3 Повторні незалежні експерименти

3.1 Формула Бернуллі

Припустимо, що якийсь експеримент може повторюватися в незмінних умовах стільки разів, скільки захочеться, і ці повторення не залежать один від одного. В даному випадку говорять про проведення послідовності незалежних експериментів. Незалежність випробувань слід розуміти в тому сенсі, що будь-які події, які можуть статися в результаті, незалежні в сукупності.

Найпростішою вважається послідовність незалежних випробувань, в кожному з яких можливі тільки 2 результату: успіх – У (1) і невдача – Н (0). Послідовність незалежних випробувань з двома результатами називається *схемою незалежних випробувань Бернуллі*.

Теорема 3.1 (Бернуллі) Імовірність того, що в n повторних незалежних випробуваннях, в кожному з яких імовірність появи випадкової події A рівна p ($0<p<1$), дана подія наступить (відбудеться) рівно k разів знаходиться за формулою Бернуллі:

$$P_n(k) = C_n^k p^k q^{n-k}, 0 \le k \le n,$$

де $q = 1 - p$ – імовірність непояви події А в кожному випробуванні.

Оскільки розглянуті повторні події $0 \le k \le n$ утворюють повну групу подій, то $\sum_{k=0}^{n} P_n(k) = 1$. Аналогічний результат можна отримати з формули біному Ньютона :

$$\sum_{k=0}^{n} P_n(k) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1.$$

Імовірності $P_n(k)$ при фіксованому n спочатку ростуть при збільшенні числа k від 0 до деякого значення k_0, а потім зменшуються при зменшенні числа k від k_0 до n.

Число успіхів k_0, при якому ймовірності $P_n(k)$ досягають максимуму, називається *найбільшою кількістю успіхів* і визначається з подвійної нерівності:

$$np - q \le k_0 \le np - q + 1 = np + p.$$

При цьому:

– якщо число $np - q$ не ціле, то існує одне найімовірніше число успіхів: $k_0 = [np - q] + 1 = [np + p]$;

– якщо число $np - q$ є цілим числом, то є два найімовірніших числа успіхів: $k'_0 = np - q$ і $k''_0 = np - q + 1 = np + p$;

– якщо число np є цілим числом, то $k_0 = np$.

Приклад 1 Що більш імовірно: виграти 3 гри з 4 або 5 з 8 проти рівного суперника (нічиї не враховуються)?

Розв'язання. У цьому прикладі ми порівнюємо дві ймовірності $P_4(3)$ і $P_8(5)$, коли $p = q = \frac{1}{2}$. Оскільки, $P_4(3) = C_4^3 \frac{1}{2^4} = 4 \cdot \frac{1}{2^4}$, а $P_8(5) = C_8^5 \frac{1}{2^8} = \frac{7}{2} \cdot \frac{1}{2^4}$, то $P_4(3) > P_8(5)$, тобто перемога в 3 іграх з 4 є більш імовірною.

Приклад 2 По цілі робиться 3 самостійних постріли. Імовірність ураження різними пострілами однакова і дорівнює $p = 0{,}9$. Яка ймовірність: а) промаху; б) одного влучання; в) двох влучань; г) трьох влучань? Розв'яжіть задачу, якщо ймовірність влучання при різних пострілах різна: $p_1 = 0{,}7, p_2 = 0{,}8, p_3 = 0{,}9$.

Розв'язання. У цьому випадку $n = 3,\ p = 0{,}9.\ q = 0{,}1$. Використовуючи формулу Бернуллі, знаходимо:

а) $P_3(0) = C_3^0 \cdot 0,9^0 \cdot 0,1^3 = 0{,}001$ – ймовірність трьох промахів;

б) $P_3(1) = C_3^1 \cdot 0,9^1 \cdot 0,1^2 = 0{,}027$ – ймовірність одного влучання;

в) $P_3(2) = C_3^2 \cdot 0,9^2 \cdot 0,1^1 = 0{,}243$ – ймовірність двох влучань;

г) $P_3(3) = C_3^3 \cdot 0,9^3 \cdot 0,1^0 = 0{,}729$ – ймовірність трьох влучань.

Якщо ймовірність влучання при різних пострілах різна, то

а) $P_3(0) = 0{,}3 \cdot 0{,}2 \cdot 0{,}1 = 0{,}006$ – ймовірність трьох промахів;

б) $P_3(1) = 0{,}7 \cdot 0{,}2 \cdot 0{,}1 + 0{,}3 \cdot 0{,}8 \cdot 0{,}1 + 0{,}3 \cdot 0{,}2 \cdot 0{,}9 = 0{,}092$ – ймовірність одного влучання;

в) $P_3(2) = 0{,}7 \cdot 0{,}8 \cdot 0{,}1 + 0{,}3 \cdot 0{,}8 \cdot 0{,}9 + 0{,}7 \cdot 0{,}2 \cdot 0{,}9 = 0{,}398$ – ймовірність двох влучань;

г) $P_3(3) = 0{,}7 \cdot 0{,}8 \cdot 0{,}9 = 0{,}504$ – ймовірність трьох влучань.

3.2 Граничні теореми в схемі Бернуллі

Використовувати формулу Бернуллі при великих значеннях k і n дуже складно, так як вона пов'язана з громіздкими обчисленнями. Так, якщо $n = 200$, $k = 116$, $p = 0{,}72$ формула Бернуллі набуває вигляду $P_{200}(116) = C_{200}^{116} \cdot 0{,}72^{116} \cdot 0{,}28^{84}$. Обчислити результат практично неможливо. Обчислення $P_n(k)$ викликає труднощі і для малих значеннях $p(q)$. Виникає необхідність пошуку наближених формул для розрахунку $P_n(k)$, які забезпечують необхідну точність. Такі формули дають нам граничні теореми; вони містять так звані асимптотичні формули, які при великій кількості випробувань дають довільно малу відносну похибку. Розглянемо три граничні теореми, що містять асимптотичні формули для розрахунку біноміальної ймовірності $P_n(k)$ при $n \to \infty$.

Теорема Пуассона

Якщо кількість випробувань нескінченно збільшується $(n \to \infty)$ і ймовірність p настання події A в кожному випробуванні нескінченно зменшується $(p \to 0)$, але така, що їх добуток np є постійною величиною $(np = a = const)$, то ймовірність $P_n(k)$ задовольняє граничну рівність

$$\lim_{n\to\infty} P_n(k) = \frac{a^k e^{-a}}{k!}$$

– асимптотична формула Пуассона.

З асимптотичної формули Пуассона для великих n і малих p випливає наближена формула Пуассона

$$P_n(k) \approx \frac{a^k e^{-a}}{k!}, a = np,\ \ k = 0,1,2,\dots .$$

Ця формула використовується, коли ймовірність успіху події $p = const$ вкрай мала, тобто сам успіх (поява A) – подія рідкісна (наприклад, виграш автомобіля за лотерейним квитком), але кількість випробувань n велика, середня кількість успіхів $np = a$ незначна. Приблизна формула Пуассона зазвичай використовується, коли $n \geq 50$, $np \leq 10$.

Приклад 3 Приватна винороббня Одеси відправила до Києва 1500 пляшок вина. Імовірність того, що пляшка може розбитися в дорозі, становить 0,002. Знайдіть ймовірність того, що в дорозі буде розбито не більше 4 пляшок (подія A).

Розв'язання. Шукана ймовірність така

$$P_{1500}(0) + P_{1500}(1) + P_{1500}(2) + P_{1500}(3) + P_{1500}(4).$$

Оскільки $n = 1500$, $p = 0{,}002$, то $a = np = 3$. ймовірність події A буде знайдена за допомогою наближеної формули Пуассона:

$$P(A) = P_{1500}(k \leq 4) \approx \frac{3^0 e^{-3}}{0!} + \frac{3^1 e^{-3}}{1!} + \frac{3^2 e^{-3}}{2!} + \frac{3^3 e^{-3}}{3!} + \frac{3^4 e^{-3}}{4!} \approx 0{,}815.$$

Формула Пуассона знаходить застосування в теорії масового обслуговування, її можна вважати математичною моделлю найпростішого потоку подій.

Потік подій – це послідовність подій, які відбуваються в випадковий час (наприклад, потік відвідувачів перукарні, потік дзвінків на телефонній станції, потік збоїв елементів, потік обслуговуваних абонентів і т.д.).

Потік подій з властивостями стаціонарності, буденності і відсутності наслідків називається *найпростішим* (*Пуассоновським*) потоком.

Властивість стаціонарності означає, що ймовірність настання k подій у відрізку часу довжини τ залежить тільки від його довжини (тобто не залежить від початку його відліку). Тому середня кількість подій, що з'являються за одиницю часу, так звана інтенсивність λ потоку, є постійною величиною: $\lambda(t) = \lambda$.

Властивість звичайності (буденності, посередності) означає, що подія з'являється не в групах, а поодинці. Іншими словами, ймовірність того, що за невеликий проміжок часу Δt відбудеться більше однієї події, мізерно мала в порівнянні з ймовірністю настання тільки однієї події (наприклад, потік човнів, що наближаються до причалу, звичайний).

Властивість відсутності наслідків означає, що ймовірність настання k подій, що відбуваються в будь-який момент проміжку часу τ, не залежить від того, скільки подій з'явилося на будь-якому іншому ділянці, що не перекривається (мовляв: «майбутнє» потоку не залежить від «минулого», наприклад, потоку людей, що заходять в супермаркет).

Ймовірність настання k подій найпростішого потоку впродовж часу t визначається за формулою Пуассона

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

Приклад 4 Телефонна станція обслуговує 2000 абонентів. Імовірність зателефонувати будь-якому абоненту протягом години становить 0,003. Яка ймовірність того, що протягом години зателефонують 5 абонентів?

Розв'язання. Середня кількість абонентів, які дзвонили протягом години, становить $2000 \cdot 0{,}003 = 6$ $(a = np = \lambda t)$. Отже, $p_5 = \frac{6^5 e^{-6}}{5!} \approx 0{,}13$.

Локальна теорема Муавра-Лапласа

У тих випадках, коли число випробувань n велике і ймовірність p не близька до нуля ($p \neq 0$, $p \neq 1$), для обчислення біноміальних ймовірностей використовуються теореми Муавра-Лапласа.

Локальна теорема Муавра-Лапласа. Якщо ймовірність настання події A в кожному випробуванні постійна і відмінна від нуля і одиниці, а кількість незалежних випробувань досить велика, то ймовірність $P_n(k)$ можна обчислити за наближеною формулою

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x),$$

де $x = \frac{k-np}{\sqrt{npq}}$.

Вираз $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ називається функцією Гаусса, а її графік називається кривою ймовірності (див. рис. 3.1).

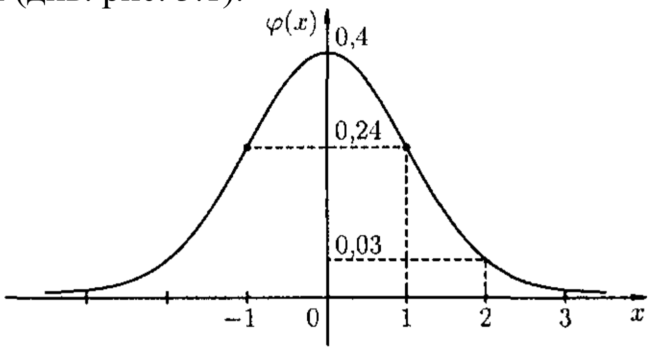

Рисунок 3.1 – Графік функції Гаусса

Для функції складено таблиці значень $\varphi(x)$ (див. додаток). Користуючись таблицею, слід враховувати, що:

а) $\varphi(x)$ парна функція, тобто $\varphi(-x) = \varphi(x)$;

б) якщо $x \geq 4$, то можна припустити, що $\varphi(x) = 0$.

Приклад 5 Імовірність ураження цілі одним пострілом для даного стрільця дорівнює 0,7. Знайдіть ймовірність того, що при 200 пострілах мішень буде вражена 160 разів.

Розв'язання. Тут $n = 200$, $p = 0{,}7$, $q = 0{,}3$, $k = 160$. Застосовуємо формулу. Маємо: $\sqrt{npq} = \sqrt{200 \cdot 0{,}7 \cdot 0{,}3} = \sqrt{42} \approx 6{,}48$, отже, $x = \frac{160-200 \cdot 0{,}7}{\sqrt{42}} = \frac{20}{6{,}48} \approx 3{,}09$. Враховуючи, що $\varphi(3{,}09) \approx 0{,}0034$, отримуємо $P_n(k) \approx \frac{1}{6{,}48} \cdot 0{,}0034 \approx 0{,}0005$.

Інтегральна теорема Муавра-Лапласа

У тих випадках, коли необхідно розрахувати ймовірність того, що в n незалежних випробуваннях подія A з'явиться хоча б k_1 разів, але не більше k_2 разів, тобто $P_n(k_1 \leq k \leq k_2)$ або $P_n(k_1; k_2)$ використовують інтегральну теорему Муавра-Лапласа.

Інтегральна теорема Муавра-Лапласа. Якщо ймовірність p настання події A в кожному випробуванні постійна і відрізняється від нуля і одиниці, то ймовірність $P_n(k_1 \leq k \leq k_2)$ можна знайти за наближеною формулою

$$P_n(k_1 \leq k \leq k_2) = \Phi(x_2) - \Phi(x_1) = \Phi_0(x_2) - \Phi_0(x_1),$$

де $x_1 = \frac{k_1-np}{\sqrt{npq}}$, $x_2 = \frac{k_2-np}{\sqrt{npq}}$.

Для спрощення обчислень при використанні формули була введена спеціальна функція $\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$, яка називається *нормованою функцією Лапласа* (графік на рис. 3.2). Поряд з нормованою функцією Лапласа використовується функція $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{t^2}{2}}dt$, яку також називають функцією Лапласа. Для неї справедлива рівність $\Phi(-x)+\Phi(x)=1$; вона пов'язана з функцією $\Phi_0(x)$ формулою $\Phi(x)=0{,}5+\Phi_0(x)$.

Існують таблиці наближених значень функцій $\Phi_0(x)$ і $\Phi(x)$ (інтеграл не береться в елементарних функціях), які наведені в більшості підручників з теорії ймовірностей (див. додаток).

Ця функція має наступні властивості:

а) функція $\Phi_0(x)$ непарна:

$$\Phi_0(-x)=\frac{1}{\sqrt{2\pi}}\int_0^{-x} e^{-\frac{t^2}{2}}dt=|t=-z|=-\frac{1}{\sqrt{2\pi}}\int_0^{x} e^{-\frac{z^2}{2}}dz=-\Phi_0(x);$$

б) якщо $x\geq 5$, то можна припустити, що $\Phi_0(x)=0{,}5$.

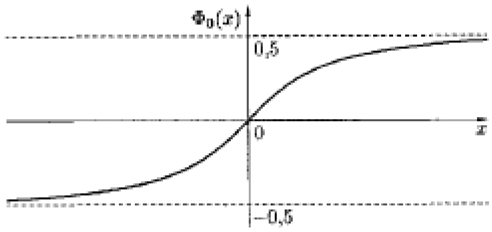

Рисунок 3.2 – Графік функції $\Phi_0(x)$

Приклад 6 Інспекція встановила, що цех виробляє в середньому 96% продукції вищого сорту. На підставі цього приймальник перевіряє 200 виробів цього цеху. Якщо серед них більше 10 виробів не вищого сорту, то вся партія продукції вибраковується, тобто повертається в цех. Яка ймовірність того, що партія буде прийнята?

Розв'язання. Тут $n=200$, $p=0{,}04$ (ймовірність непридатного продукту), $q=0{,}96$. Імовірність прийняття всієї партії, тобто $P_{200}(0\leq k\leq 10)$, можна знайти за формулою (3.4); тут $k_1=0$, $k_2=10$. Знаходимо,

$$x_1=\frac{0-200\cdot 0{,}04}{\sqrt{200\cdot 0{,}04\cdot 0{,}96}}\approx -2{,}89,\ x_2=\frac{10-200\cdot 0{,}04}{\sqrt{200\cdot 0{,}04\cdot 0{,}96}}\approx 0{,}72,$$

$$P_{200}(0\leq k\leq 10)=\Phi_0(0{,}72)-\Phi_0(-2{,}89)=0{,}26424+0{,}49807=0{,}7623.$$