ЛЕКЦІЯ 16
3.3. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

3.3.1. Поняття статистичної гіпотези. Помилки першого та другого роду

В генеральній сукупності закони або параметри розподілу досліджуваної ознаки X, як правило, невідомі, тому щодо цих законів або параметрів можуть висуватися різні припущення. Ці припущення, які перевіряються за даними вибірок, називаються *статистичними гіпотезами*. Прикладами таких гіпотез можуть бути висловлювання: ознаки X і Y мають однакові математичні сподівання, ознака X – нормально розподілена тощо.

Статистична гіпотеза висувається дослідником на основі його досвіду або одержаної статистичної інформації. Висунута гіпотеза називається *нульовою* або *основною* і позначається H_0. Поряд з основною гіпотезою прийнято висувати і іншу H_k, яка суперечить основній і називається *конкуруючою* або *альтернативною*. Як правило, події, що їх описують гіпотези H_0 і H_k, складають повну групу подій, тобто, якщо нульова гіпотеза не вірна, то має місце конкуруюча гіпотеза H_k. Справедливість однієї з цих гіпотез перевіряється методами математичної статистики, при цьому можливі *помилки першого* та *другого роду*:

– відмова від вірної гіпотези;
– прийняття невірної гіпотези.

Ці помилки, хоча і здаються на перший погляд досить близькими, можуть суттєво відрізнятись як по змісту, так і по їх можливим наслідкам. Наприклад, при підході літака до аеродрому в складних метеоумовах висувається гіпотеза про можливість посадки (власне, про значення параметра – ймовірність благополучної посадки). Нехай нульова гіпотеза – посадка можлива, тобто ймовірність благополучної посадки більша деякої заданої величини. Альтернативна їй гіпотеза – посадка неможлива, тобто ймовірність аварійної посадки все ж значна.

Помилка першого роду може виникнути, якщо нульова гіпотеза буде відхилена, в той час як умови дозволяють посадку. При цьому виникає матеріальний збиток (посадка на запасному аеродромі і пов’язані з нею витрати). Помилка другого роду може статися, якщо нульова гіпотеза буде прийнята, в той час як об’єктивні обставини не дозволяють виконати посадку. В цьому випадку помилка другого роду може привести до більш важких наслідків, ніж помилка першого роду.

3.3.2. Критерії узгодження та схема їх застосування

Для перевірки справедливості нульової статистичної гіпотези H_0 відносно закону або параметрів розподілу ознаки X генеральної сукупності по наявним експериментальним даним, які містяться у вибірці, застосовуються спеціальні випадкові величини K, які називаються *статистичними критеріями* або *критеріями узгодження* і закон розподілу яких відомий принаймні наближено. Розроблено багато різних критеріїв узгодження, серед яких найбільш широке застосування для перевірки гіпотез про закони розподілів одержали, завдяки своїй надійності, критерії К. Пірсона і А.М. Колмогорова.

Для застосування критерію обирається достатньо мала ймовірність α (зокрема, 0,001, 0,01 або 0,05), яка називається *рівнем значущості* і являє собою допустиму умовами дослідження ймовірність помилки першого роду. За прийнятим рівнем значущості α і об’ємом вибірки n (або залежним від n числом степенів свободи) по таблиці розподілу критерію K знаходяться критичні точки $k_{кр}$, які розбивають всю множину значень критерію K на дві неперетинні множини:

– критичну область, в яку K потрапляє з занадто малою ймовірністю α, тому при цьому гіпотеза H_0 може бути відхиленою;
– область прийняття гіпотези.

В залежності від змісту основної і конкуруючої гіпотез критична область може бути односторонньою (зокрема, правосторонньою (рис.3.4, а) і лівосторонньою (рис.3.4, б)) або двосторонньою (рис.3.4, в).

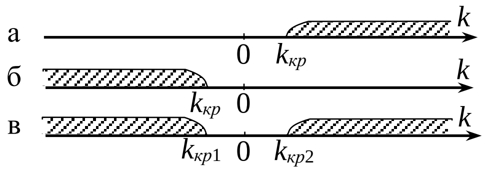

Рис.3.4

Для правосторонньої критичної області $k_{кр}$ знаходиться з умови

$$P(K \geq k_{кр}) = \alpha \qquad (k_{кр} > 0), \tag{3.35}$$

для лівосторонньої – з умови

$$P(K \leq k_{кр}) = \alpha \qquad (k_{кр} < 0), \tag{3.36}$$

а для двосторонньої – з умови

$$P(K \leq k_{кр1}) + P(K \geq k_{кр2}) = \alpha \tag{3.37}$$

В загальному випадку схема застосування критерію для перевірки статичної гіпотези H_0 складається з таких етапів:

1) вибір відповідного критерію K і обчислення за даними вибірки його значення, яке називається спостереженим значенням і позначається $K_{сп}$;

2) обрання за змістом гіпотези рівня значущості α;

3) знаходження за таблицею, відповідною розподілу обраного критерію, критичної точки $k_{кр}$ для прийнятих рівня значущості α і об’єму вибірки n або кількості степенів свободи;

4) порівняння $K_{сп}$ і $k_{кр}$ та прийняття рішення, згідно з умовами (3.35)-(3.37), наприклад, для правосторонньої і двосторонньої симетричної областей виконання умови $K_{сп} \geq k_{кр}$ є підставою для відхилення гіпотези H_0, а не виконання – для прийняття гіпотези.

3.3.3. Перевірка статистичних гіпотез. Критерій узгодження χ^2 Пірсона

Якщо за даними вибірки, зокрема, за видом емпіричної функції розподілу $F^*(x)$ можна зробити припущення про вид теоретичної функції розподілу $F_T(x)$ досліджуваної ознаки X генеральної сукупності, тобто висунути нульову гіпотезу H_0: $F_T(x) = F(x)$, то для перевірки гіпотези H_0 вибирається деяка невід’ємна міра R розбіжності між емпіричною функцією $F^*(x)$ і гіпотетичною теоретичною функцією $F(x)$:

$$R = R\{F^*(x), F(x)\} \tag{3.39}$$

Оскільки функція $F^*(x)$ залежить від вибірки, то величина R є випадковою і може знаходитись різними способами в залежності від вибраного критерію узгодження, тобто під вибором міри R фактично розуміється вибір відповідного критерію.

Якщо закон розподілу випадкової величини R відомий, то для вибраного рівня значущості α можна знайти таке число R_0, для якого виконується умова

$$P\{R > R_0\} = \alpha \qquad (3.40)$$

і яке називається *межею значущості критерію*.

Далі за даними вибірки будується емпірична функція $F^*(x)$ і обчислюється величина R (3.39), відповідна обраному критерію. Якщо $R > R_0$, то згідно з умовою (3.40) відбувається малоймовірна подія, тобто гіпотеза H_0 не узгоджується з даними вибірки і має бути відхилена. Якщо ж $R < R_0$, то гіпотеза H_0 може бути прийнята.

3.3.4. Критерій узгодження χ^2 (хі-квадрат) Пірсона

Для перевірки гіпотези $H_0 : F_T(x) = F(x)$ про закон розподілу ознаки X генеральної сукупності в випадку, коли параметри передбачуваного теоретичного розподілу невідомі і їх доводиться оцінювати за даними вибірки, застосовується критерій χ^2 К. Пірсона.

За міру R розбіжності між емпіричною функцією $F^*(x)$ і теоретичною функцією $F(x)$ в ньому обрана сума квадратів відхилень $\left(\omega_i^* - p_i\right)$, взятих з деякими ваговими коефіцієнтами:

$$R = \sum_{i=1}^{l} c_i \left(\omega_i^* - p_i\right)^2,$$

де ω_i^* – відносні частоти з інтервального статистичного розподілу вибірки (табл.3.2);

p_i – відповідні ймовірності теоретичного розподілу, вибрані з ряду розподілу дискретної ознаки X або обчислені за формулою $p_i = F(x_i) - F(x_{i-1})$ для неперервної ознаки X.

Коефіцієнти c_i призначені для вирівнювання значущості відхилень $\left(\omega_i^* - p_i\right)$, які при малих ймовірностях p_i виявляються досить значними, а при великих p_i – малозначущими, тому за коефіцієнти c_i Пірсоном були обрані величини, обернено пропорційні ймовірностям p_i:

$$c_i = \frac{n}{p_i}$$

При такому виборі коефіцієнтів міра R позначається χ^2:

$$R = \chi^2 = \sum_{i=1}^{l} \frac{n(\omega_i^* - p_i)^2}{p_i}.$$

Оскільки $\omega_i^* = \dfrac{n_i^*}{n}$, то

$$R = \chi^2 = \sum_{i=1}^{l} \frac{(n_i^* - np_i)^2}{np_i}. \qquad (3.41)$$

Пірсон показав, що при збільшенні n закон розподілу величини R наближається до розподілу χ^2 (п.2.4.4), тому значення $R_0 = \chi^2_{kp}$ для заданого рівня значущості α і числа степенів свободи k вибирається з таблиці критичних точок розподілу χ^2 (додаток 3).

Число степенів свободи обчислюється за формулою

$$k = l - r - 1,$$

де l – кількість інтервалів статистичного розподілу вибірки, r – кількість невідомих параметрів теоретичного розподілу, які оцінюються за даними вибірки. Оскільки дані вибірки підлягають обов’язковій умові

$$\sum_{i=1}^{l} \omega_i^* = 1 ,$$

то число k зменшується ще на 1.

Далі застосування критерію χ^2 виконується за загальною схемою, наведеною в п.3.3.3: якщо обчислене за формулою (3.41) спостережене значення критерію $\chi^2_{сп} > \chi^2_{кр}$, то гіпотеза H_0 відхиляється, а при $\chi^2_{сп} < \chi^2_{кр}$ – може бути прийнятою.

Зауваження. Критерій χ^2 більш ефективний для вибірок великого об’єму ($n \geq 50$) з частотами n_i^* в інтервальному статистичному розподілі, не меншими 5-8, тому при застосуванні критерію суміжні інтервали з меншими частотами слід згрупувати.

Розглянемо застосування критерію χ^2 для перевірки гіпотез про розподіл ознаки X генеральної сукупності за найбільш відомими і часто вживаними розподілами.

1) *Нормальний розподіл*. Для перевірки гіпотези про нормальний розподіл ознаки X за оцінки невідомих параметрів a і σ приймається вибіркова середня $\bar{x}_B$ і вибіркове середнє квадратичне відхилення s, тому число степенів свободи $k = l - 3$, а ймовірності p_i для i-го інтервалу статистичного розподілу обчислюються за формулою:

$$p_i = \Phi\left(\frac{x_{i+1} - \bar{x}_B}{s}\right) - \Phi\left(\frac{x_i - \bar{x}_B}{s}\right). \qquad (3.42)$$

2) *Показниковий розподіл*. Для перевірки гіпотези про показниковий розподіл ознаки X за оцінку невідомого параметра λ приймається величина $\lambda^* = 1/\bar{x}_B$, тому число степенів свободи $k = l - 2$, а ймовірності p_i для інтервалів (x_{i-1}, x_i) статистичного розподілу знаходяться за формулою:

$$p_i = e^{-\lambda^* x_{i-1}} - e^{-\lambda^* x_i} . \qquad (3.43)$$

3) *Рівномірний розподіл*. Для перевірки гіпотези про рівномірний розподіл ознаки X за оцінки невідомих параметрів a і b – меж інтервалу – приймаються величини a^* і b^*, обчислювані за формулами:

$$a^* = \bar{x}_B - \sqrt{3}s, \quad b^* = \bar{x}_B + \sqrt{3}s .$$

Тому число степенів свободи $k = l - 3$, а ймовірності p_i для інтервалів (x_{i-1}, x_i) статистичного розподілу знаходяться за формулою:

$$p_i = \frac{x_i - x_{i-1}}{b^* - a^*} .$$

4) *Біноміальний розподіл*. Для перевірки гіпотези про біноміальний розподіл ознаки X використовується дискретний статистичний розподіл вибірки, оскільки біноміальний розподіл застосовний до дискретної випадкової величини. Розглядається вибірка n серій по N випробувань, в кожному з яких подія A може відбутися з сталою ймовірністю p. Якщо ця ймовірність (як параметр розподілу) невідома, то вона оцінюється за даними вибірки величиною

$$\tilde{p} = \frac{1}{Nn}\sum_{i=1}^{m} x_i n_i \,, \tag{3.44}$$

де m – число варіант у вибірці, яке може бути меншим N. В цьому випадку число степенів свободи $k = m - 2$, а ймовірності p_i обчислюються за формулою Бернуллі:

$$p_i = C_N^i \tilde{p}^i (1 - \tilde{p})^{N-i} \,. \tag{3.45}$$

5) *Розподіл Пуассона*. Для перевірки гіпотези про розподіл дискретної ознаки X за законом Пуассона також застосовується дискретний статистичний розподіл вибірки. Оцінкою параметра λ є величина $\lambda^* = \overline{x}_B$, тому число степенів свободи $k = m - 2$, а ймовірності p_i для кожної варіанти обчислюються за формулою Пуассона:

$$p_i = \frac{\lambda^{*i} e^{-\lambda^*}}{i!} \,.$$