сенсу.

Випадок, коли кількість випробувань (n) в схемі Бернуллі велика, на практиці означає, що маємо справу з «*масовим явищам*». Тому розв’язок знаходиться в іншій площині і пов'язаний з *граничними теоремами теорії ймовірностей* (зокрема в схемі випробувань Бернуллі).

Розподіл Пуассона.

Виявляється, що якщо кількість випробувань (n) *дуже* велика (n кілька сотень), то результат кожного окремого з них вже *не має суттєвого впливу* на остаточний результат. Для великих значень n біноміальний розподіл буде в певному сенсі «*близький*» до деякого *граничного* розподілу. Фактично гранична теорема дає *наближений* розв’язок, *точність* якого є цілком прийнятною з практичної точки зору.

- *Граничні теореми* – це і є формальний вираз *закономірностей, притаманних масовим явищам*, або інакше – формальний вираз *статистичних закономірностей*.

Одним з таких *граничних* розподілів в схемі випробувань Бернуллі є розподіл *Пуассона*.

Визначення. Випадкова величина π має розподіл *Пуассона з параметром* λ, якщо вона може приймати всі цілі невід’ємні значення $\pi \in \{0, 1, 2 \ldots\}$ і при цьому:

$$p_k = P(\pi = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!},\; k = 0, 1, 2, \ldots :.$$

Гранична теорема Пуассона.

Нехай випадкова величина ξ_n визначає число «*Успіхів*» в серії з n випробувань Бернуллі. Припустимо, що ймовірність «*Успіху*» (p_n) залежить від кількості (n) випробувань і при цьому виконуються наступні умови:

1) $p_n \to 0$, при $n \to \infty$, а також 2) $\lim\limits_{n\to\infty} (p_n \cdot n) = \lambda$. То для довільного значення $k = 0, 1, 2, \ldots$ існує наступна границя: $\lim\limits_{n\to\infty} P(\xi_n = k) = e^{-\lambda} \cdot \dfrac{\lambda^k}{k!}$.

Іншими словами, припустимо, що

- $\{(k, p_n(k)),\, k = 0, 1, \ldots n\}$ – біноміальний розподіл з параметрами (n, p_n),
- $\{(k, p_\lambda(k)),\, k = 0, 1, 2, \ldots .\}$ – розподіл Пуассона з параметром $\lambda = n \cdot p_n$.

Тоді для довільного значення $k = 0, 1, 2, \ldots$ має місце приблизна рівність

$$p_n(k) \approx p_\lambda(k),\ k = 0, 1, 2, \ldots$$

якщо тільки число n є «*достатньо великим*», а при цьому p_n є «*достатньо малою*». Це і є математичний вираз ствердження:

- $B(n, p) \approx \pi$, тобто для великих значень n біноміальний розподіл буде в «*певному сенсі близький*» до деякого *граничного* розподілу.

Статистична закономірність в даному випадку полягає в тому, що при одночасному виконанні двох умов: (n – *велике*, p_n – *мале*) остаточний результат визначає число ($\lambda = n \cdot p_n$).

Можна довести, що для випадкової величини π, яка має розподіл Пуассона з параметром λ, параметри відповідно дорівнюють: 1). $E(\pi) = \lambda$. 2). $D(\pi) = \lambda$. Отже наступна *статистична закономірність* звучить наступним чином:

- Середня *очікувана кількість* рідких подій в масовому явищі дорівнює λ.

4. *Застосування граничної теореми Пуассона*.

На практиці це означає, що для біноміального розподілу $B(n, p)$ з параметрами (n, p), у випадку, коли:

1) p «*достатньо мале*», і одночасно:

2) n – «*достатньо велике*», для довільного $k = 0, 1, 2, \ldots$:

$$P(\xi_n = k) = C_n^k \cdot p^k \cdot (1-p)^{n-k} \approx e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

де число λ визначається рівністю:

$$\lambda = n \cdot\cdot p.$$

Приклад 3. (продовження). Ймовірність виграти в лотереї дорівнює 0,01. В лотереї беруть участь **200** осіб кожна з яких має 1 білет. Яка ймовірність того, що принаймні ***троє*** з них виграють?

Розв'язок. Нехай, як і раніше, випадкова величина ξ визначає кількість осіб, що виграли в лотереї. ξ має *біноміальний розподіл з параметрами* (200, 0,01): $\xi \Leftrightarrow B(200,\ 0{,}01)$. Число $n = 200$ напевно можна вважати «*достатньо великим*», а значення $p_n = 0{,}01$ без сумніву є «*достатньо мале*». Отже можна використати *граничну теорему Пуассона*:

$$\lambda = n \cdot p_n = 200 \cdot 0{,}01 = 2.$$

Скористаємось тепер таблицею розподілу Пуассона з параметром $\lambda = 2$.

$$\mathrm{P}\{\xi \geq 3\} \approx \mathrm{P}\{\pi \geq 3\} = 1 - \mathrm{P}\{\pi < 3\} =$$

$$= 1 - \mathrm{P}\{\pi = 0\} - \mathrm{P}\{\pi = 1\} - \mathrm{P}\{\pi = 2\} =$$

$$= 1 - (0{,}135 + 0{,}270 + 0{,}270\}) = 1 - 0{,}675 = 0{,}325.$$

Відповідь: *P* = 0,325.

ЗАУВАЖЕННЯ. Очевидною позитивною рисою такого підходу з практичної точки зору є те, що розв'язок не зміниться, якщо, наприклад, $n = 400$, а $p_n = 0{,}005$, або $n = 2000$, а $p_n = 0{,}001$, ітп.

Приклад 4. Аналізуючи досвід проведення екзаменів з певного предмету встановлено, що ймовірність успішного результату дорівнює **0,98**. Група складається з **80** студентів. Яка ймовірність того, що ***принаймні* 76** з них здадуть екзамен.

Розв'язок. Введемо схему випробувань Бернуллі, яка описує стохастичний експеримент, що розглядається. «УСПІХОМ» будемо називати ситуацію, коли студент ***не здасть*** екзамен. Випадкова величина ξ визначає кількість осіб, що *не здали* екзамен. Тоді згідно з умовою:

$$p = P\{\text{«УСПІХУ»}\} = 1 - 0{,}98 = 0{,}02.$$

«Випробуванням» буде складання іспиту кожним окремим студентом, тобто $n = 80$. Таким чином випадкова величина ξ має *біноміальний розподіл з параметрами* (80, 0,02): $\xi \Leftrightarrow B(80,\ 0{,}02)$. Оскільки в даному випадку $p = 0{,}02$ – «*достатньо мале*», а $n = 80$ – «*достатньо велике*», то можна використати теорему Пуассона: $\lambda = n \cdot p = 80 \cdot 0{,}02 = 1{,}6$. Отже:

$$\text{P}\{\text{Принаймні 76 студентів здасть екзамен}\} = \text{P}\{\text{Найбільше 4 особи не здадуть екзамен}\} = \text{P}\{\xi \le 4\} \approx \text{P}\{\pi \le 4\} =$$
$$= \text{P}\{\pi = 0\} + \text{P}\{\pi = 1\} + \text{P}\{\pi = 2\} = \text{P}\{\pi = 3\} + \text{P}\{\pi = 4\},$$

де випадкова величина π має розподіл Пуассона з параметром $\lambda = 1{,}6$. Використовуючи відповідну таблицю отримаємо:

$$\text{P}\{\xi \le 4\} \approx \text{P}\{\pi \le 4\} = 0{,}202 + 0{,}323 + 0{,}258 + 0{,}137 + 0{,}055 = 0{,}975.$$

Відповідь: $\boldsymbol{P = 0{,}975}$.

Приклад 5. Партію з 50 виробів, серед яких є 5 штук з певними дефектами, випадковим чином запаковано в 10 пачок по 5 виробів в кожну. Клієнт отримав одну з цих пачок.

1). Скільки виробів з дефектами отримав цей клієнт?

Відповідь: Кількість (ξ) виробів з дефектами – це випадкова величина, що має *біноміальний розподіл з параметрами* (5, 0,1): $\boldsymbol{\xi \Leftrightarrow B(5,\ 0{,}1)}$.

2). Яка ймовірність того, що кількість виробів з дефектами не менша двох?

Відповідь: $\boldsymbol{P = 0{,}081}$.

3). На яку кількість виробів з дефектами повинен *сподіватися* цей клієнт?

Відповідь: $\boldsymbol{m = 0{,}5}$.

4). На скільки ця оцінка є точною?

Відповідь: $\boldsymbol{\sigma = S(\xi) = 0{,}67}$.

Розв'язок.

1) Введемо схему випробувань Бернуллі, яка описує те, яким чином виріб з дефектом може потрапити до пачки, яку отримав клієнт. Оскільки в умові немає жодних припущень щодо процесу розміщення виробів дефектами в окремих пачках, то логічно припустити, що:

- Ймовірність потрапити кожного виробу дефектом до кожної *конкретної* пачки є *такою самою* для всіх пачок, тобто дорівнює 0,01.

«*УСПІХОМ*» будемо називати ситуацію, коли якийсь серед *п'яти* виробів з дефектами, виявиться саме в тій пачці, яку отримав клієнт. Згідно з умовою: $p = P\{\text{«УСПІХУ»}\} = \dfrac{1}{10} = 0{,}1$. Оскільки в партії є 5 виробів з дефектами і немає жодної інформації, що між ними існує якийсь зв'язок, то необхідно вважати, що кожен з них може потрапити до будь-якої пачки незалежно від інших. Тому необхідно провести $n = 5$ незалежних «*випробувань*», щоб всі вироби з дефектами були розміщені по окремим пачкам.

Таким чином пакування виробів по пачкам, про яке йдеться в прикладі 5, можна описати за допомогою схеми випробувань Бернуллі *з параметрами*:

$$n = 5,\ p = 0{,}1.$$

1). Кількість виробів з дефектами, яку може отримати клієнт, буде випадковою величиною. Позначимо її символом ξ.

Тоді приймаючи до уваги приклад 5 (ПРАКТИЧНЕ ЗАНЯТТЯ № 1), можемо стверджувати, що випадкова величина ξ має *біноміальний розподіл з параметрами* ($n = 5$, $p = 0{,}1$). $\xi \Leftrightarrow B(5,\ 0{,}1)$.

2). $P\{$кількість виробів з дефектами не менша двох$\} = P\{\xi \geq 2\}$.
Згідно з визначенням: $Q(k, n, p) = P\{B(n, p) \geq k\}$. Таким чином:

$$P\{\xi \geq 2\} = P\{B(5,\ 0{,}1) \geq 2\} = Q(2,\ 5,\ 0{,}1) = 0{,}081.$$

3). На яку кількість виробів з дефектами повинен *сподіватися* цей клієнт?

$$m = E(\xi) = E(B(5,\ 0{,}1)) = n \cdot p = 5 \cdot 0{,}1 = 0{,}5.$$

4). На скільки ця оцінка є точною? $\sigma^2 = D(\xi) = E(\xi - m)^2$.
Точність можна оцінити, обчислюючи стандартне відхилення:

$$\sigma = S(\xi) = \sqrt{D(\xi)}\ .$$

$$\sigma^2 = D(B(5,\ 0{,}1)) = n{\cdot}p{\cdot}(1 - p) = 5{\cdot}0{,}1{\cdot}(1 - 0{,}1) = 5{\cdot}0{,}1{\cdot}0{,}9 = 0{,}45.$$

$$\sigma = \sqrt{0{,}45} = 0{,}67.$$