

2. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

2.1. Основні поняття

Існує велика кількість різноманітних методів перевірки статистичних гіпотез. При виборі методу для вирішення певного конкретного завдання необхідно виходити з відповідей на такі питання:

- якою є мета перевірки гіпотези;
- у яких шкалах виміряні аналізовані дані;
- чи є аналізовані вибірки незалежними або спряженими;
- скільки вибірок необхідно порівняти.

Розглянуті в цьому розділі методи застосовують при порівнянні двох вибірок. При більшій кількості вибірок використовують методи дисперсійного аналізу.

Гіпотезу, що перевіряють, називають **нульовою гіпотезою** (H_0). Прикладами нульових гіпотез можуть бути такі твердження: “Середні значення двох вибірок суттєво не відрізняються одне від одного”; “Дисперсія першої вибірки суттєво перевищує дисперсію другої”; “Розподіл вибірки відповідає нормальному закону з певними параметрами” тощо. Гіпотезу, що суперечить нульовій, називають **конкуруючою**, або **альтернативною гіпотезою** (H_1). Для вказаних вище нульових гіпотез конкуруючими можуть бути такі твердження: “Середні значення двох вибірок суттєво розрізняються одне від одного”; “Дисперсія першої вибірки не перевищує істотно дисперсію другої”; “Розподіл вибірки не відповідає нормальному закону із вказаними параметрами”. Для однієї нульової гіпотези у загальному випадку можна сформулювати багато різних альтернативних гіпотез.

Розрізняють прості та складні гіпотези. **Простою** називають гіпотезу, що містить тільки одне твердження. **Складні гіпотези** складаються з декількох простих (при цьому кількість простих гіпотез може бути нескінченно великою).

Зазвичай при перевірці нульової гіпотези використовують певні модельні розподіли, що приблизно відповідають розподілу досліджуваного параметра. Їх називають **статистичними критеріями**. На практиці як критерії найчастіше використовують нормальний розподіл, χ^2 -розподіл, розподіли Стюдента і Фішера. **Значенням критерію, що спостерігається**, називають його величину, яку розраховують за досліджуваними вибірками.

Для перевірки гіпотези весь вибірковий простір поділяють на дві області, що не перетинаються: критичну (w) та область прийняття ($W - w$). **Критичною областю** називають сукупність значень критерію, за яких нульову гіпотезу слід відхилити. **Областю прийняття гіпотези (областю допустимих значень)** називають сукупність значень критерію, за яких нульову гіпо-

тезу приймають. Перевірка гіпотези передбачає розрахунок значення критерію і перевірку його потрапляння до області прийняття гіпотези.

Вирізняють **двобічні** й **однобічні (лівобічні, правобічні)** критичні області (рис. 2.1, 2.2). Їх використання залежить від вибору конкуруючої гіпотези.

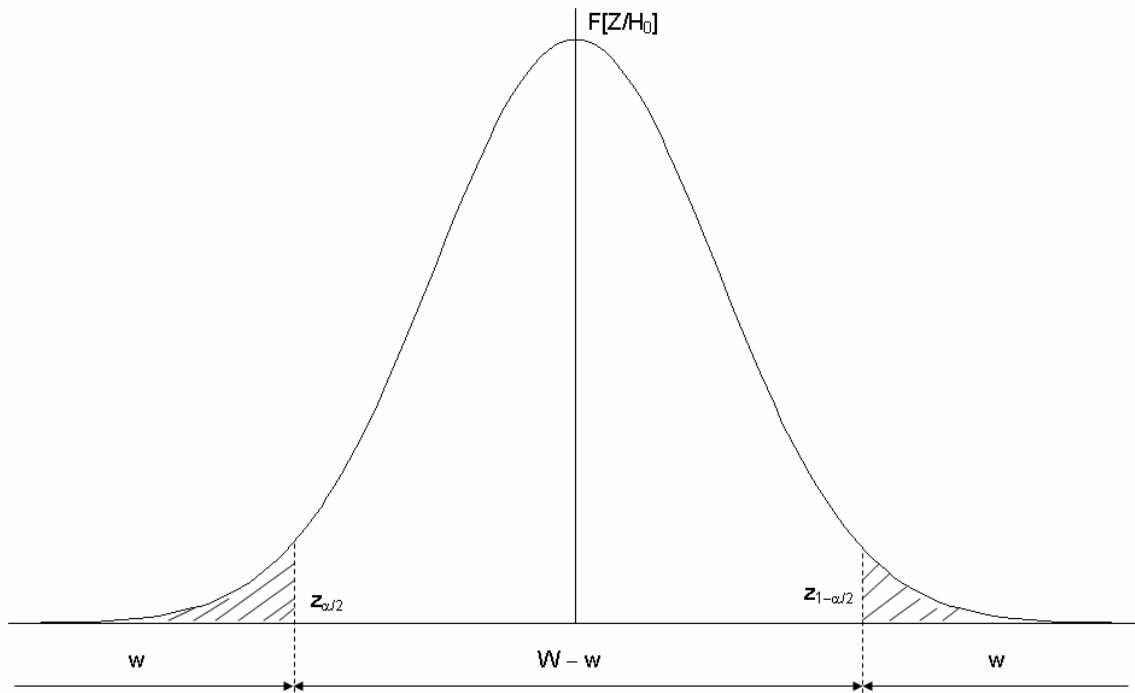


Рис. 2.1. Приклад двобічної критичної області

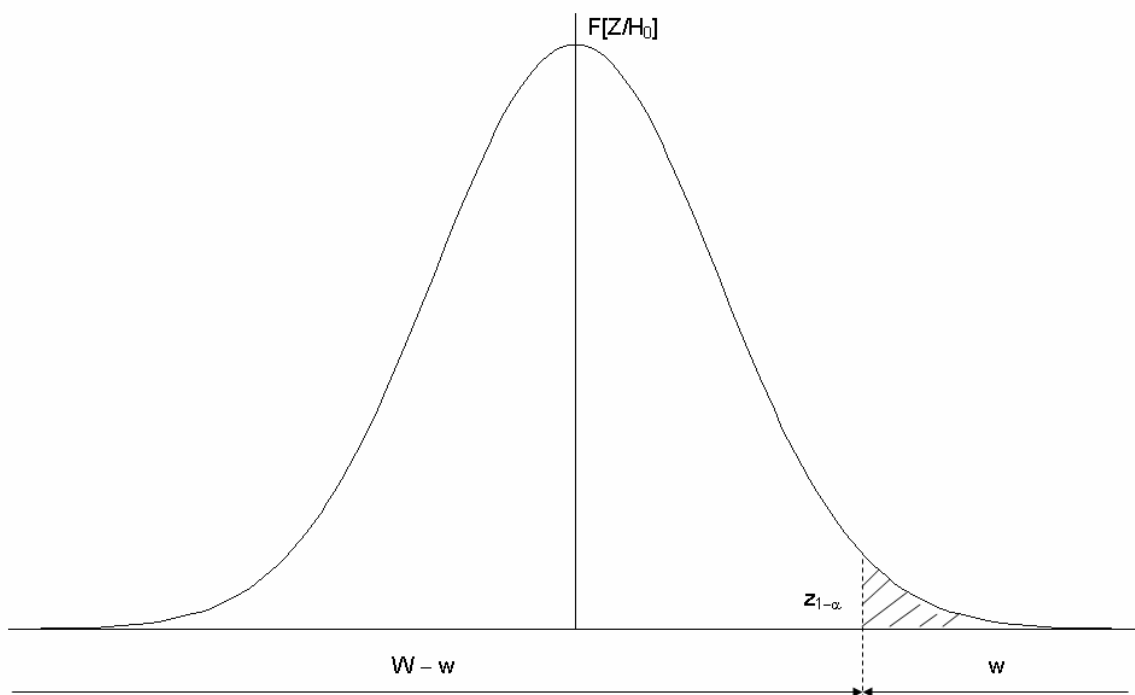


Рис. 2.2. Приклад правобічної критичної області

Якщо розподіл імовірності спостережень, що відповідає нульовій гіпотезі H_0 , є відомим, то критичну область визначають так, щоб при виконанні H_0 імовірність її відхилення була рівною заздалегідь заданій малій величині (**рівню значущості**) α .

$$P(x \in w | H_0) = \alpha, \quad (2.1)$$

Замість рівня значущості можна використовувати також **довірчий рівень** $p = 1 - \alpha$.

Критерії, що базуються на використанні заздалегідь заданого рівня значущості, називають **критеріями значущості**. Рівень значущості визначає розмір критичної області: що більшим є рівень значущості, то ширшою буде критична область.

Розглядають два типи помилок, що можуть виникати при перевірці статистичних гіпотез:

– **помилкою першого роду** є відхилення правильної нульової гіпотези, рівень значущості α є ймовірністю такої помилки;

– **помилкою другого роду** є прийняття помилкової нульової гіпотези.

У деяких застосуваннях помилки першого та другого роду називають, відповідно, **ризиком виробника** та **ризиком споживача**.

Зменшення ймовірності помилки першого роду водночас призводить до підвищення ймовірності помилки другого роду β . З огляду на це додатково вводять поняття **потужності критерію** $1 - \beta$, яка є імовірністю відхилення помилкової нульової гіпотези, тобто ймовірністю потрапляння критерію до критичної області за умови, що правильною є конкуруюча гіпотеза:

$$P\{x \in w | H_1\} = 1 - \beta. \quad (2.2)$$

Потужність критерію можна підвищити, збільшуючи обсяг вибірки. При визначенні критичної області її зазвичай будують так, щоб максимізувати потужність обраного критерію. За наявності декількох критеріїв, що можуть використовуватися для перевірки досліджуваної гіпотези, рекомендується обирати більш потужні з них, якщо їх застосування є обґрунтованим.

Важливим завданням є визначення обсягу вибірки, який дає змогу гарантувати певне значення похибки першого роду за умови, що похибка другого роду не перевищує заданого значення. Для цього необхідно розв'язати таку систему:

$$\begin{cases} P[Z \in w / H_0] = \alpha; \\ P[Z \in W / H_1] = \beta. \end{cases} \quad (2.3)$$

Її аналітичне розв'язання можливо лише у найпростіших випадках. Але, в багатьох випадках істотне спрощення можна отримати, якщо замінити ймовірності α й β значеннями меж відповідних критичних інтервалів.

Загальна методика отримання висновків при перевірці гіпотез передбачає, що на першому етапі необхідно задати рівень значущості. Найчастіше його беруть рівним 0,01; 0,05 або 0,1. Обираючи рівень значущості, слід пам'ятати, що його зменшення знижує ймовірність помилки першого роду, але збільшує ймовірність помилки другого роду. Тому, виходячи з конкретних умов, потрібно знайти певний компроміс між ймовірностями припустити помилки різного типу.

На другому етапі за даними вибірки розраховують значення критерію та порівнюють його з обчисленими для заданого рівня значущості межами критичної області. Якщо розраховане значення критерію потрапляє до них, то нульову гіпотезу відхиляють. В іншому випадку вважають, що немає підстав для відхилення нульової гіпотези і або приймають її на заданому рівні значущості, або здійснюють додаткову перевірку. Для визначення меж критичної області застосовують спеціальні таблиці або розраховують їх на основі відомих законів розподілу використовуваних критеріїв.

Можливості сучасної комп'ютерної техніки та наявного програмного забезпечення дають змогу отримувати висновки іншим шляхом. Якщо за наявними емпіричними даними розрахувати значення критерію, то на наступному етапі можна визначити, для якого рівня значущості це значення буде критичним. Ураховуючи, що рівень значущості є ймовірністю відхилення правильної нульової гіпотези, ми можемо за його значенням зробити висновок про ймовірність правильності або помилковості нульової гіпотези. Залежно від того, задовольняє нас отримана ймовірність помилки чи ні, нульову гіпотезу приймають або відхиляють.

При перевірці гіпотез доцільно застосовувати різні методи, призначені для вирішення одних й тих самих завдань та однакових типів даних. Причинами розбіжності отримуваних при цьому результатів зазвичай є:

- помилки при введенні даних;
- непридатність окремих методик для типу даних, що розглядають;
- алгоритмічні помилки у програмах, що використовують для аналізу.

Залежно від наявності або відсутності можливості визначення на пряму розбіжності порівнюваних вибірок, розрізняють **однобічні** та **двобічні критерії**. Перші застосовують, якщо наявні дані дають змогу вказати такий напрям, наприклад зробити висновок, що значення порівнюваної ознаки для одної вибірки є вищим, ніж в іншій. Двобічні критерії дають можливість зробити висновок лише про різницю вибірок за порівнюваною ознакою. Відповідно до цього говорять про однобічні й двобічні гіпотези. Для двобічних критеріїв рівень значущості є вдвічі більшим, ніж для відповідних однобічних. При використанні однобічних критеріїв рекомендується спочатку розраховувати двобічні. Якщо за двобічним критерієм різниці між вибірками немає, то наступне порівняння за однобічним є необґрунтованим.

Дані реальних експериментів можуть бути подані **незалежними** або **спряженими** вибірками. Для незалежних вибірок критерії допомагають виявити статистичну значущість різниці, що спостерігається. Прикладами незалежних вибірок є:

- мешканці двох різних населених пунктів (при демографічних дослідженнях);
- дві партії однотипної продукції, виготовлені різними працівниками на різному обладнанні (при розробці технології виробництва);
- випускники різних шкіл (при аналізі результатів зовнішнього незалежного оцінювання).

Критерії, що застосовують до вибірок з попарно спряженими даними, називають **парними**. Прикладами спряжених вибірок є:

- дані опитування громадської думки до й після певної суспільно значущої події;
- дві партії однотипної продукції, виготовлені одними й тими самими працівниками на одному й тому самому обладнанні до й після внесення певних змін до технології;
- одна й та сама партія виробів до і після певної технологічної обробки.

2.2. Параметричні тести

Критерії й тести, що застосовують для порівняння вибірок, поділяють на дві групи: параметричні й непараметричні. Особливістю параметричних критеріїв є припущення, що розподіл ознаки в генеральній сукупності підпорядковується певному відомому закону. Ця відповідність має бути доведена до застосування будь-якого з параметричних тестів. Переважна більшість параметричних тестів розроблена для нормально розподілених даних. Але для деяких типів гіпотез існують параметричні тести, призначені для вибірок, що підпорядковуються іншим законам розподілу.

Як правило, параметричні критерії є потужнішими за непараметричні. Застосування непараметричних критеріїв у випадках, коли можна використовувати параметричні, призводить до збільшення ймовірності прийняття помилкової нульової гіпотези, тобто помилки другого роду.

Якщо досліджувані вибірки підпорядковуються нормальному закону розподілу з відомими дисперсіями, то як критерій рівності їх середніх значень можна використовувати величину:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \quad (2.4)$$

де n_1, n_2 – кількість елементів у вибірках. Нульова гіпотеза полягає у рівності середніх.

Z-критерій є випадковою величиною, що підпорядковується стандартному нормальному розподілу. Якщо конкуруючою гіпотезою є $\bar{x}_1 \neq \bar{x}_2$, то праву критичну точку можна визначити з умови:

$$P(0 < Z < z_{\gamma \delta \alpha \hat{\alpha} \hat{\epsilon} \delta}) = \Phi(z) = (1 - \alpha) / 2, \quad (2.5)$$

де $\Phi(z)$ є функцією Лапласа (інтегралом ймовірностей), що пов'язана з функцією стандартного нормального розподілу $F(z)$ співвідношенням $\hat{O}(z) = F(z) - 1/2$. Ліва й права критичні точки для одного й того самого рівня значущості пов'язані умовою $z_{\beta \hat{\alpha} \hat{\epsilon} \delta} = -z_{\gamma \delta \alpha \hat{\alpha} \hat{\epsilon} \delta}$.

Нехай, наприклад, при вимірюванні певного параметра у двох серіях однакових виробів, виготовлених на різних установках, отримали такі значення: $\bar{x}_1 = 184$, $\bar{x}_2 = 181$, $\sigma_1^2 = 121$, $\sigma_2^2 = 107$, $n_1 = n_2 = 40$. Як нульову гіпотезу візьмемо твердження, що немає істотної різниці між параметрами виробів, виготовлених на різних установках. Тоді:

$$Z = \frac{184 - 181}{\sqrt{\frac{121}{40} + \frac{107}{40}}} \approx 1,26.$$

Беручи рівень значущості рівним 0,05 (5%), отримаємо для визначення критичної точки умову:

$$\Phi(z_{\gamma \delta \alpha \hat{\alpha} \hat{\epsilon} \delta}) = (1 - \alpha) / 2 = 0,475.$$

Для визначення величини $z_{\gamma \delta \alpha \hat{\alpha} \hat{\epsilon} \delta}$ можна використати функцію НОРМСТОБР електронних таблиць MS Excel, беручи як значення аргументу величину (0,5 + 0,475). Після цього одержимо: $z_{\gamma \delta \alpha \hat{\alpha} \hat{\epsilon} \delta} = 1,96$. Оскільки $Z < z_{\gamma \delta \alpha \hat{\alpha} \hat{\epsilon} \delta}$, ми можемо прийняти нульову гіпотезу на рівні значущості 0,05.

Використовуючи вбудовані функції електронних таблиць MS Excel, ми можемо отримати більш точну оцінку ймовірності відхилення правильності нульової гіпотези. Величина $z_{\gamma \delta \alpha \hat{\alpha} \hat{\epsilon} \delta} = 1,26$ для двобічної гіпотези відповідає рівню значущості $\alpha \approx 0,209$. Тобто, якщо ми не приймаємо нульову гіпотезу, то ймовірність помилки першого роду становить приблизно 21%. Таку імовірність у більшості випадків вважають занадто високою. Це пов'язано з тим, що зазвичай висновок формулюють як наявність чи відсутність достатніх підстав для відхилення (а не для прийняття) нульової гіпотези. При необхідності більш чіткого обґрунтування її прийняття виконують додаткові перевірки.

Якщо конкуруючою гіпотезою є: $\bar{x}_1 > \bar{x}_2$, то:

$$\Phi(z_{\hat{\epsilon} \delta}) = (1 - 2\alpha) / 2, \quad (2.6)$$

і нульову гіпотезу приймають, якщо $Z < z_{\hat{\epsilon} \delta}$.

Якщо конкуруючою гіпотезою є: $\bar{x}_1 < \bar{x}_2$, то критичну точку z' визначають з умови (2.6), враховуючи, що $z'_{\epsilon\delta} = -z_{\epsilon\delta}$. Нульову гіпотезу приймають, якщо $Z > -z_{\epsilon\delta}$.

Z-критерій можна застосовувати також для порівняння середніх значень довільно розподілених незалежних вибірок великого обсягу ($n_{1,2} \geq 30$), враховуючи, що в цьому разі вибіркові середні мають приблизно нормальний розподіл, а вибіркові дисперсії є достатньо точними оцінками генеральних дисперсій.

Для порівняння середніх значень вибірок застосовують **t-критерій Стьюдента**. Його запропоновано американським статистиком Уільямом Госсетом в 1908 р. за результатами дослідження проблеми скорочення кількості проб, які потрібно взяти при контролі за якістю продукції пивоварного заводу за умови забезпечення виконання вимог стандартів.

Розглядають дві незалежні нормальні вибірки з генеральних сукупностей, що мають рівні або нерівні, але відомі чи рівні невідомі дисперсії.

Значення критерію Стьюдента розраховують за формулою:

$$t = \frac{|\bar{x}_A - \bar{x}_B|}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}, \quad (2.7)$$

де σ_A^2, σ_B^2 – відомі внутрішньогрупові дисперсії;

n_A та n_B – чисельності груп. Для m груп рівної чисельності статистика має t -розподіл з кількістю степенів вільності $m(n-1)$.

У випадку, коли обсяги вибірок є малими або істотно розрізняються, а їх дисперсії є рівними, останні замінюють вибірковим середнім квадратичним відхиленням, яке розраховують за формулою:

$$s^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}; \quad (2.8)$$

якщо стандартні відхилення вибірок оцінюють за самими вибірками, або:

$$s^2 = \frac{s_1^2 n_1 + s_2^2 n_2}{n_1 + n_2 - 2}, \quad (2.9)$$

якщо їх оцінюють незалежно. Формула для визначення розрахункового значення критерію у цьому разі набуває вигляду:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}. \quad (2.10)$$

Відповідна статистика має розподіл Стьюдента з $k = n_1 + n_2 - 2$ степенями вільності.

При застосуванні критерію Стьюдента у вигляді (2.10) необхідно спочатку перевірити гіпотезу про рівність дисперсій.

Критичні точки симетричні стосовно нуля. Нульову гіпотезу відхиляють: якщо $|t| < t_{\alpha/2, k}$ при конкуруючій гіпотезі $\bar{x}_1 \neq \bar{x}_2$; якщо $t > t_{\alpha, k}$ при конкуруючій гіпотезі $\bar{x}_1 > \bar{x}_2$; якщо $t < -t_{\alpha, k}$ при конкуруючій гіпотезі $\bar{x}_1 < \bar{x}_2$.

При аналізі спряжених вибірок їх порівняння здійснюють з метою визначення наявності ефекту від певного фактора, наприклад, впливу змін у технології на якість виробленої продукції. Вимога щодо рівності дисперсій при цьому не висувається. Нульова гіпотеза полягає у відсутності різниці між середніми. Значення критерію розраховують за формулою:

$$t = \frac{\sum_{i=1}^n \delta_i}{\sqrt{\frac{n \sum_{i=1}^n \delta_i^2 - \left(\sum_{i=1}^n \delta_i\right)^2}{n-1}}}, \quad (2.11)$$

де n – кількість елементів у кожній із вибірок;

$\delta_i = x_i - y_i$, x_i та y_i – відповідні значення елементів першої та другої вибірок.

Іноді цей критерій називають **одновібірковим критерієм Стьюдента**. Відповідна статистика має розподіл Стьюдента з кількістю степенів вільності $n-1$.

Якщо дисперсії або їх відношення невідомі й припущення про рівність дисперсій є необґрунтованим, то виникає так звана **проблема Беренса – Фішера**, що полягає у перевірці нульової гіпотези про рівність вибірових середніх за таких умов. Одним з підходів до її вирішення є застосування **критерію Уелча (Крамера – Уелча)**, запропонованого Б. Уелчем в 1947 р. Його значення розраховують за формулою:

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad (2.12)$$

де s_1^2 , s_2^2 – розраховані за вибірками оцінки дисперсії.

Статистика цього критерію є приблизно такою самою, як для розподілу Стюдента з кількістю степенів вільності:

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}. \quad (2.13)$$

Порівняння з (2.7) вказує, що основною відмінністю критерію Уелча з погляду прикладного аналізу є зміна кількості степенів вільності.

F-критерій Фішера запропоновано британським біологом і статистиком Рональдом Фішером в 1920 р. Його використовують для порівняння дисперсій двох вибірок. Його значення розраховують за формулою:

$$F = s_1^2 / s_2^2, \quad (2.14)$$

де s_1^2 , s_2^2 – значення оцінок більшої та меншої дисперсії, відповідно. Кількості степенів вільності для пошуку критичного значення обирають рівними $n_1 - 1$ та $n_2 - 1$. Гіпотезу про рівність дисперсій порівнюваних сукупностей відхиляють, якщо обчислене значення перевищує табличне при заданому довірчому рівні. При цьому, якщо конкуруючою є однобічна гіпотеза $s_1^2 > s_2^2$, то як критичну точку беруть значення оберненого розподілу Фішера, що відповідає рівню значущості α при заданій кількості степенів вільності. Якщо ж конкуруючою є двобічна гіпотеза $s_1^2 \neq s_2^2$, то критичною точкою буде значення оберненого розподілу Фішера, що відповідає рівню значущості $\alpha / 2$.

Якщо перевіряють гіпотезу про рівність виправленої дисперсії вибірки з гіпотетичною генеральною дисперсією генеральної сукупності, то значення критерію розраховують як:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}. \quad (2.15)$$

При цьому, якщо конкуруючою є однобічна гіпотеза $\sigma^2 > \sigma_0^2$, то як критичну точку беруть значення оберненого χ^2 розподілу, що відповідає рівню значущості α при $k = n - 1$ степенях вільності. Нульову гіпотезу приймають, якщо $\chi^2 < \chi_{\alpha}^2(\alpha, n - 1)$.

Якщо конкуруючою є однобічна гіпотеза $\sigma^2 < \sigma_0^2$, то як критичну точку беруть значення оберненого χ^2 розподілу, що відповідає рівню значущості $1 - \alpha$ при $k = n - 1$ степенях вільності. Нульову гіпотезу приймають, якщо $\chi^2 > \chi_{1-\alpha}^2(1 - \alpha, n - 1)$.

Якщо ж конкуруючою є двобічна гіпотеза $\sigma^2 > \sigma_0^2$, то критичні точки розраховують за рівняннями:

$$\begin{aligned} P(\chi^2 < \chi_{\bar{\epsilon}\bar{\alpha}\bar{\epsilon}\bar{\delta}}^2(\alpha/2, n-1)) &= \alpha/2; \\ P(\chi^2 > \chi_{\bar{\epsilon}\bar{\delta}\bar{\alpha}\bar{\epsilon}\bar{\delta}}^2(\alpha/2, n-1)) &= \alpha/2. \end{aligned} \quad (2.16)$$

Нульову гіпотезу приймають, якщо:

$$\chi_{\bar{\epsilon}\bar{\alpha}\bar{\epsilon}\bar{\delta}}^2(\alpha/2, n-1) < \chi^2 < \chi_{\bar{\epsilon}\bar{\delta}\bar{\alpha}\bar{\epsilon}\bar{\delta}}^2(\alpha/2, n-1).$$

При цьому можна враховувати, що внаслідок симетрії розподілу:

$$P(\chi^2 < \chi_{\bar{\epsilon}\bar{\alpha}\bar{\epsilon}\bar{\delta}}^2) = 1 - P(\chi^2 > \chi_{\bar{\epsilon}\bar{\delta}\bar{\alpha}\bar{\epsilon}\bar{\delta}}^2) = 1 - \alpha/2. \quad (2.17)$$

Для оцінювання значущості отриманого значення χ^2 використовують **критерій В.І. Романовського** (запропонований радянським математиком Всеволодом Івановичем Романовським в 1928 р.):

$$R = \frac{\chi^2 - k}{\sqrt{2k}}. \quad (2.18)$$

При $R \geq 3$ значення χ^2 вважають значущим, а порівнювані вибірки – істотно різними.

Параметричні тести можна застосовувати також при **множинних порівняннях**, тобто при порівнянні двох груп вибірок одна з одною. Кожну групу задають подібно до того, як задають параметри масивів даних у методах двофакторного дисперсійного аналізу. При множинних порівняннях використовують багатовимірні узагальнення тестів, що були розглянуті вище.

2.3. Непараметричні тести

У багатьох випадках емпіричні дані не задовольняють нормальний розподіл. Тому для їх аналізу некоректно застосовувати параметричні тести. Серед непараметричних тестів важливе місце займають так звані **робастні методи**, що виявляють слабку чутливість до відхилень від стандартних умов і можуть використовуватися в широкому діапазоні реальних умов.

При перевірці нульової гіпотези про однорідність вибірок числових даних рекомендується [45] використовувати омега-квадрат критерій або (за відсутності необхідних таблиць та програмного забезпечення) критерій Смірнова.

Критерій омега-квадрат (критерій Крамера – фон Мізеса) базується на розгляді відхилення між двома емпіричними функціями розподілу (або між емпіричною й теоретичною функціями розподілу при ідентифікації закону розподілу). Вперше його запропонували у 1928–1930 р.

шведський математик Карл Харальд Крамер та американський математик і механік Ріхард фон Мізес, який народився у Львові.

У загальному вигляді критерій має вигляд:

$$\omega^2 = \int_{-\infty}^{\infty} [F_n(x) - F^*(x)]^2 dF^*(x), \quad (2.19)$$

де $F_n(x)$, $F^*(x)$ – відповідно, значення емпіричної та теоретичної функцій розподілу в точці x .

Двохвибірковий варіант критерію запропоновано в 1951 р. американським статистиком Еріхом Лео Леманом та досліджено у 1952 р. американським математиком Мюреєм Розенблаттом. Тому його іноді називають також **критерієм Лемана – Розенблатта**. На сьогодні його вважають [45] найпотужнішим з критеріїв, призначених для перевірки гіпотези про однорідність незалежних вибірок.

Спочатку визначають ранги елементів i, j для кожної вибірки. Потім елементи обох вибірок об'єднують і визначають їх ранги r_i, s_j у загальній вибірці. Розрахункове значення визначають за формулою:

$$A = m\omega^2 = \frac{1}{mn(m+n)} \left[m \sum_{i=1}^m (r_i - i)^2 + n \sum_{j=1}^n (s_j - j)^2 \right] - \frac{4mn-1}{6(m+n)}, \quad (2.20)$$

де m, n – обсяги порівнюваних вибірок. Нульову гіпотезу відхиляють, якщо розрахункове значення критерію перевищить критичне для заданого рівня значущості.

Критерій Смірнова запропонований в 1939 р. радянським математиком Миколою Васильовичем Смірновим. Він призначений для перевірки гіпотези про однорідність двох вибірок з неперервним законом розподілу. Значення статистики Смірнова можна розрахувати за формулою:

$$D_{m,n} = \sup_x |F_m(x) - G_n(x)|, \quad (2.21)$$

де $F_m(x)$, $G_n(x)$ – значення функцій розподілу вибірок у точці x .

На практиці використовують емпіричні функції розподілу вибірок і розраховують значення критерію за формулами:

$$\begin{aligned} D_{m,n}^+ &= \max_{1 \leq r \leq n} \left[\frac{r}{n} - F_m(y_r') \right] = \max_{1 \leq s \leq m} \left[G_n(x_s') - \frac{s-1}{m} \right]; \\ D_{m,n}^- &= \max_{1 \leq r \leq n} \left[F_m(y_r') - \frac{r-1}{n} \right] = \max_{1 \leq s \leq m} \left[\frac{s}{m} - G_n(x_s') \right]; \\ D_{m,n} &= \max \{ D_{m,n}^+; D_{m,n}^- \}, \end{aligned} \quad (2.22)$$

де x_s', y_r' – впорядковані за зростанням елементи досліджуваних вибірок.

Нульову гіпотезу відхиляють, якщо розрахункове значення критерію перевищить критичне для відповідного рівня значущості.

Для незалежних вибірок можна застосовувати **критерій рандомізації компонент**. Він розроблений Р. Фішером у 1920 р. для аналізу вибірок малого обсягу.

При порівнянні спряжених вибірок основою методу є перебирання можливих результатів, побудованих з різницевих оцінок. Нульова гіпотеза полягає у рівності вибірових середніх. Нехай дані вибірки x_i, y_i ($i = 1, 2, \dots, n$), де n – кількість пар експериментальних значень. Значення різницевих оцінок визначимо за формулою:

$$s_j = \sum_{i=1}^n a_{ji} |x_j - y_j| \quad (j = 1, 2, \dots, 2^n), \quad (2.23)$$

де a_{ji} ($j = 1, \dots, 2^n; i = 1, \dots, n$) – елементи матриці можливих результатів, розраховані згідно з методикою побудови повного ортогонального плану експерименту. Вона є матрицею з 2^n рядків та n стовпців. При цьому i -й стовпець містить величини $+1$ та -1 , що чергуються з кроком 2^{j-1} . Для $n = 3$ такий план має вигляд:

$$A = \begin{pmatrix} +1 & +1 & +1 \\ -1 & +1 & +1 \\ +1 & -1 & +1 \\ -1 & -1 & +1 \\ +1 & +1 & -1 \\ -1 & +1 & -1 \\ +1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

Сума масиву різницевих оцінок $S = \sum_{j=1}^{2^n} s_j$. Кількість сприятливих результатів:

$$N = \sum_{j=1}^{2^n} n_j, \quad n_j = \begin{cases} 0 & (s_j < S) \\ 1 & (s_j \geq S) \end{cases}. \quad (2.24)$$

Однобічне P -значення розраховують за формулою $p = N / 2^n$ і порівнюють із заданим значенням довірчого рівня p^* . Якщо $p > p^*$, то нульову гіпотезу приймають на рівні значущості $1 - p^*$. Для застосування однобіч-

ного критерію обсяг вибірки має бути не нижчим ніж 5 або 6 при рівнях значущості 0,01 та 0,05, відповідно. Для двобічних критеріїв і тих самих рівнів значущості мінімальні обсяги вибірки дорівнюють, відповідно, 7 та 8. За великих обсягів вибірок час обчислень швидко збільшується і доцільно використовувати інші критерії.

При порівнянні незалежних вибірок нульова гіпотеза полягає в належності двох досліджуваних вибірок до генеральних сукупностей з однаковими середніми. Нехай є дві вибірки: x_i ($i=1, 2, \dots, n_x$) та y_j ($j=1, 2, \dots, n_y$), де n_x, n_y – кількість елементів у них. Методика тесту основана на перебиранні всіх комбінацій даних. Знаходимо величину:

$$S = \min \left\{ \sum_{i=1}^{n_x} x_i; \sum_{j=1}^{n_y} y_j \right\}. \quad (2.25)$$

Кількість сприятливих результатів визначаємо за формулою:

$$N = 2 \sum_{j=1}^{C_n^m} n_j, \quad n_j = \begin{cases} 0 & (s_j < S); \\ 1 & (s_j \geq S), \end{cases} \quad (2.26)$$

де n_j – оцінка j -го результату;

C_n^m – загальна кількість результатів;

$n = n_x + n_y$ – чисельність об'єднаної вибірки;

m – чисельність вибірки, що відповідає мінімальному значенню

$$s_j = \sum_{i=1}^m z_{ji} \quad (j=1, 2, \dots, C_n^m);$$

z_{ji} – масив сполучень з об'єднаної вибірки, який будують подібно до розглянутої раніше процедури побудови матриці можливих результатів для спряжених вибірок. Однобічне значення p розраховують як $p = N / C_n^m$. Його порівнюють із заданим рівнем значущості α . Нульову гіпотезу відхиляють, якщо $p < \alpha$ або $p > 1 - \alpha$. Як і у попередньому випадку, критерій застосовують для відносно малих вибірок. Їх мінімальний допустимий обсяг є таким самим, як і для спряжених вибірок.

W-критерій Уїлкоксона (критерій рангових сум) запропонований в 1945 р. американським хіміком і статистиком Френком Уїлкоксоном. Його застосовують для порівняння двох незалежних сукупностей за їх центральною тенденцією, тобто за центрами емпіричних функцій розподілу. Сукупності можуть мати як однакові, так і різні чисельності. Критерій оперує не числовими значеннями даних, а їх рангами – місцями у впорядкованих за згасанням або зростанням рядах даних. При його застосуванні

передбачається, що розподіли вибірок є неперервними, а нульова гіпотеза полягає в тому, що функції розподілу вибірок збігаються одна з одною.

Процедура обчислення значення критерію є близькою до обчислення критерію рандомізації компонент. Різниця полягає в тому, що замість вихідних даних використовують їх ранги. Ранжирування порівнюваних вибірок здійснюють сумісно. Вихідні дані об'єднують до однієї вибірки, впорядковують, визначають ранги елементів об'єднаної вибірки. Потім формують дві нові вибірки, елементами яких є ранги відповідних елементів вихідних вибірок. Якщо деякі значення збігаються, то відповідним спостереженням призначають середній ранг. Обчислення статистики критерію здійснюють за формулою:

$$W = \min \left\{ \sum_{i=1}^{n_1} R_i, \sum_{i=1}^{n_2} S_i \right\}, \quad (2.27)$$

де R_i – ранги вибірки, що має найменшу, а S_i – вибірки, яка має найбільшу суму рангів. Для вибірок малого обсягу (до 25 елементів) суму рангів W вибірки, що має меншу кількість елементів, порівнюють з критичним значенням, яке визначають за спеціальними таблицями [25]. Нульову гіпотезу відхиляють, якщо: $W' < w_{i\alpha i.\dot{e}\delta}(\alpha/2, n_1, n_2)$ або $W' > w_{\dot{a}\dot{a}\dot{d}\dot{d}.\dot{e}\dot{d}}(\alpha/2, n_1, n_2)$ при нульовій гіпотезі $F_1(x) \neq F_2(x)$; $W' < w_{i\alpha i.\dot{e}\delta}(\alpha, n_1, n_2)$ при нульовій гіпотезі $F_1(x) > F_2(x)$; $W' > w_{\dot{a}\dot{a}\dot{d}\dot{d}.\dot{e}\dot{d}}(\alpha, n_1, n_2)$ при нульовій гіпотезі $F_1(x) < F_2(x)$.

Статистика $W^* = \left| \frac{W - \mu_W}{\sigma_W} \right|$, де $\mu_W = \frac{n_1(N+1)}{2}$ – математичне споді-

вання, $\sigma_W^2 = \frac{n_1 n_2 (N+1)}{12}$ – дисперсія, $N = n_1 + n_2$, при збільшенні N набли-

жається до стандартного нормального розподілу. Для вибірок великого обсягу (понад 25 елементів) нульову гіпотезу відхиляють на рівні значущості α , якщо $W^* > z_{1-\alpha/2}$ для двобічної гіпотези. Якщо отримане значення α перевищує 0,02, то вводять поправку на неперервність і вважають, що нове значення найменшої суми рангів дорівнює $W + 0,5$.

При застосуванні W -критерію Уїлкоксона слід мати на увазі, що згідно з [45] він належить до так званих критеріїв зсуву. Тобто найбільш потужним він є при виявленні різниці, пов'язаної з тим, що одну з вибірок отримано додаванням одного й того самого числа до всіх елементів іншої вибірки.

Він є нечутливим до різниці дисперсій порівнюваних вибірок, а також коефіцієнтів їх асиметрії та ексцесу. Зокрема, якщо дві вибірки мають симетричні функції розподілу з однаковими середніми значеннями, але різними стандартними відхиленнями, то в об'єднаної послідовності елементи однієї вибірки матимуть підвищену кількість елементів з високими та

низькими рангами. Елементи іншої вибірки будуть мати підвищену кількість елементів із середніми значеннями рангів. Але суми рангів усіх елементів для цих двох вибірок можуть бути приблизно однаковими.

U-критерій Манна – Уїтні запропонований в 1947 р. американськими математиками Г.Б. Манном та Д.Р. Уїтні. Він призначений для перевірки нульової гіпотези про однаковість розподілу досліджуваних сукупностей або для перевірки рівності окремих параметрів цих розподілів, наприклад, середніх значень. Спостереження мають бути непарними. Цей критерій є найпотужнішим непараметричним аналогом t -критерію Стьюдента для незалежних вибірок. У деяких випадках його потужність може бути навіть більшою, ніж у t -критерію.

Обчислення здійснюють за формулами:

$$\begin{aligned} U_1 &= n_1 n_2 + n_1 (n_1 + 1) / 2 - R_1; \\ U_2 &= n_1 n_2 + n_2 (n_2 + 1) / 2 - R_2; \\ U &= \max \{U_1, U_2\}, \end{aligned} \quad (2.28)$$

де R_1, R_2 – суми рангів вибірок;

n_1, n_2 – кількість елементів у них. Якщо $n_1, n_2 > 20$, то розподіл вибірки для U -статистики наближається до нормального. Правильність обчислення величин U_1, U_2 можна перевірити за формулою:

$$U_1 + U_2 = n_1 n_2. \quad (2.29)$$

Модифікована статистика $\frac{U - \mu_U}{\sigma_U}$, де $\mu_U = \frac{n_1 n_2}{2}$ – математичне сподівання,

$\sigma_U^2 = \frac{n_1 n_2 (N + 1)}{12}$ – дисперсія, $N = n_1 + n_2$, має стандартний нормальний розподіл. Результати обчислення за цим критерієм збігаються з даними, отримуваними за W -критерієм Уїлкоксона. На цьому критерії базується багатовимірний тест Джонкхієра – Терпстра.

T-критерій Уїлкоксона (одновибірковий, знаковий ранговий критерій Уїлкоксона) запропонований Ф. Уїлкоксоном в 1945 р. Його застосовують для порівняння вибірок з попарно спряженими значеннями. Він є непараметричним аналогом t -критерію Стьюдента для спряжених вибірок. Перевіряють нульову гіпотезу про симетричність розподілу різниць спряжених значень стосовно нуля. Методика розрахунку є близькою до розрахунку значення W -критерію Уїлкоксона, але в цьому випадку оперують модулями різниць відповідних значень. Масив модулів різниць, з якого вилучені нульові значення, ранжирують. Потім рангам надають знаки різниць, обчислюють суми додатних (W^+) і від'ємних (W^-) рангів, й беруть $W^* = \min(W^-, W^+)$. Цю величину порівнюють з критичним значенням.

Для перевірки правильності розрахунків можна використовувати тотожність:

$$W^+ + W^- = \frac{N(N+1)}{2}. \quad (2.30)$$

Статистика $\frac{W^* - \mu_{W^*}}{\sigma_{W^*}}$, де $\mu_{W^*} = \frac{N(N+1)}{4}$ – математичне сподівання,

$\sigma_{W^*}^2 = \frac{N(N+1)(2N+1)}{24}$ – дисперсія, N – чисельності вибірок, має стандартний нормальний розподіл.

Розглянемо докладніше дві проблеми, що виникають при застосуванні критерію Уїлкоксона та подібних йому рангових критеріїв [23]. Першою з них є проблема рівних рангів, яка існує лише при застосуванні нормальної або іншої апроксимації критерію і не виникає при точному обчисленні відповідних статистик. Для W -критерію Уїлкоксона у цьому випадку необхідно врахувати поправку до дисперсії та розраховувати її за формулою:

$$\sigma_W^2 = \frac{n_1 n_2}{12} \left[N + 1 - \frac{T}{N(N-1)} \right], \quad (2.31)$$

де T – поправка на об'єднання рангу, формулу для обчислення якої наведено у розділі 1, $N = n_1 + n_2$. Для U -критерію Манна – Уїтні модифікований вираз для дисперсії має аналогічний вигляд. Для T -критерію Уїлкоксона скореговану дисперсію визначають за формулою:

$$\sigma_{W^*}^2 = \frac{2N(N+1)(2N+1) - T}{48}, \quad (2.32)$$

де N – чисельність кожного ряду.

Інша проблема виникає тільки для T -критерію Уїлкоксона і полягає в тому, що наявність збігів одночасно в обох аналізованих спряжених вибірках призводить до нульових різниць. На сьогодні ця проблема залишається недостатньо дослідженою. Один з підходів до її вирішення полягає у викреслюванні рівних спряжених значень із порівнюваних вибірок. При цьому обсяги вибірок зменшуються на кількість викреслених значень.

Критерій χ^2 (хі-квадрат) запропонований в 1900 р. видатним британським математиком, біологом та філософом Карлом Пірсоном. Його використовують для перевірки нульової гіпотези про однаковість розподілу досліджуваних випадкових величин. Його широко застосовують у дисперсійному аналізі та інших методах аналізу даних. Цей критерій оперує не первинними даними, а їх розподілом за класами. З огляду на це необхідно враховувати вимогу щодо мінімальних обсягів вибірок та кількостей

класів. За різними оцінками, мінімальна допустима кількість класів знаходиться у межах 4–7, а кількість елементів у вибірках – у межах 20–40.

Якщо аналізовані дані вимірювали в кількісних або порядкових шкалах, то при порівнянні вибірок однакового обсягу значення критерію обчислюють за формулою:

$$\chi^2 = \sum_{i=1}^N \frac{(f_i - g_i)^2}{f_i + g_i}, \quad (2.33)$$

де f_i, g_i ($i = 1, 2, \dots, N$) – частоти розподілів порівнюваних вибірок;
 N – кількість класів.

За тих самих умов для вибірок різного обсягу значення критерію обчислюють за формулою:

$$\chi^2 = \frac{1}{n_1 n_2} \sum_{i=1}^N \frac{(n_2 f_i - n_1 g_i)^2}{f_i + g_i}, \quad (2.34)$$

де n_1, n_2 – кількості спостережень у порівнюваних масивах.

Критерій χ^2 можна застосовувати також і для порівняння вибірок значень номінальних ознак. У цьому випадку аналізують дані, подані у вигляді таблиці спряженості ознак. Елементами таблиці є числа, рівні кількостям елементів досліджуваних вибірок, для яких досліджувана ознака набуває значень, котрі відповідають певному класу. Кожний рядок таблиці характеризує розподіл елементів відповідної вибірки за класами, а кожний стовпець – наповненість певного класу в різних вибірках. Значення критерію розраховують за формулою:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(a_{ij} - e_{ij})^2}{e_{ij}}, \quad (2.35)$$

де a_{ij} ($i = 1, 2, \dots, r; j = 1, 2, \dots, c$) – елементи таблиці спряженості ознак;
 r – кількість вибірок (стовпців); c – кількість класів (рядків);
 e_{ij} – очікувані величини, що відповідають значенням a_{ij} .

Їх обчислюють як доданок i -го вектора-стовпця на j -й вектор-рядок, поділений на суму елементів усієї таблиці $\sum_{i=1}^r \sum_{j=1}^c a_{ij}$. Кількість степенів вільності при обчисленні P -значення статистики χ^2 беруть рівною $(r-1)(c-1)$.

Критерій серій Вальда – Волфовиця розроблений в 1940 р. американськими математиками Абрахамом Вальдом і Джекобом Волфовицем. Його використовують для перевірки нульової гіпотези про те, що дві неза-

лежні випадкові вибірки обсягами n_1 та n_2 не відрізняються одна від одної за досліджуваною ознакою.

Результати спостережень записують як варіаційний ряд об'єднаної вибірки, а їх належність до вихідних вибірок помічають за допомогою додаткової змінної, яка може набувати два значення, наприклад “0” та “1”. Послідовність її значень називають **послідовністю кодів**.

Серією послідовності кодів називають будь-яку послідовність її однакових значень. Наприклад, у послідовності 00101111011 є такі серії: 00, 1, 0, 1111, 0, 11. Очевидно, що за умови справедливості нульової гіпотези кількість серій N має бути великою, а за умови її помилковості – відносно малою. Якщо обсяги вибірок є достатньо великими ($n_1, n_2 > 20$) для перевірки нульової гіпотези можна використовувати статистику:

$$Z = \frac{\left| N - \left(\frac{2n_1n_2}{n_1 + n_2} \right) + 1 \right| - \frac{1}{2}}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}}, \quad (2.36)$$

яка має стандартний нормальний розподіл.

Критерій знаків запропонований Ф. Уїлкоксоном. Його використовують для перевірки нульової гіпотези про однорідність двох спряжених вибірок. Нехай x_i і y_i – значення відповідних елементів цих вибірок. Якщо вибірки є однорідними, то ймовірності появи додатних та від'ємних різниць $x_i - y_i$ є рівними. Імовірність появи нульових значень цих різниць вважається нульовою, оскільки передбачається, що розподіл досліджуваної ознаки є неперервним. Якщо внаслідок випадкових похибок або округлення результатів такі різниці з'являються, то відповідні спостереження виключають з подальшого аналізу. За умови справедливості нульової гіпотези ймовірність p появи знаків певного типу (наприклад знаків “+”) різниць $x_i - y_i$ підпорядковується біноміальному розподілу з параметрами: $p = 1/2$; m , де m – кількість різниць, що аналізують.

Нульовою гіпотезою є $H_0 : p = 1/2$. Як конкуруючі можна розглядати такі гіпотези: $H_1^{(1)} : p > 1/2$; $H_1^{(2)} : p < 1/2$; $H_1^{(3)} : p \neq 1/2$. Нульову гіпотезу відхиляють на рівні значущості α , якщо:

– при конкуруючій гіпотезі $H_1^{(1)}$ виконується нерівність:

$$\sum_{i=r}^m C_m^i \left(\frac{1}{2} \right)^m \leq \alpha, \quad (2.37)$$

де r – кількість додатних різниць;

C_m^i – кількість сполучень;

– при конкуруючій гіпотезі $H_1^{(2)}$ виконується нерівність:

$$\sum_{i=0}^r C_m^i \left(\frac{1}{2}\right)^m \leq \alpha, \quad (2.38)$$

– при конкуруючій гіпотезі $H_1^{(3)}$ виконується одна з нерівностей:

$$\sum_{i=0}^r C_m^i \left(\frac{1}{2}\right)^m \leq \frac{\alpha}{2}; \quad (2.39)$$

$$\sum_{i=r}^m C_m^i \left(\frac{1}{2}\right)^m \leq \frac{\alpha}{2}. \quad (2.40)$$

Для обчислення кількості сполучень можна використовувати такі апроксимації біноміального розподілу нормальним:

$$\sum_{i=0}^r C_m^i \left(\frac{1}{2}\right)^m \approx \Phi\left(\frac{r - m/2}{\sqrt{m/4}}\right), \text{ якщо } m > 50; \quad (2.41)$$

$$\sum_{i=0}^r C_m^i \left(\frac{1}{2}\right)^m \approx \Phi\left(\frac{r - m/2 + 0,5}{\sqrt{m/4}}\right), \text{ якщо } m \leq 50, \quad (2.42)$$

де $\Phi(x)$ – функція розподілу для стандартного нормального закону.

2.4. Визначення моделей розподілу емпіричних даних

На практиці часто виникає проблема перевірки відповідності емпіричного розподілу деякому заданому теоретичному. При цьому вирізняють прості та складні гіпотези. Якщо гіпотеза стверджує, що із ℓ параметрів розподілу k мають задані значення, то гіпотезу вважають простою, коли $k = \ell$, і складною – якщо $k < \ell$. Різницю $\ell - k$ називають **кількістю степенів вільності гіпотези**, а k – **кількістю накладених обмежень**.

Особливу роль відіграє перевірка розподілу на нормальність, оскільки її прийняття дає змогу застосовувати більш досліджені параметричні критерії перевірки наступних гіпотез.

Для перевірки відповідності емпіричного розподілу теоретичному застосовують так звані **критерії згоди**: ω^2 , Смірнова, χ^2 , Ястремського, Бернштейна та інші.

Критерій ω^2 (Крамера – фон Мізеса) запропонований в 1928–1930 р. К. Крамером та Р. фон Мізесом. Його використовують у випадках, коли необхідно перевірити нульову гіпотезу про відповідність ви-

бірки певному відомому закону розподілу. Розрахункове значення обчислюють за формулою:

$$n\omega^2 = \frac{1}{12n} + \sum_{i=1}^n \left[F(x_i) - \frac{2i-1}{2n} \right]^2, \quad (2.43)$$

де $F(x)$ – теоретична функція розподілу, n – обсяг вибірки.

При $n > 40$ критичні значення визначають згідно з табл. 2.1 [37]. Якщо значення параметрів розподілу визначають за вибіркою, то критичні значення суттєво зменшуються.

Таблиця 2.1

Критичні значення статистики ω^2

α	0,900	0,950	0,990	0,995	0,999
$n\omega^2(\alpha)$	0,3473	0,4614	0,7435	0,8694	1,1679

Критерій Смірнова у вигляді (2.21) застосовують, якщо емпіричну функцію розподілу будують за масивом частот. У випадку, коли її побудову здійснюють безпосередньо за вихідною вибіркою (при цьому чисельність вихідної вибірки та масиву функції розподілу збігаються), для розрахунку критерію при двобічній гіпотезі застосовують формули:

$$D_n = \max_{1 \leq m \leq n} \{ D_n^{(1)}, D_n^{(2)} \}; \quad (2.44)$$

$$D_n^{(1)} = \max_{1 \leq m \leq n} \left\{ \frac{m}{n} - F(x_m) \right\}; \quad D_n^{(2)} = \max_{1 \leq m \leq n} \left\{ F(x_m) - \frac{m-1}{n} \right\},$$

а при однобічній – $D_n = D_n^{(1)}$.

Функція розподілу D_n є однією й тією самою для всіх неперервних розподілів, а функція розподілу величини $K = D_n \sqrt{n}$ за великих n збігається до **статистики Колмогорова – Смірнова**, тому в літературі його часто називають **критерієм Колмогорова – Смірнова**. Критерій Смірнова за певних умов можна використовувати також для порівняння емпіричних функцій розподілу двох вибірок (перевірка гіпотези про їх однорідність).

Обмеженнями для застосування цього критерію є:

- вимога щодо неперервності теоретичної функції розподілу;
- необхідність мати достатньо представницькі вибірки ($n > 200$);
- необхідність незалежних, тобто отриманих не за самою вибіркою, оцінок параметрів розподілу (проста гіпотеза) для порівняння емпіричної й теоретичної функцій розподілу.

Якщо обсяги вибірок $n > 35$, то критичне значення статистики Колмогорова – Смірнова, що відповідає рівню значущості α , можна розрахувати за формулою:

$$K_\alpha \approx \sqrt{-\frac{\ln \frac{\alpha}{2}}{2}}. \quad (2.45)$$

При застосуванні критерію Смірнова для перевірки нульової гіпотези про однорідність двох вибірок достатньо великого обсягу ($n_1, n_2 > 40$) можна використовувати [19] такі критичні значення:

$$\text{при } \alpha = 0,05 \quad D_c = 1,36 \sqrt{\frac{n_1 n_2}{n_1 + n_2}}; \quad (2.46)$$

$$\text{при } \alpha = 0,10 \quad D_c = 1,22 \sqrt{\frac{n_1 n_2}{n_1 + n_2}}. \quad (2.47)$$

Для складних гіпотез вводять модифіковані статистики Колмогорова, але їх функції розподілу є різними для різних типів неперервних розподілів. На відміну від більшості інших критеріїв, критерій Смірнова за достатнього обсягу досліджуваної вибірки дає змогу встановити різницю емпіричної й теоретичної функцій розподілу незалежно від параметрів, що її зумовлюють. Зокрема він є чутливим до різниці як вибірових середніх, так і стандартних відхилень вибірок, коефіцієнтів їх асиметрії та ексцесу. Але така універсальність досягається за рахунок зменшення потужності критерію.

Критерій χ^2 як критерій згоди застосовують для порівняння емпіричної та теоретичної функцій розподілу. Він оперує не первинними даними, а їх розподілом за класами рівної ширини. Тому необхідно враховувати вимогу щодо мінімальних обсягів ряду спостережень і кількості класів. За різними оцінками, мінімальна допустима кількість класів знаходиться у межах 4–7, а кількість елементів у ряді спостережень – в межах 20–200.

Значення критерію розраховують за формулою:

$$\chi^2 = \sum_{i=1}^k \frac{(v_i - np'_i)^2}{np'_i}, \quad (2.48)$$

де v_i – абсолютні частоти для k класів; p'_i – теоретичні ймовірності обраного розподілу (параметри теоретичного розподілу розраховують за емпіричною вибіркою або задають); n – загальна кількість спостережень (для неперервного розподілу цю величину треба помножити на довжину класового інтервалу d). Кількість степенів вільності беруть рівною $k - r - 1$, де r – кількість параметрів теоретичного розподілу. Зокрема, при обчисленні параметрів теоретичного розподілу за інтервальним варіаційним рядом кількість степенів вільності беруть рівною $k - 2$ для біноміального і $k - 3$ – для нормального розподілу.

Загальна схема застосування критерію χ^2 є такою. Спочатку будують емпіричну функцію розподілу. Потім на основі її аналізу визначають r параметрів, які необхідні для побудови теоретичної функції розподілу. Далі визначають межі класових інтервалів, абсолютні частоти v_i й теоре-

тичні ймовірності p'_i . Після цього розраховують значення критерію χ^2 і порівнюють його з критичним.

Розглянуті вище критерії можна застосовувати лише для вибірок достатньо великого обсягу. Для перевірки нормальності вибірок обсягом 3–50 значень використовують **W-критерій Шапіро – Уїлка**. Він запропонований в 1965 р. американським статистиком Самуелом Шапіро й канадським статистиком Мартіном Уїлком і є одним з найефективніших методів вирішення цього завдання. Цей критерій базується на регресії порядкових статистик. Його обчислення здійснюють за формулами:

$$W = b^2 / S^2; S^2 = \sum_{i=1}^n (x_i - \bar{x})^2; b = \sum_{i=1}^k a_{n-i+1} (x_{n-i+1} - x_i), \quad (2.49)$$

де x_i ($i=1, 2, \dots, n$) – ранжируваний ряд; n – обсяг вибірки, параметр k беруть рівним $n/2$ для парних i $(n-1)/2$ для непарних n , a_{n-i+1} ($i=1, 2, \dots, k; n=3, 4, \dots, 50$) – константи. Гіпотезу про нормальний розподіл приймають, якщо значення критерію перевищує критичну для заданого довірчого рівня величину.

Перевірку нормальності розподілу можна здійснити також за результатами аналізу процентилей. Розподіл можна вважати близьким до нормального, якщо значення окремих процентилей є близькими до наведених нижче величин:

- 2,5% центиль $\approx \mu - 2\sigma$;
- 16% центиль $\approx \mu - \sigma$;
- 50% центиль $\approx \mu$;
- 84% центиль $\approx \mu + \sigma$;
- 97,5% центиль $\approx \mu + 2\sigma$.

Найпростіші способи перевірки нормальності вибірки базуються на розрахунку значень коефіцієнтів його асиметрії та ексцесу. Вважають, що розподіл є близьким до нормального, якщо $|A|, |E| < 0,1$, і сильно відрізняється від такого, коли $|A|, |E| > 0,5$.

Ще одним способом перевірки типу розподілу є побудова емпіричної функції розподілу в певних координатах, що лінеаризують її графік. У статистичних пакетах SPSS, Statistica та інших реалізовано спеціальні засоби такої перевірки. Як приклад на рис. 2.3, 2.4 наведені лінеаризовані графіки емпіричних функцій розподілу вибірок з генеральних сукупностей, що підпорядковуються рівномірному розподілу на відрізку $[-3, 3]$ і стандартному нормальному розподілу.

На рис. 2.3 графіки побудовано в координатах, що мають лінеаризувати функцію нормального розподілу, а на рис. 2.4 – функцію рівномірного розподілу.

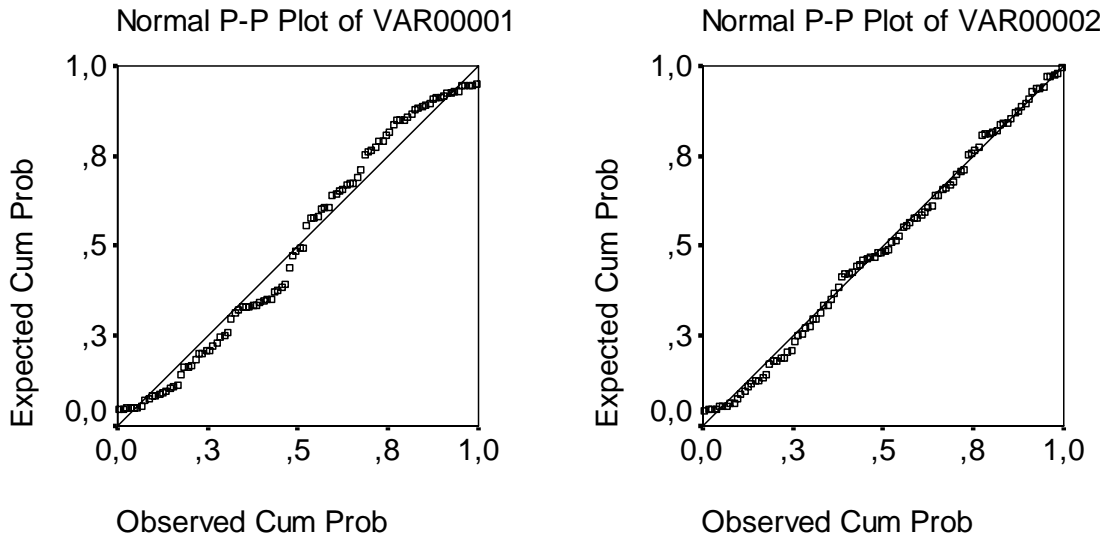


Рис. 2.3. P-P діаграми рівномірно й нормально розподілених вибірок у координатах, що лінеаризують функцію нормального розподілу

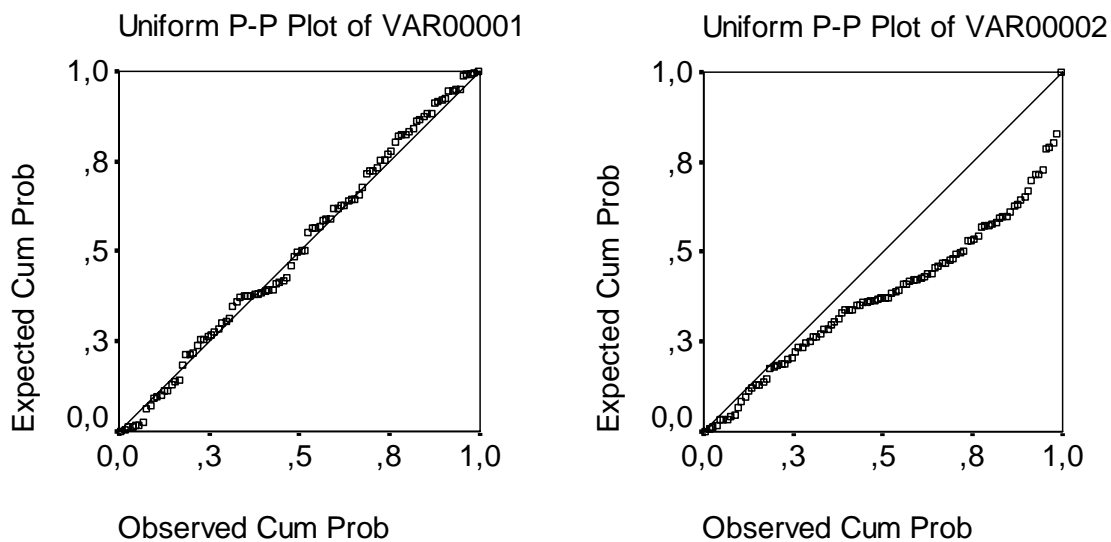


Рис. 2.4. P-P діаграми рівномірно й нормально розподілених вибірок у координатах, що лінеаризують функцію рівномірного розподілу

2.5. Приклад ідентифікації функції розподілу однорідної вибірки

Ідентифікація емпіричних функцій розподілу є відносно простою для однорідних вибірок. У цьому випадку її алгоритм може бути таким.

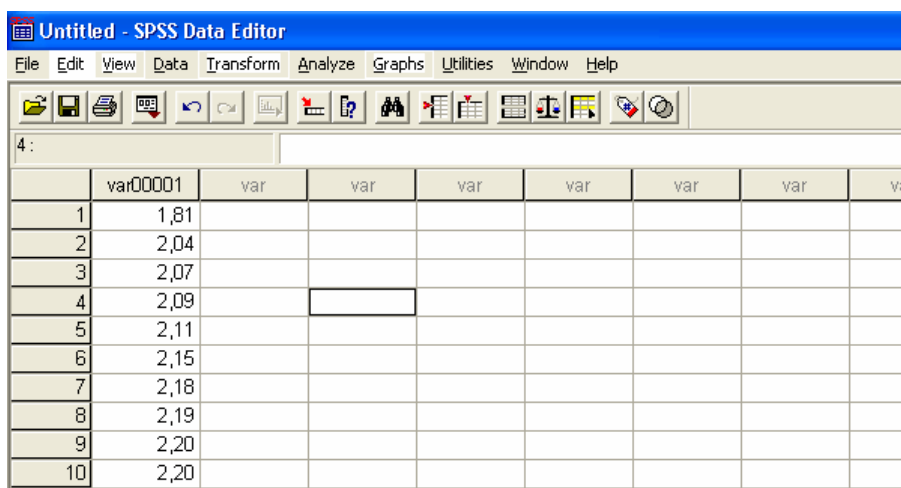
1. За допомогою P-P діаграм статистичних пакетів (SPSS, Statistica тощо) підбираємо найбільш придатний тип розподілу.

2. Використовуючи статистичні пакети або мінімізуючи суму квадратів залишків моделі за допомогою процедури “Пошук розв’язку” електронних таблиць MS Excel, уточнюємо параметри розподілу.

3. Перевіряємо адекватність підбраної моделі розподілу, використовуючи критерії Крамера – Уелча та Фішера (для нормального розподілу), ω^2 або Смирнова (для інших типів розподілу).

4. Розглянемо як приклад завдання ідентифікації моделі розподілу питомого електричного опору епітаксійних шарів кремнієвих композицій [13].

До робочого вікна пакету SPSS (рис. 2.5) вводимо значення елементів досліджуваної вибірки, що є значеннями питомого електричного опору епітаксійного шару досліджуваної серії виробів. У головному меню обираємо пункти Graphs / P-P Plots. При цьому відчиняється діалогове вікно (рис. 2.6).



4:	var00001	var	var	var	var	var	var	va
1	1,81							
2	2,04							
3	2,07							
4	2,09							
5	2,11							
6	2,15							
7	2,18							
8	2,19							
9	2,20							
10	2,20							

Рис. 2.5. Головне вікно для уведення даних пакету SPSS

У цьому вікні зазначаємо, для якої вибірки треба побудувати P-P діаграму, а також, який саме тип розподілу перевірятимемо. При цьому можливо вибрати такі типи розподілу: бета, χ^2 , експоненціальний, гама, напівнормальний, Лапласа, логістичний, логнормальний, нормальний, Парето, Стьюдента, Вейбула, рівномірний.

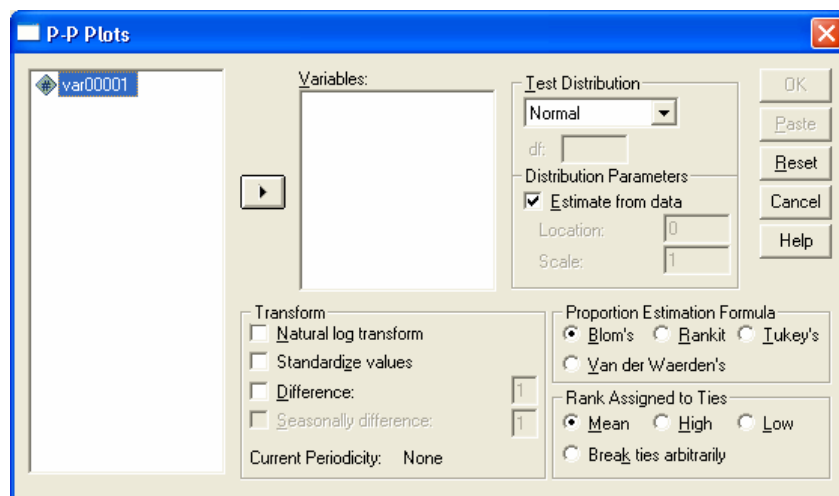


Рис. 2.6. Діалогове вікно побудови P-P діаграм

Для досліджуваної вибірки одержуємо такі результати. Бета розподіл підібрати не вдається (параметри, що визначаються при мінімізації цільового функціонала, не відповідають обмеженням цього розподілу). Розподіли χ^2 , експоненціальний, напівнормальний, Парето, Стюдента й рівномірний (рис. 2.7–2.12) суттєво відрізняються від розподілу досліджуваної вибірки.

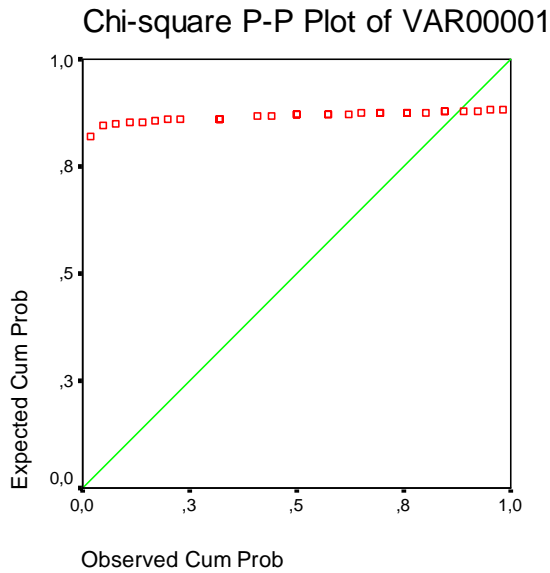


Рис. 2.7. P-P діаграма досліджуваної вибірки для χ^2 розподілу

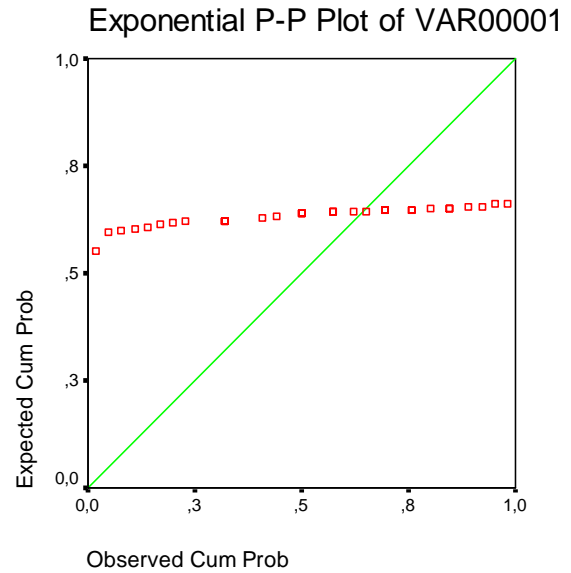


Рис. 2.8. P-P діаграма досліджуваної вибірки для експоненціального розподілу

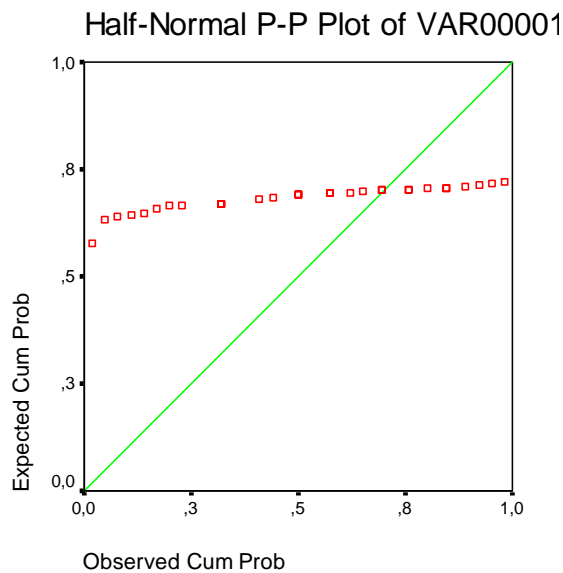


Рис. 2.9. P-P діаграма досліджуваної вибірки для напівнормального розподілу

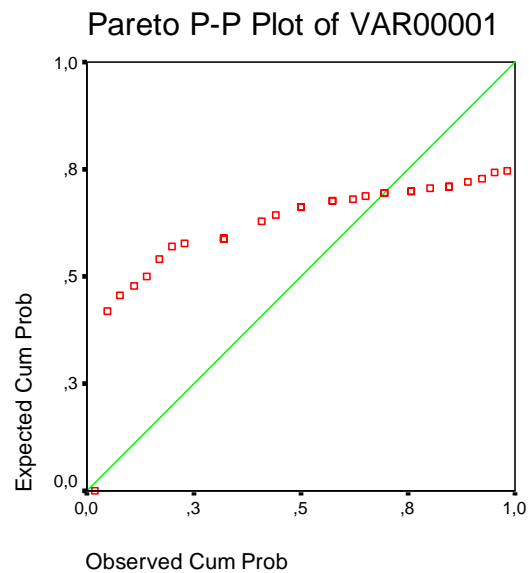


Рис. 2.10. P-P діаграма досліджуваної вибірки для розподілу Парето

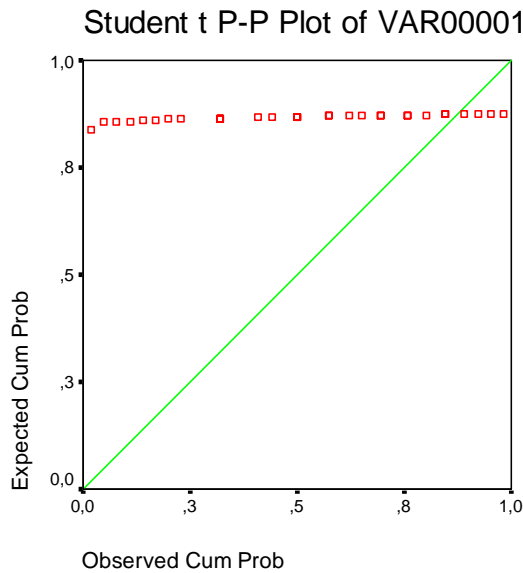


Рис. 2.11. P-P діаграма досліджуваної вибірки для розподілу Стьюдента

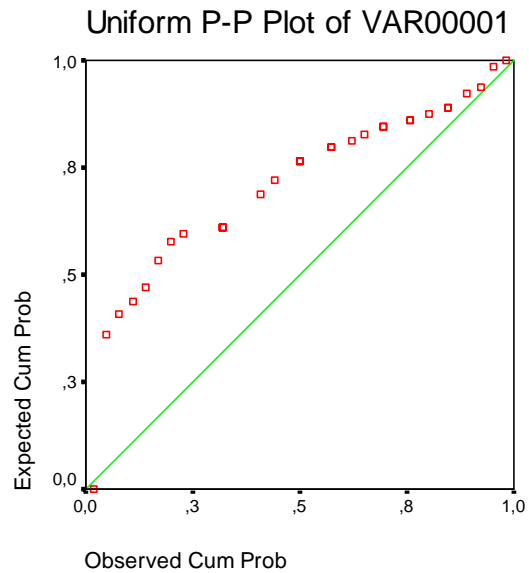


Рис. 2.12. P-P діаграма досліджуваної вибірки для рівномірного розподілу

Більш придатними є розподіли гама з параметром форми 276,9 й параметром масштабу 122,4 (рис. 2.13), Лапласа (рис. 2.14), логістичний з параметром розташування 2,26 й параметром масштабу 0,0749 (рис. 2.15), логнормальний з параметром масштабу 2,257 й параметром форми 0,0627 (рис. 2.16), нормальний (рис. 2.17) і Вейбулла з параметром масштабу 2,324 й параметром форми 19,08 (рис. 2.18). При цьому для нормального розподілу і розподілу Лапласа параметри моделі не визначаються.

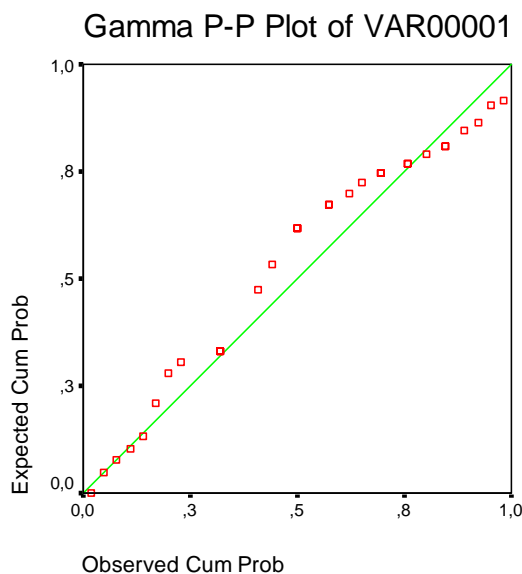


Рис. 2.13. P-P діаграма досліджуваної вибірки для гама розподілу

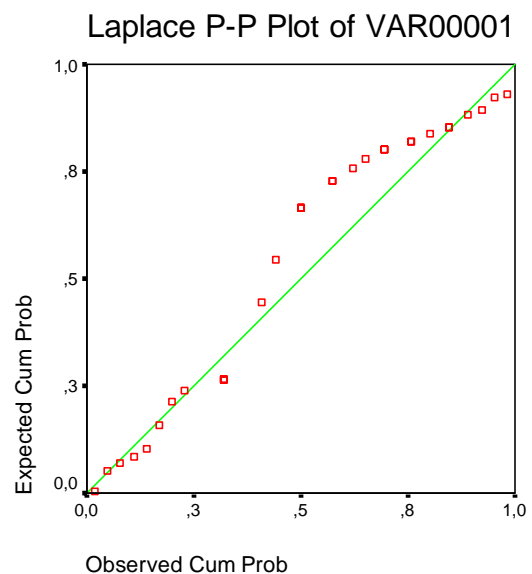


Рис. 2.14. P-P діаграма досліджуваної вибірки для розподілу Лапласа

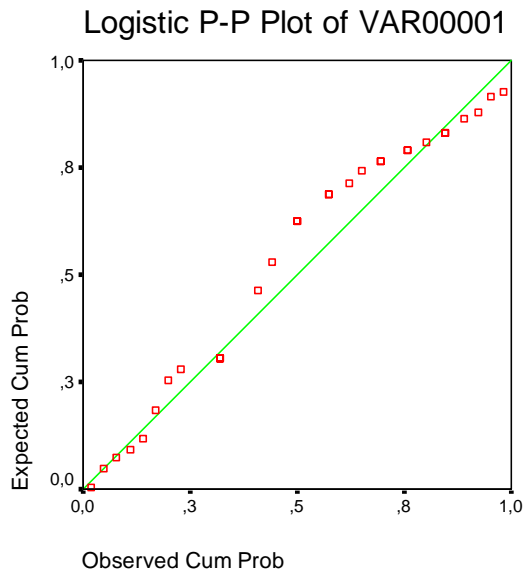


Рис. 2.15. P-P діаграма досліджуваної вибірки для логістичного розподілу

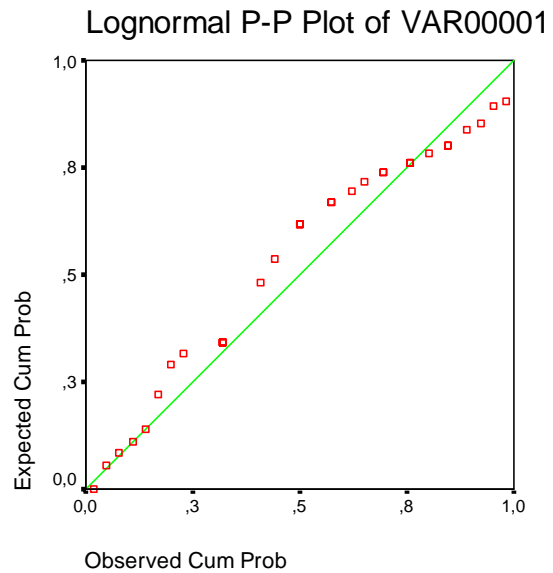


Рис. 2.16. P-P діаграма досліджуваної вибірки для логнормального розподілу

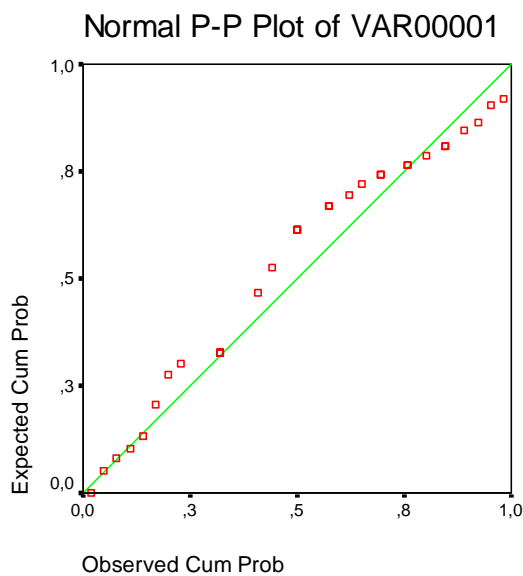


Рис. 2.17. P-P діаграма досліджуваної вибірки для нормального розподілу

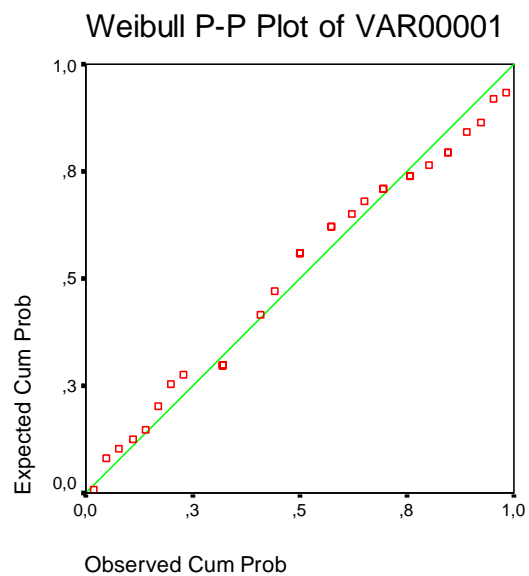


Рис. 2.18. P-P діаграма досліджуваної вибірки для розподілу Вейбулла

З наведених P-P діаграм бачимо, що найбільш придатним для опису досліджуваної вибірки є розподіл Вейбулла.

Для перевірки гіпотези, що вибірка підпорядковується розподілу Вейбулла, використовуємо електронні таблиці MS Excel. Для цього у робочому вікні сформуємо стовпчик даних, що містить елементи досліджуваної вибірки. Потім впорядкуємо дані за зростанням. Для цього у головному меню обираємо пункт “Дані” і підпункт “Сортування”. При цьому з’являється діалогове вікно (рис. 2.19).

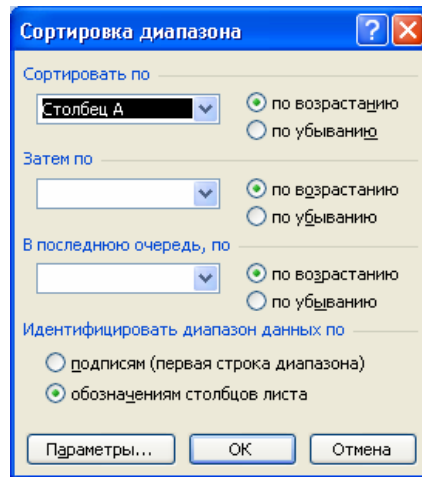


Рис. 2.19. Діалогове вікно сортування даних

Після цього у сусідньому стовпчику поруч з кожним значенням вихідної вибірки наведемо величину n / N , де n – ранг спостереження, N – обсяг вибірки. Отримані величини є значеннями емпіричної функції розподілу $f(x_i)$, а відповідні елементи досліджуваної вибірки – значеннями аргументу x_i .

Для перевірки гіпотези про тип розподілу розрахуємо значення функції розподілу Вейбулла. Для цього у третьому стовпчику кожному значенню x_i приведемо у відповідність величину, що визначається формулою:

$$=(1-\text{EXP}(-\text{СТЕПЕНЬ}(A4/\$N\$20;\$O\$20))), \quad (2.50)$$

де $\$N\20 , $\$O\20 – посилання на комірки, куди уведено значення параметрів розподілу. Як початкові, будемо використовувати значення, розраховані у пакеті SPSS. Далі формуємо стовпчик модулів різниць відповідних значень емпіричної й теоретичної функцій розподілу. В окремій комірці розраховуємо суму квадратів цих різниць. Далі використовуємо цю комірку як цільову у процедурі “Пошук розв’язку”. Для її застосування у головному меню обираємо пункт “Сервіс” і підпункт “пошук розв’язку”. При цьому з’являється діалогове вікно (рис. 2.20).

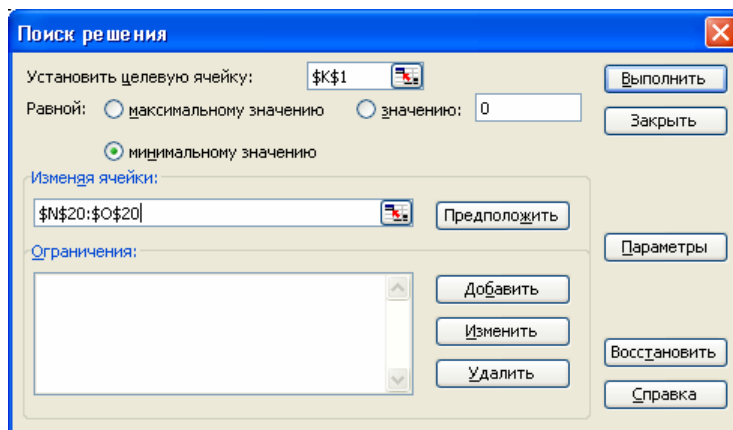


Рис. 2.20. Діалогове вікно процедури “Пошук розв’язку”

У цьому вікні зазначаємо, що шукатимемо мінімальне значення, і й робимо посилання на цільову комірку й комірки, де розташовані значення параметрів розподілу, які треба змінювати для отримання мінімального значення цільової комірки. Крім того встановлюємо значення параметрів процедури. Для цього натискаємо кнопку параметри, і з'являється діалогове вікно (рис. 2.21).

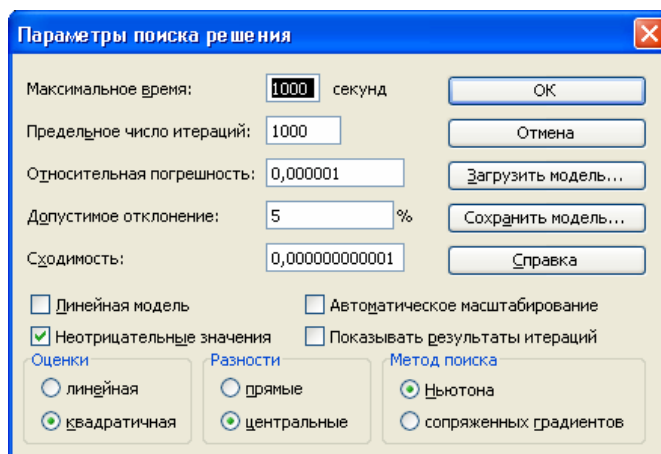


Рис. 2.21. Діалогове вікно встановлення параметрів процедури “Пошук розв’язку”

Для задачі, що розглядається, необхідно задати такі параметри:

- максимальний час;
- гранична кількість ітерацій;
- збіжність;
- невід’ємні значення;
- квадратичні оцінки;
- центральні різниці;
- метод Ньютона.

Інші параметри не задаємо, оскільки для нашої задачі вони є несуттєвими.

Максимальний час та граничну кількість ітерацій можна залишити рівними 100 (значення, що встановлюються за умовчанням). При підборі значень параметрів однорідних вибірок час та кількість ітерацій зазвичай не перевищують цих значень. У більш складних випадках ці параметри доцільно брати більшими. Значення параметра “Збіжність” дає максимальну величину різниці між двома послідовними значеннями, одержуваними у цільовій комірці, після досягнення якої ітерації зупиняються. Чим меншою буде ця величина, тим більшими будуть час та кількість ітерацій, необхідних для отримання розв’язку. Обмеження на невід’ємні значення пов’язано з тим, що у нашому випадку від’ємні значення не відповідають задачі, що розв’язується.

У досліджуваному випадку після оптимізації отримуємо значення параметра масштабу 2,319 й параметра форми – 20,67. Вони дещо відрізняються від значень, одержаних у пакеті SPSS.

Якщо мінімізувати не суму квадратів різниць значень емпіричної й теоретичної функцій розподілу, а максимальний модуль цих різниць, знов одержимо дещо інші значення параметрів розподілу: параметра масштабу 2,321 й параметра форми – 18,51. Бачимо, що результат визначення параметра масштабу є більш стійким і менше залежить від способу його визначення.

Для перевірки адекватності моделі розрахуємо значення критерію Смірнова, яке дорівнює в останньому випадку 0,49. Воно є значно нижчим, ніж критичне значення, яке при розрахунку параметрів розподілу безпосередньо за вибіркою дорівнює [21] 0,895 при рівні значущості 0,05.

Аналогічно ми можемо перевірити гіпотезу про інший тип розподілу. Зокрема для перевірки гіпотези про відповідність вибірки нормальному закону розподілу у комірки для розрахунку значень теоретичної функції розподілу уведемо формули:

$$=НОРМРАСП (А6; \$M\$3; \$M\$7; ИСТИНА), \quad (2.51)$$

де $\$M\3 , $\$M\7 – посилання на комірки, зі значеннями математичного сподівання й стандартного відхилення. Ці значення можна розрахувати за допомогою функцій =СРЗНАЧ (А1: А33) та =СТАНДОТКЛОН (А1: А33), де А1: А33 – посилання на діапазон комірок, в яких розташовано елементи досліджуваної вибірки.

У цьому випадку розрахункове значення критерію Смірнова дорівнює 0,73, що також менше, ніж критичне значення, але значно вище, ніж розрахункове значення для розподілу Вейбулла.

На рис. 2.22 наведено графіки емпіричної функції розподілу, а також підібраних теоретичних моделей. Не зважаючи на помітну різницю між ними, ми не маємо підстав для відхилення нульових гіпотез про відповідність розглянутих моделей однорідних розподілів наявним емпіричним даним, оскільки в обох випадках розрахункове значення критерію Смірнова не перевищує критичного.

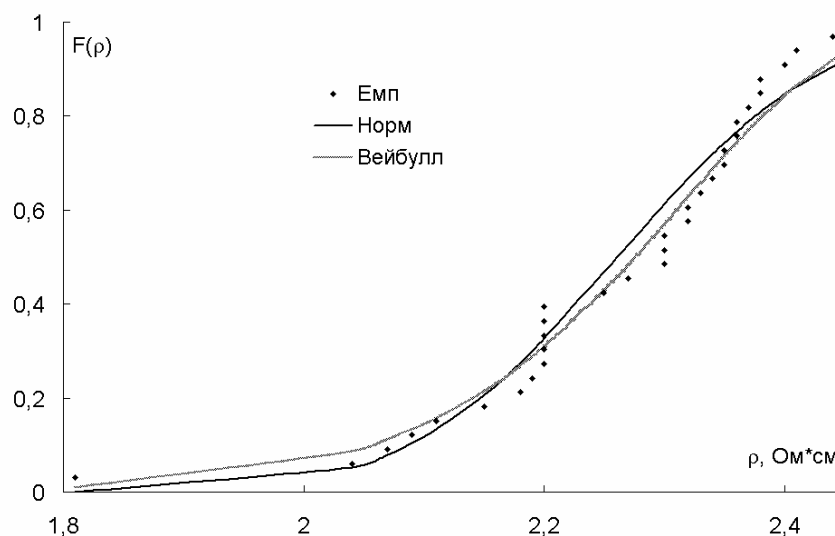


Рис. 2.22. Емпірична й теоретичні функції розподілу досліджуваної вибірки

2.6. Приклад ідентифікації функції розподілу неоднорідної вибірки

Для неоднорідних вибірок Р-Р діаграми зазвичай не дають змоги виявити більш-менш придатний для подальшої перевірки тип розподілу. Аналіз даних у цьому випадку доцільно розпочинати з побудови емпіричної функції розподілу й гістограми вибірки.

Як приклад, розглянемо результати єдиного державного екзамену з математики у Російській Федерації (ЄДЕ) у 2008 р. [14]. На рис. 2.23, 2.24 наведено функцію їх розподілу й відповідну гістограму. Гістограма свідчить про істотну неоднорідність досліджуваної вибірки. Тому далі розглядатимемо модель суміші розподілів.

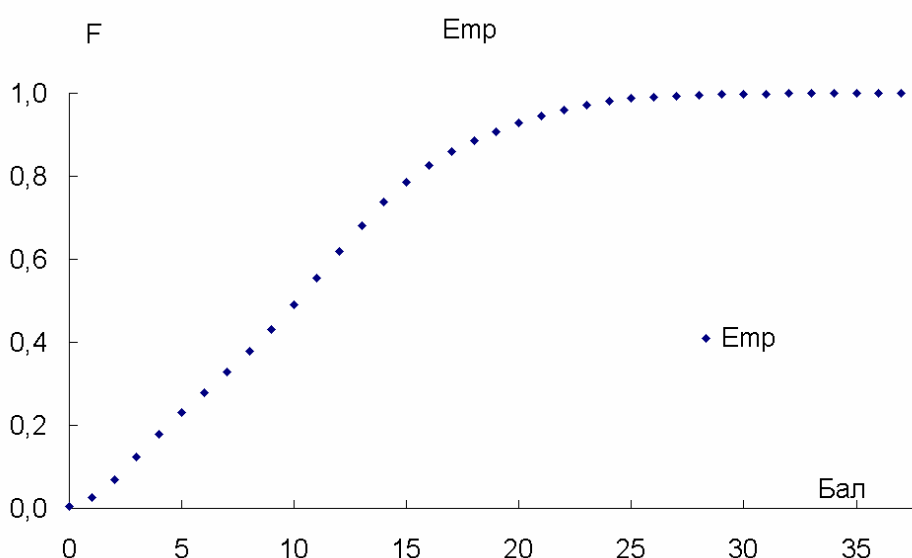


Рис. 2.23. Функція розподілу результатів ЄДЕ-2008 з математики

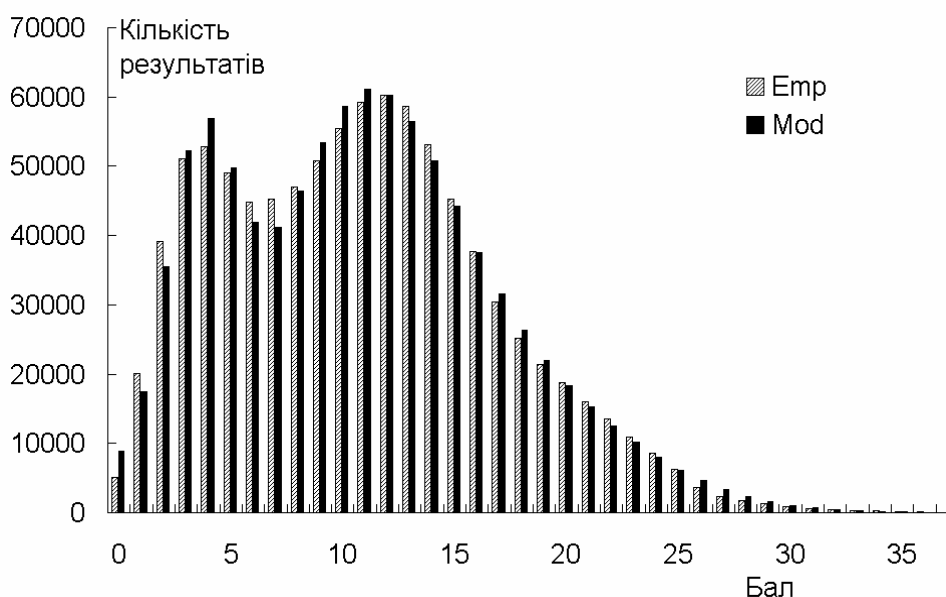


Рис. 2.24. Гістограма розподілу результатів ЄДЕ-2008 з математики

Шукатимемо модель у вигляді суміші нормально розподілених компонент:

$$F_M = \sum_i \alpha_i N(\mu_i, s_i), \quad (2.52)$$

де α_i – вагові коефіцієнти, що мають задовольняти вимогу $\sum_i \alpha_i = 1$, μ_i – математичні сподівання компонент, s_i – їх стандартні відхилення.

Для підбору параметрів моделі використовуємо процедуру “Пошук розв’язку” електронних таблиць MS Excel. Для моделей неоднорідного розподілу результат її застосування зазвичай є нестійким і залежить від вибору початкового наближення. У випадку, що розглядається, аналіз гістограми дає змогу достатньо точно визначити початкові наближення математичного сподівання і стандартного відхилення для двох компонент. Але перевірка адекватності отриманої після оптимізації моделі за критеріями Смірнова та χ^2 , а також порівняння гістограми з моделлю показує, що модель не є адекватною. Тому було перевірено модель, яка містила три компоненти. Початкові значення параметрів третьої компоненти були визначені на основі аналізу різниці між емпіричною функцією розподілу і двохкомпонентною моделлю. Після оптимізації було отримано таку модель розподілу:

$$F_M = 0,204N(3,49; 1,63) + 0,507N(10,60; 3,82) + 0,288N(16,54; 5,57) \quad (2.53)$$

Розрахункове значення критерію Смірнова дорівнює 0,03 і є значно меншим, ніж критичне. Емпірична й модельні гістограми розподілу достатньо добре відповідають одна одній (рис. 2.24). Також спостерігається добра відповідність між емпіричною й теоретичною функцією розподілу (рис. 2.25).

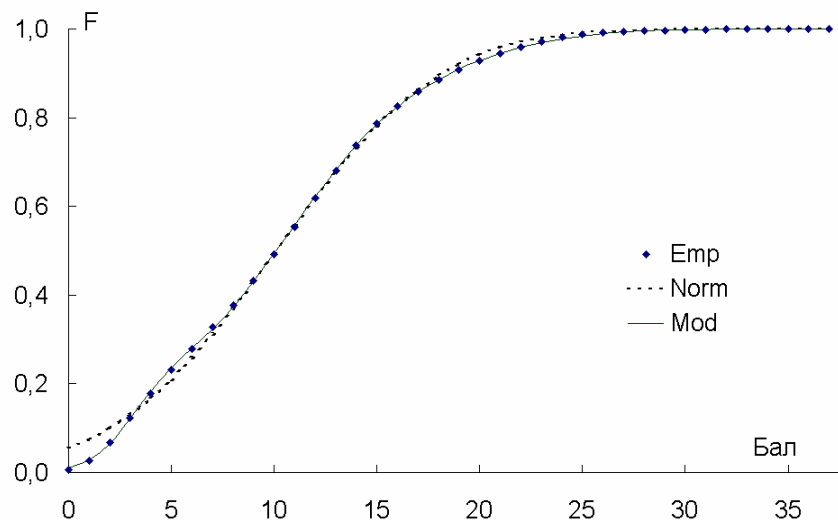
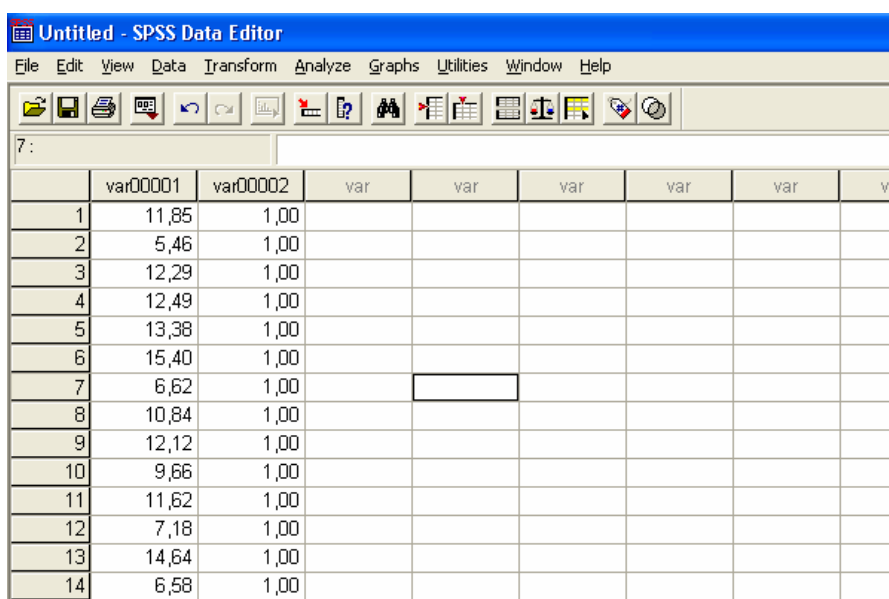


Рис. 2.25. Порівняння емпіричної та теоретичної функцій розподілу результатів єдиного державного екзамену з математики

У пакеті SPSS не передбачено спеціальних засобів для ідентифікації неоднорідних функцій розподілу. Для вирішення цієї проблеми можна використати засоби кластерного аналізу, які будуть докладно розглядатися у розділі 6.

Для ілюстрації розглянемо вибірку, побудовану, як суміш двох нормально розподілених компонент з математичними сподіваннями $a = 10$ й $b = 15$ або 20 та стандартними відхиленнями $s_{1,2} = 3$, обсягом по 100 елементів.

Уведемо досліджувані вибірки одна за одною до вікна даних пакету SPSS, як значення var00001 (рис. 2.26). Позначимо у сусідньому стовпчику (var00002), до якої з вихідних вибірок належить відповідне значення var00001. Це полегшує подальший аналіз результатів кластеризації.



	var00001	var00002	var	var	var	var	var	va
1	11,85	1,00						
2	5,46	1,00						
3	12,29	1,00						
4	12,49	1,00						
5	13,38	1,00						
6	15,40	1,00						
7	6,62	1,00						
8	10,84	1,00						
9	12,12	1,00						
10	9,66	1,00						
11	11,62	1,00						
12	7,18	1,00						
13	14,64	1,00						
14	6,58	1,00						

Рис. 2.26. Фрагмент вікна з вихідними даними

Далі у головному меню обираємо Analyze/Classify/K-means Cluster Analysis. При цьому з'являється діалогове вікно (рис. 2.27).

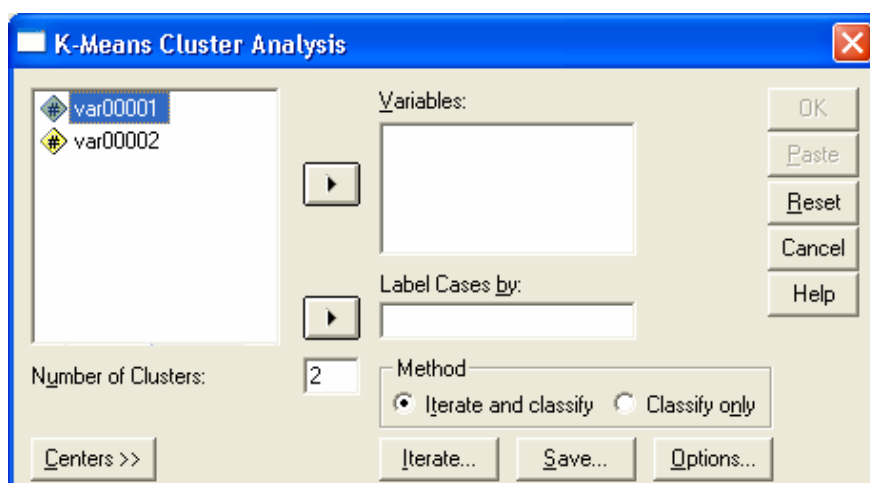


Рис. 2.27. Діалогове вікно кластерного аналізу

У цьому вікні зазначаємо змінну, значення якої необхідно поділити на кластери, метод класифікації, а також необхідність виведення окремих даних у файл або вікно результатів аналізу.

У випадку $b = 15$ отримано такі результати:

- кількість елементів першого кластера – 84, а другого – 116;
- центри кластерів 9,16 і 15,15, тобто похибки визначення центрів дорівнюють, відповідно, 8,4% й 1%;
- кількість помилково класифікованих значень першого кластера 27, другого кластера – 11;
- загальна імовірність помилкової класифікації – 19%.

У випадку $b = 20$ отримано такі результати:

- кількість елементів першого кластера – 102, а другого – 98;
- центри кластерів 10,20 і 20,45, тобто похибки визначення центрів дорівнюють, відповідно, 2% й 2,25%;
- кількість помилково класифікованих значень першого кластера – 2, другого кластера – 4;
- загальна імовірність помилкової класифікації – 3%.

Таким чином ми бачимо, що застосування цієї процедури дає прийнятні результати лише у випадку, коли кластери є достатньо добре відмежованими один від одного.

Після виокремлення кластерів для отримання більш повної інформації про їх параметри використаємо засоби побудови описової статистики. Для цього виберемо у головному меню *Analyze/Descriptive Statistics/Frequencies*. При цьому з'являється діалогове вікно побудови описової статистики (рис. 2.28).

У підменю *Statistics* (рис. 2.29) зазначимо, які саме параметри необхідно розрахувати. У підменю *Charts* (рис. 2.30) зазначимо, що необхідно побудувати гістограми розподілів й нанести на них графіки нормального розподілу.

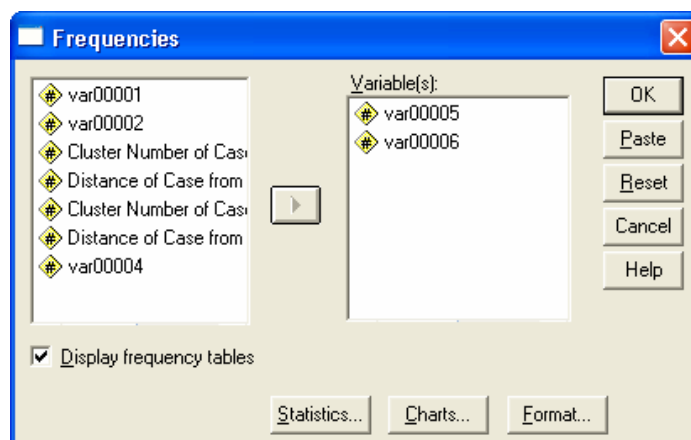


Рис. 2.28. Діалогове вікно побудови описової статистики

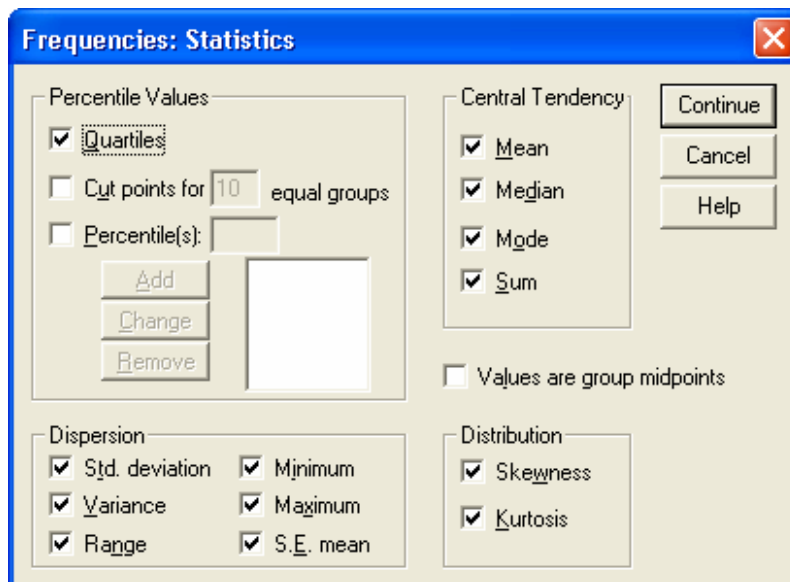


Рис. 2.29. Підменю для розрахунку параметрів описової статистики

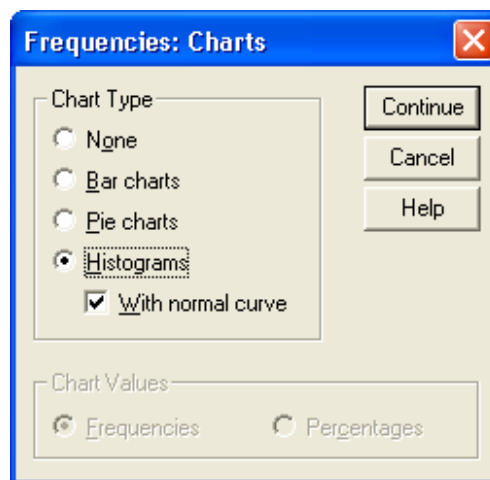


Рис. 2.30. Підменю Charts

На рис. 2.31 наведено вікно з результатами розрахунку описової статистики, а на рис. 2.32, 2.33 – гістограми розподілу частот для отриманих класів. Наведені дані свідчать, що математичні сподівання й стандартні відхилення є близькими до заданих параметрів, а розподіл значень у класах – до нормального.

На підставі отриманих результатів можна побудувати модель суміші розподілів у вигляді:

$$F(x) = 0,51N(10, 2; 2,86) + 0,49N(20, 45; 2,93), \quad (2.54)$$

де $N(\mu; s)$ – функція нормального розподілу з математичним сподіванням μ та стандартним відхиленням s .

Значення вагових коефіцієнтів отриманої формули взято пропорційними чисельності отриманих класів з урахуванням умови $\alpha_1 + \alpha_2 = 1$.

Statistics

		VAR00005	VAR00006
N	Valid	98	102
	Missing	102	98
Mean		20,4493	10,1982
Std. Error of Mean		,29559	,28354
Median		20,7644	10,8251
Mode		16,49	2,27 ^a
Std. Deviation		2,92618	2,86365
Variance		8,56251	8,20052
Skewness		,378	-,778
Std. Error of Skewness		,244	,239
Kurtosis		-,147	,381
Std. Error of Kurtosis		,483	,474
Range		13,74	12,70
Minimum		15,40	2,27
Maximum		29,14	14,97
Sum		2004,03	1040,21
Percentiles	25	18,0383	8,7034
	50	20,7644	10,8251
	75	22,1956	12,2346

a. Multiple modes exist. The smallest value is shown

Рис. 2.31. Результати розрахунку описової статистики

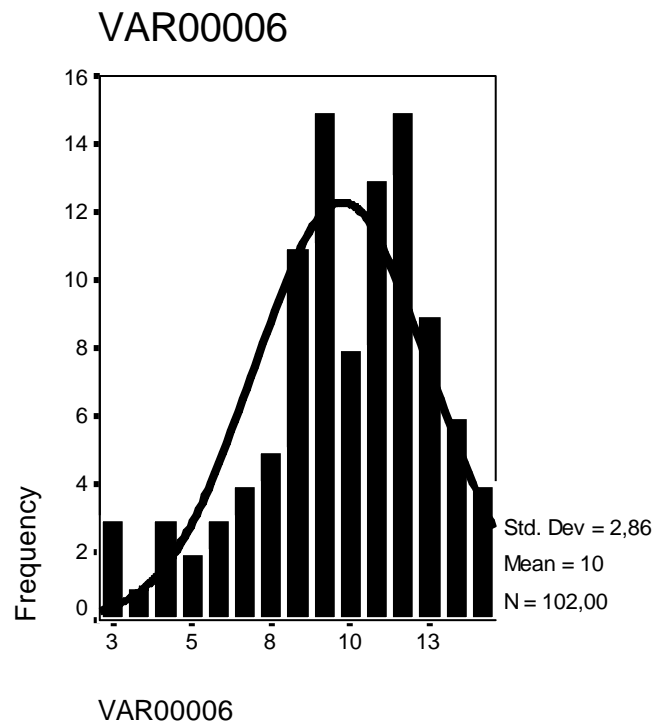


Рис. 2.32. Гістограма абсолютних частот для першого класу

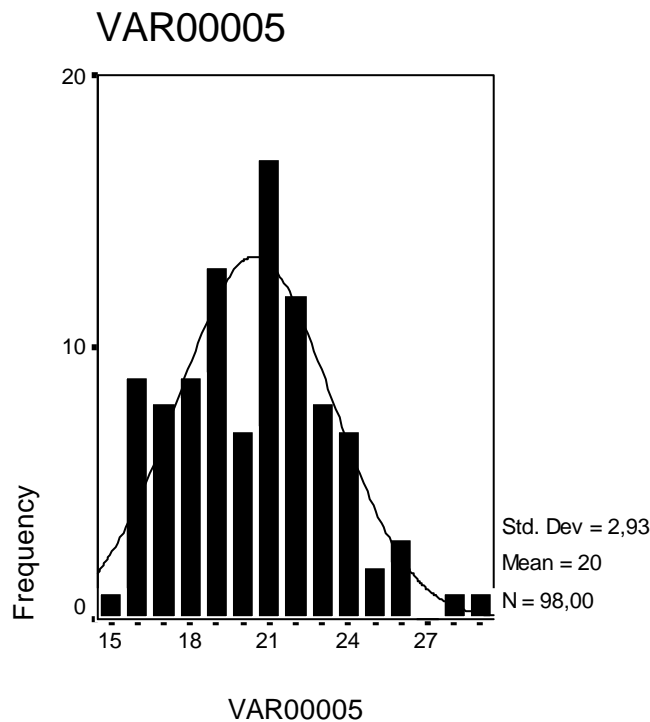


Рис. 2.33. Гістограма абсолютних частот для другого класу

Контрольні питання

1. З чого слід виходити при виборі методу перевірки статистичної гіпотези?
2. Які гіпотези називають нульовою та конкуруючою? Наведіть приклади. Скільки конкуруючих гіпотез можна сформулювати для однієї й тієї самої нульової гіпотези?
3. Які гіпотези називають простими та складними? Наведіть приклади.
4. Що називають статистичним критерієм? Які основні типи критеріїв використовують для перевірки гіпотез?
5. У чому полягає різниця між однобічними та двобічними критеріями?
6. Які величини називають рівнем значущості й довірчим рівнем? Для чого використовують ці параметри?
7. Які типи помилок розглядають при перевірці статистичних гіпотез? Як вони пов'язані між собою?
8. Що називають потужністю критерію? Для чого використовують цей параметр?
9. Як можна підвищити потужність критерію?
10. Якою є загальна методика перевірки статистичних гіпотез?
11. Які вибірки називають незалежними та спряженими? Наведіть приклади.
12. Які критерії називають параметричними? У яких випадках доцільно використовувати параметричні критерії?

13. Які гіпотези можна перевіряти за допомогою Z -критерію? Якими є умови правомірності його застосування?
14. Які гіпотези можна перевіряти за допомогою t -критерію Стьюдента? Якими є умови правомірності його застосування?
15. Які гіпотези можна перевіряти за допомогою критерію Уелча? Якими є умови правомірності його застосування?
16. Які гіпотези можна перевіряти за допомогою F -критерію Фішера? Якими є умови правомірності його застосування?
17. Які гіпотези можна перевіряти за допомогою критерію рандомізації компонент Фішера? Якими є умови правомірності його застосування?
18. Які гіпотези можна перевіряти за допомогою W -критерію Уїлкоксона? Якими є умови правомірності його застосування?
19. Які гіпотези можна перевіряти за допомогою U -критерію Манна – Уїтні? Якими є умови правомірності його застосування?
20. Які гіпотези можна перевіряти за допомогою T -критерію Уїлкоксона? Якими є умови правомірності його застосування?
21. Які гіпотези можна перевіряти за допомогою критерію χ^2 ? Якими є умови правомірності його застосування?
22. Які гіпотези можна перевіряти за допомогою критерію серій Вальда – Волфовиця? Якими є умови правомірності його застосування?
23. Які гіпотези можна перевіряти за допомогою критерію знаків? Якими є умови правомірності його застосування?
24. Які критерії можна використовувати для перевірки нормальності розподілу даних? Охарактеризуйте їх переваги й недоліки.
25. Яку гіпотезу можна перевіряти за допомогою ω^2 критерію Мізеса? Якими є умови правомірності його застосування?
26. Яку гіпотезу можна перевіряти за допомогою критерію Смірнова? Якими є умови правомірності його застосування?
27. Які гіпотези можна перевіряти за допомогою критерію χ^2 ? Якими є умови правомірності його застосування?
28. Яку гіпотезу можна перевіряти за допомогою W -критерію Шапіро – Уїлка? Якими є умови правомірності його застосування?