§9 Статистичні нульові та альтернативні гіпотези

1. Різного роду припущення відносно законів розподілу, їх параметрів та инших завдань статистичної аналізи носить назву **статистичної гіпотези**.

2. Висувають одну нульову (H_0) та одну **(H_1)** або декілька **(H_1, H_2, …, $H_к$)** альтернативних гіпотез.

3. Статистичні гіпотези бувають прості та складні. Гіпотези записують словами або математичними символами.

4. Наприклад, прості гіпотези:

4.1. Нульова гіпотеза. H_0: Реалізації (результати досліджень) випадкової величини X: $x_1, x_2, \ldots, x_i, \ldots, x_N$ **підпорядковані** нормальному закону розподілу Ґавса (н.з.р.).

Альтернативна гіпотеза H_1: Реалізації (результати досліджень) випадкової величини X: $x_1, x_2, \ldots, x_i, \ldots, x_N$ **непідпорядковані** н.з.р. Ґавса (нормальному закону розподілу Ґавса).

4.2. Нульова гіпотеза. H_0: $\sigma_1^2 = \sigma_2^2$

$$\uparrow \quad \uparrow$$

$$S_1^2 \neq S_2^2, \qquad (88)$$

або словами: H_0: дві ґенеральні дисперсії σ_1^2 і σ_2^2 від двох чинників X_1 і X_2, яким дано статистичні оцінки за вибірковими дисперсіями S_1^2 і S_2^2, відповідно, **рівні**.

Альтернативна гіпотеза H_1: $\sigma_1^2 \neq \sigma_2^2$

$$\uparrow \quad \uparrow$$

$$S_1^2 \neq S_2^2, \qquad (89)$$

або словами: H_1: дві ґенеральні дисперсії σ_1^2 і σ_2^2 від двох чинників X_1 і X_2, яким дано статистичні оцінки за вибірковими дисперсіями S_1^2 і S_2^2, відповідно, **нерівні**.

4.3. Якщо нульова гіпотеза H_0 буде прийнята, а альтернативна гіпотеза H_1 – відкинута, то з ймовірністю **$p=1-\alpha$** стверджується, що дві ґенеральні

дисперсії σ_1^2 і σ_2^2 статистично рівні, при цьому дві вибіркові дисперсії S_1^2 і S_2^2, які є оцінками генеральних дисперсій σ_1^2 і σ_2^2, статистично однорідні. Ці статистичні висновки правдиві з ймовірністю **p=1-α**, для рівня значущости **α = 1–p** та обсягів виборок $\mathbf{N_1}$ **і** $\mathbf{N_2}$, відповідно, за якими розраховані вибіркові дисперсії.

4.4. Якщо буде прийнята альтернативна гіпотеза $\mathbf{H_1}$, а нульова гіпотеза H_0 – відкинута, то з ймовірністю **p=1-α** стверджується, що дві ґенеральні дисперсії $\boldsymbol{\sigma_1^2}$ **і** $\boldsymbol{\sigma_2^2}$ статистично нерівні, а дві вибіркові дисперсії $\mathbf{S_1^2}$ **і** S_2^2 статистично неоднорідні. Ці статистичні висновки правдиві з ймовірністю **p=1-α**, для рівня значущости **α= 1 – p** та обсягів виборок $\mathbf{N_1}$ **і** $\mathbf{N_2}$, відповідно, за якими розраховані вибіркові дисперсії.

5. Наприклад, складна гіпотеза, що записана математичними символами:

- **нульова гіпотеза: $\mathbf{H_0}$:**

$$\begin{array}{c} \sigma_1^2 = \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_1^2 \neq S_\varepsilon^2 \\ \cdots\cdots\cdots \\ \sigma_2^2 = \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_2^2 \neq S_\varepsilon^2 \end{array} \qquad (90)$$

або словами: H_0: перша σ_1^2 і друга σ_2^2 ґенеральні дисперсії дорівнюють дисперсії помилки σ_ε^2, яким дано оцінки вибірковими дисперсіями S_1^2, S_2^2, S_ε^2, відповідно.

- **альтернативні гіпотези:**

$$\left.\begin{array}{c} H_1: \sigma_1^2 \neq \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_1^2 \neq S_\varepsilon^2 \\ \sigma_2^2 = \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_2^2 \neq S_\varepsilon^2 \end{array}\right\} (4) \qquad \left.\begin{array}{c} H_2: \sigma_1^2 = \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_1^2 \neq S_\varepsilon^2 \\ \sigma_2^2 \neq \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_2^2 \neq S_\varepsilon^2 \end{array}\right\} (5) \qquad \left.\begin{array}{c} H_3: \sigma_1^2 \neq \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_1^2 \neq S_\varepsilon^2 \\ \sigma_2^2 \neq \sigma_\varepsilon^2 \\ \uparrow \quad\quad \uparrow \\ S_2^2 \neq S_\varepsilon^2 \end{array}\right\} (91)$$

6. Статистичний метод перевірки нульової або альтернативної гіпотез полягає у тому, що за числовими значеннями критеріїв узгодження відомих

теоретичних розподілів знаходять: **якісну міру розходження** розподілу між результатами експерименту (емпіричними даними, даними виборки) або спостережень та теоретичними розподілами і таке инше. При цьому можна встановити і кількісну міру розходження цих розподілів.

Тобто, **прийняття** (правдивість) або **відкидання** (неправдивість) нульової гіпотези $\mathbf{H_0}$, або, відповідно, **відкидання** або прийняття **альтернативної гіпотези** $\mathbf{H_1}$, здійснюється шляхом порівняння **розрахованої певної статистики** за даними експерименту або спостережень з табличним (теоретичним) значенням (табульованими значеннями) певного теоретичного розподілу критерію узгодження.

Ці значення вибирають із таблиць, задаючи певним рівнем значущости α прийняття нульової гіпотези H_0:

$$\alpha = 1 - p, \qquad (92)$$

або рівнем ймовірности прийняття H_0:

$$p = 1 - \alpha \qquad (93)$$

та числом ступенів вільностей:

$$f = N - \xi, \qquad (94)$$

або обсягом виборки N,

де ξ - число накладених обмежень на число ступенів вільностей.

Рівень значущости α (і, відповідно, рівень ймовірности p) вибирають із ряду (табл. 1).

7. Якщо нульова гіпотеза $\mathbf{H_0}$ була **прийнята**, то рівень значущости $\alpha=(1 - p)$ – це є **максимальна ймовірність** (максимальний ризик) того, що прийнята гіпотеза H_0, як правдива, коли **у дійсности вона** може бути **неправдивою**, а $p=1 - \alpha$ – це є **мінімальна ймовірність** того, що прийнята нульова гіпотеза H_0 **дійсно є правдивою**.

Таблиця 1

Ряди рівнів значущости та рівнів ймовірности

α	α, %	p, %	α	α, %	p, %
0,0001	0,01	99,99	0,050	5,0	95,0
0,0005	0,05	99,95	0,100	10,0	90,0
0,0010	0,10	99,90	0,150	15,0	85,0
0,0025	0,25	99,75	0,200	20,0	80,0
0,0050	0,50	99,50	0,300	30,0	70,0
0,010	1,0	99,0	0,400	40,0	60,0
0,025	2,5	97,50	0,500	50,0	50,0

8. Якщо H_0 приймається, то **статистичні висновки формулюють**, наприклад, **так**:

«**Ґенеральні характеристики двох** сукупностей, що порівнюються, статистично рівні з ймовірністю $p=1 - \alpha$ [або $p(\%)=100 - \alpha(\%)$], а ймовірність ризику прийняти при цьому неправдиву H_0 становить $\alpha=1-p$ [або $\alpha(\%)=100 - p(\%)$] і ці статистичні висновки вірні за обсягами виборки N_1 і N_2 (числі ступенів вільностей $f_1=N_1-\xi$ та $f_2=N_2-\xi$, де ξ - число накладених зв'язків».

9. Якщо нульова гіпотеза H_0 була відкинута як неправдива, то рівень значущости $\alpha=(1-p)$ – це є **максимальна ймовірність** (максимальний ризик) того, що відкинута гіпотеза H_0 як неправдива, коли у дійсности вона є **правдивою**, а $p=(1-\alpha)$ – це є **мінімальна ймовірність** того, що відкинута гіпотеза H_0 **дійсно є неправдивою**.

10. Природа статистичних висновків є такою, що:

а) відкидаючи нульову гіпотезу H_0 як неправдиву, можна заздалегідь дати оцінку ймовірности можливої помилки під час відкидання істинно правдивої гіпотези H_0;

б) і, навпаки, приймаючи нульову гіпотезу як **правдиву**, то це ще **не означає**, що вона підтверджується із заданим рівнем ймовірности, а це означає

лише те, що вона узгоджується із даними виконаного статистичного досліду (і не більше!) і що можливо здійснити инший дослід, внаслідок якого нульова гіпотеза H_0 може бути відкинути.

11. Ділянки прийняття або відкидання статистичних гіпотез визначаються так: нехай маємо певний теоретичний розподіл $\varphi(z)$ випадкової величини Z з математичним сподіванням $a_z=z_0$ і числом ступенів вільностей $f=N-\xi$.

Нехай $Z(\alpha, f) = Z_p$ – квантиль випадкової величини Z, що відповідає значенню цієї величини з ймовірністю $p=1-\alpha$.

11.1. Для двобічної межі (рис.12): ймовірність попадання випадкової величини Z в интервал $[z(1 - \alpha/2); z(\alpha/2)]$ дорівнює:

$$p[z(1-\alpha/2) < Z \le z(\alpha/2)] = \int_{z(1-\alpha/2)}^{z(\alpha/2)} \varphi(z)dz = 1-\alpha = p, \qquad (95)$$

що чисельно дорівнює площі $p=(1-\alpha)$ під кривою $\varphi(z)\sim z$.

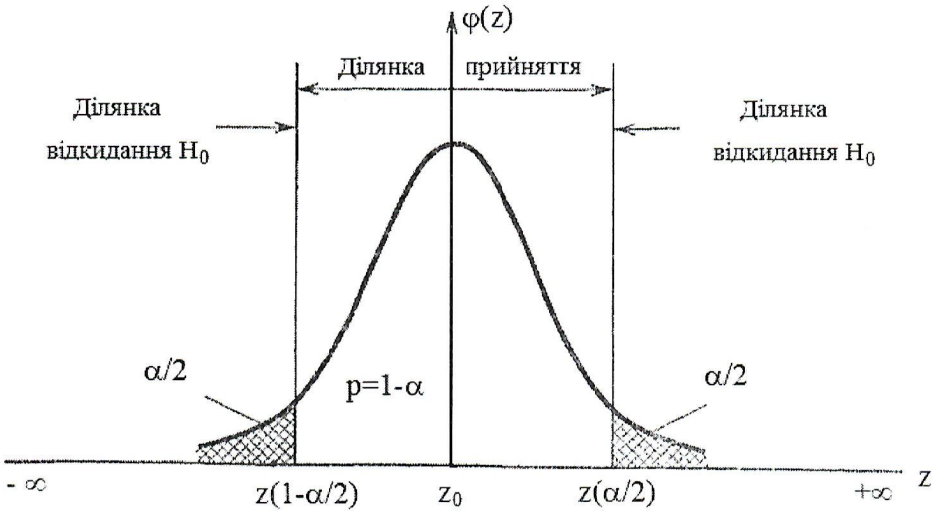


Рис. 12 Двобічна межа прийняття або відкидання нульової гіпотези.

Решта сумарної площі під кривою $\varphi(z) \sim z$ чисельно дорівнює:

$$p\,[z(1-\alpha/2) > Z > z(\alpha/2)] = \alpha, \qquad (96)$$

тобто ймовірність попадання випадкової величини Z в интервал $[z(-\infty); z(1-\alpha/2)]$ дорівнює:

$$p[z(-\infty) < Z \le z(1-\alpha/2)] = \int_{-\infty}^{z(1-\alpha/2)} \varphi(z)dz = \alpha/2, \qquad (97)$$

а ймовірність попадання випадкової величини Z в интервал [z(α/2); z (+∞)] дорівнює:

$$p[z(\alpha/2) \le Z < z(+\infty)] = \int_{z(\alpha/2)}^{+\infty} \varphi(z)dz = \alpha/2, \qquad (98)$$

що чисельно дорівнює відповідним площам під кривою φ(z)~ z.

11.2. Для однобічної межі (рис.13): ймовірність попадання випадкової величини Z в интервал [Z_α; (+∞)] дорівнює:

$$p[Z_\alpha \le Z < z(+\infty)] = \int_{Z_\alpha}^{+\infty} \varphi(z)dz = \alpha, \qquad (99)$$

а ймовірність попадання випадкової величини Z в интервал [(-∞); Z_α] дорівнює:

$$p[z(-\infty) < Z \le Z_\alpha] = \int_{-\infty}^{Z_\alpha} \varphi(z)dz = 1-\alpha = p, \qquad (100)$$

що чисельно дорівнює відповідним площам під кривою φ(z)~ z.

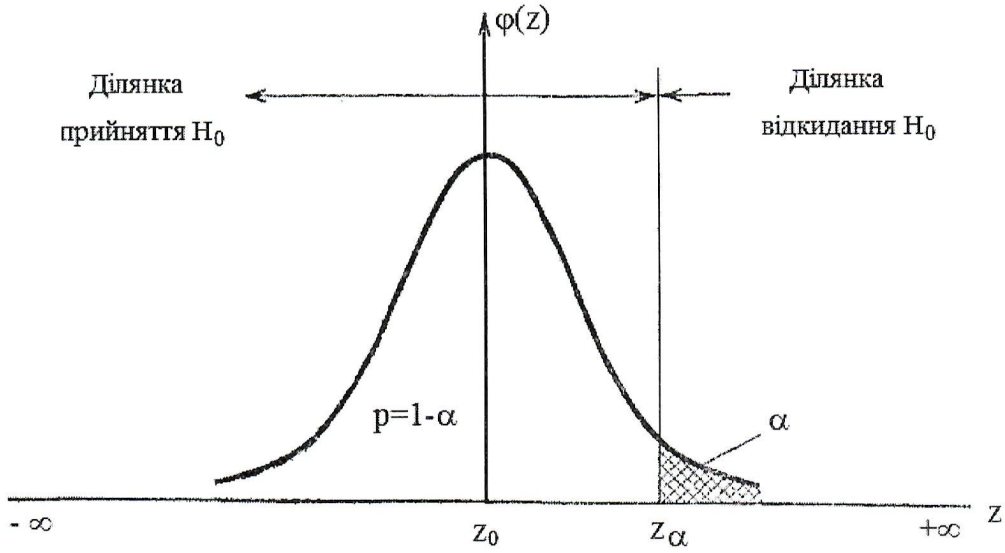
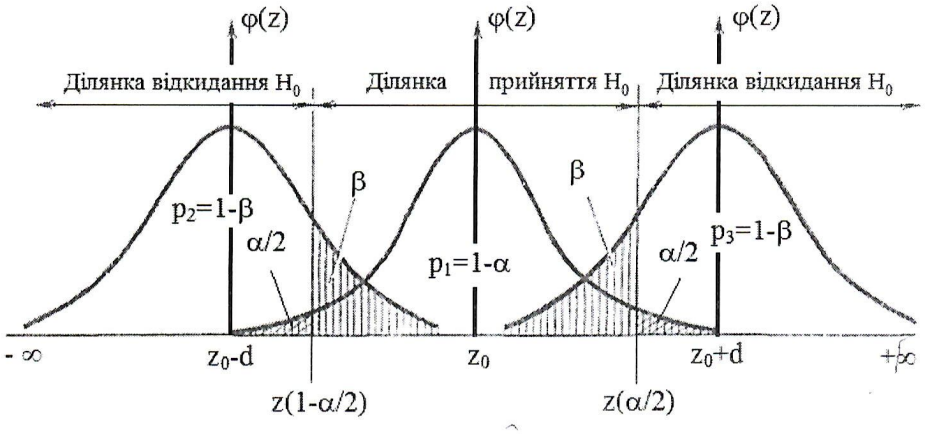


Рис.13. Однобічна межа прийняття або відкидання нульової гіпотези.

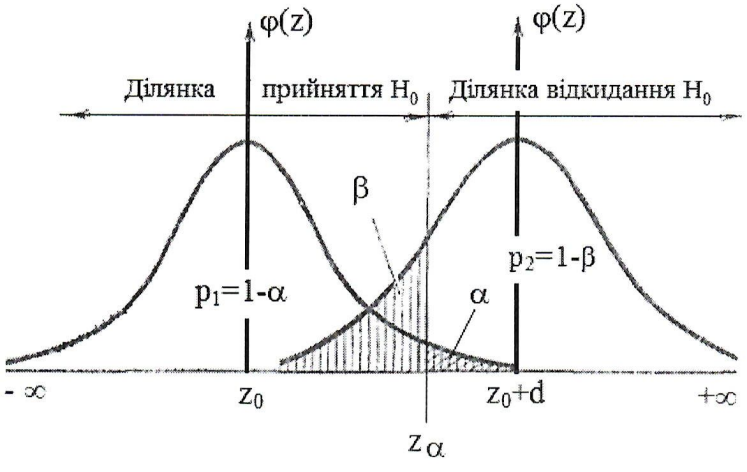
12. Під час прийняття або відкидання статистичних гіпотез можна припуститися помилок 1-го та 2-го роду:

а) помилки 1-го роду – це коли нульова гіпотеза H_0 **відкидається**, а **насправді** вона є **правдивою**. Тому ймовірність припуститися помилки 1-го роду дорівнює рівню значущости критерію – ймовірности α=1-р.

б) помилки 2-го роду – це коли нульова гіпотеза H_0 **приймається**, а **насправді** вона є **неправдивою**. Тому ймовірність припуститися помилки 2-го роду дорівнює ймовірности $\beta=1 - p_2 = 1 - p_3$, де $p_2 = (1 - \beta) = p_3$ – рівень потужности критерію.



а



б

Рис. 14. Зв'язок помилок 1-го і 2-го роду за: двобічної (а) та однобічної (б) межі.

12.1. **На рис.14 показано** співвідношення помилки 1-го роду (α) та помилки другого роду (β), які чисельно дорівнюють відповідним площам під кривою $\varphi(z) \sim z$.

12.2. На рис.14а для двобічної межі площі під кривою $\varphi(z) \sim z$ чисельно дорівнюють:

а) $\alpha/2$ – ймовірности того, що прийнята нульова гіпотеза H_0 (як правдива) насправді може бути неправдивою (причому $\alpha/2 + \alpha/2 = \alpha = 1 - p_1$); (101)

б) $p_1 = (1-\alpha)$ – ймовірности того, що прийнята нульова гіпотеза H_0 (як правдива) насправді може бути правдивою;

в) $\beta = (1 - p_2) = (1 - p_3)$ – ймовірности того, що відкинута нульова гіпотеза H_0 (як неправдива) насправді може бути правдивою;

г) $p_2 = p_3 = (1-\beta)$ – ймовірности того, що відкинута нульова гіпотеза H_0 (як неправдива) насправді може бути неправдивою.

12.3. Ймовірність того, що параметр Z вийде за верхню межу $z(\alpha/2)$ становить:

$$p[z(\alpha/2) \le Z < z(+\infty)] = \int_{Z(\alpha/2)}^{+\infty} \varphi(z)dz = \alpha/2. \qquad (102)$$

12.4. Ймовірність того, що параметр Z вийде за нижню межу $z(1-\alpha/2)$ становить:

$$p[z(-\infty) < Z \le z(1-\alpha/2)] = \int_{-\infty}^{z(1-\alpha/2)} \varphi(z)dz = \alpha/2. \qquad (103)$$

12.5. Ймовірність того, що параметр Z не вийде за межі інтервалу $[z(1-\alpha/2); z(\alpha/2)]$ становить:

$$p[z(1-\alpha/2) \le Z \le z(\alpha/2)] = \int_{z(1-\alpha/2)}^{z(\alpha/2)} \varphi(z)dz = 1-\alpha. \qquad (104)$$

12.6. На рис.14б для однобічної межі площі під кривою $\varphi(z) \sim z$ чисельно дорівнюють:

а) $\alpha = (1 - p_1)$ – ймовірности того, що прийнята нульова гіпотеза H_0 (як правдива), насправді може бути неправдивою;

б) $p_1 = (1-\alpha)$ – ймовірности того, що прийнята нульова гіпотеза H_0 (як правдива), насправді може бути правдивою;

в) $\beta = (1 - p_2)$ – ймовірности того, що відкинута нульова гіпотеза H_0 (як неправдива) насправді може бути правдивою;

г) $p_2 = (1-\beta)$ – ймовірности того, що відкинута нульова гіпотеза H_0 (як неправдива), насправді може бути неправдивою.

12.7. Ймовірність того, що параметр Z вийде за межу Z_α становить:

$$p[\, Z_\alpha \le Z < z(+\infty)] = \int_{Z_\alpha}^{+\infty} \varphi(z)dz = \alpha = 1 - p. \qquad (105)$$

Ймовірність того, що параметр Z не вийде за межу Z_α становить:

$$p[\, z(-\infty) < Z \le Z_\alpha] = \int_{-\infty}^{Z_\alpha} \varphi(z)dz = 1 - \alpha = p. \qquad (105)$$

12.8. Як видно з рис.14, помилки 1-го і 2-го роду взаємопов'язані: зменшення помилки 1-го роду (α) приводить до збільшення помилки 2-го роду (β) і навпаки.

12.9. Для одночасного зменшення помилок 1-го (α) і 2-го (β) роду необхідно збільшити об'єм вибірки N.