	МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

КОМИССАРОВ В. В.

ПРАКТИКУМ ПО МАТЕМАТИЧЕСКИМ МЕТОДАМ В ПСИХОЛОГИИ

Новосибирск 2011

Работа подготовлена на кафедре
высшей математики для студентов,
обучающихся по специальности психология ФГО НГТУ

Комиссаров В. В.
Практикум по математическим методам в психологии: Учеб. пособие. – Новосибирск: НГТУ, 2011. – 102 с.

Данное учебное пособие содержит разработки 4 лабораторных работы по математическим методам в психологии.

Каждая разработка содержит краткие теоретические сведения, необходимые таблицы и варианты индивидуальных лабораторных работ.

Оглавление

Введение

В данном пособии содержатся теоретические сведения, необходимые таблицы и варианты индивидуальных лабораторных работ по математическим методам в психологии.

Приведен список литературы, рекомендованной для более глубокого изучения рассмотренных вопросов.

Это пособие содержит разработки 4 лабораторных работы, достаточно полно характеризующих основные подходы и критерии, используемые при использовании математических методов в психологии.

Правила оформления лабораторных работ

Отчет по лабораторным работам (кроме лабораторной работы №1) сдается на листах формата А4 в рукописном виде или распечатанном на принтере и должен содержать

1. Титульный лист, на котором указывается номер лабораторной работы, её название, вариант, ФИО и номер группы студента.
2. Постановку задачи.
3. Предварительный анализ данных, содержащий обоснование применяемого критерия и проверка 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 применения правила, позволяющего принять истинную и отклонить ложную гипотезы.
8. Выводы по лабораторной работе (математическая статистика).
9. Выводы по решаемой задаче (психология).
10. Список использованной литературы.

Лабораторная работа №1

По данной выборке (см. приложение к лабораторной работе №1) построить:

1) вариационный ряд;
2) статистическое распределение, разбив выборку на 10 равных интервалов;
3) построить полигон относительных частот;
4) выборочное среднее, выборочную дисперсию, исправленную выборочную дисперсию, асимметрию, эксцесс;
5) 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) найти минимальный объём выборки при котором с надёжностью $\gamma = 0,95$ точность оценки математического ожидания a генеральной совокупности по выборочной средней равна $\delta = 0,5$.

Краткая теоретическая справка

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

Статистическим распределением выборки называют перечень вариант 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), $(x_2,n_2),\ldots,(x_k,n_k)$. 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 прямых и получают полигон частот.

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

Выборочной средней $\overline{x}_B$ называют среднее арифметическое значение признака выборочной совокупности.

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

$$\bar{x}_в = \frac{1}{n}(x_1 + x_2 + \ldots + x_n).$$

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

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

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

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

$$D_B = \left(\sum_{i=1}^{k} n_i (x_i - \bar{x}_Г)^2\right)\Big/ n = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{k} n_i x_i\right)^2$$

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

$a_s = m_3/\sigma^3_в$, $e_k = m_4/\sigma^4_в - 3$;

здесь $\sigma_в$ – выборочное среднее квадратическое отклонение; m_3 и m_4 – центральные эмпирические моменты третьего и четвёртого порядков:

$$m_3 = \left(\sum n_i (x_i - \bar{x})^3\right)/n,\ m_4 = \left(\sum n_i (x_i - \bar{x})^4\right)/n,$$

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 $M_k = \left(\sum n_i x_i^k\right) / n$ – 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 находят аргумент t, которому соответствует значение функции Лапласа, равное $\gamma/2$.

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

Пример выполнения лабораторной работы в Ms Excel

86	81	76	80	84	85			
95	77	85	95	89	83			
82	77	76	71	87	68			
89	64	81	90	72	97			
91	75	80	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			
X_{min}=	64	64						
X_{max}=	103	104						
R=	39	40						
h=		4						
Хл	Хпр	Хс	n_i	ν_i	$X_c n_i$	$X^2_c n_i$	$X^3_c n_i$	$X^4_c n_i$
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**

$X_{выб}$=	82,40	M_1=	82,4	m_3=	41,728			
$D_{выб}$=	66,51	M_2=	6856,26667	m_4=	11176			
S^2=	67,63	M_3=	575958,4					
$\sigma_{выб}$=	8,16	M_4=	48835156,3					
S=	8,22							
A_s=	0,08							
E_s=	-0,47							
γ=	0,90							
t=	1,64		Доверительный интервал					
δ=	1,75	(	80,65	84,15	)			
γ=	0,95							
t=	1,96							
δ=	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), в которой первый аргумент – блок в котором располагаются исходные данные, второй аргумент – правые границы частичных интервалов. Функция матричная, для корректного её использования необходимо воспользоваться следующим алгоритмом: 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 воспользовавшись функцией =НОРМОБР(0,5+НАДЕЖНОСТЬ/2;0;1), где НАДЕЖНОСТЬ – значение γ.

Приложение к лабораторной работе №1. Варианты заданий

Вариант 1 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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 (γ = 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

Лабораторная работа №2

Оценки различий между выборками

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, который позволяет 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

$\rho \le 0.01$, 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 варианте все значения первого ряда выше всех значений второго ряда. Различия, безусловно, достоверны, при соблюдении условия, что n_1, $n_2 \ge 11$.

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

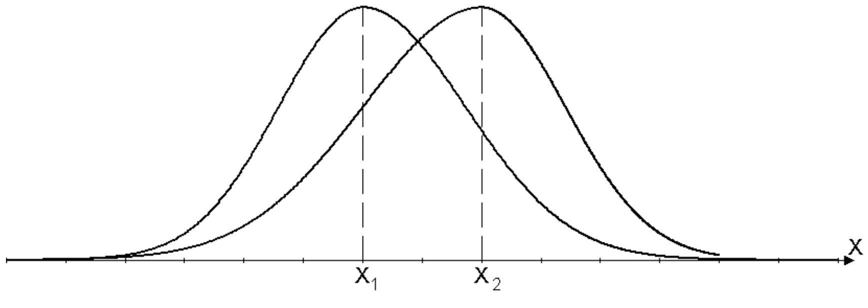


Рис. 1. Вариант соотношения распределений признака в двух выборках, при котором критерий *Q* беспомощен

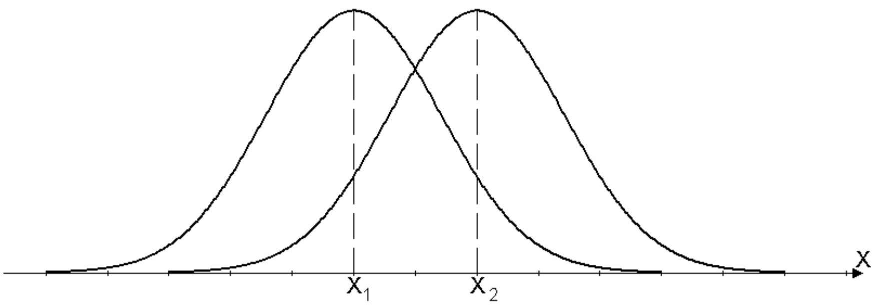


Рис. 2. Вариант соотношения распределений признака в двух выборках,
при котором критерий Q может быть могущественным .

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 должны не совпадать между собой, 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 (рис. 2).

АЛГОРИТМ Подсчет критерия Q Розенбаума

1. Проверить, выполняются ли ограничения: n_1, $n_2 \geq$ ll, $n_1 \approx n_2$.
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. Обозначить полученную величину как S_1.
5. Определить самое низкое (минимальное) значение в выборке 1.
6. Подсчитать количество значений в выборке 2, которые ниже минимального значения выборки 1. Обозначить полученную величину как S_2.

7. Подсчитать эмпирическое значение Q по формуле: $Q_{эмп} = S_1 + S_2$;.

8. По Табл. I определить критические значения Q для данных n_1 и n_2. Если $Q_{эмп}$ равно $Q_{0,05}$ или превышает его, H_0 отвергается.

9. При n_1, $n_2 > 26$ сопоставить полученное эмпирическое значение с $Q_{кр} = 8$ ($\rho \leq 0,05$) и $Q_{кр}=10$ ($\rho \leq 0,01$). Если $Q_{эмп}$ превышает или по крайней мере равняется $Q_{кр} = 8$ H_0 отвергается.

Таблица 1

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

n	**11**	**12**	**13**	**14**	***15***	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**
ρ=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
ρ=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 $Q_{эмп}$ равен или выше критического значения $Q_{0,05}$- и тем более достоверными ($\rho \leq 0{,}01$), если $Q_{эмп}$ равен или выше критического значения $Q_{0,01}$

Задача *1*

У предполагаемых участников психологического эксперимента был 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 факультета и психологического факультета. Показатели вербального интеллекта представлены в Табл. 2.

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

Таблица 2.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									

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 1-м рядом (выборкой, группой) мы называем тот ряд значений, в котором значения, по предварительной оценке, выше, а 2-м рядом - тот, где они предположительно ниже.

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

Эмпирическое значение критерия U отражает то, насколько велика зона совпадения между рядами. Поэтому чем меньше $U_{эмп}$, 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 трудоемким.

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 звучит сложно, 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. В случае, 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 полученные нами 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 расчетной сумм рангов будет свидетельствовать об ошибке, допущенной при начислении рангов или их суммировании. Прежде чем продолжить работу, необходимо найти ошибку и устранить ее.

АЛГОРИТМ Подсчет критерия U Манна-Уитни.

1. Объединить все данные в единый ряд, пометив данные, принадлежащие разным выборкам.
2. Проранжировать значения на карточках, 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 Табл. 3. Если $U_{эмп} > U_{кр0,05}$. H_0 принимается. Если $U_{эмп} \leq U_{кр0,05}$ H_0 отвергается. Чем меньше значения U, тем достоверность различий выше.

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
		$\rho=0{,}01$																	
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											
1	-	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 $U_{эмп}$ ниже или равен $U_{0,05}$ и тем более достоверными ($\rho < 0{,}01$), если $U_{эмп}$ ниже или равен $U_{0,01}$

Задача *2*

Имеются результатам обследования студентов физического и психологического факультетов с помощью методики Д. Векслера для измерения вербального и невербального интеллекта. Показатели невербального интеллекта представлены в Табл. 4.

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

Таблица 4

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			

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									

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, 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 сколько-нибудь 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 ее, то различия между выборками оказываются недостоверными.

Ограничения критерия H

1. При сопоставлении 3-х 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 ($\rho \le 0{,}05$).

Для того, чтобы оказалось возможным диагностировать различия на более высоком уровнем значимости *(*$\rho \le 0{,}01$*)*, 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. Критические значения критерия H и соответствующие им уровни значимости приведены в Табл. III. Таблица предусмотрена только для трех выборок и $\{n_1, n_2, n_3\} \leq 5$. При большем количестве выборок и испытуемых в каждой выборке необходимо пользоваться Таблицей критических значений критерия χ^2 (Табл. IV), поскольку критерий Крускала-Уоллиса асимптотически приближается к распределению χ^2. Количество степеней свободы при этом определяется по формуле:

$\nu = c - 1$, 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 Подсчет 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 на карточках, приписывая меньшему значению меньший ранг. Общее количество рангов будет равняться количеству испытуемых в объединенной выборке.
4. Подсчитать суммы рангов отдельно по каждой группе. Проверить совпадение общей суммы рангов с расчетной.
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 объединенной 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 Табл. 5. Если $H_{эмп}$, равен или превышает критическое значение χ^2, H_0 отвергается. Количество степеней свободы при этом определяется по формуле: $\nu = c – 1$, где c - количество сопоставляемых выборок.

Таблица 5
Критические значения критерия χ^2 для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,0l$ при разном числе степеней свободы ν. Различия между двумя распределениями могут считаться достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,05}$- и тем более достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,01}$.

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			

Задача 3

В эксперименте по исследованию интеллектуальной настойчивости испытуемым предъявлялись сначала разрешимые четырехбуквенные, пятибуквенные- и шестибуквенные анаграммы, а затем неразрешимые анаграммы, время работы над которыми не ограничивалось. Эксперимент проводился индивидуально с каждым испытуемым. Использовалось 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 Табл. 6.

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

Таблица 6.

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						
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			

S - *критерии тенденций Джонкира*

Назначение критерия S

Критерий *S* предназначен для выявления *тенденций* изменения Признака при переходе от выборки к выборке при сопоставлении *трех 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 расчета, 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.

При упорядочивании выборок 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 существенна, то большая часть значений справа должна быть выше. Критерий *S* позволяет определить, преобладают ли 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 $S_{эмп}$ 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 неперекрещивающихся 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 придется искусственно уравнивать выборки, утрачивая при этом часть полученных наблюдений.

Если исследователь хочет избежать этого, ему следует 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. Табл. V для определения критических значений *S*). При большем количестве выборок или наблюдений в них придется пользоваться критерием *H* Крускала-Уоллиса.

АЛГОРИТМ Подсчет критерия S Джонкира

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

2. Занести данные в таблицу, упорядочив их в каждой группе в порядке возрастания 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, просуммировав 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 Табл. 7 для данного количества групп (c) и количества испытуемых в каждой группе (n). Если эмпирическое значение S превышает или по крайней мере равняется критическому значению, H_0 отвергается.

Таблица 7. Критические значения, критерия тенденций S Джонкира для количества групп (c) от трех до шести ($3 \le c \le 6$) и количества испытуемых в каждой группе от двух до десяти ($2 \le n \le 10$)

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

c	*n*								
	2	3	4	5	6	7	8	9	10
ρ = 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
ρ = 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

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. Одним из определяемых качеств была "Авторитетность". В конце 8-часового сеанса диагностических ролевых игр и упражнений проводился социометрический опрос участников группы, в котором они должны были ответить на вопрос: "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 Табл. 8.

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

Таблица **8**

Показатели по шкале Авторитетности в группах с разным социометрическим статусом

Вариант 4

№	Группа 1	Группа 2	Группа 3	Группа 4
1	3	6	7	7
2	6	7	6	9
3	5	7	5	8
4	3	3	5	9
5	6	4	7	7

Вариант 5

№	Группа 1	Группа 2	Группа 3	Группа 4
1	4	7	4	7
2	4	5	4	7
3	3	7	7	7
4	3	7	5	9
5	5	3	6	8

Вариант 2

№	Группа 1	Группа 2	Группа 3	Группа 4
1	6	6	7	8
2	6	7	5	10
3	6	6	8	7
4	4	6	7	8
5	6	4	7	7

Вариант 6

№	Группа 1	Группа 2	Группа 3	Группа 4
1	5	4	5	10
2	3	4	6	9
3	5	4	8	9
4	4	5	7	6
5	4	6	6	9

Вариант 3

№	Группа 1	Группа 2	Группа 3	Группа 4
1	6	6	7	8
2	6	7	5	10
3	6	6	8	7
4	4	6	7	8
5	6	4	7	7

Вариант 7

№	Группа 1	Группа 2	Группа 3	Группа 4
1	4	5	6	8
2	3	6	6	9
3	4	6	6	8
4	5	3	5	7
5	5	7	6	8

Вариант 4

№	Группа 1	Группа 2	Группа 3	Группа 4
1	5	5	7	7
2	5	4	6	9
3	2	5	8	7
4	6	6	5	10
5	2	6	5	9

Вариант 8

№	Группа 1	Группа 2	Группа 3	Группа 4
1	2	7	5	6
2	6	7	5	9
3	5	4	4	7
4	4	3	7	10
5	5	7	8	6

Вариант 9

№	Группа 1	Группа 2	Группа 3	Группа 4
1	6	4	5	6
2	2	5	5	7
3	6	4	5	9
4	3	7	8	6
5	4	3	6	8

Вариант 10

№	Группа 1	Группа 2	Группа 3	Группа 4
1	6	6	8	8
2	3	5	5	10
3	4	3	8	9
4	4	6	7	6
5	3	6	8	9

Вариант 11

№	Группа 1	Группа 2	Группа 3	Группа 4
1	4	7	6	10
2	6	4	4	9
3	3	3	7	9
4	4	4	5	9
5	5	6	4	8

Вариант 12

№	Группа 1	Группа 2	Группа 3	Группа 4
1	3	5	7	9
2	4	6	7	8
3	3	3	5	7
4	6	5	5	6
5	2	4	8	6

Вариант 13

№	Группа 1	Группа 2	Группа 3	Группа 4
1	5	6	6	10
2	3	4	8	9
3	3	4	5	6
4	4	6	7	6
5	4	6	4	6

Вариант 14

№	Группа 1	Группа 2	Группа 3	Группа 4
1	3	6	6	8
2	4	3	8	7
3	4	5	6	6
4	6	6	6	6
5	2	6	7	7

Вариант 15

№	Группа 1	Группа 2	Группа 3	Группа 4
1	5	4	6	8
2	4	6	8	6
3	4	4	7	7
4	6	5	7	9
5	2	6	5	8

Вариант 16

№	Группа 1	Группа 2	Группа 3	Группа 4
1	5	4	5	8
2	4	5	7	5
3	4	4	7	8
4	6	5	9	9
5	3	6	5	8

Лабораторная работа №3
Выявление различий в распределении признака

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*, *трех или более эмпирических* распределений одного и того же признака.

Описание критерия

Критерий χ^2 отвечает на вопрос о том, 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 т. п. мы уже можем применить критерий χ2.

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. Объем выборки должен быть достаточно большим: $n \geq 30$. При n<30 критерий χ^2 дает весьма приближенные 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 накопив определенного минимального числа наблюдений. Если, например, мы хотим проверить наши предположения о том, что частота обращений в телефонную службу Доверия неравномерно распределяются по 7 дням недели, то нам потребуется 5 * 7 = 35 обращений. Таким образом, если количество разрядов *(k)* задано заранее, как в данном случае, минимальное число наблюдений (n_{min}) определяется по формуле: $n_{min} = k * 5$.

3. Выбранные разряды должны "вычерпывать" все 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 неперекрещивающимися: 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

Расчет критерия χ^2

1. Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты (первый столбец).
2. Рядом с каждой эмпирической частотой записать теоретическую частоту (второй столбец).
3. Подсчитать разности между эмпирической и теоретической частотой по каждому разряду (строке) и записать их в третий столбец.
4. Определить число степеней свободы по формуле: $\nu = k - 1$, где k – количество разрядов признака.

   Если $\nu = 1$, внести поправку на "непрерывность".
5. Возвести в квадрат полученные разности, и занести их в четвертый столбец.

6. Разделить полученные квадраты разностей на теоретическую частоту и записать результаты в пятый столбец.
7. Просуммировать значения пятого столбца. Полученную сумму обозначить как $\chi^2_{эмп}$

$$\chi^2 = \sum_{j=1}^{k} \frac{(f_{Эj} - f_T)^2}{f_T}$$

8. Определить по Табл. 9 критические значения для данного числа степеней свободы ν.
   Если $\chi^2_{эмп}$ меньше критического значения, расхождения между распределениями статистически недостоверны. Если $\chi^2_{эмп}$ равно критическому значению или превышает его, расхождения между распределениями статистически достоверны.

При сопоставлении двух и более эмпирических распределений (второй, третий варианты гипотез) все данные объединяются в одну выборку, и все расчеты делаются согласно рассмотренному алгоритму. При этом отдельно выделяется задача нахождения теоретических частот. В общем случае число значений в первой и второй выборках различны и требуется учитывать эту пропорцию. Общая формула подсчёта $f_{теор}$ для сопоставления двух или более эмпирических распределений: $f_{Tij} = \dfrac{\sum_j f_{i,j} \cdot \sum_i f_{i,j}}{\sum_i \sum_j f_{i,j}}$, здесь *i* – номер разряда в выборке; *j* – номер выборки.

Если расположить исходные данные в виде таблицы, где значения в каждой выборке располагаются по столбцам, то в числители полученной формулы имеем: (Сумма частот по соответствующей строке)*(Сумма частот по соответствующему столбцу). В знаменателе (Общее количество наблюдений).

Таблица 9

Критические значения критерия χ^2 для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,0l$ при разном числе степеней свободы ν. Различия между двумя распределениями могут считаться достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,05}$- и тем более достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,01}$.

	ρ			ρ			ρ	
ν	0,05	0,01	ν	0,05	0,01	ν	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	100,425
3	7,815	11,345	37	52,192	59,892	71	91,670	101,621
4	9,488	13,277	38	53,384	61,162	72	92,808	102,816
5	11,070	15,086	39	54,572	62,428	73	93,945	104,010
6	12,592	16,812	40	55,758	63,691	74	95,081	105,202
7	14.067	18,475	41	56,942	64,950	75	96,217	106,393
8	15,507	20,090	42	58,124	66,206	76	97.351	107,582
9	16,919	21,666	43	59,304	67,459	77	98,484	108,771
10	18.307	23,209	44	60,481	68,709	78	99,617	109,958
11	19,675	24,725	45	61,656	69,957	79	100,74	111,144
12	21,026	26,217	46	62,830	71,201	80	101,879	112,329
13	22,362	27,688	47	64,001	72,443	81	103,010	113,512
14	23,685	29,141	48	65,171	73,683	82	104,139	114,695
15	24,996	30,578	49	66,339	74,919	83	105,267	115,876
16	26.296	32,000	50	67,505	76,154	84	106,395	117,057
17	27,587	33,409	51	68,669	77,386	85	107.522	118,236
18	28,869	34,805	52	69,832	78,616	86	108,648	119,414
19	30,144	36,191	53	70,993	79,84'3	87	109,773	120,591
20	31,410	37,566	54	72.153	81,069	88	110,898	121,767
21	32,671	38,932	55	73,311	82,292	89	112,022	122,942
22	33,924	40,289	56	74,468	83,513	90	113,145	124,116
23	35,172	41,638	57	75,624	84,733	91	114,268	125,289
24	36,415	42,980	58	76,778	85.950	92	115,390	126,462
25	37,652	44,314	59	77,931	87.166	93	116,511	127,633
26	38,885	45,642	60	79,082	88,379	94	117,632	128,803
27	40.113	46,963	61	80,232	89,591	95	118.752	129,973
28	41,337	48,278	62	81,381	90,802	96	119,871	131,141
29	42,557	49,588	63	82,529	92,010	97	120,990	132,309
30	43,773	50,892	64	83,675	93,217	98	122,108	133,476
31	44,985	52,191	65	84,821	94,422	99	123,225	134,642
32	46,194	53,486	66	85,965	95,626	100	124,342	135,807
33	47,400	54,776	67	87,108	96,828			
34	48,602	56,061	68	88,250	98,028			

Задача 1

Перед героиней комедии Н. В. Гоголя «Женитьба» Агафьей Тихоновной стоит задача выбора между стоит задача выбора между пятью женихами. Одного она сразу исключила из рассмотрения, потому что он был купеческого звания, как и она сама. А из остальных она не знала, кого выбрать.

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 предпочтенья Агафья Тихоновна от равномерного распределения. Данные по вариантам приведены в таблице 10.

Таблица 10

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.

По условию предыдущей задачи ответить на вопрос, одинаковая ли система предпочтений проявляется во взгляде Агафьи Тихоновны и ее словах? Данные по вариантам приведены в таблице 11.

Таблица 11

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

λ - критерий Колмогорова-Смирнова

Назначение критерия

Критерий X 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 ча*стоты.

Если различия между двумя распределениями существенны, то в какой-то момент разность накопленных частот достигнет критического значения, и мы сможем признать различия статистически достоверными. 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. В то же время, 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 пе-

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

АЛГОРИТМ

Расчет 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: $f^*_{эмп}=f_{эмп}/n$:
   где $f^*_{эмп}$ - эмпирическая частота по данному разряду;
   n - общее количество наблюдений. Занести результаты во второй столбец.
3. Подсчитать накопленные эмпирические частости $\sum f^*_j$ 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^* - эмпирическая частость данного *j-го* разряда. Занести результаты в третий столбец таблицы.
4. Подсчитать накопленные теоретические частости для каждого разряда по формуле:
   $F^*_m=\sum f^*_{Tj} = \sum f^*_{Tj-1} + f^*_{Tj}$,
   где $\sum f^*_{Tj-1}$ - теоретическая частость, накопленная на предыдущие 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-го и 4-го столбцов).
6. Записать в пятый столбец абсолютные величины полученных разностей. Обозначить их как d.
7. Определить по пятому столбцу наибольшую абсолютную величину разности: $d_{max}= d_{эмп}$
8. По Табл. 12 определить или рассчитать критические значения d_{max} для данного количества наблюдений n. Если d_{max} равно критическому значению d или превышает его, различия между распределениями достоверны.

Таблица 12

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

Различия между распределениями могут считаться достоверными, если абсолютная величина максимальной разности $d_{эмп}$ достигает или превышает $d_{0,05}$ и тем более достоверными, если $d_{эмп}$ достигает или превышает $d_{0,01}$.

n	Максимальный модуль разности накопленных		n	Максимальный модуль разности накопленных	
	ρ=0,05	ρ=0,01		ρ=0,05	ρ=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	$1.36/\sqrt{n}$	$1.63/\sqrt{n}$

АЛГОРИТМ
Расчет критерия λ при сопоставлении двух эмпирических распределении

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 частости по каждому разряду для распределения 1 по формуле:
   $f^{*}_{эмп}=f_{эмп}/n_1$, где $f_{эмп}$ - эмпирическая частота в данном разряде;
   n_1 - количество наблюдений в выборке.
   Занести эмпирические частости распределения 1 в третий столбец.
3. Подсчитать эмпирические частости по каждому разряду для распределения 2 по формуле:
   $f^{*}_{эмп}=f_{эмп}/n_2$, где $f_{эмп}$ - эмпирическая частота в данном разряде;
   n_2 - количество наблюдений во 2-й выборке.
   Занести эмпирические частости распределения 2 в четвертый столбец таблицы.
4. Подсчитать накопленные эмпирические частости для распределения 1 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^{*} - эмпирическая частость данного *j-го* разряда.
   Полученные результаты записать в пятый столбец.
5. Подсчитать накопленные эмпирические частости для распределения 2 по той же формуле и записать результат в шестой столбец.
6. Подсчитать разности между накопленными частостями 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:

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

   где n_1 - количество наблюдений в первой выборке;
   n_2 - количество наблюдений во второй выборке.
9. По Табл. 13 определить, какому уровню статистической значимости соответствует полученное значение λ.
   Если $\lambda > \lambda_{0,05}$, различия между распределениями достоверны.

Таблица 13

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 $\lambda_{эмп}$.

λ	λ последний десятичный знак									
	0	1	2	3	4	5	6	7	8	9
	ρ - десятичные знаки									
0,3	99999	99998	99995	99991	99983	99970	99949	99917	99872	99807
0,4	99719	99603	99452	99262	99027	98741	98400	97998	97532	96998
0.5	96394	95719	94969	94147	93250	92282	91242	90134	88960	87724
0,6	86428	85077	83678	82225	80732	79201	77636	76042	74422	72781
0,7	71124	69453	67774	66089	64402	62717	61036	59363	57700	56050
0,8	54414	52796	51197	49619	48063	46532	45026	43545	42093	40668
0,9	39273	37907	36571	35266	33992	32748	31536	30356	29206	28087
1,0	27000	25943	24917	23922	22957	22021	21114	20236	19387	18566
1,1	17772	17005	16264	15550	14861	14196	13556	12939	12345	11774
1,2	11225	10697	10190	09703	09235	08787	08357	07944	07550	07171
1.3	06809	06463	06132	05815	05513	05224	04949	04686	04435	04196
1,4	03968	03751	03545	03348	03162	02984	02815	02655	02503	02359
1,5	02222	02092	01969	01852	01742	01638	01539	01446	01357	01274
1,6	01195	01121	01051	00985	00922	00864	00808	00756	00707	00661
1,7	00618	00577	00539	00503	00469	00438	00408	00380	00354	00330
1,8	00307	00285	00265	00247	00229	00213	00198	00186	00170	00158
1,9	00146	00136	00126	00116	00108	00100	00092	00085	00079	00073
2,0	00016	00062	00057	00053	00048	00045	00041	00038	00035	00032
2,1	00030	00027	00025	00023	00021	00019	00018	00016	00015	00014
2,2	00013	00011	00010	00010	00009	00008	00007	00007	00006	00006
2,3	00005	00005	00004	00004	00004	00003	00003	00003	00002	00002
2,4	00002	00002	00002	00001	00001	00001	00001	00001	00001	00001

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

Задача 3: сопоставление эмпирического распределения с теоретическим

В выборке здоровых лиц мужского пола, студентов технических и военно-технических вузов в возрасте от 19-ти до 22 лет, средний возраст 20 лет, проводился тест Люшера в 8-цветном варианте. Установлено, что желтый цвет предпочитается испытуемыми чаще, чем отвергается (Табл. 14). Можно ли утверждать, что распределение желтого цвета по 8-и позициям у здоровых испытуемых отличается от равномерного распределения?

Таблица 14

Эмпирические частоты попадания желтого цвета на каждую из 8 позиций

Вариант	Позиции желтого 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, полученные в двух выборках, полученных различными исследователями при проведении теста Люшера в 8-цветном варианте.

Таблица 15

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

Разряды	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

Разряды	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

Разряды	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

Разряды	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

Лабораторная работа №4

Аппрксимация опытных данных методом наименьших квадратов

Постановка задачи.

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

$P(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + ... + c_M x^M, M \le N$ такую, что

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

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

$$\begin{aligned}
&c_0(N+1) + c_1\sum_{i=0}^{N} x_i + c_2\sum_{i=0}^{N} x_i^2 + c_3\sum_{i=0}^{N} x_i^3 + ... + c_M\sum_{i=0}^{N} x_i^M = \sum_{i=0}^{N} y_i,\\
&c_0\sum_{i=0}^{N} x_i + c_1\sum_{i=0}^{N} x_i^2 + c_2\sum_{i=0}^{N} x_i^3 + c_3\sum_{i=0}^{N} x_i^4 + ... + c_M\sum_{i=0}^{N} x_i^{M+1} = \sum_{i=0}^{N} x_i y_i,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&c_0\sum_{i=0}^{N} x_i^M + c_1\sum_{i=0}^{N} x_i^{M+1} + c_2\sum_{i=0}^{N} x_i^{M+2} + c_3\sum_{i=0}^{N} x_i^{M+3} + ... + c_M\sum_{i=0}^{N} x_i^{2M} = \sum_{i=0}^{N} x_i^M y_i.
\end{aligned} \qquad (1)$$

В матричном виде система (1) имеет вид $A_M^T A_M C = A_M^T Y$, где C - вектор коэффициентов $\vec{C} = \{c_i\}_{i=0,M}$, Y - вектор значений функции $Y = \{f(x_i)\}_{i=0,N}$. Матрица A_M состоит из M+1 столбца матрицы Вандермонда $A = \begin{pmatrix} 1\, x_0\; x_0^2\; x_0^3 ... x_0^N \\ 1\, x_1\; x_1^2\; x_1^3 ... x_1^N \\ .... \\ 1\, x_N\; x_N^2\; x_N^3 ... x_N^N \end{pmatrix}$.

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) возможны 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, нужным для возможности построения 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* задается уравнением $y = ae^{bx}$, то логарифмирование его дает
$ln(y) = bx + ln(a)$.

Полагая $ln(y) = Y$, $ln(a) = A$, $b = B$ получим линейную функцию: $Y=A + Bx$, откуда видно, что, 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 изобразятся прямыми линиями.

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 16 приведены различные виды функций и линеаризующие их преобразования.

Таблица 16. 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 17) 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 17. 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
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	8,66
2	7,14
3	6,58
4	6,53
5	5,55
6	5,58
7	5,48
8	5,37
9	5,03
10	4,74

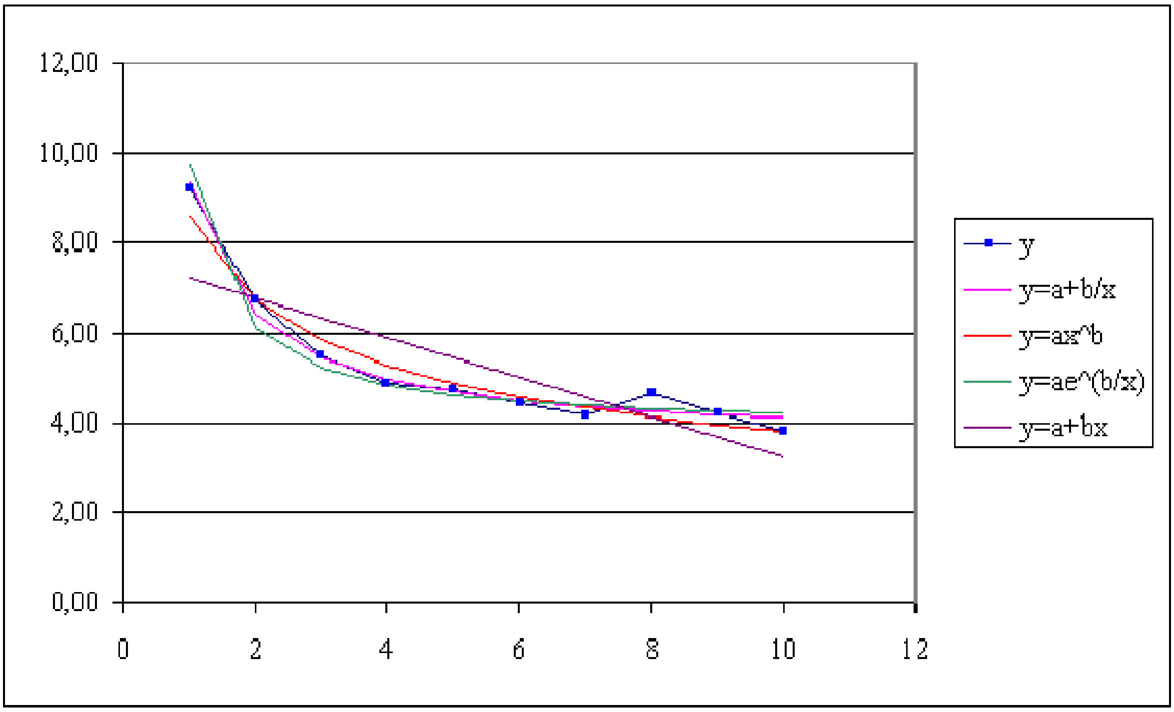
x	y	X=1/x	Y						
						y=a+b/x			
1	9,22	1,00	9,22	1,00	1,00		10	2,92897	52,2539
2	6,72	0,50	6,72	1,00	0,50		2,928968	1,54977	19,3316
3	5,48	0,33	5,48	1,00	0,33				
4	4,88	0,25	4,88	1,00	0,25	a=	3,52082		
5	4,72	0,20	4,72	1,00	0,20	b=	5,819707		
6	4,44	0,17	4,44	1,00	0,17				
7	4,18	0,14	4,18	1,00	0,14	I15:=МУМНОЖ(ТРАНСП(F15:G24);F15:G24)			
8	4,63	0,13	4,63	1,00	0,13	K15:=МУМНОЖ(ТРАНСП(F15:G24);D15:D24)			
9	4,19	0,11	4,19	1,00	0,11	I18:=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	52,2539
2,00	6,72	1,00	2,00		55	385	250,812
3,00	5,48	1,00	3,00				
4,00	4,88	1,00	4,00	a=	7,664373		
5,00	4,72	1,00	5,00	b=	-0,443451		
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 функцией y=a+b/x

Линейная корреляция. Линии регрессии

Говорят, что 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 задается 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 $\overline{y}_{x_i} = \dfrac{\sum_j n_{i,j} y_j}{n_{x_i}}$. Если теперь точки $\left(x_i ; \overline{y}_{x_i}\right)$ выровнять по методу наименьших квадратов вдоль кривой $y=f(x,c_1, c_2, \ldots, c_m)$, то последняя называется *линией регрессии* y на x, а ее уравнение — *уравнением регрессии* y на x.

Аналогично определяется линия регрессии x на y. Наиболее простыми и наиболее важными случаями кривых регрессий являются прямые линии.

Угловой коэффициент 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:

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

где $\overline{XY} = \dfrac{\sum_i \sum_j n_{i,j} x_i y_j}{N}$ среднее значение произведения признаков X и Y; $\overline{X}$ и $\overline{Y}$ — их средние значения, а σ_x^2 и σ_y^2 - их дисперсии.

Уравнения прямых регрессий имеют вид: $y - \overline{Y} = \rho_{y/x}(x - \overline{X}); \; x - \overline{X} = \rho_{x/y}(y - \overline{Y}).$

Коэффициентом линейной корреляции признаков X и Y называется величина:

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

В этом случае уравнения прямых регрессий можно записать:

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

Коэффициент линейной корреляции г обладает следующими свойствами:

1) $-1 \le r_{xy} \le 1$;

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

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

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

Сила связи не зависит от ее направленности и определяется по абсолютному значению коэффициента корреляции. Максимальное возможное абсолютное значение коэффициента корреляции $r = 1{,}00$; минимальное $r = 0$.

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

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

Частная классификация корреляционных связей:

1) высокая значимая корреляция	при r, соответствующем уровню статистической значимости $\rho \le 0{,}01$
2) значимая корреляция	при r, соответствующем уровню статистической значимости $\rho \le 0{,}05$
3) тенденция достоверной связи	при r, соответствующем уровню статистической значимости $\rho \le 0.10$
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 объеме выборки. Чем больше объем выборки, Тем меньшей величины коэффициента корреляции оказывается достаточно, чтобы корреляция была признана достоверной. В результате при малом объеме выборки может оказаться так, что сильная корреляция окажется недостоверной. В то же время при больших объемах выборки даже слабая корреляция может оказаться 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 $r_в$, который является случайной величиной.

Пусть двумерная генеральная совокупность *(X, Y)* распределена нормально. Из этой совокупности извлечена выборка объема *n* и по ней найден выборочный коэффициент корреляции $r_в \neq 0$. Требуется проверить нулевую гипотезу H_0: $r_в = 0$ 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: $r_в \neq 0$, *надо вычислить наблюдаемое значение критерия*

$$T_{набл} = r_в\sqrt{n-2}\,/\,\sqrt{1-r_в^2}$$

и по таблице 18 *критических точек распределения Стьюдента, по заданному уровню значимости α и числу степеней свободы* $k = n - 2$ *найти критическую точку* $t_{кр}(\alpha; k)$ *двусторонней критической области. Если* $|T_{набл}| < t_{кр}$ *— нет оснований отвергнуть нулевую гипотезу. Если* $|T_{набл}| > t_{кр}$ *— нулевую гипотезу отвергают.*

Таблица 18. Критические точки распределения Стьюдента

Число степеней свободы	Уровень значимости α (двусторонняя критическая об-				
	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)					

Если связь между *X* и *Y* установлена, то линейное приближение *Y* на *X* дается формулой *линейной регрессии:* $y-\overline{Y}=\rho_{y/x}(x-\overline{X})$ или $y = ax + b$. Линейное приближение *X* на *Y* дается формулой линейной регрессии: $y-\overline{Y}=\rho_{y/x}(x-\overline{X})$ 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 служит выборочный коэффициент корреляции. Ля оценки 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:

$\eta_{xy} = \sigma_{межгр} / \sigma_{общ}$, или в других обозначениях:

$$\eta_{yx} = \sigma_{\bar{y}_x} / \bar{\sigma}_y \ .$$

Здесь $\sigma_{\bar{y}_x} = \sqrt{D_M} = \sqrt{\left(\sum n_x \left(\bar{y}_x - \bar{y}\right)^2 \ ) / n\right)}$; $\sigma_{\bar{y}} = \sqrt{D_o} = \sqrt{\left(\sum n_y \left(y - \bar{y}\right)^2 \ ) / n\right)}$,

где n – объем 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.

Аналогично вычисляется выборочное корреляционное отношение X к Y: $\eta_{xy} = \sigma_{\bar{x}_y} / \bar{\sigma}_x$.

Выборочное корреляционное отношение обладает следующими свойствами:

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

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: $\eta \ge |r_в|$;

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

Задача 2

По данным корреляционной таблицы (Таблица 19)

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;*

4) оценить тесноту корреляционной связи с помощью корреляционного отношения.

Таблица 19. Корреляционные таблицы

Вариант 1

X\Y	130	150	170	190	210	230	250	m_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
m_y	*7*	*9*	*12*	*13*	*13*	*7*	*3*	64

Вариант 2

X\Y	70	80	90	100	110	120	130	m_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
m_y	*9*	*10*	*15*	*19*	*16*	*11*	*5*	85

Вариант 3

X\Y	100	115	130	145	160	175	190	m_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
m_y	*12*	*8*	*25*	*17*	*25*	*15*	*3*	105

Вариант 4

X\Y	25	35	45	55	65	75	85	m_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
m_y	*20*	*30*	*6*	*29*	*29*	*25*	*6*	145

Вариант 5

X\Y	50	65	80	95	110	125	140	m_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
m_y	*20*	*28*	*16*	*21*	*20*	*23*	*4*	132

Вариант 6

X\Y	30	50	70	90	110	130	150	m_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
m_y	*31*	*34*	*41*	*29*	*40*	*19*	*6*	200

Вариант 7

X\Y	100	115	130	145	160	175	190	m_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
m_y	*12*	*13*	*13*	*25*	*15*	*16*	*4*	98

Вариант 8

X\Y	10	20	30	40	50	60	70	m_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
m_y	*10*	*22*	*16*	*18*	*24*	*20*	*6*	116

Вариант 9

X\Y	70	85	100	115	130	145	160	m_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
m_y	*4*	*20*	*16*	*19*	*34*	*18*	*7*	118

Вариант 10

X\Y	35	50	65	80	95	110	125	m_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
m_y	*42*	*18*	*31*	*38*	*26*	*19*	*4*	178

Вариант 11

X\Y	50	65	80	95	110	125	140	m_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
m_y	*20*	*20*	*21*	*19*	*18*	*17*	*3*	118

Вариант 12

X\Y	25	40	55	70	85	100	115	m_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
m_y	*22*	*19*	*22*	*21*	*26*	*20*	*3*	133

Вариант 13

X\Y	80	95	110	125	140	155	170	m_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
m_y	*35*	*18*	*30*	*38*	*33*	*14*	*12*	180

Вариант 14

X\Y	25	35	45	55	65	75	85	m_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
m_y	*13*	*37*	*16*	*32*	*24*	*20*	*15*	157

Вариант 15

X\Y	50	65	80	95	110	125	140	m_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
m_y	*9*	*14*	*16*	*13*	*24*	*21*	*5*	102

Вариант 16

X\Y	30	50	70	90	110	130	150	m_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
m_y	*22*	*23*	*12*	*17*	*24*	*8*	*1*	107

Пример нахождения коэффициента корреляции и уравнений линейной регрессии в Ms Excel.

Y \ X	10	15	20	25	30	35	Сумма:
35	5	1	0	0	0	0	**6**
45	0	6	2	0	0	0	**8**
55	0	0	5	40	5	0	**50**
65	0	0	2	8	7	0	**17**
75	0	0	0	4	7	8	**19**
Сумма:	**5**	**7**	**9**	**52**	**19**	**8**	**100**

Y \ X	10	15	20	25	30	35	Условные средние X_y:
35	1750	525	0	0	0	0	*10,833*
45	0	4050	1800	0	0	0	*16,250*
55	0	0	5500	55000	8250	0	*25,000*
65	0	0	2600	13000	13650	0	*26,471*
75	0	0	0	7500	15750	21000	*31,053*
Условные средние Y_x:	*35,000*	*43,571*	*55,000*	*58,077*	*66,053*	*75,000*	**150375**

x	n	xn	(x^2)n
10	5	50	500
15	7	105	1575
20	9	180	3600
25	52	1300	32500
30	19	570	17100
35	8	280	9800
Сумма:	**100**	**2485**	**65075**

y	n	yn	(y^2)n
35	6	210	7350
45	8	360	16200
55	50	2750	151250
65	17	1105	71825
75	19	1425	106875
Сумма:	**100**	**5850**	**353500**

Ячейка	**Ф-ла, операция**
A1:G6	Исходные данные
B7	=СУММ(B2:B6), копировать в C7:G7
H2	=СУММ(B2:G2), копировать в H3:H7
B10	=B2*$A10*B$9 копировать в B10:G14
B15	=СУММ(B10:B14)/(B7*B9) копировать в C15:G15
H10	=СУММ(B10:G10)/(A10*H2) копировать в H11:H14
H15	=СУММ(B10:G14)

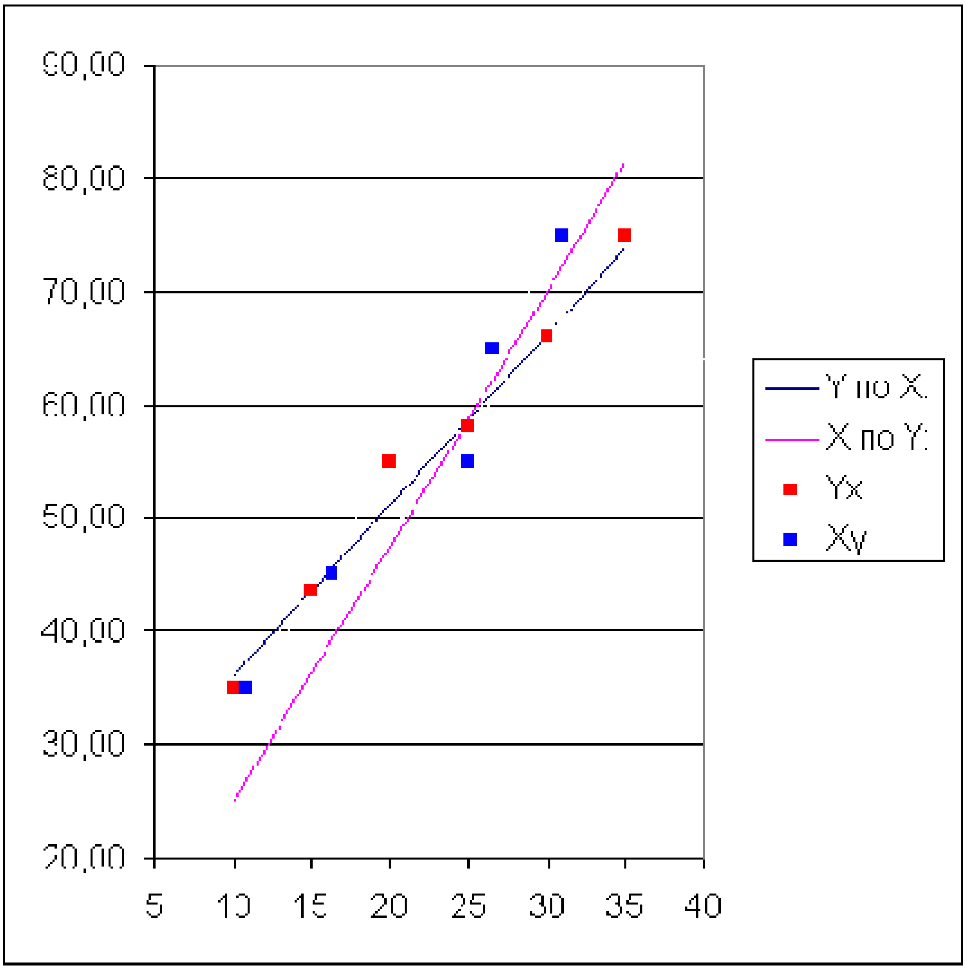
В результате:

M(X)=	24,85	=C24/B24
M(Y)=	58,50	=C33/B33
M(XY)=	1503,75	=H15/H7
M(X^2)=	650,75	=D24/B24
M(Y^2)=	3535,00	=D33/B33
D(X)=	33,23	=B39-B36^2
D(Y)=	112,75	=B40-B37^2
s(X)=	5,76	=КОРЕНЬ(B41)
s(Y)=	10,62	=КОРЕНЬ(B42)
r=	**0,817**	=(B38-B36*B37)/(B43*B44)

Коэффициенты прямых регрессий:

X по Y:	k=0,444	b=-1,105
Y по X:	k=1,506	b=21,088

X	Y по X:
10	36,14
15	43,67
20	51,20
25	58,73
30	66,25
35	73,78
X	X по Y:
10	25,03
15	36,30
20	47,57
25	58,84
30	70,11
35	81,38
X	Y_x
10	35,00
15	43,57
20	55,00
25	58,08
30	66,05
35	75,00
X_y	Y
10,83	35
16,25	45
25,00	55
26,47	65
31,05	75



Ранговая корреляция. Коэффициент ранговой корреляции r_s *Спирмена*

Коэффициент ранговой корреляции рекомендуется применять в тех случаях, когда нам 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 подсчета 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 (Табл-XVI), а именно $N<40$.
2. Коэффициент ранговой корреляции Спирмена r_s при большом количестве одинаковых рангов по одной или обеим сопоставляемым переменным дает огрубленные значения. В идеале оба коррелируемых ряда должны представлять собой две последовательности несовпадающих

значений. В случае, если это условие не соблюдается, необходимо вносить поправку на одинаковые ранги.

АЛГОРИТМ. Расчет 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 столбец таблицы по порядку номеров испытуемых или признаков.
3. Проранжировать значения переменной В, в соответствии с теми же правилами. Занести ранги во второй столбец таблицы по порядку номеров испытуемых или признаков.
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 четвертый 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 - объем каждой группы одинаковых рангов в ранговом ряду А;

   b - объем каждой группы одинаковых рангов в ранговом ряду В.
8. Рассчитать коэффициент ранговой корреляции r_s. по формуле:

   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. 20 критические значения r_s для данного N. Если r_s превышает критическое значение или по крайней мере равен ему, корреляция достоверно отличается от 0.

Таблица 20

Критические значения выборочного коэффициента корреляции рангов. Связь достоверна, если $r_{s\ эмп} \geq r_{s\ 0.05}$ и тем более достоверна, если $r_{s\ эмп} \geq r_{s\ 0.01}$

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

Задача 3

В исследовании, моделирующем деятельность авиадиспетчера, группа испытуемых, студентов физического факультета проходила подготовку перед началом работы на тренажере. 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 21. Количество ошибок, допущенных испытуемыми в тренировочной сессии, с показателями вербального и невербального интеллекта, измеренными по методике Д. Векслера

Вариант 1

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	15	127	96
2	44	126	109
3	50	139	91
4	7	132	94
5	47	130	93
6	26	123	106
7	49	139	110
8	5	130	105
9	10	124	102
10	44	125	100

Вариант 2

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	33	135	96
2	17	127	96
3	33	130	110
4	25	130	96
5	39	124	109
6	10	125	109
7	6	135	97
8	39	137	97
9	15	136	100
10	15	123	102

Вариант 3

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	8	122	105
2	42	132	96
3	2	126	97
4	11	138	93
5	8	127	97
6	41	138	107
7	3	130	103
8	34	120	104
9	18	127	92
10	22	133	98

Вариант 4

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	26	137	103
2	36	121	98
3	4	131	94
4	27	132	95
5	17	124	107
6	27	124	95
7	13	130	106
8	47	133	99
9	44	137	94
10	22	138	108

Вариант 5

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	16	132	97
2	14	137	94
3	19	131	102
4	34	130	104
5	28	123	108
6	26	126	91
7	16	138	101
8	23	133	101
9	31	124	91
10	9	134	101

Вариант 6

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	24	130	101
2	28	136	97
3	47	121	100
4	48	133	98
5	5	134	104
6	6	130	95
7	34	140	108
8	27	126	107
9	8	133	107
10	43	129	109

Вариант 7

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	37	130	104
2	49	123	93
3	21	131	107
4	1	134	102
5	12	138	97
6	16	129	94
7	50	125	96
8	24	140	92
9	50	133	98
10	17	129	92

Вариант 8

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	26	128	100
2	35	138	100
3	14	121	100
4	20	127	100
5	34	122	103
6	40	126	100
7	19	130	103
8	20	138	106
9	21	129	94
10	8	135	95

Вариант 9

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	20	129	96
2	44	121	92
3	46	124	93
4	4	135	100
5	13	138	96
6	30	134	99
7	3	126	104
8	50	133	106
9	21	121	110
10	36	123	108

Вариант 10

Испытуемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	3	126	105
2	23	120	107
3	46	123	91
4	36	127	109
5	19	131	108
6	21	136	103
7	46	130	107
8	27	124	102
9	14	124	91
10	7	120	93

Вариант 11

Испы-туемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	30	121	100
2	16	128	106
3	20	135	109
4	11	124	93
5	28	138	106
6	42	122	97
7	28	134	92
8	10	137	102
9	18	131	92
10	47	128	101

Вариант 12

Испы-туемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	27	128	103
2	29	133	94
3	25	127	99
4	16	126	102
5	3	139	103
6	5	122	95
7	49	126	104
8	17	137	91
9	30	131	96
10	46	129	96

Вариант 13

Испы-туемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	43	137	105
2	26	138	109
3	24	124	96
4	3	139	102
5	36	138	99
6	48	129	93
7	43	133	93
8	10	137	91
9	35	138	107
10	9	123	102

Вариант 14

Испы-туемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	15	127	96
2	3	138	101
3	30	123	98
4	35	124	103
5	44	124	96
6	28	126	92
7	0	135	107
8	31	134	94
9	13	139	94
10	2	134	109

Вариант 15

Испы-туемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	14	131	106
2	29	126	94
3	47	129	91
4	45	131	98
5	31	135	97
6	43	121	104
7	6	130	97
8	24	132	93
9	37	131	93
10	30	130	102

Вариант 16

Испы-туемый	Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	36	128	95
2	20	138	92
3	12	128	103
4	19	130	92
5	36	123	91
6	11	129	97
7	49	138	96
8	11	122	95
9	39	129	98
10	31	131	105

Литература

1. *Глосс Дяс„ Стенли Длс.* Статистические методы в педагогике и психологии. М.,1976.
2. *Готтсданкер Роберт.* Основы психологического эксперимента. М., 1982.
3. *Гусев А.М., Измайлов Ч.А., Михалевская М.Б.* Измерение в психологии. М., 1997.
4. *Корнилова Т.В.* Введение в психологический эксперимент. М., 1997.
5. *Стивенс С.С.* Экспериментальная психология. М., 1960. Т. I.
6. *Суходольский Г.В.* Основы математической статистики для психологов. Л., 1972.
7. *Урбах В.Ю.* Математическая статистика для биологов и медиков. М., 1963.
8. *Яноши Л.* Теория и практика обработки результатов измерений. М., 1968.
9. *Битинас Б.* Многомерный анализ в педагогике и педагогической психологии. Вильнюс, 1971.
10. *Урбах В.Ю.* Биометрические методы. Статистическая обработка опытных данных в биологии, сельском хозяйстве и медицине. М., 1964.
11. *Артемьева Е.Ю., Мартынов Е.М.* Вероятностные методы в психологии. М., 1975.
12. *Кендэл М.* Ранговые корреляции. М., 1975.
13. *Сидоренко Е.В.* Методы математической обработки в психологии. СПб, 2000
14. *Урбах В.Ю.* Статистический анализ в биологических и медицинских исследованиях. М., 1975.
15. *Михеев В. И.* Методика получения и обработки экспериментальных данных в психолого-педагогических исследованиях. М., 1986.
16. *Наследов А. Д.* Математические методы психологического исследования. – СПб : Речь, 2007.
17. *Кутейников А. Н.* Математические методы в психологии. Учебное пособие. – СПб : Речь, 2008. – 172 с.