

Основні поняття математичної статистики.

У психології часто використовують різні числа. Змінні, які є результатами замірів, називають варіантами (вони варіюються, тобто змінюються) і їх позначають X_i . Усі значення змінної, які розташовані в одному ряду у порядку зростання або зменшення, утворюють варіаційний ряд. Кількість повторень однакових результатів у складі варіаційного ряду називається частотою цього значення змінної.

Нормальний розподіл даних. Одиначний розподіл даних як стандарт.

Для того, щоб від розподілу випадкових подій перейти до кривої нормального розподілу, необхідно зробити лише один крок. Нормальна крива — це частотний розподіл подій, коли кількість їх дуже велика.

Для вибірових даних формула підібраної нормальної кривої виглядає так:

$$Y = \frac{1}{\sqrt{2\pi S^2}} e^{-\frac{1}{2} \left(\frac{X-M}{S} \right)^2}.$$

Приклад. $M = 4,0$; $S = 1,0$; $e^0 = 1,0$; $\pi = 3,14$.

$$Y = \frac{1}{\sqrt{2\pi \cdot 1^2}} e^{-\frac{1}{2} \left(\frac{4-4}{1} \right)^2} = \frac{1}{\sqrt{2 \cdot 3,1461 \cdot 1}} e^{-\frac{1}{2} (0)^2} = \frac{1}{\sqrt{6,2832}} e^0 = \frac{1}{2,5066} = 0,399.$$

X	1	2	3	4	5	6	7
Y	0,004	0,054	0,242	0,399	0,242	0,054	0,004

На практиці часто використовують робочу формулу:

$$p' = \frac{n \cdot k}{\sigma} \cdot f(x),$$

де p' — теоретична частота; n — об'єм ряду; k — класовий проміжок; σ — середнє квадратичне відхилення; $f(x)$ — перша функція нормального відхилення, яку знаходять за таблицями ординат нормальної кривої [22];

$$X = \frac{W - M}{\sigma},$$

де M — нормоване відхилення середин класів.

Розглянемо на прикладі вирівнювання емпіричних варіаційних кривих за нормальним законом Муавра, Лапласа, Гауса. Досліджували вибірку ($N = 90$) працівників на визначення рівня інтелектуального розвитку (табл. 5.5; табл. 5.6; табл. 5.7).

Таблювання даних: ранговий порядок, розподіл частот.

Таблица 5.5. Результати досліджень рівня інтелектуального розвитку працівників-регулювальників кольорових телевізорів

№ з/п	Працівник	<i>IQ</i>	№ з/п	Працівник	<i>IQ</i>	№ з/п	Працівник	<i>IQ</i>
1	К-ч О.	115	31	Б-о Т.	116	61	Г-с П.	85
2	Б-а М.	142	32	Б-к С.	116	62	Г-ч Р.	122
3	Д-а М.	110	33	В-а М.	110	63	Д-а О.	133
4	М-н С.	105	34	В-а Г.	132	64	Д-к С.	102
5	Л-к О.	109	35	Г-а Р.	104	65	І-й І.	95
6	П-в З.	99	36	Д-к Р.	90	66	Н-к В.	115
7	Г-ь О.	125	37	О-а Л.	93	67	П-д Г.	128
8	Б-к Н.	118	38	К-к В.	102	68	Р-ч М.	89
9	Г-к Ю.	136	39	К-й В.	107	69	Д-ч Г.	114
10	Б-й А.	107	40	К-к М.	112	70	Г-о В.	101
11	Н-к Д.	126	41	М-к С.	122	71	Д-в І.	97
12	В-н О.	90	42	С-а Г.	117	72	Ж-к С.	86
13	Д-р І.	120	43	Г-ч Л.	131	73	З-ч О.	118
14	К-т Н.	117	44	Н-а І.	117	74	В-о І.	116
15	Д-в І.	112	45	Р-й Г.	148	75	К-а С.	110
16	Г-й А.	106	46	П-а Р.	106	76	К-а О.	134
17	В-н А.	129	47	Ш-н Ю.	126	77	К-в Б.	96
18	К-а І.	94	48	Ш-к М.	91	78	К-к М.	117
19	Ф-а О.	116	49	Ю-в І.	94	79	К-к І.	88
20	Г-а І.	144	50	М-к І.	100	80	В-й С.	127
21	Я-к В.	111	51	В-й С.	108	81	Г-о Н.	111
22	Г-ч А.	108	52	Х-а Г.	113	82	В-а І.	103
23	Д-й М.	106	53	Ш-а Н.	123	83	Б-к Н.	96
24	Д-а Г.	93	54	Н-д Л.	116	84	Г-к А.	114
25	С-а І.	115	55	Ж-д Л.	117	85	Г-й П.	87
26	Р-ч І.	121	56	Г-а Б.	132	86	Г-о О.	113
27	П-а Г.	138	57	К-н О.	112	87	Г-о О.	100
28	Л-а Л.	106	58	Ю-н В.	100	88	Н-р Т.	95
29	Ш-х І.	127	59	Б-ц І.	128	89	О-к В.	75
30	В-н Г.	91	60	В-н В.	117	90	Л-о М.	76

Таблиця 5.6. Складання варіаційного ряду первинних даних

Варіаційний ряд первинних даних									
1	2	3	4	5	6	7	8	9	10
148	144	142	138	136	134	133	132	132	131
129	128	128	127	127	126	126	125	123	122
122	121	120	118	118	117	117	117	117	117

1	2	3	4	5	6	7	8	9	10
117	116	116	116	116	116	115	115	115	114
114	113	113	112	112	111	111	110	110	110
109	108	108	107	107	106	106	106	106	105
104	103	102	102	101	100	100	100	99	97
96	96	95	95	94	94	93	93	91	91
90	90	89	88	87	86	85	76	75	75

Таблиця 5.7. Групування класів

Дати	Кількість класів
6—11	4
12—22	5
23—46	6
47—93	7
94—187	8

Розмах

$$\rho = \max - \min = 148 - 76 = 72,$$

$$r = 1 + 3,3 \log 90 = 7.$$

Величина класів:

$$k = \frac{\rho}{r} = \frac{90}{8} = 11,25;$$

$$p' = \frac{n \cdot k}{\sigma} \cdot f(x),$$

де p' — теоретична частота; n — об'єм ряду; k — класовий проміжок; σ — стандартне відхилення; $f(x)$ — перша функція нормального відхилення, яку знаходять за таблицею ординат нормальної кривої [24].

Вирівнювання емпіричної варіаційної кривої за нормальним законом (табл. 5.8).

Спосіб вирівнювання для тих випадків, коли емпіричний розподіл попередньо прийнятий за випадкову форму виявлення закону нормального розподілу Гауса-Лапласа, ґрунтується на застосуванні значень ординат нормальної кривої $f(x)$ і робочої формули $p' = \frac{N_k}{\sigma} \cdot f(x)$. Значення $f(x)$ першої функції нормального відхилення наведені в математичній табл. 5.5 [22, с. 63].

Таблица 6.8. Вирівнювання емпіричної варіаційної кривої за нормальним законом

Класи		Рахунок	Емпіричні частоти, p	$W - M$	$x = \frac{W - M}{\sigma}$	$f(x)$, див. [22, табл. 5]	Теоретичні частоти	
межі класів	середина варіації, W						$\frac{n \cdot h}{\sigma} f(x)$	p'
148—137	142,5	xxxx	4	42,5	2,93	0,0054	0,366	0,4
136—125	130,5	xxxxxxxxxxxxxx	14	30,5	2,10	0,043	2,91	3
124—113	118,5	xxxxxxxxxxxxxxxxxxxxxx	25	18,5	1,27	0,178	12,08	13
112—101	106,5	xxxxxxxxxxxxxxxxxxxxxx	22	6,5	0,448	0,363	24,57	25
100—89	94,5	xxxxxxxxxxxxxxxxxxxxxx	18	-5,5	0,379	0,156	13,59	13
88—77	82,5	xxxx	4	-17,5	1,20	0,194	10,17	10
76—65	70,5	xxx	3	-29,5	2,03	0,050	3,39	4
64—53	58,5		0	-41,5	2,86	0,0066	0,448	0,5

Алгоритм вирівнювання емпіричних варіаційних кривих за нормальним законом Гауса-Лапласа (рис. 5.1):

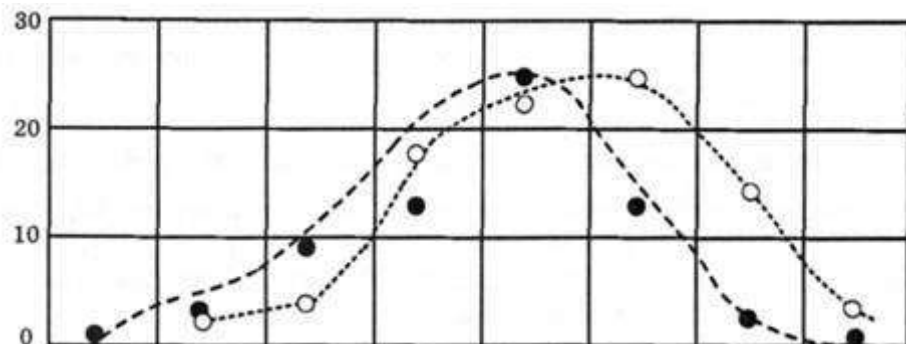


Рис. 5.1. Обчислення емпіричної варіаційної кривої за нормальним законом

Приклад. $n = 90$; $k = 11$; $M = 100,5$.

$$\sigma = \frac{X_{\max} - X_{\min}}{k} = \frac{72}{4,94} = 14,574;$$

$$\frac{n \cdot k}{\sigma} = \frac{90 \cdot 11}{14,574} = 67,929.$$

Співвідношення різних типів шкал оцінювання подано в табл. 5.9:

Таблица 5.9. Співвідношення різних типів шкал оцінювання

Варіації	58	70	82	94	106	118	130	142
Емпіричні частоти, p	—	3	4	18	22	25	14	4
Теоретичні частоти, p'	0,5	4	10	13	25	13	3	0,4
Стандартне відхилення	-4σ	-3σ	-2σ	-1σ	0	$+1\sigma$	$+2\sigma$	$+3\sigma$
Z-оцінки	-4	-3	-2	-1	0+1	+2	+2	+3
T-оцінки	10	20	30	40	50	60	70	80
IQ-стандарт	40	55	70	85	100	115	130	145
ST-стандарт	0	1	2,5	4	5,5	7	8,5	10

Середнє арифметичне (середнє значення або вибіркове середнє) дорівнює сумі всіх значень варіанти, поділений на кількість членів варіаційного ряду (n). Його визначають за формулою:

$$\bar{X} = (X_1 + X_2 + X_3 + \dots + X_N) / N \text{ або } \bar{X} = \left(\sum_{i=1}^N X_i \right) / N.$$

Медіана — це значення варіанти, яке ділить варіантний ряд навпіл. Місце розташування цього значення визначають за формулою:

$$\text{Місце медіани} = (N + 1) / 2.$$

Мода — значення варіаційного ряду, яке трапляється найчастіше.

Перелічені три величини є мірами центральної тенденції і їх використовують залежно від характеру розподілу значень досліджуваної змінної і завдань, що стоять перед діагностом.

Для оцінки зміни значень змінної використовують такі характеристики, як дисперсія і середньоквадратичне (або стандартне відхилення).

Дисперсія дорівнює квадрату середніх відхилень значення варіанти від середнього значення. Вона є однією з характеристик індивідуальних результатів розкиду значень досліджуваної змінної біля середнього значення.

Значення дисперсії використовують у різних статистичних розрахунках, але не має безпосереднього спостережливого характеру. Величиною, яка безпосередньо пов'язана зі змістом змінної, що спостерігається, є середньоквадратичне відхилення. Воно дорівнює квадратному кореню з дисперсії і визначається за формулою:

$$\sigma = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N (\bar{X} - X_i)^2},$$

(в літературі часто позначають як *SD*).

Стандартне відхилення широко використовують як міру розкиду для різних характеристик.

Побудова гикали за даними експерименту.

Для того щоб надати експериментальним даним наочну форму, часто треба подати їх в одній із шкал оцінювання. Розглянемо таку процедуру на прикладі.

Припустимо, що Ви провели обстеження 65 працівників за допомогою тесту інтелекту й отримали такі результати (в балах):

59	48	53	47	57	64	62	62	65	57	57
	81	83	48	65	76	53	61	60		
37	51	51	63	81	60	77	71	57	82	66
	54	47	61	76	50	57	58	52		
57	40	53	66	71	61	61	55	73	50	70

	59	50	59	83	69	66	67	47		
56	60	43	54	47	81	76	69			

Щоб побудувати шкалу, треба реалізувати такий алгоритм.

- Знайти граничні значення, які набуває змінна:
 а) $X_{\max} = 83$;
 б) $X_{\min} = 37$.
- Вибрати оптимальну кількість інтервалів ($k = 10$) (табл. 5.10).

Таблица 5.10. Шкала інтервалів

№з/п	Межі інтервалів	Кількість осіб, які потрапили в цей інтервал
01	86—40	2
02	41—45	1
08	46—50	9
04	51—55	9
05	56—60	13
06	61 — 65	10
07	66—70	7
08	71—75	3

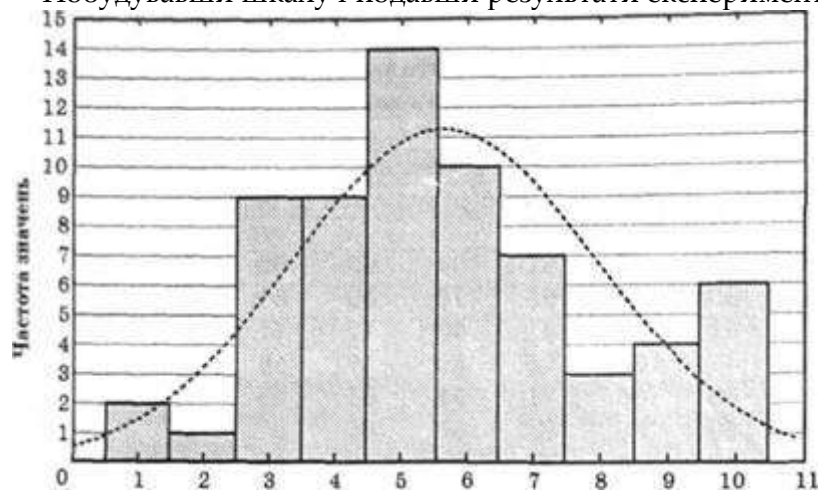
09	76—80	4
10	81—85	6

3. Визначити розмір інтервалу (крок шкали):

$$R = \frac{\Delta}{k} = \frac{(83-37)}{10} = 4,6 \approx 5.$$

- 4. Побудувати інтервали шкали.
- 5. Побудувати гістограму (рис. 5.2).

Побудувавши шкалу і подавши результати експерименту на гістограмі, ми можемо зробити якісний аналіз даних тесту. В нашому прикладі розподіл



Класи

Рис. 5.2. Крива побудови розподілу емпіричних даних

даних значно відрізняється від нормального розподілу. Це зумовлено двома причинами:

- 1) поганий тест;

о 2) опитування проведено на нерепрезентативній вибірці (є дві підгрупи обстежуваних: із середніми та високими показниками інтелекту).

Підбір суб'єктів дослідження.

Для визначення об'єму вибірки за допомогою математичної формули треба хоча б приблизно знати величину середнього квадратичного відхилення (σ) ознаки, яку вивчають.

Помилка вибіркового дослідження зменшується зі збільшенням вибірки. Ця залежність є в основі вирішення "зворотного" завдання — скільки треба взяти людей або провести дослідження, щоб можна було гарантувати достовірний результат.

Сформульоване завдання розв'язують за допомогою формули:

$$n = \frac{t^2 \cdot \sigma^2}{m^2},$$

де t — достовірний коефіцієнт; σ — середнє квадратичне відхилення; m — задана міра точності.

Послідовність операцій при використанні згаданої формули ми покажемо на типовому прикладі.

Припустимо, що вимагають визначити кількість досліджень, потрібних для отримання достовірних результатів показників успішності навчання молодого працівника двома різними методами пошуку пошкоджень у радіоапаратурі — цілісним методом і частинами.

Прийнято, що в психологічних дослідженнях мінімальною припустимою достовірною ймовірністю є 95 %, тобто тільки в п'яти випадках зі ста можуть виникнути показники, які не підтверджують прийняту гіпотезу. Такій ймовірності відповідає коефіцієнт достовірності $t = 1,96 \approx 2$.

Припустимо, що за умовами завдання дослідження можна скористатися величиною стандартного квадратичного відхилення, яку отримали в аналогічних попередніх експериментах, і воно становитиме 1.1. Вважатимемо, що для розв'язання поставленого завдання в експерименті буде потрібна міра точності в 0,2 бала. Іншими словами, коливання середньої величини оцінки успішності не повинні перевищувати 0,2 бала.

Знайдені значення підставляють у формулу:

$$n = \frac{2^2 \cdot 1,1^2}{0,2^2} = \frac{4 \cdot 1,21}{0,04} = 121.$$

Отже, надійність результатів дослідження можна досягнути тільки при об'ємі матеріалу, який дорівнює мінімум 121 показнику. В цьому випадку його можна отримати лише за рахунок кількості досліджуваних (121 особа).

Теорія статистичного висновку (статистична перевірка наукової гіпотези).

Без підготовки в галузі теорії і методів перевірки статистичних гіпотез неможливо читати наукові звіти й розуміти їх. Теорія статистичного висновку — це формалізована система методів виконання завдань через виведення властивостей великого масиву (генеральної сукупності) даних шляхом обстеження вибірки. Завдання полягає в тому, щоб передбачити властивості всієї сукупності, знаючи лише властивості вибірки із цієї сукупності.

Які висновки можна зробити про властивості генеральної сукупності за вибірковим спостереженням?

Є два види гіпотез: наукові та статистичні. Наукова гіпотеза — це запропоноване розв'язання проблеми. Це розумно обґрунтоване й розвинуте допущення. Його перевіряють результатами експерименту, які відповідають на запитання: буде гіпотеза істиною чи ні. Статистична гіпотеза — просто судження стосовно невідомого параметра. Статистичною називають гіпотезу про вид невідомого розподілу або про параметри відомих розподілів. Статистичною називають гіпотезу про те, що змінна в генеральній сукупності розподілена за нормальним законом. Гіпотезу, що перевіряють, називають нульовою і позначають H_0 . Поряд з нульовою розглядають конкуруючу (альтернативну) гіпотезу H_1 яка її заперечує.

Статистичні критерії. Для перевірки H_0 використовують спеціально підібрану випадкову величину, точний розподіл якої відомий і здебільшого зведено в таблиці. Ця величина називається критичною $K_{кр}$, якщо вона віддаляє критичну ділянку від ділянки прийняття гіпотези.

Практично прийняття або відкидання H_0 гіпотези відбувається так: вибирають відповідний критерій (t , F , χ^2 , λ), обчислюють значення, яке простежується у вибірці критерію λ , або t , враховуючи емпіричний розподіл; вибирають рівень статистичної значимості (0,05 або 0,01). За таблицею розподілу K для заданого рівня значимості знаходимо критичну точку $K_{кр(таб)}$. Якщо $K_{емп} > K_{кр(таб)}$, то нульову (H_0) гіпотезу відкидають, якщо ж $K_{емп} < K_{кр(таб)}$, то її відкидати немає підстав. Для більшої впевненості потрібно її перевірити іншим способом (наприклад, збільшити вибірку).

Як конкретний приклад розглянемо застосування критерію λ Колмогорова про вид розподілу ознаки, яку вивчаємо (IQ).

Оцінка відмінності розподілів емпіричного від теоретичного за критерієм λ подано в табл. 5.11:

Таблиця 5.11. Емпіричний і теоретичний розподіл даних

W	P	P'	Накопичені частоти		d
			$\sum P$	$\sum P'$	
142	4	0,4	90	79	11
130	14	13	87	78	9
118	25	13	73	65	8
106	22	25	48	52	4
94	18	13	26	27	1
82	4	10	8	14	6
70	3	4	4	4	0
Σ	90	78,4	—	—	—

Критерій λ придатний для будь-яких розподілів.

Відмінності можна вважати випадковими, якщо емпіричний критерій не досягає потрібного порога ймовірності.

$n = 90$ — обсяг кожної з груп (емпіричної та теоретичної).

$$\lambda = \frac{|d|}{\sqrt{n}} = \frac{11}{\sqrt{90}} = \frac{11}{9,48} = 1,16,$$

$$1,16 < 1,36 \text{ (таб.)},$$

$$K_{ем} < K_{таб}; 0,3 < 1,36.$$

Відмінності недостовірні, оскільки немає достатніх підстав вважати, що вибірки взяті з двох генеральних сукупностей, які відрізняються своїми розподілами:

$$K_{\text{ем}} > K_{\text{таб}}, 2,45 > 1,95.$$

Відмінності не можуть вважатися випадковими, вони достовірні. Вибірки взяті з двох генеральних сукупностей, які явно розрізняють за своїми розподілами.

Статистична оцінка достовірності різниці між середнім показником двох груп експериментальних даних.

Завдання цього типу дуже часто трапляються в практиці психологів різних профілів. Питання про те, чи відрізняються між собою дві групи спостережень виникає буквально всюди.

Розглянемо t -критерій Стюдента. Цей критерій використовують для оцінки статистичної значимості різних вибірових середніх двох розподілів первинних величин. Статистична значимість різниці середніх арифметичних величин вираховується за формулою:

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}},$$

де t — критерій Стюдента; M_1 і M_2 — середні арифметичні величини вибірки N_1 і N_2 ; m_1 і m_2 — квадрати середніх помилок, які обчислюють за формулою:

$$m = \frac{\sigma}{\sqrt{N}},$$

де σ — стандартне відхилення; N — обсяг вибірки.

Різниця середніх арифметичних величин вважається статистично значимою, якщо $t \geq t_{st}$, де t — критерій Стюдента, t_{st} — стандартне значення.

Стандартне значення (t_{st}) визначають з врахуванням обсягу вибірки (N) або ступеня свободи ($N = N_1 + N_2 - 2$). Критичне значення t_{st} для трьох порогів ймовірності наведено в табл. 5.12:

Таблиця 5.12. Стандартні значення t -критерію Стюдента

N	$t_1 = 0,05$	$t_2 = 0,01$	$t_3 = 0,001$	N	$t_1 = 0,05$	$t_2 = 0,01$	$t_3 = 0,001$
1	12,7	63,7	636,7	13	2,2	3,0	4,2
2	4,3	9,9	31,6	14 — 15	2,1	3,0	4,1
3	3,2	5,8	12,9	16 — 17	2,1	2,9	4,0
4	2,8	4,6	8,6	18 — 20	2,1	2,9	3,9
5	2,6	4,0	6,9	21 — 24	2,1	2,8	3,8
6	2,4	3,7	6,0	25 — 28	2,1	2,8	3,7
7	2,4	3,5	5,4	29 — 30	2,0	2,8	3,7
8	2,3	3,4	5,0	31 — 34	2,0	2,7	3,7
9	2,3	3,3	4,8	35 — 42	2,0	2,7	3,6
10	2,2	3,2	4,6	43 — 62	2,0	2,7	3,5
11	2,2	3,1	4,4	63 — 175	2,0	2,6	3,4
12	2,2	3,1	4,3	176 і більше	—	2,6	3,3

Приклад. У табл. 5.18 наведено середньогрупові величини шістнадцяти змінних методики Кеттела. Ці дані отримано усередненням трьох статистично однорідних вибірок, які відрізняються лише мірою вираженості психічної напруги. У кожній вибірці було представлено по 28 осіб від 26 до 80 років. Завдання аналізу — виявити індивідуально-особистісні властивості регулювальників кольорових телевізорів, які суттєво впливають на виникнення психічної напруги. Відправним пунктом аналізу були середньогрупові величини, отримані методом профілювання.

Таблиця 5.18. Статистичні параметри показників для груп, які розрізняють за рівнем психічної напруженості

Показники	$M \pm m$ "ВН"- групи	σ "ВН"- групи	$M \pm m$ "СН"- групи	σ "СН"- групи	$M \pm m$ "НН"- групи	σ "НН"- групи
A	$6,3 \pm 0,36$	1,9	$6,8 \pm 0,24$	1,2	$6,8 \pm 0,28$	1,4
B	$5,1 \pm 0,37$	1,9	$5,5 \pm 0,32$	1,6	$6,1 \pm 0,41$	2,1
C	$5,0 \pm 0,36$	1,5	$4,8 \pm 0,28$	1,4	$5,6 \pm 0,30$	1,56
E	$4,9 \pm 0,30$	1,5	$5,0 \pm 0,28$	1,4	$5,2 \pm 0,28$	1,50
F	$4,3 \pm 0,24$	1,2	$4,5 \pm 0,19$	0,9	$4,3 \pm 0,17$	0,9
G	$5,8 \pm 0,33$	1,7	$5,5 \pm 0,26$	1,39	$5,5 \pm 0,26$	1,4
H	$4,3 \pm 0,35$	1,8	$4,1 \pm 0,34$	1,7	$4,3 \pm 0,34$	1,7
I	$4,9 \pm 0,26$	1,3	$4,3 \pm 0,26$	1,3	$5,2 \pm 0,27$	1,3
L	$3,5 \pm 0,26$	1,3	$3,3 \pm 0,24$	1,2	$3,7 \pm 0,32$	1,6
M	$4,4 \pm 0,20$	1,0	$4,7 \pm 0,17$	0,8	$5,2 \pm 0,18$	0,9
N	$5,3 \pm 0,27$	1,4	$5,7 \pm 0,28$	1,4	$5,5 \pm 0,28$	1,5
O	$4,7 \pm 0,30$	1,05	$5,7 \pm 0,21$	1,1	$7,2 \pm 0,23$	1,2
Q1	$3,4 \pm 0,19$	1,0	$3,7 \pm 0,21$	1,1	$3,6 \pm 0,22$	1,1
Q2	$4,7 \pm 0,37$	1,9	$3,7 \pm 0,23$	1,1	$3,7 \pm 0,23$	1,2
Q3	$4,8 \pm 0,22$	1,1	$4,0 \pm 0,23$	1,2	$4,5 \pm 0,19$	1,0
Q4	$5,2 \pm 0,27$	1,4	$5,6 \pm 0,17$	0,9	$6,5 \pm 0,25$	1,3

З табл. 5.14 видно, що значимі відмінності при порівнянні високо- і низьконапружених ("ВН" і "НН") регулювальників виявлено у п'яти особистісних властивостях (групу середньо-напружених надалі позначатимемо як "СЯ"). Різниця середніх статистично значима на рівні $p \leq 0,001, 0,01, 0,05$.

Аналіз експериментального матеріалу показав, що найбільш суттєві відмінності виявлено перш за все за показником самоконтролю (Q2) й емоційної нестійкості (C). Гірші показники — за рівнем інтелекту (B), суб'єктивності, замкненості (M), тривоги, схильності до самозвинувачення (O), фрустрованості, напруженості інстинктивної сфери (Q4).

Особистісний профіль за методикою 16PF Кеттела подано на рис. 5.3.

Таблица 6.14. Значення величин t-критерію Стьюдента* при порівнянні показників груп, які розрізняють за рівнем психічної напруги

* Для $p \leq 0,05$ $t_{\alpha} = 2,01$; для $p \leq 0,01$ $t_{\alpha} = 2,63$; для $p \leq 0,001$ $t_{\alpha} = 3,0$.

Групи, що порівнюються	Показники															
	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4
BH:HH	1,07	1,81	0,67	0,67	0,24	0,66	0,14	0,95	0,34	2,84	0,63	6,57	0,59	2,35	1,16	3,49
CH:BN	0,47	1,1	2,07	0,49	0,82	0,09	0,59	2,55	0,79	1,98	0,60	4,83	0,39	0,10	1,40	2,99
HH:CH	1,14	0,76	0,43	0,18	0,68	0,57	0,43	1,64	0,49	1,07	1,24	2,66	0,98	2,29	2,36	1,13

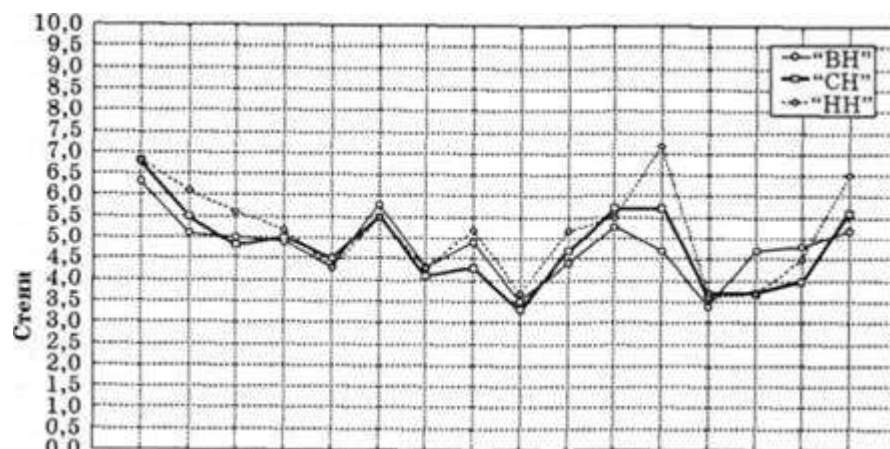
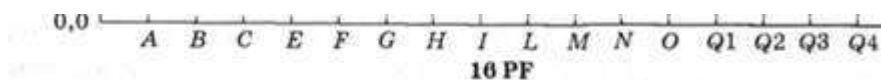


Рис. 5.3. Особистісний профіль регулювальників різного рівня психічної напруженості: A — замкнений — контактний; B — низький — високий рівень інтелекту; C — емоційно лабільний — емоційно стабільний; E — підлеглий — незалежний; F — стриманий — імпульсивний; G — безпосередній — совісний; H — несміливий — авантюрист; I — брутальний — делікатний; L — довірливий — підозріливий; M — з конкретним мисленням — з творчою уявою; N — безхитрісний — хитрий; O — впевнений в собі — тривожний; Q1 — консервативний — схильний до експерименту; Q2 — залежний від групи — індивідуаліст; Q3 — недисциплінований з внутрішньою конфліктністю — врівноважений; Q4 — розслаблений — напружений



Однак психодіагностику треба пам'ятати, що існування статистично значимої різниці середніх значень є важливим, але не єдиним аргументом на

користь наявності або відсутності зв'язку (залежності) між явищами або змінними. Тому потрібно використовувати й інші аргументи кількісного та змістовного обґрунтування можливого зв'язку.

Статистично значимий достовірний зв'язок — необхідна й недостатня умова існування змістовної взаємозалежності змінних.

Статистичне дослідження міри взаємозв'язку між двома явищами. Міра зв'язку між експериментальними даними (r_{xy}).

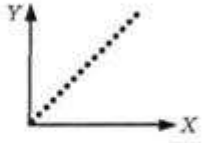
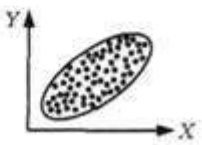
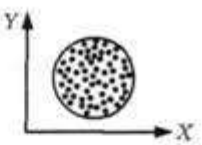
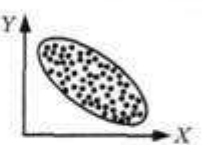
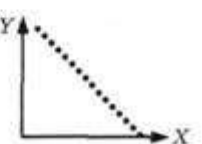
У будь-якому трудовому процесі фактори, які його становлять, перебувають у тісному взаємозв'язку. Вміння змінити один фактор так, щоб отримати відповідну зміну іншого, зробить трудовий процес оптимальнішим.

Будь-яка галузь сучасної науки прагне здебільшого виразити відкриті нею закони у формі математичної моделі, тобто у виді тих чи інших співвідношень між показниками, які характеризують різні сторони пізнаних явищ. Таку модель відповідних співвідношень між показниками називають функціональною залежністю. Функціональний зв'язок відображає чітку однозначну залежність, при якій зміна одного певного фактора обов'язково призводить до однозначної зміни іншого. Знання функціональної залежності дає змогу передбачити знання залежної величини (функції) при будь-якому значенні "керуючої" нею величини (аргументу).

Навіть у точних науках, які володіють більшими можливостями, стандартизувати всі фактори, крім тих, які вивчають, не завжди вдається зблизити умови проведення суміжних дослідів. Ще важче досягти того в психологічних (педагогічних) дослідженнях. Причин дуже багато: по-перше, багато факторів, які впливають на хід навчання і виховання, невідомо (наприклад, певні події в житті досліджуваного, неможливо детально врахувати його життєвий досвід, неможливо повністю ізолювати людину в експерименті від впливу середовища, яке постійно змінюється та ін.); по-друге, практично неможливо підібрати зовсім однакових людей для порівняльного експерименту; по-третє, неможливо знайти двох однакових педагогів для проведення занять в експериментальних і контрольних групах; по-четверте, суб'єктивні переживання досліджуваних, їх ставлення до занять, експерименту недоступні для безпосереднього вивчення. Людина як об'єкт дослідження дуже складна у своїх виявленнях, щоб її поведінку можна було вкласти в якусь формулу. Саме тому як в біології, так і в психології й педагогіці, здебільшого уникають говорити про функціональні зв'язки й основну увагу приділяють дослідженню статистичних зв'язків або кореляції.

Кореляція дає змогу знаходити статистично достовірні кількісні міри зв'язку (табл. 5.15) в тих випадках, коли певному фактору відповідає не одне, а кілька значень якого-небудь іншого фактора, причому варіюючих у певних межах. Зв'язок у цьому випадку виражатиметься середніми значеннями, які отримали багатьма змінами (замірів).

Таблиця 5.15. Інтерпретація значень r_{xy} . Ілюстрація типів лінійних зв'язків, які є між X і Y для різних значень r_{xy}

Величина r_{xy}	Описання лінійного зв'язку	Діаграма розсіювання
+1,00	Суворий прямий зв'язок, при якому збільшення причинного фактора спричиняє збільшення наслідкового фактора. Є функціональний зв'язок	
Близько +0,50	Слабкий прямий зв'язок, коефіцієнт кореляції, до 0,30, середній зв'язок — від 0,31 до 0,69 і сильний зв'язок — від 0,70 до 0,99	
Близько 0,00	Немає зв'язку, тобто варіації X і $Y = 0$	
Близько -0,50	Слабкий зворотний зв'язок	
-1,00	Суворий зворотний зв'язок, при якому збільшення причинного фактора спричиняє зменшення наслідкового. Є функціональний зв'язок	

Корелюючи фактори поділяють на причинні, які видозмінюються і першими, спричиняють зміну інших факторів; наслідкові, які видозмінюються під впливом причинних факторів. Наслідкові фактори можуть набувати багато значень у відповідних межах.

Дослідників часто цікавить, як пов'язані між собою дві змінні в певній групі осіб (класі, школі, нації тощо). Наприклад, чи мають учні, які навчилися читати раніше за інших, тенденцію до кращої успішності в шостому класі? Чи простежуються у великих класах менші успіхи в набуванні знань за семестр, ніж у менших класах? Чи пов'язана тривалість роботи педагога в школі безпосередньо зі середньою зарплатою? Щоб відповісти на такі

запитання, ми повинні простежити за кожною змінною для груп об'єктів (типових представників, якими можуть бути класи, школи, райони тощо).

Дані, зібрані для відповіді на одне з подібних запитань, можуть виглядати так, як наведено в табл. 5.16. У цьому прикладі змінними, які вивчали у 12 учнів, були оцінки IQ, що визначали за допомогою тесту інтелекту Векслера в шостому класі, і успішність у середній школі з хімії, яку оцінювали на підставі тесту, складеного з 35 запитань.

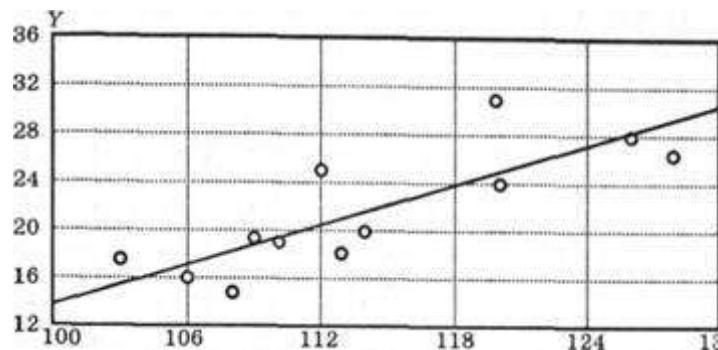
Таблиця 5.16. Дані оцінок рівня інтелекту й оцінок тесту успішності з хімії

Порядковий номер	IQ(x)	Оцінка тесту успішності з хімії
1	120	31
2	112	26
3	110	19
4	120	24
5	103	17
6	126	28
7	113	18
8	114	20
9	106	16
10	108	15
11	128	27

Зв'язок між двома змінними можна зобразити графічною діаграмою розсіювання. Діаграма розсіювання для цього прикладу зображена на рис. 5.4. На діаграмі розсіювання показники кожного учня позначені крапкою. Мітка розміщена у місці перетину прямих ліній, які проведені через відмітку показника IQ перпендикулярно до осі X і через відмітку показника тесту з хімії перпендикулярно до осі Y для кожного учня. З діаграми на рис. 5.4 видно, що є слабкий позитивний зв'язок між цими показниками. Узагальненою мірою цього зв'язку є коефіцієнт коореляції Пірсона.

Формула для розрахунку коефіцієнта Пірсона (r_{xy}):

$$r_{x,y} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \cdot \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}}$$



X

Рис. 5.4. Діаграма розсіювання, яка відображає зв'язок IQ(x) з успішністю з хімії (Y) для 12 школярів

Приклад обчислення $r_{x,y}$. (табл. 5.17) Щоб проілюструвати обчислення $r_{x,y}$ за формулою, треба використати деякі дані тесту на виявлення загальних і спеціальних здібностей. Дослідник вивчає зв'язок двох типів розумових здібностей учнів неповної середньої школи: абстрактне і вербальне мислення. Розроблено два тести: для замірів нахилів до абстрактного мислення (X) і до вербального мислення (Y). Обидва тести запропонували 40 школярам молодшого класу неповної середньої школи в одному місті з 30 тис. жителів. Результати досліджень 40 учнів наведено в табл. 5.18. Кожен тест мав 50 запитань, а результат — кількість правильних відповідей.

Таблиця 5.17. Приклад обчислення $r_{x,y}$ за даними таблиці

Проміжні результати	Кінцеві обчислення
$n = 40$	$r_{x,y} = \frac{40 \cdot 40798 - (-1465) \cdot 1057}{\sqrt{[40 \cdot 55725 - (1465)^2] \cdot [40 \cdot 32551 - (1057)^2]}} =$ $= \frac{83415}{123761,128} = 0,67$
$\sum_{i=1}^{40} x_i = 1465$	
$\sum_{i=1}^{40} y_i = 1057$	
$\sum_{i=1}^{40} x_i^2 = 55725$	
$\sum_{i=1}^{40} y_i^2 = 32551$	
$\sum_{i=1}^{40} x_i y_i = 40798$	

Таблиця 5.18. Вихідні оцінки здатності до абстрактного та вербального мислення 40 школярів неповної середньої школи

Порядковий номер	Х, абстрактне мислення	У, вербальне мислення	Порядковий номер	Х, абстрактне мислення	У, вербальне мислення	Порядковий номер	Х, абстрактне мислення	У, вербальне мислення
1	19	17	16	82	24	81	31	16
2	32	7	17	48	45	32	41	87
3	88	17	18	48	26	33	42	37
4	44	28	19	38	16	34	24	14
5	28	27	20	47	26	35	43	41
6	85	81	21	38	30	36	36	19

7	39	20	22	25	18	37	39	18
8	39	17	23	35	26	38	39	39
9	44	35	24	22	17	39	39	37
10	44	48	25	40	17	40	48	47
11	24	10	26	42	26			
12	87	28	27	41	16			
18	29	13	28	41	87			
14	40	48	29	87	26			
15	42	45	30	30	21			