Лабораторна робота

Схема Бернуллі. Асимптотична формула Пуассона. Локальна та інтегральна теореми Лапласа

СХЕМА БЕРНУЛЛІ

Ймовірність того, що число «успіхів» в серії з n послідовних незалежних випробувань за схемою Бернуллі дорівнює k обчислюється за формулою $P_n(k) = C_n^k p^k q^{n-k}$. Обчислити $P_n(k)$ в R можна наступним чином:

```
> choose(n,k)*p^k*q^(n-k)
```

або використати функцію $\boldsymbol{dbinom(k, n, p)}$. Набір $P_n(k)$ називається біномним (біноміальним) розподілом, а відповідна формула – формулою Бернуллі.

Приклад 1. Із партії, в якій є **12** стандартних і **4** нестандартних деталей, навмання беруть **3** деталі з поверненням. Знайти ймовірність того, що серед узятих деталей: 1) усі три стандартні; 2) не більш як одна нестандартна; 3) принаймі одна нестандартна.

Зразок виконання в R:

```
n <- 3; p <- 12 / (4 + 12)

# probA - ймовірність вибору трьох стандартних деталей
probA <- choose(n,3) * p ^ 3 * (1 - p) ^ (n - 3)
#або
probA1 <- dbinom(3,n,p)

# probB - ймовірність вибору не більш як одної нестандартної деталі
probB <- probA + dbinom(2,n,p)

# probC - ймовірність вибору принаймі однієї нестандартної деталі
probC <- 1 - probA
```

Результат:

n	3
p	0.75
probA	0.421875
probA1	0.421875
probB	0.84375
probC	0.578125

АСИМПТОТИЧНА ФОРМУЛА ПУАССОНА

При великому числі випробувань обчислення за формулою Бернуллі стають складними. Однак у ряді випадків їх можна замінити простішими асимптотичними формулами. Одна з них ґрунтується на теоремі Пуассона.

Теорема Пуассона. *Якщо ймовірність появи події A в кожному з n незалежних випробувань за схемою Бернуллі рівна p і якщо для $n \to \infty$ $p \to 0$ так, що $np \to \lambda$ $(0 < \lambda < \infty)$, то*

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

для будь-якого $k = 0,1,2,\dots$, де $P_n(k)$ – ймовірність появи k разів події A в n випробуваннях.

Отже, при великих n $(n > 100)$ і малих p $(np < 30)$ можна користуватися наближеними формулами:

$$P_n(k) \approx \frac{\lambda^k}{k!}e^{-\lambda}.$$

Для обчислення $\frac{\lambda^k}{k!}e^{-\lambda}$ в R використовується функція $\boldsymbol{dpois(k,\lambda)}$.

$$P_n(k_1 \le k \le k_2) = \sum_{k=k_1}^{k_2} P_n(k) \approx \sum_{k=k_1}^{k_2} \frac{\lambda^k}{k!}e^{-\lambda}, \lambda = np.$$

Для $P_n(k_1 \le k \le k_2)$ в R:

$$\boldsymbol{pbinom(k_2,n,p) - pbinom(k_1-1,n,p)},$$

а для $\sum_{k=k_1}^{k_2} \frac{\lambda^k}{k!}e^{-\lambda}$:

$$\boldsymbol{ppois(k_2,\lambda) - ppois(k_1-1,\lambda)},$$

при $k_1 > 0$.

Приклад 2. Дослідити точність асимптотичної формули Пуассона на прикладі розв’язку такої задачі:

З умов випуску лотереї відомо, що виграшними є **1/40** всіх випущених лотерейних квитків. Яка ймовірність, що з **200** куплених лотерейних квитків виграшними будуть **5**? Не менше **5**?

Зразок виконання в R:

```r
n <- 200; p <- 1/40
l <- n*p

# Подія A - серед куплених квитків 5 виграшних
# Обчислення ймовірності події A за формулою Бернуллі (точне значення)
probA <- choose(n,5) * p ^ 5 * (1 - p) ^ (n - 5)
#або
probA1 <- dbinom(5,n,p)

# Обчислення ймовірності події A за наближеною формулою Пуассона
aproxProbA <- dpois(5,l)

# Похибка
difference1 <- abs(probA - aproxProbA)

# Подія B - серед куплених квитків не менше 5-ти виграшних
# Обчислення ймовірності події B (точне значення)
probB <- 1 - pbinom(4,n,p)
#або
probB1 <- 1 - sum(dbinom(0:4,n,p))

# Обчислення ймовірності події B за наближеною формулою
aproxProbB <- 1 - ppois(4,l)
#або
aproxProbB1 <- 1 - sum(dpois(0:4,l))

# Похибка
difference2 <- abs(probB - aproxProbB)
```

Результат:

aproxProbA	0.175467369767851
aproxProbB	0.559506714934788
aproxProbB1	0.559506714934788
difference1	0.0022334447164295
difference2	0.00222407184902118
l	5
n	200
p	0.025
probA	0.17770081448428
probA1	0.177700814484281
probB	0.561730786783809
probB1	0.561730786783809

ЛОКАЛЬНА ТА ІНТЕГРАЛЬНА ТЕОРЕМИ ЛАПЛАСА

Якщо $npq \geq 10$, то для розрахунків використовують наближення у відповідності з теоремою Муавра-Лапласа.

Локальна теорема Муавра-Лапласа. *Нехай* $0 < p < 1$ *і величина* $x_k = \frac{k-np}{\sqrt{npq}}$ *обмежена при* $n \to \infty$*, тоді*

$$P_n(k) = C_n^k p^k q^{n-k} \to \frac{e^{\frac{-x_k^2}{2}}}{\sqrt{2\pi npq}}$$

Вимога обмеженості величини x_k означає, що при $n \to \infty$ величина k також повинна зростати разом з величиною n. Точність формули

$$P_n(k) \approx \frac{e^{\frac{-x_k^2}{2}}}{\sqrt{2\pi npq}}$$

зростає як з ростом величин n і k, так і в міру наближення величин p і q до 0,5.

Приклад 3. Кубик кидають **5000** разів. Знайти ймовірність того, що «двійка» випаде **900** разів. Провести обчислення за формулою Бернуллі і за наближеною формулою Муавра-Лапласа. Порівняти результати.

Зразок виконання в R:

```r
n <- 5000; k <- 900
p <- 1 / 6; q <- 1 - p
l1 <- n * p; l2 <- n * p * q

# Використовуємо локальну формулу Муавра-Лапласа
xk <- (k - l1)/sqrt(l2)
aproxProbA <- 1 / sqrt(2 * pi * l2) * exp(-xk ^ 2 / 2)
#або
aproxProbA1 <- dnorm(xk) / sqrt(l2)

# За формулою Бернуллі
probA <- dbinom(k,n,p)

# Похибка
difference <- abs(probA - aproxProbA)
```

Результат:

aproxProbA	0.000617090655162208
aproxProbA1	0.000617090655162208
difference	2.19913318569256e-05
k	900
l1	833.333333333333
l2	694.444444444444
n	5000
p	0.166666666666667
probA	0.000639081987019134
q	0.833333333333333
xk	2.52982212813471

Інтегральна теорема Муавра-Лапласа. *Якщо ймовірність p появи події A в кожному випробуванні стала і $0 < p < 1$, то ймовірність $P_n(k_1, k_2)$ того, що подія A настане не менше k_1 і не більше k_2 разів, наближено дорівнює визначеному інтегралу*

$$P_n(k_1, k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-\frac{z^2}{2}} dz,$$

де $x' = \frac{k_1 - np}{\sqrt{npq}}$, $x'' = \frac{k_2 - np}{\sqrt{npq}}$.

Функція $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$ називається **функцією Лапласа**.

$$P_n(k_1, k_2) \approx \Phi(x'') - \Phi(x'), \qquad x' = \frac{k_1 - np}{\sqrt{npq}}, \qquad x'' = \frac{k_2 - np}{\sqrt{npq}}.$$

Точність цієї наближеної формули зростає разом зі зростанням n. В R для обчислення значень $\Phi(x)$ призначена функція $\boldsymbol{pnorm(x, 0, 1) - 1/2}$.

Приклад 4. Дослідити точність інтегральної формули Муавра-Лапласа на прикладі розв’язку такої задачі:

Зерна пшениці проростають з ймовірністю **0,95**. Знайти ймовірність того, що із **2000** посіяних зерен зійде від **1700** до **1900**. Провести обчислення за формулою Бернуллі і за наближеною інтегральною формулою Муавра-Лапласа.

Зразок виконання в R:

```
p <- 0.95; q <- 1 - p; n <- 2000; k1 <- 1700; k2 <- 1900
l1 <- n * p; l2 <- n * p * q

# Використовуємо формулу з інтегральної теореми Муавра-Лапласа
x1 <- (k1 - l1)/sqrt(l2)
x2 <- (k2 - l1)/sqrt(l2)
aproxProbA <- pnorm(x2,0,1) - pnorm(x1,0,1)

# Обчислюємо точне значення ймовірності
probA <- pbinom(k2,n,p) - pbinom(k1-1,n,p)
#або
probA1 <- sum(dbinom(1700:1900,n,p))

# Похибка
difference <- abs(probA - aproxProbA)
```

Результат:

aproxProbA	0.5
difference	0.0143254915923848
k1	1700
k2	1900
l1	1900
l2	95.0000000000001
n	2000
p	0.95
probA	0.514325491592385
probA1	0.514325491592385
q	0.05
x1	-20.5195670417031
x2	0

Індивідульні завдання

Завдання 1

Із партії, в якій є **L+12** стандартних і **L** нестандартних деталей, навмання беруть **5** деталей з поверненням. Знайти ймовірність того, що серед узятих деталей: 1) три стандартні; 2) не більш як одна нестандартна; 3) принаймі одна нестандартна. (**L** – номер варіанту)

Завдання 2

Дослідити точність асимптотичної формули Пуассона на прикладі розв'язку такої задачі:

З умов випуску лотереї відомо, що виграшними є **1/40** всіх випущених лотерейних квитків. Яка ймовірність, що з **[200*(L+1)/L]** куплених лотерейних квитків виграшними будуть **L**? Не менше **L**?

(**L** – номер варіанту, [·] – ціла частина числа)

Завдання 3

Кубик кидають **6000** разів. Знайти ймовірність того, що «двійка» випаде **1000-10L** разів. Провести обчислення за формулою Бернуллі і за наближеною формулою Муавра-Лапласа. Порівняти результати. (**L** – номер варіанту)

Завдання 4

Дослідити точність інтегральної формули Муавра-Лапласа на прикладі розв'язку такої задачі:

Ймовірність влучення стрілка в мішень дорівнює **1-(L+7)/(L+40)**. Знайти ймовірність того, що зі **120** вистрілів він влучить в мішень від **85-L** до **100-L** разів. Провести обчислення за формулою Бернуллі і за наближеною інтегральною формулою Муавра-Лапласа. (**L** – номер варіанту)