

оновлення роботи програмного модулю моніторингу метеорологічних даних та автоматичне створення архівів заповненої бази даних за кожен день. Створено журнал подій та журнал помилок. Програмний модуль парсингу метеорологічних даних було інтегровано в роботу сайту моніторингу теплозбереження будівель СумДУ. Розроблений програмний продукт може бути інтегровано в роботу інших підприємств та організацій, де потрібні кліматичні показники для специфічних розрахунків.

Список літератури: 1. Шендрік В.В. Система збирання, розміщення та аналізу даних / В.В. Шендрік, С.М. Ващенко // Вісник Національного університету "Львівська політехніка". Серія "Інформаційні системи та мережі" – 2011. – № 715. – С. 343-353.

Поступила в редколегію 15.06.2012

УДК 519.234

В. О. ШЕВЧЕНКО, асис., ХНАДУ, Харьков

ПРОВЕРКА ДОСТОВЕРНОСТИ РЕЗУЛЬТАТОВ МОДЕЛИРОВАНИЯ МЕТОДАМИ НЕПАРАМЕТРИЧЕСКОЙ СТАТИСТИКИ

Описана методика перевірки достовірності імітаційної моделі придбання знань студентами з дисципліни "Інформатика" за допомогою методів непараметричної статистики: медіанним критерієм та критерієм Вілкоксона-Манна-Уїтні

Ключові слова: непараметрична статистика, нульова гіпотеза, альтернативна гіпотеза, критерій, моделювання

Описана методика проверки достоверности имитационной модели приобретения знаний студентами по дисциплине «Информатика» с помощью методов непараметрической статистики: медианным критерием и критерием Вилкоксона-Манна-Уитни

Ключевые слова: непараметрическая статистика, нулевая гипотеза, альтернативная гипотеза, критерий, моделирование

A technique for validating a simulation model of knowledge acquisition by students on academic discipline "Computer Science" by means of nonparametric statistics methods application : the median test and the Wilcoxon-Mann-Whitney test is described.

Keywords: nonparametric statistics, null hypothesis, alternative hypothesis, test, simulation.

1. Введение

Результаты педагогических измерений, полученные в процессе педагогических исследований, нуждаются в математической обработке. Во многих случаях для сравнения полученных результатов различных измерений используют средние значения. Главным недостатком этого метода является его качественный характер и неопределенность. Статистические методы позволяют численно оценить достоверность произведенных сравнений. Глубокая статистическая обработка результатов измерений допустима в случае, когда известны законы распределения измеряемых случайных величин. К сожалению, такие законы при педагогических измерениях практически неизвестны. В этом случае, единственным математическим аппаратом, позволяющим, получить количественные оценки является непараметрическая статистика [1].

2. Цель и постановка задачи

Выдвинута гипотеза, что результаты моделирования зачетного балла построенной модели приобретения знаний студентами по дисциплине «Информатика» соответствуют реальным зачетным баллам студентов, полученным ими на зачете по информатике.

Цель: подтвердить выдвинутую гипотезу методами непараметрической статистики.

Для подтверждения гипотезы воспользовались двумя методами непараметрической статистики: медианным критерием и критерием Вилкоксона-Манна-Уитни.

3. Проверка достоверности результатов моделирования по медианному критерию

Проверку проведем по методике, описанной в работе [2]. Исходные данные: вектор $X = \{x_1, x_2, \dots, x_i, \dots, x_N\}$ – вектор баллов, полученных студентами на зачете; вектор $Y = \{y_1, y_2, \dots, y_i, \dots, y_N\}$ – вектор зачетных баллов, рассчитанных для тех же студентов при помощи модели. Шкала измерений 100-балльная ($M=100$). Анализируются данные 136 студентов ($N = 136$).

Алгоритм применения медианного критерия:

1) Формируем вектор Z путем объединения векторов X и Y :

$$Z = X \cup Y = \{x_1, x_2, \dots, x_i, \dots, x_N, y_1, y_2, \dots, y_i, \dots, y_N\}, \quad (1)$$

где $i = \overline{1, N}$; N – длина векторов X и Y .

2) путем суммирования количества одинаковых баллов для векторов X , Y и Z строятся соответствующие совокупности абсолютных частот появления баллов:

$$\begin{aligned} XA &= \{xa_1, xa_2, \dots, xa_j, \dots, xa_M\}, \\ YA &= \{ya_1, ya_2, \dots, ya_j, \dots, ya_M\}, \\ ZA &= \{za_1, za_2, \dots, za_j, \dots, za_M\}, \end{aligned} \quad (2)$$

где $j = \overline{1, M}$; M – количество градаций в выбранной шкале измерений; xa_j, ya_j, za_j – частоты появления баллов в векторах X и Y соответственно:

$$\begin{aligned} xa_j &= \sum_{i=1}^N (x_i = g_j), \\ ya_j &= \sum_{i=1}^N (y_i = g_j), \\ za_j &= \sum_{i=1}^N (z_i = g_j), \end{aligned} \quad (3)$$

где g_j – j -ое значение в выбранной шкале.

3) Строится вектор Z_Σ путем последовательного суммирования частот вектора ZA :

$$Z_\Sigma = \sum_{i=1}^N za_i. \quad (4)$$

4) Сформированные данные заносятся в таблицу.

5) Определяется строка MED , в которой находится медиана объединенной выборки MED_Z :

$$MED = \frac{\max(Z_{\Sigma})}{2}. \quad (5)$$

$$MED = \frac{271}{2} = 135,5.$$

6) В таблице выделяется строка, в которой находится значение MED . В нашем случае значение $MED = 135,5$ лежит между значениями шкалы 68 и 69. Следовательно, $MED_Z = 68,5$.

7) Для векторов XA и YA суммируются частоты над и не над строкой MED :

$$\begin{aligned} Q_{1X} &= \sum_{j > MED_Z} xa_j, \\ Q_{2X} &= \sum_{j \leq MED_Z} xa_j, \\ Q_{1Y} &= \sum_{j > MED_Z} ya_j, \\ Q_{2Y} &= \sum_{j \leq MED_Z} ya_j. \end{aligned} \quad (6)$$

Результаты заносятся в таблице.

Таблица - Суммы частот над и не над строкой MED

	XA	YA
$MED_Z >$	$Q_{1X} = 69$	$Q_{1Y} = 68$
$\leq MED_Z$	$Q_{2X} = 66$	$Q_{2Y} = 68$

8) Так как все $Q_{ij} > 5$ (см. табл. 1), медианный критерий применим. Вычисляется наблюдаемое значение статистики:

$$T_i = \frac{\max(Z_{\Sigma}) (|Q_{1X} \cdot Q_{2Y} - Q_{2X} \cdot Q_{1Y}| - \max(Z_{\Sigma})/2)^2}{(Q_{1X} + Q_{1Y})(Q_{2X} + Q_{2Y})(Q_{1X} + Q_{2X})(Q_{1Y} + Q_{2Y})}, \quad (7)$$

где T_n – наблюдаемое значение статистики.

$$T_i = \frac{271(|69 \cdot 68 - 66 \cdot 68| - 271/2)^2}{(69 + 68)(66 + 68)(69 + 66)(68 + 68)} \approx 0,003773.$$

При уровне значимости $\alpha = 0,05$ критическое значение статистики $T_{кр} = 3,841$ [2]. Так как $T_n < T_{кр}$, проверяемая гипотеза принимается.

4. Проверка достоверности результатов моделирования по критерию Вилкоксона-Манна-Уитни

Алгоритм применения критерия Вилкоксона-Манна-Уитни:

1) Элементам вектора X устанавливается метка $R_i = 0, i = \overline{1, N}$; элементам вектора Y устанавливается метка $R_i = 1, i = \overline{1, N}$. Элементы с метками $R_i = 0$ и $R_i = 1$ объединяются в вектора $ПХ$ и $ПУ$ соответственно.

2) Векторы X и Y объединяются в вектор $Z = X \cup Y = \{x_1, x_2, \dots, x_i, \dots, x_N, y_1, y_2, \dots, y_i, \dots, y_N\}$, векторы $ПХ$ и $ПУ$ объединяются в вектор $ПZ = ПХ \cup ПУ$.

3) Из вектора Z формируется вектор ZR путем упорядочивания по возрастанию элементов вектора Z . При этом элементы вектора $ПZ$ перемещаются

соответственно перемещению элементов вектора Z при упорядочивании, так, чтобы все элементы вектора ZR остались помеченными своими метками.

4) Производится ранжирование вектора ZR . Каждому элементу вектора ZR присваивается порядковый номер этого элемента в векторе ZR . Если в векторе ZR имеются несколько элементов, одинаковых по значению, каждому из этих элементов присваивается ранг, равный усредненному значению этих элементов. Из результатов ранжирования формируется вектор R .

5) Определяется наблюдаемое значение статистики по формуле (8):

$$T_i = S - \frac{N(N-1)}{2}, \quad (8)$$

где T_n – наблюдаемое значение статистики; N – длина вектора X или Y ; S – промежуточная величина, вычисляемая по формуле (9):

$$S = \sum_{i=1}^{2N} R_i. \quad (9)$$

Так как длина исходных векторов X и Y одинакова, для вычисления суммы по формуле (9) можно выбрать любой признак R_i . Принимаем $R_i = 0$. В этом случае получаем:

$$S = \sum_{i=1}^{2N} (R_i = 0) = 18415.$$

Полученное значение подставляется в формулу (8):

$$T_i = 18415 - \frac{136(136-1)}{2} = 9235.$$

6) Так как $N > 20$, $T_{кр}$ вычисляется по формуле (10):

$$T_{кр} = \frac{N^2}{2} + Q \sqrt{\frac{N^2(2N+1)}{12}}, \quad (10)$$

где Q – квантиль.

При уровне значимости $\alpha = 0,05$, $Q = 1,96$. Получаем:

$$T_{кр} = \frac{136^2}{2} + 1,96 \sqrt{\frac{136^2(2 \cdot 136 + 1)}{12}} = 10519,41.$$

7) Определяется $T_{кр1}$ и $T_{кр2}$ по соотношениям (11) и (12):

$$T_{кр1} = N^2 - T_{кр}, \quad (11)$$

$$T_{кр1} = 136^2 - 10519,41 = 7976,59.$$

$$T_{кр2} = T_{кр}, \quad (12)$$

$$T_{кр2} = 10519,41.$$

Наблюдаемое значение статистики не попадает в критические области ($T_n \geq T_{кр1}$ и $T_n \leq T_{кр2}$). Следовательно, проверяемая гипотеза принимается.

5. Вывод

Расчеты, проведенные с помощью методов непараметрической статистики, позволяют утверждать, что результаты моделирования и педагогического эксперимента подчиняются одинаковым законам, и это подтверждает достоверность разработанной модели.

Список литературы: 1. Метешкин, А. А. Педагогические исследования в высших учебных заведениях / А. А. Метешкин – Харьков, ХНАДУ, 2006. – 97 с. 2. Грабарь, М. И. Применение математической статистики в педагогических исследованиях. Непараметрические методы / М. И. Грабарь, К. А. Краснянская – М., Педагогика, 1977. – 136 с. 3. Паповян, С. С. Математические методы в социальной психологии / С. С. Паповян – М., Наука, 1983. – 343 с.

Поступила в редколлегию 09.06.2012

УДК 378

В.Л. МАРЧЕНКО, докт. техн. наук, проф., «ХНУ», Хмельницький,
Л.С. ДРАПАК, канд.фіз.-мат.наук, доц., «ХНУ», Хмельницький

ОСОБЛИВОСТІ ЗАСТОСУВАННЯ ТАБЛИЧНОГО ПРОЦЕСОРА MS EXCEL ДЛЯ РОЗВ'ЯЗУВАННЯ НЕЛІНІЙНИХ РІВНЯНЬ

Розглянуто використання засобу табличного процесора MS Excel «Пошук рішення» при розв'язуванні нелінійних рівнянь. Показано, що існують випадки отримання помилкових результатів та досліджено їх причину. Рекомендовано можливості використання програмного засобу для аналізу складних функцій.

Ключові слова: надбудова процесора, нелінійні функції, обмеження.

Рассмотрено использование средства табличного процессора MS Excel «Поиск решения» при решении нелинейных уравнений. Показано, что существуют случаи получения ложных результатов и исследованы их причины. Рекомендовано возможности использования программного средства для анализа сложных функций.

Ключевые слова: надстройка процессора, нелинейные функции, ограничения.

The use of means of tabular processor MS Excel «Finding solutions » in the solution of nonlinear equations. It is shown that there are cases of a false result and investigated their cause. Recommended possibility of using software for analysis of complex functions

Keywords: Superstructure processor, nonlinear functions, limits of authority

Вступ

Дослідження даної статті відносяться до галузі науки “Прикладна математика та комп'ютерні технології”. Практика застосування та подальшого розв'язування рівнянь будь – якого типу показує, що цей процес є досить трудомістким і таким, при якому не завжди вдається отримати результат в аналітичному вигляді. В цьому випадку доцільно використовувати чисельні методи розрахунку. Табличний процесор MS Excel має досить багато засобів для проведення таких розрахунків, серед яких добре відомі надбудови «Підбір параметрів», «Сценарії» та «Пошук рішення». Останній з цих програмних засобів, хоча й відомий як засіб, призначений для розв'язування задач оптимізації, має можливість і для розв'язування рівнянь. Саме таке застосування надбудови розглядається в даній роботі.

Аналіз літературних даних та постановка проблеми

«Пошук рішення» вимагає виконання певної послідовності дій, серед яких є наступні: введення цільової функції, призначення комірок електронної таблиці невідомим, задання значення правої частини рівняння, введення обмежень. Ця схема дій в більшості випадків забезпечує отримання результату при коректній