

РОЗРОБКА УЗАГАЛЬНЕНОГО АЛГОРИТМУ АНАЛІЗУ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ ДЛЯ НАВЧАННЯ МАЙБУТНІХ ІНЖЕНЕРІВ-ПЕДАГОГІВ

Розглянуто процес розробки алгоритму аналізу експериментальних даних в умовах задач з різних предметних галузей, який становитиме підґрунтя для проведення коректного аналізу експериментальних даних.

Ключові слова: алгоритм, експериментальні дані, інженери-педагоги, модель, ознаки експериментальних даних.

Постановка проблеми. На сьогодні не існує науково-обґрунтованої ефективної методики проведення аналізу даних з різних предметних галузей для студентів інженерно-педагогічних спеціальностей. Якщо розглядаються деякі методи то вони мають поверхневий хаотичний характер загального змісту, тому проблема аналізу даних та вибору відповідного статистичного методу з використанням комп'ютерних технологій опрацювання експериментальних даних є актуальною та потребує глибокої розробки.

Аналіз останніх досліджень та публікацій. Проблемою структурування змісту, моделювання методів навчання для студентів інженерно-педагогічних спеціальностей займалися такі дослідники Галазунова А.Т. [1], Донеллі-мл. Р. [3], Дорошенка Ю.О. та Ротаєнка П.А. [4], Жалдака М.І., Кізьмінової Н.М. та Михаліна Г.О. [5], Лазарева М.І. [7], , Томашевського О.В. [10].

Багато праць присвячено проблемам змісту та методів навчання майбутніх інженерів-педагогів, але на жаль не приділяється достатньої уваги питанням опрацювання експериментальних даних та проведенню попереднього їх аналізу.

Постановка завдання. Основною метою даної роботи є розробка модулів алгоритму аналізу експериментальних даних на основі раніше створеної узагальненої ознакової моделі аналізу даних [8, с. 129] для навчання майбутніх інженерів-педагогів проведенню коректного аналізу експериментальних даних в умові задачі.

Виклад основного матеріалу. Розглянемо процес аналізу експериментальних даних в умові задачі. На основі раніше розробленої узагальненої ознакової моделі аналізу експериментальних даних в умовах задач [8, с. 129], яка має таку структуру: встановлення факторної та результативної змінних, опис даних на основі шкал виміру, визначення закону розподілу, якому підпорядковуються дані, вивчення мети проведення аналізу експериментальних даних.

Зазначимо вимоги до моделювання процесу аналізу експериментальних даних на основі використання узагальненої ознакової моделі аналізу експериментальних даних:

1. Модель має складатися з чітко окреслених модулів, які відповідають кожному окремому виду ознак даних.
2. При роботі з кожним модулем необхідно здійснити відповідну кількість кроків для побудови ознакової моделі яка описується цим модулем.
3. Модель має забезпечити неперервне виконання кожного кроку кожного модуля алгоритму для побудови ознакової моделі експериментальних даних.
4. Модель процесу аналізу експериментальних даних повинна відображати відповідну діяльність інженера-педагога.

Результати аналізу праць П. Гальперіна [2], О. Матюшкіна [7], Н. Талізінної [9] свідчать, що всім вищезазначеним вимогам відповідає модель процесу аналізу експериментальних даних у вигляді алгоритму.

Розв'язання задач з різних предметних галузей діяльності майбутнього інженера-педагога слід починати з визначення зв'язків між змінними, в залежності від постановки умови задачі та особливостей предметних галузей, так як змінні можуть змінювати свій вплив, результативна змінна може стати факторною, а факторна – результативною. При цьому слід бути дуже уважному, неправильний вибір напряму змінних приведе до безглузких результатів.

Встановлення зв'язку між змінними, як в умові задачі, так і при виборі статистичного методу аналізу слід починати з визначення факторної змінної – z_1 , якщо сукупність являється однорідною (одна чи декілька ознак, що вивчаються, є загальними для всіх одиниць, і різнорідна сукупність об'єднує явища різного типу), то факторна змінна завжди наявна. Якщо сукупність є різнорідною то факторну змінну визначити не можливо, необхідно провести уточнення умови задачі. У переважній більшості задач з усіх предметних галузей експериментальні дані є однорідними і ми можемо ввести z_1 до ознакової моделі даних задачі.

Аналогічно визначаємо наявність результативної змінної в даних задачі – z_2 . Якщо результативна ознака наявна в експериментальних даних умови задачі, то її необхідно включити до ознакової моделі, що представляє зв'язки між даними.

В результаті визначення факторної та результативної ознак даних в умові задачі, отримуємо таку модель:

$$z(X,Y)=\{z_1, z_2\}.$$

Процес визначення факторної і результативної змінних можна представити у вигляді алгоритму з використанням відповідних позначень ознак даних (рис. 1).

Отже, розроблений модуль алгоритму дозволяє визначити наявність змінних в задачі та встановити їх роль та в подальшому співставити дані ознаки умови задачі з відповідними виділеними ознаками статистичних методів.

Проаналізуємо змінну X , яка наявна в умові задачі та визначимо характер типу даних до якого вона належить відповідно до ознакової моделі аналізу даних умови задачі на основі шкал виміру.

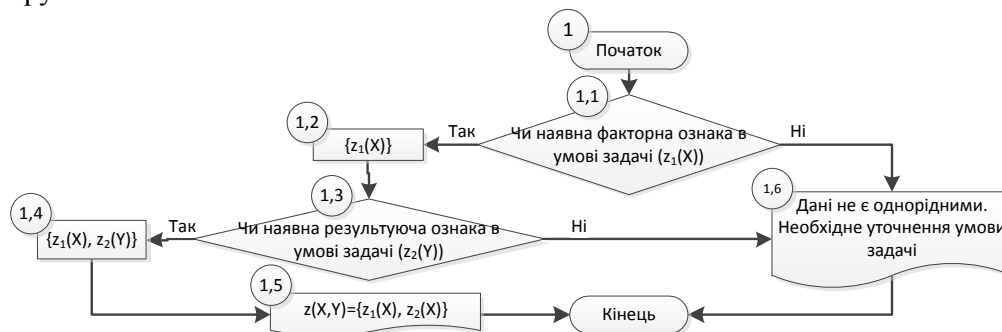


Рис. 1. Перший модуль алгоритму побудови ознакової моделі експериментальних даних

Для того, щоб спроектувати другий модуль алгоритму необхідно детально перевірити кожен ознаку шкали виміру та встановити тип шкали і визначити тип даних (кількісні – h_2 або якісні – h_1 дані).

Розглянемо шкалу найменувань ($h_{1,1}$) та проаналізуємо ознаки, які її характеризують. Дослідимо ознаку $h_{1,1,1}$ (відсутність числових характеристик), якщо дані змінної X відрізняються одне від одного тільки якісною характеристикою (фізика, математика, хімія – предмети, що вивчаються учнями) то вони відповідають даній ознаці. Якщо ця умова виконується то переходимо до розгляду наступної ознаки $h_{1,1,2}$ (неможливо проводити з даними математичні операції крім співставлення (дорівнює «=», не дорівнює «≠»)), в протилежному випадку необхідно перейти до розгляду ознак порядкової шкали ($h_{1,2}$). Тобто будь-яка не відповідність ознак шкали виміру даних задачі призводить до переходу і аналізу наступної шкали виміру.

Розглянемо випадок, коли ознака $h_{1,1,1}$ шкали найменувань виконується і переходимо до дослідження ознаки $h_{1,1,2}$. Якщо з даними змінної X неможливо виконувати математичні операції крім співставлення (математика \neq фізика, хімія = хімія) то умова даної ознаки виконується. Переходимо до наступної ознаки цієї ж шкали виміру – $h_{1,1,3}$ (в номінальній шкалі досліджуваним об'єктам приписуються певні літерні або кодові (літерно-цифрові) значення), наприклад дані змінної X представлені так: математика (літерне) або 1А (літерно-цифрове) або 4-11-23 (кодове).

Проаналізуємо наступну ознаку шкали $h_{1,1} - h_{1,1,4}$ (неможливо впорядкувати об'єкти ознаки та відразу визначити їх кількість). Перевіряємо змінну X на наявність цієї ознаки, якщо в даних не можливо визначити більший або менший елементи то дані відповідають цій ознаці

(номер футболіста в команді, шкільні предмети).

Розглянемо ознаку $h_{1,1,5}$ (дані приймають лише два протилежних значення (дані виміряні в дихотомічній шкалі)), дана ознака є підмножиною номінальної шкали. Наприклад: ліворукий – праворукий; стать: чоловіча – жіноча; 0 або 1.

Проаналізувавши всі ознаки номінальної шкали виміру і з'ясувавши, що змінна X задовольняє всім ознакам цієї шкали та належить до якісного типу даних. Можемо записати таку наступну модель:

$$h_1 = h_{1,1} = \{h_{1,1,1}, h_{1,1,2}, h_{1,1,3}, h_{1,1,4}, h_{1,1,5}\}.$$

Записану модель можна представити у вигляді алгоритму з використанням відповідних позначень ознак даних (рис. 2).

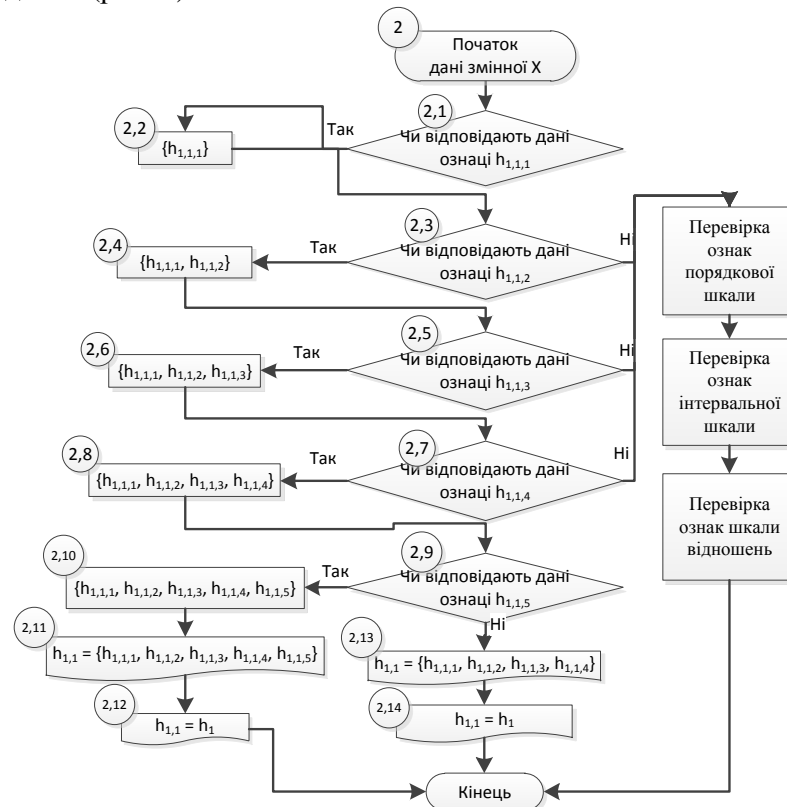


Рис. 2. Алгоритм аналізу змінної X на основі ознак номінальної шкали

Аналіз змінної X на відповідність ознакам порядкової шкали здійснюється за аналогічним алгоритмом. Модель для порядкової шкали матиме наступний вигляд:

$$h_1 = h_{1,2} = \{h_{1,2,1}, h_{1,2,2}, h_{1,2,3}, h_{1,2,4}, h_{1,2,5}\}.$$

Тобто, перевіряється наявність ознак, які належать до даної моделі:

- $h_{1,2,1}$ – відсутність числових характеристик, які мають одиницю виміру.
- $h_{1,2,2}$ – можливе здійснення порівняння та співставлення об'єктів за величиною ознаки (дорівнює « \equiv », не дорівнює « \neq », більше « $>$ », менше « $<$ »).
- $h_{1,2,3}$ – можливе приписування об'єктам певних літерних, кодових або числових значень.
- $h_{1,2,4}$ – дані можливо впорядкувати (за зростанням або спаданням значень даних).
- $h_{1,2,5}$ – дані можливо представити у вигляді рангів (проранжувати).

Записану модель можна представити у вигляді алгоритму з використанням відповідних позначень ознак даних (рис. 3).

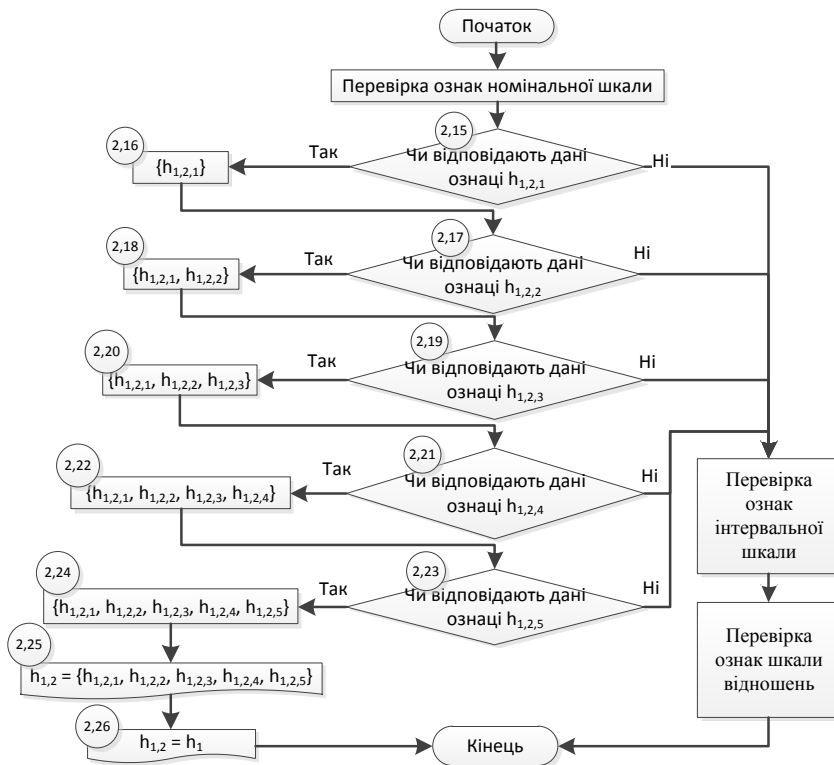


Рис. 3. Алгоритм аналізу змінної X на основі порядкової шкали

Якщо змінна X не відповідає ознакам порядкової шкали переходимо до аналізу цієї змінної на основі ознак кількісних типів даних, до яких відносяться інтервальна шкала та шкала відношень.

Перші ознаки ($h_{2,1,1}$ та $h_{2,2,1}$), які характеризують ці дві шкали виміру (дані виражені кількісно і мають одиницю виміру) – співпадають, тому їх аналіз здійснюється одночасно. Якщо дані змінної X в умові задачі виражені за допомогою числових значень то можемо ввести $h_{2,1,1}$ та $h_{2,2,1}$ до ознакової моделі даних задачі і перейти до розгляду наступних ознак.

Ознаки $h_{2,1,2}$ та $h_{2,2,2}$ (з динаміки можливо проводити порівняння в кількісній мірі (дорівнює «=», не дорівнює «≠», більше «>», менше «<»)) також співпадають і свідчать про можливість здійснення логічних операцій з даними змінної X. Якщо ці ознаки виконуються то можна перейти до розгляду наступної ознаки.

Розглянемо ознаку $h_{2,1,3}$ (відсутня точка абсолютного відліку) яка виконується тільки для інтервальної шкали виміру ($h_{2,1}$). Якщо в даних змінної X відсутня точка абсолютного відліку то ознаку $h_{2,1,3}$ можемо ввести до ознакової моделі даних задачі. У протилежному випадку переходимо до розгляду шкали відношень ($h_{2,2}$) і вводимо в ознакову модель даних задачі ознаку $h_{2,2,3}$ (наявна точка абсолютного відліку).

Повертаємось до інтервальної шкали виміру даних ($h_{2,1}$) і перевіряємо ознаку $h_{2,1,4}$ (над даними можна виконувати наступні математичні операції: додавання «+», віднімання «-»). Якщо умова ознаки виконується то відносимо змінну X до інтервальної шкали і кількісного типу даних.

Отже, можемо записати таку модель:

$$h_2 = h_{2,1} = \{h_{2,1,1}, h_{2,1,2}, h_{2,1,3}, h_{2,1,4}\}.$$

Розглянемо етап коли виконується ознака $h_{2,2,3}$, в даних наявна точка абсолютного відліку та перейдемо до розгляду наступної ознаки шкали відношень ($h_{2,2}$) – $h_{2,2,4}$ (над даними можна виконувати наступні математичні операції: додавання «+», віднімання «-», множення «*», ділення «/»). Якщо умова ознаки виконується то відносимо змінну X до шкали відношень і кількісного типу даних.

Отже, можемо записати таку модель:

$$h_2 = h_{2,2} = \{h_{2,2,1}, h_{2,2,2}, h_{2,2,3}, h_{2,2,4}\}.$$

Отримані моделі представимо у вигляді алгоритму аналізу кількісних даних, що виміряні в інтервальній шкалі або шкалі відношень (рис. 4).

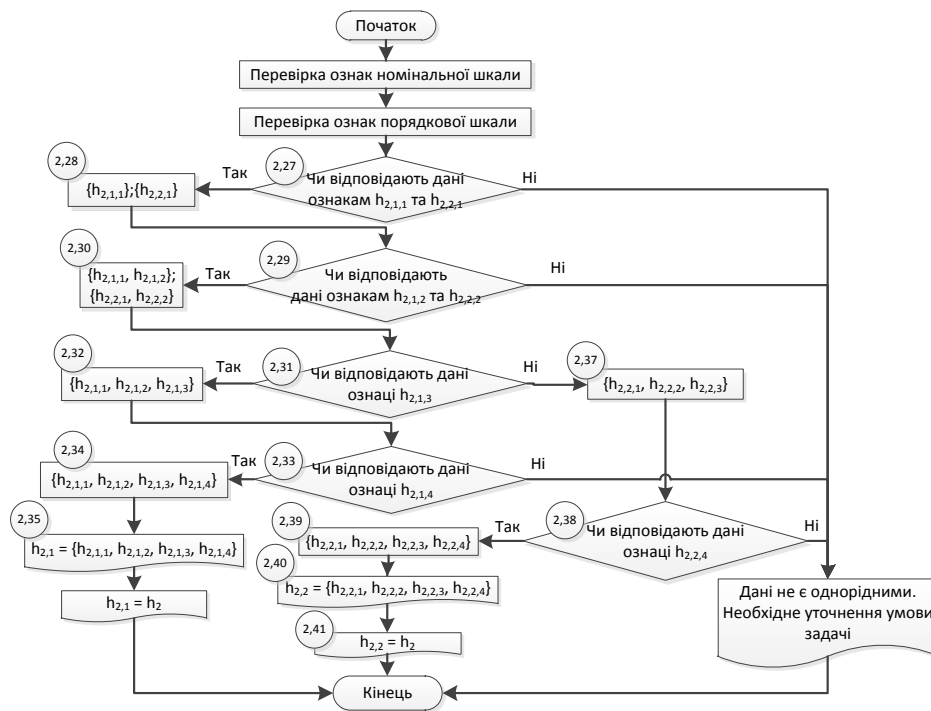


Рис. 4. Алгоритм аналізу змінної X на основі ознак інтервальної шкали та шкали відношень

На основі вищерозглянутих частин алгоритму будується узагальнений алгоритм аналізу даних змінної X на основі ознак типів шкал виміру.

Наступний третій модуль алгоритму полягає в аналізі змінної Y в умові задачі та виконується за аналогічним алгоритмом на основі ознак шкал виміру.

У четвертому модулі алгоритму виконується аналіз даних в умові задачі на визначення закону розподілу експериментальних даних (g).

Аналіз закону розподілу слід починати з визначення того, чи виміряні дані змінної X в шкалі найменувань або порядковій шкалі. Якщо ці дані належать до цих шкал то вони не підпорядковуються нормальному закону розподілу. У протилежно випадку необхідно перейти до визначення кількості даних в умові задачі (ознака $g_{1,1}$). Якщо кількість даних перевищує 25 спостережень то можна вважати, що дані підпорядковуються нормальному закону розподілу та одержуємо таку модель:

$$g(X) = g_1(X) = \{g_{1,1}(X)\}$$

Якщо кількість спостережень є меншою за 25, тобто ознака $g_{1,1}$ не виконується, то необхідно розглянути ознаку $g_{1,2}$ (побудова нормального ймовірнісного графіку). У випадку коли побудований нормальний ймовірнісний графік засвідчує що розподіл даних підпорядковуються нормальному закону то одержуємо наступну модель:

$$g(X) = g_1 = \{g_{1,2}\}.$$

В протилежному випадку дані не підпорядковуються нормальному закону розподілу і модель набуває наступного вигляду:

$$g(X) = g_2(X).$$

Розподіл даних змінної Y перевіряється аналогічним чином.

Процес визначення закону розподілу даних можна представити у вигляді алгоритму з використанням відповідних позначень ознак (рис. 5)

У п'ятому модулі узагальненого алгоритму необхідно здійснити визначення мети проведення аналізу експериментальних даних – v.

Всі вище розглянуті модулі утворюють цілісну структуру узагальненого алгоритму аналізу експериментальних даних.

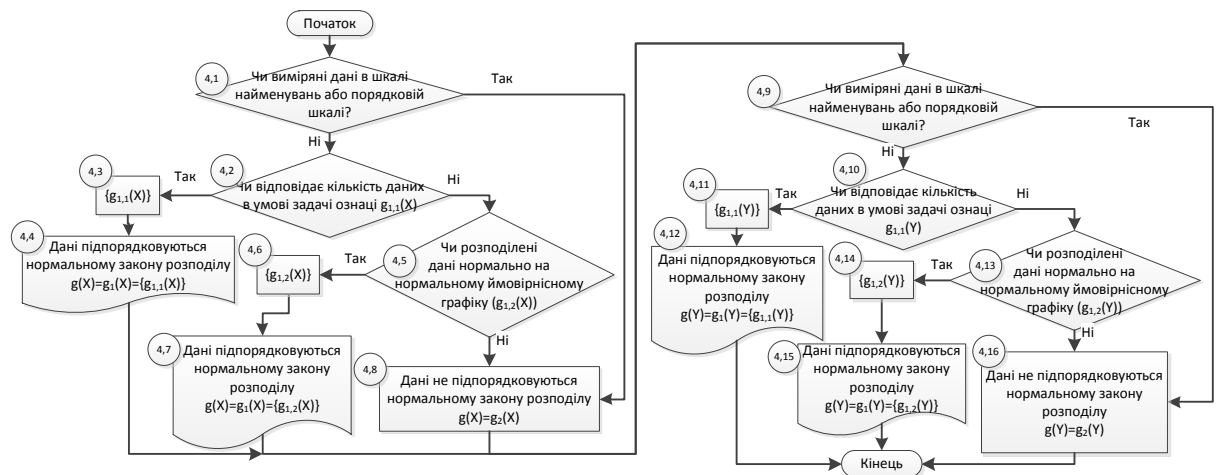


Рис. 5. Алгоритм визначення закону розподілу, якому підпорядковуються змінні в умові задачі

Висновок. Після детального, обґрунтованого, коректного аналізу даних умови задачі студентам необхідно перейти до вибору статистичного методу аналізу експериментальних даних. Існує дуже велика кількість статистичних методів аналізу експериментальних даних які є необхідними у майбутній діяльності інженерів-педагогів. На основі виділених ознак аналізу даних умови відповідної задачі та детального проходження кожного етапу алгоритму, студенти зможуть обрати адекватний метод аналізу експериментальних даних, співставивши ознакову модель даних умови задачі та ознакові моделі статистичних методів аналізу які ми плануємо розробити в подальших дослідженнях.

Список використаної літератури

1. Галазунов А.Т. Педагогические исследования: содержание, организация и обработка результатов / А.Т. Галазунов. – М.: Издательский центр АПО, 2003. – 41 с.
2. Гальперин П.Я. Психология как объективная наука: Избр. психол. тр. / П.Я. Гальперин; под ред. А.И. Подольского. - М.: Москв. психол.-соц. ин-т; Воронеж : МОДЭК, 2003. - 478 с.
3. Донелли-мл. Р. Статистика / Роберт А. Донелли-мл.; пер. с англ. Н.А. Ворониной. – М.: Астрель: АСТ, 2007. – XIV, 367, [3] с. ил.
4. Дорошенко Ю.О. Достовірність комп'ютерного тестування: Навч.-метод. посіб. / Ю.О. Дорошенко, П.А. Ротаєнко; за ред. Ю.О. Дорошенка. – К.: Педагогічна думка, 2007. – 176 с.
5. Жалдак М.І. Теорія ймовірностей і математична статистика: підручник для студентів педагогічних університетів / М.І. Жалдак, Н.М. Кузьміна, Г.О. Михалін. – К.: Полтава-Довкілля, 2009. – 498 с.
6. Лазарев М.І. Полісистемне моделювання змісту технологій навчання загальноінженерних дисциплін: монографія / М.І. Лазарев. – Х.: Вид-во НФаУ, 2003. – 356 с.
7. Матюшкин А.М. Мышление, обучение, творчество / А.М. Матюшкин. - М.: Изд-во Моск. психол.-соц. ин-та; Воронеж: Изд-во НПО "Модек", 2003. - 718, [1] с.
8. Павленко Л. В. Методика створення ознакових моделей аналізу психолого-педагогічних, економічних та технічних даних в умовах задач / Л. В. Павленко // Комп'ютерно-інтегровані технології: освіта, наука, виробництво". - 2011. – Вип. № 4 - С. 123-129.
9. Талызина Н.Ф. Управление процессом усвоения знаний: (Психол. основа) / Н. Ф. Талызина. - [2-е изд.], доп. и испр. - М.: Изд-во МГУ, 1984. - 344 с.;
10. Томашевський О. В. Комп'ютерні технології статистичної обробки даних: навч. посібник [для студ. вищ. навч. закладів, які навчаються за спец. «Якість, стандартизація та сертифікація»] / О. В. Томашевський, В. П. Рисіков. – Запоріжжя: Запорізький національний технічний ун-т, 2006. – 174 с.

Стаття надійшла до редакції 28.02.2012.

Л. В. Павленко. Разработка обобщенного алгоритма анализа экспериментальных данных для обучения будущих инженеров-педагогов.

Рассмотрен процесс разработки алгоритма анализа экспериментальных данных в условиях задач разных предметных областей, который будет представлять основу для проведения корректного анализа экспериментальных данных.

Ключевые слова: алгоритм, экспериментальные данные, инженеры-педагоги, модель, признаки экспериментальных данных.

L. Pavlenko. The Generalized Algorithm Development of Experimental Data Analysis for Future Engineers-Teachers Training.

A development of algorithm of experimental data analysis process is considered in the conditions of tasks of different subject domains, which will present basis for realization of correct analysis of experimental data.

Keywords: algorithm, experimental data, engineers-teachers, model, signs of experimental data.