

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ”  
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ**

**Кафедра математичних методів захисту інформації**

**ЗВІТ**

**З ВИРОБНИЧОЇ ПРАКТИКИ**

Напрямок підготовки: 6.040301 «Прикладна математика»

Тема: «Розробка автоматизованого тестуючого комплексу,  
що враховує психологічні особливості студентів»

Виконав студент 4 курсу

групи ФІ-13

Кригін Валерій Михайлович

Науковий керівник:

Доктор фізико-математичних наук, професор

Дороговцев Андрій Анатолійович

---

*(підпис)*

**Київ 2015**

## ЗМІСТ

1 Вступ . . . . .	3
1.1 Обґрунтування та актуальність роботи . . . . .	3
1.2 Мета та завдання . . . . .	3
2 Основна частина . . . . .	4
2.1 Теоретичні відомості . . . . .	4
2.1.1 Метод головних компонент . . . . .	4
2.1.2 Гістограма . . . . .	7
2.1.3 Критерій узгодженості Пірсона $\chi^2$ . . . . .	8
2.1.4 Застосування критерія узгодженості Пірсона $\chi^2$ . . . . .	13
Перелік посилань . . . . .	15

## **1 ВСТУП**

### **1.1 Обґрунтування та актуальність роботи**

Існуючі на даний момент системи тестування недостатньо гнучкі: вони аналізують лише відповіді на запитання, відносячи їх до вірних або невірних, а на цій базі роблять кінцевий висновок щодо знань студента. Стрімкий розвиток комп'ютерної техніки й інформаційних технологій надає можливість визначати ритм складання тесту, а також індивідуальні особливості людини. Дані психологічних досліджень допоможуть правильно трактувати отримані значення, а добре вивчені та перевірені часом математичні методи надають великі можливості для систематизації та обробки результатів вимірювання.

### **1.2 Мета та завдання**

Завдання наступні:

- 1) Вивчити математичні методи та розділи психології, що дозволять розв'язати поставлену задачу, пояснити та обґрунтувати отримані результати
- 2) Ознайомитися з правилами побудови тестових завдань для найбільш ефективної та об'єктивної процедури оцінки знань студентів
- 3) Розробити програмний комплекс тестування й обробки результатів
- 4) Моделювання

За мету поставлено збільшення об'єктивності тестування, а також покращення якості навчання за допомогою порад студентам і викладачам практичних занять.

## 2 ОСНОВНА ЧАСТИНА

### 2.1 Теоретичні відомості

#### 2.1.1 Метод головних компонент

Метод головних компонент (Principal component analysis) — метод, що дозволяє зменшити розмірність досліджуваної вибірки з мінімальними втратами інформації.

Маємо  $m$  об'єктів, з яких треба зняти по  $n$  певних властивостей. На вході в нас є виборки  $\vec{X}_k$ , кожна з яких відповідає сукупності властивостей  $k$ -го об'єкту

$$\vec{X}_k = \begin{bmatrix} x_k^1 \\ x_k^2 \\ \vdots \\ x_k^n \end{bmatrix}, \quad k = \overline{1, m}$$

Згрупуємо всі вимірювання в одну матрицю  $X$

$$X = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_m^1 \\ x_1^2 & x_2^2 & \dots & x_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & \dots & x_m^n \end{bmatrix}$$

Спочатку нам знадобиться знайти вибіркові середні значення для кожної властивості

$$a_i = \frac{1}{m} \cdot \sum_{k=1}^m x_k^i, \quad i = \overline{1, n}$$

Маємо вектор вибірових середніх значень

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Центруємо отримані дані, що містяться в матриці  $X$ , віднявши від кожного стовбця вектор вибірових середніх  $\vec{a}$

$$\tilde{X} = \begin{bmatrix} \tilde{x}_1^1 & \tilde{x}_2^1 & \dots & \tilde{x}_m^1 \\ \tilde{x}_1^2 & \tilde{x}_2^2 & \dots & \tilde{x}_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_1^n & \tilde{x}_2^n & \dots & \tilde{x}_m^n \end{bmatrix} = \begin{bmatrix} x_1^1 - a_1 & x_2^1 - a_1 & \dots & x_m^1 - a_1 \\ x_1^2 - a_2 & x_2^2 - a_2 & \dots & x_m^2 - a_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n - a_n & x_2^n - a_n & \dots & x_m^n - a_n \end{bmatrix}$$

Обчислюємо вибірову коваріаційну матрицю властивостей. Вибіркову коваріацію  $i$  та  $j$  властивості рахуємо за формулою

$$\sigma_i^j = \frac{1}{m} \cdot \sum_{k=1}^m \tilde{x}_k^i \cdot \tilde{x}_k^j = \frac{1}{m} \cdot \sum_{k=1}^m \left[ (x_k^i - a_i) \cdot (x_k^j - a_j) \right], \quad i, j = \overline{1, n}$$

Маємо вибірову коваріаційну матрицю

$$K = \begin{bmatrix} \sigma_1^1 & \sigma_1^2 & \dots & \sigma_1^n \\ \sigma_2^1 & \sigma_2^2 & \dots & \sigma_2^n \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_n^1 & \sigma_n^2 & \dots & \sigma_n^n \end{bmatrix}$$

Щоб отримувати лише потрібну інформацію, ми хочемо знайти таке ортогональне лінійне перетворення  $L$  вхідної матриці  $\tilde{X}$ , щоб отримати матрицю

$Y = L \cdot \tilde{X}$ , яка має діагональну вибірку коваріаційну матрицю  $K'$  з незростаючими зверху вниз значеннями. Діагональна вибірка коваріаційна матриця гарантує той факт, що отримані значення  $Y$  будуть некорельованими. Рангування значень діагональних елементів матриці  $K'$  за величиною дасть більш наглядне представлення про будову досліджуваних об'єктів, адже діагональні елементи — вибірккові дисперсії; а чим більше дисперсія, тим більше відповідна властивість змінюється від об'єкту до об'єкту і тим більше корисної інформації вона нам надає.

Вибіркова коваріаційна матриця  $K'$  для  $Y = L \cdot \tilde{X}$  має вигляд

$$K' = L \cdot K \cdot L^* = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

З лінійної алгебри відомо, що матриця  $L$  складається з координат власних векторів матриці  $K$ , а елементи  $\lambda_k$  — її власні числа, які існують і є невід'ємними через невід'ємну означеність матриці  $K$ . Вважаємо що числа  $\lambda_1, \dots, \lambda_n$  впорядковані від більшого до меншого для зручності подальших дій. Позначимо власний вектор матриці  $K$ , що відповідає власному числу  $\lambda_k$ , як  $\vec{l}_k$ . Тоді

$$\vec{l}_k = [l_k^1, l_k^2, \dots, l_k^n], \quad k = \overline{1, n}$$

Матриця  $L$  має вигляд

$$L = \begin{bmatrix} l_1^1 & l_1^2 & \dots & l_1^n \\ l_2^1 & l_2^2 & \dots & l_2^n \\ \vdots & \vdots & \ddots & \vdots \\ l_n^1 & l_n^2 & \dots & l_n^n \end{bmatrix}$$

Треба зменшити розмірність простору досліджуваних параметрів системи з  $n$  до  $p < n$ , але при цьому втратити якомога менше відомостей про досліджувані об'єкти. Введемо міру інформації, що залишається при зменшенні кількості компонент, що розглядаються

$$I = \frac{\lambda_1 + \dots + \lambda_p}{\lambda_1 + \dots + \lambda_n}$$

Будемо вважати, що діємо продуктивно, тому починаємо обирати з перших компонент, адже саме вони є найбільш інформативними. Також бачимо, що інформативність змінюється в межах від 0 (нічого не дізнаємось) до 1 (зберегли усю інформацію).

Надалі буде розглядатися матриця головних компонент  $Y$

$$Y = \begin{bmatrix} y_1^1 & y_2^1 & \dots & y_m^1 \\ y_1^2 & y_2^2 & \dots & y_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ y_1^p & y_2^p & \dots & y_m^p \end{bmatrix}$$

### 2.1.2 Гістограма

Для подальшого аналізу потрібно здобути щільність розподілу головних компонент. Оскільки маємо справу з вибіркою і вибірковими характеристиками, потрібно побудувати гістограму, адже це і є вибіркова характеристика, що відповідає щільності.

Побудуємо  $j$ -й стовбець гістограми для виборки з  $k$ -ї строки матриці  $Y$

$$h_j^k = \frac{1}{m} \cdot \sum_{i=1}^m \mathbb{1}(y_i^k \in I_j^k), \quad j = \overline{1, N}, \quad k = \overline{1, p}$$

де  $I^k$  — набір напівінтервалів, що розбиває відрізок  $\left[ \min_{i=\overline{1, m}} y_i^k; \max_{i=\overline{1, m}} y_i^k \right]$  на  $N$

рівних частин. Для вибору  $N$  можна скористатися досить відомою формулою Стюрджеса (Sturges' formula) [1]

$$N = \lfloor \log_2 m \rfloor + 1$$

Маємо матрицю гістограм

$$H = \begin{bmatrix} h_1^1 & h_2^1 & \dots & h_N^1 \\ h_1^2 & h_2^2 & \dots & h_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ h_1^p & h_2^p & \dots & h_N^p \end{bmatrix}$$

і напівінтервалів, що відповідають кожному стовбчику кожної гістограми

$$I = \begin{bmatrix} I_1^1 & I_2^1 & \dots & I_N^1 \\ I_1^2 & I_2^2 & \dots & I_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ I_1^p & I_2^p & \dots & I_N^p \end{bmatrix}$$

### 2.1.3 Критерій узгодженості Пірсона $\chi^2$

Гістограма може використовуватися не тільки для графічної інтерпретації отриманих даних, але й для віднесення вибірки до якогось відомого розподілу. Відповідь на питання “Чи дійсно вибірка  $y_1^k, \dots, y_p^k$  має розподіл  $F^k$ ?” може надати критерій узгодженості Пірсона.

Розглянемо вектор

$$\eta^k = \left[ \frac{\nu_1^k - m \cdot \rho_1^k}{\sqrt{m \cdot \rho_1^k}}, \dots, \frac{\nu_N^k - m \cdot \rho_N^k}{\sqrt{m \cdot \rho_N^k}} \right]$$



Знайдемо його характеристичну функцію

$$\varphi_{\eta^k}(\lambda) = \mathbb{M} e^{i \cdot \lambda \cdot \eta^k}, \quad \lambda \in \mathbb{R}^N$$

Для зручності перепозначимо індикатор

$$\mathfrak{I}_{i,j}^k = \mathbb{1}\{y_i^k \in I_j^k\}$$

Подивимось, чому дорівнює скалярний добуток в експоненті

$$\begin{aligned} (\lambda, \eta^k) &= \sum_{j=1}^N \lambda_j \cdot \frac{\nu_j^k - m \cdot \rho_j^k}{\sqrt{m \cdot \rho_j^k}} = \sum_{j=1}^N \frac{\lambda_j}{\sqrt{m \cdot \rho_j^k}} \cdot \sum_{i=1}^m (\mathfrak{I}_{i,j}^k - \rho_j^k) = \\ &= \sum_{j=1}^N \sum_{i=1}^m \frac{\lambda_j}{\sqrt{m \cdot \rho_j^k}} \cdot (\mathfrak{I}_{i,j}^k - \rho_j^k) = \sum_{i=1}^m \sum_{j=1}^N \lambda_j \cdot \frac{\mathfrak{I}_{i,j}^k - \rho_j^k}{\sqrt{m \cdot \rho_j^k}} \end{aligned}$$

Бачимо суму  $m$  незалежних однаково розподілених випадкових величин. Введемо позначення

$$\mathfrak{I}_j^k = \mathbb{1}\{y_1^k \in I_j^k\}$$

А також позначимо новий випадковий вектор

$$\zeta^k = \left[ \frac{\mathfrak{I}_1^k - \rho_1^k}{\sqrt{m \cdot \rho_1^k}}, \dots, \frac{\mathfrak{I}_N^k - \rho_N^k}{\sqrt{m \cdot \rho_N^k}} \right]$$

Тоді скалярний добуток прийме вигляд

$$(\lambda, \eta^k) = \sum_{i=1}^m \sum_{j=1}^N \lambda_j \cdot \zeta_j^k = \sum_{i=1}^m \lambda \cdot \zeta^k = m \cdot \lambda \cdot \zeta^k = m \cdot (\lambda, \zeta^k)$$

За рахунок незалежності випадкових величин  $\zeta_j^k$  маємо

$$\varphi_{\eta^k}(\lambda) = \mathbb{M} e^{i \cdot \lambda \cdot \eta^k} = \mathbb{M} e^{m \cdot i \cdot \lambda \cdot \zeta^k} = \left( \mathbb{M} e^{i \cdot \lambda \cdot \zeta^k} \right)^m \quad (2.1)$$

Розглянемо характеристичну функцію випадкового вектора  $\zeta^k$

$$\varphi_{\zeta^k}(\lambda) = \mathbb{M} \left[ \exp \left\{ i \cdot \sum_{j=1}^N \lambda_j \cdot \zeta_j^k \right\} \right] \quad (2.2)$$

Легко побачити, що

$$\begin{aligned} (\lambda, \zeta^k) &= \sum_{j=1}^N \lambda_j \cdot \zeta_j^k = \sum_{j=1}^N \lambda_j \cdot \frac{\mathfrak{I}_j^k - \rho_j^k}{\sqrt{m \cdot \rho_j^k}} = \sum_{j=1}^N \left( \frac{\lambda_j}{\sqrt{m \cdot \rho_j^k}} \cdot \mathfrak{I}_j^k - \frac{\sqrt{\rho_j^k} \cdot \lambda_j}{\sqrt{m}} \right) = \\ &= \sum_{j=1}^N \mathfrak{I}_j^k \cdot \left( \frac{\lambda_j}{\sqrt{m \cdot \rho_j^k}} - \sum_{l=1}^N \frac{\sqrt{\rho_l^k} \cdot \lambda_l}{\sqrt{m}} \right) \end{aligned}$$

Тобто характеристична функція (2.2) приймає вигляд

$$\varphi_{\zeta^k}(\lambda) = \mathbb{M} \left[ \sum_{j=1}^N \mathfrak{I}_j^k \cdot \exp \left\{ \frac{i}{\sqrt{m}} \left( \frac{\lambda_j}{\sqrt{\rho_j^k}} - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right) \right\} \right]$$

Перепозначимо вираз в круглих дужках

$$\mathfrak{z}^k = \frac{\lambda_j}{\sqrt{\rho_j^k}} - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l$$

Математичне очікування індикатора — ймовірність події, яку він перевіряє. Отже

$$\varphi_{\zeta^k}(\lambda) = \sum_{j=1}^N \rho_j^k \cdot \exp \left\{ \frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} \right\}$$

Якщо розмір вибірки  $m$  буде зростати, то характеристична функція  $\eta^k$  (2.1) буде поводитись наступним чином

$$\begin{aligned} \lim_{m \rightarrow \infty} \varphi_{\eta^k}(\lambda) &= \lim_{m \rightarrow \infty} \left( 1 + \sum_{k=1}^N p_k \cdot \left[ \exp \left\{ \frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} \right\} - 1 \right] \cdot \frac{m}{m} \right)^m = \\ &= \lim_{m \rightarrow \infty} \exp \left\{ m \cdot \sum_{k=1}^N p_k \cdot \left[ \exp \left\{ \frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} \right\} - 1 \right] \right\} \end{aligned}$$

Так як

$$e^\alpha - 1 \approx \alpha + \frac{\alpha^2}{2}, \quad \alpha \ll 1$$

Перший ступінь наближення зникає

$$\begin{aligned} \sum_{j=1}^N \rho_j^k \cdot \frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} &= \sum_{j=1}^N \rho_j^k \cdot \frac{i}{\sqrt{m}} \cdot \left( \frac{\lambda_j}{\sqrt{\rho_j^k}} - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right) = \\ &= \frac{i}{\sqrt{m}} \cdot \left( \sum_{j=1}^N \sqrt{\rho_j^k} \cdot \lambda_j - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right) = 0 \end{aligned}$$

Отже, маємо справу лише з другим

$$\begin{aligned} \sum_{j=1}^N \rho_j^k \cdot \left( \frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} \right)^2 &= - \sum_{j=1}^N \frac{\rho_j^k}{m} \cdot \left( \frac{\lambda_j}{\sqrt{\rho_j^k}} - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right)^2 = \\ &= -\frac{1}{m} \cdot \left[ \sum_{j=1}^N \lambda_j - \left( \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right)^2 \right] \end{aligned}$$

Перепишемо характеристичну функцію вектора  $\eta^k$

$$\begin{aligned} \lim_{m \rightarrow \infty} \varphi_{\eta^k}(\lambda) &= \lim_{m \rightarrow \infty} \exp \left\{ -\frac{m}{m \cdot 2} \cdot \left[ \sum_{j=1}^N \lambda_j - \left( \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right)^2 \right] \right\} = \\ &= \exp \left\{ -\frac{1}{2} \cdot \left[ \sum_{j=1}^N \lambda_j - \left( \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right)^2 \right] \right\} = e^{-\frac{1}{2} \cdot (A^k \lambda, \lambda)} \end{aligned}$$

Матриця  $A^k$  побудована наступним чином

$$A^k = \left\| \delta_{ij} - \sqrt{\rho_i^k} \cdot \sqrt{\rho_j^k} \right\|_{i=1}^n$$

Симетричність матриці очевидна, тому треба довести її невід'ємну визначеність, щоб стверджувати, що вона є коваріаційною. Для цього візьмемо вектор

$$e^k = \left[ \sqrt{\rho_1^k}, \dots, \sqrt{\rho_N^k} \right], \quad \|e^k\| = 1$$

Тоді бачимо, що

$$(A^k \lambda, \lambda) = \|\lambda\|^2 - (\lambda, e^k)^2 \quad (2.3)$$

З нерівності Коші маємо

$$\|(\lambda, e^k)\| \leq \|\lambda\| \cdot \|e^k\| = \|\lambda\|$$

Тобто матриця є дійсно невід’ємно визначеною і вектор  $\eta^k$  розподілений за нормальним законом з нульовим середнім і коваріаційною матрицею  $A^k$ .

Для подальших розрахунків розглянемо стандартний гаусівський вектор як суму випадкових нормально розподілених випадкових величин в стандартному базисі  $\mathbb{R}^N$ , який позначимо  $[e_1, \dots, e_N]$

$$\xi = \sum_{j=1}^N \xi_j \cdot e_j \sim N(\vec{0}, I)$$

Згадуємо, що ортогональні перетворення  $U$  зберігають відстані, а також справедливо наступне

$$U\xi \sim N(0, UIU^{-1}) \sim N(\vec{0}, I)$$

Також ортонормований базис залишається ортонормованим базисом після ортогонального перетворення  $U$ . Оберемо такий оператор  $U$ , щоб набір  $[e_1, \dots, e_N]$  під його дією перетворився на  $[f_1, \dots, f_N]$ , де

$$f_1 = e^k = \left[ \sqrt{\rho_1^k}, \dots, \sqrt{\rho_N^k} \right]$$

Тоді маємо вектор

$$U\xi = \hat{\xi} = \sum_{j=1}^N \hat{\xi}_j \cdot f_j \sim N(\vec{0}, I)$$

Подивимось, яке розподілення має наступний вектор

$$\Upsilon = \sum_{j=2}^N \hat{\xi}_j \cdot f_j = \hat{\xi} - \xi_1 \cdot e^k$$

Для цього розглянемо квадратичну форму

$$M[(\Upsilon, \lambda) \cdot (\Upsilon, \lambda)] = \sum_{j=2}^N (\lambda, f_j)^2 = \sum_{j=1}^N (\lambda, f_j)^2 - (\lambda, f_1)^2 = \|\lambda\|^2 - (\lambda, e^k)^2 = (A^k \lambda, \lambda)$$

З рівності (2.3) бачимо, що випадкові вектори  $\eta^k$  і  $\Upsilon$  мають однаковий розподіл. Отже, розподіли їх норм теж співпадають. Оскільки сума  $N - 1$  квадратів незалежних стандартних гаусових випадкових величин має розподіл Пірсона з  $N - 1$  ступенями вільності

$$\|\Upsilon\|^2 = \sum_{j=2}^N \xi_j^2 \sim \chi_{N-1}^2$$

Маємо

$$\|\eta^k\| = \sum_{j=1}^N \frac{(\nu_j^k - m \cdot \rho_j^k)^2}{m \cdot \rho_j^k} = m \cdot \sum_{j=1}^N \frac{(h_j^k - \rho_j^k)^2}{\rho_j^k} \sim \chi_{N-1}^2 \quad (2.4)$$

#### 2.1.4 Застосування критерія узгодженості Пірсона $\chi^2$

Розглянемо випадкову величину

$$R^k = m \cdot \sum_{j=1}^N \frac{(h_j^k - \rho_j^k)^2}{\rho_j^k} \quad (2.5)$$

З таблиці для функції розподілу  $\chi_{N-1}^2$  обираємо рівень значущості  $\alpha$  і шукаємо відповідне до кількості ступенів вільності  $r_\alpha$ . Рівень значущості — ймовірність помилки першого роду, тобто ймовірність того, що буде відкинута вірна гіпотеза

$$\mathbb{P}(\chi_{N-1}^2 \geq r_\alpha) = \alpha$$

Якщо  $R^k \leq r_\alpha$ , то гіпотеза про те, що вибірка  $Y^k$  дійсно має розподіл  $F^k$ , не відхиляється.

Розглянемо той випадок, коли ймовірність  $\rho_i^k$  відгадана невірно. Повернемося до формули (2.5)

$$R^k = \sum_{j=1}^N \frac{(\nu_j^k - m \cdot \rho_j^k)^2}{m \cdot \rho_j^k}$$

Всі члени суми є невід'ємними. Якщо хоча б один елемент буде завеликим, то

великою буде вся сума. Маємо випадкову величину  $\eta$

$$\eta = \nu_i^k - m \cdot \rho_i^k = \sum_{j=1}^m (\xi_j - \rho_i^k), \quad \mathbb{1}(y_j^k \in I_i^k) = \xi_j$$

Якщо  $\rho_i^k$  вгадано невірно, то воно не дорівнює математичному очікуванню індикатора. Додамо та віднімемо справжнє математичне очікування

$$\eta = \sum_{j=1}^m (\xi_j - M \xi_1 + M \xi_1 - \rho_i^k) = \sum_{j=1}^m (\xi_j - M \xi_1) + \sum_{j=1}^m (M \xi_1 - \rho_i^k)$$

Останній доданок є просто різницею, помноженою на  $m$

$$\eta = \sum_{j=1}^m (\xi_j - M \xi_1) + m \cdot (M \xi_1 - \rho_i^k)$$

Поділимо на  $\sqrt{m}$ , щоб скористатися центральною граничною теоремою

$$\frac{\eta}{\sqrt{m}} = \frac{1}{\sqrt{m}} \cdot \sum_{j=1}^m (\xi_j - M \xi_1) + \frac{1}{\sqrt{m}} \cdot m \cdot (M \xi_1 - \rho_i^k)$$

Перший доданок має розподіл  $N(0, \sigma^2)$ , де  $\sigma^2$  — дисперсія випадкової величини  $\xi_1$  для достатньо великих  $m$ . Отже, вся сума зростає пропорційно  $\sqrt{m}$

$$\frac{\eta}{\sqrt{m}} = \frac{1}{\sqrt{m}} \cdot \sum_{j=1}^m (\xi_j - M \xi_1) + \sqrt{m} \cdot (M \xi_1 - \rho_i^k) \sim \sqrt{m} \cdot (M \xi_1 - \rho_i^k)$$

Тобто якщо гіпотеза невірна, то  $R^k$  буде зростати пропорційно до величини  $\sqrt{m}$ , що є достатньо великою швидкістю.

Чим більше рівень значущості, тим менше значення  $r_\alpha$ , а отже і проміжок, в який дозволяється потрапити значенню  $R^k$ . В цьому випадку зростає ймовірність відхилити вірну гіпотезу щодо розподілу, але при цьому є більше впевненості в правильності результату. Зазвичай  $\alpha$  обирають рівним 0.1, 0.05, 0.01.

**ПЕРЕЛІК ПОСИЛАНЬ**

1. *Sturges, Herbert A.* The Choice of a Class Interval / Herbert A. Sturges // *J-AM-STAT-ASSOC.* — 1926. — March. — Vol. 21, no. 153. — Pp. 65–66.