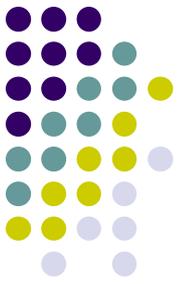


U-критерий Манна – Уитни



Используется для оценки различий между двумя несвязными выборками по уровню количественно измеренного признака

Позволяет выявлять различия в значении признака между малыми выборками

Гипотезы (для случая, когда уровень первой группы на первый взгляд выше)

- H_0 : «Уровень признака в группе 2 не ниже уровня признака в группе 1»
- H_1 : «Уровень признака в группе 2 ниже уровня признака в группе 1»

U-критерий Манна – Уитни



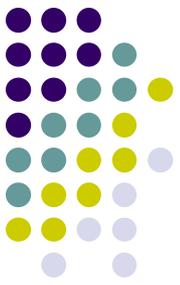
Алгоритм U-теста:

- 1) обе выборки соединяются в единую выборку. При этом запоминается, к какой выборке относится каждый элемент массива
- 2) общая выборка ранжируется по возрастанию

Правила ранжирования:

- меньшему значению начисляется меньший ранг
- наименьшему значению начисляется ранг 1
- наибольшему значению начисляется ранг, соответствующий количеству ранжируемых значений $n_1 + n_2$

U-критерий Манна – Уитни



Правила ранжирования:

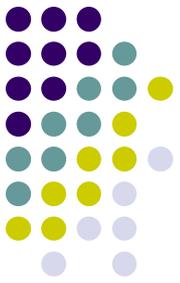
- в случае, если несколько значений равны, им начисляется ранг, представляющий собой среднее значение тех рангов, которые они получили бы, если бы они не были равны
- общая сумма рангов должна совпадать с расчетной,

которая определяется по формуле $\sum R_i = \frac{N \cdot (N + 1)}{2}$

где R_i – ранги,

$N = n_1 + n_2$ – общее количество ранжируемых наблюдений

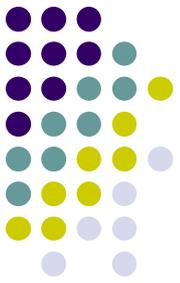
U-критерий Манна – Уитни



Алгоритм U-теста:

- 3) разбить единую выборку на прежние две выборки
- 4) подсчитать сумму рангов отдельно по каждой выборке.
Проверить, совпадает ли общая сумма рангов с расчетной

U-критерий Манна – Уитни



Алгоритм U-теста:

5) найти эмпирическое значение критерия. Для этого подсчитать:

$$U_1 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - R_1,$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - R_2,$$

$$U_{\text{эмп}} = \min \{U_1, U_2\},$$

где n_1 – объем первой выборки,

n_2 – объем второй выборки,

R_1 и R_2 – ранговые суммы групп

U-критерий Манна – Уитни



Алгоритм U-теста:

5) найти критическое значение критерия $U_{крит}$ по таблицам

Таблица II

Критические значения критерия U Манна-Уитни для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$ (по Гублеру Е.В., Генкину А.А., 1973)

Различия между двумя выборками можно считать значимыми ($\rho < 0,05$), если $U_{наб}$ ниже или равен $U_{0,05}$, и тем более достоверными ($\rho < 0,01$), если $U_{наб}$ ниже или равен $U_{0,01}$.

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	-	0																	
4	-	0	1																
5	0	1	2	4															
6	0	2	3	5	7														
7	0	2	4	6	8	11													
8	1	3	5	8	10	13	15												
9	1	4	7	11	14	17	20	24	27										
10	1	4	7	11	14	17	20	24	27										
11	1	5	8	12	16	19	23	27	31	34									
12	2	5	9	13	17	21	26	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	-	-	0	1															
6	-	-	1	2	3														
7	-	0	1	3	4	6													
8	-	0	2	4	6	7	9												
9	-	1	3	5	7	9	11	14											
10	-	1	3	6	8	11	13	16	19										
11	-	1	4	7	9	12	15	18	22	25									
12	-	2	5	8	11	14	17	21	24	28	31								
13	0	2	5	9	12	16	20	23	27	31	35	39							
14	0	2	6	10	13	17	22	26	30	34	38	43	47						
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56					
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77			
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88		
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Таблица II. Продолжение

n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
$\rho \leq 0,05$																		
21	19	26	34	41	49	57	65	73	81	89	97	105	113	121	130	138	146	154
22	20	28	36	44	52	60	69	77	85	94	102	111	119	128	136	145	154	162
23	21	29	37	46	55	63	72	81	90	99	107	116	125	134	143	152	161	170
24	22	31	39	48	57	66	75	85	94	103	113	122	131	141	150	160	169	179
25	23	32	41	50	60	69	79	89	98	108	118	128	137	147	157	167	177	187
26	24	33	43	53	62	72	82	93	103	113	123	133	143	154	164	174	185	195
27	25	35	45	55	65	75	86	96	107	118	128	139	150	160	171	182	193	203
28	26	36	47	57	68	79	89	100	111	122	133	144	156	167	178	189	200	212
29	27	38	48	59	70	82	93	104	116	127	139	150	162	173	185	196	208	220
30	28	39	50	62	73	85	96	108	120	132	144	156	168	180	192	204	216	228
31	29	41	52	64	76	88	100	112	124	137	149	161	174	186	199	211	224	236
32	30	42	54	66	78	91	103	116	129	141	154	167	180	193	206	219	232	245
33	31	43	56	68	81	94	107	120	133	146	159	173	186	199	213	226	239	253
34	32	45	58	71	84	97	110	124	137	151	164	178	192	206	219	233	247	261
35	33	46	59	73	86	100	114	128	142	156	170	184	198	212	226	241	255	269
36	35	48	61	75	89	103	117	132	146	160	175	189	204	219	233	248	263	278
37	36	49	63	77	92	106	121	135	150	165	180	195	210	225	240	255	271	286
38	37	51	65	79	94	109	124	139	155	170	185	201	216	232	247	263	278	294
39	38	52	67	82	97	112	128	143	159	175	190	206	222	238	254	270	286	302
40	39	53	69	84	100	115	131	147	163	179	196	212	228	245	261	278	294	311
$\rho \leq 0,01$																		
21	10	16	22	29	35	42	49	56	63	70	77	84	91	98	105	113	120	127
22	10	17	23	30	37	45	52	59	66	74	81	89	96	104	111	119	127	134
23	11	18	25	32	39	47	55	62	70	78	86	94	102	109	117	125	133	141
24	12	19	26	34	42	49	57	66	74	82	90	98	107	115	123	132	140	149
25	12	20	27	35	44	52	60	69	77	86	95	103	112	121	130	138	147	156
26	13	21	29	37	46	54	63	72	81	90	99	108	117	126	136	145	154	163
27	14	22	30	39	48	57	66	75	85	94	103	113	122	132	142	151	161	171
28	14	23	32	41	50	59	69	78	88	98	108	118	128	138	148	158	168	178
29	15	24	33	42	52	62	72	82	92	102	112	123	133	143	154	164	175	185
30	15	25	34	44	54	64	75	85	95	106	117	127	138	149	160	171	182	192
31	16	26	35	46	56	67	77	88	99	110	121	132	143	155	166	177	188	200
32	17	27	37	47	58	69	80	91	103	114	126	137	149	160	172	184	195	207
33	17	27	37	47	58	69	80	91	103	114	126	137	149	160	172	184	195	207
34	18	28	38	49	60	72	83	95	106	118	130	142	154	166	178	190	202	214
35	19	30	41	53	64	77	89	101	114	126	139	152	164	177	190	203	216	229
36	19	31	42	54	67	79	92	104	117	130	143	156	170	183	196	210	223	236
37	20	32	44	56	69	81	95	108	121	134	148	161	175	189	202	216	230	244
38	21	33	45	58	71	84	97	111	125	138	152	166	180	194	208	223	237	251
39	21	34	46	59	73	86	100	114	128	142	157	171	185	200	214	229	244	258
40	22	35	48	61	75	89	103	117	132	146	161	176	191	206	221	236	251	266

U-критерий Манна – Уитни



Алгоритм U-теста:

6) сравнить эмпирическое и критическое значение критерия

Если $U_{эмп} > U_{крит}$, то нулевая гипотеза принимается

Если $U_{эмп} \leq U_{крит}$, то нулевая гипотеза отклоняется

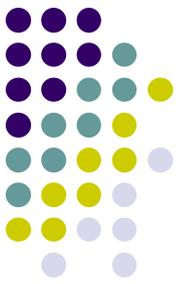
U-критерий Манна – Уитни



Ограничения применимости критерия

- в каждой из выборок должно быть не менее 3 значений признака. Допускается, чтобы в одной выборке было два значения, но во второй тогда не менее пяти.
- в каждой выборке должно быть не более 60 наблюдений (ограничение не является строгим для случая компьютерной обработки)

U-критерий Манна – Уитни

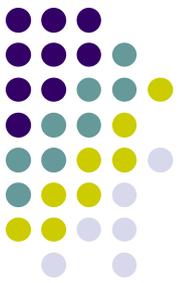


Пример

Результаты обследования студентов физического и психологического факультетов Ленинградского университета:

Студенты - физики		Студенты-психологи	
Код имени испытуемого	Показатель невербального интеллекта	Код имени испытуемого	Показатель невербального интеллекта
1. И.А.	111	1. Н.Т.	113
2. К.А.	104	2. О.В.	107
3. К.Е.	107	3. Е.В.	123
4. П.А.	90	4. Ф.О.	122
5. С.А.	115	5. И.Н.	117
6. С.А.	107	6. И.Ч.	112
7. Т.А.	106	7. И.В.	105
8. Ф.А.	107	8. К.О.	108
9. Ч.И.	95	9. Р.Р.	111
10. Ц.А.	116	10. Р.И.	114
11. С.М.А.	127	11. О.К.	102
12. К.Ан.	115	12. Н.К.	104
13. Б.Л.	102		
14. Ф.В.	99		

Можно ли утверждать, что одна из выборок превосходит другую по уровню невербального интеллекта?

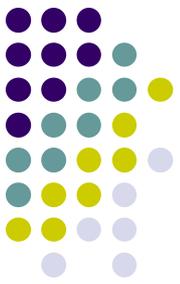


Проверка гипотез в Excel при помощи *U*-критерия Манна – Уитни

- 1) Занести показатели в первый и второй столбец
Найти объем каждой выборки

	А	В
1	Показатель невербального интеллекта	
2	Студенты - физики	Студенты-психологи
3	111	113
4	104	107
5	107	123
6	90	122
7	115	117
8	107	112
9	106	105
10	107	108
11	95	111
12	116	114
13	127	102
14	115	104
15	102	
16	99	
17		
18	n_1	n_2
19	14	12
20		

fx =СЧЁТ(А3:А16)



Проверка гипотез в Excel при помощи *U*-критерия Манна – Уитни

2) Выполнить операцию ранжирования

$$= \text{РАНГ}(A3; \$A\$3: \$B\$16; 1) + (\text{СЧЁТЕСЛИ}(\$A\$3: \$B\$16; A3) - 1) / 2$$

Показатель невербального интеллекта		Ранги	
Студенты - физики	Студенты-психологи	Студенты - физики	Студенты-психологи
111	113	15,5	18
104	107	6,5	11,5
107	123	11,5	25
90	122	1	24
115	117	20,5	23
107	112	11,5	17
106	105	9	8
107	108	11,5	14
95	111	2	15,5
116	114	22	19
127	102	26	4,5
115	104	20,5	6,5
102		4,5	
99		3	



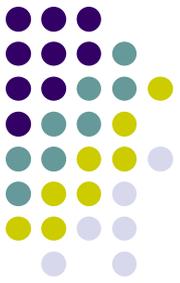
Проверка гипотез в Excel при помощи *U*-критерия Манна – Уитни

3) Найти ранговые суммы каждой выборки

	A	B	C	D	E
1	Показатель невербального интеллекта			Ранги	
2	Студенты - физики	Студенты-психологи		Студенты - физики	Студенты-психологи
3	111	113		15,5	18
4	104	107		6,5	11,5
5	107	123		11,5	25
6	90	122		1	24
7	115	117		20,5	23
8	107	112		11,5	17
9	106	105		9	8
10	107	108		11,5	14
11	95	111		2	15,5
12	116	114		22	19
13	127	102		26	4,5
14	115	104		20,5	6,5
15	102			4,5	
16	99			3	
17					
18	n_1	n_2		R_1	R_2
19	14	12		165	186

f_x =СУММ(E3:E16)

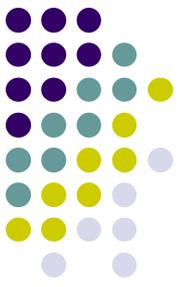
Проверка гипотез в Excel при помощи U -критерия Манна – Уитни



Так как $R_1=165 < 186=R_2$, то сформулируем такие гипотезы:

H_0 : «Группа студентов-психологов не превосходит группу студентов-физиков по уровню невербального интеллекта»

H_1 : «Группа студентов-психологов превосходит группу студентов-физиков по уровню невербального интеллекта»



Проверка гипотез в Excel при помощи U -критерия Манна – Уитни

5) Подсчитать значение U -критерия по каждой выборке

Показатель невербального интеллекта		Ранги			
Студенты - физики	Студенты-психологи	Студенты - физики	Студенты-психологи		
111	113	15,5	18		
104	107	6,5	11,5		
107	123	11,5	25		
90	122	1	24		
115	117	20,5	23		
107	112	11,5	17		
106	105	9	8		
107	108	11,5	14		
95	111	2	15,5		
116	114	22	19		
127	102	26	4,5		
115	104	20,5	6,5		
102		4,5			
99		3			
n_1	n_2	R_1	R_2	U_1	U_2
14	12	165	186	108	60

$$U_1 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - R_1,$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - R_2$$

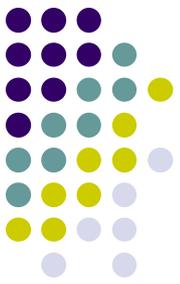


Проверка гипотез в Excel при помощи U -критерия Манна – Уитни

6) Найти эмпирическое значение критерия $U_{\text{эмп}}$

	G	H	I	J
17				
18		U_1	U_2	$U_{\text{эмп}}$
19		108	60	60
20				
21				

f_x =МИН(G19:H19)



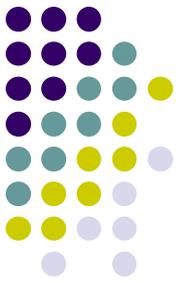
Проверка гипотез в Excel при помощи U -критерия Манна – Уитни

7) По таблице найти критическое значение критерия $U_{крит}$

U_1	U_2		$U_{эмп}$	$U_{кр}$
108	60		60	51
уровень значимости α				0,05

Критические значения критерия U Манна-Уитни												
	N=2	N=3	N=4	N=5	N=6	N=7	N=8	N=9	N=10	N=11	N=12	N
							p=	0,05				
3	-		0									
4	-		0	1								
5		0	1	2	4							
6		0	2	3	5	7						
7		0	2	4	6	8	11					
8		1	3	5	8	10	13	15				
9		1	4	6	9	12	15	18	21			
10		1	4	7	11	14	17	20	24	27		
11		1	5	8	12	16	19	23	27	31	34	
12		2	5	9	13	17	21	26	30	34	38	42
13		2	6	10	15	19	24	28	33	37	42	47
14		3	7	11	16	21	26	31	36	41	46	51
15		3	7	11	16	21	26	31	36	41	46	51

Проверка гипотез в Excel при помощи U -критерия Манна – Уитни

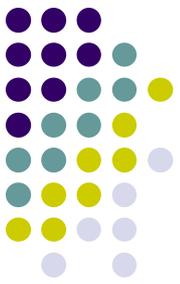


8) Сравнить эмпирическое $U_{эмп}$ и критическое $U_{крит}$ значение критерия

Так как $U_{эмп} = 60 > 51 = U_{крит}$, то нулевая гипотеза H_0 принимается на уровне значимости $\alpha=0,05$

Вывод: статистически значимое превосходство группы студентов-психологов над группой студентов-физиков по уровню невербального интеллекта не обнаружено

Проверка гипотез в Excel при помощи U -критерия Манна – Уитни



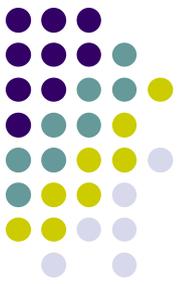
Примечание

Если нет возможности найти критическое значение критерия $U_{крит}$, то нужно подсчитать по эмпирическому значению критерия уровень $U_{эмп}$ значимости p

$$p = 2 \left(1 - \text{НОРМСТРАСП} \left(\frac{\frac{n_1 n_2}{2} - U_{эмп}}{\sqrt{\frac{n_1 n_2}{12} (n_1 + n_2 + 1)}} \right) \right)$$

Если уровень значимости $p \leq \alpha$, то нулевая гипотеза H_0 принимается

Если уровень значимости $p > \alpha$, то нулевая гипотеза H_0 отклоняется



Проверка гипотез в Excel при помощи *U*-критерия Манна – Уитни

Примечание

$$=2*(1-НОРМСТРАСП((A19*B19/2-J19)/КОРЕНЬ(A19*B19/12*(A19+B19+1))))$$

	A	B	C	D	E	F	G	H	I	J	K
17											
18	n_1	n_2		R_1	R_2		U_1	U_2		$U_{эмп}$	
19	14	12		165	186		108	60		60	
20											
21										уровень значимости α	0,05
22										уровень значимости p	0,21704
23											

f_x =2*(1-НОРМСТРАСП((A19*B19/2-J19)/КОРЕНЬ(A19*B19/12*(A19+B19+1))))

Проверка гипотез в STATISTICA при помощи U -критерия Манна – Уитни



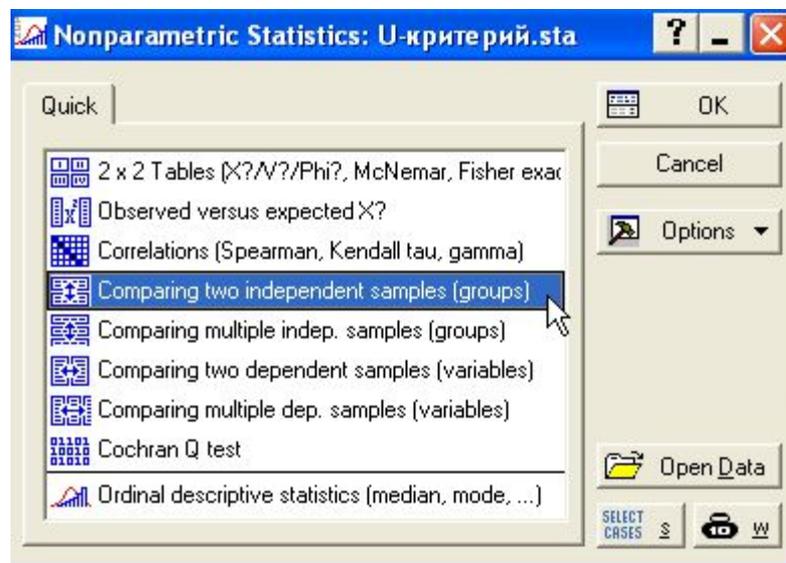
- 1) Занести все показатели в один столбец
Промаркировать показатели:
1 – для первой выборки,
2 – для второй выборки

	1 Показатели	2 Маркеры
1	111	1
2	104	1
3	107	1
4	90	1
5	115	1
6	107	1
7	106	1
8	107	1
9	95	1
10	116	1
11	127	1
12	115	1
13	102	1
14	99	1
15	113	2
16	107	2
17	123	2
18	122	2
19	117	2
20	112	2
21	105	2
22	108	2
23	111	2
24	114	2
25	102	2
26	104	2

Проверка гипотез в STATISTICA при помощи *U*-критерия Манна – Уитни



- 2) Выполнить последовательность команд *Статистика* → *Непараметрические данные*
- 3) В окне *Nonparametric Statistics* выбрать *Comparing two independent samples (groups)*

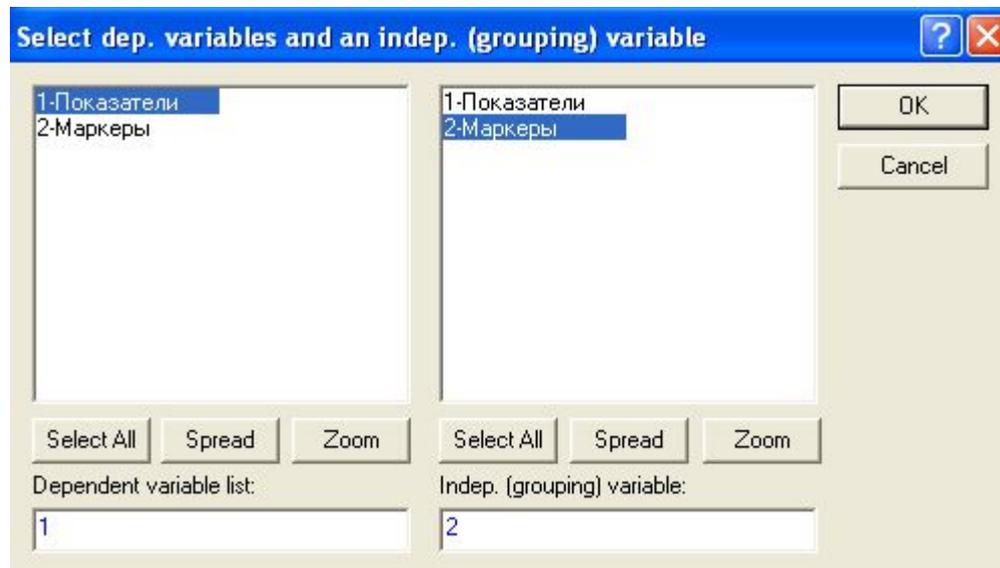


и нажать **ОК**

Проверка гипотез в STATISTICA при помощи U -критерия Манна – Уитни



- 4) Нажать кнопку **Variable** и выбрать в левом окне (зависимая переменная) – первую переменную, в правом (группированная переменная) – вторую переменную

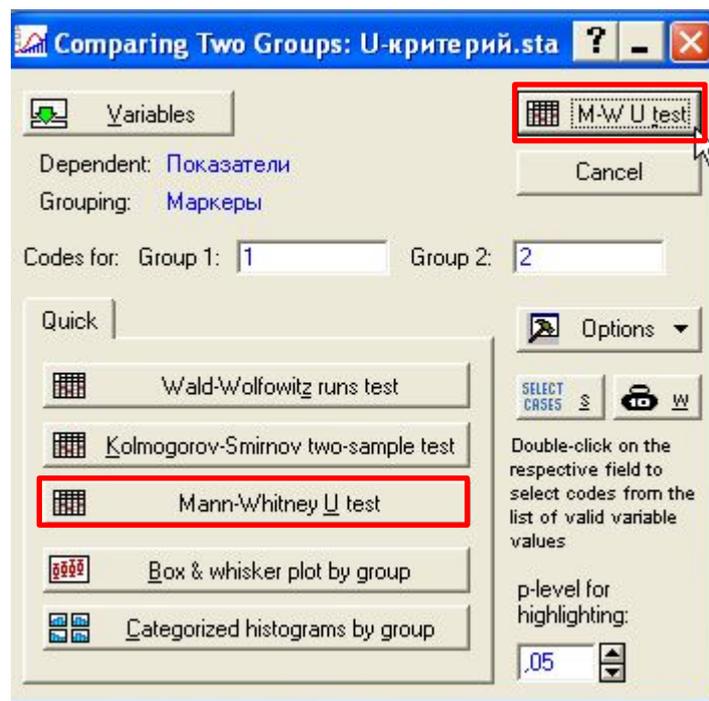


и нажать **ОК**

Проверка гипотез в STATISTICA при помощи U -критерия Манна – Уитни



5) Нажать на одну из кнопок *Mann-Whitney U test*



Проверка гипотез в STATISTICA при помощи U -критерия Манна – Уитни



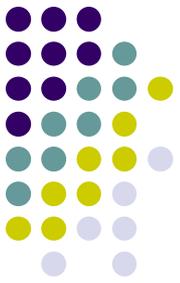
6) Результаты

variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-level	Z adjusted	p-level	Valid N Group 1	Valid N Group 2	2*1 sided exact p
Показатели	165,0000	186,0000	60,00000	-1,23443	0,217045	-1,23739	0,215943	14	12	0,231155

Если значения в этой таблице будут отмечены красным цветом, это будет означать, что нулевая гипотеза H_0 отклоняется

Так как в данном случае нет отмеченных красным результатов, то нулевая гипотеза H_0 принимается на уровне значимости $\alpha=0,05$

Проверка гипотез в STATISTICA при помощи U -критерия Манна – Уитни



7) **Вывод:** статистически значимое превосходство группы студентов-психологов над группой студентов-физиков по уровню невербального интеллекта не обнаружено