

Министерство образования и науки Российской Федерации
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

В. В. КОМИССАРОВ
Н. В. КОМИССАРОВА

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ПСИХОЛОГИИ

Лабораторный практикум

Новосибирск2013

УДК 159.9.07
ББК 88.4
К 634

Рецензенты:

Е. В. Казанцева, кандидат физ. мат. наук,
доцент кафедры высшей математики НГТУ

Н. Г. Орлова, кандидат физ. мат. наук,
доцент кафедры статистики и математики СибУПК

Комиссаров В. В.

К 634 Математические методы в психологии: / канд. физ. – мат. наук В. В. Комиссаров, ст. преподаватель Н. В. Комиссарова; – НГТУ. – Новосибирск, 2013. – 108 с.

УДК 159.9.07
ББК 88.4

Работа подготовлена на кафедре высшей математики для студентов,
обучающихся по направлению 0303000.62 «Психология»

© Комиссаров В. В., Комиссарова Н. В., 2013
© Новосибирский государственный
технический университет, 2013

СОДЕРЖАНИЕ

ОБЩИЕ ПОЛОЖЕНИЯ.....	4
ПРАВИЛА ОФОРМЛЕНИЯ ЛАБОРАТОРНЫХ РАБОТ	4
Лабораторное занятие 1	5
ОПРЕДЕЛЕНИЕ ТОЧЕЧНЫХ И ИНТЕРВАЛЬНЫХ ОЦЕНОК ВЫБОРКИ.....	5
Лабораторное занятие 2	14
ОЦЕНКИ РАЗЛИЧИЙ МЕЖДУ ВЫБОРКАМИ.....	14
Лабораторное занятие 3	34
ОЦЕНКА ДОСТОВЕРНОСТИ СДВИГА В ЗНАЧЕНИЯХ ИССЛЕДУЕМОГО ПРИЗНАКА.....	34
Лабораторное занятие 4	51
ВЫЯВЛЕНИЕ РАЗЛИЧИЙ В РАСПРЕДЕЛЕНИИ ПРИЗНАКА	51
Лабораторное занятие 5	64
СПИСОК ЛИТЕРАТУРЫ.....	90
Приложение	91

ОБЩИЕ ПОЛОЖЕНИЯ

Практикум включает 5 лабораторных занятий, достаточно полно характеризует основные подходы и критерии математических методов используемых в психологии.

ПРАВИЛА ОФОРМЛЕНИЯ ЛАБОРАТОРНЫХ РАБОТ

По лабораторным занятиям (кроме лабораторного занятия 1) следует составить отчёт. Отчёт, выполненный на листах формата А4 от руки или распечатанный на принтере и должен содержать:

- 1) титульный лист, на котором указывается номер лабораторного занятия, его тема, вариант, ФИО и номер группы студента;
- 2) постановку задачи;
- 3) предварительный анализ данных, содержащий обоснование применяемого критерия и проверки ограничений;
- 4) формулировку статистических гипотез;
- 5) промежуточные результаты расчетов;
- 6) эмпирическое значение критерия и теоретические значения критерия для уровней значимости 0,05 и 0,01;
- 7) сравнение эмпирического значения критерия с теоретическими и применение правила, позволяющего принять истинную гипотезу и отклонить ложную;
- 8) выводы по лабораторному занятию (математическая статистика);
- 9) выводы по решаемой задаче (психология);
- 10) список использованной литературы.

Лабораторное занятие 1

ОПРЕДЕЛЕНИЕ ТОЧЕЧНЫХ И ИНТЕРВАЛЬНЫХ ОЦЕНОК ВЫБОРКИ

Методические указания

Пусть из генеральной совокупности извлечена выборка, причем x_1 наблюдалось n_1 раз, x_2 – n_2 раз, x_k – n_k раз и $\sum n_i = n$ – объем выборки. Наблюдаемые значения x_i называют *вариантами*, а последовательность вариантов, записанных в возрастающем порядке, – *вариационным рядом*. Числа наблюдений называют *частотами*, а их отношения к объему выборки $\frac{n_i}{n} = W_i$ – *относительными частотами*.

Статистическим распределением выборки называют перечень вариантов и соответствующих им частот или относительных частот. Статистическое распределение можно задать также в виде последовательности интервалов и соответствующих им частот (в качестве частоты, соответствующей интервалу, принимают сумму частот, попавших в этот интервал).

Полигоном частот называют ломаную, отрезки которой соединяют точки $(x_1, n_1), (x_2, n_2), \dots, (x_k, n_k)$. Для построения полигона частот на оси абсцисс откладывают варианты x_i , а на оси ординат – соответствующие им частоты n_i . Точки (x_i, n_i) соединяют отрезками и получают полигон частот.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки $(x_1, W_1), (x_2, W_2), \dots, (x_k, W_k)$.

Выборочной средней $\bar{X}_B$ называют среднее арифметическое значение признака выборочной совокупности.

Если все значения $x_1, x_2, \dots, x_n$ признака выборки объема n различны, то

$$\bar{X}_B = \frac{1}{n}(x_1 + x_2 + \dots + x_n).$$

Если же значения признака $x_1, x_2, \dots, x_k$ имеют соответственно частоты $n_1, n_2, \dots, n_k$, причем $n_1 + n_2 + \dots + n_k = n$, то $\bar{X}_B = \frac{1}{n} \sum_{i=1}^k n_i x_i$.

Выборочной дисперсией D_B называют среднее арифметическое квадратов отклонений значений признака генеральной совокупности от их среднего значения $\bar{X}_B$.

Если все значения $x_1, x_2, \dots, x_n$ признака выборки объема n различны, то

$$D_B = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X}_B)^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2.$$

Если же значения признака $x_1, x_2, \dots, x_k$ имеют соответственно частоты $n_1, n_2, \dots, n_k$, причем $n_1 + n_2 + \dots + n_k = n$, то

$$D_B = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{X}_B)^2 = \frac{1}{n} \sum_{i=1}^k n_i x_i^2 - \left(\frac{1}{n} \sum_{i=1}^k n_i x_i \right)^2.$$

Исправленная дисперсия, которую обычно обозначают через S^2 :

$$S^2 = \frac{n}{n-1} D_B = \frac{\sum_{i=1}^k n_i (x_i - \bar{X}_B)^2}{n-1}.$$

Исправленная дисперсия является несмещенной оценкой генеральной дисперсии.

Для оценки среднего квадратического отклонения генеральной совокупности используют «исправленное» среднее квадратическое отклонение S , которое равно квадратному корню из исправленной дисперсии.

Асимметрия и эксцесс эмпирического распределения определяются равенствами

$$A_s = m_3/(\sigma_B)^3, E_s = m_4/(\sigma_B)^4 - 3,$$

где σ_B – выборочное среднее квадратическое отклонение; m_3 и m_4 – центральные эмпирические моменты третьего и четвертого порядков:

$$m_3 = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{X}_B)^3, m_4 = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{X}_B)^4,$$

или

$$m_3 = M_3 - 3M_2M_1 + 2M_1^3.$$

$$m_4 = M_4 - 4M_3M_1 + 6M_2M_1^2 - 3M_1^4,$$

где $M_k = (\sum n_i x_i^k) / n$ – начальные моменты k -го порядка.

Доверительный интервал, покрывающий неизвестный параметр a с надёжностью γ : $(\bar{x} - t\sigma / \sqrt{n}; \bar{x} + t\sigma / \sqrt{n})$; точность оценки $\delta = t\sigma / \sqrt{n}$.

Число t определяется из равенства $2\Phi(t) = \gamma$, или $\Phi(t) = \gamma/2$; по таблице функции Лапласа (таблица 1 приложения) находят аргумент t , которому соответствует значение функции Лапласа, равное $\gamma/2$.

Замечание. Если требуется оценить математическое ожидание a наперед заданной точностью δ и надёжностью γ , то минимальный объём выборки, который обеспечит эту точность, находят по формуле $n = t^2 \sigma^2 / \delta^2$.

Задание

По данной выборке (см. варианты заданий к лабораторному занятию 1):

- 1) постройте вариационный ряд;
- 2) постройте статистическое распределение, разбив выборку на 10 равных интервалов;
- 3) постройте полигон относительных частот;
- 4) найдите выборочное среднее, выборочную дисперсию, исправленную выборочную дисперсию, асимметрию, эксцесс;
- 5) найдите доверительный интервал для оценки с надёжностью γ (см. вариант) неизвестного математического ожидания a , если генеральное среднее квадратическое отклонение σ равно «исправленному» среднему квадратическому отклонению S ;
- 6) найдите минимальный объём выборки при котором с надёжностью $\gamma = 0,95$ точность оценки математического ожидания a генеральной совокупности по выборочной средней равна $\delta = 0,5$.

Пример выполнения лабораторной работы в MS Excel

86	81	Полигон относительных частот						
95	77							
82	77							
89	64							
91	75							
78	94							
90	70							
84	79							
81	76							
89	85							
$X_{min}=$	64	$X_{max}=$	104	$R=$	40	$h=$	4	
$X_{л}$	X_{np}	X_c	n_i	v_i	$X_c n_i$	$X_c^2 n_i$	$X_c^3 n_i$	$X_c^4 n_i$
64	68	66	2	0,033	132	8712	574992	37949472
68	72	70	5	0,083	350	24500	1715000	120050000
72	76	74	7	0,117	518	38332	2836568	209906032
76	80	78	8	0,133	624	48672	3796416	296120448
80	84	82	14	0,233	1148	94136	7719152	632970464
84	88	86	8	0,133	688	59168	5088448	437606528
88	92	90	9	0,150	810	72900	6561000	590490000
92	96	94	4	0,067	376	35344	3322336	312299584
96	100	98	2	0,033	196	19208	1882384	184473632
100	104	102	1	0,017	102	10404	1061208	108243216
			60		4944	411376	3,5E+07	2,93E+09



$X_{\text{выб}}=$	82,40	$M_1=$	82,4	$m_3=$	41,728		
$D_{\text{выб}}=$	66,51	$M_2=$	6856,27	$m_4=$	11176		
$S^2=$	67,63	$M_3=$	575958,4				
$\sigma_{\text{выб}}=$	8,16	$M_4=$	48835156,3				
$S=$	8,22						
$A_s=$	0,08						
$E_s=$	-0,47						
$\gamma=$	0,90						
$t=$	1,64	Доверительный интервал					
$\delta=$	1,75	(	80,65	84,15	)		
$\gamma=$	0,95						
$t=$	1,96						
$\delta=$	0,50						
$N=$	1039,25						

Минимальный объем выборки, обеспечивающий заданную точность:
1040

Замечание. Для нахождения абсолютных частот в MS Excel можно воспользоваться функцией =ЧАСТОТА(A1:F10;B18:B27), где первый аргумент – блок, в котором располагаются исходные данные, второй аргумент – правые границы частичных интервалов. Функция матричная, для корректного её использования необходимо прибегнуть к следующему алгоритму: а) выделяется блок, в котором должны располагаться результаты; б) вводится формула; в) завершается операция нажатием [Ctrl]+[Shift]+[Enter]. Найти решение уравнения $\Phi(t) = \gamma/2$ относительно переменной t можно найти по таблице I приложения или воспользовавшись функцией =НОРМОБР(0,5+НАДЕЖНОСТЬ/2;0;1), где НАДЕЖНОСТЬ – значение γ .

Варианты заданий к лабораторной работе №1

Вариант 1 ($\gamma = 0,95$)

81	81	86	89	77	82
85	85	85	78	85	81
77	89	74	91	94	74
84	89	73	83	91	89
90	79	69	99	69	92
82	87	81	89	85	102
82	77	84	79	78	83
90	82	89	85	75	90
81	82	100	94	94	86
63	79	87	75	93	70

Вариант 2 ($\gamma = 0,99$)

123	121	134	121	115	105
135	118	112	130	106	119
104	129	118	128	123	123
128	119	113	126	114	119
100	121	135	127	119	120
114	114	131	109	95	130
123	122	125	124	127	117
110	126	141	132	116	129
97	119	128	144	129	121
123	122	105	112	133	119

Вариант 3 ($\gamma = 0,9$)

91	81	103	94	70	92
79	93	70	108	81	90
86	72	102	75	112	75
96	71	87	98	89	95
79	94	88	93	95	98
90	104	86	93	68	54
105	79	87	73	69	85
82	86	85	55	73	78
109	89	89	90	89	88
66	98	79	90	108	79

Вариант 4 ($\gamma = 0,95$)

71	62	69	77	67	62
78	63	65	65	63	66
60	73	66	56	59	73
63	58	65	69	64	54
72	64	61	59	68	68
60	66	67	76	69	62
56	69	66	60	49	61
67	67	55	62	71	70
68	62	72	62	74	74
53	67	59	56	63	70

Вариант 5 ($\gamma = 0,9$)

97	96	99	98	91	90
87	95	97	96	88	120
107	87	100	90	88	100
79	102	94	102	100	101
97	99	104	103	86	105
81	98	104	105	94	103
100	86	91	97	105	104
83	101	110	87	97	90
86	98	80	93	116	87
94	102	101	83	100	98

Вариант 6 ($\gamma = 0,95$)

74	81	76	80	84	85
95	77	72	95	89	83
82	77	76	71	87	68
89	64	81	90	72	97
91	75	83	79	85	83
78	94	87	103	70	87
90	70	82	99	81	89
84	79	78	74	81	75
81	76	73	81	89	93
89	85	83	92	84	72

Вариант 7 ($\gamma = 0,99$)

114	104	84	84	103	105
101	81	91	76	89	89
103	92	80	96	92	103
104	96	87	111	93	93
78	113	104	74	109	98
110	90	92	101	100	84
91	75	93	97	102	112
96	96	90	95	113	88
83	90	99	95	84	98
83	103	91	87	89	96

Вариант 8 ($\gamma = 0,9$)

77	89	90	107	99	87
92	105	102	90	130	74
102	98	80	115	98	93
98	101	98	57	77	88
105	97	94	92	95	103
90	107	92	75	98	101
101	111	107	107	96	93
97	100	99	108	87	102
98	93	110	94	97	87
97	113	98	110	80	78

Вариант 9 ($\gamma = 0,95$)

137	130	150	131	122	133
144	142	117	125	131	126
110	122	141	121	123	151
131	141	133	136	135	127
133	139	116	128	110	143
140	131	113	133	122	124
135	127	150	138	128	115
125	102	121	127	113	130
135	134	118	141	131	133
112	138	144	125	125	137

Вариант 10 ($\gamma = 0,95$)

83	97	75	100	101	97
93	101	86	99	88	105
90	101	105	99	100	75
99	106	85	96	105	92
81	101	109	126	100	82
86	112	92	101	92	81
100	103	88	88	115	92
102	117	93	94	108	92
101	95	86	93	110	91
83	96	85	89	107	92

Вариант 11 ($\gamma = 0,9$)

81	81	86	89	77	82
85	85	85	78	85	81
77	89	74	91	94	74
84	89	73	83	91	89
90	79	69	99	69	92
82	87	81	89	85	102
82	77	84	79	78	83
90	82	89	85	75	90
81	82	100	94	94	86
63	79	87	75	93	70

Вариант 12 ($\gamma = 0,99$)

123	121	134	121	115	105
135	118	112	130	106	119
104	129	118	128	123	123
128	119	113	126	114	119
100	121	135	127	119	120
114	114	131	109	95	130
123	122	125	124	127	117
110	126	141	132	116	129
97	119	128	144	129	121
123	122	105	112	133	119

Вариант 13 ($\gamma = 0,95$)

91	81	103	94	70	92
79	93	70	108	81	90
86	72	102	75	112	75
96	71	87	98	89	95
79	94	88	93	95	98
90	104	86	93	68	54
105	79	87	73	69	85
82	86	85	55	73	78
109	89	89	90	89	88
66	98	79	90	108	79

Вариант 14 ($\gamma = 0,9$)

71	62	69	77	67	62
78	63	65	65	63	66
60	73	66	56	59	73
63	58	65	69	64	54
72	64	61	59	68	68
60	66	67	76	69	62
56	69	66	60	49	61
67	67	55	62	71	70
68	62	72	62	74	74
53	67	59	56	63	70

Вариант 15 ($\gamma = 0,99$)

97	96	99	98	91	90
87	95	97	96	88	120
107	87	100	90	88	100
79	102	94	102	100	101
97	99	104	103	86	105
81	98	104	105	94	103
100	86	91	97	105	104
83	101	110	87	97	90
86	98	80	93	116	87
94	102	101	83	100	98

Вариант 16 ($\gamma = 0,95$)

74	81	76	80	84	85
95	77	72	95	89	83
82	77	76	71	87	68
89	64	81	90	72	97
91	75	83	79	85	83
78	94	87	103	70	87
90	70	82	99	81	89
84	79	78	74	81	75
81	76	73	81	89	93
89	85	83	92	84	72

Лабораторное занятие 2

ОЦЕНКИ РАЗЛИЧИЙ МЕЖДУ ВЫБОРКАМИ

Методические указания

***Q*-критерий Розенбаума**

Назначение критерия. Критерий используется для оценки различий между *двумя* выборками по *уровню* какого-либо признака, количественно измеренного. В каждой из выборок должно быть не менее 11 испытуемых.

Описание критерия. Это очень простой непараметрический критерий, позволяющий быстро оценить различия между двумя выборками по какому-либо признаку. Критерий применяется в тех случаях, когда данные представлены по крайней мере в порядковой шкале. Однако если критерий *Q* не выявляет достоверных различий, это еще не означает, что их действительно нет.

В этом случае стоит применить критерий φ^* Фишера. Если же *Q*-критерий выявляет достоверные различия между выборками с уровнем значимости $\rho \leq 0,01$, можно ограничиться только им и избежать трудностей применения других критериев.

Применение критерия начинаем с того, что упорядочиваем значения признака в обеих выборках по нарастанию (или убыванию) признака.

Гипотезы:

H_0 : уровень признака в выборке 1 не превышает уровня признака в выборке 2;

H_1 : уровень признака в выборке 1 превышает уровень признака в выборке 2.

Графическое представление критерия *Q*. Рассмотрим три варианта соотношения рядов значений в двух выборках. В первом варианте все значения первого ряда выше всех значений второго ряда. Различия, безусловно, достоверны, при соблюдении условия $n_1, n_2 \geq 11$.

В варианте на рис. 1, напротив, оба ряда находятся на одном и том же уровне: различия недостоверны. В варианте на рис. 2 ряды частично перекрещиваются, но все же первый ряд оказывается гораздо выше второго.

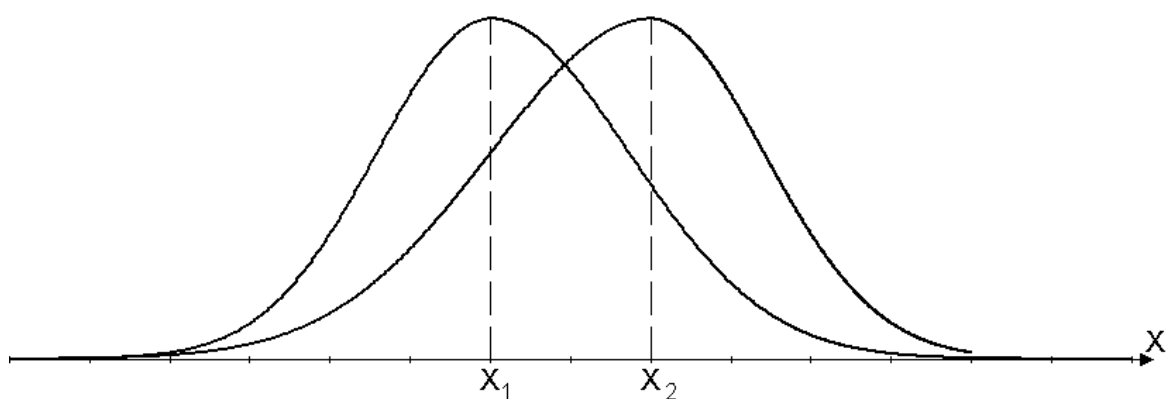


Рис. 1. Вариант соотношения распределений признака в двух выборках, при котором критерий Q беспомощен

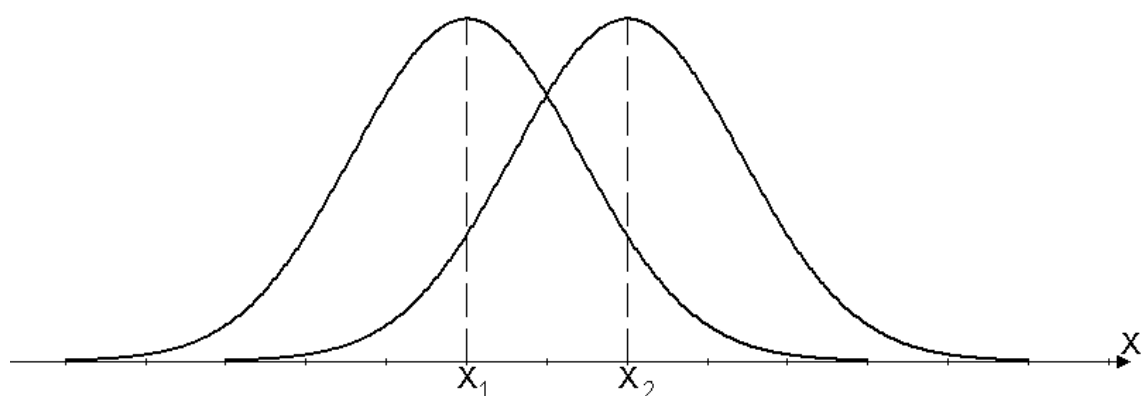


Рис. 2. Вариант соотношения распределений признака в двух выборках, при котором критерий Q может быть могущественным.

Ограничения критерия Q

1. В каждой из сопоставляемых выборок должно быть не менее 11 наблюдений; при этом объемы выборок должны примерно совпадать.

При этом учитываются следующие правила:

а) если в обеих выборках меньше 50 наблюдений, то абсолютная величина разности между n_1 и n_2 не должна быть больше 10 наблюдений;

б) если в каждой из выборок больше 51 наблюдения, но меньше 100, то абсолютная величина разности между n_1 и n_2 не должна быть больше 20 наблюдений;

в) если в каждой из выборок больше 100 наблюдений, то допускается, чтобы одна из выборок была больше другой не более чем в 1,5–2 раза.

2. Диапазоны разброса значений в двух выборках не должны совпадать, в противном случае применение критерия бессмысленно. Между тем возможны случаи, когда диапазоны разброса значений совпадают, но вследствие разносторонней асимметрии двух распределений различия в средних величинах признаков существенны (см. рис. 1).

Алгоритм подсчёта критерия Q Розенбаума

1. Проверить, выполняются ли ограничения: $n_1, n_2 \geq 11, n_1 \approx n_2$.
2. Упорядочить значения отдельно в каждой выборке по степени возрастания признака. Считать выборкой 1 ту выборку, значения в которой предположительно выше, а выборкой 2 – ту, где значения предположительно ниже.
3. Определить самое высокое (максимальное) значение в выборке 2.
4. Подсчитать количество значений в выборке 1, которые выше максимального значения в выборке 2. Полученную величину обозначить S_1 .
5. Определить самое низкое (минимальное) значение в выборке 1.
6. Подсчитать количество значений в выборке 2, которые ниже минимального значения выборки 1. Полученную величину обозначить S_2 .
7. Подсчитать эмпирическое значение Q по формуле: $Q_{\text{эмп}} = S_1 + S_2$.
8. По таблице 2 приложения определить критические значения Q для данных n_1 и n_2 . Различия между двумя выборками можно считать достоверными ($\rho \leq 0,05$), если $Q_{\text{эмп}}$ равен или выше критического значения $Q_{0,05}$ и тем более достоверными ($\rho \leq 0,01$), если $Q_{\text{эмп}}$ равен или выше критического значения $Q_{0,01}$.
9. При $n_1, n_2 > 26$ сопоставить полученное эмпирическое значение с $Q_{\text{кр}} = 8$ ($\rho \leq 0,05$) и $Q_{\text{кр}} = 10$ ($\rho \leq 0,01$). Если $Q_{\text{эмп}}$ превышает или по крайней мере равняется $Q_{\text{кр}} = 8$, H_0 отвергается.

U -критерий Манна–Уитни

Назначение критерия. Критерий предназначен для оценки различий между двумя выборками по уровню какого-либо признака, количественно измеренного. Он позволяет выявлять различия между малыми выборками, когда $n_1, n_2 \geq 3$ или $n_1 = 2, n_2 \geq 5$, и является более мощным, чем критерий Розенбаума.

Описание критерия. Этот метод определяет, достаточно ли мала зона перекрещивающихся значений между двумя рядами. Положим, что первым рядом (выборкой, группой) мы называем тот ряд значений, в котором значения, по предварительной оценке, выше, а вторым рядом – тот, где они предположительно ниже.

Чем меньше область перекрещивающихся значений, тем более вероятно, что различия достоверны. Иногда их называют различиями в расположении двух выборок.

Эмпирическое значение критерия U отражает то, насколько велика зона совпадения между рядами. Поэтому чем меньше $U_{\text{эмп}}$, тем более вероятно, что различия достоверны.

Гипотезы:

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.

Ограничения критерия U

1. В каждой выборке должно быть не менее 3 наблюдений: $n_1, n_2 \geq 3$; допускается, чтобы в одной выборке было 2 наблюдения, но тогда во второй их должно быть не менее 5.

2. В каждой выборке должно быть не более 60 наблюдений; $n_1, n_2 \leq 60$.

Однако уже при $n_1, n_2 > 20$ ранжирование становится достаточно трудоёмким.

Замечание. В случае, если $n_1, n_2 > 20$, лучше использовать другой критерий, а именно, угловое преобразование Фишера в комбинации с критерием λ , позволяющим выявить критическую точку, в которой накапливаются максимальные различия между двумя сопоставляемыми выборками. Формулировка сложная, но сам метод достаточно прост.

Правила ранжирования

1. Меньшему значению начисляется меньший ранг (наименьшему значению начисляется ранг 1). Наибольшему значению начисляется ранг, соответствующий количеству ранжируемых значений. Например, если $n = 7$, то наибольшее значение получит ранг 7, за возможным исключением для тех случаев, которые предусмотрены правилом 2.

2. Если несколько значений равны, им начисляется ранг, представляющий собой среднее значение из тех рангов, которые они получили бы, если бы не были равны. Например, 3 наименьших значения равны 10 секундам. Если бы мы измеряли время более точно, то

эти значения могли бы различаться и составляли бы, скажем, 10,2 сек; 10,5 сек; 10,7 сек. В этом случае они получили бы ранги, соответственно, 1, 2 и 3. Но поскольку значения равны, каждое из них получает средний ранг: $(1+2+3)/3 = 6/3 = 2$

Допустим, следующие 2 значения равны 12 сек. Они должны были бы получить ранги 4 и 5, но, поскольку они равны, то получают средний ранг: $(4 + 5)/2 = 4,5$

3. Общая сумма рангов должна совпадать с расчетной, которая определяется по формуле

$$\sum(R_i) = \frac{N(N+1)}{2},$$

где N – общее количество ранжируемых наблюдений (значений). Несовпадение реальной и расчётной суммы рангов будет свидетельствовать об ошибке, допущенной при начислении рангов или их суммировании. Прежде чем продолжить работу, необходимо найти ошибку и устранить её.

Алгоритм подсчёта критерия U Манна–Уитни

1. Объединить все данные в единый ряд, пометив данные, принадлежащие разным выборкам.

2. Проранжировать значения, приписывая меньшему значению меньший ранг. Всего рангов получится столько, сколько у нас: $(n_1 + n_2)$.

3. Подсчитать сумму рангов отдельно для каждой выборки.

4. Определить большую из двух ранговых сумм.

5. Определить значение U по формуле

$$U = n_1 \cdot n_2 + n_x \cdot (n_x + 1)/2 - T_x,$$

где n_1 – количество испытуемых в выборке 1;

n_2 – количество испытуемых в выборке 2;

T_x – большая из двух ранговых сумм;

n_x – количество испытуемых в группе с большей суммой рангов.

6. Определить критические значения U по таблице 3 приложения.

Если $U_{\text{эмп}} > U_{\text{кр}0,05}$, H_0 принимается. Если $U_{\text{эмп}} \leq U_{\text{кр}0,05}$ H_0 отвергается. Чем меньше значения U , тем достоверность различий выше.

***H*–критерий Крускала–Уоллиса**

Назначение критерия. Критерий предназначен для оценки различий одновременно между тремя (четырьмя и т. д.) выборками по уровню какого-либо признака.

Он позволяет установить, то что уровень признака изменяется при переходе от группы к группе, но не указывает на направление этих изменений.

Описание критерия. Критерий *H* иногда рассматривается как непараметрический аналог метода дисперсионного однофакторного анализа для несвязных выборок. Иногда его называют критерием «суммы рангов».

Данный критерий является продолжением критерия *U* на большее, чем 2, количество сопоставляемых выборок. Все индивидуальные значения ранжируются так, как если бы это была одна большая выборка. Затем все индивидуальные значения возвращаются в свои первоначальные выборки, и мы подсчитываем суммы полученных ими рангов отдельно по каждой выборке. Если различия между выборками случайны, суммы рангов не будут различаться существенно, так как высокие и низкие ранги равномерно распределятся между выборками. Но если в одной из выборок будут преобладать низкие значения рангов, в другой – высокие, а в третьей – средние, то критерий *H* позволит установить эти различия.

Гипотезы:

H_0 : Между выборками 1, 2, 3 (и т. д.) существуют лишь случайные различия по уровню исследуемого признака.

H_1 : Между выборками 1, 2, 3 (и т. д.) существуют неслучайные различия по уровню исследуемого признака.

Графическое представление критерия *H*. Критерий *H* оценивает общую сумму перекрещивающихся зон при сопоставлении всех обследованных выборок. Если суммарная область наложения мала, то различия достоверны; если она достигает определенной критической величины и превосходит её, то различия между выборками оказываются недостоверными.

Ограничения критерия *H*

1. При сопоставлении трёх выборок допускается, чтобы в одной из них $n = 3$, а двух других $n = 2$. Но при таких численных составах выборок мы сможем установить различия лишь на низшем уровне значимости ($p \leq 0,05$).

Для того, чтобы стало возможным диагностирование различия на более высоком уровне значимости ($p \leq 0,01$), необходимо, чтобы в каждой выборке было не менее 3 наблюдений, или чтобы по крайней мере в одной из них было 4 наблюдения, а в двух других – по 2; при этом неважно, в какой именно выборке сколько испытуемых, а важно соотношение 4 : 2 : 2.

2. Критерий Крускала–Уоллиса асимптотически приближается к распределению χ^2 . Критические значения χ^2 приведены в таблице 4 приложения. Количество степеней свободы при этом определяется по формуле $v = c - 1$, где c – количество сопоставляемых выборок.

3. При множественном сопоставлении выборок достоверные различия между какой-либо конкретной парой (или парами) их могут оказаться стертыми. Это ограничение можно преодолеть, если провести все возможные попарные сопоставления, число которых будет равняться $c \cdot (c - 1) / 2$. Для таких попарных сопоставлений используется, естественно, критерий для двух выборок, например U или ϕ^* .

Алгоритм подсчёта критерия Н Крускала–Уоллиса

1. Объединить все данные в единый ряд, пометив данные, принадлежащие разным выборкам.

2. Упорядочить данные по степени нарастания признака.

3. Проранжировать значения, приписывая меньшему значению меньший ранг. Общее количество рангов будет равняться количеству испытуемых в объединённой выборке.

4. Подсчитать суммы рангов отдельно по каждой группе. Проверить совпадение общей суммы рангов с расчётной.

5. Подсчитать значение критерия H по формуле

$$H = \left[\frac{12}{N(N+1)} \cdot \sum_j \frac{T_j^2}{n_j} \right] - 3(N+1)$$

где N – общее количество испытуемых в объединённой выборке;

n_j – количество испытуемых в каждой группе;

T_j – суммы рангов по каждой группе.

6. При количестве групп $c > 3$ или количестве испытуемых $n_1, n_2, n_3 > 5$ определить критические значения χ^2 по таблице 4 приложения. Различия между двумя распределениями могут считаться достоверными, если $H_{\text{эмп}} = \chi^2_{\text{эмп}}$ достигает или превышает $\chi^2_{0,05}$, и тем более достоверными, если $\chi^2_{\text{эмп}}$ достигает или превышает $\chi^2_{0,01}$. Количество степеней свободы при этом определяется по формуле: $v = c - 1$, где c – количество сопоставляемых выборок.

S –критерии тенденций Джонкира

Назначение критерия S . Критерий S предназначен для выявления тенденций изменения Признака при переходе от выборки к выборке при сопоставлении трёх и более выборок.

Описание критерия S . Критерий S позволяет нам упорядочить обследованные выборки по какому-либо признаку, например, по креативности, фрустрационной толерантности, гибкости и т. п.

Критерий S основан на способе расчёта, близком к принципу критерия Q Розенбаума. Все выборки располагаются в порядке возрастания исследуемого признака, при этом выборку, в которой значения в общем ниже, мы помещаем слева, выборку, в которой значения выше, правее, и так далее в порядке возрастания значений. Таким образом, все выборки выстраиваются слева направо в порядке возрастания значений исследуемого признака.

Упорядочивая выборки, мы можем опираться на средние значения в каждой выборке или даже на суммы всех значений в каждой выборке, потому что в каждой выборке должно быть одинаковое количество значений. В противном случае критерий S неприменим.

Для каждого индивидуального значения подсчитывается количество значений справа, превышающих его по величине. Если тенденция возрастания признака слева направо существенная, то большая часть значений справа должна быть выше. Критерий S позволяет определить, преобладают справа более высокие значения или нет. Статистика S отражает степень этого преобладания. Чем выше эмпирическое значение S , тем тенденция возрастания признака является более существенной.

Следовательно, если $S_{\text{эмп}}$ равняется критическому значению или превышает его, нулевая гипотеза может быть отвергнута.

Гипотезы:

H_0 : Тенденция возрастания значений признака при переходе от выборки к выборке является случайной;

H_1 : тенденция возрастания значений признака при переходе от выборки к выборке не является случайной.

Графическое представление критерия. Фактически критерий S позволяет определить, достаточно ли велика суммарная зона непрекращающихся значений в сопоставляемых выборках: действительно ли в первом ряду значения в общем ниже, чем в последующих, во втором – ниже, чем в оставшихся справа последующих, и т. д.

Ограничения критерия S

1. В каждой из сопоставляемых выборок должно быть одинаковое число наблюдений. Если число наблюдений неодинаково, то придётся искусственно уравнивать выборки, утрачивая при этом часть полученных наблюдений.

Если исследователь хочет избежать этого, ему надо воспользоваться критерием H , позволяющим выявить различия между тремя и более выборками без указания на направление этих различий.

2. Нижний порог: не менее 3 выборок и не менее 2 наблюдений в каждой выборке. Верхний порог в существующих таблицах: не более 6 выборок и не более 10 наблюдений в каждой выборке (см. таблицу 5 приложения для определения критических значений S). При большем количестве выборок или наблюдений в них придётся пользоваться критерием H Крускала–Уоллиса.

Алгоритм подсчёта критерия S Джонкира

1. Если количества испытуемых в группах не совпадают, уравнивать группы, ориентируясь на количество наблюдений в меньшей из групп.

2. Занести данные в таблицу, упорядочив их в каждой группе по возрастанию признака.

3. Начиная с крайнего левого столбца подсчитать для каждого индивидуального значения количество превышающих его значений во всех столбцах справа S_i .

4. Подсчитать суммы показателей в скобках по столбцам.

5. Подсчитать общую сумму, т. е. сложить все суммы по столбцам. Эту общую сумму обозначить как A .

6. Подсчитать максимально возможное количество превышающих значений B , которое мы получили бы, если бы все значения справа были выше значений слева:

$$B = \frac{c(c-1)}{2} \cdot n^2,$$

где c – количество столбцов (сопоставляемых групп);

n – количество наблюдений в каждом столбце (группе).

7. Определить эмпирическое значение S по формуле $S = 2A - B$.

8. Определить критические значения S по таблице 5 приложения для данного количества групп c и количества испытуемых в каждой группе n . Если эмпирическое значение S превышает критическое значение или равно ему, H_0 отвергается.

Тенденция является достоверной, если $S_{\text{эмп}}$ достигает $S_{0,05}$ или превышает его ($\rho \leq 0,05$), и ещё более достоверной, если $S_{\text{эмп}}$ достигает $S_{0,01}$ или превышает его ($\rho \leq 0,01$).

Задача 1

У предполагаемых участников психологического эксперимента измерен уровень вербального и невербального интеллекта с помощью методики Д. Векслера. Было обследовано две группы юношей в возрасте от 18 до 24 лет студентов физического и психологического факультета. Показатели вербального интеллекта представлены в таблице 1.

Можно ли утверждать, что одна группа превосходит другую по уровню вербального интеллекта?

Таблица 1

Показатели вербального интеллекта в группах

Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5	
Ф	П	Ф	П	Ф	П	Ф	П	Ф	П
135	130	124	115	133	125	135	115	132	124
130	129	137	120	126	117	130	125	129	124
131	121	125	118	130	128	134	119	132	122
128	129	126	122	134	124	132	125	122	126
127	119	135	118	122	117	137	125	130	124
137	124	130	118	125	128	134	118	132	126
126	125	126	122	125	129	128	128	136	116
137	129	131	131	121	122	136	121	134	119
131	129	132	121	129	129	131	116	121	118
137	130	127	124	128	132	133	113	134	125
137	131	137	119	133	131	129	115	135	128
127	123	123	127	121	131	133	129	127	128
133		126	119	127		129	121	130	116
125		133	132			131	124	131	
		126						126	

Вариант 6		Вариант 7		Вариант 8		Вариант 9		Вариант 10	
Ф	П	Ф	П	Ф	П	Ф	П	Ф	П
128	123	122	129	124	128	134	121	122	116
129	128	128	115	130	119	121	119	136	128
124	116	134	123	134	125	128	127	124	124
130	125	124	120	124	115	128	122	135	115
121	125	134	123	132	123	123	127	128	118
134	120	123	124	126	129	134	118	123	114
134	121	125	126	121	127	135	117	123	121
124	114	126	115	132	118	131	121	126	119
131	126	134	126	136	125	121	117	131	116
134	115	132	127	135	130	129	128	125	123
136	114	123	114	128	118	125	122	130	125
129	122	130	126	125	115	126	128	136	126
123	113	128	128	129		125	122	127	
125	120	134	119	133		130	118	131	
134									
Вариант 11		Вариант 12		Вариант 13		Вариант 14		Вариант 15	
Ф	П	Ф	П	Ф	П	Ф	П	Ф	П
135	124	125	128	126	118	134	130	123	133
125	126	129	125	129	127	123	120	128	132
134	124	136	130	135	123	126	120	136	132
131	116	126	127	133	121	126	124	134	118
126	118	131	117	133	118	128	124	129	122
123	116	131	131	124	117	131	132	124	125
129	127	129	125	127	125	130	118	133	118
135	121	134	127	127	129	124	126	128	123
122	118	126	131	126	119	126	126	133	132
135	118	129	121	131	121	128	117	125	123
126	128	127	131	131	119	128	132	133	115
130	120	122	120	127	126	135	115	129	127
133		135		131	127	128	120	136	
129				129		125	123		
125									

Задача 2

Имеются результаты обследования студентов физического и психологического факультетов с помощью методики Д. Векслера, предназначенной для измерения вербального и невербального интеллекта. Показатели невербального интеллекта представлены в таблице 2.

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

Таблица 2

Показатели невербального интеллекта в группах

Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5	
Ф	П	Ф	П	Ф	П	Ф	П	Ф	П
123	104	106	125	93	123	107	115	114	108
124	104	107	114	100	121	127	117	101	115
107	121	94	115	111	124	93	118	110	101
100	111	129	124	99	118	105	105	99	103
91	111	112	110	111	106	90	105	106	100
94	112	90	113	111	105	100	118	107	118
123	119	94	122	130	115	105	108	93	120
114	117	126	121	94	111	102	116	101	105
129	123	100	102	128	110	101	103	97	120
101	111	96	105	92	106	127	125	124	119
109	109	116	115	90	120	102	115	104	105
107	102	115		109	109	103		95	114
99	112	98		105		115		109	123
110						108			
96						111			

Продолжение табл. 2

Вариант 6		Вариант 7		Вариант 8		Вариант 9		Вариант 10	
Ф	П	Ф	П	Ф	П	Ф	П	Ф	П
113	124	94	105	128	121	110	121	119	110
119	121	113	117	93	119	91	122	103	105
112	114	124	124	122	106	101	116	104	122
99	117	120	124	118	117	123	110	104	113
128	123	126	123	94	119	106	105	92	110
119	106	92	100	123	116	101	118	118	107
104	102	118	122	101	122	121	114	123	123
105	118	124	102	125	114	129	105	102	113
125	107	107	119	121	119	118	112	112	116
115	102	120	121	107	121	93	120	104	106
91	125	111	109	123	106	95	114	115	113
108	107	112	111	113	119	117		98	107
97		121		126		99		103	
		96						123	
		117							

Вариант 11		Вариант 12		Вариант 13		Вариант 14		Вариант 15	
Ф	П	Ф	П	Ф	П	Ф	П	Ф	П
100	117	110	109	126	104	124	104	116	105
105	113	125	116	103	113	91	108	128	122
103	105	113	121	116	116	123	100	100	109
123	119	115	119	98	118	104	125	130	117
107	109	100	114	121	122	112	111	116	100
101	110	109	101	120	104	113	101	123	101
121	109	124	104	130	114	97	101	91	104
129	118	129	110	101	101	102	118	100	109
94	109	116	105	112	103	103	107	116	110
102	115	107	114	104	114	116	110	95	110
105	104	105	121	92	109	95	120	112	114
92	109	102		97	107	103		106	115
101	116			130		128			
102									
126									

Задача 3

В эксперименте по исследованию интеллектуальной настойчивости испытуемым предлагали сначала разрешимые четырехбуквенные, пятибуквенные и шестибуквенные анаграммы, а затем неразрешимые анаграммы, время работы над которыми не ограничивалось. Эксперимент проводился индивидуально с каждым испытуемым. Было использовано 4 комплекта анаграмм. У исследователя возникло впечатление, что над некоторыми неразрешимыми анаграммами испытуемые продолжали работать дольше, чем над другими, и, возможно, необходимо будет делать поправку на то, какая именно неразрешимая анаграмма предъявлялась тому или иному испытуемому. Показатели длительности попыток в решении неразрешимых анаграмм представлены в таблице 3.

Можно ли утверждать, что длительность попыток решения каждой из 4 неразрешимых анаграмм примерно одинакова?

Таблица 3

**Показатели длительности попыток
в решении неразрешимых анаграмм**

Вариант 1				Вариант 2			
Ф	К	С	Г	Ф	К	С	Г
289	100	124	69	58	135	255	59
667	135	354	1465	281	1228	549	318
794	966	449	2011	565	1336	1081	715
1467	994	1206	3073	705	1535	1462	1280
1798	1188	1298		1729	1700	1571	2281
	1643	2143		1944		1902	
	1931			2154		2084	
	2064						
Вариант 3				Вариант 4			
Ф	К	С	Г	Ф	К	С	Г
190	127	73	54	465	395	155	118
233	413	349	168	634	544	591	485
957	491	982	977	1091	1007	690	523
1311	936	1108	1062	1294	1585	1040	936
1644	1014	1489	1153	1923	2057	1424	1457
	1117	1515	1510			1866	2986
	1262	2838	2752			2208	
	2180						

Продолжение табл. 3

Вариант 5				Вариант 6			
Ф	К	С	Г	Ф	К	С	Г
275	789	144	437	456	59	59	316
353	948	152	564	561	64	216	349
913	1057	827	663	873	165	610	776
1107	1645	1251	1445	1593	277	665	1086
1511	2257	1546	1752	1731	588	1012	1757
2025	2272	2128	2509	2078	722	1053	1994
		2540		2181	1061	1402	
		2972			1317	1889	
					1537	1956	
					2347		
					2700		
Вариант 7				Вариант 8			
Ф	К	С	Г	Ф	К	С	Г
239	254	89	154	116	332	118	185
278	379	110	237	522	665	135	428
830	632	764	287	984	755	446	891
927	800	1068	372	1757	919	639	1221
1646	820	1372	574	2078	1179	920	1940
	1383	1734	1205		2032	939	
	2212		1780			1532	
			2547			1606	
						1696	
Вариант 9				Вариант 10			
Ф	К	С	Г	Ф	К	С	Г
167	120	150	110	272	370	226	586
564	151	267	196	378	562	426	732
1319	654	837	745	479	1035	1115	1018
1599	1136	1034	2114	682	1446	1129	1059
1765	2189	1670		905	1668	1239	1073
2060		2518		1484		1627	2101
				1958			
				2203			

Вариант 11				Вариант12			
Ф	К	С	Г	Ф	К	С	Г
257	308	76	212	392	114	293	76
812	437	100	242	408	363	630	184
1043	599	484	417	673	957	745	1705
1122	1109	572	1079	751	1051	825	2193
1792	1228	2024	1711	2039	1630	1031	
	1490				1837	1713	
	1603				2008	2106	
	1704					2645	
Вариант13				Вариант14			
Ф	К	С	Г	Ф	К	С	Г
86	181	218	81	74	248	79	237
97	370	242	245	306	330	605	238
758	443	550	404	655	862	1025	316
1171	759	917	537	1019	1153	1068	575
1889	1082	1363	1002	1326	2031	1174	813
	1104	2044	1051	1531		1513	864
	1227		1348	2037			1201
	1487		1895				1805
	1996		1935				
	2042						
Вариант15				Вариант16			
Ф	К	С	Г	Ф	К	С	Г
67	225	231	91	243	263	83	104
139	313	448	269	314	471	587	333
234	422	858	487	708	748	1019	565
314	2292	1192	1053	824	940	1330	745
1058		1341	2092	937	1183	1865	2162
1393		1590		1257	1359		
1844		2332		1710	1921		
				2334			

Задача 4

Выборка претендентов на должность коммерческого директора фирмы была обследована с помощью Оксфордской методики экспресс-видеодиагностики, использующей диагностические ролевые игры. Были обследованы 20 мужчин в возрасте от 25 до 40 лет, средний возраст – 31,5 года. Одно из определяемых качеств – «Авторитетность». В конце 8-часового сеанса диагностических ролевых игр и упражнений проводился социометрический опрос участников группы. Они должны были определить: «Если бы я сам был представителем фирмы, я выбрал бы на должность коммерческого директора: 1)..., 2)..., 3)....». В результате этой процедуры каждый участник получил то или иное количество выборов от других участников, отражающее его социометрический статус в группе претендентов.

Результаты исследования представлены в таблице 4.

Можно ли считать, что группы с разным статусом различаются и по уровню авторитетности, определявшейся независимо от социометрии с помощью экспресс-видеодиагностики?

Таблица 4

Показатели по шкале авторитетности в группах с разным социометрическим статусом

Вариант 1					Вариант 2				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	3	6	7	7	1	6	6	7	8
2	6	7	6	9	2	6	7	5	10
3	5	7	5	8	3	6	6	8	7
4	3	3	5	9	4	4	6	7	8
5	6	4	7	7	5	6	4	7	7
Вариант 3					Вариант 4				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	6	6	7	8	1	5	5	7	7
2	6	7	5	10	2	5	4	6	9
3	6	6	8	7	3	2	5	8	7
4	4	6	7	8	4	6	6	5	10
5	6	4	7	7	5	2	6	5	9

Вариант 5					Вариант 6				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	4	7	4	7	1	5	4	5	10
2	4	5	4	7	2	3	4	6	9
3	3	7	7	7	3	5	4	8	9
4	3	7	5	9	4	4	5	7	6
5	5	3	6	8	5	4	6	6	9
Вариант 7					Вариант 8				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	4	5	6	8	1	2	7	5	6
2	3	6	6	9	2	6	7	5	9
3	4	6	6	8	3	5	4	4	7
4	5	3	5	7	4	4	3	7	10
5	5	7	6	8	5	5	7	8	6
Вариант 9					Вариант 10				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	6	4	5	6	1	6	6	8	8
2	2	5	5	7	2	3	5	5	10
3	6	4	5	9	3	4	3	8	9
4	3	7	8	6	4	4	6	7	6
5	4	3	6	8	5	3	6	8	9
Вариант 11					Вариант 12				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	4	7	6	10	1	3	5	7	9
2	6	4	4	9	2	4	6	7	8
3	3	3	7	9	3	3	3	5	7
4	4	4	5	9	4	6	5	5	6
5	5	6	4	8	5	2	4	8	6

Вариант 13					Вариант 14				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	5	6	6	10	1	3	6	6	8
2	3	4	8	9	2	4	3	8	7
3	3	4	5	6	3	4	5	6	6
4	4	6	7	6	4	6	6	6	6
5	4	6	4	6	5	2	6	7	7
Вариант 15					Вариант 16				
№	Гр. 1	Гр. 2	Гр. 3	Гр. 4	№	Гр. 1	Гр. 2	Гр. 3	Гр. 4
1	5	4	6	8	1	5	4	5	8
2	4	6	8	6	2	4	5	7	5
3	4	4	7	7	3	4	4	7	8
4	6	5	7	9	4	6	5	9	9
5	2	6	5	8	5	3	6	5	8

Лабораторное занятие 3

ОЦЕНКА ДОСТОВЕРНОСТИ СДВИГА В ЗНАЧЕНИЯХ ИССЛЕДУЕМОГО ПРИЗНАКА

Методические указания

G-критерий знаков

Назначение критерия G. Критерий знаков G предназначен для установления общего *направления* сдвига исследуемого признака.

Он позволяет определить, в какую сторону в выборке в целом изменяются значения признака при переходе от первого измерения ко второму: изменяются ли показатели в сторону улучшения, повышения или усиления или, наоборот, в сторону ухудшения, понижения или ослабления.

Описание критерия G. Критерий знаков применим и к тем сдвигам, которые можно определить лишь *качественно* (например, изменение отрицательного отношения к чему-либо на положительное), и к тем, которые могут быть измерены *количественно* (например, сокращение времени работы над заданием после экспериментального воздействия).

Во втором случае, однако, если сдвиги варьируют в достаточно широком диапазоне, лучше применять критерий T Вилкоксона. Он учитывает не только направление, но и интенсивность сдвигов и может оказаться более мощным в определении достоверности сдвигов, чем критерий знаков.

Суть критерия знаков состоит в том, что он определяет, не слишком ли много наблюдается «нетипичных сдвигов», чтобы сдвиг в «типичном» направлении считать преобладающим? Ясно, чем меньше «нетипичных сдвигов», тем вероятнее, что преобладание «типичного» сдвига. $G_{\text{эмп}}$ — это количество «нетипичных» сдвигов. Чем меньше $G_{\text{эмп}}$, тем более вероятно, что сдвиг в «типичном» направлении статистически достоверен.

Гипотезы:

H_0 : преобладание типичного направления сдвига является случайным;

H_1 : преобладание типичного направления сдвига не является случайным.

В таблице 6 приложения даны критические значения критерия знаков для разных n .

Зона значимости простирается влево, в сторону более низких значений, поскольку чем меньше «нетипичных» знаков, тем достовернее «типичный» сдвиг. Зона незначимости, напротив, простирается вправо, в сторону более высоких значений G . Постепенно «нетипичных» сдвигов становится так много, что теряется само ощущение какого-то преобладания в направленности сдвигов. Зона незначимости характеризует ситуацию, когда сдвиги обоих направлений перемешаны.

Ограничения критерия знаков. Количество наблюдений в обоих замерах – не менее 5 и не более 300.

Алгоритм расчёта критерия знаков G

1. Подсчитать количество нулевых реакций и исключить их из рассмотрения. В результате n уменьшится на число нулевых реакций.

2. Определить преобладающее направление изменений. Считать сдвиги в преобладающем направлении «типичными».

3. Определить количество «нетипичных» сдвигов. Считать это число эмпирическим значением G ($G_{\text{эмп}}$).

4. По таблице 6 приложения определить критические значения G для данного n ($G_{\text{кр}}$).

5. Сопоставить $G_{\text{эмп}}$ с $G_{\text{кр}}$. Если $G_{\text{эмп}}$ меньше $G_{\text{кр}}$ или равен ему, сдвиг в типичную сторону может считаться достоверным.

***T*-критерий Вилкоксона**

Назначение критерия. Критерий применяется для сопоставления показателей, измеренных в двух разных условиях на одной и той же выборке испытуемых.

Он позволяет установить не только *направленность* изменений, но и их *выраженность*. С его помощью мы определяем, является ли сдвиг показателей в каком-то одном направлении более интенсивным, чем в другом.

Описание критерия T . Этот критерий применим в тех случаях, когда признаки измерены по крайней мере по шкале порядка и сдвиги между вторым и первым замерами тоже могут быть упорядочены. Для этого они должны варьировать в достаточно широком диапазоне. Критерий T можно применять и в тех случаях, когда сдвиги принимают только три значения: -1 , 0 и $+1$, но тогда критерий T вряд ли

добавит что-нибудь новое к тем выводам, которые можно получить с помощью критерия знаков. Но если сдвиги изменяются, скажем, от -30 до $+45$, тогда имеет смысл их ранжировать и потом суммировать ранги.

Суть метода состоит в том, что мы сопоставляем выраженность сдвигов в том и ином направлении по абсолютной величине. Для этого мы сначала ранжируем все абсолютные величины сдвигов, а потом суммируем ранги. Если сдвиги в положительную и в отрицательную сторону происходят случайно, то суммы рангов абсолютных значений их будут примерно равны. Если же интенсивность сдвига в одном из направлений перевешивает, то сумма рангов абсолютных значений сдвигов в противоположную сторону будет значительно меньше, чем это могло бы быть при случайных изменениях.

Первоначально мы исходим из предположения о том, что типичным сдвигом будет сдвиг в наиболее часто встречающемся направлении, а нетипичным, или редким, – сдвиг в наиболее редко встречающемся направлении.

Гипотезы:

H_0 : интенсивность сдвигов в типичном направлении не превосходит интенсивности сдвигов в нетипичном направлении;

H_1 : интенсивность сдвигов в типичном направлении превышает интенсивность сдвигов в нетипичном направлении.

Ограничения в применении критерия T Вилкоксона

1. Минимальное количество испытуемых, прошедших измерения в двух условиях, 5 человек. Максимальное количество испытуемых – 50 человек, что диктуется верхней границей имеющихся таблиц. Критические значения T приведены в таблице 7 приложения.

2. Нулевые сдвиги из рассмотрения исключаются, и количество наблюдений n уменьшается на количество этих нулевых сдвигов. Можно обойти это ограничение, сформулировав гипотезы, включающие отсутствие изменений, например: «Сдвиг в сторону увеличения значений превышает сдвиг в сторону уменьшения значений и имеет место тенденция сохранения их на прежнем уровне».

Алгоритм подсчёта критерия T Вилкоксона

1. Составить список испытуемых в любом порядке, например алфавитном.

2. Вычислить разность между индивидуальными значениями во втором и первом замерах («после» – «до»). Определить, что будет

считаться «типичным» сдвигом, и сформулировать соответствующие гипотезы.

3. Перевести разности в абсолютные величины и записать их отдельным столбцом (иначе трудно отвлечься от знака разности).

4. Проранжировать абсолютные величины разностей, начисляя меньшему значению меньший ранг. Проверить совпадение полученной суммы рангов с расчётной.

5. Отметить ранги, соответствующие сдвигам в «нетипичном» направлении.

6. Подсчитать сумму этих рангов по формуле: $T = \sum R_r$, где R_r – ранговые значения сдвигов с более редким знаком.

7. Определить критические значения T для данного n по таблице 7 приложения. Если $T_{\text{эмп}}$ меньше или равен $T_{\text{кр}}$, сдвиг в «типичную» сторону по интенсивности достоверно преобладает.

Критерий χ_r^2 Фридмана

Назначение критерия. Критерий χ_r^2 применяется для сопоставления показателей, измеренных в *трех или более* условиях на одной и той же выборке испытуемых.

Критерий позволяет установить, что величины показателей от условия к условию *изменяются*, но при этом не указывает на направление изменений.

Описание критерия. Данный критерий является распространением критерия T Вилкоксона на большее, чем 2, количество условий измерения. Однако здесь мы ранжируем не абсолютные величины сдвигов, а сами индивидуальные значения, полученные данным испытуемым в 1, 2, 3 (и т. д.) замерах.

Например, если у испытуемого в первом замере определена скорость прохождения графического лабиринта 54 с, во втором замере – 42 с, а в третьем замере – 63 с, то эти показатели получают ранги, соответственно, 2, 1, 3, поскольку меньшему значению, полученному во втором замере, мы начислим ранг 1, среднему, полученному в первом замере, – ранг 2, а наибольшему значению, полученному в третьем замере, – ранг 3.

После того, как все значения будут проранжированы, подсчитываются суммы рангов по столбцам для каждого из произведённых замеров.

Если различия между значениями признака, полученными в разных условиях, случайны, то суммы рангов по разным условиям бу-

дуг приблизительно равны. Но если значения признака изменяются в разных условиях каким-то закономерным образом, то в одних условиях будут преобладать высокие ранги, а в других низкие. Суммы рангов будут достоверно различаться между собой. Эмпирическое значение критерия χ_r^2 и указывает на то, насколько различаются суммы рангов. Чем больше эмпирическое значение χ_r^2 , тем более существенные расхождения сумм рангов оно отражает.

Если χ_r^2 равняется критическому значению или превышает его, различия статистически достоверны.

Гипотезы:

H_0 : между показателями, полученными (измеренными) в разных условиях, существуют лишь случайные различия;

H_1 : между показателями, полученными в разных условиях, существуют неслучайные различия.

Ограничения критерия

1. Нижний порог: не менее двух испытуемых ($n \geq 2$), каждый из которых прошел не менее 3-х замеров ($c \geq 3$).

2. При $c = 3$, $n \leq 9$ уровень значимости полученного эмпирического значения χ_r^2 определяется по таблице 8а приложения; при $c = 4$, $n \leq 4$ – по таблице 8б приложения; при большом количестве испытуемых или условий полученные эмпирические значения χ_r^2 сопоставляются с критическими значениями χ^2 определяемыми по таблице 4 приложения. Это объясняется тем, что χ_r^2 имеет распределение, сходное с распределением χ^2 . Число степеней свободы ν определяется по формуле $\nu = c - 1$, где c – количество условий измерения (замеров).

Алгоритм подсчёта критерия χ_r^2 Фридмана

1. Проранжировать индивидуальные значения первого испытуемого, полученные им в 1, 2, 3 (и т. д.) замерах.

2. Прodelать то же самое в отношении всех другим испытуемых.

3. Просуммировать ранги по условиям, в которых осуществлялись замеры. Проверить совпадение общей суммы рангов с расчётной суммой.

4. Определить эмпирическое значение χ_r^2 по формуле

$$\chi_r^2 = \left[\frac{12}{n \cdot c \cdot (c + 1)} \cdot \sum (T_j^2) \right] - 3 \cdot n \cdot (c + 1)$$

где c – количество условий;

n – количество испытуемых;

T_j – суммы рангов по каждому из условий.

5. Определить уровни статистической значимости для χ^2

а) при $c = 3$, $n \leq 9$ – по таблице 8а;

б) при $c = 4$, $n \leq 4$ – по таблице 8б.

6. При большем количестве условий и/или испытуемых определить количество степеней свободы v по формуле: $v = c - 1$, где c – количество условий (замеров). По таблице 4 приложения определить критические значения критерия χ^2 при данном числе степеней свободы v . Если χ^2 равен критическому значению χ^2 или превышает его, различия достоверны.

L-критерий тенденций Пейджа

Назначение критерия. Критерий L Пейджа применяется для сопоставления показателей, измеренных в *трёх и более* условиях на *одной* и той же выборке испытуемых.

Критерий позволяет выявить *тенденции* в изменении величин признака при переходе от условия к условию. Его можно рассматривать как продолжение теста Фридмана, поскольку он не только констатирует различия, но и указывает на *направление* изменений.

Описание критерия тенденций L . Критерий позволяет проверить наши предположения об определенной возрастной или ситуативно обусловленной динамике тех или иных признаков. Он позволяет объединить несколько произведённых замеров единой гипотезой о тенденции изменения значений признака при переходе от замера к замеру. Если бы не ограничения, критерий был бы незаменим в «продольных», или лонгитюдинальных, исследованиях.

К сожалению, имеющиеся таблицы критических значений рассчитаны только на небольшую выборку ($n \leq 12$) и ограниченное количество сопоставляемых замеров ($c \leq 6$).

Если эти ограничения не выполняются, приходится использовать критерий χ^2 Фридмана.

В критерии L применяется такое же ранжирование условий по каждому испытуемому, как и в критерии χ^2 . После того, как значения всех испытуемых будут проранжированы, подсчитываются суммы рангов по каждому условию. Затем все условия располагаются в порядке возрастания ранговых сумм: на первом месте слева окажется условие с меньшей ранговой суммой, за ним – условие со следующей по величине ранговой суммой и так далее, пока справа не окажется условие с самой большой ранговой суммой. Затем мы с помощью специальной формулы подсчёта L проверяем, действительно ли значения возрастают слева направо. Эмпирическое значе-

ние критерия L отражает степень различия между ранговыми суммами, поэтому чем выше значение L , тем существеннее различия.

Гипотезы:

H_0 : увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему, и далее, случайно;

H_1 : увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему, и далее, неслучайно.

При формулировании гипотез мы имеем в виду новую нумерацию условий, соответствующую предполагаемым тенденциям.

Ограничения критерия Пейджа

1. Нижний порог – 2 испытуемых, каждый из которых прошёл не менее 3 замеров в разных условиях. Верхний порог – 12 испытуемых и 6 условий ($n \leq 12$, $c \leq 6$). Критические значения критерия L предусматривают три уровня статистической значимости: $p \leq 0,05$; $p \leq 0,01$; $p \leq 0,001$.

2. Необходимым условием применения теста является упорядоченность столбцов данных: слева должен располагаться столбец с наименьшей ранговой суммой показателей, справа – с наибольшей. Можно просто пронумеровать заново все столбцы, а потом вести расчёты не слева направо, а по номерам, но так легко запутаться.

Исследуются не среднегрупповые тенденции, а степень *совпадения индивидуальных тенденций*. Важен именно порядок, а не абсолютные показатели времени. Поэтому и формулируемые нами гипотезы – это гипотезы о тенденциях изменения индивидуальных показателей.

Алгоритм подсчёта критерия тенденций L Пейджа

1. Проранжировать индивидуальные значения первого испытуемого, полученные им в 1, 2, 3 (и т. д.) замерах. При этом первым может быть любой испытуемый, например тот, чьё имя стоит первым по алфавиту.

2. Прodelать то же самое по отношению ко всем другим испытуемым.

3. Просуммировать ранги по условиям, в которых осуществлялись замеры. Проверить совпадение общей суммы рангов с расчётной суммой.

4. Расположить все условия в порядке возрастания их ранговых сумм в таблице.

5. Определить эмпирическое значение L по формуле
$$L = \sum (T_j \cdot j),$$

где T_j – сумма рангов по данному условию;

j – порядковый номер, приписанный данному условию в упорядоченной последовательности условий.

По таблице 9 приложения определить критические значения L для данного количества испытуемых n и данного количества условий s . Если $L_{эмп}$ равен критическому значению или превышает его, тенденция достоверна.

Задача 1

В группе студентов проводится тренинг по коррекции уровня тревожности. Показатели уровня тревожности до тренинга и после тренинга представлены в таблице 5.

Используя G -критерий знаков, узнать, произошли ли изменения уровня тревожности у студентов в результате тренинга.

Таблица 5

Показатели уровня тревожности до и после тренинга

Вариант											
1		2		3		4		5		6	
До	После	До	После	До	После	До	После	До	После	До	После
51	49	51	44	48	50	49	48	51	44	47	48
49	45	48	49	49	46	51	44	49	50	52	49
49	45	50	45	47	49	49	47	52	50	52	47
50	46	49	47	52	50	51	45	49	47	52	44
46	48	46	46	49	48	50	44	51	45	51	45
51	47	47	46	46	49	49	48	47	50	51	44
46	50	50	46	46	49	50	50	51	44	50	44
48	45	47	47	47	47	50	47	47	45	46	47
51	47	50	46	51	48	50	48	52	46	50	46
49	46	47	45	49	50	51	50	52	47	47	44
46	44	51	45	50	50	52	47	49	48	52	50
50	49	52	45	50	44	49	49	52	48	51	50
46	50	46	50	47	48	52	50	46	49	47	49
52	49	48	45	46	46	48	49	47	44	52	48
49	46	46	48	49	47	50	50	47	49	47	46
50	47	47	49	46	47	47	48	50	48	49	44
46	47	49	48	48	47	50	49	48	50	46	44
52	46	46	47	51	45	50	48	50	50	52	50
49	44	50	44	50	49	52	48	46	46	47	48
51	45	49	48	47	45	51	47	46	50	49	45
46	50	52	44	48	49	51	47	48	48	52	49
47	50	48	49	51	49	47	48	49	44	49	50
50	50	49	45	49	50	49	49	52	46	52	50
52	49	46	44	51	50	48	45	51	49	49	47
47	48	47	44	50	46	49	44	46	49	52	47
50	44	48	49	51	45	52	45	49	50	47	46
48	49	49	47	48	46	49	48	50	47	48	47
50	46	47	44	49	49	46	46	51	44	50	44
52	48	51	44	48	49	46	46	48	47	46	45
46	47	52	49	47	46	46	45	47	48	47	47

Вариант											
7		8		9		10		11		12	
До	После	До	После	До	После	До	После	До	После	До	После
46	44	50	50	52	45	52	45	51	45	48	44
48	47	46	45	50	46	50	45	47	50	52	44
46	49	49	45	46	46	48	48	46	48	50	50
52	49	49	45	50	50	50	45	51	47	49	48
50	44	48	44	48	45	52	45	52	47	51	48
47	48	51	49	51	46	52	50	49	47	47	46
48	48	47	49	51	46	49	47	50	44	51	44
51	46	48	45	52	45	52	47	49	47	47	45
49	49	47	45	52	49	49	45	47	47	51	47
47	46	50	46	46	46	52	47	47	48	49	44
47	49	51	49	47	44	52	45	48	49	50	45
50	47	48	49	47	47	50	47	49	48	46	44
46	47	49	44	52	45	50	44	46	44	49	50
46	47	49	44	50	46	49	46	48	45	46	49
49	49	48	47	49	47	52	45	49	47	46	50
52	50	51	46	48	46	48	46	47	49	46	50
49	50	48	47	46	45	51	47	47	45	50	47
48	48	51	50	48	47	50	46	52	50	48	49
48	45	47	47	51	45	47	47	51	44	46	46
50	50	50	49	49	47	48	48	46	47	50	49
46	46	50	50	51	46	46	45	48	45	47	49
51	44	47	50	49	44	52	45	48	48	50	49
48	48	51	46	52	49	51	50	51	45	46	46
50	46	52	49	51	46	46	50	51	50	51	50
47	50	48	50	47	44	52	44	49	47	48	46
48	49	46	47	46	44	47	46	51	47	51	44
52	47	48	44	49	48	50	45	49	46	51	46
52	48	52	47	46	44	46	49	49	46	49	48
48	49	46	48	49	45	51	47	49	47	50	46
46	46	50	46	50	46	50	45	50	44	48	47

Вариант											
13		14		15		16		17		18	
До	После	До	После	До	После	До	После	До	После	До	После
49	47	46	48	46	49	49	46	47	48	47	45
49	48	48	44	50	44	46	47	52	47	51	49
49	50	50	48	49	48	46	48	48	46	49	48
50	50	47	46	51	46	48	50	49	46	46	49
47	50	50	46	46	50	51	49	49	49	52	44
48	45	47	50	47	48	50	44	47	47	46	50
48	46	49	44	49	48	46	47	46	50	47	44
49	49	46	48	46	49	52	46	49	50	48	46
49	48	52	48	49	50	51	47	46	47	51	48
48	44	49	45	49	45	48	44	49	48	48	46
46	44	48	48	51	49	47	45	51	46	50	48
46	50	46	50	51	47	50	50	49	45	47	48
50	45	50	46	46	48	48	46	47	50	51	50
50	50	47	45	50	47	52	45	52	48	48	48
52	49	52	48	46	44	52	45	52	44	48	48
51	45	47	49	50	47	52	50	51	48	49	46
47	44	46	44	51	47	47	50	52	49	46	45
46	47	49	46	52	44	49	46	52	49	48	47
52	47	51	44	46	45	52	47	48	49	46	47
52	44	47	48	46	45	52	49	49	46	47	45
52	45	47	48	49	50	47	45	49	45	51	44
50	44	46	49	49	45	48	44	47	48	50	45
46	48	49	50	50	45	48	44	49	44	51	46
50	47	46	48	48	47	49	50	50	47	50	49
48	47	48	48	51	49	48	44	48	47	52	46
48	49	47	48	48	48	47	46	48	50	51	46
50	44	48	45	47	48	47	49	50	49	50	44
46	48	49	45	51	47	46	45	52	44	48	45
52	50	52	49	51	49	50	44	51	44	52	44
51	48	51	47	48	45	46	46	51	50	47	44

Задача 2

Психолог проводит с младшими школьниками коррекционную работу по формированию навыков внимания, используя для оценки результатов корректурную пробу. Задача состоит в том, чтобы, используя критерий T Вилкоксона, определить, будет ли уменьшаться количество ошибок внимания у младших школьников после специальных коррекционных упражнений.

В таблице 6 приведены соответствующие экспериментальные данные.

Таблица 6

Результаты корректурных проб при коррекционной работе по формированию навыков внимания

Вариант 1			Вариант 2			Вариант 3			Вариант 4		
№	До	После	№	До	После	№	До	После	№	До	После
1	24	22	1	27	21	1	17	11	1	14	11
2	12	12	2	32	26	2	15	13	2	17	18
3	42	41	3	43	47	3	28	25	3	28	27
4	30	31	4	24	19	4	32	26	4	27	27
5	40	32	5	34	36	5	22	15	5	23	26
6	55	44	6	35	30	6	30	26	6	33	36
7	50	50	7	22	15	7	14	11	7	19	19
8	52	32	8	26	28	8	42	40	8	35	29
9	50	32	9	32	28	9	24	26	9	33	27
10	22	21	10	35	34	10	18	13	10	29	21
11	33	34	11	25	21	11	14	16	11	36	32
12	78	56	12	38	39	12	25	25	12	43	44
13	79	78	13	43	42	13	14	17	13	32	26
14	25	23	14	43	44	14	42	36	14	15	16
15	28	22	15	24	27	15	44	42	15	16	11
16	16	12	16	19	17	16	29	23	16	18	12
17	17	16	17	21	24	17	38	38	17	35	29
18	12	18	18	24	22	18	18	20	18	16	8
19	25	25	19	21	25	19	16	13	19	43	38
			20	33	32	20	21	20	20	22	25
			21	44	40	21	42	46			
			22	16	12						
			23	22	23						

Продолжение таблицы 6

Вариант 5			Вариант 6			Вариант 7			Вариант 8		
№	До	После	№	До	После	№	До	После	№	До	После
1	14	7	1	29	27	1	26	24	1	14	7
2	21	18	2	34	32	2	39	31	2	39	35
3	40	36	3	22	20	3	33	35	3	23	25
4	36	31	4	35	28	4	43	37	4	42	36
5	15	15	5	39	33	5	22	15	5	39	33
6	29	24	6	18	21	6	26	28	6	35	38
7	24	20	7	17	17	7	39	41	7	44	46
8	33	35	8	44	45	8	23	15	8	43	44
9	21	15	9	43	45	9	37	40	9	21	15
10	18	11	10	34	30	10	26	26	10	39	31
11	42	38	11	16	20	11	23	16	11	40	39
12	20	22	12	19	22	12	21	24	12	17	14
13	20	14	13	15	17	13	44	42	13	24	21
14	30	31	14	30	30	14	19	20	14	42	38
15	14	10	15	33	27	15	30	26	15	38	32
16	37	33	16	23	18	16	22	24	16	32	26
17	25	26	17	39	38	17	37	39	17	23	25
18	35	37	18	39	41	18	18	20	18	28	23
19	43	36	19	27	19	19	18	19	19	31	25
20	40	43				20	37	29			
						21	24	17			
						22	26	25			

Продолжение таблицы 6

Вариант 9			Вариант 10			Вариант 11			Вариант 12		
№	До	После	№	До	После	№	До	После	№	До	После
1	23	19	1	19	14	1	41	44	1	27	22
2	30	22	2	18	14	2	24	26	2	41	45
3	44	42	3	27	23	3	15	13	3	34	36
4	14	8	4	26	22	4	32	34	4	33	34
5	36	29	5	44	37	5	43	45	5	33	28
6	39	31	6	26	27	6	43	40	6	16	17
7	30	34	7	43	41	7	36	28	7	21	25
8	41	40	8	39	42	8	42	40	8	41	41
9	15	17	9	28	28	9	29	26	9	14	9
10	15	13	10	21	20	10	22	25	10	39	40
11	16	16	11	18	13	11	20	21	11	17	19
12	14	16	12	39	31	12	17	15	12	37	30
13	36	39	13	28	24	13	24	28	13	37	34
14	20	17	14	24	18	14	34	30	14	36	31
15	35	37	15	31	33	15	15	18	15	18	15
16	41	38	16	39	42	16	23	22	16	44	43
17	26	19	17	17	20	17	18	12	17	36	36
18	33	29	18	18	16	18	25	21	18	15	10
19	33	25	19	23	21	19	15	13	19	41	34
20	25	27	20	22	16	20	18	12	20	15	9
21	31	35	21	31	34	21	37	36	21	32	35
			22	17	18				22	35	39

Окончание таблицы 6

Вариант 13			Вариант 14			Вариант 15			Вариант 16		
№	До	После	№	До	После	№	До	После	№	До	После
1	16	9	1	32	33	1	24	22	1	32	27
2	37	41	2	37	34	2	33	31	2	29	25
3	28	30	3	26	26	3	30	28	3	24	18
4	24	26	4	42	43	4	38	36	4	33	27
5	22	21	5	41	37	5	35	36	5	20	18
6	30	29	6	36	39	6	29	23	6	32	33
7	37	36	7	19	19	7	20	21	7	24	26
8	20	15	8	36	39	8	38	31	8	42	45
9	34	27	9	20	15	9	32	26	9	34	31
10	24	18	10	16	13	10	34	34	10	42	46
11	28	22	11	35	31	11	30	26	11	29	33
12	31	30	12	20	16	12	18	12	12	21	25
13	16	9	13	30	23	13	29	32	13	39	43
14	42	34	14	18	22	14	35	31	14	35	32
15	34	30	15	44	47	15	43	45	15	43	39
16	41	39	16	21	14	16	34	29	16	21	25
17	26	30	17	32	31	17	30	31	17	19	14
18	39	35	18	32	24	18	30	30	18	23	15
19	19	21	19	16	13	19	33	34	19	44	47
20	30	24				20	15	16	20	21	19
21	17	15				21	18	12	21	39	34
22	36	30				22	29	29			
23	34	31									

Задача 3

Семи школьникам предъявляют тест Равенна. Фиксируется время решения каждого задания (А, В, С). Выясняется вопрос: будут ли найдены статистически значимые различия между временем решения первых трёх заданий теста (Критерий Фридмана).

В таблице 7 приведены соответствующие экспериментальные данные.

Таблица 7

Время решения тестовых заданий

Вариант 1				Вариант 2				Вариант 3			
№	А	В	С	№	А	В	С	№	А	В	С
1	11	15	13	1	10	12	10	1	8	14	10
2	11	13	13	2	10	12	15	2	10	14	14
3	12	13	9	3	9	13	14	3	10	12	13
4	9	13	14	4	7	16	14	4	12	12	13
5	9	15	10	5	8	14	10	5	12	14	14
6	8	16	13	6	11	17	15	6	11	16	12
7	8	14	10	7	9	14	15	7	8	12	14
Вариант 4				Вариант 5				Вариант 6			
№	А	В	С	№	А	В	С	№	А	В	С
1	11	15	9	1	7	15	11	1	12	15	10
2	11	15	10	2	10	12	14	2	9	13	12
3	10	11	14	3	9	15	13	3	10	14	13
4	10	11	12	4	11	16	11	4	8	12	12
5	8	11	11	5	13	16	12	5	8	15	10
6	8	16	12	6	7	13	13	6	13	15	11
7	8	17	13	7	11	13	11	7	7	12	10
Вариант 7				Вариант 8				Вариант 9			
№	А	В	С	№	А	В	С	№	А	В	С
1	10	11	13	1	12	16	10	1	10	14	15
2	9	14	14	2	8	16	9	2	9	10	12
3	7	15	12	3	10	15	13	3	9	12	13
4	7	14	13	4	12	14	9	4	10	11	17
5	8	11	14	5	8	15	15	5	11	11	14
6	11	16	11	6	11	15	13	6	8	13	12
7	11	16	13	7	13	11	13	7	8	14	17

Вариант 10				Вариант 11				Вариант 12			
№	A	B	C	№	A	B	C	№	A	B	C
1	14	12	9	1	11	11	9	1	16	12	9
2	16	12	9	2	13	14	10	2	13	12	9
3	13	14	7	3	12	10	9	3	16	14	11
4	13	15	10	4	15	9	10	4	16	12	11
5	14	10	10	5	14	11	7	5	16	15	12
6	12	14	7	6	12	11	9	6	11	11	11
7	14	10	7	7	13	9	12	7	15	15	12
Вариант 13				Вариант 14				Вариант 15			
№	A	B	C	№	A	B	C	№	A	B	C
1	12	12	15	1	11	8	15	1	13	9	12
2	14	10	13	2	12	12	16	2	13	11	11
3	12	7	15	3	15	7	13	3	12	8	13
4	12	11	17	4	11	9	16	4	13	9	16
5	14	10	11	5	13	9	16	5	15	10	15
6	11	8	14	6	12	11	16	6	14	12	12
7	12	8	17	7	11	12	13	7	11	8	11

Задача 4

Продолжая исследования, психолог решил проверить гипотезу: действительно ли время решения задач (см. данные задачи 3) увеличивается (Критерий Пейджа).

Лабораторное занятие 4

ВЫЯВЛЕНИЕ РАЗЛИЧИЙ В РАСПРЕДЕЛЕНИИ ПРИЗНАКА

Методические указания

χ^2 – критерий Пирсона

Назначения критерия. Критерий χ^2 применяется в двух целях;

- 1) для сопоставления *эмпирического* распределения признака с теоретическим – равномерным, нормальным или каким-то иным;
- 2) для сопоставления двух, трёх или более эмпирических распределений одного и того же признака.

Описание критерия. Критерий χ^2 отвечает на вопрос: с одинаковой ли частотой встречаются разные значения признака в эмпирическом и теоретическом распределениях или в двух и более эмпирических распределениях.

Преимущество метода состоит в том, что он позволяет сопоставлять распределения признаков, представленных в любой шкале, начиная от шкалы наименований. В самом простом случае альтернативного распределения «да – нет», «допустил брак – не допустил брака», «решил задачу – не решил задачу» и др. уже можно применить критерий χ^2 .

При сопоставлении эмпирического распределения с теоретическим мы определяем степень расхождения между эмпирическими и теоретическими частотами.

При сопоставлении двух эмпирических распределений мы определяем степень расхождения между эмпирическими частотами и теоретическими частотами, которые наблюдались бы в случае совпадения двух этих эмпирических распределений (формулы расчета теоретических частот будут специально даны для каждого варианта сопоставлений).

Чем больше расхождение между двумя сопоставляемыми распределениями, тем больше эмпирическое значение χ^2 .

Гипотезы. Возможны несколько вариантов гипотез – в зависимости от задач, которые мы перед собой ставим.

Первый вариант

H_0 : полученное эмпирическое распределение признака не отличается от теоретического (например равномерного) распределения;

H_1 : полученное эмпирическое распределение признака отличается от теоретического распределения.

Второй вариант

H_0 : эмпирическое распределение 1 не отличается от эмпирического распределения 2;

H_1 : эмпирическое распределение 1 отличается от эмпирического распределения 2.

Третий вариант

H_0 : эмпирические распределения 1, 2, 3, ... не различаются между собой;

H_1 : эмпирические распределения 1, 2, 3, ... различаются между собой.

Критерий χ^2 позволяет проверить все три варианта гипотез.

Ограничения критерия

1. Объём выборки должен быть достаточно большим: $n \geq 30$. При $n < 30$ критерий χ^2 даёт весьма приближённые значения. Точность критерия повышается при больших n .

2. Теоретическая частота для каждой ячейки таблицы не должна быть меньше 5: $f \geq 5$. Это означает, что если число разрядов задано заранее и не может быть изменено, то мы не можем применять метод χ^2 , не имея определённого минимального числа наблюдений. Если, например, мы хотим проверить наши предположения о том, что частота обращений в телефонную службу доверия неравномерно распределяется по 7 дням недели, то нам потребуется $5 \cdot 7 = 35$ обращений. Таким образом, если количество разрядов (k) задано заранее, как в данном случае, минимальное число наблюдений (n_{min}) определяется по формуле $n_{min} = k \cdot 5$.

3. Выбранные разряды должны «вычерпывать» всё распределение, то есть охватывать весь диапазон вариативности признаков. При этом группировка на разряды должна быть одинаковой во всех сопоставляемых распределениях.

4. Необходимо вносить «поправку на непрерывность» при сопоставлении распределений признаков, которые принимают всего 2 значения. При внесении поправки значение χ^2 уменьшается.

5. Разряды должны быть непереключающимися: если наблюдение отнесено к одному разряду, то оно уже не может быть отнесено ни к какому другому разряду.

Сумма наблюдений по разрядам всегда должна быть равна общему количеству наблюдений.

Алгоритм расчёта критерия χ^2

1. Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты ($f_{эj}$) (первый столбец).

2. Рядом с каждой эмпирической частотой записать теоретическую частоту (f_{Tj}) (второй столбец).

3. Подсчитать разности между эмпирической и теоретической частотой по каждому разряду (строке) и записать их в третий столбец ($f_{эj} - f_{Tj}$).

4. Возвести в квадрат полученные разности, и занести их в четвертый столбец $(f_{эj} - f_{Tj})^2$.

5. Разделить полученные квадраты разностей на теоретическую частоту и записать результаты в пятый столбец.

6. Просуммировать значения пятого столбца. Полученную сумму обозначить как $\chi^2_{эмп}$:

$$\chi^2_{эмп} = \sum_{j=1}^k \frac{(f_{эj} - f_{Tj})^2}{f_{Tj}}$$

7. Определить число степеней свободы по формуле: $\nu = k - 1$, где k – количество разрядов признака. Если $\nu = 1$, внести поправку на «непрерывность». Все расчёты производятся по предложенному алгоритму, эмпирическое значение критерия вычисляют по формуле

$$\chi^2_{эмп} = \sum_{j=1}^k \frac{(f_{эj} - f_{Tj} - 0,5)^2}{f_{Tj}}.$$

8. Определить по таблице 4 приложения критические значения для данного числа степеней свободы ν . Различия между двумя распределениями могут считаться статистически достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,05}$, и ещё более достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,01}$.

При сопоставлении двух и более эмпирических распределений (второй, третий варианты гипотез) все данные объединяют в одну выборку, и все расчёты выполняют согласно рассмотренному алгоритму. При этом отдельно выделяется задача нахождения теоретических частот. В общем случае число значений в первой и второй выборке различно и требуется учитывать эту пропорцию. Общая формула подсчёта $f_{теор} = f_{T,i,j}$ для сопоставления двух или более эмпирических распределений

$$f_{Ti,j} = \frac{\sum_j f_{i,j} \sum_i f_{i,j}}{\sum_i \sum_j f_{i,j}},$$

где i – номер разряда в выборке;

j – номер выборки.

Если представить исходные данные в виде таблицы, где значения каждой выборки располагаются по столбцам, то в числителе полученной формулы будем иметь: *Сумма частот по соответствующей строке* × *Сумма частот по соответствующему столбцу*. В знаменателе – *Общее количество наблюдений*.

Число степеней свободы при сопоставлении эмпирических распределений определяется по формуле

$$v = (k - 1)(c - 1),$$

где k – количество разрядов признака (строк в таблице эмпирических частот);

c – количество сравниваемых распределений (столбцов в таблице эмпирических частот).

Задача 1

Перед героиней комедии Н. В. Гоголя «Женитьба» Агафьей Тихоновной стоит задача выбора между пятью женихами. Причём одного она сразу исключила из рассмотрения, потому что он был купеческого звания, как и она сама.

Допустим, за полчаса смотрин ею зафиксированы следующие наблюдения.

Агафья Тихоновна:

сидела с опущенными глазами	25 минут;
благосклонно смотрела на Никанора Ивановича	k_1 раз;
благосклонно смотрела на Ивана Кузьмина	k_2 раз;
благосклонно смотрела на Ивана Павловича	k_3 раз;
благосклонно смотрела на Балтазара Балтазарыча	k_4 раз.

Проверить отличаются ли предпочтения Агафи Тихоновны от равномерного распределения. Данные по вариантам приведены в таблице 8.

Таблица 8

Вариант (№)	1	2	3	4	5	6	7	8
Н.И. (k_1)	12	6	3	12	5	6	7	4
И.К. (k_2)	14	8	8	5	6	4	9	9
И.П. (k_3)	4	12	7	11	8	4	7	11
Б.Б. (k_4)	3	8	6	7	11	9	5	9
Вариант (№)	9	10	11	12	13	14	15	16
Н.И. (k_1)	10	14	13	3	11	4	4	5
И.К. (k_2)	3	12	7	8	6	12	11	14
И.П. (k_3)	3	7	13	7	3	9	10	6
Б.Б. (k_4)	10	3	9	9	12	3	5	5

Задача 2.

Используя условие предыдущей задачи, ответить на вопрос: одинаковая ли система предпочтений во взглядах Агафьи Тихоновны и её словах? Данные по вариантам приведены в таблице 9.

Таблица 90

Вариант	1		2		3		4	
Женихи	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.
Н.Н.	12	12	6	8	3	17	12	6
И.К.	14	8	8	20	8	22	5	16
И.П.	4	17	12	11	7	9	11	17
Б.Б.	3	22	8	5	6	10	7	6
Вариант	5		6		7		8	
Женихи	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.
Н.Н.	6	20	10	24	13	7	11	13
И.К.	3	5	3	10	6	16	13	16
И.П.	13	16	13	13	11	23	13	5
Б.Б.	6	23	13	16	3	15	10	15

Вариант	9		10		11		12	
Женихи	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.
Н.Н.	5	19	5	10	9	5	11	12
И.К.	9	14	8	9	5	5	10	20
И.П.	12	4	3	6	13	14	8	18
Б.Б.	4	11	6	14	11	13	9	21
Вариант	13		14		15		16	
Женихи	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.	Взгл.	Упом. в разг.
Н.Н.	10	6	14	12	4	21	6	23
И.К.	11	5	8	7	9	8	8	20
И.П.	7	18	6	18	6	4	3	15
Б.Б.	6	23	6	7	2	7	9	16

Методические указания

λ-критерий Колмогорова–Смирнова

Назначение критерия. Критерий предназначен для сопоставления двух распределений:

- а) *эмпирического с теоретическим*, например, равномерным или нормальным;
- б) *одного эмпирического распределения с другими эмпирическим распределением.*

Критерий позволяет найти точку, в которой сумма накопленных *расхождений* между двумя распределениями является наибольшей, и оценить достоверность этого расхождения.

Описание критерия. Если в методе χ^2 мы сопоставляли частоты двух распределений отдельно по каждому разряду, то здесь мы сопоставляем сначала частоты по первому разряду, потом по сумме первого и второго разрядов, потом по сумме первого, второго и третьего разрядов, и т. д. Таким образом, мы *сопоставляем* всякий раз *накопленные к данному разряду частоты*.

Если различия между двумя распределениями существенны, то в какой-то момент разность накопленных частот достигнет критического значения, и мы сможем признать различия статистически дос-

товерными. В формулу критерия λ включается эта разность. Чем больше эмпирическое значение λ , тем более существенны различия.

Гипотезы:

H_0 : различия между двумя распределениями недостоверны (судя по точке максимального накопленного расхождения между ними);

H_1 : различия между двумя распределениями достоверны (судя по точке максимального накопленного расхождения между ними).

Ограничения критерия λ .

1. Критерий требует, чтобы выборка была достаточно большой. При сопоставлении двух эмпирических распределений необходимо, чтобы $n_{1,2} \geq 50$. Сопоставление эмпирического распределения с теоретическим иногда допускается при $n \geq 5$.

2. Разряды должны быть упорядочены по нарастанию или убыванию какого-либо признака. Они обязательно должны отражать какое-то однонаправленное его изменение. Например, мы можем за разряды принять дни недели, 1-й, 2-й, 3-й месяцы после прохождения курса терапии, повышение температуры тела, усиление чувства недостаточности и т. д. Если мы возьмем разряды, которые случайно оказались выстроенными в данную последовательность, то и накопление частот будет отражать лишь этот элемент случайного соседства разрядов. Например, если шесть стимульных картин в методике Хекхаузена разным испытуемым предъявляются в разном порядке, мы не вправе говорить о накоплении реакций при переходе от картины № 1 стандартного набора к картине № 2 и т. д. Мы не можем говорить об однонаправленном изменении признака при сопоставлении категорий «очередность рождения», «национальность», «специфика полученного образования» и т. п. Эти данные представляют собой номинативные шкалы: в них нет никакого однозначного однонаправленного изменения признака.

Итак, мы не можем накапливать частоты по разрядам, которые отличаются лишь качественно и не представляют собой шкалы порядка.

Во всех случаях, когда разряды представляют собой не упорядоченные по возрастанию или убыванию какого-либо признака категории, нам следует применять метод χ^2 .

Алгоритм расчёта абсолютной величины разности d между эмпирическим и равномерным распределениями

1. Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты (первый столбец).

2. Подсчитать относительные эмпирические частоты (частости) для каждого разряда по формуле

$$f_{\text{эм}}^* = f_{\text{эм}} / n,$$

где $f_{\text{эм}}^*$ – эмпирическая частота по данному разряду;
 n – общее количество наблюдений.

Занести результаты во второй столбец.

3. Подсчитать накопленные эмпирические частоты $\sum f_j^*$ по формуле

$$F^* = \sum f_j^* = \sum f_{j-1}^* + f_j^*,$$

где $\sum f_{j-1}^*$ – частость, накопленная на предыдущих разрядах;

j – порядковый номер разряда;

f_j^* – эмпирическая частость данного j -го разряда.

Занести результаты в третий столбец таблицы.

4. Подсчитать накопленные теоретические частоты для каждого разряда по формуле

$$F_T^* = \sum f_{Tj}^* = \sum f_{Tj-1}^* + f_{Tj}^*,$$

где $\sum f_{Tj-1}^*$ – теоретическая частость, накопленная на предыдущих разрядах;

j – порядковый номер разряда;

f_{Tj}^* – теоретическая частость данного разряда.

Занести результаты в четвёртый столбец таблицы.

5. Вычислить разности между эмпирическими и теоретическими накопленными частостями по каждому разряду (между значениями 3 и 4 столбца).

6. Записать в пятый столбец абсолютные величины полученных разностей. Обозначить их как d .

7. Определить по пятому столбцу наибольшую абсолютную величину разности: $d_{\text{max}} = d_{\text{эм}}$.

8. По таблице 9 приложения определить или рассчитать критические значения d_{max} для данного количества наблюдений n . Различия между распределениями могут считаться достоверными, если абсолютная величина максимальной разности $d_{\text{эм}}$ достигает или превышает $d_{0,05}$, и ещё более достоверными, если $d_{\text{эм}}$ достигает или превышает $d_{0,01}$.

Алгоритм расчёта критерия λ при сопоставлении двух эмпирических распределений

1. Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты, полученные в распределении 1 (первый столбец) и в распределении 2 (второй столбец).

2. Подсчитать эмпирические частоты по каждому разряду для распределения 1 по формуле

$$f_{\text{эмп}}^* = f_{\text{эмп}} / n_1,$$

где $f_{\text{эмп}}$ – эмпирическая частота в данном разряде;

n_1 – количество наблюдений в выборке.

Занести эмпирические частоты распределения 1 в третий столбец.

3. Подсчитать эмпирические частоты по каждому разряду для распределения 2 по формуле:

$$f_{\text{эмп}}^* = f_{\text{эмп}} / n_2,$$

где $f_{\text{эмп}}$ – эмпирическая частота в данном разряде;

n_2 – количество наблюдений во 2-й выборке.

Занести эмпирические частоты распределения 2 в четвёртый столбец таблицы.

4. Подсчитать накопленные эмпирические частоты для распределения 1 по формуле

$$F^* = \sum f_j^* = \sum f_{j-1}^* + f_j^*,$$

где $\sum f_{j-1}^*$ – частоты, накопленные на предыдущих разрядах;

j – порядковый номер разряда;

f_j^* – эмпирическая частота данного j -го разряда.

Полученные результаты записать в пятый столбец.

5. Подсчитать накопленные эмпирические частоты для распределения 2 по той же формуле и записать результат в шестой столбец.

6. Подсчитать разности между накопленными частотами по каждому разряду. Записать в седьмой столбец абсолютные величины разностей. Обозначить их как d .

7. Определить по седьмому столбцу наибольшую абсолютную величину разности d_{max} .

8. Подсчитать значение критерия λ по формуле

$$\lambda = d_{\text{max}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

где n_1 – количество наблюдений в первой выборке;
 n_2 – количество наблюдений во второй выборке.

9. По таблице 10 приложения определить, какому уровню статистической значимости соответствует полученное значение λ . Если $\lambda > \lambda_{0,05}$, различия между распределениями достоверны.

Задача 3 (сопоставление эмпирического распределения с теоретическим)

В выборке здоровых лиц мужского пола, студентов технических и военно-технических вузов в возрасте от 19 до 22 лет, средний возраст 20 лет, проводился тест Люшера в 8-цветном варианте. Установлено, что жёлтый цвет предпочитается испытуемыми чаще, чем отвергается (табл. 10). Можно ли утверждать, что распределение желтого цвета по 8 позициям у здоровых испытуемых отличается от равномерного распределения?

Таблица 10

**Эмпирические частоты попадания желтого цвета
на каждую из 8 позиций**

Вариант	Позиция жёлтого цвета							
	1	2	3	4	5	6	7	8
1	14	13	9	19	8	22	21	14
2	9	20	5	20	22	5	11	23
3	12	24	18	22	16	20	12	11
4	16	21	7	11	5	6	11	16
5	13	15	17	16	22	6	4	20
6	20	17	12	18	7	7	6	23
7	20	21	16	14	18	18	10	6
8	5	22	5	6	14	8	16	5
9	5	17	7	5	17	8	20	20
10	21	4	12	18	9	6	22	21
11	17	6	9	11	10	14	12	9
12	16	11	16	18	22	17	16	19
13	19	10	5	17	4	7	11	15
14	8	20	23	18	12	7	20	24
15	19	22	9	13	20	16	12	8
16	9	21	11	7	19	21	16	5
17	6	22	21	7	21	18	22	6
18	6	13	9	23	21	9	13	21
19	9	14	9	15	16	5	17	6
20	10	10	24	14	10	14	14	19

Задача 4 (сопоставление двух эмпирических распределений)

Сопоставить данные в двух выборках, полученных различными исследователями при проведении теста Люшера в 8-цветном варианте (таблица 121).

Таблица 12

Эмпирические частоты попадания жёлтого цвета на каждую из 8 позиций в двух исследованиях

Разряд	Вариант 1		Вариант 2		Вариант 3		Вариант 4	
	f_1	f_2	f_1	f_2	f_1	f_2	f_1	f_2
1	127	30	118	21	118	33	114	32
2	114	22	89	14	118	35	72	9
3	125	17	124	25	91	28	117	8
4	103	29	102	29	72	28	121	20
5	65	13	95	12	77	14	70	10
6	71	10	123	25	75	31	127	28
7	71	28	74	34	92	20	123	31
8	81	9	129	26	96	29	73	9
Разряд	Вариант 5		Вариант 6		Вариант 7		Вариант 8	
	f_1	f_2	f_1	f_2	f_1	f_2	f_1	f_2
1	98	14	119	24	120	14	114	36
2	106	25	108	12	99	10	123	19
3	128	35	103	37	69	26	113	15
4	113	17	107	9	122	24	113	20
5	130	29	104	14	114	25	75	8
6	78	15	108	32	96	29	128	18
7	67	28	105	23	98	15	93	22
8	64	35	128	23	107	27	78	28

Окончание таблицы 11

Разряд	Вариант 9		Вариант 10		Вариант 11		Вариант 12	
	f_1	f_2	f_1	f_2	f_1	f_2	f_1	f_2
1	120	9	79	28	116	15	105	11
2	78	16	110	35	79	15	76	8
3	110	9	90	27	85	36	81	26
4	91	27	102	13	126	30	118	26
5	68	25	92	33	68	8	122	36
6	123	28	106	24	86	12	89	21
7	127	33	101	30	123	26	99	24
8	70	18	96	31	89	29	69	31
Разряд	Вариант 13		Вариант 14		Вариант 15		Вариант 16	
	f_1	f_2	f_1	f_2	f_1	f_2	f_1	f_2
1	129	29	97	25	96	27	108	15
2	68	22	102	33	86	25	110	27
3	100	16	101	30	89	24	81	21
4	85	9	93	22	78	15	63	12
5	61	26	114	31	122	16	73	10
6	120	22	89	14	102	16	68	28
7	123	26	66	7	110	18	62	8
8	121	24	65	26	68	36	73	18

Лабораторное занятие 5

АППРОКСИМАЦИЯ ОПЫТНЫХ ДАННЫХ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

Методические указания

Пусть таблично задана некоторая функция $y_i = f(x_i), i = \overline{0, N}$. Требуется построить функцию вида

$$P(x) = c_0 + c_1x + c_2x^2 + c_3x^3 + \dots + c_Mx^M, M \leq N \text{ такую, что}$$

$$Q(x) = \sum_{i=0}^N (P(x_i) - f(x_i))^2 \rightarrow \min.$$

Неизвестные коэффициенты могут быть найдены путём решения системы линейных алгебраических уравнений из $M + 1$ уравнения с $M + 1$ неизвестными вида

$$c_0(N+1) + c_1 \sum_{i=0}^N x_i + c_2 \sum_{i=0}^N x_i^2 + c_3 \sum_{i=0}^N x_i^3 + \dots + c_M \sum_{i=0}^N x_i^M = \sum_{i=0}^N y_i,$$

$$c_0 \sum_{i=0}^N x_i + c_1 \sum_{i=0}^N x_i^2 + c_2 \sum_{i=0}^N x_i^3 + c_3 \sum_{i=0}^N x_i^4 + \dots + c_M \sum_{i=0}^N x_i^{M+1} = \sum_{i=0}^N x_i y_i,$$

.....

$$c_0 \sum_{i=0}^N x_i^M + c_1 \sum_{i=0}^N x_i^{M+1} + c_2 \sum_{i=0}^N x_i^{M+2} + c_3 \sum_{i=0}^N x_i^{M+3} + \dots + c_M \sum_{i=0}^N x_i^{2M} = \sum_{i=0}^N x_i^M y_i.$$

В матричном виде данная система имеет вид

$$A_M^T A_M C = A_M^T Y,$$

где C – вектор коэффициентов $\vec{C} = \{c_i\}, i = \overline{0, M}$,

Y – вектор значений функции $Y = \{f(x_i)\}, i = \overline{0, N}$.

Матрица A_M состоит из $M + 1$ столбца матрицы

$$\text{Вандермонда } A = \begin{pmatrix} 1 & x_0 & x_0^2 & x_0^3 & \dots & x_0^M \\ 1 & x_1 & x_1^2 & x_1^3 & \dots & x_1^M \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_N & x_N^2 & x_N^3 & \dots & x_N^M \end{pmatrix}.$$

Решение системы линейных алгебраических уравнений $(A_M^T A_M)^{-1} A_M^T \vec{Y} = \vec{C}$.

В случае $M = 1$ аппроксимационная функция представляет собой прямую линию $y = ax + b$; для нахождения её коэффициентов имеем систему из двух уравнений с двумя неизвестными:

$$b(N+1) + a \sum_{i=0}^N x_i = \sum_{i=0}^N y_i,$$

$$b \sum_{i=0}^N x_i + a \sum_{i=0}^N x_i^2 = \sum_{i=0}^N x_i y_i.$$

Очевидно, что решение задачи аппроксимации (построения такого рода преобразований) возможно в гораздо более общем случае. Именно, всякая неявная функция, заданная соотношением вида $a\varphi(x) + b\psi(y) + c = 0$, где a, b, c – постоянные, будет изображаться прямой линией на функциональной сетке, где на оси Ox построена шкала функции $\varphi(x)$, а на оси Oy – шкала функции $\psi(y)$. Разумеется, функции $\varphi(x)$ и $\psi(y)$ должны удовлетворять условиям, обеспечивающим построение функциональных шкал (непрерывность и монотонность).

Особенно часто используются различные логарифмические сетки, с помощью которых можно «выпрямлять» графики степенных и показательных функций. Если зависимость между x и y задаётся уравнением $y = ae^{bx}$, то логарифмирование его даёт $\ln(y) = bx + \ln(a)$.

Полагая $\ln(y) = Y$, $\ln(a) = A$, $b = B$ получим линейную функцию: $Y = A + Bx$. Оставив равномерную шкалу на оси Ox и построив на оси Oy логарифмическую шкалу, мы изобразим аппроксимационную функцию прямой линией.

Описанная выше сетка называется *полулогарифмической*. *Логарифмической* называется функциональная сетка, у которой на каждой из осей Ox и Oy построена логарифмическая шкала. На такой сетке графики степенных функций представляют собой прямые линии.

Действительно, если $y = ax^b$, то $\ln(y) = \ln(a) + b \ln(x)$.

Полагая теперь $\ln(y) = Y$, $\ln(a) = A$, $\ln(x) = X$, запишем уравнение в виде $Y = A + BX$.

Выбор функциональных шкал, позволяющих линеаризовать аппроксимационную функцию $a\varphi(x) + b\psi(y) + c = 0$, достаточно разнообразен. В таблице 12 приведены различные виды функций и линеаризующие их преобразования.

Таблица 12

**Различные виды аппроксимационных функций и
линеаризующие их преобразования**

№	Функция	Y	X	A	B
1	$y = a + b/x$	y	$1/x$	a	b
2	$y = 1/(a + bx)$	$1/y$	x	a	b
3	$y = x/(a + bx)$	x/y	x	a	b
4	$y = a \cdot b^x$	$\ln(y)$	x	$\ln(a)$	$\ln(b)$
5	$y = a \cdot e^{bx}$	$\ln(y)$	x	$\ln(a)$	b
6	$y = 1/(a + b \cdot e^{-x})$	$1/y$	e^{-x}	a	b
7	$y = a \cdot x^b$	$\ln(y)$	$\ln(x)$	$\ln(a)$	b
8	$y = a + b \cdot \ln(x)$	y	$\ln(x)$	a	b
9	$y = a/(b + x)$	$1/y$	x	b/a	$1/a$
10	$y = a \cdot x/(b+x)$	$1/y$	$1/x$	$1/a$	b/a
11	$y = a \cdot e^{b/x}$	$\ln(y)$	$1/x$	$\ln(a)$	b
12	$y = a + b \cdot x^n$	y	x^n	a	b

Задача 1

Аппроксимировать функцию, заданную таблично (таблица 13) с помощью функций: $y = a + b/x$; $y = ax^b$; $y = ae^{b/x}$; $y = a + bx$. Выбрать среди них наилучшую. Построить графики аппроксимационных функций и исходной функции.

Таблица 13

Варианты заданий

Вариант 1		Вариант 2		Вариант 3		Вариант 4	
X	Y	X	Y	X	Y	X	Y
1	9,22	1	10,54	1	23,88	1	6,22
2	6,72	2	25,05	2	9,12	2	8,28
3	5,48	3	39,99	3	6,52	3	9,48
4	4,88	4	57,64	4	5,34	4	11,26
5	4,72	5	71,47	5	5,02	5	12,78
6	4,44	6	90,81	6	4,50	6	13,51
7	4,18	7	106,87	7	5,86	7	15,33
8	4,63	8	125,55	8	5,67	8	16,69
9	4,19	9	141,48	9	5,74	9	18,61
10	3,79	10	161,89	10	4,28	10	19,33

Окончание таблицы 13

Вариант 5		Вариант 6		Вариант 7		Вариант 8	
X	Y	X	Y	X	Y	X	Y
2	6,942	2	23,18	2	10,00	2	7,84
3,5	5,555	3,5	49,37	3,5	5,39	3,5	9,65
5	4,587	5	69,22	5	6,20	5	12,31
6,5	4,046	6,5	94,59	6,5	5,24	6,5	14,39
8	4,146	8	122,65	8	4,49	8	16,33
9,5	4,483	9,5	152,64	9,5	4,88	9,5	18,57
11	3,745	11	179,93	11	3,76	11	21,11
12,5	3,481	12,5	208,46	12,5	4,34	12,5	22,87
14	3,655	14	238,49	14	5,39	14	25,83
15,5	4,199	15,5	271,95	15,5	4,65	15,5	28,63
Вариант 9		Вариант 10		Вариант 11		Вариант 12	
X	Y	X	Y	X	Y	X	Y
0,5	15,749	0,5	4,94	0,5	163,91	0,5	4,87
1,5	7,882	1,5	16,82	1,5	12,82	1,5	6,32
2,5	5,901	2,5	33,12	2,5	8,66	2,5	7,76
3,5	5,393	3,5	47,39	3,5	7,25	3,5	10,69
4,5	5,199	4,5	62,76	4,5	5,43	4,5	11,88
5,5	4,620	5,5	79,14	5,5	6,02	5,5	13,43
6,5	4,093	6,5	94,53	6,5	4,25	6,5	14,44
7,5	4,690	7,5	115,14	7,5	5,70	7,5	15,73
8,5	3,876	8,5	134,64	8,5	5,00	8,5	17,89
9,5	3,841	9,5	149,31	9,5	5,27	9,5	19,06
Вариант 13		Вариант 14		Вариант 15		Вариант 16	
X	Y	X	Y	X	Y	X	Y
3	9,090	3	41,52	3	7,07	3	9,52
4	7,300	4	54,01	4	5,95	4	10,22
5	7,142	5	73,04	5	5,85	5	12,68
6	6,304	6	86,38	6	4,92	6	14,30
7	6,015	7	106,45	7	5,19	7	15,69
8	5,187	8	123,83	8	5,46	8	16,08
9	5,640	9	143,26	9	4,16	9	17,86
10	4,868	10	161,36	10	5,18	10	19,89
11	5,178	11	181,42	11	4,10	11	21,01
12	4,415	12	201,63	12	4,89	12	22,20

Пример решения задачи аппроксимации в MS Excel

Вариант 0

X	Y
1	9,22
2	6,72
3	5,48
4	4,88
5	4,72
6	4,44
7	4,18
8	4,63
9	4,19
10	3,79

x	y	$X=1/x$	Y	$y=a+b/x$					
1	9,22	1,00	9,22	1,00	1,00	10	2,92897	52,2539	
2	6,72	0,50	6,72	1,00	0,50	2,928968	1,54977	19,3316	
3	5,48	0,33	5,48	1,00	0,33				
4	4,88	0,25	4,88	1,00	0,25	$a=$	3,52082		
5	4,72	0,20	4,72	1,00	0,20	$b=$	5,819707		
6	4,44	0,17	4,44	1,00	0,17				
7	4,18	0,14	4,18	1,00	0,14	$=МУМНОЖ(ТРАНСП(F15:G24);$			
						$F15:G24)$			
8	4,63	0,13	4,63	1,00	0,13	$=МУМНОЖ(ТРАНСП(F15:G24);$			
						$D15:D24)$			
9	4,19	0,11	4,19	1,00	0,11	$=МУМНОЖ(МОБР(I15:J16);$			
						$K15:K16)$			
10	3,79	0,10	3,79	1,00	0,10				

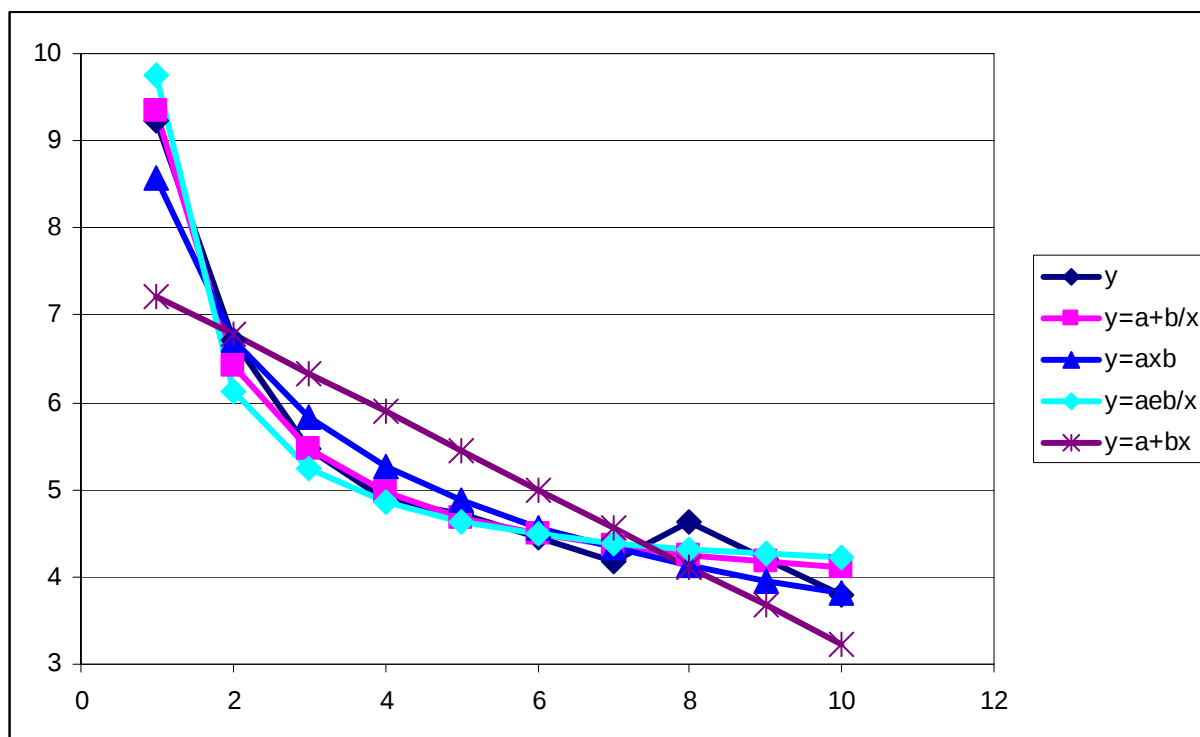
$X = \ln(x)$	$Y = \ln(y)$			$y = ax^b$			
0,00	2,22	1,00	0,00		10	15,1044	16,1841
0,69	1,91	1,00	0,69		15,10441	27,6502	22,7436
1,10	1,70	1,00	1,10				
1,39	1,59	1,00	1,39	$\ln(a) =$	2,149856	8,58362	$=a$
1,61	1,55	1,00	1,61	$b =$	-0,351848	-0,3518	$=b$
1,79	1,49	1,00	1,79				
1,95	1,43	1,00	1,95				
2,08	1,53	1,00	2,08				
2,20	1,43	1,00	2,20				
2,30	1,33	1,00	2,30				

$X = 1/x$	$Y = \ln(y)$			$y = ae^{b/x}$			
1,00	2,22	1,00	1,00		10	2,92897	16,1841
0,50	1,91	1,00	0,50		2,928968	1,54977	5,38446
0,33	1,70	1,00	0,33				
0,25	1,59	1,00	0,25	$\ln(a) =$	1,345701	3,84088	$=a$
0,20	1,55	1,00	0,20	$b =$	0,931074	0,93107	$=b$
0,17	1,49	1,00	0,17				
0,14	1,43	1,00	0,14				
0,13	1,53	1,00	0,13				
0,11	1,43	1,00	0,11				
0,10	1,33	1,00	0,10				

x	y			$y = a + bx$			
1,00	9,22	1,00	1,00		10	55	52,259
2,00	6,72	1,00	2,00		55	385	250,82
3,00	5,48	1,00	3,00				
4,00	4,88	1,00	4,00	$a =$	7,6643		
5,00	4,72	1,00	5,00	$b =$	-0,4434		
6,00	4,44	1,00	6,00				
7,00	4,18	1,00	7,00				
8,00	4,63	1,00	8,00				
9,00	4,19	1,00	9,00				
10,00	3,79	1,00	10,00				

x	y	$y=a+b/x$	$y=ax^b$	$y=ae^{b/x}$	$y=a+bx$	$y=a+b/x$	$y=ax^b$	$y=ae^{b/x}$	$y=a+bx$
1	9,22	9,341	8,584	9,745	7,221	0,015	0,405	0,276	3,996
2	6,72	6,431	6,726	6,118	6,777	0,084	0,000	0,364	0,003
3	5,48	5,461	5,832	5,239	6,334	0,000	0,125	0,058	0,732
4	4,88	4,976	5,270	4,848	5,891	0,009	0,151	0,001	1,018
5	4,72	4,685	4,872	4,627	5,447	0,001	0,024	0,008	0,533
6	4,44	4,491	4,570	4,486	5,004	0,003	0,018	0,002	0,321
7	4,18	4,352	4,328	4,387	4,560	0,029	0,021	0,042	0,142
8	4,63	4,248	4,130	4,315	4,117	0,149	0,255	0,102	0,268
9	4,19	4,167	3,962	4,260	3,673	0,000	0,050	0,005	0,263
10	3,79	4,103	3,818	4,216	3,230	0,095	0,001	0,177	0,319

Сумма квадратов уклонений: **0,386 1,049 1,036 7,595**



Наилучшая аппроксимация исходных данных представлена функцией $y=a+b/x$, так как для неё приходится наименьшая сумма квадратов уклонений.

Методические указания

Линейная корреляция. Линии регрессии

Два признака X и Y находятся в *корреляционной зависимости*, если каждому значению одного из них соответствует некоторое распределение другого. Корреляционная зависимость между признаками X и Y обычно задаётся с помощью *корреляционной таблицы*: $n = \{n_{i,j}\}$, где $n_{i,j}$ – частота, с которой встречается пара (x_i, y_j) ; $N = \sum_i \sum_j n_{i,j}$

Корреляционная зависимость между признаками X и Y может быть заменена *функциональной*, если каждому значению признака X поставить в соответствие условное среднее признака Y , т. е. $X = x_i$ поставить в соответствие величину

$$\bar{y}_{x_i} = \frac{\sum_j n_{i,j} y_j}{n_{x_i}}.$$

Если теперь точки $(x_i; \bar{y}_{x_i})$ выровнять по методу наименьших квадратов вдоль кривой $y = f(x, c_1, c_2, \dots, c_m)$, то последняя будет называется *линией регрессии Y на X* , а её уравнение – *уравнением регрессии Y на X* .

Аналогично определяется линия регрессии X на Y . Каждому значению признака $Y = y_j$ ставится в соответствие величина условного среднего

$$\bar{x}_{y_j} = \frac{\sum_i n_{i,j} x_i}{n_{y_j}}.$$

Наиболее простыми и наиболее важными случаями кривых регрессий являются прямые линии.

Угловым коэффициентом прямой регрессии Y на X (X на Y) называют *коэффициентом регрессии Y на X* (X на Y) и обозначают обычно через $\rho_{y/x}$ ($\rho_{x/y}$). Коэффициенты регрессии могут быть вычислены по формулам:

$$\rho_{x/y} = \frac{\overline{XY} - \bar{X}\bar{Y}}{\sigma_x^2}; \quad \rho_{y/x} = \frac{\overline{XY} - \bar{X}\bar{Y}}{\sigma_y^2},$$

где $\bar{X}$ и $\bar{Y}$ – средние значения признаков;

$\overline{XY} = \frac{\sum_i \sum_j n_{ij} x_i y_j}{N}$ среднее значение произведения признаков X и Y ;

σ_x^2 и σ_y^2 – дисперсии признаков:

$$\bar{X} = \frac{1}{N} \sum_i x_i n_{xi}, \quad \bar{Y} = \frac{1}{N} \sum_j y_j n_{yj},$$

$$\sigma_x^2 = \frac{1}{N} \sum_i x_i^2 n_{xi} - \left(\frac{1}{N} \sum_i x_i n_{xi} \right)^2, \quad \sigma_y^2 = \frac{1}{N} \sum_j y_j^2 n_{yj} - \left(\frac{1}{N} \sum_j y_j n_{yj} \right)^2.$$

Уравнения прямых регрессий имеют вид:

$$y - \bar{Y} = \rho_{y/x} (x - \bar{X}); \quad x - \bar{X} = \rho_{x/y} (y - \bar{Y}).$$

Коэффициентом линейной корреляции признаков X и Y называется величина:

$$r_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} = \frac{\overline{XY} - \bar{X} \bar{Y}}{\sigma_x \sigma_y} = \pm \sqrt{\rho_{y/x} \rho_{x/y}}.$$

В этом случае уравнения прямых регрессий можно записать:

$$y - \bar{Y} = r_{xy} \frac{\sigma_x}{\sigma_y} (x - \bar{X}); \quad x - \bar{X} = r_{xy} \frac{\sigma_y}{\sigma_x} (y - \bar{Y}).$$

Коэффициент линейной корреляции r обладает следующими свойствами:

- 1) $-1 \leq r_{xy} \leq 1$;
- 2) если $r_{xy} = \pm 1$, то между признаками X и Y существует линейная функциональная зависимость (при $r_{xy} = 1$ прямая, а при $r_{xy} = -1$ обратная);
- 3) если $r_{xy} = 0$, то между признаками X и Y отсутствует линейная корреляционная зависимость;

$$4) \rho_{x/y} = r_{xy} \frac{\sigma_x}{\sigma_y}; \quad \rho_{y/x} = r_{xy} \frac{\sigma_y}{\sigma_x}.$$

Квадрат коэффициента линейной корреляции даёт коэффициент детерминации, который измеряет долю вариации Y , объясняемую влиянием X , и наоборот.

Степень, сила или теснота корреляционной связи определяются по величине коэффициента корреляции.

Сила связи не зависит от её направленности и определяется по абсолютному значению коэффициента корреляции. Максимальное

возможное абсолютное значение коэффициента корреляции $r = 1$; минимальное $r = 0$.

Используется две системы классификации корреляционных связей по их силе – общая и частная. Общая классификация корреляционных связей осуществляется по шкале Чаддока (профессор Колумбийского университета Роберт Чаддок (Robert Emmet Chaddock, 1879 – 1940)):

- | | |
|-------------------|--|
| 1) весьма высокая | при коэффициенте корреляции $ r > 0,90$; |
| 2) высокая | при $0,70 < r < 0,9$; |
| 3) заметная | при $0,50 < r < 0,7$; |
| 4) умеренная | при $0,30 < r < 0,5$; |
| 5) слабая | при $ r < 0,3$. |

Частная классификация корреляционных связей:

- | | |
|--------------------------------|---|
| 1) высокая значимая корреляция | при r , соответствующем уровню статистической значимости $\rho \leq 0,01$ |
| 2) значимая корреляция | при r , соответствующем уровню статистической значимости $\rho \leq 0,05$ |
| 3) тенденция достоверной связи | при r , соответствующем уровню статистической значимости $\rho \leq 0,10$ |
| 4) незначимая корреляция | при r , не достигающем уровня статистической значимости |

Две эти классификации не совпадают. Первая ориентирована только на величину коэффициента корреляции, а вторая определяет, какого уровня значимости достигает данная величина коэффициента корреляции при данном объёме выборки. Чем больше объём выборки, тем меньшего по величине коэффициента корреляции оказывается достаточно, чтобы корреляция была признана достоверной. В результате при малом объёме выборки сильная корреляция может оказаться недостоверной. В то же время при больших объёмах выборки даже слабая корреляция может быть достоверной.

На практике о распределении признаков X и Y в генеральной совокупности судят по данным выборки. По этим данным может быть найден *выборочный коэффициент* линейной корреляции r_{ϵ} , который является случайной величиной.

Итак, пусть двумерная генеральная совокупность (X, Y) распределена нормально. Из этой совокупности извлечена выборка объёма n и по ней найден выборочный коэффициент корреляции $r_{\epsilon} \neq 0$. Требуется проверить нулевую гипотезу $H_0: r_{\epsilon} = 0$ о равенстве нулю генерального коэффициента корреляции.

Если нулевая гипотеза принимается, то это означает, что X и Y некоррелированы; в противном случае — коррелированы.

Правило. Для того чтобы при уровне значимости α проверить нулевую гипотезу о равенстве нулю генерального коэффициента корреляции нормальной двумерной случайной величины при конкурирующей гипотезе $H_1: r_{\theta} \neq 0$, надо вычислить наблюдаемое значение критерия $T_{\text{набл}} = r_{\theta} \sqrt{n-2} / \sqrt{1-r_{\theta}^2}$ и по таблице 12 приложения критических точек распределения Стьюдента, по заданному уровню значимости α и числу степеней свободы $k = n - 2$ найти критическую точку $t_{\text{кр}}(\alpha; k)$ двусторонней критической области. Если $|T_{\text{набл}}| < t_{\text{кр}}$, нет оснований отвергнуть нулевую гипотезу. Если $|T_{\text{набл}}| > t_{\text{кр}}$, нулевую гипотезу отвергают.

Если связь между X и Y установлена, то линейное приближение Y на X даётся формулой линейной регрессии: $y - \bar{Y} = r_{xy} \frac{\sigma_x}{\sigma_y} (x - \bar{X})$ или $y = ax + b$. Линейное приближение X на Y дается формулой линейной регрессии: $x - \bar{X} = r_{xy} \frac{\sigma_y}{\sigma_x} (y - \bar{Y})$ или $x = cy + d$.

Следует иметь в виду, что $y = ax + b$ и $x = cy + d$ — различные прямые. Первая прямая получается в результате решения задачи о минимизации суммы квадратов отклонений по вертикали, а вторая — при решении задачи о минимизации суммы квадратов отклонений по горизонтали.

Для оценки тесноты линейной корреляционной связи между признаками в выборке используется выборочный коэффициент корреляции.

Слабая линейная связь не исключает наличие между признаками X и Y нелинейной (криволинейной) связи. Оценке тесноты нелинейной корреляционной связи служат:

η_{xy} — выборочное корреляционное отношение Y к X ;

η_{yx} — выборочное корреляционное отношение X к Y .

Выборочное корреляционное отношение Y к X называется отношение межгруппового среднего квадратического отклонения к общему среднему квадратическому отклонению признака Y : $\eta_{yx} = \sigma_{\text{межгр}} / \sigma_{\text{общ}}$, или в других обозначениях: $\eta_{yx} = \delta_y / \sigma_y$.

Аналогично вычисляется выборочное корреляционное отношение X к Y : $\eta_{xy} = \delta_x / \sigma_x$.

Здесь σ_x^2, σ_y^2 – внутригрупповые дисперсии, определённые ранее и характеризующие разброс фактических значений от общих средних. Величины δ_x^2 и δ_y^2 называются межгрупповыми дисперсиями и вычисляются по формулам:

$$\delta_x^2 = \frac{1}{N} \sum_j (\bar{x}_{y_j} - \bar{X})^2 n_{y_j}, \quad \delta_y^2 = \frac{1}{N} \sum_i (\bar{y}_{x_i} - \bar{Y})^2 n_{x_i},$$

где N – объём выборки;

n_x – частота значения x признака;

n_y – частота значения y признака;

$\bar{Y}$ – общая средняя признака Y ;

$\bar{y}_x$ – условная средняя признака Y ;

$\bar{X}$ – общая средняя признака X ;

$\bar{x}_y$ – условная средняя признака X .

Выборочное корреляционное отношение обладает следующими свойствами:

1) корреляционное отношение удовлетворяет двойному неравенству $0 \leq \eta \leq 1$;

2) если $\eta = 0$, признак Y с признаком X корреляционной зависимостью не связан;

3) если $\eta = 1$, признак Y связан с признаком X функциональной зависимостью;

4) выборочное корреляционное отношение не меньше абсолютной величины выборочного коэффициента корреляции: $\eta \geq |r_s|$;

5) если выборочное корреляционное отношение равно абсолютной величины выборочного коэффициента корреляции, то имеет место точная линейная корреляционная зависимость.

Задача 2

По данным корреляционной таблицы 14

1) оценить тесноту линейной связи между признаками X и Y с помощью коэффициента корреляции;

2) проверить гипотезу о значимости выборочного коэффициента корреляции;

3) составить уравнения линий регрессии Y на X и X на Y и построить их графики, нанести на них значения условных средних $\bar{y}_x$ и $\bar{x}_y$;

4) оценить тесноту корреляционной связи с помощью корреляционного отношения.

Корреляционные таблицы

Вариант 1								
$X \backslash Y$	130	150	170	190	210	230	250	n_x
15					7	5	3	15
25				5	4	2		11
35			3	4	2			9
45		2	6	4				12
55	3	5	3					11
65	4	2						6
n_y	7	9	12	13	13	7	3	64
Вариант 2								
$X \backslash Y$	70	80	90	100	110	120	130	n_x
5				6	7	5	3	21
15			2	5	4	2	2	15
25		1	3	4	2	4		14
35	2	2	7	4	3			18
45	3	5	3					11
55	4	2						6
n_y	9	10	15	19	16	11	5	85
Вариант 3								
$X \backslash Y$	100	115	130	145	160	175	190	n_x
70	8	1	0	1	0	0	0	10
80	4	1	1	0	1	0	0	7
90	0	5	15	3	2	0	0	25
100	0	0	6	8	7	1	0	22
110	0	1	3	2	12	1	2	21
120	0	0	0	3	3	13	1	20
n_y	12	8	25	17	25	15	3	105
Вариант 4								
$X \backslash Y$	25	35	45	55	65	75	85	n_x
35	16	6	2	1	0	0	0	25
45	3	18	1	1	1	0	0	24
55	0	6	2	3	1	0	0	12
65	1	0	0	19	6	2	0	28
75	0	0	1	4	18	4	1	28
85	0	0	0	1	3	19	5	28
n_y	20	30	6	29	29	25	6	145

Вариант 5									
$X \backslash Y$	50	65	80	95	110	125	140	n_x	
35	0	0	1	1	2	17	2	23	
45	0	0	1	5	7	5	1	19	
55	1	3	1	7	8	1	1	22	
65	0	9	7	4	3	0	0	23	
75	8	14	4	4	0	0	0	30	
85	11	2	2	0	0	0	0	15	
n_y	20	28	16	21	20	23	4	132	
Вариант 6									
$X \backslash Y$	30	50	70	90	110	130	150	n_x	
50	0	0	1	1	9	8	2	21	
65	0	1	4	2	18	7	3	35	
80	0	0	3	19	10	3	1	36	
95	2	5	20	6	2	1	0	36	
110	2	22	10	0	0	0	0	34	
125	27	6	3	1	1	0	0	38	
n_y	31	34	41	29	40	19	6	200	
Вариант 7									
$X \backslash Y$	100	115	130	145	160	175	190	n_x	
70	9	1	0	0	0	0	0	10	
80	1	7	3	1	0	0	0	12	
90	2	3	9	3	2	0	0	19	
100	0	2	1	19	2	1	1	26	
110	0	0	0	1	9	7	2	19	
120	0	0	0	1	2	8	1	12	
n_y	12	13	13	25	15	16	4	98	
Вариант 8									
$X \backslash Y$	10	20	30	40	50	60	70	n_x	
10	7	1	1	0	0	0	0	9	
20	3	10	2	2	1	0	0	18	
30	0	7	6	4	0	1	0	18	
40	0	3	5	8	1	1	1	19	
50	0	1	2	4	18	1	1	27	
60	0	0	0	0	4	17	4	25	
n_y	10	22	16	18	24	20	6	116	

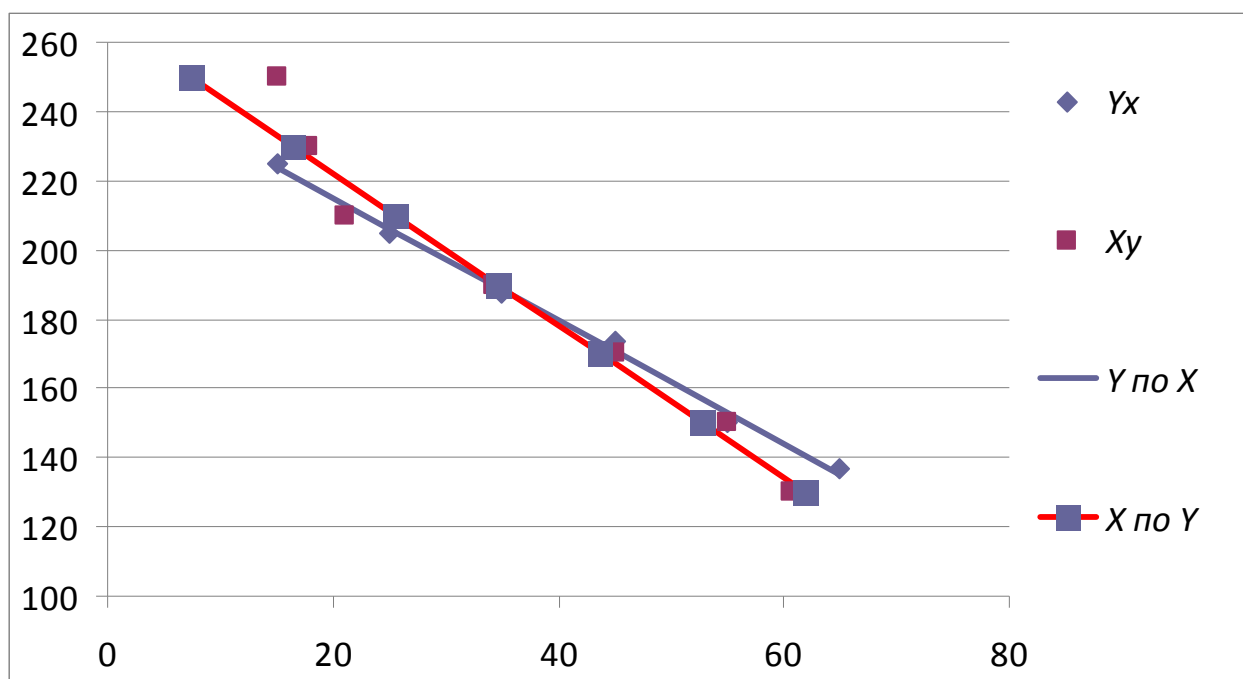
Вариант 9									
$X \backslash Y$	70	85	100	115	130	145	160	n_x	
50	0	0	1	1	7	14	3	26	
60	0	1	0	10	16	1	3	31	
70	0	1	2	3	8	3	1	18	
80	1	3	9	1	2	0	0	16	
90	0	8	2	3	1	0	0	14	
100	3	7	2	1	0	0	0	13	
n_y	4	20	16	19	34	18	7	118	
Вариант 10									
$X \backslash Y$	35	50	65	80	95	110	125	n_x	
45	0	0	1	3	1	10	3	18	
55	0	1	2	4	21	6	1	35	
65	0	3	1	20	0	2	0	26	
75	3	4	20	10	3	1	0	41	
85	10	6	4	1	1	0	0	22	
95	29	4	3	0	0	0	0	36	
n_y	42	18	31	38	26	19	4	178	
Вариант 11									
$X \backslash Y$	50	65	80	95	110	125	140	n_x	
35	13	6	1	0	0	0	0	20	
45	5	5	5	2	1	0	0	18	
55	1	6	8	6	0	1	0	22	
65	1	2	5	7	0	0	1	16	
75	0	1	1	2	10	6	1	21	
85	0	0	1	2	7	10	1	21	
n_y	20	20	21	19	18	17	3	118	
Вариант 12									
$X \backslash Y$	25	40	55	70	85	100	115	n_x	
50	12	5	1	1	0	0	0	19	
65	7	11	7	0	1	0	0	26	
80	2	1	11	2	2	1	0	19	
95	1	2	1	10	5	0	0	19	
110	0	0	2	7	14	5	0	28	
125	0	0	0	1	4	14	3	22	
n_y	22	19	22	21	26	20	3	133	

Вариант 13									
$X \backslash Y$	80	95	110	125	140	155	170	n_x	
70	0	0	1	1	7	12	8	29	
80	0	1	2	5	16	2	3	29	
90	1	2	3	29	7	0	1	43	
100	1	8	23	2	2	0	0	36	
110	9	5	0	0	1	0	0	15	
120	24	2	1	1	0	0	0	28	
n_y	35	18	30	38	33	14	12	180	
Вариант 14									
$X \backslash Y$	25	35	45	55	65	75	85	n_x	
35	0	0	0	4	2	13	11	30	
45	0	1	1	8	19	4	2	35	
55	1	1	6	17	2	2	1	30	
65	0	4	5	2	1	1	1	14	
75	2	22	2	0	0	0	0	26	
85	10	9	2	1	0	0	0	22	
n_y	13	37	16	32	24	20	15	157	
Вариант 15									
$X \backslash Y$	50	65	80	95	110	125	140	n_x	
35	2	2	1	0	0	0	0	5	
45	4	4	3	1	1	0	0	13	
55	2	7	9	2	0	0	0	20	
65	1	1	0	7	2	1	1	13	
75	0	0	2	2	16	7	1	28	
85	0	0	1	1	5	13	3	23	
n_y	9	14	16	13	24	21	5	102	
Вариант 16									
$X \backslash Y$	30	50	70	90	110	130	150	n_x	
50	17	1	1	0	0	0	0	19	
65	3	14	1	2	1	0	0	21	
80	2	5	5	5	0	1	0	18	
95	0	2	2	7	4	1	1	17	
110	0	1	3	2	19	5	0	30	
125	0	0	0	1	0	1	0	2	
n_y	22	23	12	17	24	8	1	107	

Пример нахождения коэффициента корреляции и уравнений линейной регрессии в MS Excel

$X \backslash Y$	130	150	170	190	210	230	250	
15					7	5	3	15
25				5	4	2		11
35			3	4	2			9
45		2	6	4				12
55	3	5	3					11
65	4	2						6
	7	9	12	13	13	7	3	64
X	N	NX	NX^2					
15	15	225	3375	$X_e =$	36,71875			
25	11	275	6875	$X_e^2 =$	1628,125			
35	9	315	11025	$Y_e =$	185,3125			
45	12	540	24300	$Y_e^2 =$	35437,5			
55	11	605	33275	$XY =$	6306,25			
65	6	390	25350	$\sigma_x =$	16,72897			
	64	2350	104200	$\sigma_y =$	33,11763			
				$C_{xy} =$	-498,193			
Y	N	NY	NY^2	$r_{xy} =$	-0,89923			
130	7	910	118300	$\rho_{xy} =$	-1,78016			
150	9	1350	202500	$\rho_{yx} =$	-0,45423			
170	12	2040	346800					
190	13	2470	469300					
210	13	2730	573300					
230	7	1610	370300					
250	3	750	187500					
	64	11860	2268000					

$X \backslash Y$	130	150	170	190	210	230	250	Y_x
15	0	0	0	0	22050	17250	11250	224,667
25	0	0	0	23750	21000	11500	0	204,546
35	0	0	17850	26600	14700	0	0	187,778
45	0	13500	45900	34200	0	0	0	173,333
55	21450	41250	28050	0	0	0	0	150
65	33800	19500	0	0	0	0	0	136,667
X_y	60,714	55	45	34,231	21,1539	17,857	15	403600
X	Y_x	X	Y по X			X_y	Y	X по Y
15	224,67	15	223,975		130	60,71429	130	61,844
25	204,546	25	206,174		150	55	150	52,759
35	187,778	35	188,372		170	45	170	43,674
45	173,333	45	170,571		190	34,23077	190	34,590
55	150	55	152,769		210	21,15385	210	25,505
65	136,67	65	134,967		230	17,85714	230	16,420
					250	15	250	7,336



Оценим тесноту корреляционной связи с помощью корреляционного отношения

$$X_{\bar{c}} = 36,72 \quad Y_{\bar{c}} = 185,31$$

Y_x	n_x	$(Y_x - Y_{\bar{B}})^2 n_x$	X_y	n_y	$(X_y - X_{\bar{B}})^2 n_y$
224,67	15	23231,26	60,71	7	4030,5
204,55	11	4068,97	55	9	3007,84
187,78	9	54,70	45	12	822,95
173,33	12	1722,00	34,23	13	80,47
150	11	13716,7	21,15	13	3149,46
136,67	6	14198,5	17,86	7	2490,32
		56992,13	15	3	1415,11
					14996,65

$$\begin{aligned} \delta_y^2 &= 890,50 & \sigma_y^2 &= 1096,78 \\ \delta_x^2 &= 234,32 & \sigma_x^2 &= 279,86 \end{aligned}$$

$$\eta_{yx} = 0,901$$

$$\eta_{xy} = 0,915$$

Условие $\eta \geq |r_{xy}|$ выполнено.

Если расчёты и построения выполнены верно, то прямая регрессии Y на X должна пройти вблизи всех точек $(x; y_x)$ так, чтобы эти точки располагались по обе стороны от прямой. Аналогично, прямая регрессии X на Y наилучшим образом приближена к точкам $(x_y; y)$.

Обе прямые регрессии пересекаются в точке $(X; Y)$. В примере это точка $(36,72; 185,32)$. Чем больше сила связи между признаками, тем ближе расположены обе прямые друг к другу (угол между ними мал). В приведённом примере высокая значимая корреляционная связь, что хорошо видно на графике.

Методические указания

Ранговая корреляция.

Коэффициент ранговой корреляции r_s Спирмена

Коэффициент ранговой корреляции рекомендуется применять в тех случаях, когда необходимо проверить, согласованно ли изменяются разные признаки у одного и того же испытуемого и насколько совпадают индивидуальные ранговые показатели у двух отдельных испытуемых или у испытуемого и группы.

Назначение рангового коэффициента корреляции. Метод ранговой корреляции Спирмена позволяет определить тесноту (силу) и направление корреляционной связи между *двумя признаками* или *двумя профилями (иерархиями)* признаков.

Для подсчёта ранговой корреляции необходимо располагать двумя рядами значений, которые могут быть проранжированы. Такими рядами значений могут быть:

1) *два признака*, измеренные в одной и той же группе испытуемых;

2) *две индивидуальные иерархии признаков*, выявленные у двух испытуемых по одному и тому же набору признаков (например, личностные профили по 16-факторному опроснику Р. Б. Кеттелла, иерархии ценностей по методике Р. Рокича, последовательности предпочтений в выборе из нескольких альтернатив и др.);

3) *две групповые иерархии признаков*;

4) *индивидуальная и групповая иерархии признаков*.

Вначале показатели ранжируются отдельно по каждому из признаков. Как правило, меньшему значению признака начисляется меньший ранг.

Гипотезы.

Возможны два варианта гипотез. Первый относится к случаю 1, второй – к трем остальным случаям.

Первый вариант гипотез:

H_0 : корреляция между переменными А и Б не отличается от нуля;

H_1 : корреляция между переменными А и Б достоверно отличается от нуля.

Второй вариант гипотез:

H_0 : корреляция между иерархиями А и Б не отличается от нуля;

H_1 : корреляция между иерархиями А и Б достоверно отличается от нуля.

Ограничения коэффициента ранговой корреляции

1. По каждой переменной должно быть представлено не менее 5 наблюдений. Верхняя граница выборки определяется имеющимися таблицами критических значений (приложение, таблица 13), а именно $N < 40$.

2. Коэффициент ранговой корреляции Спирмена r_s при большом количестве одинаковых рангов по одной или обоим сопоставляемым переменным даёт огрублённые значения. В идеале оба коррелируемых ряда должны представлять собой две последовательности несовпадающих значений. Если это условие не соблюдается, необходимо вносить поправку на одинаковые ранги.

Алгоритм расчёта коэффициента ранговой корреляции Спирмена r_s

1. Определить, какие два признака или две иерархии признаков будут участвовать в сопоставлении как переменные А и В.

2. Проранжировать значения переменной А, начисляя ранг 1 наименьшему значению, в соответствии с правилами ранжирования. Занести ранги в первый (по порядку номеров испытуемых или признаков) столбец таблицы.

3. Проранжировать значения переменной В – в соответствии с теми же правилами. Занести ранги во второй (по порядку номеров испытуемых или признаков) столбец таблицы.

4. Подсчитать разности d между рангами А и В по каждой строке таблицы и занести в третий столбец таблицы.

5. Возвести каждую разность в квадрат: d^2 . Эти значения занести в четвёртый столбец таблицы.

6. Подсчитать сумму квадратов $\sum d^2$.

7. При наличии одинаковых рангов рассчитать поправки:

$$T_a = \sum(a^3 - a)/12, T_b = \sum(b^3 - b)/12,$$

где a – объём каждой группы одинаковых рангов в ранговом ряду А;

b – объём каждой группы одинаковых рангов в ранговом ряду В.

8. Рассчитать коэффициент ранговой корреляции r_s :

а) при отсутствии одинаковых рангов

$$r_s = 1 - 6 \frac{\sum d^2}{N(N^2 - 1)}$$

б) при наличии одинаковых рангов

$$r_s = 1 - 6 \frac{\sum d^2 + T_a + T_b}{N(N^2 - 1)},$$

где $\sum d^2$ - сумма квадратов разностей между рангами;

T_a и T_b – поправки на одинаковые ранги;

N – количество испытуемых или признаков, участвовавших в ранжировании.

9. Определить по табл. 13 приложения критические значения r_s для данного N . Если r_s превышает критическое значение или по крайней мере равен ему, корреляция достоверно отличается от 0. Связь достоверна, если $r_{s\text{ эмн}} \geq r_{s\ 0,05}$, и ещё более достоверна, если $r_{s\text{ эмн}} \geq r_{s\ 0,01}$.

Задача 3

Группа испытуемых студентов физического факультета – проходила подготовку на тренажере, моделирующем деятельность авиадиспетчера. Испытуемые должны были решать задачи по выбору оптимального типа взлетно-посадочной полосы для заданного типа самолета. Связано ли количество ошибок, допущенных испытуемыми в тренировочной сессии, с показателями вербального и невербального интеллекта, измеренными по методике Д. Векслера? (таблица 15)

**Количество ошибок, допущенных испытуемыми
в тренировочной сессии и показатели вербального и невербаль-
ного интеллекта, измеренные по методике Д. Векслера**

Вариант 1				Вариант 2			
Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта	Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	15	127	96	1	33	135	96
2	44	126	109	2	17	127	96
3	50	139	91	3	33	130	110
4	7	132	94	4	25	130	96
5	47	130	93	5	39	124	109
6	26	123	106	6	10	125	109
7	49	139	110	7	6	135	97
8	5	130	105	8	39	137	97
9	10	124	102	9	15	136	100
10	44	125	100	10	15	123	102
Вариант 3				Вариант 4			
1	8	122	105	1	26	137	103
2	42	132	96	2	36	121	98
3	2	126	97	3	4	131	94
4	11	138	93	4	27	132	95
5	8	127	97	5	17	124	107
6	41	138	107	6	27	124	95
7	3	130	103	7	13	130	106
8	34	120	104	8	47	133	99
9	18	127	92	9	44	137	94
10	22	133	98	10	22	138	108

Вариант 5				Вариант 6			
Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта	Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	16	132	97	1	24	130	101
2	14	137	94	2	28	136	97
3	19	131	102	3	47	121	100
4	34	130	104	4	48	133	98
5	28	123	108	5	5	134	104
6	26	126	91	6	6	130	95
7	16	138	101	7	34	140	108
8	23	133	101	8	27	126	107
9	31	124	91	9	8	133	107
10	9	134	101	10	43	129	109
Вариант 7				Вариант 8			
1	37	130	104	1	26	128	100
2	49	123	93	2	35	138	100
3	21	131	107	3	14	121	100
4	1	134	102	4	20	127	100
5	12	138	97	5	34	122	103
6	16	129	94	6	40	126	100
7	50	125	96	7	19	130	103
8	24	140	92	8	20	138	106
9	50	133	98	9	21	129	94
10	17	129	92	10	8	135	95

Вариант 9				Вариант 10			
Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта	Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	20	129	96	1	3	126	105
2	44	121	92	2	23	120	107
3	46	124	93	3	46	123	91
4	4	135	100	4	36	127	109
5	13	138	96	5	19	131	108
6	30	134	99	6	21	136	103
7	3	126	104	7	46	130	107
8	50	133	106	8	27	124	102
9	21	121	110	9	14	124	91
10	36	123	108	10	7	120	93
Вариант 11				Вариант 12			
1	30	121	100	1	27	128	103
2	16	128	106	2	29	133	94
3	20	135	109	3	25	127	99
4	11	124	93	4	16	126	102
5	28	138	106	5	3	139	103
6	42	122	97	6	5	122	95
7	28	134	92	7	49	126	104
8	10	137	102	8	17	137	91
9	18	131	92	9	30	131	96
10	47	128	101	10	46	129	96

Вариант 13				Вариант 14			
Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта	Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	43	137	105	1	15	127	96
2	26	138	109	2	3	138	101
3	24	124	96	3	30	123	98
4	3	139	102	4	35	124	103
5	36	138	99	5	44	124	96
6	48	129	93	6	28	126	92
7	43	133	93	7	0	135	107
8	10	137	91	8	31	134	94
9	35	138	107	9	13	139	94
10	9	123	102	10	2	134	109
Вариант 15				Вариант 16			
1	14	131	106	1	36	128	95
2	29	126	94	2	20	138	92
3	47	129	91	3	12	128	103
4	45	131	98	4	19	130	92
5	31	135	97	5	36	123	91
6	43	121	104	6	11	129	97
7	6	130	97	7	49	138	96
8	24	132	93	8	11	122	95
9	37	131	93	9	39	129	98
10	30	130	102	10	31	131	105

СПИСОК ЛИТЕРАТУРЫ

1. *Сидоренко, Е. В.* Методы математической обработки в психологии: / Е. В. Сидоренко. – СПб. : Речь, 2006. – 349 с.
2. *Наследов, А. Д.* Математические методы психологического исследования: / А. Д. Наследов. – СПб. : Речь, 2007. – 398 с.
3. *Кутейников, А. Н.* Математические методы в психологии: учебное пособие / А. Н. Кутейников. – СПб. : Речь, 2008. – 172 с.
4. *Вуколов, Э. А.* Основы статистического анализа, Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: учебное пособие / Э. А. Вуколов. – 2-е изд., испр. и доп. – М. : ФОРУМ, 2008. – 464 с. – (Высшее образование).
5. Математические методы в психологии и педагогике: практикум / В. В. Комиссаров. – Новосибирск: Изд-во НГТУ, 2013. – 104 с.

Приложение

Таблица 1

Таблица значений интегральной функции Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,51	0,1950	1,01	0,3438	1,51	0,4345	2,01	0,4778	2,51	0,4940	3,10	0,49903
0,01	0,0040	0,52	0,1985	1,02	0,3461	1,52	0,4357	2,02	0,4783	2,52	0,4941	3,20	0,49931
0,02	0,0080	0,53	0,2019	1,03	0,3485	1,53	0,4370	2,03	0,4788	2,53	0,4943	3,30	0,49952
0,03	0,0120	0,54	0,2054	1,04	0,3508	1,54	0,4382	2,04	0,4793	2,54	0,4945	3,40	0,49966
0,04	0,0160	0,55	0,2088	1,05	0,3531	1,55	0,4394	2,05	0,4798	2,55	0,4946	3,50	0,49977
0,05	0,0199	0,56	0,2123	1,06	0,3554	1,56	0,4406	2,06	0,4803	2,56	0,4948	3,60	0,49984
0,06	0,0239	0,57	0,2157	1,07	0,3577	1,57	0,4418	2,07	0,4808	2,57	0,4949	3,70	0,499892
0,07	0,0279	0,58	0,2190	1,08	0,3599	1,58	0,4429	2,08	0,4812	2,58	0,4951	3,80	0,499928
0,08	0,0319	0,59	0,2224	1,09	0,3621	1,59	0,4441	2,09	0,4817	2,59	0,4952	3,90	0,499952
0,09	0,0359	0,60	0,2257	1,10	0,3643	1,60	0,4452	2,10	0,4821	2,60	0,4953	4,00	0,499968
0,10	0,0398	0,61	0,2291	1,11	0,3665	1,61	0,4463	2,11	0,4826	2,61	0,4955	4,10	0,499979
0,11	0,0438	0,62	0,2324	1,12	0,3686	1,62	0,4474	2,12	0,4830	2,62	0,4956	4,20	0,499987
0,12	0,0478	0,63	0,2357	1,13	0,3708	1,63	0,4484	2,13	0,4834	2,63	0,4957	4,30	0,499991
0,13	0,0517	0,64	0,2389	1,14	0,3729	1,64	0,4495	2,14	0,4838	2,64	0,4959	4,40	0,499995
0,14	0,0557	0,65	0,2422	1,15	0,3749	1,65	0,4505	2,15	0,4842	2,65	0,4960	4,50	0,499997
0,15	0,0596	0,66	0,2454	1,16	0,3770	1,66	0,4515	2,16	0,4846	2,66	0,4961	4,60	0,499998
0,16	0,0636	0,67	0,2486	1,17	0,3790	1,67	0,4525	2,17	0,4850	2,67	0,4962	4,70	0,499999
0,17	0,0675	0,68	0,2517	1,18	0,3810	1,68	0,4535	2,18	0,4854	2,68	0,4963	4,80	0,499999
0,18	0,0714	0,69	0,2549	1,19	0,3830	1,69	0,4545	2,19	0,4857	2,69	0,4964	4,90	0,500000
0,19	0,0753	0,70	0,2580	1,20	0,3849	1,70	0,4554	2,20	0,4861	2,70	0,4965	5,00	0,500000
0,20	0,0793	0,71	0,2611	1,21	0,3869	1,71	0,4564	2,21	0,4864	2,71	0,4966		
0,21	0,0832	0,72	0,2642	1,22	0,3888	1,72	0,4573	2,22	0,4868	2,72	0,4967		
0,22	0,0871	0,73	0,2673	1,23	0,3907	1,73	0,4582	2,23	0,4871	2,73	0,4968		
0,23	0,0910	0,74	0,2704	1,24	0,3925	1,74	0,4591	2,24	0,4875	2,74	0,4969		
0,24	0,0948	0,75	0,2734	1,25	0,3944	1,75	0,4599	2,25	0,4878	2,75	0,4970		
0,25	0,0987	0,76	0,2764	1,26	0,3962	1,76	0,4608	2,26	0,4881	2,76	0,4971		
0,26	0,1026	0,77	0,2794	1,27	0,3980	1,77	0,4616	2,27	0,4884	2,77	0,4972		
0,27	0,1064	0,78	0,2823	1,28	0,3997	1,78	0,4625	2,28	0,4887	2,78	0,4973		

0,28	0,1103	0,79	0,2852	1,29	0,4015	1,79	0,4633	2,29	0,4890	2,79	0,4974
0,29	0,1141	0,80	0,2881	1,30	0,4032	1,80	0,4641	2,30	0,4893	2,80	0,4974
0,30	0,1179	0,81	0,2910	1,31	0,4049	1,81	0,4649	2,31	0,4896	2,81	0,4975
0,31	0,1217	0,82	0,2939	1,32	0,4066	1,82	0,4656	2,32	0,4898	2,82	0,4976
0,32	0,1255	0,83	0,2967	1,33	0,4082	1,83	0,4664	2,33	0,4901	2,83	0,4977
0,33	0,1293	0,84	0,2995	1,34	0,4099	1,84	0,4671	2,34	0,4904	2,84	0,4977
0,34	0,1331	0,85	0,3023	1,35	0,4115	1,85	0,4678	2,35	0,4906	2,85	0,4978
0,35	0,1368	0,86	0,3051	1,36	0,4131	1,86	0,4686	2,36	0,4909	2,86	0,4979
0,36	0,1406	0,87	0,3078	1,37	0,4147	1,87	0,4693	2,37	0,4911	2,87	0,4979
0,37	0,1443	0,88	0,3106	1,38	0,4162	1,88	0,4699	2,38	0,4913	2,88	0,4980
0,38	0,1480	0,89	0,3133	1,39	0,4177	1,89	0,4706	2,39	0,4916	2,89	0,4981
0,39	0,1517	0,90	0,3159	1,40	0,4192	1,90	0,4713	2,40	0,4918	2,90	0,4981
0,40	0,1554	0,91	0,3186	1,41	0,4207	1,91	0,4719	2,41	0,4920	2,91	0,4982
0,41	0,1591	0,92	0,3212	1,42	0,4222	1,92	0,4726	2,42	0,4922	2,92	0,4982
0,42	0,1628	0,93	0,3238	1,43	0,4236	1,93	0,4732	2,43	0,4925	2,93	0,4983
0,43	0,1664	0,94	0,3264	1,44	0,4251	1,94	0,4738	2,44	0,4927	2,94	0,4984
0,44	0,1700	0,95	0,3289	1,45	0,4265	1,95	0,4744	2,45	0,4929	2,95	0,4984
0,45	0,1736	0,96	0,3315	1,46	0,4279	1,96	0,4750	2,46	0,4931	2,96	0,4985
0,46	0,1772	0,97	0,3340	1,47	0,4292	1,97	0,4756	2,47	0,4932	2,97	0,4985
0,47	0,1808	0,98	0,3365	1,48	0,4306	1,98	0,4761	2,48	0,4934	2,98	0,4986
0,48	0,1844	0,99	0,3389	1,49	0,4319	1,99	0,4767	2,49	0,4936	2,99	0,4986
0,49	0,1879	1,00	0,3413	1,50	0,4332	2,00	0,4772	2,50	0,4938	3,00	0,4987
0,50	0,1915	1,01	0,3438	1,51	0,4345	2,01	0,4778	2,51	0,4940	3,01	0,4987

$\Phi(-x) = \Phi(-x), \Phi(x > 5) = 0,5.$

Таблица 2

**Критические значения критерия Q Розенбаума
для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$**

n	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
$\rho = 0,05$																
11	6															
12	6	6														
13	6	6	6													
14	7	7	6	6												
15	7	7	6	6	6											
16	8	7	7	7	6	6										
17	7	7	7	7	7	7	7									
18	7	7	7	7	7	7	7	7								
19	7	7	7	7	7	7	7	7	7							
20	7	7	7	7	7	7	7	7	7	7						
21	8	7	7	7	7	7	7	7	7	7	7					
22	8	7	7	7	7	7	7	7	7	7	7	7				
23	8	8	7	7	7	7	7	7	7	7	7	7	7			
24	8	8	8	8	8	8	8	8	8	8	7	7	7	7		
25	8	8	8	8	8	8	8	8	8	7	7	7	7	7	7	
26	8	8	8	8	8	8	8	8	8	8	7	7	7	7	7	7
$\rho = 0,01$																
11	9															
12	9	9														
13	9	9	9													
14	9	9	9	9												
15	9	9	9	9	9											
16	9	9	9	9	9	9										
17	10	9	9	9	9	9	9									
18	10	10	9	9	9	9	9	9								
19	10	10	10	9	9	9	9	9	9							
20	10	10	10	10	9	9	9	9	9	9						
21	11	10	10	10	9	9	9	9	9	9	9					
22	11	11	10	10	10	9	9	9	9	9	9	9				
23	11	11	10	10	10	10	9	9	9	9	9	9	9			
24	12	11	11	10	10	10	10	9	9	9	9	9	9	9		
25	12	11	11	10	10	10	10	10	9	9	9	9	9	9	9	
26	12	12	11	11	10	10	10	10	10	9	9	9	9	9	9	9

Различия между двумя выборками можно считать достоверными ($\rho \leq 0,05$), если $Q_{\text{эмп}}$ равен или выше критического значения $Q_{0,05}$; и ещё более достоверными ($\rho \leq 0,01$), если $Q_{\text{эмп}}$ равен или выше критического значения $Q_{0,01}$.

Таблица 3

**Критические значения критерия U Манна–Уитни
для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$**

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_2	$\rho = 0,05$																		
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	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	10	10		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	10	10	11	12	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	10	10	11	12	13	13
n_2	$\rho = 0,01$																		
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	11	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	10	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	10	10	11

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

Таблица 4

Критические значения критерия χ^2 для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$ при разном числе степеней свободы ν

ν	ρ		ν	ρ		ν	ρ	
	0,05	0,01		0,05	0,01		0,05	0,01
1	3,841	6,635	35	49,802	57,342	69	89,391	99,227
2	5,991	9,210	36	50,998	58,619	70	90,631	100,42
3	7,815	11,345	37	52,192	59,892	71	91,670	101,62
4	9,488	13,277	38	53,384	61,162	72	92,808	102,81
5	11,070	15,086	39	54,572	62,428	73	93,945	104,01
6	12,592	16,812	40	55,758	63,691	74	95,081	105,20
7	14,067	18,475	41	56,942	64,950	75	96,217	106,39
8	15,507	20,090	42	58,124	66,206	76	97,351	107,58
9	16,919	21,666	43	59,304	67,459	77	98,484	108,77
10	18,307	23,209	44	60,481	68,709	78	99,617	109,95
11	19,675	24,725	45	61,656	69,957	79	100,74	111,14
12	21,026	26,217	46	62,830	71,201	80	101,87	112,32
13	22,362	27,688	47	64,001	72,443	81	103,01	113,51
14	23,685	29,141	48	65,171	73,683	82	104,13	114,69
15	24,996	30,578	49	66,339	74,919	83	105,26	115,87
16	26,296	32,000	50	67,505	76,154	84	106,39	117,05
17	27,587	33,409	51	68,669	77,386	85	107,52	118,23
18	28,869	34,805	52	69,832	78,616	86	108,64	119,41
19	30,144	36,191	53	70,993	79,843	87	109,77	120,59
20	31,410	37,566	54	72,153	81,069	88	110,89	121,76
21	32,671	38,932	55	73,311	82,292	89	112,02	122,94
22	33,924	40,289	56	74,468	83,513	90	113,14	124,11
23	35,172	41,638	57	75,624	84,733	91	114,26	125,28
24	36,415	42,980	58	76,778	85,950	92	115,39	126,46
25	37,652	44,314	59	77,931	87,166	93	116,51	127,63
26	38,885	45,642	60	79,082	88,379	94	117,63	128,80
27	40,113	46,963	61	80,232	89,591	95	118,75	129,97
28	41,337	48,278	62	81,381	90,802	96	119,87	131,14
29	42,557	49,588	63	82,529	92,010	97	120,99	132,30
30	43,773	50,892	64	83,675	93,217	98	122,10	133,47
31	44,985	52,191	65	84,821	94,422	99	123,22	134,64
32	46,194	53,486	66	85,965	95,626	100	124,34	135,80
33	47,400	54,776	67	87,108	96,828			
34	48,602	56,061	68	88,250	98,028			

Таблица 5

Критические значения критерия тенденций S Джонкира для количества групп (c) от трех до шести ($3 \leq c \leq 6$) и количества испытуемых в каждой группе от двух до десяти ($2 \leq n \leq 10$)

c	n								
	2	3	4	5	6	7	8	9	10
$\rho = 0,05$									
3	10	17	24	33	42	53	64	76	88
4	14	26	38	51	66	82	100	118	138
5	20	34	51	71	92	115	140	166	194
6	26	44	67	93	121	151	184	219	256
$\rho = 0,01$									
3		23	32	45	59	74	90	106	124
4	20	34	50	71	92	115	140	167	195
5	26	48	72	99	129	162	197	234	274
6	34	62	94	130	170	213	260	309	361

Тенденция является достоверной, если $S_{эмп}$ достигает $S_{0,05}$ или превышает его ($\rho \leq 0,05$), и ещё более достоверной, если $S_{эмп}$ достигает $S_{0,01}$ или превышает его ($\rho \leq 0,01$).

Таблица 6

Критические значения критерия знаков G для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$

n	ρ		n	ρ		n	ρ		n	ρ	
	0,05	0,01		0,05	0,01		0,05	0,01		0,05	0,01
5	0	—	27	8	7	49	18	15	92	37	34
6	0	—	28	8	7	50	18	16	94	38	35
7	0	0	29	9	7	52	19	17	96	39	36
8	1	0	30	10	8	54	20	18	98	40	37
9	1	0	31	10	8	56	21	18	100	41	37
10	1	0	32	10	8	58	22	19	110	45	42
11	2	1	33	11	9	60	23	20	120	50	46
12	2	1	34	11	9	62	24	21	130	55	51
13	3	1	35	12	10	64	24	22	140	59	55
14	3	2	36	12	10	66	25	23	150	64	60

n	ρ		n	ρ		n	ρ		n	ρ	
	0,05	0,01		0,05	0,01		0,05	0,01		0,05	0,01
15	3	2	37	13	10	68	26	23	160	69	64
16	4	2	38	13	11	70	27	24	170	73	69
17	4	3	39	13	11	72	28	25	180	78	73
18	5	3	40	14	12	74	29	26	190	83	78
19	5	4	41	14	12	76	30	27	200	87	83
20	5	4	42	15	13	78	31	28	220	97	92
21	6	4	43	15	13	80	32	29	240	106	101
22	6	5	44	16	13	82	33	30	260	116	110
23	7	5	45	16	14	84	33	30	280	125	120
24	7	5	46	16	14	86	34	31	300	135	129
25	7	6	47	17	15	88	35	32			
26	8	6	48	17	15	90	36	33			

Преобладание «типичного» сдвига является достоверным, если $G_{\text{эмп}}$ ниже или равен $G_{0,05}$, и ещё более достоверным, если $G_{\text{эмп}}$ ниже или равен $G_{0,01}$.

Таблица 7

**Критические значения критерия T Вилкоксона
для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$**

n	ρ		n	ρ	
	0,05	0,01		0,05	0,01
5	0	—	28	130	101
6	2	—	29	140	110
7	3	0	30	151	120
8	5	1	31	163	130
9	8	3	32	175	140
10	10	5	33	187	151
11	13	7	34	200	162

n	ρ		n	ρ	
	0,05	0,01		0,05	0,01
12	17	9	35	213	173
13	21	12	36	227	185
14	25	15	37	241	198
15	30	19	38	256	211
16	35	23	39	271	224
17	41	27	40	286	238
18	47	32	41	302	252
19	53	37	42	319	266
20	60	43	43	336	281
21	67	49	44	353	296
22	75	55	45	371	312
23	83	62	46	389	328
24	91	69	47	407	345
25	100	76	48	426	362
26	110	84	49	446	379
27	119	92	50	466	397

«Типичный» сдвиг является достоверно преобладающим по интенсивности, если $T_{\text{эмп}}$ ниже или равен $T_{0,05}$, и ещё более достоверно преобладающим, если $T_{\text{эмп}}$, ниже или равен $T_{0,01}$.

Таблица 8а

Критические значения критерия χ^2 Фридмана
для количества условий $c = 3$ и количества испытуемых
от двух до девяти ($2 \leq n \leq 9$).

$n = 2$		$n = 3$		$n = 4$		$n = 5$	
χ^2	ρ	χ^2	ρ	χ^2	ρ	χ^2	ρ
0	1,000	0,000	1,000	0,0	1,000	0,0	1,000
1	0,833	0,667	0,944	0,5	0,931	0,4	0,954
3	0,500	2,000	0,528	1,5	0,653	1,2	0,691
4	0,167	2,667	0,361	2,0	0,431	1,6	0,522
		4,667	0,194	3,5	0,273	2,8	0,367
		6,000	0,028	4,5	0,125	3,6	0,182
				6,0	0,069	4,8	0,124
				6,5	0,042	5,2	0,093
				8,0	0,0046	6,4	0,039
						7,6	0,024
						8,4	0,0085
						10,0	0,00077

Окончание таблицы 8а

$n = 6$		$n = 7$		$n = 8$		$n = 9$	
χ^2	ρ	χ^2	ρ	χ^2	ρ	χ^2	ρ
0,00	1,000	0,000	1,000	0,00	1,000	0,000	1,000
0,33	0,956	0,286	0,964	0,25	0,967	0,222	0,971
1,00	0,740	0,857	0,768	0,75	0,794	0,667	0,814
1,33	0,570	1,143	0,620	1,00	0,654	0,889	0,865
2,33	0,430	2,000	0,486	1,75	0,531	1,556	0,569
3,00	0,252	2,571	0,305	2,25	0,355	2,000	0,398
4,00	0,184	3,429	0,237	3,00	0,285	2,667	0,328
4,33	0,142	3,714	0,192	3,25	0,236	2,889	0,278
5,33	0,072	4,571	0,112	4,00	0,149	3,556	0,187
6,33	0,052	5,429	0,085	4,75	0,120	4,222	0,154
7,00	0,029	6,000	0,052	5,25	0,079	4,667	0,107
8,33	0,012	7,143	0,027	6,25	0,047	5,556	0,069
9,00	0,0081	7,714	0,021	6,75	0,038	6,000	0,057
9,33	0,0055	8,000	0,016	7,00	0,030	6,222	0,048
10,33	0,0017	8,857	0,0084	7,75	0,018	6,889	0,031
12,00	0,0001	10,286	0,0036	9,00	0,0099	8,000	0,019
		10,571	0,0027	9,25	0,0080	8,222	0,016
		11,143	0,0012	9,75	0,0048	8,667	0,010
		12,286	0,00032	10,75	0,0024	9,556	0,0060
		14,000	0,00002	12,00	0,0011	10,667	0,0035
				12,25	0,00086	10,889	0,0029
				13,00	0,00026	11,556	0,0013
				14,25	0,00006	12,667	0,00066
				16,00	0,00000	13,556	0,00035
						14,000	0,00020
						14,222	0,00009
						14,889	0,00005
						16,222	0,00001
						18,000	0,00000

Таблица 8б

**Критические значения критерия χ_r^2 Фридмана
для количества условий $c = 4, 2 \leq n \leq 4$**

$n = 2$		$n = 3$		$n = 4$			
χ_r^2	ρ	χ_r^2	ρ	χ_r^2	ρ	χ_r^2	ρ
0,0	1,000	0,0	1,000	0,0	1,000	5,7	0,141
0,6	0,958	0,6	0,958	0,3	0,992	6,0	0,105
1,2	0,834	1,0	0,910	0,6	0,928	6,3	0,094
1,8	0,792	1,8	0,727	0,9	0,900	6,6	0,077
2,4	0,625	2,2	0,608	1,2	0,800	6,9	0,068
3,0	0,542	2,6	0,524	1,5	0,754	7,2	0,054
3,6	0,458	3,4	0,446	1,8	0,677	7,5	0,052
4,2	0,375	3,8	0,342	2,1	0,649	7,8	0,036
4,8	0,208	4,2	0,300	2,4	0,524	8,1	0,033
5,4	0,167	5,0	0,207	2,7	0,508	8,4	0,019
6,0	0,042	5,4	0,175	3,0	0,432	8,7	0,014
		5,8	0,148	3,3	0,389	9,3	0,012
		6,6	0,075	3,6	0,355	9,6	0,0069
		7,0	0,054	3,9	0,324	9,9	0,0062
		7,4	0,033	4,5	0,242	10,2	0,0027
		8,2	0,017	4,8	0,200	10,8	0,0016
		9,0	0,0017	5,1	0,190	11,1	0,00094
				5,4	0,158	12,0	0,000072

Различия между условиями можно считать достоверными на указанном в таблице уровне значимости, если $\chi_{г\text{ эмп}}^2$ достигает соответствующего, критического значения или превышает его.

Таблица 9

**Критические значения критерия тенденций L Пейджа
для количества условий от трех до шести ($3 \leq c \leq 6$) и количества
испытуемых от двух до двенадцати ($2 \leq n \leq 12$)**

n	c (количество условий)				
	3	4	5	6	ρ
2	—	—	109	178	0,001
	—	60	106	173	0,01
	28	58	103	166	0,05
3	—	89	160	260	0,001
	42	87	155	252	0,01
	41	84	150	244	0,05
4	56	117	210	341	0,001
	55	114	204	331	0,01
	54	111	197	321	0,05
5	70	145	259	420	0,001
	68	141	251	409	0,01
	66	137	244	397	0,05
6	83	172	307	499	0,001
	81	167	299	486	0,01
	79	163	291	474	0,05
7	96	198	355	577	0,001
	93	193	346	563	0,01
	91	189	338	550	0,05
8	109	225	403	655	0,001
	106	220	393	640	0,01
	104	214	384	625	0,05
9	121	252	451	733	0,001
	119	246	441	717	0,01
	116	240	431	701	0,05

n	c (количество условий)				
	3	4	5	6	ρ
10	134	278	499	811	0,001
	131	272	487	793	0,01
	128	266	477	777	0,05
11	147	305	546	888	0,001
	144	298	534	869	0,01
	141	292	523	852	0,05
12	160	331	593	965	0,001
	156	324	581	946	0,01
	153	317	570	928	0,05

Тенденция является достоверной, если $L_{\text{эмп}}$ достигает или превышает $L_{0,05}$ и ещё более достоверной, если $L_{\text{эмп}}$ достигает или превышает $L_{0,01}$.

Таблица 10

Критические значения $d_{\max}$ соответствующие уровням статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$ при сопоставлении эмпирического распределения с теоретическим

n	Максимальный модуль разности накопленных частот $d_{\max}$		n	Максимальный модуль разности накопленных частот $d_{\max}$	
	$\rho=0,05$	$\rho=0,01$		$\rho=0,05$	$\rho=0,01$
5	0,6074	0,7279	50	0,1921	0,2302
10	0,4295	0,5147	60	0,1753	0,2101
15	0,3507	0,4202	70	0,1623	0,1945
20	0,3037	0,3639	80	0,1518	0,1820
25	0,2716	0,3255	90	0,1432	
30	0,2480	0,2972	100	0,1358	
40	0,2147	0,2574	>100	$1,36/\sqrt{n}$	$1,63/\sqrt{n}$

Таблица 11

Критерий λ Колмогорова–Смирнова для сопоставления эмпирического распределения с теоретическим (при $n > 50$) или двух эмпирических распределений между собой (при $n > 50$):
уровни статистической значимости разных значений $\lambda_{\text{эмп}}$

λ	λ – последний десятичный знак									
	0	1	2	3	4	5	6	7	8	9
	ρ – десятичные знаки									
0,3	9999	9999	9999	9999	9998	9997	9994	9991	9987	9980
0,4	9971	9960	9945	9926	9902	9874	9840	9799	9753	9699
0,5	9639	9571	9496	9414	9325	9228	9124	9013	8896	8772
0,6	8642	8507	8367	8222	8073	7920	7763	7604	7442	7278
0,7	7112	6945	6777	6608	6440	6271	6103	5936	5770	5605
0,8	5441	5279	5119	4961	4806	4653	4502	4354	4209	4066
0,9	3927	3790	3657	3526	3399	3274	3153	3035	2920	2808
1,0	2700	2594	2491	2392	2295	2202	2111	2023	1938	1856
1,1	1777	1700	1626	1555	1486	1419	1355	1293	1234	1177
1,2	1122	1069	1019	0970	0923	0878	0835	0794	0755	0717
1,3	0680	0646	0613	0581	0551	0522	0494	0468	0443	0419
1,4	0396	0375	0354	0334	0316	0298	0281	0265	0250	0235
1,5	0222	0209	0196	0185	0174	0163	0153	0144	0135	0127
1,6	0119	0112	0105	0098	0092	0086	0080	0075	0070	0066
1,7	0061	0057	0053	0050	0046	0043	0040	0038	0035	0033
1,8	0030	0028	0026	0024	0022	0021	0019	0018	0017	0015
1,9	0014	0013	0012	0011	0010	0010	0009	0008	0007	0007
2,0	0001	0006	0005	0005	0004	0004	0004	0003	0003	0003
2,1	0003	0002	0002	0002	0002	0001	0001	0001	0001	0001
2,2	0001	0001	0001	0001	0000	0000	0000	0000	0000	0000
2,3	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
2,4	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000

По полученному значению $\lambda_{\text{эмп}}$ определяется уровень значимости различий между двумя распределениями.

Таблица 12

Критические точки распределения Стьюдента

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,95
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Уровень значимости α (односторонняя критическая область)					

**Критические значения выборочного
коэффициента корреляции рангов**

n	ρ		n	ρ		n	ρ	
	0,05	0,01		0,05	0,01		0,05	0,01
5	0,94	—	17	0,48	0,62	29	0,37	0,48
6	0,85	—	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,58	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,49	0,51	37	0,33	0,43
14	0,54	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,66	27	0,38	0,49	39	0,32	0,41
16	0,50	0,64	28	0,38	0,48	40	0,31	0,40