Лекція 4.
Схема Бернуллі. Граничні теореми

План лекції

Питання, що розглядаються:

Схема Бернуллі, непостійна ймовірність появи подій в випробуваннях, твірна функція, наймовірніше число настання подій в схемі Бернуллі, граничні теореми, теорема Пуассона, потік подій, інтенсивність потоку, простий потік, локальна теорема Муавра-Лапласа, інтегральна теорема Муавра-Лапласа, функція Гауса, функція Лапласа, інтеграл Лапласа.

1. Схема Бернуллі

Схемою Бернуллі називається серія повторних незалежних випробувань, в кожному з яких ця подія А має одну і ту ж ймовірність $P(A) = p$, не залежну від номера випробування.

Таким чином, в схемі Бернуллі для кожного випробування є тільки два результати: подія A (успіх), ймовірність якої $P(A) = p$ і подія $\overline{A}$ (невдача), ймовірність якої $P(\overline{A}) = q = 1 - p$.

Імовірність, що подія A наступить в m випробуваннях, визначається за формулою Бернуллі

$$P_n(m) = C_n^m p^m q^{n-m}.$$

*Приклад***.** Ймовірність враження цілі стрільцем при одному пострілі дорівнює 0,8. Знайти ймовірність, що при 10 пострілах ціль вражається 8 разів.

Розвязання. $$P_{10}(8) = \frac{10!}{2!\,8!} \cdot (0{,}8)^8 \cdot (0{,}2)^2 = 0{,}302.$$

2. Випадок непостійної ймовірності появи події в випробуваннях

Ми припускали, що ймовірність настання події в кожному з випробувань постійна. На практиці часто доводиться зустрічатися із складнішим випадком, коли випробування проводяться в неоднакових умовах, і ймовірність події від випробування до випробуванняу змінюється. Наприклад, проводиться серія пострілів при дальності, що змінюється**.**

Спосіб обчислення ймовірності заданого числа появ подій в таких умовах дає загальна теорема про повторення випробувань.

Нехай проводиться n незалежних випробувань, в кожному з яких може з'явитися або не з'явитися деяка подія A, причому ймовірність появи цієї події в i-му випробуванні дорівнює p_i, а ймовірність її непояви відповідно $q_i = 1 - p_i$ $(i = 1, \ldots, n)$. Потрібно знайти ймовірність $P_{m,n}$ того, що в результаті n випробувань подія A з'явиться рівно m разів.

Розвязання цієї задачі проводиться за допомогою так званої твірної функції, що має вигляд:

$$\varphi_n(z) = \prod_{i=1}^{n} (q_i + p_i z).$$

Для того, щоб знайти ймовірність появи події A рівно m разів в серії n випробувань, досить перемножити співмножники у твірній функції. Коефіцієнт при членові z^m і дасть шукану ймовірність.

*Приклад***.** Проводиться 4 незалежні постріли по одній і тій же цілі з різних відстаней. Ймовірність попадання при цих пострілах дорівнює відповідно $p_1 = 0{,}1$ $p_2 = 0{,}2$ $p_3 = 0{,}3$ $p_4 = 0{,}4$. Знайти ймовірність трьох попадань.

Розвязання: Складемо твірну функцію

$$\varphi_n(z) = \prod_{i=1}^{n}(q_i + p_i\, z) = (0{,}9 + 0{,}1z)\cdot(0{,}8 + 0{,}2z)\cdot(0{,}7 + 0{,}3z)\cdot(0{,}6 + 0{,}4z) =$$

$$= 0{,}302 + 0{,}440z + 0{,}215z^2 + 0{,}040z^3 + 0{,}002z^4.$$

Звідси ймовірність трьох попадань дорівнює 0,040. Легко знайти і ймовірність жодного, одного, двох і чотирьох попадань, виписуючи коефіцієнти при z^0, z^1, z^2 і z^4.

3. Наймовірніше число настання подій в схемі Бернуллі

Число настання події A називається ***наймовірнішим***, якщо воно має найбільшу ймовірність в порівнянні з ймовірністю настання A будь-якої іншої кількість разів.

Теорема. Наймовірніше число настання події A в n незалежних випробуваннях знаходиться між числами $np - q$ і $np + p$.

Слід зазначити, що наймовірніших чисел може бути два або одне залежно від того, є $np + p$ цілим числом або ні. Якщо це число неціле, то наймовірніше число $m_0 = [np + p]$ (ціла частина), інакше є два значення $m_0' = np - q$ і $m_0'' = np + p$

Приклад. Скільки разів потрібно кинути гральний кубик, щоб наймовірніше число появ трійки дорівнювало 55?

Розвязання. За умовою задачі m_0=55, ймовірність появи трійки дорівнює $p = \frac{1}{6}$, отже, ймовірність непояви трійки дорівнює $q = \frac{5}{6}$. Складемо нерівність

$$n\frac{1}{6} - \frac{5}{6} \le 55 \le n\frac{1}{6} + \frac{1}{6}$$

Отримали лінійну систему нерівностей:

$$n\frac{1}{6} - \frac{5}{6} \le 55, \qquad n - 5 \le 330, \qquad n \le 335,$$
$$n\frac{1}{6} + \frac{5}{6} \ge 55 \qquad n + 1 \ge 330 \qquad n \ge 329$$

Отже, кубик необхідно кинути від 329 до 335 разів.

4. Теорема Пуассона

Якщо число випробувань велике, формулу Бернуллі застосовувати незручно. В цьому випадку можна застосовувати наближені формули, точність яких збільшується із зростанням n.

***Теорема*:** Якщо ймовірність p настання події A в кожному випробуванні постійна і мала, а число незалежних випробувань n досить велике, то ймовірність того, що подія A наступить m разів, наближено дорівнює

$$P_n(m) = \frac{\lambda^m}{m!} e^{-\lambda},$$

де $\lambda = np$.

Доведення. Введемо позначення $\lambda = np$, виразимо звідси $p = \frac{\lambda}{n}$ і підставимо цей вираз у формулу Бернуллі :

$$P_n(m) = C_n^m \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} =$$

$$= \frac{\lambda^m}{n^m} \frac{n(n-1)(n-2)...(n-m+1)}{m!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} =$$

$$= \frac{\lambda^m}{m!} \frac{n-1}{n} \frac{n-2}{n} ... \frac{n-m+1}{n} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} =$$

$$= \frac{\lambda^m}{m!} \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)...\left(1 - \frac{m-1}{n}\right)\left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m}.$$

При $m << n$ усі вирази в дужках, за винятком передостаннього, можна прийняти рівними одиниці, тобто

$$P_n(m) = \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^n.$$

При $n \to \infty$

$$\lim_{n \to \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda},$$

тому

$$P_n(m) = \frac{\lambda^n}{m!} e^{-\lambda},$$

що і вимагалося довести.

Приклад. Ткаля обслуговує 1000 веретен. Ймовірність обриву нитки на одному веретені протягом однієї хвилини дорівнює 0,005. Знайти ймовірність того, що протягом однієї хвилини обрив нитки відбудеться на 7 веретенах.

Розвязання. Так як кількість випробувань велика (n = 1000), а ймовірність окремого випробування дуже мала (p = 0,005), то для обчислення шуканої ймовірності скористаємося формулою Пуассона. Параметр розподілу λ= 1000×0,005 = 5, тоді шукана ймовірність дорівнює:

$$P_{1000}(7) = \frac{5^7}{7!} e^{-5} = \frac{78125}{5040} \cdot 0\,006738 = 0{,}1044.$$

5. Поняття потоку подій

Формула Пуассона знаходить застосування в теорії масового обслуговування. Вона може розглядатися як математична модель простого потоку подій з інтенсивністю $\lambda = np$. Параметр $\lambda = np$ є при цьому середнє число успіхів.

Потоком подій *називають послідовність подій, які настають у випадкові моменти часу.*

Інтенсивністю потоку λ називають середнє число подій за одиницю часу

Простим **(*пуассонівским*)** називають потік подій, який має властивості стаціонарності, відсутності післядії і ординарності.

Властивість стаціонарності характеризується тим, що ймовірність $P_t(m)$ появи m подій на будь-якому проміжку часу залежить тільки від числа m і від тривалості проміжку часу t і не залежить від початку його відліку.

Властивість відсутності післядії характеризується тим, що ймовірність появи m подій на будь-якому проміжку часу не залежить від того, з'являлися або не з'являлися події в моменти часу, передуючі початку даного проміжку, тобто передісторія потоку не позначається на ймовірності появи подій.

Властивість ординарності характеризується тим, що поява двох і більше подій за малий проміжок часу малоймовірна в порівнянні з ймовірністю появи тільки однієї події.

Якщо інтенсивність простого потоку λ відома, то ймовірність появи m подій за час t визначається формулою

$$P_t(m) = \frac{(\lambda t)^m \cdot e^{-\lambda t}}{m!}.$$

Приклад. Середнє число замовлень таксі, що поступають на диспетчерський пункт за одну хвилину, дорівнює трьом. Знайти ймовірність того, що за 2 хвилини поступить 4 виклики.

Розвязання. Підставляючи в наведену вище формулу $\lambda = 3,\ t = 2,\ m = 4$, отримаємо:

$$P_2(4) = \frac{(3 \cdot 2) \cdot e^{-3 \cdot 2}}{4!} = 0{,}135.$$

6. Локальна теорема Муавра-Лапласа

Лапласом була отримана важлива наближена формула для ймовірності $P_n(m)$ появи події A рівно m разів, за умови, що n досить велике. На відміну від формули Пуассона тут немає обмеження на мализну величини p в окремому випробуванні, тобто область застосовності формули Лапласа ширша.

Теорема. Ймовірність того, що в умовах схеми Бернуллі подія A при n випробуваннях з'явиться рівно m разів, виражається наближеною формулою Лапласа

$$P_n(m) \approx \frac{1}{\sqrt{2\pi npq}} e^{\frac{-t^2}{2}}, \text{ де } \qquad t = \frac{m-np}{\sqrt{npq}}.$$

Формулу Лапласа іноді називають асимптотичною формулою, оскільки доведено, що відносна помилка формули Лапласа прямує до нуля при $n \to \infty$.

7. Інтегральна (глобальна) теорема Муавра-Лапласа

Інтегральна теорема Муавра-Лапласа містить наближену формулу для ймовірності $P_n(m_1, m_2)$ того, що подія A з'явиться не менше m_1 разів і не більше m_2 разів.

Теорема. Ймовірність того, що подія A з'явиться в n випробуваннях від m_1 до m_2 разів, приблизно дорівнює визначеному інтегралу

$$P_n(m_1, m_2) \approx \int_{t_{m_1}}^{t_{m_2}} \varphi_0(t)dt = \frac{1}{\sqrt{2\pi}} \int_{t_{m_1}}^{t_{m_2}} e^{-\frac{t^2}{2}} dt,$$

де: $\quad t_{m_1} = \dfrac{m_1 - np}{\sqrt{npq}}; \;\; t_{m_2} = \dfrac{m_2 - np}{\sqrt{npq}}.$

Доведення. На підставі теореми додавання ймовірностей несумісних подій

$$P_n(m_1, m_2) = \sum_{m=m_1}^{m_2} P_n(m).$$

Або, використовуючи локальну теорему Лапласа

$$P_n(m_1, m_2) \approx \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{npq}} \varphi_0(t_m).$$

Введемо позначення $\Delta t_m = t_{m+1} - t_m = \dfrac{(m+1)-np}{\sqrt{npq}} - \dfrac{m-np}{\sqrt{npq}} = \dfrac{1}{\sqrt{npq}}.$

Запишемо P_n у вигляді $P_n(m_1, m_2) \approx \sum\limits_{m=m_1}^{m_2} \varphi_0(t_m)\Delta t_m$.

Очевидно, при $n \to \infty$ величина $\Delta t_m \to 0$, і остання сума прямує до визначеного інтегралу:

$$P_n(m_1, m_2) \approx \int_{t_{m_1}}^{t_{m_2}} \varphi_0(t)dt = \frac{1}{\sqrt{2\pi}} \int_{t_{m_1}}^{t_{m_2}} e^{-\frac{t^2}{2}} dt,$$

що і вимагалося довести.

Введемо стандартний інтеграл Лапласа (функцію Лапласа):

$$\Phi_0(x) = \int_0^x \varphi_0(t)dt = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt,$$

який, очевидно, є первісною функції Гауса

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Тоді на підставі формули Ньютона - Лейбніца можна записати

$$P_m(m_1, m_2) \approx \Phi_0(t_{m_2}) - \Phi_0(t_{m_1}).$$

Значення функцій $\Phi_0(x)$ і $\varphi_0(x)$ зазвичай знаходяться з таблиць, причому таблиці зазвичай дані лише для невідємних значень x, оскільки $\varphi_0(x)$ - парна функція, а $\Phi_0(x)$ - непарна.

Приклад. Ймовірність присутності студента на лекції дорівнює 0,8. Знайти ймовірність того, що з 100 студентів на лекції будуть присутні не менше 75 і не більше 90.

Розв'язання. Так як кількість випробувань велика (n = 100), то для знаходження ймовірності, що подія A зявиться від 75 до 90 разів, скористаємося інтегральною теоремою Лапласа. Визначимо аргументи інтегральної функції Лапласа x_1 і x_2:

$$x_1 = \frac{75 - 100 \cdot 0{,}8}{100 \cdot 0{,}8 \cdot 0{,}2} = -1{,}25, \qquad x_2 = \frac{90 - 100 \cdot 0{,}8}{100 \cdot 0{,}8 \cdot 0{,}2} = 2{,}5.$$

Враховуючи, що функція $\Phi(x)$ є непарною, тобто $\Phi(-x) = -\Phi(x)$ за таблицею значень інтегральної функції Лапласа знаходимо:

$\Phi(-1{,}25) = -\Phi(1{,}25) = -0{,}39435$ і $\Phi(2{,}5) = 0{,}49379$,

тоді

$P_{100}(75 \le k \le 90) = \Phi(2{,}5) - \Phi(-1{,}25) = 0{,}49379 + 0{,}39435 = 0{,}888.$

Питання для самоперевірки

1. Що називають схемою Бернуллі?
2. Записати формулу Бернуллі.
3. Що таке твірна функція?
4. Коли застосовується твірна функція?
5. В яких межах лежить наймовірніше число появи подій в незалежних випробуваннях?
6. Записати формулу Пуассона.
7. Коли застосовують формулу Пуассона?
8. Що називають потоком подій?
9. Що таке інтенсивність потоку?
10. Що називають простим потоком подій?
11. Сформулювати властивості простого потоку подій.
12. Записати локальну теорему Лапласа.
13. Записати інтегральну теорему Лапласа.
14. Записати функції Гаусса та Лапласа.