

Методичні рекомендації

У студентів першого та четвертого курсів було досліджено рівень самооцінки. Перевірити гіпотезу про наявність відмінностей у самооцінці студентів 1-го та 4-го курсів. Результати тестування наведено у таблиці.

Таблиця 1 – Вихідні дані

№	Курс 1	Курс 4
1	0,12	0,84
2	0,47	0,63
3	0,38	0,23
4	0,13	0,95
5	0,19	0,98
6	0,63	0,73
7	0,56	0,95
8	0,55	0,96
9	0,58	0,95
10	0,45	0,54
11	0,35	0,28
12	0,73	0,23
13	0,15	
14	0,35	

1. Порівняння незалежних вибірок.

Провести перевірку гіпотези за критерієм Стьюдента.

Сформулюємо гіпотези:

H_0 : Відмінності між показниками самооцінки студентів 1 курсу та 4 курси недостовірні;

H_1 : Відмінності між показниками самооцінки студентів 1 курсу та 4 курси достовірні.

В даному випадку перевірка достовірності відмінностей між показниками самооцінки студентів 1 та 4 курсів передбачає порівняння середніх та дисперсій.

Спочатку перевіримо розподіли двох вибірок на близькість до нормального розподілу. Для цього скористаємося критерієм асиметрії та ексцесу. Для обчислення емпіричних значень асиметрії та ексцесу застосовуються функції SKEW і KURT, а критичні обчислювалися по формулам (3.9)–(3.10) [див. посібник]. Розрахунки наведені на рисунках 3.18–3.19. Оскільки емпіричні значення асиметрії та ексцесу в обох групах (1 курс і 4 курс) менші за критичні, то розподіл в обох групах близький до нормального.

	A	B	C	D	E	F	G	H	I	J	K
1	№	Курс 1	Курс 4				Перевірка розподілу на близькість до нормального				
2	1	0,12	0,84				Курс 1		Курс 4		
3	2	0,47	0,63				Аемп	-0,11	-0,65		
4	3	0,38	0,23				Еемп	-1,14	-1,29		
5	4	0,13	0,95				n	14	12		
6	5	0,19	0,98				Акр	1,66	1,75		
7	6	0,63	0,73				Екр	3,91	3,88		
8	7	0,56	0,95								
9	8	0,55	0,96								
10	9	0,58	0,95								
11	10	0,45	0,54								
12	11	0,35	0,28								
13	12	0,73	0,23								
14	13	0,15									
15	14	0,35									
16											
17											

Рисунок 1 – Розрахунок показників асиметрії та ексцесу

	A	B	C	D	E	F	G	H	I
1	№	Курс 1	Курс 4				Перевірка розподілу на близькість до нормального		
2	1	0,12	0,84				Курс 1		Курс 4
3	2	0,47	0,63				Аемп	=SKEW(B2:B15)	=SKEW(C2:C13)
4	3	0,38	0,23				Еемп	=KURT(B2:B16)	=KURT(C2:C13)
5	4	0,13	0,95				n	14	12
6	5	0,19	0,98				Акр	=3*SQRT((6*(H5-1))/((H5+1)*(H5+3)))	=3*SQRT((6*(I5-1))/((I5+1)*(I5+3)))
7	6	0,63	0,73				Екр	=5*SQRT((24*H5*(H5-2)*(H5-3))/((H5+1)^2*(H5+3)*(H5+5)))	=5*SQRT((24*I5*(I5-2)*(I5-3))/((I5+1)^2*(I5+3)*(I5+5)))
8	7	0,56	0,95						
9	8	0,55	0,96						
10	9	0,58	0,95						
11	10	0,45	0,54						
12	11	0,35	0,28						
13	12	0,73	0,23						
14	13	0,15							
15	14	0,35							
16									

Рисунок 2 – Розрахунок показників асиметрії та ексцесу

Далі проводимо обчислення середніх та дисперсії. Для цього застосовуються функції AVERAGE і VAR.S (). Після цього по формулі (3.11) [див. посібник] обчислено емпіричне значення критерію Фішера. Критичне значення (знайдене по таблицям) дорівнює 3,05. Оскільки емпіричне значення критерію Фішера менше за критичне, то дисперсії двох вибірок однорідні.

	A	B	C
1	№	Курс 1	Курс 4
2	1	0,12	0,84
3	2	0,47	0,63
4	3	0,38	0,23
5	4	0,13	0,95
6	5	0,19	0,98
7	6	0,63	0,73
8	7	0,56	0,95
9	8	0,55	0,96
10	9	0,58	0,95
11	10	0,45	0,54
12	11	0,35	0,28
13	12	0,73	0,23
14	13	0,15	
15	14	0,35	
16	M	0,40	0,69
17	D	0,04	0,09
18	F	2,31	

	A	B	C
1	№	Курс 1	Курс 4
2	1	0,12	0,84
3	2	0,47	0,63
4	3	0,38	0,23
5	4	0,13	0,95
6	5	0,19	0,98
7	6	0,63	0,73
8	7	0,56	0,95
9	8	0,55	0,96
10	9	0,58	0,95
11	10	0,45	0,54
12	11	0,35	0,28
13	12	0,73	0,23
14	13	0,15	
15	14	0,35	
16	M	=AVERAGE(B2:B15)	=AVERAGE(C2:C13)
17	D	=VAR.S(B2:B15)	=VAR.S(C2:C13)
18	F	=C17/B17	
19			

Рисунок 3 – Розрахунок показників середнє і дисперсія

Емпіричне значення критерію Стюдента обчислюється по формулі (3.12), а критичні значення знайдені за допомогою функції T.INV. Розрахунок представлений на рисунку 4.

	A	B	C		A	B	C
1	№	Курс 1	Курс 4	1	№	Курс 1	Курс 4
2	1	0,12	0,84	2	1	0,12	0
3	2	0,47	0,63	3	2	0,47	0
4	3	0,38	0,23	4	3	0,38	0
5	4	0,13	0,95	5	4	0,13	0
6	5	0,19	0,98	6	5	0,19	0
7	6	0,63	0,73	7	6	0,63	0
8	7	0,56	0,95	8	7	0,56	0
9	8	0,55	0,96	9	8	0,55	0
10	9	0,58	0,95	10	9	0,58	0
11	10	0,45	0,54	11	10	0,45	0
12	11	0,35	0,28	12	11	0,35	0
13	12	0,73	0,23	13	12	0,73	0
14	13	0,15		14	13	0,15	
15	14	0,35		15	14	0,35	
16	M	0,40	0,69	16	M	=AVERAGE(B2:B15)	=AVERAGE(C2:C13)
17	D	0,04	0,09	17	D	=VAR.S(B2:B15)	=VAR.S(C2:C13)
18	F	2,31		18	F	=C17/B17	
19				19			
20	темп	0,52		20	темп	=ABS(B14-C14)/SQRT(((14-1)*B16+(12-1)*C16)/(14+12-2)*(1/14+1/12))	
21	ткр1	1,71		21	ткр1	=abs(T.INV(0,05;14+12-2))	
22	ткр2	2,49		22	ткр2	=abs(T.INV(0,01;14+12-2))	
23				23			

Рисунок 4 – Розрахунок емпіричного та критичних значень критерія Стюдента

Будуємо вісь значущості (рисунок 5). Оскільки емпіричне значення критерія Стьюдента потрапило в зону незначущості, то у нас немає підстав відхиляти нульову гіпотезу.

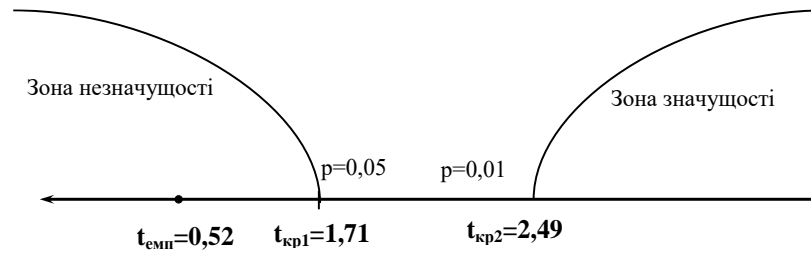


Рисунок 5 – Вісь значущості для критерія Стюдента

Отже, за рівнем самооцінки студенти 1 і 4 курсів не розрізняються.

Провести перевірку гіпотези за критерієм Манна-Вітні.

Для визначення першої та другої вибірок були знайдені максимальне та мінімальне значення в кожній вибірці. Як видно з результатів (рисунок 6) мінімальне і максимальне значення в групі Курс 1 менші ніж в групі Курс 4. Отже, першою вибіркою є Курс 4, а другою – Курс 1.

	A	B	C	D		A	B	C
1	№	Курс 1	Курс 4		1	№	Курс 1	Курс 4
2	1	0,12	0,84		2	1	0,12	0,84
3	2	0,47	0,63		3	2	0,47	0,63
4	3	0,38	0,23		4	3	0,38	0,23
5	4	0,13	0,95		5	4	0,13	0,95
6	5	0,19	0,98		6	5	0,19	0,98
7	6	0,63	0,73		7	6	0,63	0,73
8	7	0,56	0,95		8	7	0,56	0,95
9	8	0,55	0,96		9	8	0,55	0,96
10	9	0,58	0,95		10	9	0,58	0,95
11	10	0,45	0,54		11	10	0,45	0,54
12	11	0,35	0,28		12	11	0,35	0,28
13	12	0,73	0,23		13	12	0,73	0,23
14	13	0,15			14	13	0,15	
15	14	0,35			15	14	0,35	
16					16			
17	min	0,12	0,23		17	min	=МИН(B2:B15)	=МИН(C2:C15)
18	max	0,73	0,98		18	max	=МАКС(B2:B15)	=МАКС(C2:C15)

Рисунок 6 – Визначення першої та другої вибірок

Сформулюємо статистичні гіпотези.

H_0 : Рівень самооцінки студентів четвертого курсу не перевищує рівня самооцінки студентів першого курсу.

H_1 : Рівень самооцінки студентів четвертого курсу перевищує рівень самооцінки студентів першого курсу.

Об'єднаємо дві вибірки в одну, позначивши їх різними кольорами, для розрізнення. Поєднання здійснено завдяки використанню формул (рисунок 7).

Розділимо ранги об'єднаної вибірки на дві вибірки, скориставшись формулами (рисунок 9).

E	F	G	H	I	J
об'єднана вибірка	Ранги			Перша вибірка (Курс 4)	Друга вибірка (Курс 1)
=B2	=РАНГ.СР(E2:\$E\$2:\$E\$27:0)			=F16	=F2
=B3	=РАНГ.СР(E3:\$E\$2:\$E\$27:0)			=F17	=F3
=B4	=РАНГ.СР(E4:\$E\$2:\$E\$27:0)			=F18	=F4
=B5	=РАНГ.СР(E5:\$E\$2:\$E\$27:0)			=F19	=F5
=B6	=РАНГ.СР(E6:\$E\$2:\$E\$27:0)			=F20	=F6
=B7	=РАНГ.СР(E7:\$E\$2:\$E\$27:0)			=F21	=F7
=B8	=РАНГ.СР(E8:\$E\$2:\$E\$27:0)			=F22	=F8
=B9	=РАНГ.СР(E9:\$E\$2:\$E\$27:0)			=F23	=F9
=B10	=РАНГ.СР(E10:\$E\$2:\$E\$27:0)			=F24	=F10
=B11	=РАНГ.СР(E11:\$E\$2:\$E\$27:0)			=F25	=F11
=B12	=РАНГ.СР(E12:\$E\$2:\$E\$27:0)			=F26	=F12
=B13	=РАНГ.СР(E13:\$E\$2:\$E\$27:0)			=F27	=F13
=B14	=РАНГ.СР(E14:\$E\$2:\$E\$27:0)				=F14
=B15	=РАНГ.СР(E15:\$E\$2:\$E\$27:0)				=F15
=C2	=РАНГ.СР(E16:\$E\$2:\$E\$27:0)				
=C3	=РАНГ.СР(E17:\$E\$2:\$E\$27:0)				
=C4	=РАНГ.СР(E18:\$E\$2:\$E\$27:0)				
=C5	=РАНГ.СР(E19:\$E\$2:\$E\$27:0)				
=C6	=РАНГ.СР(E20:\$E\$2:\$E\$27:0)				
=C7	=РАНГ.СР(E21:\$E\$2:\$E\$27:0)				
=C8	=РАНГ.СР(E22:\$E\$2:\$E\$27:0)				
=C9	=РАНГ.СР(E23:\$E\$2:\$E\$27:0)				
=C10	=РАНГ.СР(E24:\$E\$2:\$E\$27:0)				
=C11	=РАНГ.СР(E25:\$E\$2:\$E\$27:0)				
=C12	=РАНГ.СР(E26:\$E\$2:\$E\$27:0)				
=C13	=РАНГ.СР(E27:\$E\$2:\$E\$27:0)				

Рисунок 9 – Вибірки, створені з рангів

Для кожного ряду обчислено суму рангів. Це значення 115 та 236. Обсяг першої вибірки $n_1 = 12$, обсяг другої – $n_2 = 14$; обсяг об'єднаної вибірки – $n = n_1 + n_2 = 26$. Також проведена перевірка суми рангів (рисунок 10).

	H	I	J	K	L
		Перша вибірка (Курс 4)	Друга вибірка (Курс 1)		
1					
2		6	26		
3		9,5	15		
4		21,5	17		
5		4	25		
6		1	23		
7		7,5	9,5		
8		4	12		
9		2	13		
10		4	11		
11		14	16		
12		20	18,5		
13		21,5	7,5		
14			24		
15			18,5	Перевірка	
16	T=	115	236	351	
17	n=	12	14	26	351

	H	I	J	K	L
		Перша вибірка (Курс 4)	Друга вибірка (Курс 1)		
1					
2		=F16	=F2		
3		=F17	=F3		
4		=F18	=F4		
5		=F19	=F5		
6		=F20	=F6		
7		=F21	=F7		
8		=F22	=F8		
9		=F23	=F9		
10		=F24	=F10		
11		=F25	=F11		
12		=F26	=F12		
13		=F27	=F13		
14			=F14		
15			=F15		
16	T=	=СУММ(I2:I15)	=СУММ(J2:J15)	Перевірка	=СУММ(I16:J16)
17	n=	12	14	=СУММ(I17:J17)	=K17*(K17+1)/2

Рисунок 10 – Перевірка суми рангів

Визначимо найбільшу рангову суму: $T_x = 236$, відповідно обсяг вибірки з найбільшою ранговою сумою – $n_x = 14$. На рисунку 3.28 представлено обчислення цих показників.

	Н	І	Ж		Н	І	Ж
		Перша вбірка (Курс 4)	Друга вбірка (Курс 1)			Перша вибірка (Курс 4)	Друга вибірка (Курс 1)
1							
2		6	26		=F16	=F2	
3		9,5	15		=F17	=F3	
4		21,5	17		=F18	=F4	
5		4	25		=F19	=F5	
6		1	23		=F20	=F6	
7		7,5	9,5		=F21	=F7	
8		4	12		=F22	=F8	
9		2	13		=F23	=F9	
10		4	11		=F24	=F10	
11		14	16		=F25	=F11	
12		20	18,5		=F26	=F12	
13		21,5	7,5		=F27	=F13	
14			24			=F14	
15			18,5			=F15	
16	T=	115	236		T=	=СУММ(I2:I15)	=СУММ(J2:J15)
17	n=	12	14		n=	12	14
18							
19	T _x =	236			T _x =	=МАКС(I16:J16)	
20	n _x =	14			n _x =	=ЕСЛИ(I19=I16;I17;J17)	

Рисунок 11 – Обчислення T_x та n_x

Нарешті для обчислення $U_{емп}$ скористалися формулою (1), представленої в алгоритмі (рисунку 12).

15		18,5	Перевірка	15		=F15	Перевірка
16	T=	115	236	16	T=	=СУММ(I2:I15)	=СУММ(J2:J15)
17	n=	12	14	17	n=	12	14
18			26	18			=СУММ(I17:J17)
19	T _x =	236	351	19	T _x =	=МАКС(I16:J16)	=СУММ(I16:J16)
20	n _x =	14	351	20	n _x =	=ЕСЛИ(I19=I16;I17;J17)	=K17*(K17+1)/2
21				21			
22				22			
23	U _{емп} =	37		23	U _{емп} =	=I17*J17+(I20*(I20+1))/2-J16	

Рисунок 12 – Обчислення $U_{емп}$

По таблицям критичних значень критерія U -Манна-Уитні для рівнів значущості $\alpha = 0,05$ й $\alpha = 0,01$ та $n_1 = 12$ й $n_2 = 14$: $U_{1кр} = 51$ та $U_{2кр} = 38$.

Будуємо вісь значущості.

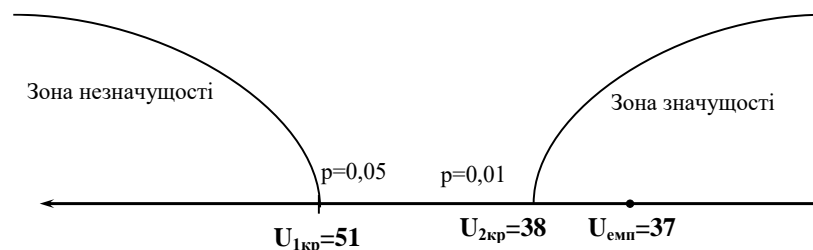


Рисунок 13 – Вісь значущості для критерія Манна-Вітні

$U_{\text{емп}}$ потрапило до зони значущості, тому гіпотеза H_0 відхиляється та приймається H_1 .

Висновок: Згідно з критерієм Манна-Вітні можна стверджувати, що студенти 4 курсу перевершують студентів 1 курсу за рівнем самооцінки.

2. Перевірити гіпотезу про відмінність значень показника, виміряного двічі на одній і тій же вибірці (Вимірювання 1 і Вимірювання 2). Результати тестування наведено у таблиці 2.

Таблиця 2 – Вихідні дані для приклада

№	Вимір 1	Вимір 2
1	6	14
2	11	5
3	12	8
4	8	10
5	5	14
6	10	7
7	7	12
8	6	13
9	3	11
10	9	9
11	4	15
12	5	16

Сформулюємо гіпотези:

H_0 : Значення показника, виміряного двічі на одній і тій же вибірці, не відрізняються;

H_1 : Значення показника, виміряного двічі одній і тій же вибірці, різняться.

Визначається різниця значень ознак. Знайти модуль значення різниці, для чого скористатися математичною функцією **ABS**. Нульовий зсув видалити (комірка **E11**). Відповідно **n=11**.

Визначити ранги за допомогою статистичної функції **РАНГ.СР**.

Визначити суми рангів позитивних та негативних зрушень. Підрахувати кількість позитивних (n^+) та негативних (n^-) у діапазоні **D2:D13**. Для цього скористатися статистичною функцією **СУММЕСЛИ** (рисунок 14, рисунок 15).

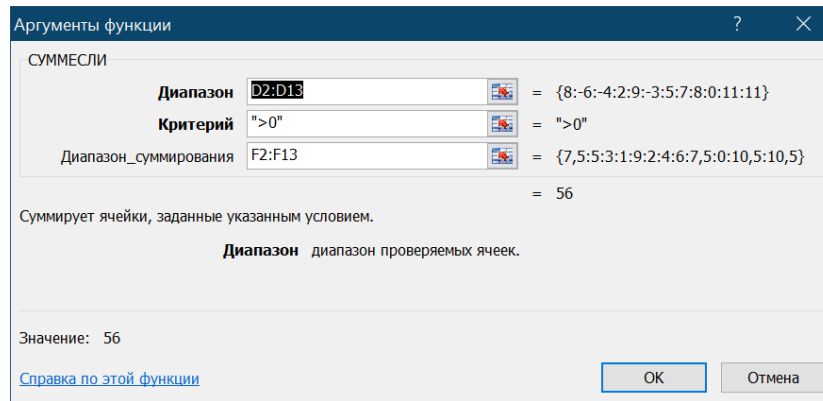


Рисунок 14 – Вікно функції СУММЕСЛИ для обчислення позитивних зрушень

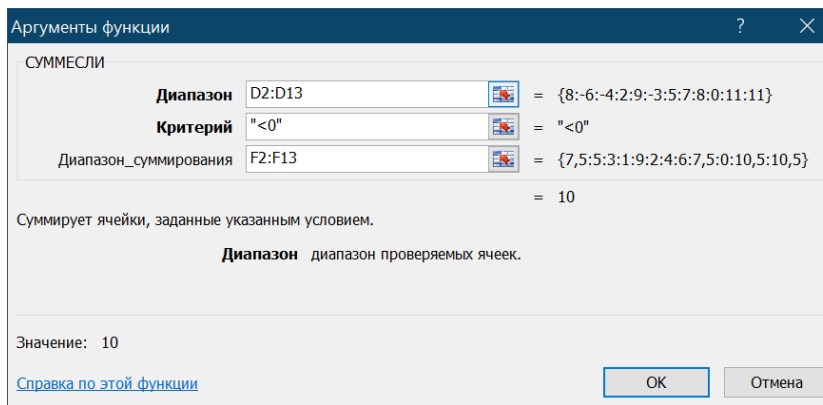


Рисунок 15 – Вікно функції СУММЕСЛИ для обчислення негативних зрушень

Визначити типові й нетипові зсуви за допомогою логічної функції ЕСЛИ (рисунок 16).

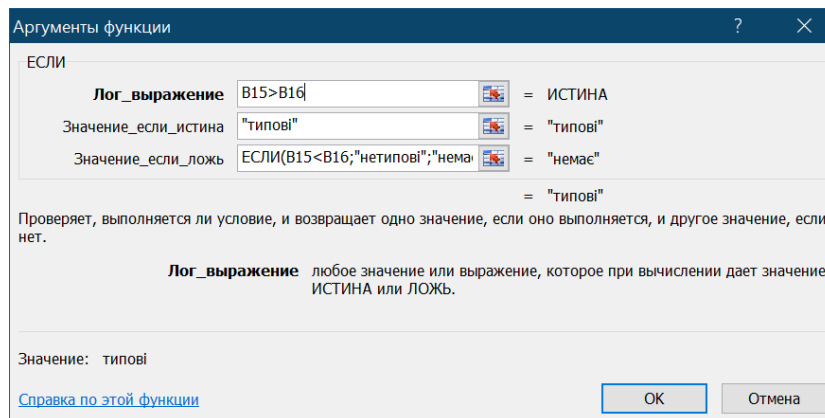


Рисунок 16 – Вікно функції ЕСЛИ для визначення типу зсуву

Нетиповим знаком є "-" і тому емпіричне значення критерію дорівнює сумі рангів нетипових зсувів $T_{\text{емп}} = 10$.

Це можна визначити за допомогою функції ЕСЛИ (рисунок 17).

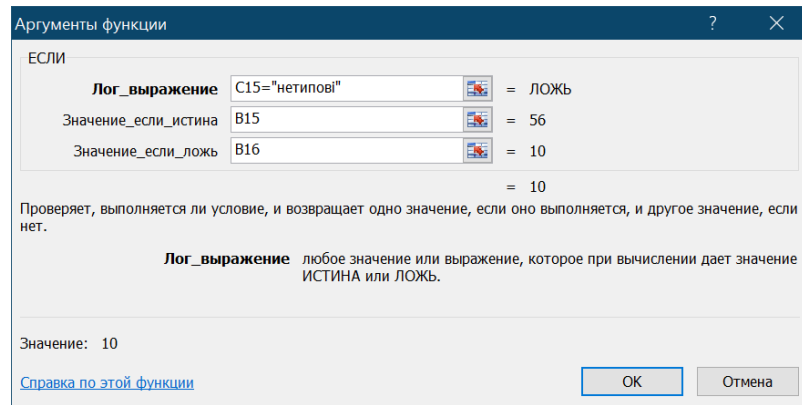


Рисунок 17 – Визначення емпіричного значення критерія Вілкоксона

Розрахунки наведені на рисунках 18 і 19.

	A	B	C	D	E	F
1	№	Вимір 1	Вимір 2	d	d	Ранги d
2	1	6	14	8	8	7,5
3	2	11	5	-6	6	5
4	3	12	8	-4	4	3
5	4	8	10	2	2	1
6	5	5	14	9	9	9
7	6	10	7	-3	3	2
8	7	7	12	5	5	4
9	8	6	13	7	7	6
10	9	3	11	8	8	7,5
11	10	9	9	0		
12	11	4	15	11	11	10,5
13	12	5	16	11	11	10,5
14	Розрахунки					66
15	R ₊ =	56	типові			
16	R ₋ =	10	нетипові			
17	n=	11				
18	Темп=	10				
19						

Рисунок 18 – Результати обчислення емпіричного значення критерія Вілкоксона

	A	B	C	D	E	F
1	№	Вимір 1	Вимір 2	d	d	Ранги d
2	1	6	14	=C2-B2	=ABS(D2)	=РАНГ.СР(E2:\$E\$2:\$E\$13;1)
3	2	11	5	=C3-B3	=ABS(D3)	=РАНГ.СР(E3:\$E\$2:\$E\$13;1)
4	3	12	8	=C4-B4	=ABS(D4)	=РАНГ.СР(E4:\$E\$2:\$E\$13;1)
5	4	8	10	=C5-B5	=ABS(D5)	=РАНГ.СР(E5:\$E\$2:\$E\$13;1)
6	5	5	14	=C6-B6	=ABS(D6)	=РАНГ.СР(E6:\$E\$2:\$E\$13;1)
7	6	10	7	=C7-B7	=ABS(D7)	=РАНГ.СР(E7:\$E\$2:\$E\$13;1)
8	7	7	12	=C8-B8	=ABS(D8)	=РАНГ.СР(E8:\$E\$2:\$E\$13;1)
9	8	6	13	=C9-B9	=ABS(D9)	=РАНГ.СР(E9:\$E\$2:\$E\$13;1)
10	9	3	11	=C10-B10	=ABS(D10)	=РАНГ.СР(E10:\$E\$2:\$E\$13;1)
11	10	9	9	=C11-B11		
12	11	4	15	=C12-B12	=ABS(D12)	=РАНГ.СР(E12:\$E\$2:\$E\$13;1)
13	12	5	16	=C13-B13	=ABS(D13)	=РАНГ.СР(E13:\$E\$2:\$E\$13;1)
14	Розрахунки					=СУММ(F2:F13)
15	R ₊ =	=СУММЕСЛИ(D2:D13;">0";F2:F13)	=ЕСЛИ(B15>B16;"типові";ЕСЛИ(B15<B16;"нетипові";"немає"))			
16	R ₋ =	=СУММЕСЛИ(D2:D13;"<0";F2:F13)	=ЕСЛИ(B16>B15;"типові";ЕСЛИ(B16<B15;"нетипові";"немає"))			
17	n=	11				
18	Темп=	=ЕСЛИ(C15="нетипові";B15;B16)				
19						

Рисунок 19 – Формули обчислення емпіричного значення критерія Вілкоксона

По таблиці критичних значень знайдені критичні значення для $n = 11$:

$$T_{кр} = \begin{cases} 13, \text{ для } \alpha \leq 0,05 \\ 7 \text{ для } \alpha \leq 0,01 \end{cases}$$

Побудуємо вісь значущості (рисунок 20)

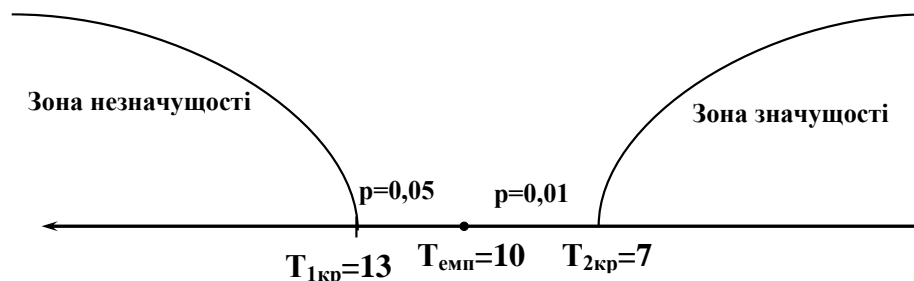


Рисунок 20 – Вісь значущості для критерія Вілкоксона

$T_{емп}$ знаходиться в зоні невизначеності, тому необхідно скористатися іншим критерієм або збільшити вибірку. Хоча можна просто порівняти $T_{емп}$ з $T_{1кр}$. Оскільки $T_{емп} < T_{1кр}$, то на рівні значущості $\alpha \leq 0,05$ гіпотеза H_0 відхиляється та приймається альтернативна гіпотеза H_1 .