

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

Кафедра математичних методів захисту інформації

«До захисту допущено»

Завідувач кафедри

(підпис) (ініціали, прізвище)
“ ” 2015 р.

Дипломна робота

освітньо-кваліфікаційного рівня «бакалавр»

з напрямку підготовки 6.040301 «Прикладна математика»

на тему: «Розробка автоматизованого тестуючого комплексу,

що враховує психологічні особливості студентів»

Виконав: студент 4 курсу групи ФІ-13 Кригін Валерій Михайлович

Керівник д-р ф.-м. наук, професор Дороговцев Андрій Анатолійович

Рецензент к. т. н., доцент Головенкін Володимир Павлович

(підпис)

(підпис)

(підпис)

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент _____

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 χ^2	8
2.1.4 Типи вищої нервової діяльності	14
2.1.5 Теппінг-тест (Tapping rate)	15
2.1.6 Адаптованість	18
2.1.7 Відпочинок як процес Пуассона	19
2.1.8 Гідравліка	21
Перелік посилань	26

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. [1]

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 наочне уявлення про будову досліджуваних об'єктів, адже діагональні елементи — вибірккові дисперсії. 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 λ_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 λ_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) [2]

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

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 χ^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_m^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

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

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

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

За рахунок незалежності випадкових величин ζ_j^k маємо

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

Розглянемо характеристичну функцію випадкового вектора ζ^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

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

Перепознаємо 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 η^k (2.1) буде поводитись наступним чином

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

Для $\exp \left\{ \frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} \right\}$ використаємо співвідношення

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

Маємо

$$\begin{aligned} \sum_{j=1}^N \rho_j^k \cdot \frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} &= \sum_{j=1}^N \rho_j^k \cdot \frac{i}{\sqrt{m}} \cdot \left(\frac{\lambda_j}{\sqrt{\rho_j^k}} - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right) = \\ &= \frac{i}{\sqrt{m}} \cdot \left(\sum_{j=1}^N \sqrt{\rho_j^k} \cdot \lambda_j - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right) = 0, \\ \sum_{j=1}^N \rho_j^k \cdot \left(\frac{i \cdot \mathfrak{z}^k}{\sqrt{m}} \right)^2 &= - \sum_{j=1}^N \frac{\rho_j^k}{m} \cdot \left(\frac{\lambda_j}{\sqrt{\rho_j^k}} - \sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right)^2 = \\ &= -\frac{1}{m} \cdot \left[\sum_{j=1}^N \lambda_j - \left(\sum_{l=1}^N \sqrt{\rho_l^k} \cdot \lambda_l \right)^2 \right] \end{aligned}$$

Тому

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

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 η^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, який розподіл має наступний вектор

$$\Upsilon = \sum_{j=2}^N \hat{\xi}_j \cdot f_j = \hat{\xi} - \hat{\xi}_1 \cdot e^k$$

Для цього розглянемо квадратичну форму

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

З рівності (2.3) бачимо, що випадкові вектори η^k та Υ 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

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

Останнє співвідношення дає змогу перевіряти належність виборки y_1^k, \dots, y_m^k до розподілу F^k . Перевірка виглядає наступним чином.

Розглянемо випадкову величину

$$R^k = m \cdot \sum_{j=1}^N \frac{(h_j^k - \rho_j^k)^2}{\rho_j^k} \quad (2.4)$$

Обираємо рівень значущості α для функції розподілу χ_{N-1}^2 і шукаємо відповідне до кількості ступенів вільності r_α . 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 , приймається.

Розглянемо той випадок, коли ймовірність ρ_i^k відгадана невірно. Повернемося до формули (2.4)

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

Всі члени суми є невід'ємними. Якщо хоча б один елемент буде завеликим, то великою буде вся сума. Маємо випадкову величину η

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

Якщо ρ_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 асимптотично має розподіл $N(0, \sigma^2)$, де σ^2 — дисперсія випадкової величини ξ_1 для достатньо великих m . Отже, вся сума зростає пропорційно до \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 зараз R^k буде зростати пропорційно до величини m , і буде великим у порівнянні з r_α , що призведе до відхилення невірної гіпотези.

2.1.4 Типи вищої нервової діяльності

Для визначення того, які показники вимірювати і яким чином, скористуємось відомою класифікацією типів вищої нервової діяльності.

Згідно з Павловим[3] типи вищої нервової діяльності характеризуються трьома показниками: сила нервової системи (сильна або слабка), врівноваженість (врівноважена або неврівноважена) та рухливість (рухлива або інертна). Павлов розглядає 4 комбінації цих показників з 8 можливих:

- 1) Слабка
- 2) Сильна та неврівноважена
- 3) Сильна, врівноважена та інертна
- 4) Сильна, врівноважена та рухлива

Далі ці класи (комбінації) будуть називатися відповідно слабкий, неврівноважений, інертний та рухливий.

2.1.5 Теппінг-тест (Tapping rate)

Існують відомі залежності між типом вищої нервової діяльності та зміною максимального темпу рухів кистю руки з часом. Протягом 30 секунд людина намагається притримуватися максимально можливого для себе темпу. Показники темпу фіксуються через кожні 5 секунд, а далі по 6 отриманим точкам будується крива темпа руху. [4]

Для тесту можна використовувати ручку (олівець) і папір, або телеграф. Сучасні технології дозволяють проводити тест за допомогою клавіатури комп'ютера або екрану планшета.

З олівцем і папіром тест поводитьься наступним чином:

- 1) На папері креслиться 6 квадратів
- 2) Людина починає ставити якомога більше точок в першому квадраті впродовж перших 5 секунд
- 3) Коли проходить 5 секунд, потрібно перейти до наступного квадрату і ставити точки там
- 4) Процедура повторюється до тих пір, доки не пройде 30 секунд — в кінці буде заповнено всі 6 квадратів

Далі підраховується кількість точок в кожному квадраті та малюється ламана, де горизонтальна вісь відповідає номеру часового проміжку (номеру квадрата), а вертикальна відповідає кількості точок в квадраті.

Трактуються отримані дані наступним чином:

- 1) Спадна ламана відповідає слабкому типу (рис. 2.1д). Вона спадає після перших 5 секунд тесту і не повертається до початкового рівня
- 2) Ламана, що спочатку зростає, а після 10-15 секунд спадає нижче початкового рівня (проміжна між рівною та опуклою) відповідає невірноваженому типу (рис. 2.1б).
- 3) Ввігнута ламана відповідає інертному типу (рис. 2.1г). Вона спочатку спадає, а на 25-30 секундах може зрости до початкового темпу
- 4) Опукла ламана відповідає рухливому типу (рис. 2.1а). Це така ламана, що зростає в перші 10-15 секунд тесту, а після 25-30 секунд повертається або падає нижче початкового рівня
- 5) Також темп може залишатися приблизно на одному рівні протягом усього тесту, що є оптимальним для складання іспитів (рис. 2.1в).

Оскільки цей тест заснований на вимірюванні витривалості нервової системи людини за умови максимального навантаження та перевіряє темп реагування (натиснення) на подразнювачі (внутрішній подразнювач — команда собі “треба тиснути”), було вирішено використовувати відомі вигляди кривих (рис. 2.1) при моделюванні результатів виконання завдань однакової складності.

Потрібно зауважити, що швидкість розв’язування задач може змінюватися з досвідом. Тобто, якщо студент зі слабкою нервовою системою буде тренуватися виконувати завдання, то його показники з часом перейдуть на якісно новий рівень. Щодо студентів з сильною нервовою системою: швидкість не завжди означає якість виконання завдань.

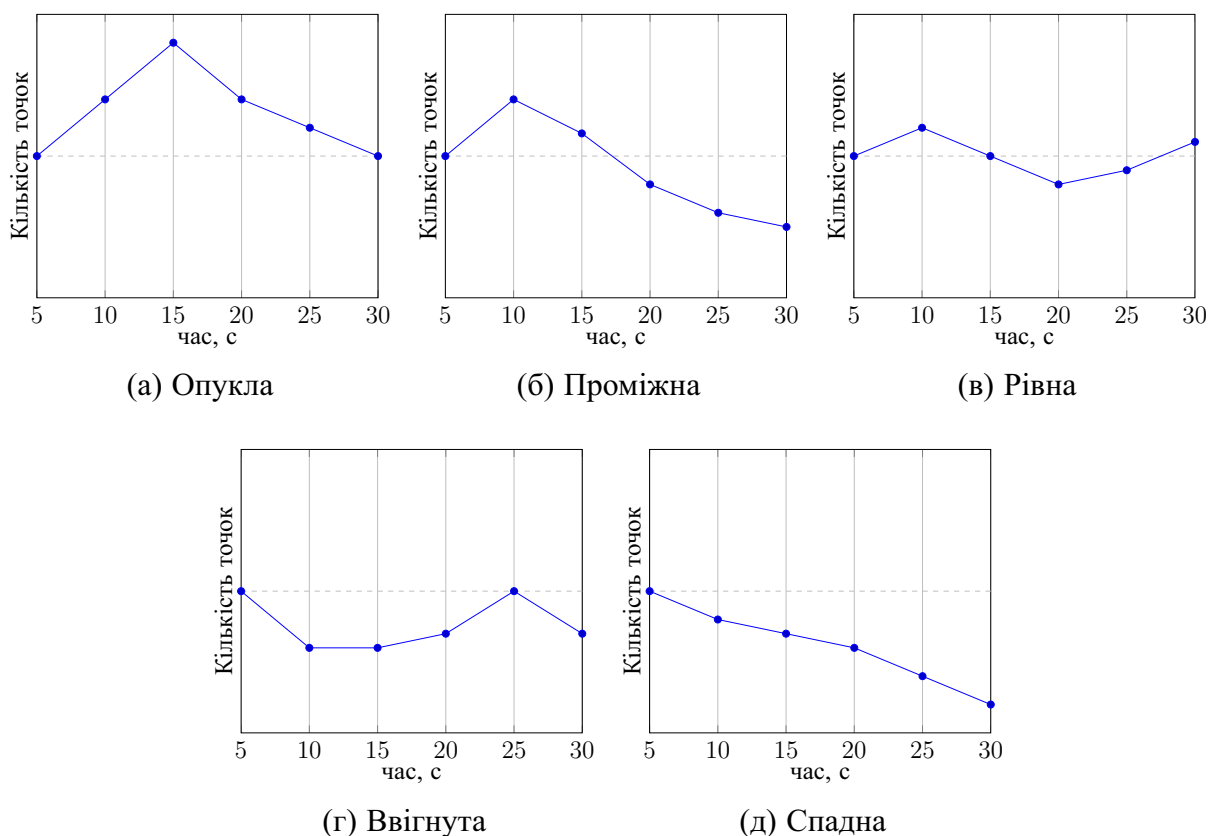


Рисунок 2.1 — Загальний вигляд залежностей кількості поставлених точок від часу. Пунктирна лінія — кількість точок в перші 5 секунд

Мета системи, що буде створено, — виявляти в автоматичному режимі слабкі місця студентів при виконанні завдань, а потім надавати поради викладачам практичних занять. Наприклад, тим, що надто швидко втомлюються, потрібно розв'язувати якомога більше базових завдань, що не є складними, але розв'язок яких повинен бути на рівні рефлексів. Студентам, які поспішають, буде наказано ретельно коментувати у письмовій формі хід своїх думок, щоб вгамуватися та підвищити свою уважність.

При використанні теплінг-тесту в якості сировини виникає потреба у диференціюванні студентів за знанням галузей, до яких відносяться задачі. Тобто, якщо завдання однакової складності, то вони можуть перевіряти знання лише в тому випадку, коли відносяться до різних тем, які студент може знати або ні. В обох випадках студенту може знадобитися час на те, щоб згенерувати відповідь

або принаймні спробувати. З’являється складність: в межах одного завдання у студента виникає кілька підзадач, кількість яких залежить від методу розв’язання, що він обере, і неможливо наперед вирахувати всі можливі способи вирішення тієї чи іншої проблеми.

2.1.6 Адаптованість

Якщо розглянути ламані на рис. 2.1 і поєднати їх, то отримаємо щось схоже на рис. 2.2, де на осі абсцис цифрами 1, 2, 3 і 4 позначено початок відповідно ввігнутої, опуклої, проміжної та спадної кривої.

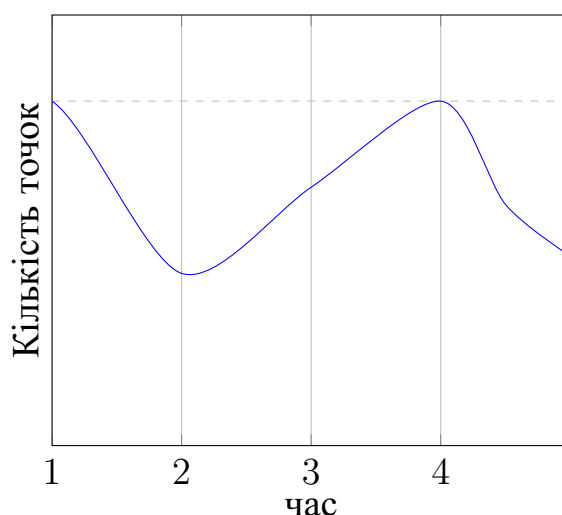


Рисунок 2.2 — Загальна картина активності

З рисунку видно, що в людей з сильною нервовою системою спостерігається ефект адаптованості — так зване “друге дихання”, чого немає у власників слабкої. Також бачимо, що крива прямого типу може бути отримана, якщо зменшити впадину між двома горбами на графіку — примусити швидше працювати механізми, що збуджують адаптованість.

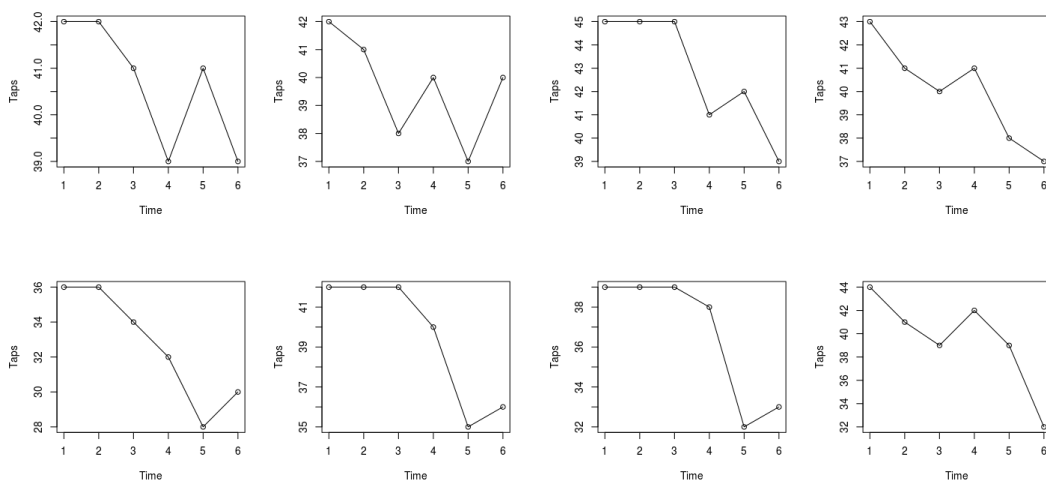
Щоб почати моделювання результатів тестів на основі відомих ламаних результатів теппінг-тесту, потрібно змоделювати результати самого теппінг-тесту.

Форма графіку, що зображено на рис. 2.2, спочатку нагадує синусоїду, але як його продовжити, якщо скористатися тим припущенням, що далі крива не буде зростати?

2.1.7 Відпочинок як процес Пуассона

Була спроба моделювати поведінку студента простим методом: в якості параметра брати його початковий ритм, а потім віднімати від нього реалізації процесу Пуассона з інтенсивністю $c \cdot t$ у відповідні проміжки часу. Тобто, було припущення, що в студента є певний ритм роботи, а його змушує сповільнюватися потреба у відпочинку, яка зростає з часом. Ідея розглядати кількість розв'язаних задач як процес Пуассона виявилася поганою як математично, так і ідеологічно: був дуже великий розкид значень, та взагалі хотілось би розглядати вирішення завдань як результат цілеспрямованої роботи студента, а не випадковий процес.

Для $c = \frac{1}{5}$ було отримано як досить цікаві для використання результати (рис. 2.3а), так і не дуже (рис. 2.3б). Моделювання проводилося за допомогою пакету NHPoisson мови програмування R.



(а) Задовільні результати

(б) Незадовільні результати

Рисунок 2.3 — Моделювання результатів теппінг-тесту

З рис. 2.3 видно, що серед гідних для подальшого аналізу ламаних трапляються ті, що мають занадто великі перепади. Також за допомогою такого підходу не вийде змодельювати різноманітні вигляди майже рівної кривої, адже мінімальне значення кількості “відпочинків” є нульовим. Якщо до цієї моделі додати шум (наприклад, замість гаусового використовувати незалежні випадкові, що розподілені за законом Бернуллі) для вирішення проблеми від’ємних значень, то це створить додаткові завеликі стрибки ламаної (рис. 2.4)

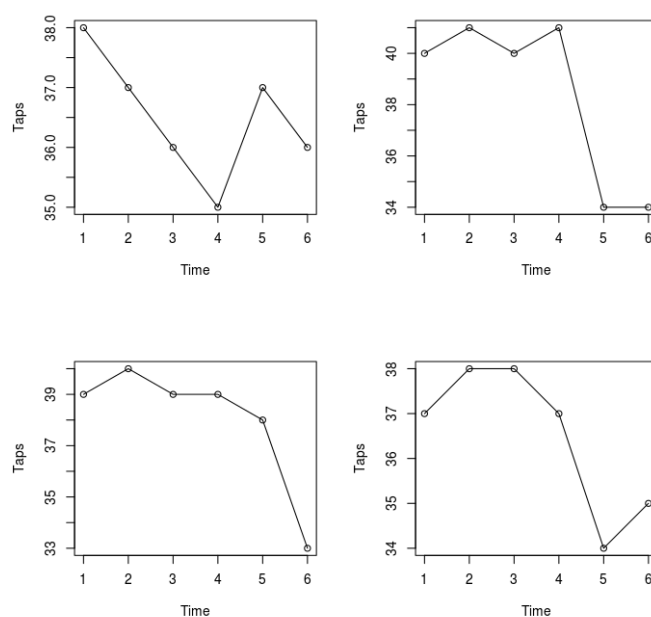


Рисунок 2.4 — Результати з додаванням центрованих Бернуллієвських випадкових величин з параметрами $(2, 0.5)$

Дана модель демонструє себе як просту, але не завжди адекватну: її результати потрібно фільтрувати. Тим не менш, нестандартні результати можна трактувати як ненормальну поведінку студента, якому потрібно пояснити, що під час контрольної роботи або екзамену він повинен бути більш зібраним та повністю концентруватися тій на роботі, що потрібно виконати.

2.1.8 Гідравліка

Для моделювання результатів теплінг-тесту постулюємо такі правила:

- 1) У студента є певний запас енергії для виконання роботи, який з часом вичерпується, а накопичується лише після відпочинку (зміну роду занять)
- 2) Коли студент працює на повну силу, він втрачає енергію
- 3) Чим менше енергії залишається, тим повільніше вона витрачається
- 4) Адаптованість — поява додаткової енергії

Це нагадує процес витікання рідини при змінному напорі води — коли її рівень в посудині змінюється. Якщо рівень рідини в цистерні змінюється достатньо повільно, можна скористатися наступним співвідношенням [5]

$$A \cdot dh = -\mu \cdot S \cdot \sqrt{h} \cdot dt \quad (2.5)$$

A — площа поперечного перерізу посудини на висоті h від отвору;

S — площа поперечного перерізу отвору внизу посудини;

dh — зміна висоти рідини за час dt ;

μ — константа, яка включає в себе прискорення вільного падіння, стиснення рідини біля отвору тощо, але в поточній задачі її сенс, звісно, буде іншим.

Моментальна швидкість зміни об'єму рідини буде рахуватися за формулою

$$v(t) = \frac{dV}{dt} = \frac{d(h \cdot A)}{dt} \quad (2.6)$$

Коли площі поперечних перерізів цистерни та отвору постійні, маємо таке рівняння для зміни висоти рідини

$$\dot{h} = -\mu \cdot \frac{S}{A} \cdot \sqrt{h}, \quad \dot{A} = \dot{S} = 0$$

Отримуємо лінійне диференціальне рівняння з розділеними змінними

$$h^{-\frac{1}{2}} \cdot dh = -\mu \cdot \frac{S}{A} \cdot dt$$

Інтегруємо в межах від $t_0 = 0$ до t

$$\int_{h_0}^{h(t)} h^{-\frac{1}{2}} dh = -\mu \cdot \frac{S}{A} \cdot \int_0^t dt$$

Та маємо розв'язок

$$h(t) = \left(\sqrt{h_0} - \frac{\mu}{2} \cdot \frac{S}{A} \cdot t \right)^2 \quad (2.7)$$

Швидкість зміни об'єму

$$v(t) = A \cdot \dot{h} = -\mu \cdot S \cdot \sqrt{h_0} \cdot \left(1 - \frac{\mu \cdot S}{2 \cdot \sqrt{h_0}} \cdot t \right) \quad (2.8)$$

Оскільки вважається, що студент працює максимально продуктивно, а це означає, що швидкість спорожнення цистерни максимальна, то з розв'язків бачимо, що це досягається тоді, коли площа отвору S максимальна. Оскільки вона не може бути більша за A , то залишається лише прирівняти $A = S$. Оскільки рівень рідини повинен зменшуватися повільно, то висота h_0 повинна бути набагато більшою за S — маємо витікання рідини з трубки. В такому разі залежність висоти рідини від часу (2.7) приймає наступний вигляд

$$h(t) = \left(\sqrt{h_0} - \frac{\mu}{2} \cdot t \right)^2$$

Щоб враховувати адаптованість, потрібно додати ще одне джерело рідини, з якого вона буде виливатися в досліджувану цистерну.

Нехай V_2 — об'єм додаткової цистерни, тоді v_2 — швидкість зміни об'єму рідини, що в ній знаходиться. Відомо, що

- 1) Нехай швидкість витікання води з основної цистерни не може перевищувати початкове значення $v(0)$ (умовилися з вигляду результатів теппінг-тесту)

- 2) В певний момент часу τ перетікання води зупиняється
- 3) В момент часу τ в основній посудині знову знаходиться така кількість рідини, що була спочатку

$$\int_0^{\tau} v_2 dt = \int_0^{\tau} v dt$$

З останнього факту легко знайти об'єм додаткової рідини V_a звичайним інтегруванням

$$V_a = - \int_0^{\tau} v dt = \mu \cdot S \cdot \sqrt{h_0} \cdot \left(1 - \frac{\mu \cdot S}{4 \cdot \sqrt{h_0}} \cdot \tau \right) \cdot \tau$$

Бачимо, що на даний момент важко сказати ще щось окрім того, що коли додаткова рідина поступає нерівномірно (не одразу активується механізм адаптованості), то

- 1) Якщо швидкість витікання води з основної цистерни не перевищує початкове значення через те, що її висота дорівнює h_0 і вона фізично не може містити більше рідини, ніж спочатку, то площа отвору S повинна бути меншою за площу перерізу посудини A , інакше неможливо буде вливати додаткову рідину таким чином, щоб її швидкістю можна було знехтувати
- 2) Друга цистерна повинна мати регульовану площу отвору, щоб була можливість моделювати опуклі ламані

Використаємо рівняння (2.5), щоб підрахувати зміну площі отвору, яка повинна бути для підтримання постійної швидкості зміни об'єму протягом певного часу τ (позначення ті ж самі, але використаємо результати для другої цистерни)

$$\begin{cases} \dot{h} = -\mu \cdot \frac{S}{A} \cdot \sqrt{h} \\ \dot{v} = 0 \end{cases} \Rightarrow \begin{cases} \dot{h} = -\mu \cdot \frac{S}{A} \cdot \sqrt{h} \\ \ddot{h} = 0 \end{cases}, \quad \dot{A} = 0 \quad (2.9)$$

Об'єднаємо систему рівнянь в одне

$$\frac{d(S \cdot \sqrt{h})}{dt} = 0$$

Диференціюємо добуток. Це не складно, але буде використано далі

$$\dot{S} = -\frac{S}{2 \cdot h} \cdot \dot{h} \quad (2.10)$$

Маємо залежність S від h

$$S(h) = S_0 \cdot \sqrt{\frac{h_0}{h}} \quad (2.11)$$

Тобто, швидкість зміни висоти рідини дорівнює

$$\dot{h} = -\mu \cdot \frac{S_0}{A_0} \cdot \sqrt{h_0}$$

Швидкість зміни об'єму рідини, як нескладно здогадатися з (2.8), рахується за формулою

$$v = -\mu \cdot S_0 \cdot \sqrt{h_0}$$

Скористуємось (2.10), щоб знайти залежність висоти від часу, адже нам потрібно знати, скільки рідини повинно бути у другій цистерні

$$2 \cdot \frac{\dot{S}}{S} = -\frac{\dot{h}}{h}$$

Застосуємо (2.9)

$$2 \cdot \frac{\dot{S}}{S} = \frac{\mu \cdot S \cdot \sqrt{h}}{A \cdot h}$$

Розділяємо змінні та згадуємо розв'язок рівняння (2.11) відносно \sqrt{h}

$$2 \cdot \sqrt{h_0} \cdot S_0 \cdot S^{-3} \cdot dS = \frac{\mu}{A} \cdot dt$$

Після інтегрування в межах від 0 до τ Маємо розв'язок

$$S(\tau) = S_0 \cdot \left[1 - \mu \cdot \frac{S_0}{A \cdot \sqrt{h_0}} \cdot \tau \right]^{-\frac{1}{2}} = \frac{S_0}{\sqrt{1 - \mu \cdot \frac{S_0}{A \cdot \sqrt{h_0}} \cdot \tau}}$$

Автоматично з (2.11) отримали

$$h(\tau) = h_0 \cdot \left[1 - \mu \cdot \frac{S_0}{A \cdot \sqrt{h_0}} \cdot \tau \right]$$

Маємо залежності між параметрами другої цистерни для побудови ламаної рівного типу. Залишилося розв'язати задачу для змінної швидкості, щоб будувати опуклі ламані.

Далі потрібно врахувати такі факти:

- 1) Площа отвору додаткової цистерни не може перевищувати площу перерізу основної, бо це не має сенсу
- 2) Площа отвору додаткової цистерни не може перевищувати площі свого перерізу, адже це теж не має сенсу

Сформулюємо обмеження математично

$$\begin{cases} S_2 & \leq A_2 \\ S_2 & \leq A \\ S & \leq A \end{cases}$$

ПЕРЕЛІК ПОСИЛАНЬ

1. Айвазян, С.А. Прикладная статистика: Классификация и снижение размерности : Справочное издание / С.А. Айвазян. — Финансы и статистика, 1989.
2. *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.
3. Павлов, И.П. Двадцатилетний опыт объективного изучения высшей деятельности (поведения) животных / И.П. Павлов. — Государственное издательство Москва-Петроград, 1923.
4. Ильин, Е.П. Дифференциальная психофизиология / Е.П. Ильин. Серия Учебник нового века. — Питер, 2001.
5. Шейпак, А.А. Основы механики жидкости и газа / А.А. Шейпак. — МГИУ, 2007.