Теорія ймовірностей, статистика та ймовірнісні процеси

Лекція 6.

Розора Ірина Василівна

Київ, 2020

Зміст I

Зміст

В теорії йм. часто говорять про в.в.ξ із законом розподілу певного виду, не вказуючи при цьому ні йм. простір, ні саму функцію $\xi(\omega)$, що задає цю випадкову величину.
У зв'язку з цим розглянемо модель вибіркового ймовірнісного простору. В цій моделі за законом розподілу стандартним чином будується йм. простір і випадкова величина, закон розподілу якої співпадає із заданими законом.

Нехай закон розподілу задається множиною значень

$$X = \{x_1, x_2, \cdots\} \qquad \text{та ймов. } p_i,\ i \geq 1. \tag{1}$$

Побудуємо йм. простір і випадкову величину ξ, для якої (1) є законом розподілу.

Покладемо

$$\Omega = X = \{x_1, x_2, \cdots\}.$$

$F = 2^X$—множина усіх підмножин з X і

$$\forall A \in F \quad P(A) = \sum_{i: x_i \in A} p_i.$$

Означення

Трійка

$$(\Omega, F, P) = (Z, 2^X, P)$$

називається вибірковим ймовірнісним простором.

Для будь-якого $x \in X$ задамо в.в. ξ як тотожне переворення

$$\xi(x) = x.$$

Тоді $\xi(x)$ має закон розподілу, що співпадає з початковим.
За законом розподілу в.в. відновлюється неоднозначно.
Можуть існувати декілька в.в. з одним законом розподілу.

Зміст

Розподіл Бернуллі

Означення

В.В. ξ має розподіл Бернуллі з параметром p, $p \in (0,1)$, якщо вона приймає тільки два значення: 1 та 0 з ймовірностями p, $1 - p = q$ відповідно.

ξ	0	1
P	q	p

$$p + q = 1, \quad p > 0, \quad q > 0.$$

Позначення. $\xi \sim Be(p)$.

Біноміальний розподіл

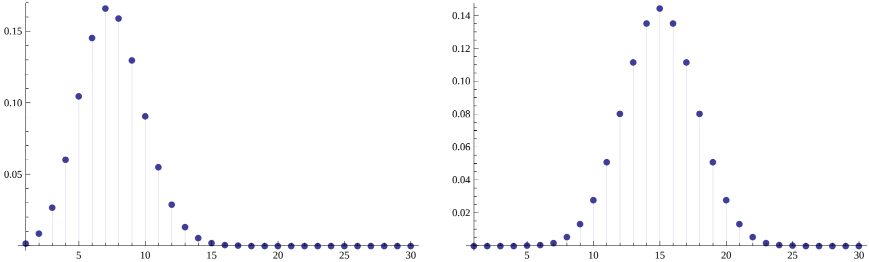
Означення

Кажуть, що в.в. ξ має біноміальний розподіл з параметрами n та p, $n \in \mathbb{N}$, $p \in (0,1)$,якщо

$$\mathbf{P}\{\xi = k\} = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n, \quad 0 < p < 1.$$

Позначення. $\xi \sim Bi(n, p)$.

Біноміальний розподіл



$n = 30,\ p = \frac{1}{4}$ $\qquad$ $n = 30,\ p = \frac{1}{2}$

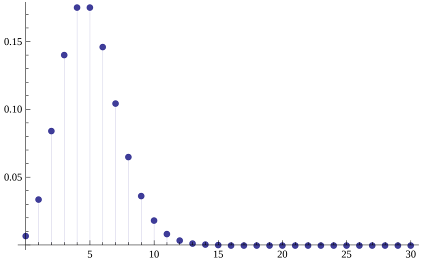
Рис. Біноміальні розподіли

Розподіл Пуассона

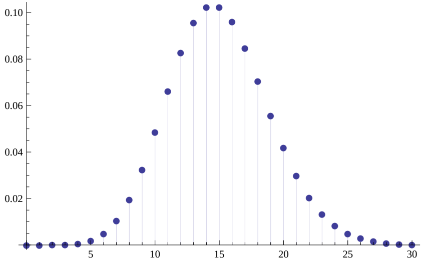
В.в. ξ має розподіл Пуассона з параметром $\lambda > 0$, якщо

$$\mathbf{P}(\xi = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots$$

Позначення. $\xi \sim \Pi(\lambda)$.



$\lambda = 5$



$\lambda = 15$

Рис. Розподіли Пуассона

Зміст

Означення

Послідовністю з n випробувань Бернуллі називається послідовність з n випробувань (експериментів) з властивостями

1. кожне випробування дихотомічне — закінчується одним з двох результатів: УСПІХ(У, 1) - НЕУСПІХ (Н, 0);
2. ймовірність У не залежить від номера випробування;
3. незалежність у сукупності.

Нехай p, $p \in (0,1)$ —ймовірність успіху в одному випробуванні. Тоді $q = 1 - p$ — ймовірність невдачі.
Позначимо через μ_n— в.в., яка дорівнює к-сті успіхів в n випробуваннях Бернуллі. Тоді ймовірність того, що в цій схемі відбулося рівно k успіхів (це так звана біноміальна ймовірність), дорівнює

$$P\{\mu_n = k\} = p_n(k) = C_n^k p^k q^{n-k}. \qquad (2)$$

Зміст

При великих значеннях n обчислювати йм. у формулі (2) важко. Тому у цих випадках можна використати наближені значення. Вважаємо, що від n залежить не тільки число успіхів, але й ймовірність успіху $p = p_n$.

Теорема (Теорема Пуассона)

Якщо

$$\lim_{n\to\infty} n \cdot p_n = \lambda > 0,$$

то для будь-якого $k = 0, 1, \cdots$

$$p_n(k) = P\{\mu_n = k\} \to_{n\to\infty} \frac{\lambda^k e^{-\lambda}}{k!}.$$

Доведення

Покладемо $\lambda_n = n \cdot p_n$. Тоді за умовою теореми $\lambda_n \to \lambda, \quad n \to \infty$.

$$p_n(k) = C_n^k p_n^k (1 - p_n)^{n-k} =$$

$$= \frac{n!}{(n-k)!k!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} =$$

$$= \frac{n(n-1)\dots(n-k+1)}{n^k} \frac{\lambda_n^k}{k!} \frac{(1-\lambda_n/n)^n}{(1-\lambda_n/n)^k} =$$

Доведення

$$= (1 - \frac{1}{n}) \dots (1 - \frac{k-1}{n}) \frac{\lambda_n^k}{k!} \frac{(1 - \lambda_n/n)^{-\frac{n}{\lambda_n} \cdot (-\lambda_n)}}{(1 - \lambda_n/n)^k} \to$$

$$e^{-\lambda} \frac{\lambda^k}{k!}, \quad n \to \infty.$$

Використали одну з чудових границь

$$\lim_{x \to 0} (1 + x)^{\frac{1}{x}} = e.$$

Приклад

Розглядається короткострокове страхування життя, коли страхова компанія сплачує 250 000 грн. у випадку смерті застрахованого протягом року, і нічого не виплачує в протилежному випадку. Припустимо, що на цих умовах застраховано 3000 клієнтів в одному віці. З таблиць смертності відомо, що йм. смерті протягом року для них дорівнює 0.003. Знайти ймовірність того, що компанія витратить не більше ніж 2 500 000 грн.

Розв'язання

Позначимо $250000 = 1$ у.о. Тоді $2500000 = 10$ у.о.
Тоді виплати страхової компанії за i-тим договором страхування ($i = \overline{1, n}$) можна описати такою в.в.

$$\xi_i = \begin{cases} 1, & \text{якщо відбувся страховий випадок;} \\ 0, & \text{якщо не відбувся страховий випадок.} \end{cases}$$

В.в.

$$\mu_n = \sum_{i=1}^{n} \xi_i$$

характеризує сумарні виплати стр. компанії за даним портфелем страхвання.
Причому всі власт. послідовноті Бернуллі виконуються.

Розв'язання

Позначимо $250000 = 1$ у.о. Тоді $2500000 = 10$ у.о.
Тоді виплати страхової компанії за i-тим договором страхування $(i = \overline{1, n})$ можна описати такою в.в.

$$\xi_i = \begin{cases} 1, & \text{якщо відбувся страховий випадок;} \\ 0, & \text{якщо не відбувся страховий випадок.} \end{cases}$$

В.в.

$$\mu_n = \sum_{i=1}^{n} \xi_i$$

характеризує сумарні виплати стр. компанії за даним портфелем страхвання.
Причому всі власт. послідовноті Бернуллі виконуються.

Розв'язування

Потрібно знайти

$$P\{\mu_n \leq 10\} = \sum_{k=0}^{10} P\{\mu_n = k\}.$$

Для підрахунку $P\{\mu_n = k\}$ використаємо Т. Пуассона. За умовою

$$n = 3000, \quad p_n = 0,003.$$

Тоді

$$\lambda \approx 3000 \cdot 0,003 = 9.$$

Отже,

Розв'язування

$$P\{\mu_n = k\} \approx \frac{9^k e^{-9}}{k!} \Rightarrow$$

$$P\{\mu_n \leq 10\} = \sum_{k=0}^{10} P\{\mu_n = k\} \approx \sum_{k=0}^{10} \frac{9^k e^{-9}}{k!} = 0,706.$$

Теорема (Локальна теорема Муавра-Лапласа)

Нехай в схемі незалежних виробувань Бернуллі $p \in (0,1)$—стала, $q = 1 - p$. Тоді при $n \to \infty$, $npq \to \infty$ і k таких, що

$$\left\| \frac{k - np}{\sqrt{npq}} \right\| := |x_k| \leq c,$$

де c— деяка стала, то

$$\lim_{n\to\infty} \frac{\sqrt{npq}p_n(k)}{\varphi(x_k)} = 1, \qquad (3)$$

$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ — щільність стандартного гауссового розподілу.

Зауваження

Теорема Пуассона діє у випадку, коли ймовірність успіху $p_n \to 0$. У схемі, коли вона віддалена від 0, варто застосовувати локальну теорему Муавра-Лапласа.

Зауваження

(3) означає асимптотичну еквівалентність

$$p_n(k) = C_n^k p^k q^{n-k} \approx \frac{1}{\sqrt{npq}}\varphi(x_k) = \frac{1}{\sqrt{npq}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x_k^2}{2}}.$$

Доведення.

Є два способи доведення теореми:

- використати формулу Стірлінга для наближеного значення $n!$ (без д-ня);
- дана теорема є наслідком центральної граничної теореми (ЦГТ), яку будемо доводити пізніше.

$\square$

Теорема (Інтегральна теорема Муавра-Лапласа)

Нехай виконуються умови локальної теореми Муавра-Лапласа. μ_n— кількість успіхів в n випробуваннях Бернуллі. Тоді для всіх $a < b$ при $n \to \infty$

$$P\{a < \frac{\mu_n - np}{\sqrt{npq}} < b\} \to \int_a^b \varphi(x)dx, \qquad (4)$$

де $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ — щільність стандартного гауссового розподілу.

Доведення

Позначимо

$$x_k = \frac{k - pn}{\sqrt{npq}} \quad \Rightarrow \quad \Delta_k = x_{k+1} - x_k = \frac{1}{\sqrt{npq}}.$$

Тоді, використовуючи схему незалежних випробувань Бернуллі,

$$P\{a < \frac{\mu_n - np}{\sqrt{npq}} < b\} = \sum_{k: a < x_k < b} p_n(k) =$$

$$= \sum_{k: a < x_k < b} (\sqrt{npq} p_n(k)) \frac{1}{\sqrt{npq}} \approx$$

За локальною Т. Муавра-Лапласа $\sqrt{npq} p_n(k) \approx \varphi(x_k)$.

$$\approx \sum_{k: a < x_k < b} \varphi(x_k) \Delta_k \to \int_a^b \varphi(x) dx.$$

Доведення

Позначимо

$$x_k = \frac{k - pn}{\sqrt{npq}} \quad \Rightarrow \quad \Delta_k = x_{k+1} - x_k = \frac{1}{\sqrt{npq}}.$$

Тоді, використовуючи схему незалежних випробувань Бернуллі,

$$P\{a < \frac{\mu_n - np}{\sqrt{npq}} < b\} = \sum_{k: a < x_k < b} p_n(k) =$$

$$= \sum_{k: a < x_k < b} (\sqrt{npq} p_n(k)) \frac{1}{\sqrt{npq}} \approx$$

За локальною Т. Муавра-Лапласа $\sqrt{npq} p_n(k) \approx \varphi(x_k)$.

$$\approx \sum_{k: a < x_k < b} \varphi(x_k) \Delta_k \to \int_a^b \varphi(x) dx.$$

Доведення

Позначимо

$$x_k = \frac{k - pn}{\sqrt{npq}} \quad \Rightarrow \quad \Delta_k = x_{k+1} - x_k = \frac{1}{\sqrt{npq}}.$$

Тоді, використовуючи схему незалежних випробувань Бернуллі,

$$P\{a < \frac{\mu_n - np}{\sqrt{npq}} < b\} = \sum_{k:a<x_k<b} p_n(k) =$$

$$= \sum_{k:a<x_k<b} (\sqrt{npq}p_n(k)) \frac{1}{\sqrt{npq}} \approx$$

За локальною Т. Муавра-Лапласа $\sqrt{npq}p_n(k) \approx \varphi(x_k)$.

$$\approx \sum_{k:a<x_k<b} \varphi(x_k)\Delta_k \to \int_a^b \varphi(x)dx.$$

Зауваження

Умову (4) можна замінити на

$$P\{a < \frac{\mu_n - np}{\sqrt{npq}} < b\} \to \Phi(b) - \Phi(a),$$

де

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx -$$

функція розподілу для стандартної гаауссової випадкової величини. Її значення можна знайти в стат. таблицях.

Приклад

Ймовірність деякого виробу бути бракованим дорівнює $p = 0,005$. Чому дорівнює ймовірність того, що серед 10 000 навмання взятих виробів число бракованих становитиме 40? Розв'язати задачу трьома способами.

Розв'язання

$$n = 10000, \quad p = 0,005, \quad q = 0,995.$$

1. Використаємо формулу для к-сті успіхів у випробуваннях Бернуллі.

$$p_{10000}(40) = C_{10000}^{40} 0,005^{40} 0,995^{10000-40} = 0,0197.$$

2. За теоремою Пуассона $\lambda \approx np = 10000 \cdot 0,005 = 50$

$$p_{10000}(40) \approx \frac{e^{-50}(50)^{40}}{40!} \approx 0,021.$$

Розв'язання

$$n = 10000, \quad p = 0,005, \quad q = 0,995.$$

1. Використаємо формулу для к-сті успіхів у випробуваннях Бернуллі.

$$p_{10000}(40) = C_{10000}^{40} 0,005^{40} 0,995^{10000-40} = 0,0197.$$

2. За теоремою Пуассона $\lambda \approx np = 10000 \cdot 0,005 = 50$

$$p_{10000}(40) \approx \frac{e^{-50}(50)^{40}}{40!} \approx 0,021.$$

Розв'язання

$$n = 10000, \quad p = 0,005, \quad q = 0,995.$$

1. Використаємо формулу для к-сті успіхів у випробуваннях Бернуллі.

$$p_{10000}(40) = C_{10000}^{40} 0,005^{40} 0,995^{10000-40} = 0,0197.$$

2. За теоремою Пуассона $\lambda \approx np = 10000 \cdot 0,005 = 50$

$$p_{10000}(40) \approx \frac{e^{-50}(50)^{40}}{40!} \approx 0,021.$$

Розв'язання

3) За локальною Т. Муавра-Лапласа

$$\sqrt{npq} = \sqrt{10000 \cdot 0,005 \cdot 0,995} \approx 7,05.$$

$$x_k = \frac{k - np}{\sqrt{npq}} = \frac{40 - 50}{7,05} = -1,42.$$

$$p_n(k) \approx \frac{1}{\sqrt{npq}}\varphi(x_k) = \frac{1}{7,05\sqrt{2\pi}}e^{-\frac{1,42^2}{2}} = 0,0206.$$

Приклад

Підкидається гральний кубик 720 разів. Знайти йм. того, що шістка ("6") з'явиться від 90 до 150 разів.

Застосуємо інтегральну Т. Муавра-Лапласа.

$$p = \frac{1}{6}, \quad q = \frac{5}{6}, \quad n = 720.$$

$$np = 120, \quad \sqrt{npq} = \sqrt{720 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 10.$$

Приклад

Підкидається гральний кубик 720 разів. Знайти йм. того, що шістка ("6") з'явиться від 90 до 150 разів.

Застосуємо інтегральну Т. Муавра-Лапласа.

$$p = \frac{1}{6}, \quad q = \frac{5}{6}, \quad n = 720.$$

$$np = 120, \quad \sqrt{npq} = \sqrt{720 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 10.$$

Приклад

Підкидається гральний кубик 720 разів. Знайти йм. того, що шістка ("6") з'явиться від 90 до 150 разів.

Застосуємо інтегральну Т. Муавра-Лапласа.

$$p = \frac{1}{6}, \quad q = \frac{5}{6}, \quad n = 720.$$

$$np = 120, \quad \sqrt{npq} = \sqrt{720 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 10.$$

Нехай μ_{720} — кількість "6" у 720 підкиданнях. Потрібно підрахувати

$$P\{90 \le \mu_{720} \le 150\} = P\{\frac{90 - np}{\sqrt{npq}} \le \frac{\mu_{720} - np}{\sqrt{npq}} \le \frac{150 - np}{\sqrt{npq}}\} =$$

$$= P\{\frac{90 - 120}{10} \le \frac{\mu_{720} - np}{\sqrt{npq}} \le \frac{150 - 120}{10}\} =$$

$$= P\{-3 \le \frac{\mu_{720} - np}{\sqrt{npq}} \le 3\} \approx \Phi(3) - \Phi(-3) = 0,997.$$

Зміст

Розглянемо ймов. простір (Ω, F, P), на якому задана д.в.в. $\xi : \Omega \to X = \{x_1, x_2, \dots\}$, $x_k \in \mathbf{R}$—різні.
Позначимо

$$A_k = \{\omega \in \Omega : \xi(\omega) = x_k\}, \quad p_k = P(A_k).$$

Тоді д.в.в зображається як

$$\xi(\omega) = \xi = \sum_{k=1}^{\infty} x_k \cdot \mathbf{I}_{A_k}(\omega),$$

де індикаторна функція події $A_k \in \mathcal{F}$:

$$\mathbf{I}_{A_k}(\omega) = \begin{cases} 1, & \text{коли } \omega \in A_k; \\ 0, & \text{коли } \omega \notin A_k. \end{cases}$$

Означення

Д.в.в. ξ називається сумовною, якщо ряд

$$\sum_{k=1}^{\infty} |x_k| p_k < \infty.$$

Означення

Нехай ξ — сумовна д.в.в. з множиною значень $X = \{x_1, x_2 \ldots\}$ та розподілом $p_k = P(A_k) = P\{\xi = x_k\}$, $k \geq 1$. Математичним сподіванням д.в.в. ξ називається число

$$M\xi = E\xi = \sum_{k=1}^{\infty} x_k \cdot p_k.$$

Зауваження

При визначенні д.в.в. для нас не важливий порядок нумерації її можливих значень, тому природно вважати, що і $M\xi$ не залежить від порядку доданків. А це можливо при абсолютній збіжності ряду.

Означення (альтернативне означення)

Математичним сподіванням в.в. ξ, заданої на дискретному йм. просторі (Ω, F, P), будемо називати суму ряду

$$M\xi = E\xi = \sum_{\omega \in \Omega}^{\infty} \xi(\omega) \cdot P(\omega).$$

Зауваження

Якщо в.в. ξ невід'ємна, то ряд в $M\xi$ або абсолютно збігається, або розбігається до $+\infty$. В останньому випадку вважаємо, що $M\xi = +\infty$.

Зауваження

Якщо в точках з координатами $x_1, x_2 \dots$ знаходяться тіла з масами $p_1, p_2 \dots$ відповідно, то $M\xi$— центр ваги системи мас.

Теорема (про обчислення м.сп.функції від д.в.в)

Нехай $\xi : \Omega \to X = \{x_1, x_2 \cdots\}$*—д.в.в., а* g*—довільна числова функція, що визначена на* X*. Тоді*

$$Mg(\xi) = \sum_{k=1}^{\infty} g(x_k)p_k = \sum_{k=1}^{\infty} g(x_k)P\{\xi = x_k\}.$$

Властивості матем. сподівання для д.в.в.

ξ, η — сумовні д.в.в.

- **йм. індикатора:** $M\mathbf{I}_A(\omega) = P(A)$
- **нормованість:** $Mc = c$ для довільної сталої $c \in \mathbf{R}$;
- **невід'ємність:** якщо $\xi \geq 0$, то $M\xi \geq 0$;
- **однорідність:** $M(c\xi) = cM\xi$ для довільної сталої $c \in \mathbf{R}$;
- **адитивність:** $M(\xi + \eta) = M\xi + M\eta$;
- **монотонність:** якщо $\xi \leq \eta$, то $M\xi \leq M\eta$;
- **додатність:** якщо в.в. $\xi \geq 0$ та $M\xi = 0$, то $\xi = 0$;
- **інваріантність:** якщо $\xi = \eta$, то $M\xi = M\eta$;
- **опуклість:** $|M\xi| \leq M|\xi|$.

Доведення I

- **йм. індикатора.**

$$\mathbf{I}_A(\omega) = \begin{cases} 1, & \text{коли } \omega \in A; \\ 0, & \text{коли } \omega \notin A. \end{cases}$$

$$M\mathbf{I}_A(\omega) = 1 \cdot P(A) + 0 \cdot P(\overline{A}) = P(A).$$

- **нормованість.** Вправа.

Доведення II

- **невід'ємність.**

$$\xi \geq 0 \quad \Leftrightarrow \forall k \geq 1\, x_k \geq 0 \Rightarrow$$

$$M\xi = \sum_{k=1}^{\infty} x_k p_k \geq 0.$$

- **однорідність.** Вправа.
- **адитивність.** Покажемо, що $M(\xi + \eta) = M\xi + M\eta$. Нехай розподіли ξ, η відповідно дорівнюють

Доведення III

ξ	x_1	x_2	x_3	$\cdots$
P	p_1	p_2	p_3	$\cdots$

η	y_1	y_2	y_3	$\cdots$
P	q_1	q_2	q_3	$\cdots$

Через z_k позначимо усі різні значення $(x_i + y_j)$ і

$$I_k = \{(i,j) | x_i + y_j = z_k\}.$$

Тоді

$$P\{\xi + \eta = z_k\} = \sum_{(i,j)\in I_k} p_i q_j.$$

$$M\xi + M\eta = \sum_{i=1}^{\infty} x_i p_i + \sum_{j=1}^{\infty} y_j q_j =$$

Доведення IV

$$= \sum_{i=1}^{\infty} x_i p_i \left(\sum_{j=1}^{\infty} q_j \right) + \sum_{j=1}^{\infty} y_j q_j \left(\sum_{i=1}^{\infty} p_i \right) =$$

$$= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (x_i p_i q_j + y_j q_j p_i) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (x_i + y_j)\, p_i q_j =$$

$$= \sum_{k=1}^{\infty} z_k \left(\sum_{(i,j) \in I_k} p_i q_j \right) = \sum_{k=1}^{\infty} z_k P\{\xi + \eta = z_k\} = M(\xi + \eta).$$

Доведення V

- **монотонність.** Доведемо, що якщо $\xi \leq \eta$, то $M\xi \leq M\eta$. Оскільки $\eta - \xi \geq 0$, то з власт. невід'ємності ($M(\eta - \xi) \geq 0$), однорідності та адитивності випливає

$$M(\eta - \xi) = M(\eta) + M(-\xi) = M\eta - M\xi \geq 0.$$

ПИТАННЯ?