УДК 378.14:330

Д.А. Найко, Вінниця, Україна / D. Naiko, Vinnytsia, Ukraine
e-mail: dmnaiko@ukr.net

ПЕРЕВІРКА НЕПАРАМЕТРИЧНИХ ГІПОТЕЗ У ПЕДАГОГІЧНОМУ ЕКСПЕРИМЕНТІ

Анотація. *У роботі розкривається суть деяких найпростіших непараметричних методів математичної статистики перевірки статистичних гіпотез. Вказано на певні можливості їх застосування при перевірці статистичних гіпотез в педагогічних дослідженнях. Наводиться приклад перевірки непараметричних гіпотез у експерименті, пов'язаному з вивченням комунікативної компетентності випускників ВНЗ. При формуванні контрольних та експериментальних груп, для забезпечення достеменності результатів експерименту одним із найважливіших моментів є однорідність цих груп. Ми показуємо методику перевірки однорідності експериментальних та контрольних груп, яка спирається зокрема на статистичний критерій Вілкоксона та поняття нормального розподілу випадкової величини . Ми розглядаємо також інші непараматричні критерії, такі як критерій серій та критерій знаків. Наводимо приклад застосування критерію знаків при порівнянні двох методик.*

Ключові слова*: нульова гіпотеза, альтернативна гіпотеза, розподіл випадкової величини, вибірка, статистичний критерій, комунікативна компетентність, критерій Вілкоксона.*

Check of nonparametric hypothesis in a pedagogical experiment

Annotation. *The paper reveals the essence of some of the simplest nonparametric methods of mathematical statistics verification of statistical hypotheses. Certain possibility of their use in verifying statistical hypotheses in pedagogical research are pointed out. An example of nonparametric hypotheses check in the experiment, related to the study of communicative competence of university graduates, is given.*

In forming the control and experimental groups for ensuring the validity of experimental results one of the most important things is homogeneity of these groups. We demonstrate the technique of checking the homogeneity of the experimental and control groups, which is based on statistical Wilcoxon criterion and the concept of normal distribution of a random variable.

We are considering other nonparametric criteria, such as the run test and the sign test. An example of nonparametric hypotheses check for the sign test when comparing the two methods is given.

Key words: *null hypothesis, alternative hypothesis, variable distribution, sample, statistic criterion, communicative competence, Wilcoxon criterion.*

Вступ. У деяких наукових працях, зокрема працях молодих науковців, легко помітити слабкість чи навіть помилковість методики обробки результатів педагогічного дослідження. Це стосується як констатувального, так і заключного етапів експерименту. У цій роботі ми зосереджуємо свою увагу, значною мірою, лише на методиці перевірки однорідності експериментальної та контрольної груп на початковому етапі педагогічного експерименту за критерієм Вілкоксона, проте беремо до розгляду і такі відомі непараметричні методи як метод серій та метод знаків [1, с. 281].

При обробці результатів досліджень, зокрема педагогічних, розподіл генеральної сукупності часто є невідомим або відрізняється від нормального (у випадку неперервних випадкових величин). Тому застосування відомих параметричних методів є неприйнятним, бо це може призвести до помилок. У такому разі використовують методи, які не залежать від характеру розподілу генеральної сукупності, це так звані *непараметричні методи*.

Непараметричні методи передбачають використання не самих числових значень вибірки, а структурних властивостей вибірки (наприклад, відношення порядку між її елементами тощо). Зрозуміло, що при цьому втрачається частина інформації, яка міститься у вибірці. Це означає, що потужність непараметричних методів є меншою аніж потужність параметричних. Проте непараметричні методи можуть бути застосовані при загальніших припущення щодо розподілів і є простішими з точки зору обчислювальної роботи.

Цілий ряд непараметричних методів використовують для перевірки гіпотези про належність двох вибірок $x_1, x_2, ..., x_{n_1}$ та $y_1, y_2, ..., y_{n_2}$ до однієї і тієї ж генеральної сукупності, тобто гіпотези про те, що функції розподілу $F_1(x)$ (для випадкової величини X) та $F_2(x)$ (для випадкової величини Y) двох генеральних сукупностей рівні: $F_1(x) \equiv F_2(x)$.

Такі генеральні сукупності називаються *однорідними*. Необхідною умовою однорідності є рівність числових характеристик досліджуваних генеральних сукупностях, зокрема середніх значень, дисперсій, медіан тощо. В якості основного припущення, при використанні непараметричних критеріїв, беруть лише *неперервність розподілу* генеральної сукупності.

Розглянемо найпростіші критерії такого типу – *критерій знаків*, *критерій Вілкоксона*, *критерій серій*.

Основна частина. 1. *Критерій знаків*. Задача перевірки нульової гіпотези H_0 про однорідність генеральних сукупностей за їх попарно пов'язаними вибірками виникає, наприклад, при *порівнянні двох методик* визначення одного й того самого показника. При дослідженні кожного з n об'єктів за першою та другою методиками отримують певні числові значення x_i та y_i ($i = 1, 2, ..., n$) відповідно. Якщо вибірки, які порівнюються, отримано з однорідних сукупностей, то значення x_i та y_i взаємозамінні, а отже, ймовірності появи додатних та від'ємних різниць $x_i - y_i$ рівні. Ймовірності появи нульових різниць дорівнюють нулеві, в силу припущення про неперервність розподілу вимірюваної ознаки. (Справді, якщо $X -$ неперервна випадкова величина, значення якої є x_i, $Y -$ неперервна випадкова величина, значення якої є y_i і $y_i -$ певне, отримане в результаті вимірювання значення , то

$$P(x_i - y_i = 0) = P(x_i = y_i) = P(X = y_i) = 0,$$

бо ймовірність набути неперервною випадковою величиною X певного значення y_i дорівнює нулеві.)

Таким чином,

$$P(x_i - y_i > 0) = P(x_i - y_i < 0) = 1/2, \quad i = 1, 2, ..., l,$$

де $l -$ число ненульових різниць, $l \leq n$. Нульові різниці можуть з'явитися через випадкові похибки або помилки заокруглень і тому пари, що їм відповідають, вилучаються з розгляду.

Статистикою критерію знаків є число знаків «+» або «−» в послідовності знаків різниць парних вибірок $(x_i; y_i)$, $i = 1, 2, ..., l$. Для визначеності можна домовитися брати до уваги, наприклад, число знаків «+».

За умови, що нульова гіпотеза H_0 істинна, а дослідні пари чисел $(x_i; y_i)$, отже і знаки різниць $x_i - y_i$, *незалежні* , то число знаків «+» має *біноміальний розподіл* з параметрами $p = 1/2$ та l. Тому задача зводиться до перевірки нульової гіпотези $H_0 : p = 1/2$ за однієї з альтернативних гіпотез:

$$H_1 : p > 1/2 \text{ або } H_1 : p < 1/2 \text{ або } H_1 : p \neq 1/2.$$

Нехай $r -$ отримане число знаків «+», а $\alpha -$ заданий рівень значущості. Гіпотезу H_0 відхиляють, якщо при $H_1 : p > 1/2$ виконується нерівність:

$$\sum_{i=r+1}^{l} C_l^i \left(1/2\right)^l < \alpha ; \quad (1)$$

або при $H_1 : p < 1/2$ виконується нерівність:

$$\sum_{i=0}^{r} C_l^i \left(1/2\right)^l < \alpha ; \quad (2)$$

або при $H_1 : p \neq 1/2$ виконується одна з нерівностей:

$$\sum_{i=r+1}^{l} C_l^i \left(1/2\right)^l < \alpha / 2 \text{ чи } \sum_{i=0}^{r} C_l^i \left(1/2\right)^l < \alpha / 2. \quad (3)$$

Якщо за відповідних альтернативних гіпотез нерівності (1) – (3) не виконуються, то гіпотеза H_0 не суперечить результатам спостережень і приймається на рівні значущості α .

Часто перевірку гіпотези $H_0 : p = 1/2$ проводять, використовуючи відому статистику Фішера F . Гіпотеза H_0 відхиляється, якщо при $H_1 : p > 1/2$

виконується нерівність:

$$F_{досл} = \frac{r}{l - r + 1} \geq F_{кр}(\alpha; k_1, k_2) , \quad (4)$$

де $k_1 = 2(l - r + 1)$, $k_2 = 2r$, $F_{кр}(\alpha; k_1, k_2)$ – критична точка розподілу Фішера (див. [1, с. 147]; або при $H_1 : p < 1/2$ виконується нерівність:

$$F_{досл} = \frac{l - r}{r + 1} \geq F_{кр}(\alpha; k_1, k_2) , \quad (5)$$

де $k_1 = 2(r + 1)$, $k_2 = 2(l - r)$; або при $H_1 : p \neq \frac{1}{2}$ виконується одна з нерівностей:

$F_{досл} = \dfrac{r}{l - r + 1} \geq F_{кр}(\alpha / 2; k_1, k_2)$ чи $F_{досл} = \dfrac{l - r}{r + 1} \geq F_{кр}(\alpha / 2; k_1, k_2)$.

Приклад. Нехай у результаті дослідження за двома методиками отримано дві пов'язані (однакових обсягів) вибірки:

x_i	42	65	60	53	75	85	70	45	55	44
y_i	43	68	62	52	76	80	70	47	53	45

Використовуючи критерій знаків, взявши рівень значущості $\alpha = 0{,}1$, покажемо, що ці дві вибірки є однорідними.

Послідовність знаків різниць $x_i - y_i$ має вигляд:

$- , - , - , + , - , + , 0 , - , + , -$.

Число ненульових різниць $l = 9$, число додатних різниць $r = 3$. Перевіряємо гіпотезу про те, що відмінності в отриманих результатах викликані випадковими помилками, тобто гіпотезу $H_0 : p = 1/2$. Альтернативна гіпотеза полягає в тому, що результати другої методики мають додатне зміщення; іншими словами, ймовірність появи додатних різниць повинна бути меншою за $1/2$. Отже, альтернативна гіпотеза має вигляд: $H_1 : p < 1/2$.

Для перевірки гіпотези $H_0 : p = 1/2$ використовуємо нерівність (5). Насамперед, маємо:

$k_1 = 2 \cdot (3 + 1) = 8$, $k_2 = 2 \cdot (9 - 3) = 12$,

$F_{досл} = \dfrac{9 - 3}{3 + 1} = 1{,}5$.

Оскільки за таблицею [3, с. 354] критичних точок статистики Фішера $F_{кр}(0{,}1; 8; 12) = 2{,}24$, то гіпотеза H_0 не суперечить результатам спостережень. Треба вважати, що відмінності у результатах цих двох вибірок спричинені випадковими помилками, тому генеральні сукупності X, Y є однорідними.

2. *Критерій Вілкоксона*. Нехай на однорідність потрібно перевірити дві незалежні вибірки різних обсягів. Тобто, нехай X та Y – неперервні випадкові величини , а

$$x_1, x_2, ..., x_{n_1} \quad (6)$$

та

$$y_1, y_2, ..., y_{n_2} \quad (7)$$

їхні незалежні вибірки обсягів відповідно n_1 та n_2 .

Перевірку гіпотези про однорідність вибірок (6) та (7) можна провести за критерієм Вілкоксона (W - критерієм).

Нульова гіпотеза полягає у тому, що при всіх значеннях аргументу (тут аргумент завжди позначаємо через x) функції розподілу $F_1(x)$ (для випадкової величини X) та $F_2(x)$ (для випадкової величини Y) рівні між собою: $F_1(x) = F_2(x)$.

Альтернативні гіпотези мають вигляд:

$F_1(x) \neq F_2(x)$, $F_1(x) < F_2(x)$, $F_1(x) > F_2(x)$.

Відразу зауважимо, що приймання альтернативної гіпотези H_1 : $F_1(x) < F_2(x)$ означає, що $X > Y$, з огляду на зростання інтегральної функції розподілу. Аналогічно, справедливість альтернативної гіпотези H_1 :

$F_1(x) > F_2(x)$, означає, що $X < Y$.

Далі в наших міркуваннях вважаємо, що у вибірках (6) та (7) $n_1 \leq n_2$.

1-й випадок, коли обсяги обох вибірок не перевищують 25.

Для того, щоб на рівні значущості α перевірити нульову гіпотезу $H_0 : F_1(x) = F_2(x)$ про однорідність двох незалежних вибірок (6) та (7) обсягів n_1 та n_2 ($n_1 \leq n_2$) за альтернативної гіпотези $H_1 : F_1(x) \neq F_2(x)$, треба діяти за таким правилом:

1) записати обидві вибірки у вигляді одного варіаційного ряду (у зростаючому порядку) і знайти для нього $w_{досл}$ – суму порядкових номерів варіант першої вибірки (вибірки меншого обсягу);

2) за таблицею значень критерію Вілкоксона знайти ліву критичну точку критерію W $w_{ліва\ кр}(Q, n_1, n_2)$, де $Q = \alpha / 2$;

3) за формулою $w_{права\ кр} = (n_1 + n_2 + 1)n_1 - w_{ліва\ кр}$ знайти праву критичну точку.

Якщо $w_{ліва\ кр}$ < $w_{досл}$ < $w_{права\ кр}$, то немає підстав відхиляти нульову гіпотезу. Якщо $w_{досл}$ < $w_{ліва\ кр}$ або $w_{досл}$ > $w_{права\ кр}$, то нульову гіпотезу відхиляють.

Зауваження 1. За альтернативної гіпотези: $F_1(x) > F_2(x)$ потрібно за таблицею знайти ліву критичну точку $w_{ліва\ кр}(Q,n_1,n_2)$, де $Q=\alpha$. Тоді, якщо $w_{досл}$ > $w_{ліва\ кр}$, то немає підстав відхиляти нульову гіпотезу. Якщо $w_{досл}$ < $w_{ліва\ кр}$, то нульову гіпотезу відхиляють.

За альтернативної гіпотези $F_1(x)<F_2(x)$ потрібно знайти праву критичну точку $w_{права\ кр}=(n_1+n_2+1)n_1-w_{ліва\ кр}$, де $Q=\alpha$. Тоді, якщо $w_{досл}$ < $w_{права\ кр}$, то немає підстав відхиляти нульову гіпотезу. Якщо $w_{досл}$ > $w_{права\ кр}$, то нульову гіпотезу відхиляють.

Зауваження 2. Якщо декілька варіант однієї вибірки однакові, то у спільному варіаційному ряду їм приписують порядкові номери так, наче б то ці варіанти є різними. Якщо ж збігаються варіанти різних вибірок, то всім їм приписують однаковий порядковий номер, який дорівнює середньому арифметичному порядкових номерів, який мали б ці варіанти до збігу.

2-й випадок, коли обсяг хоча б однієї з двох вибірок перевищує 25.

Для того, щоб на рівні значущості α перевірити нульову гіпотезу H_0:

$F_1(x)=F_2(x)$ про однорідність двох незалежних вибірок (6) та (7) обсягів n_1 та n_2 ($n_1 \le n_2$) за альтернативної гіпотези H_1: $F_1(x)\ne F_2(x)$, треба діяти за такою схемою:

1) з рівності: $\Phi(z_{кр})=(1-\alpha)/2$ за таблицею значень функції Лапласа $\Phi(z)$ знайти число $z_{кр}$;

2) ліву критичну точку знайти з рівності:

$$w_{ліва\ кр}(Q,n_1,n_2)=\left[\frac{(n_1+n_2+1)\cdot n_1-1}{2}-z_{кр}\sqrt{\frac{n_1n_2(n_1+n_2+1)}{12}}\right], \quad (8)$$

де $Q=\alpha/2$; $[a]$ – ціла частина числа a;

3) за формулою $w_{права\ кр}=(n_1+n_2+1)n_1-w_{ліва\ кр}$ знайти праву критичну точку;

Якщо $w_{ліва\ кр}$ < $w_{досл}$ < $w_{права\ кр}$, то немає підстав відхиляти нульову гіпотезу. Якщо $w_{досл}$ < $w_{ліва\ кр}$ або $w_{досл}$ > $w_{права\ кр}$, то нульову гіпотезу відхиляють.

За альтернативної гіпотези $F_1(x)<F_2(x)$ або $F_1(x)>F_2(x)$:

1) число $z_{кр}$ знаходимо з рівності: $\Phi(z_{кр})=(1-2\alpha)/2$;

2) прийнявши $Q=\alpha$, з рівності (8) знаходимо ліву критичну точку;

3) за формулою: $w_{права\ кр}=(n_1+n_2+1)n_1-w_{ліва\ кр}$ знаходимо праву критичну точку;

Якщо $w_{ліва\ кр}$ < $w_{досл}$ < $w_{права\ кр}$, то немає підстав відхиляти нульову гіпотезу. Якщо ж $w_{досл}$ < $w_{ліва\ кр}$ або $w_{досл}$ > $w_{права\ кр}$, то нульову гіпотезу відхиляємо.

Описаний критерій був, наприклад, застосований нами в роботі [2, с. 70] для перевірки статистичної гіпотези про однорідність експериментальної та контрольної груп студентів при вивченні ряду питань, пов’язаних з комунікативною компетентністю майбутніх менеджерів-аграріїв. Опишемо один з епізодів цієї роботи.

До проведення експерименту було залучено студентів контрольної та експериментальної груп (№1 та №2). Ми брали до уваги рейтингові оцінки, за якими студенти вступали до університету.

Для групи №1, обсягом $n_1=78$, така вибірка мала вигляд:

707, 703, 702, 700, 699, 697, 691, 687, 687, 686, 684, 678, 673, 672, 672, 671, 670, 670, 668, 664, 662, 658, 657, 654, 649, 646, 646, 645, 645, 643, 641, 641, 640, 640, 639, 638, 638, 637, 637, 632, 631, 629, 629, 628, 626, 624, 623, 623, 622, 621, 620, 615, 614, 608, 608, 607, 606, 605, 603, 602, 601, 601, 599, 597, 597, 597, 596, 595, 584, 584, 582, 578, 572, 569, 568, 550, 548, 540.

Середнє значення цієї вибірки $\bar{x}_1=632{,}96$, вибіркове середнє квадратичне відхилення $\sigma_1=39{,}46$, $\Sigma x_1=49371$, $n_1=78$.

Для групи №2 обсягом $n_2=74$ відповідна вибірка мала вигляд:

720, 708, 707, 692, 689, 687, 684, 684, 683, 680, 679, 678, 675, 672, 670, 669, 668, 665, 660, 659, 656, 655, 652, 651, 650, 649, 648, 645, 645, 645, 644, 644, 643, 642, 642, 642, 641, 640, 640, 639, 638, 638, 637, 634, 633, 629, 629, 627, 625, 624, 624, 622, 620, 617, 616, 614, 612, 610, 606, 604, 603, 601, 600, 598, 597, 592, 587, 585, 583, 574, 570, 565, 547, 544.

Середнє значення цієї вибірки $\bar{x}_2=637{,}12$, вибіркове середнє квадратичне відхилення $\sigma_2=36{,}74$, $n_2=74$, $\Sigma x_2=47147$.

Отже, *середні значення та середні квадратичні відхилення обох вибірок є відповідно досить близькими числами*.

Ми висловлюємо *основну гіпотезу*: ці дві вибірки однорідні, тобто мають однаковий розподіл.

Альтернативна гіпотеза: ці дві вибірки не є однорідними.

Для перевірки правильності основної гіпотези вибираємо рівень значущості $\alpha=0{,}05$.

Для отриманих вибірок *об’єднаний варіаційний ряд* має вигляд:

720, ***708***, 707, ***707***, 703, 702, 700, 699, 697, ***692***, 691, ***689***, ***687***, 687, 687, 686, 684, ***684***, ***684***, ***683***, ***680***, ***679***, ***678***, 678, ***675***, 673, 672, ***672***, 672, 671, 670, 670, ***670***, ***669***, ***668***, 668, ***665***, 664, 662, ***660***, ***659***, 658, 657, ***656***, ***655***, 654, ***652***, ***651***, ***650***, ***649***, 649, ***648***, 646, 646, 645, 645, ***645***, ***645***, ***645***, ***644***, ***644***, ***643***, 643, ***642***, ***642***, ***642***, ***641***, 641, 641, 640, 640, ***640***, ***640***, ***639***, 639, 638, 638, ***638***, ***638***, ***637***, 637, 637, ***634***, ***633***, 632, 631, 629, ***629***, ***629***, 629, 628, ***627***, 626, ***625***, ***624***, ***624***, 624, 623, 623, 622, ***622***, 621, 620, ***620***, ***617***, ***616***, 615, 614, ***614***, ***612***, ***610***, 608, 608, 607, ***606***, 606, 605, ***604***, ***603***, 603, 602, 601, 601, ***601***, ***600***, 599, ***598***, ***597***, 597, 597, 597, 596, 595, ***592***, ***587***, ***585***, 584, 584, ***583***, 582, 578, ***574***, 572, ***570***, 569, 568, ***565***, 550, 548, ***547***, ***544***, 540.

Сума порядкових номерів варіант вибірки для групи №1 дорівнює **5777**, а для групи №2 ця сума дорівнює **5785**. Оскільки у даній ситуації обсяги вибірок $74 = n_1 \le n_2 = 78$, то треба взяти $w_{досл} = 5785$.

При $\alpha = 0,05$ з рівності $\Phi(z_{кр}) = (1-\alpha)/2 = (1-0,05)/2 = 0,475$

за таблицею значень функції Лапласа знаходимо , що $z_{кр} = 1,96$.

Знаходимо межі критичної області для величини w. Оскільки $n_1 = 74$, $n_2 = 78$, $w_{досл} = 5785$, $\alpha = 0,05$, то

$$w_{ліва\ кр}(Q, n_1, n_2) = \left[\frac{(n_1+n_2+1)\cdot n_1 - 1}{2} - z_{кр}\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}} \right] =$$

$$= \left[\frac{(74+78+1)\cdot 74 - 1}{2} - 1,96\cdot\sqrt{\frac{74\cdot 78\cdot(74+78+1)}{12}} \right] =$$

$$= [5660,5 - 532,709] = 5128,$$

де $[a]-$ціла частина числа a. Звідси отримуємо праву межу критичної області

$w_{права\ кр} = (n_1+n_2+1)n_1 - w_{ліва\ кр} = (75+78+1)74 - 5128 = 6194$.

Оскільки $5128 < 5785 < 6194$, тобто $w_{ліва\ кр} <$ $w_{досл}$ $<$ $w_{права\ кр}$, то основна гіпотеза про однорідність вибірок приймається. Отже, групи №1 та №2 є однорідними.

3. *Критерій серій*. Для перевірки *нульової гіпотези* H_0 *про випадковість та незалежність елементів вибірки* застосовують критерій серій.

Нехай $x_1, x_2, ..., x_n$ – вибірка результатів спостережень, $\bar{h}_X -$ вибіркова медіана, визначена за цими даними. Кожному елементу вибірки ставимо у відповідність знак «$+$» або «$-$» залежно від того, більшим чи меншим є цей елемент за медіану (нульові значення до уваги не беруться). Позначимо через n_1 число знаків «$+$», а через n_2 – число знаків «$-$». Отже, всій вибірці поставлено у відповідність деякий набір (послідовність) знаків.

Серією у цьому наборі називається будь-яка послідовність, яка складається з однакових знаків і обмежена однаковим знаками з обох боків або знаходиться наприкінці чи на початку цього набору. Наприклад, в наборі

$$+-+++-----++++$$

міститься 5 серій: $(+), (-), (+++), (-----), (++)$, $n_1 = 8, n_2 = 6$.

Статистикою критерію серій є число N. Критична область визначається нерівностями $N \le N_1$ і $N \ge N_2$. Значення меж N_1 та N_2 критичної області на заданому рівні значущості наводяться у відповідних таблицях критерію серій [3].

Перевіримо, наприклад, чи можна вважати послідовність чисел 31, 39, 40, 45, 27, 28, 35, 55, 21, 33, 42, 36 випадковою, якщо за рівень значущості прийняти $\alpha = 0,05$.

Запишемо вибірку у вигляді варіаційного ряду: 21, 27, 28, 31, 33, 35, 36, 39, 40, 42, 45, 55. Тоді вибіркова медіана $\bar{h}_B = \frac{35+36}{2} = 35,5$. Отриманому ряду спостережень відповідає такий набір знаків:

$$-, +, +, +, -, -, -, +, -, -, +, +,$$

звідки $n_1 = 6,\ n_2 = 6$, число серій $N = 6$. За таблицею значень критерію серій (3 с.375), на рівні значущості $\alpha = 0,05$, знаходимо $N_1 = 3,\ N_2 = 11$. Оскільки $N_1 < N < N_2$ $(3 < 6 < 11)$, то нульова гіпотеза H_0 приймається: отримані значення можна вважати випадковими.

За великих вибірок, коли або $n_1 > 20$ або $n_2 > 20$ або $n_1 > 20$ і $n_2 > 20$, для перевірки гіпотези H_0 можна використати статистику U, дослідне значення якої обчислюється за формулою:

$$u_{досл} = \left(N - \frac{2n_1n_2}{n_1+n_2} \right) - \frac{1}{2} \Bigg/ \sqrt{\frac{2n_1n_2[2n_1n_2 - (n_1+n_2)]}{(n_1+n_2)^2(n_1+n_2-1)}}.$$

За умови істинності гіпотези H_0 статистика U має приблизно нормальний розподіл з математичним сподіванням 0 та дисперсією 1. У такому разі критична область визначається нерівностями: $u_{досл} \le u_{ліва\ кр}$ або $u_{досл} \ge u_{права\ кр}$, де $u_{права\ кр} = u_{(1-\alpha)/2}$, $u_{ліва\ кр} = -u_{(1-\alpha)/2}$.

Зауваження. Ми навели найпростіші непараметричні критерії, проте є й інші непараметричні критерії, наприклад, критерії Колмогорова, Уайта, які теж широко використовуються у різних галузях науки.

Література:

1. Гмурман В. Е. Теория вероятностей и математическая статистика: учебное пособие для вузов. – М.: Высшая школа, 2003. – 479 с.

2. Краєвська О. Д. Визначення однорідності груп при дослідженні процесу формування комунікативної компетентності майбутніх менеджерів-аграріїв /

О. Д. Краєвська, Д. А. Найко // Економіка. Фінанси. Менеджмент: актуальні питання науки і практики. / Збірник наукових праць Вінницького національного аграрного університету. – 2015. – №1. – С. 70 – 81.

3. Найко Д. А. Теорія ймовірностей та математична статистика: навчальний посібник. – Вінниця: ВНАУ, 2016. – 377с.