ПРАКТИЧНЕ ЗАНЯТТЯ № 5(*Додаток*)
Граничні теореми.

(ПРАКТИЧНЕ ЗАНЯТТЯ № 2.)

2. *Властивості математичного сподівання та дисперсії*.

Приклад 4. Довести наступні властивості математичного сподівання:

1) $E(a\cdot\xi) = a\cdot E(\xi)$.

2) $E(\xi + \eta) = E(\xi) + E(\eta)$.

3) $E(c_1\cdot\xi_1 + c_2\cdot\xi_2 + \ldots + c_n\cdot\xi_n) = c_1\cdot E(\xi_1) + c_2\cdot E(\xi_2) + \ldots + c_n\cdot E(\xi_n)$.

4) Якщо ξ та η ***незалежні випадкові величини***, то справедлива наступна рівність:

$$E(\xi\cdot\eta) = E(\xi)\cdot E(\eta).$$

Приклад 5. Довести наступні властивості дисперсії:

1) Формула для підрахунку: $D(\xi) = E(\xi)^2 - (E(\xi))^2$.

2) $D(a\cdot\xi + b) = a^2\cdot D(\xi)$.

3) Нерівність Чебишева

$$P\{|\xi - E(\xi)| > \varepsilon\} \le \frac{D(\xi)}{\varepsilon^2}.$$

(ЛЕКЦІЯ 6(Б))

Лема 3. (***Нерівність Чебишева.***) Для довільної випадкової величини ξ та довільного додатного числа $\varepsilon > 0$ справедлива наступна нерівність:

$$P\{|\xi - E(\xi)| > \varepsilon\} \le \frac{D(\xi)}{\varepsilon^2}.$$

Доведення. Доведемо нерівність для неперервної випадкової величини ξ, що має щільність $f_\xi(x)$, $-\infty < x < \infty$. В цьому випадку (як це випливає безпосередньо з визначення) дисперсія обчислюється за формулою:

$$D(\xi) = \int_{-\infty}^{\infty} (x - E(\xi))^2 \cdot f_\xi(x)dx.$$

Оскільки щільність $f_\xi(x) \ge 0$, є невід'ємною функцією для довільного дійсного числа $-\infty < x < \infty$, то інтеграл від функції:

$$(x - E(\xi))^2\cdot f_\xi(x)$$

може лише зменшитись, якщо область інтегрування зменшується.

Тому замінимо в інтегралі правої частини рівності всю числову вісь

$$-\infty < x < \infty$$

лише такою її підмножиною, для якої виконується умова:

$$|x - E(\xi)| > \varepsilon.$$

Тоді:

$$\int_{-\infty}^{\infty} (x - E(\xi))^2 \cdot f_\xi(x)dx \ge \int_{|x-E(\xi)|>\varepsilon} (x - E(\xi))^2 f_\xi(x)dx \ge$$

$$\ge \int_{|x-E(\xi)|>\varepsilon} \varepsilon^2 \cdot f_\xi(x)dx \ge \varepsilon^2\cdot \int_{|x-E(\xi)|>\varepsilon} f_\xi(x)dx.$$

Пояснюючи в лекції 8 фізичний зміст щільності ми встановили справедливість наступної *формальної* рівності:

$$P\{\xi = x\} = f_\xi(x)\cdot dx.$$

Точніше її можна записати наступним чином:

- Для довільної *борелівської* підмножини $A \subset (-\infty; \infty)$ справедлива наступна рівність:

$$\int_A f_\xi(x)dx = P\{\xi \in A\}.$$

В нашому випадку такою множиною є:

A = {множина чисел x для яких виконується нерівність: $|x - E(\xi)| > \varepsilon$}.

Тому:

$$D(\xi) \geq \varepsilon^2 \cdot \int_{|x-E(\xi)|>\varepsilon} f_\xi(x)dx = \varepsilon^2 \cdot P\{|\xi - E(\xi)| > \varepsilon\}.$$

Тобто

$$P\{|\xi - E(\xi)| > \varepsilon\} \leq \frac{D(\xi)}{\varepsilon^2},$$

що й треба було довести.

(ЛЕКЦІЯ 8(А))

7. *Типи збіжності випадкових величин.*

Сама назва «*граничні теореми*» передбачає певну «*збіжність*». Тому перш ніж познайомитись з головними *типами граничних теорем*, необхідно визначити відповідні їм *поняття збіжності*.

Формальне визначення випадкової величини передбачало два рівнозначних з практичної точки зору підходи.

1. Перший з них дає абстрактне визначення $\xi = \xi(\omega)$, $\omega \in \Omega$, як $\mathfrak{J}$-вимірної функції, визначеної на вимірному просторі $(\Omega, \mathfrak{J}, P)$.

2. Другий пропонує однозначно представляти випадкову величину ξ шляхом визначення її функція розподілу

$$F_\xi(x) = P\{\xi < x\}, -\infty < x < \infty.$$

Тому серед великої кількості різноманітних типів збіжності виберемо ті, що відповідають названим підходам до визначення випадкової величин.

1. Перша з них – це збіжність *за ймовірністю*, (або *стохастична*), що відповідає визначенню випадкової величини, як функції $\xi = \xi(\omega)$, $\omega \in \Omega$, на ймовірнісному просторі $(\Omega, \mathfrak{J}, P)$.

2. Друга – це слабка *збіжність*, що безпосередньо пов'язана із збіжністю *функцій розподілу* випадкових величин, які вивчаються.

Припустимо, що на ймовірнісному просторі $(\Omega, \mathfrak{J}, P)$ визначено послідовність випадкових величин:

$$\xi_1 = \xi_1(\omega), \xi_2 = \xi_2(\omega), \ldots, \xi_n = \xi_n(\omega), \ldots, \omega \in \Omega.$$

Визначення. Послідовність випадкових величин ξ_n, $n = 1, 2, \ldots$, збігається при $n \to \propto$ *за ймовірністю* до випадкової величини $\xi = \xi(\omega)$, $\omega \in \Omega$, що визначена *на тому самому* ймовірнісному просторі $(\Omega, \mathfrak{J}, P)$, якщо для будь-якого додатного числа $\varepsilon > 0$ виконується наступна рівність:

$$\lim_{n\to\infty} P\{|\xi_n - \xi| > \varepsilon\} = 0$$

Інколи подібна збіжність називається *стохастичною*. Для її позначення використовується наступний символ:

$$\xi_n \underset{n\to\infty}{\overset{P}{\to}} \xi.$$

(ЛЕКЦІЯ 8(Б))

4. *Гранична теорема Бернуллі.*

В перших лекціях серед базових для теорії ймовірностей понять називався також встановлений *дослідним* шляхом «*принцип стійкості частот*».

Ця «*статистична закономірність*» набула статусу «*закону*» в побудованому на формальних аксіоматичних принципах розділі математики під назвою «*Теорія ймовірностей*».

Відповідне строге математичне твердження, під назвою «*гранична теорема Бернуллі*», що відображає ім'я її автора, є історично першим *граничним законом* і його можна вважати *першою* теоремою, характерною виключно для «*Теорії ймовірностей*», тобто результатом, який *фактично* започаткував нову математичну дисципліну.

В лекції 2 було введено поняття частоти $\nu_n(A) = \dfrac{\mu_n(A)}{n}$ випадкової події A, де $\mu_n(A)$ означає кількість появ події A в n незалежних повтореннях того самого експерименту. Як тоді підкреслювалось:

- Закономірність «*стійкості частот*» полягає в тому, що у великих серіях (для великих значень n) справедливе наступне співвідношення:

$$\nu_n(A) \approx p,$$

тобто частота $\nu_n(A)$ мало відрізняється від *деякого числа* p.

Причому експерименти свідчать, що при *збільшенні* n різниця між $\nu_n(A)$ та p, як правило, *зменшується*.

Закономірність, яку помітило багато математиків, Бернуллі першим сформулював та строго довів у вигляді наступної теореми.

Теорема. (*Бернуллі*). Якщо ймовірність випадкової події A дорівнює p:

$$P\{A\} = p,$$

то для будь-якого числа $\varepsilon > 0$ виконується наступна рівність:

$$\lim_{n\to\infty} P\{|\nu_n(A) - p| > \varepsilon\} = 0,$$

тобто

$$\frac{\mu_n(A)}{n} \underset{n\to\infty}{\overset{P}{\to}} p.$$

Доведення. Пов'яжемо з k-тим по порядку повторенням експерименту індикатор (η_k) події A, тобто випадкову величину η_k, $k = 1, 2, \dots, n$, що приймає два значення:

$$\eta_k \in \{0, 1\},$$

причому

$$\eta_k = 1, \text{ якщо поді } A \text{ відбулась;}$$
$$\eta_k = 0, \text{ якщо } A \text{ – не відбулась.}$$

Тоді, очевидно, справедлива наступна стохастична рівність:

$$\mu_n(A) = \eta_1 + \eta_2 + \dots + \eta_n.$$

З іншого боку, η_k, $k = 1, 2, \dots, n$, – незалежні випадкові величини, що мають такий самий розподіл $\{(0, 1-p), (1, p)\}$.

Тобто:

$$P\{\eta_k = 1\} = p, P\{\eta_k = 0\} = 1 - p, k = 1, 2, \dots, n.$$

Отже

$$E(\eta_n) = p;\ D(\eta_n) = p\cdot(1-p).$$

Тобто:

$$\nu_n(A) = \frac{\mu_n(A)}{n} = \frac{\eta_1 + \eta_2 + \dots + \eta_n}{n}$$

Використовуючи властивості математичного сподівання та дисперсії отримаємо:

$$E(\nu_n(A)) = E\left(\frac{\mu_n(A)}{n}\right) = \frac{1}{n}\cdot E(\mu_n(A)) = \frac{1}{n}\cdot E(\eta_1 + \eta_2 + \ldots + \eta_n) =$$

$$= \frac{1}{n}\cdot(E(\eta_1) + E(\eta_2) + \ldots + E(\eta_n)) = \frac{1}{n}\cdot n\cdot p = p.$$

$$D(\nu_n(A)) = D\left(\frac{\mu_n(A)}{n}\right) = \frac{1}{n^2}\cdot E(\mu_n(A)) = \frac{1}{n^2}\cdot D(\eta_1 + \eta_2 + \ldots + \eta_n) =$$

$$= \frac{1}{n^2}\cdot(D(\eta_1) + D(\eta_2) + \ldots + D(\eta_n)) = \frac{1}{n^2}\cdot n\cdot p\cdot(1-p) = \frac{p\cdot(1-p)}{n}.$$

Позначимо $\xi = \nu_n(A)$, та застосуємо нерівність Чебишева:

$$P\{|\xi - E(\xi)| > \varepsilon\} = P\{|\nu_n(A) - p\,| > \varepsilon\} \le \frac{D(\nu_n(A))}{\varepsilon^2} = \frac{p\cdot(1-p)}{n\cdot\varepsilon^2} \underset{n\to\infty}{\to} 0.$$

тобто

$$\nu_n(A) = \overset{P}{\underset{n\to\infty}{\to}} p.$$

Що й треба було довести.

Бернуллі змушений був написати цілу книжку, щоб довести цю теорему. Ми ж отримаємо її як очевидний наслідок знаменитої нерівності Чебишева.

Другий тип збіжності – це слабка *збіжність*, що безпосередньо пов’язана із збіжністю *функцій розподілу* випадкових величин, які вивчаються.

Припустимо, що маємо послідовність випадкових величин ξ_1, ξ_2, …, (*не обов’язково* визначених *на тому самому* ймовірнісному просторі).

Нехай $F_1(x)$, $F_2(x)$, …, відповідні їм функції розподілу:

$$F_n(x) = P\{\xi_n < x\},\ -\infty < x < \infty,\ n = 1, 2, \ldots$$

Крім того маємо випадкову величину ξ з функцією розподілу:

$$F(x) = P\{\xi < x\},\ -\infty < x < \infty.$$

Визначення. Послідовність випадкових величин ξ_n, $n = 1, 2, \ldots$, *слабо* збігається при $n \to \propto$ до випадкової величини ξ, якщо для будь-якої точки x, в якій функція розподілу $F(x)$ є неперервною, виконується наступна рівність:

$$\lim_{n\to\infty} F_n(x) = F(x),\ -\infty < x < \infty.$$

Для позначення слабкої збіжності використовується наступні символи:

$$\xi_n \underset{n\to\infty}{\Rightarrow} \xi, \text{ або } \xi_n \overset{сл.}{\underset{n\to\infty}{\to}} \xi.$$

(ЛЕКЦІЯ 8(А))

Збіжність випадкових величин.

1. Коментар до слабкої збіжності.

Варто зауважити, що у випадку, коли гранична функція $F(x)$ є неперервною, слабка збіжність $\xi_n \overset{sl}{\underset{n\to\infty}{\Rightarrow}} \xi$ випадкових величин означає збіжність послідовності функцій розподілу $\{F_n(x), n = 1, 2, \ldots\}$ до $F(x)$ для *всіх* значень x.

Більше того, оскільки функції розподілу *обмежені*, то ця *збіжність* одночасно буде і *рівномірною*.

Припустимо, що функція розподілу $F(x)$ випадкової величини ξ *не є неперервною*.

Оскільки функція розподілу є *неспадною* і *обмеженою* функцією, то вона може мати лише *стрибки*, а їх кількість – *щонайбільше злічена*.

Отже, слабка збіжність випадкових величин означає збіжність відповідних функція розподілу для *майже всіх x*. Тобто у випадку слабкої збіжності

$$\xi_n \underset{n\to\infty}{\overset{sl}{\Rightarrow}} \xi$$

множина тих x, для яких виконується співвідношення:

$$\lim_{n\to\infty} F_n(x) \neq F(x)$$

щонайбільше злічена.

Приведемо простий приклад, який показує, для чого у визначенні слабкої збіжність присутнє обмеження на точки розриву функції $F(x)$, чому вимога збіжності послідовності функцій розподілу $\{F_n(x),\ n = 1, 2, \ldots\}$ до $F(x)$ для всіх значень $x \in (-\infty;\infty)$ буде «*занадто сильним*», щоб можна було визначити корисний з практичної точки зору *тип збіжності* випадкових величин, що базується на основі збіжності функцій розподілу.

Приклад. Припустимо, що певна статистична ознака має дискретний розподіл і може приймати лише два значення: $x_1 = 0{,}5$ та $x_2 = 1$, при цьому обидва ці значення з однаково ймовірними. Для вимірювання значень цієї ознаки використовуються різні прилади, які дають трохи *занижений* результат. При цьому чим *більший* номер пристрою (n), тим *точніший* результат.

А саме, нехай випадкова величина ξ означає *реальне*, тобто *точне* значення ознаки для випадково взятої статистичної одиниці. Тоді результат (ξ_n) вимірювання значення ознаки приладом з номером (n) буде дорівнювати:

$$\xi_n = \xi - \frac{1}{n},\ n = 1, 2, \ldots$$

Покажемо, що:

$$\lim_{n\to\infty} P\{\xi_n < 0{,}5\} \neq P\{\xi < 0{,}5\};$$

і

$$\lim_{n\to\infty} P\{\xi_n < 1\} \neq P\{\xi < 1\}.$$

Згідно з умовою для будь-якого значення $n = 1, 2, \ldots$.:

$$P\{\xi < 0{,}5\} = 0;$$

а

$$P\left\{\xi < 0{,}5 + \frac{1}{n}\right\} = 0{,}5.$$

Отже для будь-якого значення $n = 1, 2, \ldots$:

$$P\{\xi_n < 0{,}5\} = P\left\{\xi - \frac{1}{n} < 0{,}5\right\} = P\left\{\xi < 0{,}5 + \frac{1}{n}\right\} = 0{,}5.$$

Тому

$$\lim_{n\to\infty} P\{\xi_n < 0{,}5\} = 0{,}5 \neq 0 = P\{\xi < 0{,}5\}.$$

Аналогічно, для будь-якого значення $n = 1, 2, \ldots$:

$$P\{\xi < 1\} = 0{,}5;$$

а

$$P\left\{\xi < 1 + \frac{1}{n}\right\} = 1,$$

Отже для будь-якого значення $n = 1, 2, \ldots$:

$$P\{\xi_n < 1\} = P\left\{\xi - \frac{1}{n} < 1\right\} = P\left\{\xi < 1 + \frac{1}{n}\right\} = 1.$$

Тому

$$\lim_{n\to\infty} P\{\xi_n < 1\} = 1 \neq 0{,}5 = P\{\xi < 1\}.$$

Іншими словами, нехай ξ – випадкова величина, що має розподіл:

$$\{(x_1 = 0{,}5;\ p_1 = 0{,}5);\ (x_2 = 1;\ p_2 = 0{,}5)\}.$$

$$F(x) = P\{\xi < x\},\ -\infty < x < \infty\ ,$$

функція розподілу випадкової величин ξ.

Визначимо послідовність випадкових величин ξ_n, $n = 1, 2, \ldots$ за формулою:

$$\xi_n = \xi - \frac{1}{n},\ n = 1, 2, \ldots\ .$$

Нехай

$$F_n(x) = P\{\xi_n < x\},\ -\infty < x < \infty,\ n = 1, 2, \ldots$$

означає функція розподілу випадкової величин ξ_n, $n = 1, 2, \ldots$.

Очевидно, що визначена таким чином послідовність випадкових величин ξ_n повинна збігатися до випадкової величини ξ для *будь-якого* типу збіжності, що має *практичний сенс*.

З іншого боку, якщо вимагати збіжність послідовності функцій розподілу $\{F_n(x),\ n = 1, 2, \ldots\}$ до $F(x)$ для *всіх* значень $x \in (-\infty;\infty)$, то побудована в прикладі послідовність випадкових величин ξ_n, $n = 1, 2, \ldots$ не буде збігатися до випадкової величини ξ:

- Як було показано, в точках $x_1 = 0{,}5$; та $x_2 = 1$, така збіжність відсутня

Причому збіжність послідовності $\{F_n(x),\ n = 1, 2, \ldots\}$ не буде виконуватись лише в точках $x_1 = 0{,}5$; та $x_2 = 1$, в яких гранична функція розподілу $F(x)$ має розриви.

Отже для того, щоб збіжність випадкових величин, яка ґрунтується на поведінці послідовності їх функцій розподілів, мала сенс, необхідно не враховувати точки можливих стрибків граничної функції розподілу $F(x)$, що і відображено у визначення слабкої збіжності.

(ЛЕКЦІЯ 9)

4. *Теорема Муавра-Лапласа.*

Теорему Муавра-Лапласа можна вважати логічним продовженням і уточненням теореми *Бернуллі*. Формулюється вона наступним чином.

Теорема. (*Муавра-Лапласа***).** Нехай $\mu_n(A)$ означає кількість появ події A в схемі випробувань *Бернуллі*. При цьому ймовірність події A дорівнює p. Тоді для довільних чисел a та b:

$$\lim_{n\to\infty} P\left(a < \frac{(\mu_n(A) - n\cdot p)}{\sqrt{n\cdot p\cdot(1-p)}} < b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx\,.$$

Іншими словами, розподіл випадкової величини

$$\frac{(\mu_n(A) - n\cdot p)}{\sqrt{n\cdot p\cdot(1-p)}}$$

збігається слабо при $n \to \infty$ до *стандартного нормального розподілу*:

$$\frac{(\mu_n(A) - n \cdot p)}{\sqrt{n \cdot p \cdot (1-p)}} \underset{n\to\infty}{\Rightarrow} N(0,\ 1).$$

Теорема *Бернуллі* стверджує, що різниці між частотою

$$\nu_n(A) = \frac{\mu_n(A)}{n}$$

випадкової події A, та ймовірністю p події A дорівнює прямує до нуля:

$$\frac{\mu_n(A)}{n} - p \underset{n\to\infty}{\to} 0\,.$$

Тобто

$$\frac{\mu_n(A) - n \cdot p}{n} \underset{n\to\infty}{\to} 0\,.$$

Використовуючи теорему *Муавра-Лапласа* це співвідношення можемо переписати наступним чином:

$$\frac{\mu_n(A) - n \cdot p}{n} = \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \cdot \frac{(\mu_n(A) - n \cdot p)}{\sqrt{n \cdot p \cdot (1-p)}} \approx \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \cdot N(0,\ 1) \approx 0.$$

А це означає, що швидкість збіжності різниці ($\nu_n(A) - p$) до нуля дорівнює $1\big/\sqrt{n}$.

Використовуючи властивості нормального розподілу, можемо записати наступне наближення:

$$\nu_n(A) \approx N\left(p; \frac{p \cdot (1-p)}{n}\right).$$