МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ВОСТОЧНОУКРАИНСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
имени ВЛАДИМИРА ДАЛЯ

Лилия Григорьевна Зубова

ОСНОВЫ МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Учебное пособие

Луганск
Издательство «Ноулидж»
2013

УДК 519.24
ББК 22.172
З 91

Рекомендовано Ученым советом
Восточноукраинского национального университета имени Владимира Даля
(протокол № 6 от 22 февраля 2013 г.)

Рецензенты:
Голубничий П. И., доктор физико-математических наук, профессор,
Заслуженный деятель науки и техники Украины
Филоненко А.Д, доктор физико-математических наук, профессор

Зубова Л. Г.
З 91 **Основы математической обработки экспериментальных данных**: учебное пособие / Л. Г. Зубова. – Луганск: Изд-во «Ноулидж», 2013. – 60с., табл. 33; рис. 4; библиогр. назв. 5.

ISBN 978-617-579-759-4

В учебном пособии дано общее представление о расчете простых статистических показателей, которые фактически являются «азбукой» в дальнейшем определении репрезентативности информации и математическом анализе. Показано как с помощью метода Монте-Карло можно дополнить недостающий для правильной математической обработки объем данных. Раскрыт механизм определения однородности данных одной выборочной совокупности. Подробно изложена последовательность проверки полученной информации на достоверность. Показано, какими статистическими методами осуществляется преобразовывание данных в случае их несоответствия нормальному закону распределения. Дано представление об одном из видов математического анализа - корреляционно-регрессионном, позволяющем определять зависимость одной выборочной совокупности от другой. Описаны методы сравнения между собой различного количества выборочных совокупностей.

Для студентов, аспирантов и научных сотрудников.

У навчальному посібнику дано загальне уявлення про розрахунок простих статистичних показників, які фактично є «азбукою» в подальшому визначенні репрезентативності інформації і математичному аналізі. Показано як за допомогою методу Монте-Карло можна доповнити бракуючий для правильної математичної обробки об'єм даних. Розкритий механізм визначення однорідності даних однієї вибіркової сукупності. Детально викладена послідовність перевірки отриманої інформації на достовірність. Показано якими статистичними методами здійснюється перетворювання даних у разі їх невідповідності нормальному закону розподілу. Дано уявлення про один з видів математичного аналізу - кореляційно-регресійний, дозволяючий визначати залежність однієї вибіркової сукупності від іншої. Описані методи порівняння між собою різної кількості вибіркових сукупностей.

Для студентів, аспірантів і наукових співробітників.

УДК 519.24
ББК 22.172

ISBN 978-617-579-759-4

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ

В течение значительного количества лет автору этого пособия довелось принять участие в заседаниях различных специализированных советов по защитам кандидатских и докторских диссертаций, познакомиться с большим количеством авторефератов диссертаций по самым различным специальностям. Анализ этих защит и авторефератов показал, что только в редких диссертациях присутствовала достаточно полная математическая обработка экспериментальных данных с проверкой цифровой информации на ***однородность*** и ***достоверность***. Как правило, в диссертационных работах присутствовал расчет статистических показателей с доверительным интервалом, приводились графики зависимостей с указанием коэффициента детерминации, построенные с использованием стандартного аппарата программы Excel. В некоторых работах присутствовала более полная математическая обработка, выполненная с помощью различных компьютерных программ, например "Statistika", но без осмысления диссертантом сущности выполненных им расчетов. Выявленные недостатки математической обработки данных стали одним из побудительных мотивов для написания этого учебного пособия. Анализ имеющихся литературных источников, посвященных математической обработке данных, показал, что они зачастую, несмотря на свой неоспоримо высокий научный уровень, попросту «отпугивают» не вполне подготовленную аудиторию большим объемом материала и сложностью его восприятия. Работая с аспирантами и студентами, взыскательно относясь к уровню выполняемой ими математической обработки данных, автор старался выработать наиболее доступные подходы к обучению, которые и нашли отражение в предлагаемом учебном пособии, написанном им по просьбе учеников и коллег.

В пособии изложена последовательность обработки экспериментальных данных с расчетом статистических показателей, с проверкой данных на однородность и достоверность. Показано, какими статистическими методами осуществляется преобразование данных в случае их несоответствия нормальному закону распределения. Дано общее представление об одном из видов математического анализа - корреляционно-регрессионном, позволяющем определять зависимость одной выборочной совокупности от другой. Описаны методы сравнения между собой различного количества выборочных совокупностей.

В основе данного учебного пособия лежат общеизвестные математические методы, изложенные в работах Гмурмана В.Е., Доспехова Б. А., Куценко М.В., Самнера Г., Школьного Е. П., Гончаровой Л. Д., Миротворской Н. К. и др., которым автор глубоко признателен за их неоценимую заочную помощь в постижении тонкостей математической обработки экспериментальных данных. Все эти методы подробно рассмотрены на предложенных конкретных примерах и расчетах.

Надеемся, что предложенное учебное пособие поможет сделать первые шаги к правильной обработке результатов проведенных экспериментов и наблюдений, к умению проверять их однородность и достоверность (соответствие нормальному закону распределения). Следующий шаг будет включать знакомство с другими законами распределения, с детерминированным математическим анализом, математическим моделированием и др., которые автор надеется осветить в дальнейших работах.

РАЗДЕЛ 1. ОБЩИЕ СВЕДЕНИЯ

Любой аспирант или научный сотрудник - «начинающий» или «с опытом» получает в результате своих исследований значительный объем данных. Это могут быть результаты наблюдений (исследователь наблюдает «без вмешательства» за каким-либо явлением или процессом) или экспериментов (при активном вмешательстве исследователя в изучаемый процесс, осуществляя постановку опытов в определенном направлении и с определенными условиями).

Все полученные данные (цифровой материал, качественные показатели) обязательно должны иметь определенный объем, должны быть правильно обработаны и проанализированы. Чтобы такой анализ не казался трудным, вначале нужно усвоить несколько несложных, но очень необходимых математических понятий и научиться рассчитывать для своих полученных значений простые статистические показатели. Это своего рода «азбука», на которой строятся все остальные расчеты.

Формулы для расчета простых статистических показателей [1]

1) Средняя арифметическая X_{cp}:

$$X_{cp} = \overline{X} = \frac{\sum X_i}{n},$$

где $\sum X_i$ - сумма всех значений изучаемой величины,

n – количество значений (количество вариант);

2) дисперсия S^2:

$$S^2 = \frac{\sum (X_i - \overline{X})^2}{n-1},$$

где $\sum (X_i - \overline{X})^2$ - сумма квадратов разностей всех значений и средней арифметической;

3) стандартное (среднее квадратическое) отклонение S:

$$S = \sqrt{S^2};$$

4) коэффициент вариации V:

$$V = \frac{S}{\overline{X}} \cdot 100\%;$$

5) абсолютная ошибка средней арифметической:

$$S_{\overline{X}} = \frac{S}{\sqrt{n}};$$

6) относительная ошибка средней арифметической:

$$S_{\overline{X}}\% = \frac{S_{\overline{X}}}{\overline{X}} \cdot 100\%;$$

7) доверительный интервал для средней арифметической:

$$\overline{X} \pm t \cdot S_{\overline{X}} ,$$

где t – критерий Стьюдента (табл. 1.1) при (n - 1) степенях свободы и 0,05 (5%-ном) уровне значимости.

Таблица 1.1

Значения *t*-критерия Стьюдента на уровнях значимости 0,10; 0,05; 0,01 [1]

Число степеней свободы (n - 1)	Уровень значимости			Число степеней свободы (n - 1)	Уровень значимости		
	0,10	0,05	0,01		0,10	0,05	0,01
1	6,3138	12,706	63,657	18	1,7341	2,1009	2,8784
2	2,9200	4,3027	9,9248	19	1,7291	2,0930	2,8609
3	2,3534	3,1825	5,8409	20	1,7247	2,0860	2,8453
4	2,1318	2,7764	4,6041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,4995	24	1,7109	2,0639	2,7969
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7530	2,1315	2,9467	60	1,6707	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6577	1,9799	2,6174
17	1,7396	2,1098	2,8982	∞	1,6449	1,9600	2,5758

Для расчета простых статистических показателей используется вспомогательная таблица (табл. 1.2).

Таблица 1.2

Образец вспомогательной таблицы для расчета простых статистических показателей

№	X_i (значения)	$X_i - X_{ср.}$	$(X_i - X_{ср.})^2$
1	5	-1	1
2	7	1	1
…	…	…	
n	…	…	
	$\sum X_i$	$\sum (X_i - X_{ср.}) \approx 0$	$\sum (X_i - X_{ср.})^2$

Пример расчета простых статистических показателей

В результате экспериментов аспирантом N был получен ряд значений изучаемого показателя X. Для расчета простых статистических показателей им вначале была составлена вспомогательная таблица 1.3.

Таблица 1.3

Вспомогательная таблица для расчета простых статистических показателей

№ пп	X_i	$X_i - X_{ср.}$	$(X_i - X_{ср.})^2$	№ пп	X_i	$X_i - X_{ср.}$	$(X_i - X_{ср.})^2$
1	6,3	-0,22	0,05	44	5,8	-0,72	0,52
2	6,2	-0,32	0,10	45	7,6	1,08	1,17
3	7,5	0,98	0,97	46	6,6	0,08	0,01
4	3,6	-2,92	8,51	47	4,3	-2,22	4,92
5	6,3	-0,22	0,05	48	8,3	1,78	3,18
6	6,8	0,28	0,08	49	7,3	0,78	0,61
7	7,2	0,68	0,47	50	6,6	0,08	0,01
8	8	1,48	2,20	51	7,1	0,58	0,34
9	11	4,48	20,09	52	4,5	-2,02	4,07
10	7,8	1,28	1,64	53	9,3	2,78	7,74
11	8,1	1,58	2,50	54	5,5	-1,02	1,04
12	5,1	-1,42	2,01	55	4,7	-1,82	3,30
13	6,1	-0,42	0,17	56	5,5	-1,02	1,04
14	6,8	0,28	0,08	57	6,3	-0,22	0,05
15	6,2	-0,32	0,10	58	6	-0,52	0,27
16	7,1	0,58	0,34	59	2,6	-3,92	15,35
17	4,8	-1,72	2,95	60	5,1	-1,42	2,01
18	9,4	2,88	8,31	61	5,1	-1,42	2,01
19	6,2	-0,32	0,10	62	4	-2,52	6,34
20	6,8	0,28	0,08	63	8,9	2,38	5,68
21	7,1	0,58	0,34	64	5	-1,52	2,30
22	6,2	-0,32	0,10	65	4,1	-2,42	5,85
23	7,2	0,68	0,47	66	5,2	-1,32	1,74
24	6,9	0,38	0,15	67	7,8	1,28	1,64
25	6,3	-0,22	0,05	68	7	0,48	0,23
26	5,2	-1,32	1,74	69	3,5	-3,02	9,11
27	4	-2,52	6,34	70	3,2	-3,32	11,01
28	7,7	1,18	1,40	71	8	1,48	2,20
29	10,2	3,68	13,56	72	6,3	-0,22	0,05
30	4,5	-2,02	4,07	73	9	2,48	6,16
31	6	-0,52	0,27	74	7,7	1,18	1,40

Продолжение таблицы 1.3

32	8,6	2,08	4,34	75	5,6	-0,92	0,84
33	6,5	-0,02	0,00	76	4,3	-2,22	4,92
34	8,8	2,28	5,21	77	7	0,48	0,23
35	5,6	-0,92	0,84	78	8	1,48	2,20
36	6,3	-0,22	0,05	79	7,2	0,68	0,47
37	5,9	-0,62	0,38	80	7,4	0,88	0,78
38	5	-1,52	2,30	81	6,2	-0,32	0,10
39	8,6	2,08	4,34	82	9,2	2,68	7,20
40	5,1	-1,42	2,01	83	8,1	1,58	2,50
41	7,9	1,38	1,91	84	6	-0,52	0,27
42	7,4	0,88	0,78	85	3,7	-2,82	7,94
43	9,1	2,58	6,67				
				$n = 85$	$\Sigma X_i = 554$	$\Sigma(X_i – X_{ср}) =$ $= 0,0$	$\sum(X_i - X_{ср})^2 =$ $= 236,84$

1) Средняя арифметическая: $X_{ср} = \dfrac{\sum X_i}{n}$ = 6,52 см

2) Дисперсия $S^2 = \dfrac{\sum(X - \bar{X})^2}{n-1} = 2,82$

3) Стандартное (среднее квадратическое) отклонение:

$$S = \sqrt{2,82} = 1,68$$

4) Коэффициент вариации:

$$V = \frac{S}{\bar{X}} \cdot 100\% \; = \frac{1,68}{6,52} \cdot 100\% = 25,7\%$$

5) Абсолютная ошибка средней:

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} \; = \frac{1,68}{\sqrt{85}} = 0,18$$

6) Относительная ошибка средней:

$$S_{\bar{X}}\% = \frac{S_{\bar{X}}}{\bar{X}} \cdot 100\% \; = \frac{0,18}{6,52} \cdot 100\% = 2,8\%$$

7) Доверительный интервал

$$\bar{X} \pm t \cdot S_{\bar{X}} = 6,52 \pm 1,99 \cdot 0,18 = 6,52 \pm 0,36$$

Вспомогательные таблицы помещать в материалы диссертации (научного отчета) не рекомендуется, так как они значительно увеличат их объем.

В материалах диссертации (отчета) - в тексте или в приложениях размещаются таблицы с уже рассчитанными простыми статистическими показателями. Пример оформления результатов расчета простых статистических показателей по вариантам эксперимента, проведенного аспирантом М представлен в табл. 1.4.

Таблица 1.4

Статистические показатели результатов эксперимента

Показатели	№ варианта эксперимента и уровни изучаемого параметра								
	1			2			3		
	0	25	50	0	25	50	0	25	50
$\bar{X} = \frac{\sum X}{n}$	63,63	32,08	25,0	42,03	30,1	24,25	73,0	41,93	36,93
$S^2 = \frac{\sum (X - \bar{X})^2}{n-1}$	346,1	127,1	18,42	137,8	51,86	8,41	442,1	236,7	183,3
$S = \sqrt{S^2}$	18,60	11,27	4,29	11,74	7,20	2,90	21,03	15,38	13,54
$V = \frac{S}{\bar{X}} \cdot 100\%$	29,24	35,15	17,20	27,94	23,92	11,96	28,80	36,70	36,67
$S_{\bar{X}} = \frac{S}{\sqrt{n}}$	4,16	2,52	0,96	2,63	1,61	0,65	4,70	3,44	3,03
$S_{\bar{X}}\% = \frac{S_{\bar{X}}}{\bar{X}} \cdot 100\%$	6,54	7,86	3,85	6,25	5,35	2,67	6,44	8,21	8,20
$\bar{X} \pm t \cdot S_{\bar{X}}$	63,63 ± 8,69	32,08 ± 5,27	25,00 ± 2,01	42,03 ± 5,49	30,10 ± 3,37	24,25 ± 1,36	73,00 ± 9,83	41,93 ± 7,19	36,93 ± 6,33

Генеральная и выборочная совокупности данных

Предположим, имеется некая метеорологическая станция, существующая 175 лет. Если исследователя интересует среднегодовая температура, полученная на этой станции за любые 30 лет, он будет работать с выборочной совокупностью данных. Если же исследователя интересуют данные среднегодовой температуры воздуха за 175 лет (за весь период существования метеорологической станции) он будет обрабатывать генеральную совокупность данных.

Не всегда для получения какого-либо вывода имеется возможность обработать генеральную совокупность данных. Это может быть вызвано очень большим ее объемом. В таком случае выводы строятся на основании математической обработки выборочной совокупности. Данные выборочной совокупности обязательно должны быть однородны и достоверны. Но об этом речь пойдет далее.

Так как исследователь может работать и с выборочной, и с генеральной совокупностью данных, он должен правильно обозначать их статистические показатели. В таблице 1.5 приведены обозначения трех статистических показателей. Остальные обозначаются идентично.

Таблица 1.5

Обозначения статистических показателей

Статистические показатели	Совокупности	
	выборочная	генеральная
Средняя арифметическая – для выборочной совокупности, *математическое ожидание* – для генеральной совокупности	$X_{ср}$.	m
Дисперсия	S^2	δ^2
Среднее квадратическое (стандартное) отклонение	S	δ

Среднее квадратическое (стандартное) отклонение и дисперсия являются основными характеристиками разброса (варьирования) отдельных значений некоторой величины вокруг ее среднего значения или математического ожидания.

Объем данных

Важным моментом сбора информации является достаточность количества полученных значений изучаемой величины. Нормальный закон распределения (он характеризует достоверность данных), о котором пойдет речь далее, требует, чтобы количество значений величины в одной повторности варианта (при 3 и более повторностях) в проведенных наблюдениях или экспериментах было не менее 20 - 30. Это связано с тем, что при малом количестве значений выборочной совокупности ее стандартное отклонение будет значительно отличаться от стандартного отклонения для всей генеральной совокупности. А выводы, сделанные на основании такой выборочной совокупности и переложенные на генеральную совокупность, будут недостоверными. Хотя в математической науке и существует такое направление как статистика малых выборок, для полноты и правильности научных выводов диссертации или научного отчета необходимо, чтобы количество значений одной повторности (при 3 и более повторностях) было не менее 20-30.

В научной практике приходится сталкиваться со случаями, когда количество данных ограниченное. Например, некая метеорологическая станция существует 15 лет, и, соответственно, среднегодовое количество наблюдений за метеорологическими величинами на этой станции менее требуемых 20-30 значений. В этом случае для дополнения данных исследователь может использовать известный математический метод Монте-Карло [2], позволяющий смоделировать дополнительные данные в пределах закона распределения фактических (полученных исследователем) значений.

Использование метода Монте-Карло для дополнения данных.

Пусть имеются данные за 15 лет по среднегодовому количеству осадков X, выраженных в миллиметрах: 450; 450; 450; 420; 410; 450; 450; 450; 450; 420; 450; 450; 450; 420; 450.

Закон распределения полученных значений имеет следующий общий и конкретный виды:

$$X\colon\ X_1, X_2, \ldots, X_n,$$
$$P\colon\ P_1,\ P_2, \ldots, P_n$$

Имеющиеся варианты значений величины X	410	420	450
Встречаемость значений величины (абсолютная частота)	1	3	11
P (относительная частота)	0,07	0,20	0,73

Величина P характеризует вероятность (относительную частоту - отношение числа случаев какого-либо события к численности совокупности: в нашем случае – к 15).

Далее все частоты разбиваются на *n* частотных интервалов. Для нашего случая это 3 интервала с координатами: 1-й интервал – от 0 до 0,07; 2-й интервал – от 0,07 до 0,27; 3-й интервал – от 0,27 до 1, где 0,27 и 1 – суммы относительных частот нарастающим итогом.

На следующем этапе из таблицы равномерно распределенных случайных чисел (табл. 1.6) произвольно выбирается строка из необходимого количества случайных чисел. В нашем случае выбрана строка из пяти чисел: 34, 07, 27, 68, 50.

Таблица 1.6

Равномерно распределенные случайные числа [2]

10 09 73 25 33	76 52 01 35 86	34 67 35 48 76	80 95 90 91 17
37 54 20 48 05	64 89 47 42 96	24 80 52 40 37	20 63 61 04 02
08 42 26 89 53	19 64 50 93 03	23 20 90 25 60	15 95 33 47 64
99 01 90 25 29	09 37 67 07 15	38 31 13 11 65	88 67 67 43 97
12 80 79 99 70	80 15 73 61 47	64 03 23 66 53	98 95 11 68 77
66 06 57 47 17	**34 07 27 68 50**	36 69 73 61 70	65 81 33 98 85
31 06 01 08 05	45 57 18 24 06	35 30 34 26 14	86 79 90 74 39
85 26 97 76 02	02 05 16 56 92	68 66 57 48 18	73 05 38 52 47
63 57 33 21 35	05 32 54 70 48	90 55 35 75 48	28 45 82 87 09
73 79 64 57 53	03 52 96 47 78	35 80 83 42 82	60 93 52 03 44
98 52 01 77 67	14 90 56 86 07	22 10 94 05 58	60 97 09 34 33
11 80 50 54 31	39 80 82 77 32	50 72 56 82 48	29 40 52 42 01
83 45 29 96 34	06 28 89 80 83	13 74 67 00 78	18 47 54 06 10

Продолжение таблицы 1.6

88 68 54 02 00	86 50 75 84 01	36 76 66 79 51	90 36 47 64 93
99 59 46 73 48	87 51 76 49 69	91 82 60 89 28	93 78 56 13 68
65 48 11 76 74	17 46 85 09 50	58 04 77 69 74	73 03 95 71 86
80 12 43 56 35	17 72 70 80 15	45 31 82 23 74	21 11 57 82 53
74 35 09 98 17	77 40 27 72 14	43 23 60 02 10	45 52 16 42 37
69 91 62 68 03	66 25 22 91 48	36 93 68 72 03	76 62 11 39 90
09 89 32 05 05	14 22 56 85 14	46 42 75 67 88	96 29 77 88 22
91 49 91 45 23	68 47 92 76 86	46 16 28 35 54	94 75 08 99 23
80 33 69 45 98	26 94 03 68 58	70 29 73 41 35	53 14 03 33 40
44 10 48 19 49	85 15 74 79 54	32 97 92 65 75	57 60 04 08 81
12 55 07 37 42	11 10 00 20 40	12 86 07 46 97	96 64 48 94 39
63 60 64 93 29	16 50 53 44 84	40 21 95 25 63	43 65 17 70 82
61 19 69 04 46	26 45 74 77 74	51 92 43 37 29	65 39 45 95 93
15 47 44 52 66	95 27 07 99 53	59 36 78 38 48	82 39 61 01 18
94 55 72 85 73	67 89 75 43 87	54 62 24 44 31	91 19 04 25 92
42 48 11 62 13	97 34 40 87 21	16 86 84 87 67	03 07 11 20 59
23 52 37 83 17	73 20 88 98 37	68 93 59 14 16	26 25 22 96 63
04 49 35 24 94	75 24 63 38 24	45 86 25 10 25	61 96 27 93 35
00 54 99 76 54	64 05 18 81 59	96 11 96 38 96	54 69 28 23 91
35 96 31 53 07	26 89 80 93 54	33 35 13 54 62	77 97 45 00 24
59 80 80 83 91	45 42 72 68 42	83 60 94 97 00	13 02 12 48 92
46 05 88 52 36	01 39 09 22 86	77 28 14 40 77	93 91 08 36 47
32 17 90 05 97	87 37 92 52 41	05 56 70 70 07	86 74 31 71 57
69 23 46 14 06	20 11 74 52 04	15 95 66 00 00	18 74 39 24 23
19 56 54 14 30	01 75 87 53 79	40 41 92 15 85	66 67 43 68 06
45 15 51 49 38	19 47 60 72 46	43 66 79 45 43	59 04 79 00 33
94 86 43 19 94	36 16 81 08 51	34 88 88 15 53	01 54 03 54 56
98 08 62 48 26	45 24 02 84 04	44 99 90 88 96	39 09 47 34 07
33 18 51 62 32	41 94 15 09 49	89 43 54 85 81	88 69 54 19 94
80 95 10 04 06	96 38 27 07 74	20 15 12 33 87	25 01 62 52 98
79 75 24 91 40	71 96 12 82 96	69 86 10 25 91	74 85 22 05 39
18 63 33 25 37	98 14 50 65 71	31 01 02 46 74	05 45 56 14 27
74 02 94 39 02	77 55 73 22 70	97 79 01 71 19	52 52 75 80 21
54 17 84 58 11	80 99 33 71 43	05 33 51 29 69	56 12 71 92 55
11 66 44 98 83	52 07 98 48 27	59 38 17 15 39	09 97 33 34 40
48 32 47 79 28	31 24 66 47 10	02 29 53 68 70	32 30 75 75 46
69 07 49 41 38	87 63 79 19 76	35 58 40 44 01	10 51 82 16 15

Числа таблицы варьируют от 00 до 99 и представляют собой вероятности, выраженные в процентах. Далее, для удобства расчета, эти числа преобразовываются в десятичную дробь и приобретают следующий вид: 0,34; 0,07; 0,27; 0,68; 0,50.

Первое число - 0,34 входит в интервал №3, соответственно X в данном случае принимает значение 450; второе число 0,07 входит в интервал №1 и X для него принимает значение 410. Все остальные числа определяются аналогично. Полученные результаты оформляются в виде таблицы (табл. 1.7).

Таблица 1.7

Фактические данные, дополненные по методу Монте-Карло

<table>
<tr><th rowspan="2">Вариант</th><th colspan="2">Исходные фактические данные</th><th rowspan="2">Частотные интервалы</th><th rowspan="2">Данные, выбранные из таблицы случайных чисел</th><th rowspan="2">Преобразованный ряд значений величины X (дополненный новыми - расчетными значениями)</th></tr>
<tr><th>X_i</th><th>P_i</th></tr>
<tr><td rowspan="3">1</td><td>410</td><td>0,07</td><td>№1
(0-0,07)</td><td rowspan="3">0,34; 0,07; 0,27; 0,68; 0,50</td><td rowspan="3">450; 450; 450; 420; 410; 450; 450; 450; 450; 420; 450; 450; 450; 420; 450; 450; 410;
420; 450; 450</td></tr>
<tr><td>420</td><td>0,2</td><td>№2
(0,07-0,27)</td></tr>
<tr><td>450</td><td>0,73</td><td>№3
(0,27-1)</td></tr>
</table>

Примечания: полужирным шрифтом выделены рассчитанные значения

РАЗДЕЛ 2. ПРОВЕРКА ДАННЫХ НА ОДНОРОДНОСТЬ

Все полученные в процессе наблюдений и экспериментов данные должны обязательно проверяться на однородность, то есть на отсутствие «сомнительных» значений.

В процессе научных исследований для получения необходимых данных используется значительное количество приборов. Работа с приборами не исключает ошибки измерений. Ошибки, как известно, бывают грубые, систематические и случайные.

Систематическими называют ошибки, которые по знаку или величине однообразно повторяются в многократных измерениях (например из-за неисправности прибора, недостаточного умения специалиста работать с ним и т.д.).

Влияние систематических ошибок стремятся ослабить путем калибровки или ремонта прибора, применением соответствующей методики измерений, а также введением поправок в результаты измерений.

Случайные ошибки – это ошибки, размер и влияние которых на каждый отдельный результат измерения остается неизвестным. Величину и знак случайной ошибки заранее установить нельзя. Однако теоретические исследования и многолетний опыт измерений показывают, что случайные ошибки подчинены определенным вероятностным закономерностям, изучение которых дает возможность получить наиболее надежный результат и оценить его точность.

Грубыми называют ошибки, превосходящие по абсолютной величине некоторый, установленный для данных условий измерений, предел. В большинстве случаев они происходят в результате промахов и просчетов исполнителя, а также по другим причинам. Такие ошибки обнаруживают повторными измерениями, а результаты, содержащие их, бракуют и заменяют новыми. Грубые ошибки приводят к появлению данных, характеризуемых в математической статистике как «сомнительные».

К сомнительным данным относятся, прежде всего, те, которые имеют либо резко максимальное, либо резко минимальное значения. И проверяются на однородность, прежде всего, именно эти значения.

Для проверки на однородность данных можно использовать различные критерии. Наиболее часто употребляемыми являются критерий Стьюдента (t) [3] и критерий Тау (τ) [1]. Можно использовать любой из них. При использовании этих критериев вначале рассчитывается их фактическое значение, которое впоследствии сравнивается с табличным (теоретическим). Для нахождения этих критериев в специальных таблицах используется 5%-ный уровень значимости и степень свободы

$(n-1)$ для критерия Стьюдента, а для критерия Тау – степень свободы n, где n – объем выборки.

Сомнительные - экстремальные (максимальные и минимальные) значения подлежат выбраковке, если фактическое (расчетное) значение используемого критерия больше или равно теоретическому (табличному) значению этого критерия. Эта гипотеза носит название H_1.

Данные считаются однородными, а сомнительные (экстремальные) значения не выбраковываются в том случае, когда фактическое (расчетное) значение критерия меньше табличного (теоретического) значения данного критерия. Это гипотеза H_0.

Критерий Стьюдента (t): если $t_ф \geq t_т$, имеет место гипотеза H_1, и проверяемое сомнительное значение выбраковывается из генеральной или выборочной совокупности;

если $t_ф < t_т$, имеет место гипотеза H_0, то есть проверяемое значение однородно со всей (генеральной) выборочной совокупностью и не выбраковывается.

Критерий Тау (τ): если $\tau_ф \geq \tau_т$ имеет место гипотеза H_1 – значение выбраковывается из (генеральной) выборочной совокупности;

если $\tau_ф < \tau_т$ - имеет место гипотезаH_0 – значение однородно со всей (генеральной) выборочной совокупностью и не выбраковывается.

Этапы проверки данных на однородность с помощью критерия Стьюдента

1. Значения совокупности сортируют в порядке возрастания.
2. Рассчитывают фактическое значение критерия ($t_ф$) по формуле:

$$t_ф = (X_{экстр} - X_{ср.})/S,$$

где $X_{экстр}$ - максимальные и минимальные значения,

$X_{ср.}$ – среднее арифметическое значение,

S - среднее квадратическое (стандартное) отклонение.

3. По специальной таблице (табл. 1.1) определяют теоретическое значение ($t_т$).
4. Сравнивают фактическое и теоретическое значения критерия t для принятия решения о необходимости выбраковки сомнительных данных.
5. При необходимости, если $\tau_ф \geq \tau_т$ (гипотеза H_1), выбраковывают сомнительные данные, определяют среднюю арифметическую и стандартное (среднее квадратическое) отклонения уже для меньшего количества значений.
6. Выполняют повторный расчет фактического значения критерия τ и определяют по таблице его теоретическое значение, но уже для меньшего количества значений.

Расчеты выполняют, пока не будет достигнута полная однородность данных, то есть когда все максимальные и минимальные значения будут соответствовать гипотезе H_0

Этапы проверки данных на однородность с помощью критерия Тау

1. Сортировка значений совокупности в порядке возрастания.
2. Расчет фактического значения критерия Тау ($\tau_ф$) по формулам:

$$\tau_ф\,(\max) = (X_n - X_{n-1})/(X_n - X_2),$$

$$\tau_ф\,(\min) = (X_2 - X_1)/(X_{n-1} - X_1).$$

3. Определение теоретических значений критерия Тау ($\tau_т$) по табл. 2.1.

Таблица 2.1

Теоретические значения критерия Тау$_т$ ($\tau_т$) [1]

Степень свободы (количество значений, *n*)	Уровень значимости, 5% (0,05)
4	0,955
5	0,807
6	0,689
7	0,610
8	0,554
9	0,512
10	0,477
11	0,450
12	0,428
14	0,395
16	0,369
18	0.349
20	0,334
22	0,320
24	0,309
26	0,299
28	0,291
30	0,283

4. Сравнение фактических и теоретических значений для принятия решения о необходимости выбраковки сомнительных данных.
5. Выбраковка при необходимости, если $\tau_ф \geq \tau_т$ (H_1), сомнительных данных.
6. Повторение расчетов фактического значения критерия с определением по таблице его теоретического значения, но уже для меньшего

количества значений до полной однородности данных, то есть пока и максимальные, и минимальные значения не будут соответствовать гипотезе H_0

Ниже приведены примеры заполнения для диссертации или научного отчета таблиц с проверкой данных на однородность при использовании критерия Стьюдента (t) (табл. 2.2) и критерия Тау (τ) (табл. 2.3).

Таблица 2.2

Проверка однородности данных с помощью критерия Стьюдента (t) с удалением сомнительных данных [3]

Показатели	Этапы расчета, убывающее количество данных совокупности и изменяющиеся значения показателей		
	I	II	III
	50	49	48
$X_{ср.}$	8,6	8,7	8,8
S(стандарт. отклон.)	1,5	1,5	1,4
X_n (max)	11,2		
$t_ф$	1,72		
$t_т$	2,01		
Вывод (вид гипотезы)	H_0		
X_1 (min)	5,0	5,4	6,1
$t_ф$	2,4	2,2	1,93
$t_т$	2,01	2,01	2
Вывод (вид гипотезы)	H_1	H_1	H_0

Таблица 2.3

Результаты проверки однородности параметров объектов с помощью критерия Тау (τ) (без сомнительных значений)

Параметры объектов	Значение экстремальных показателей		Принадлежность сомнительных значений к совокупности	
	$\tau_\phi(\min)=\frac{(X_2-X_1)}{(X_{n-1}-X_1)}$	$\tau_\phi(\max)=\frac{(X_n-X_{n-1})}{(X_n-X_2)}$	X_1т	X_nт
Высота (H), м	0,01	0,012	+	+
Площадь (S), га	0,013	0,12	+	+
Объем (V), м^3	0,00008	0,22	+	+
Масса (M), тыс. т	0,0001	0,189	+	+

РАЗДЕЛ 3. ПРОВЕРКА ДАННЫХ НА ДОСТОВЕРНОСТЬ

3.1 ОБЩИЕ СВЕДЕНИЯ О НОРМАЛЬНОМ ЗАКОНЕ РАСПРЕДЕЛЕНИЯ

Проверить данные на достоверность – это значит определить соответствие их нормальному закону распределения (закону Гаусса**).** Полученные в результате наблюдений (экспериментов) данные могут соответствовать различным законам распределения. Закон Гаусса (нормальный закон распределения) является пограничным, а все остальные законы приближаются к нему. Исходя из этого, все значения всегда вначале (но только после проверки на однородность) проверяются на соответствие закону Гаусса. В случае несоответствия, означающего «недостоверность» данных, вначале осуществляется их преобразование с повторной проверкой соответствия, а затем (если после преобразования повторилось несоответствие данных нормальному закону распределения) к полученным данным подбирается другой закон распределения (гамма, Пуассона, биноминальный и т.д.).

Закон Гаусса характеризуют два известных нам статистических показателя: математическое ожидание (m) и среднее квадратическое отклонение (δ).

При нормальном законе распределения в пределах одного положительного и одного отрицательного среднего квадратического отклонения (при расположении генеральной средней в центре) располагаются ≈68% значений, двух стандартных отклонений ≈95% значений, трех ≈99% значений. Внешне кривая Гаусса выглядит следующим образом (рис. 3.1) [1].

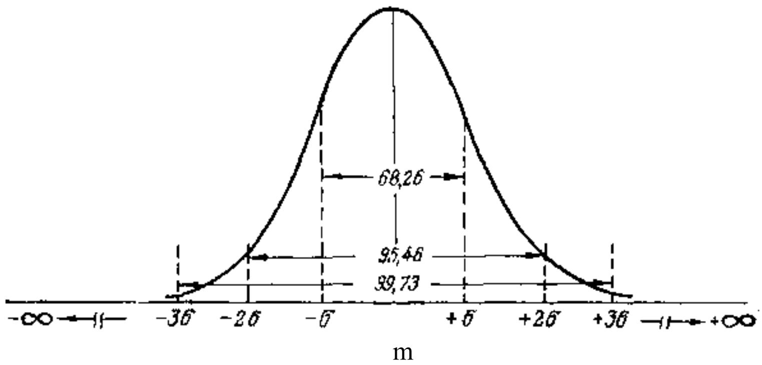


Рис.3.1. Внешний вид кривой Гаусса

Площадь, заключенная под кривой Гаусса в пределах ±σ, ±2σ или ±3σ, где расположена большая часть значений, выраженная в процентах от всей площади под кривой, называется *статистической надежностью* или *уровнем вероятности Р*. Меньшая часть значений, которая не попадает в эти границы, характеризуется другим показателем - уровнем значимости P_1. При одном стандартном отклонении σ от генеральной средней составляет уровень значимости ≈ 32% (0,32); при двух σ ≈5% (0,05); при трех σ ≈ 1% (0,01).

Мы уже использовали 5% (0,05) уровень значимости при нахождении табличного (теоретического) значения критерия Стьюдента (t) (раздел 1).

Уровень вероятности $P_{95\%}$ и уровень значимости $P_{1\ 5\%}$ – это достаточно точные (стандартные) уровни, широко используемые при обработке и анализе информации.

3.2 ВЫДВИЖЕНИЕ ГИПОТЕЗ О ДОСТОВЕРНОСТИ ДАННЫХ

Для проверки данных на достоверность вначале выдвигаются прикидочная и основная гипотезы.

Прикидочная гипотеза базируется на значении коэффициента вариации $V = \frac{S}{\overline{X}} \cdot 100\%$ (раздел 1). Коэффициент вариации может быть уточнен после проверки на однородность (раздел 2).

Основная гипотеза выдвигается по значениям двух показателей: асимметрии и эксцесса.

Асимметрия – величина, характеризующая асимметрию (несимметричность) распределения полученных значений. Рассчитывается она через простые статистические показатели по формуле:

$$A = \frac{\sum (x_i - \overline{x})^3}{n \cdot S^3},$$

где x_i – i-й элемент выборки, $\overline{x}$ – выборочная средняя,

S – стандартное отклонение $\left(S = \sqrt{S^2}\right)$, n – объем выборки.

Эксцесс – мера остроты пика распределения полученных значений.

Эксцесс также определяется через простые статистические показатели по формуле:

$$E = \frac{\sum (x_i - \overline{x})^4}{n \cdot S^4} - 3$$

Для расчета асимметрии и эксцесса используется вспомогательная таблица (табл. 3.1)

Таблица 3.1

Вспомогательная таблица для расчета асимметрии и эксцесса

№ пп	X_i	$X_i - X_{cp}$	$(X_i - X_{cp})^2$	$(X_i - X_{cp})^3$	$(X_i - X_{cp})^4$
1	X_1				
2	X_2				
…	…				
n	X_n				
	$X_{cp} = \overline{X} = \dfrac{\sum X_i}{n}$	$\sum(X_i - X_{cp}) \approx 0$	$\sum(X_i - X_{cp})^2$	$\sum(X_i - X_{cp})^3$	$\sum(X_i - X_{cp})^4$

Вспомогательные таблицы помещать в материалы диссертации (отчета) не рекомендуется, так как они значительно увеличат их объем. В материалах диссертации (отчета) - в тексте или в приложениях размещают таблицы с уже рассчитанными показателями (табл. 3.2 или табл. 3.3).

Таблица 3.2

Результаты расчета коэффициента вариации, асимметрии и эксцесса

№ п/п	Варианты опыта	Расчетные показатели		
		$V = \dfrac{S}{\overline{X}} \cdot 100\%$	$A = \dfrac{\sum (x_i - \overline{x})^3}{n \cdot S^3}$	$E = \dfrac{\sum (x_i - \overline{x})^4}{n \cdot S^4} - 3$
1	a_0b_0	3,63	0,200	-1,473
2	a_1b_0	17,8	0,199	-1,619
3	a_2b_0	26,0	0,199	-1,437
4	a_3b_0	1,82	0,200	-1,407
5	a_4b_0	8,81	0,199	-1,885
6	a_5b_0	39,8	0,166	-0,756
7	a_0b_1	1,37	0,200	-1,742
8	a_1b_1	1,0	0,199	-0,736
9	a_2b_1	1,72	0,199	3,740
10	a_3b_1	2,41	0,200	-0,422

Таблица 3.3

Выдвижение основной гипотезы о достоверности данных (о соответствии их нормальному закону распределения)

Варианты опыта	Показатели и выдвигаемая гипотеза							
	A	E	$3\sigma_a$	$3\sigma_e$	по значению А		по значению Е	
					H_0, если $A < 3\sigma_a$	H_1, $(A \geq 3\sigma_a)$	H_0, $(E < 3\sigma_e)$	H_1, $(E \geq 3\sigma_e)$
1	0,109	-1,633	1,225	2,449	+	-	+	-
2	-0,099	-1,841	1,225	2,449	+	-	+	-
3	-0,169	-0,934	1,225	2,449	+	-	+	-
4	-0,779	0,415	1,414	2,828	+	-	+	-

Для выдвижения прикидочной гипотезы о нормальном законе распределения коэффициент вариации, рассчитанный по данным наблюдений (экспериментов), должен быть менее 33%. Обычно считается, что при нормальном законе распределения асимметрия равна нулю, а эксцесс равен трем, но допускается некоторое отклонение. Основная гипотеза о соответствии полученных данных нормальному закону распределения выдвигается при условии:

$$A < 3\sigma_a \text{ и } E < 3\sigma_e,$$

где σ_a и σ_e определяются по формулам:

$$\sigma_a = \sqrt{\frac{6}{n}}, \ \sigma_в = \sqrt{\frac{24}{n}}.$$

Пример расчета асимметрии и эксцесса

В результате экспериментов аспирантом был получен ряд значений изучаемого показателя X. Для расчета асимметрии и эксцесса он составил вспомогательную таблицу 3.4.

Таблица 3.4

Вспомогательная таблица для расчета асимметрии и эксцесса

№	X_i	$X_i - X_{cp}$	$(X_i - X_{cp})^2$	$(X_i - X_{cp})^3$	$(X_i - X_{cp})^4$
1	19,7	8,32	69,14	574,95	4780,91
2	14,7	3,32	10,99	36,44	120,81
3	10,6	-0,78	0,62	-0,48	0,38
4	3,1	-8,28	68,64	-568,63	4710,95
5	14,2	2,82	7,93	22,31	62,82
6	14,5	3,12	9,71	30,23	94,19
7	15,1	3,72	13,80	51,28	190,53

Продолжение таблицы 3.4

№	X_i	$X_i – X_{cp}$	$(X_i - X_{cp})^2$	$(X_i - X_{cp})^3$	$(X_i - X_{cp})^4$
8	14,8	3,42	11,66	39,84	136,05
9	19,5	8,12	65,86	534,46	4337,28
10	15,4	4,02	16,12	64,74	259,94
11	14,6	3,22	10,34	33,24	106,88
12	6,9	-4,48	20,11	-90,20	404,52
13	13,7	2,32	5,36	12,41	28,74
14	9,6	-1,78	3,19	-5,68	10,15
15	6,7	-4,68	21,95	-102,81	481,65
16	7,5	-3,88	15,09	-58,62	227,74
17	5,5	-5,88	34,63	-203,79	1199,22
18	16,9	5,52	30,42	167,77	925,28
19	11	-0,38	0,15	-0,06	0,02
20	12,9	1,52	2,30	3,48	5,27
21	14,8	3,42	11,66	39,84	136,05
22	8,8	-2,58	6,68	-17,27	44,63
23	16	4,62	21,30	98,31	453,73
24	11,4	0,02	0,00	0,00	0,00
25	11	-0,38	0,15	-0,06	0,02
26	5,3	-6,08	37,02	-225,28	1370,75
27	**1,2**	-10,18	103,73	-1056,44	10759,55
28	15,6	4,22	17,77	74,90	315,73
29	7,9	-3,48	12,14	-42,32	147,46
30	10,4	-0,98	0,97	-0,95	0,94
31	12,9	1,52	2,30	3,48	5,27
32	15,8	4,42	19,49	86,08	380,05
33	12,3	0,92	0,84	0,77	0,70
34	12,4	1,02	1,03	1,05	1,06
35	10,7	-0,68	0,47	-0,32	0,22
36	7,2	-4,18	17,51	-73,28	306,66
37	9	-2,38	5,69	-13,56	32,34
38	12,7	1,32	1,73	2,28	2,99
39	14,8	3,42	11,66	39,84	136,05
40	3,3	-8,08	65,36	-528,44	4272,25
41	13,5	2,12	4,47	9,46	20,02
42	12	0,62	0,38	0,23	0,14
43	17,6	6,22	38,63	240,10	1492,27

Продолжение таблицы 3.4

№	X_i	$X_i - X_{cp}$	$(X_i - X_{cp})^2$	$(X_i - X_{cp})^3$	$(X_i - X_{cp})^4$
44	4,2	-7,18	51,62	-370,87	2664,62
45	12	0,62	0,38	0,23	0,14
46	9,5	-1,88	3,55	-6,69	12,62
47	4,1	-7,28	53,07	-386,58	2816,10
48	15,8	4,42	19,49	86,08	380,05
49	15,3	3,92	15,33	60,02	234,99
50	13,8	2,42	5,83	14,09	34,03
51	7,5	-3,88	15,09	-58,62	227,74
52	6	-5,38	29,00	-156,13	840,71
53	14,6	3,22	10,34	33,24	106,88
54	11,8	0,42	0,17	0,07	0,03
55	9,8	-1,58	2,51	-3,98	6,31
56	9	-2,38	5,69	-13,56	32,34
57	14,5	3,12	9,71	30,23	94,19
58	10,3	-1,08	1,18	-1,28	1,38
59	**2**	-9,38	88,07	-826,54	7756,80
60	10,6	-0,78	0,62	-0,48	0,38
61	11,5	0,12	0,01	0,00	0,00
62	7,8	-3,58	12,85	-46,06	165,13
63	16,9	5,52	30,42	167,77	925,28
64	4,7	-6,68	44,69	-298,71	1996,78
65	**2,3**	-9,08	82,53	-749,78	6811,51
66	12,5	1,12	1,24	1,39	1,55
67	15,7	4,32	18,62	80,36	346,77
68	15,3	3,92	15,33	60,02	234,99
69	8,1	-3,28	10,79	-35,44	116,41
70	4,3	-7,08	50,19	-355,60	2519,34
71	17,1	5,72	32,66	186,69	1066,98
72	15,8	4,42	19,49	86,08	380,05
73	17,8	6,42	41,16	264,03	1693,82
74	12,3	0,92	0,84	0,77	0,70
75	9,2	-2,18	4,77	-10,43	22,78
76	5,4	-5,98	35,82	-214,35	1282,84
77	12,6	1,22	1,48	1,79	2,18
78	14,2	2,82	7,93	22,31	62,82
79	15,2	3,82	14,56	55,54	211,89
80	16,4	5,02	25,15	126,15	632,68

Продолжение таблицы 3.4

№	X_i	$X_i - X_{cp}$	$(X_i - X_{cp})^2$	$(X_i - X_{cp})^3$	$(X_i - X_{cp})^4$
81	14,7	3,32	10,99	36,44	120,81
82	13,8	2,42	5,83	14,09	34,03
83	17,1	5,72	32,66	186,69	1066,98
84	9,4	-1,98	3,94	-7,82	15,52
85	3,3	-8,08	65,36	-528,44	4272,25
Σ	967,7	0,0	1679,95	-3378,01	77155,59

1) $\bar{X} = \dfrac{967,7}{85} = 11,4$

2) $S^2 = \dfrac{1679.3}{85-1} = 19.99$

3) $S = \sqrt{19.99} = 4.47$

4) $V = \dfrac{4.47}{11.4} \cdot 100\% = 39.27\%$

5) Асимметрия $A = \dfrac{\sum (x_i - \bar{x})^3}{n \cdot S^3}$ = - 0,44;

6) Эксцесс $E = \dfrac{\sum (x_i - \bar{x})^4}{n \cdot S^4} - 3$ = - 0,73.

Для того, чтобы эмпирические точки соответствовали нормальному закону распределения, должны соблюдаться неравенства:

$A/\sigma_a < 3$ и $E/\sigma_e < 3$,

где $\sigma_a = \sqrt{\dfrac{6}{n}}$, $\sigma_e = \sqrt{\dfrac{24}{n}}$

$$\sigma_a = \sqrt{\frac{6}{85}} = 0,266\,; \quad \sigma_e = \sqrt{\frac{24}{85}} = 0,531\,;$$

$$\frac{A}{\sigma_a} = \frac{-0,44}{0,266} = -1,7\,; \quad \frac{E}{\sigma_e} = \frac{-0,73}{0,531} = -1,4$$

В приведенном примере асимметрия и эксцесс (A и E) рассчитывались с использованием простых статистических показателей, однако они могут быть рассчитаны, также, с помощью метода моментов. Выбор одного из этих двух методов для расчета A и E – это решение самого исследователя.

Расчет асимметрии и эксцесса методом моментов [3]

Различают три момента: начальный (ν), центральный (μ) и основной (r). Все эти моменты могут иметь различные порядки (l) (1, 2, 3, 4, 5 и т.д.). Чаще всего в расчетах используют первые четыре порядка.

Для расчета начального момента (ν) используется следующая общая формула:

$$\hat{\nu}_l = \frac{1}{n}\sum_{i=1}^{k}\tilde{x}_i^l\, m_i \ ,$$

где n - объем выборочной совокупности, $\tilde{x}_i$ - средина i-го интервала, m_i- эмпирические частоты.

Для первых четырех порядков начальный момент (ν) определяется по следующим формулам:

$$\hat{\nu}_1 = \hat{m}_x = \overline{x} = \frac{1}{n}\sum_{i=1}^{k}\tilde{x}_i m_i \ ,$$

$$\hat{\nu}_2 = \frac{1}{n}\sum_{i=1}^{k}\tilde{x}_i^2\, m_i = \overline{x^2},$$

$$\hat{\nu}_3 = \frac{1}{n}\sum_{i=1}^{k}\tilde{x}_i^3\, m_i = \overline{x^3},$$

$$\hat{\nu}_4 = \frac{1}{n}\sum_{i=1}^{k}\tilde{x}_i^4\, m_i = \overline{x^4} \ .$$

Общая формула для определения центрального момента (μ) выглядит следующим образом:

$$\hat{\mu}_l = \frac{1}{n}\sum_{i=1}^{k}\left(\tilde{x}_i - \overline{x}\right)^l m_i \ ,$$

где $\overline{X} = \dfrac{\sum X_i}{n}$ - средняя арифметическая.

Для первых четырех порядков центральный момент (μ) определяется по следующим формулам:

$$\hat{\mu}_l = \frac{1}{n}\sum_{i=1}^{k}\left(\tilde{x}_i - \overline{x}\right)^l m_i \ ,$$

$$\hat{\mu}_2 = \hat{\sigma}_x^2 = \frac{1}{n}\sum_{i=1}^{k}\left(\tilde{x}_i - \overline{x}\right)^2 m_i \ ,$$

$$\hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^{k}\left(\tilde{x}_i - \overline{x}\right)^3 m_i \ ,$$

$$\hat{\mu}_4 = \frac{1}{n}\sum_{i=1}^{k}\left(\tilde{x}_i - \overline{x}\right)^4 m_i \ .$$

Для расчета основного момента (r) используется стандартное отклонение $S = \sqrt{S^2}$, но дисперсия в методе моментов рассчитывается по следующей формуле:

$$S_x^2 = \frac{n}{n-1}\hat{\mu}_2 .$$

Общая формула для расчета основного момента (r) выглядит следующим образом:

$$\hat{r}_l = \frac{\hat{\mu}_l}{S_x^l} .$$

Основные моменты (r) первого и второго порядков равны:

$$r_1 = 0, \text{ а } r_2 = 1.$$

Асимметрия и эксцесс определяются по основным моментам r_3 и r_4:

$$\hat{r}_3 = \frac{\hat{\mu}_3}{S_x^3} = \text{А},$$

$$E = \hat{r}_4 - 3 \text{, где}$$

$$\hat{r}_4 = \frac{\hat{\mu}_4}{S_x^4} .$$

Для расчета асимметрии и эксцесса используются вспомогательные таблицы (табл. 3.5 и табл. 3.6) [3].

Таблица 3.5

Расчет начальных моментов

№ интервала, i	Границы i -го интервала		Средина интервала i , $\tilde{x}_i$	Эмпирические частоты, m_i	$\tilde{x}_i m_i$	$\tilde{x}_i^2 m_i$	$\tilde{x}_i^3 m_i$	$\tilde{x}_i^4 m_i$
	начало x_{i-1}	конец x_{i+1}						
1	-9,4($X_{\min}$)	-7,6	-8,5	4	-34,00	289,00	-2456,50	20880,25
2	-7,6	-5,8	-6,7	2	-13,40	89,78	-601,53	4030,22
3	-5,8	-4,0	-4,9	3	-14,70	72,03	-352,95	1729,44
4	-4,0	-2,2	-3,1	7	-21,70	67,27	-208,54	646,66
5	-2,2	-0,4	-7,8	6	-7,8	10,14	-13,18	17,14
6	-0,4	1,4	0,5	11	5,50	2,75	1,38	0,69
7	1,4	3,2	2,3	2	4,60	10,58	24,33	55,97
				Сумма	-81,50	541,55	-3606,99	27360,17

Таблица 3.6

Расчет центральных моментов

№ интер-вала, i	Границы интервала i		Средина интервала i $\tilde{x}_i$	Эмпирические частоты m_i	$(\tilde{x}_i - \bar{x})m_i$	$(\tilde{x}_i - \bar{x})^2 m_i$	$(\tilde{x}_i - \bar{x})^3 m_i$	$(\tilde{x}_i - \bar{x})^4 m_i$
	начало x_{i-1}	конец x_{i+1}						
1	-9,4($X_{\min}$)	-7,6	-8,5	4	-6,2	153,76	-953,31	5910,53
2	-7,6	-5,8	-6,7	2	-4,4	38,72	-170,37	749,62
3	-5,8	-4,0	-4,9	3	2,6	20,28	-53,73	137,09
4	-4,0	-2,2	-3,1	7	-0,8	4,48	-3,58	2,87
5	-2,2	-0,4	-7,8	6	1,0	6,0	6,0	6,0
6	-0,4	1,4	0,5	11	2,8	86,24	241,47	676,12
7	1,4	3,2	2,3	2	4,6	42,32	194,67	895,49
					Сумма	351,80	-737,85	8377,72

Заполнение обеих таблиц начинается одинаково - с определения области значений (размаха варьирования):

$$R = X_{\max} - X_{\min}$$

где $X_{\max}$ – максимальное значение показателя выборочной совокупности,

$X_{\min}$ – минимальное значение показателя выборочной совокупности.

Далее определяется количество групп (k) по формуле:

$$k = 5lg\ n,$$

где n - объем выборочной совокупности

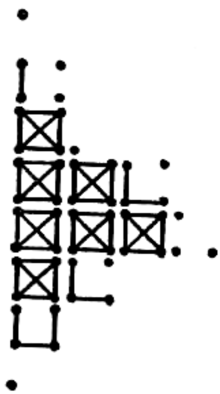
На основании области значений и количества групп определяется интервал $i = \dfrac{R}{k}$.

Эмпирические частоты (m_i), то есть частоты встречаемости значений в каждом интервале устанавливают путем разноски исходных данных. Для разноски используют либо способ «штрихов», либо способ «конвертиков».

Способ «штрихов» заключается в том, что значение, попадающее в определенный интервал обозначают штрихами по 5 штук. По второму способу – способу «конвертиков» первые четыре значения, попадающие в один интервал изображают точками по углам квадрата; следующие четыре значения №5 - №8 отмечают в виде сторон квадрата, соединяющих ранее нанесенные точки, 9-е и 10-е значение – в виде диагоналей (табл. 3.7) [1].

Таблица 3.7

Разноска исходных данных

Группа	Способ «штрихов»	Способ «конвертиков»	Частота f	Групповые варианты
1. 40,0—49,9			1	45
2. 50,0—59,9			5	55
3. 60,0—69,9			11	65
4. 70,0—79,9			26	75
5. 80,0—89,9			33	85
6. 90,0—99,9			16	95
7. 100,0—109,0			7	105
8. 110,0—120,0			1	115
Сумма			100	

Результаты расчетов асимметрии и эксцесса с помощью методов моментов можно привести в диссертации или отчете в таком виде (табл. 3.8).

Таблица 3.8

Расчетные значения моментов, асимметрия и эксцесс

Пока-зате-ли	Вид радиоактивности, варианты опыта					
	$\beta+\gamma$			γ		
	1	2	3	1	2	3
1	2	3	4	5	6	7
ν_1	10,96	18,99	16,07	9	13,54	12,23
ν_2	2743,35	380,25	367,59	187,88	201,58	202,80
ν_3	7167,73	7994,19	8895,22	4107,50	3228,42	3620,32
ν_4	194242,7	175705,8	224898,20	93683,53	54963,34	68047,97
μ_2	13,26	19,47	26,23	9,46	16,13	21,64

Продолжение таблицы 3.8

1	2	3	4	5	6	7
μ_3	-5,44	30,78	12,18	8,92	15,13	-88,12
μ_4	979,02	967,48	2170,79	483,75	566,54	1246,01
S^2	32,03	22,87	19,81	24,83	12,99	27,29
S	5,66	4,78	4,45	4,98	3,61	4,93
r_3	-0,03	0,28	0,14	0,07	0,32	-0,74
r_4	0,95	1,85	5,53	0,78	3,35	2,11
E	-2,05	-1,15	2,53	-2,22	0,35	-0,89
$3\delta_a$	0,49	0,87	0,64	0,49	0,78	0,68
$3\delta_a$	0,98	1,74	1,28	0,99	1,57	1,36
$A<3\delta_a$	+	+	+	+	+	+
$E<3\delta_a$	+	+	+	+	+	+

3.3 ПОДТВЕРЖДЕНИЕ ГИПОТЕЗЫ О ДОСТОВЕРНОСТИ ДАННЫХ

Подтверждение гипотез о достоверности данных может осуществляться качественно и количественно. Качественное подтверждение можно проводить:

1) используя метод спрямленных диаграмм [2];

2) путем построения гистограммы и полигона (кривой распределения) и визуального сравнения этой кривой с кривой Гаусса;

3) путем сравнения между собой средней арифметической, моды и медианы.

3.3.1. Подтверждение достоверности данных с помощью метода спрямленных диаграмм [2]

При использовании этого метода составляется расчетная таблица, которая впоследствии может быть размещена в материалах отчета или диссертации.

Пример. Эмпирическое распределение выборки задано в таблице в виде последовательности значений x_i, расположенных в возрастающем порядке, и соответствующих им частот. Эмпирические частоты вначале преобразуются в накопленные, а затем в относительные накопленные в долях единицы и процентах (табл. 3.9).

В последней колонке таблицы выставляется значение квантиля. Его значение определяется по табл. 3.10 и 3.11 путем использования относительной накопленной частоты.

Таблица 3.9

Определение относительной накопленной частоты и квантилей [2]

Номер i	Значения x_i	Частоты n_i	Накопленные частоты $N_i = \sum_{r=1}^{i} n_r - \frac{1}{2}$	Относительные накопленные частоты $F^*(x_i) = \frac{N_i}{n}$	Относительные накопленные частоты, % $P_i = F^*(x_i) \cdot 10$	Квантили u_{pi}
1-2	7	2	1,5	0,04	4	-1,751
3	9	1	2,5	0,07	7	-1,476
4	12	1	3,5	0,10	10	-1,282
5	14	1	4,5	0,13	13	-1,126
6	15	1	5,5	0,15	15	-1,036
7-9	18	3	8,5	0,24	24	-0,706
10-11	20	2	10,5	0,29	29	-0,553
12	22	1	11,5	0,32	32	-0,468
13-14	26	2	13,5	0,38	38	-0,305
15-16	28	2	15,5	0,43	43	-0,176
17-18	29	2	17,5	0,49	49	-0,025
19	30	1	18,5	0,51	51	0,025
20-21	31	2	20,5	0,57	57	0,176
22-23	32	2	22,5	0,63	63	0,332
24-25	34	2	24,5	0,68	68	0,468
26-28	35	3	27,5	0,76	76	0,706
29	36	1	28,5	0,79	79	0,806
30	41	1	29,5	0,82	82	0,915
31	42	1	30,5	0,85	85	1,036
32	44	1	31,5	0,88	88	1,175
33-34	45	2	33,5	0,93	93	1,476
35	46	1	34,5	0,96	96	1,751
36	48	1	35,5	0,99	99	2,326

Таблица 3.10

Квантили нормального распределения u_p [2]

(перед всеми значениями квантилей нужно поставить минус)

100 Pi, %	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0	-	3,090	2,878	2,748	2,652	2,576	2,512	2,457	2,409	2,360
1	2,326	2,290	2,257	2,226	2,197	2,170	2,144	2,120	2,097	2,075
2	2,054	2,034	2,014	1,995	1,977	1,960	1,943	1,927	1,911	1,896
3	1,881	1,866	1,852	1,838	1,825	1,812	1,799	1,787	1,774	1,762
4	1,751	1,739	1,728	1,717	1,706	1,695	1,685	1,675	1,665	1,655
5	1,645	1,635	1,626	1,616	1,607	1,598	1,589	1,580	1,572	1,563
6	1,555	1,546	1,538	1,530	1,522	1,514	1,506	1,499	1,491	1,483
7	1,476	1,468	1,461	1,454	1,447	1,440	1,433	1,426	1,419	1,412
8	1,405	1,398	1,392	1,385	1,379	1,372	1,366	1,359	1,353	1,347
9	1,341	1,335	1,329	1,323	1,317	1,311	1,305	1,299	1,293	1,287
10	1,282	1,276	1,270	1,265	1,259	1,254	1,248	1,243	1,237	1,232
11	1,227	1.221	1,216	1,211	1,206	1,200	1,195	1,190	1,185	1,180
12	1,175	1,170	1,165	1,160	1,155	1,150	1,146	1,141	1,136	1,131
13	1,126	1,122	1,117	1,112	1,108	1,103	1,098	1,094	0,089	1,085
14	1,080	1,076	1,071	1,067	1,063	1,058	1,054	1,049	1,045	1,041
15	1,036	1,032	1,028	1,024	1,019	1,015	1,011	1,007	1,003	0,99
16	0,994	0,990	0,986	0,982	0,978	0,974	0,970	0,966	0,962	0,958
17	0,954	0,950	0,946	0,942	0,938	0,935	0,931	0,927	0,923	0,919
18	0,915	0,912	0,908	0,904	0,900	0,896	0,893	0,889	0,885	0,882
19	0,878	0,874	0,871	0,867	0,863	0,860	0,856	0,852	0,849	0,846
20	0,842	0,838	0,834	0,831	0,827	0,824	0,820	0,817	0,813	0,810
21	0,806	0,803	0,800	0,796	0,793	0,789	0,786	0,782	0,779	0,776
22	0,772	0,769	0,765	0,762	0,759	0,755	0,752	0,749	0,745	0,742
23	0,739	0,736	0,732	0,729	0,726	0,722	0,719	0,716	0,713	0,710
24	0,706	0,703	0,700	0,697	0,693	0,690	0,687	0,684	0,681	0,678
25	0,674	0,671	0,668	0,665	0,662	0,659	0,656	0,653	0,650	0,646
26	0,643	0,640	0,637	0,634	0,631	0,628	0,625	0,622	0,619	0,616
27	0,613	0,610	0,607	0,604	0,601	0,598	0,595	0,592	0,589	0,586
28	0,583	0,580	0,577	0,574	0,571	0,568	0,565	0,562	0,559	0,556
29	0,553	0,550	0,548	0,545	0,542	0,539	0,536	0,533	0,530	0,527
30	0,524	0,522	0,519	0,516	0,513	0,510	0,507	0,504	0,502	0,499
31	0,496	0,493	0,490	0,487	0,485	0,482	0,479	0,476	0,473	0,470
32	0,468	0,465	0,462	0,459	0,457	0,454	0,451	0,448	0,445	0,443
33	0,440	0,437	0,434	0,432	0,429	0,426	0,423	0,421	0,418	0,415
34	0,412	0,410	0,407	0,404	0,402	0,399	0,396	0,393	0,391	0,388
35	0,385	0,383	0,380	0,377	0,375	0,372	0,369	0,366	0,364	0,361

Продолжение таблицы 3.10

36	0,358	0,356	0,353	0,350	0,348	0,345	0,342	0,340	0,337	0,335
37	0,332	0,329	0,327	0,324	0,321	0,319	0,316	0,313	0,311	0,308
38	0,305	0,303	0,300	0,298	0,295	0,292	0,290	0,287	0,285	0,282
39	0,279	0,277	0,274	0,272	0,269	0,266	0,264	0,261	0,259	0,256
40	0,253	0,251	0,248	0,246	0,243	0,240	0,238	0,235	0,233	0,230
41	0,228	0,225	0,222	0,220	0,217	0,215	0,212	0,210	0,207	0,204
42	0,202	0,199	0,197	0,194	0,192	0,189	0,187	0,184	0,181	0,179
43	0,176	0,174	0,171	0,169	0,166	0,164	0,161	0,159	0,156	0,154
44	0,151	0,148	0,146	0,143	0,141	0,138	0,136	0,133	0,131	0,128
45	0,126	0,123	0,121	0,118	0,116	0,113	0,111	0,108	0,105	0,103
46	0,100	0,098	0,095	0,093	0,090	0,088	0,085	0,083	0,080	0,078
47	0,075	0,073	0,070	0,068	0,065	0,063	0,060	0,058	0,055	0,053
48	0,050	0,048	0,045	0,043	0,040	0,038	0,035	0,033	0,030	0,028
49	0,025	0,023	0,020	0,018	0,015	0,013	0,010	0,008	0,005	0,003

Таблица 3.11

Квантили нормального распределения u_p [2]

(все значения квантилей положительные)

100 Pi, %	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
50	-	0,003	0,005	0,008	0,010	0,013	0,015	0,018	0,020	0,023
51	0,025	0,028	0,030	0,033	0,035	0,038	0,040	0,043	0,045	0,048
52	0,050	0,053	0,055	0,058	0,060	0,063	0,065	0,068	0,070	0,073
53	0,075	0,078	0,080	0,083	0,085	0,088	0,090	0,093	0,095	0,098
54	0,100	0,103	0,105	0,108	0,111	0,113	0,116	0,118	0,121	0,123
55	0,126	0,128	0,131	0,133	0,136	0,138	0,141	0,143	0,146	0,148
56	0,151	0,154	0,156	0,159	0,161	0,164	0,166	0,169	0,171	0,174
57	0,176	0,179	0,181	0,184	0,187	0,189	0,192	0,194	0,197	0,199
58	0,202	0,204	0,207	0,210	0,212	0,215	0,217	0,220	0,222	0,225
59	0,228	0,230	0,233	0,235	0,238	0,240	0,243	0,246	0,248	0,251
60	0,253	0,256	0,259	0,261	0,264	0,266	0,269	0,272	0,274	0,277
61	0,279	0,282	0,285	0,287	0,290	0,292	0,295	0,298	0,300	0,303
62	0,305	0,308	0,311	0,313	0,316	0,319	0,321	0,324	0,327	0,329
63	0,332	0,335	0,337	0,340	0,342	0,345	0,348	0,350	0,353	0,356
64	0,358	0,361	0,364	0,366	0,369	0,372	0,375	0,377	0,380	0,383
65	0,385	0,388	0,391	0,393	0,396	0,399	0,402	0,404	0,407	0,410
66	0,412	0,415	0,418	0,421	0,423	0,426	0,429	0,432	0,434	0,437
67	0,440	0,443	0,445	0,448	0,451	0,454	0,457	0,459	0,462	0,465
68	0,468	0,470	0,473	0,476	0,479	0,482	0,485	0,487	0,490	0,493
69	0,496	0,499	0,502	0,504	0,507	0,510	0,513	0,516	0,519	0,522
70	0,524	0,527	0,530	0,533	0,536	0,539	0,542	0,545	0,548	0,550

Продолжение таблицы 3.11

100 Pi, %	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
71	0,553	0,556	0,559	0,562	0,565	0,568	0,571	0,574	0,577	0,580
72	0,583	0,586	0,589	0,592	0,595	0,598	0,601	0,604	0,607	0,610
73	0,613	0,616	0,619	0,622	0,625	0,628	0,631	0,634	0,637	0,640
74	0,643	0,646	0,650	0,653	0,656	0,659	0,662	0,665	0,668	0,671
75	0,674	0,678	0,681	0,684	0,687	0,690	0,693	0,697	0,700	0,703
76	0,706	0,710	0,713	0,716	0,719	0,722	0,726	0,729	0,732	0,736
77	0,739	0,742	0,745	0,749	0,752	0,755	0,759	0,762	0,765	0,769
78	0,772	0,776	0,779	0,782	0,786	0,789	0,793	0,796	0,800	0,803
79	0,806	0,810	0,813	0,817	0,820	0,824	0,827	0,831	0,834	0,838
80	0,842	0,845	0,849	0,852	0,856	0,860	0,863	0,867	0,871	0,874
81	0,878	0,882	0,885	0,889	0,893	0,896	0,900	0,904	0,908	0,912
82	0,915	0,919	0,923	0,927	0,931	0,935	0,938	0,942	0,946	0,950
83	0,954	0,958	0,962	0,966	0,970	0,974	0,978	0,982	0,986	0,990
84	0,994	0,999	1,003	1,007	1,011	1,015	1,019	1,024	1,028	1,032
85	1,036	1,041	1,045	1,049	1,054	1,058	1,063	1,067	1,071	1,076
86	1,080	1,085	1,089	1,094	1,098	1,103	1,108	1,112	1,117	1,122
87	1,126	1,131	1,136	1,141	1,146	1,150	1,155	1,160	1,165	1,170
88	1,175	1,180	1,185	1,190	1,195	1,200	1,206	1,211	1,216	1,221
89	1,227	1,232	1,237	1,243	1,248	1,254	1,259	1,265	1,270	1,276
90	1,282	1,287	1,293	1,299	1,305	1,311	1,317	1,323	1,329	1,335
91	1,341	1,347	1,353	1,359	1,366	1,372	1,379	1,385	1,392	1,398
92	1,405	1,412	1,419	1,426	1,433	1,440	1,447	1,454	1,464	1,468
93	1,476	1,483	1,491	1,499	1,506	1,514	1,522	1,530	1,538	1,546
94	1,555	1,563	1,572	1,580	1,589	1,598	1,607	1,616	1,626	1,635
95	1,645	1,655	1,665	1,675	1,685	1,695	1,706	1,717	1,728	1,739
96	1,751	1,762	1,774	1,778	1,799	1,812	1,825	1,838	1,852	1,866
97	1,881	1,896	1,911	1,927	1,943	1,960	1,977	1,995	2,014	2,034
98	2,054	2,075	2,097	2,120	2,144	2,170	2,197	2,226	2,257	2,290
99	2,326	2,366	2,409	2,457	2,512	2,576	2,652	2,748	2,878	3,090

На рис. 4.2 в прямоугольной системе координат построены точки (x_i; u_{pi}), то есть по оси абсцисс располагаются значения x_i, по оси ординат – квантили u_{pi}.

Построенные точки лежат вблизи прямой, поэтому нет оснований отвергнуть гипотезу о нормальном распределении данных.

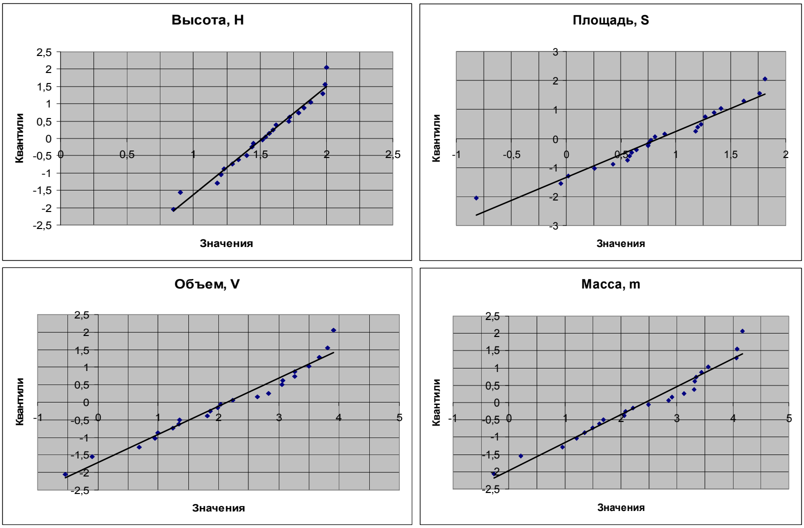


Рис. 4.2. Графическая проверка соответствия данных нормальному закону распределения

3.3.2. Подтверждение достоверности данных путем визуального сравнения эмпирической кривой с кривой распределения Гаусса

На основании полученных данных строится кривая, по которой определяется соответствующий закон распределения. Для подтверждения достоверности данных полученная кривая должна быть внешне похожа на кривую Гаусса. Построение этой кривой начинается с уже известного (по методу моментов) определения области значений (размаха варьирования) [3].

$$R = X_{\max} - X_{\min}$$

где $X_{\max}$ – максимальное значение показателя выборочной совокупности,

$X_{\min}$ – минимальное значение показателя выборочной совокупности,

n– объем выборочной совокупности.

Далее определяется количество групп (k) по формуле:

$$k = 5lg\, n,$$

где n - объем выборочной совокупности.

На основании области значений и количества групп определяется интервал $i = \frac{R}{k}$, который заносится во вспомогательную таблицу 3.12.

Таблица 3.12

Расчет интервалов и эмпирических частот

Номер интервала i	Границы интервала i		Средина интервала i, $\tilde{x}_i$	Эмпирические частоты, m_i
	начало x_{i-1}	конец x_{i+1}		
1	5,7 ($X_{\min}$)	11,4	8,55	1
2	11,4	17,1	14,25	6
3	17,1	22,8	19,95	9
4	22,8	28,5	25,65	9
5	28,5	34,2	31,35	6
6	34,2	39,9	37,05	3
7	39,9	45,6 ($X_{\max}$)	42,75	2

Эмпирические частоты (m_i), то есть частоты встречаемости в каждом интервале устанавливают путем разноски исходных данных. Для разноски используют либо способ «штрихов», либо способ «конвертиков» (табл. 3.7) [1].

На основании рассчитанных интервалов и эмпирических частот строятся гистограмма и полигон. Гистограмма это ступенчатый график в виде столбиков, имеющих высоту, пропорциональную частотам, а ширину, равную интервалам. Полигон (кривая распределения) – линия, соединяющая средины верха каждого столбика гистограммы.

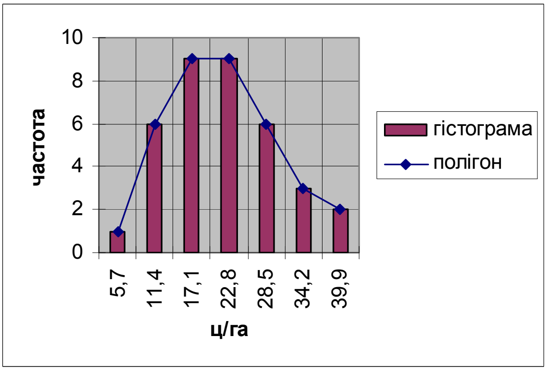


Рис. 4.3 Гистограмма и полигон эмпирических частот

Теперь производится визуальное сравнение полученной кривой распределения (полигона) с кривой Гаусса (рис. 4.3) и делается вывод относительно соответствия данных нормальному закону распределения.

3.3.3. Подтверждение достоверности данных путем сравнения между собой средней арифметической, моды и медианы

Для использования этого способа прежде всего необходимо разобраться с такими понятиями как мода (*Mo*) и медиана (*Me*).

Мода (Mo) - это интервал с самым большим значением эмпирической частоты. Встречаются случаи, когда кривая распределения имеет две и более моды, то есть является двухмодальной или трехмодальной.

Медиана (Me) - это средина интервала, в который входит средняя арифметическая.

Мода и медиана рассчитываются по специальным формулам [3]:

$$M_o = x_0 + c\frac{(m_i - m_{i-1})}{(2m_i - m_{i-1} - m_{i+1})},$$

где x_0, c, m_i – соответственно начало, длина и эмпирическая частота модального интервала;

$m_{i-1},\ m_{i+1}$ – частоты предыдущего и последующего за модальным интервалов.

$$Me = x_e + \frac{c\left(\frac{n}{2} - m^*\right)}{m_e},$$

где n – объем выборки (при условии непарного значения берется $n+1$);

m^* – накопленная частота до медианного интервала;

x_e, c, m_e – соответственно начало, длина и эмпирическая частота медианного интервала.

При нормальном распределении данных наблюдается примерное равенство средней арифметической, моды и медианы:

$$X_{cp} \approx Mo \approx Me$$

3.3.4. Количественное подтверждение гипотезы о достоверности данных с помощью критерия χ^2 Пирсона

Наиболее точным доказательством соответствия данных нормальному закону распределения является количественное подтверждение гипотезы о достоверности с помощью критерия χ^2 Пирсона:

$\chi^2{}_{ф} < \chi^2{}_{т}$ – гипотеза H_0 – данные достоверны;

$\chi^2{}_{ф} \geq \chi^2{}_{т}$ – гипотеза H_1 – данные недостоверны.

Для расчета фактического значения критерия χ^2 Пирсона ($\chi^2{}_{ф}$) необходимы значения эмпирических и теоретических частот. Определение эмпирических частот представлено в подразделе 3.3.2.

Теоретические частоты можно рассчитать с помощью интеграла вероятности $\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-t^2/2} dt$ или плотности вероятности $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$.

Для этих расчетов необходимы эмпирические границы: начало интервала (x_{i-1}), конец интервала (x_{i+1}), среднее значение интервала ($\tilde{x}_i$), а также новые левые и правые границы каждого интервала (t_{i-1}, t_{i+1}). Они могут быть рассчитаны по следующим формулам [3]:

$$t_{i-1} = \frac{x_{i-1} - \overline{x}}{S},$$

$$t_{i+1} = \frac{x_{i+1} - \overline{x}}{S},$$

$$t_i = \frac{t_{i+1} + t_{i-1}}{2}.$$

При использовании интеграла вероятности $\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-t^2/2} dt$ теоретические частоты расчитываются следующим образом: $\tilde{m}_i = n\, p_i$,

где $p_i = P\left(t_{i-1} < t < t_{i+1}\right) = \frac{1}{2}\left[\Phi\left(t_{i+1}\right) - \Phi\left(t_{i-1}\right)\right]$.

При использовании плотности вероятности $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ теоретические частоты определяются по следующей формуле :

$$\tilde{m}_i = \frac{nc}{S} f(t_i),$$

где n – объем выборочной совокупности;

c – длина интервала;

S – среднее квадратическое отклонение;

$f(t_i)$ – плотность вероятности (t_i).

Ниже представлены примеры вспомогательных таблиц для расчета теоретических частот (табл. 3.13, 3.14).

Таблица 3.13

Расчет теоретических частот с помощью $\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-t^2/2} dt$

Исх. границы интервалов		Ср. знач. интер-валов, $\tilde{x}_i$	Эмпири-ческие часто-ты, m_i	Новые границы		t_i	$\Phi(t_{i-1})$	$\Phi(t_{i+1})$	P_i	Теорети-ческие частоты, $\tilde{m}_i$
x_{i-1}	x_{i+1}			t_{i-1}	t_{i+1}					
4,50	6,86	5,68	10	-1,18	-0,62	-0,9	0,762	0,464	-0,149	-4,77
6,86	9,22	8,04	9	-0,62	-0,05	-0,34	0,465	0,040	-0,212	-6,78
9,22	11,58	10,40	5	-0,05	0,51	0,23	0,040	0,399	0,35	11,20
11,58	13,94	12,76	3	0,51	1,07	0,79	0,399	0,715	0,163	5,22
13,94	16,30	15,12	2	1,07	1,63	1,35	0,715	0,897	0,09	2,88
16,30	18,66	17,48	2	1,63	2,19	1,91	0,897	0,972	0,037	1,18
18,66	21,02	19,84	1	2,19	2,75	2,47	0,972	0,986	0,008	0,256

Таблица 3.14

Расчет теоретических частот с помощью $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$

i	Исходные границы		$\tilde{x}_i$	Эмпириче-ские частоты, m_i	Новые границы		t_i	$f(t_i)$	Теоретичес-кие частоты, $\tilde{m}_i$
	x_{i-1}	x_{i+1}			t_{i-1}	t_{i+1}			
1	5,7	11,4	8,55	1	-2,17	-1,52	-1,84	0,0734	1,7
2	11,4	17,1	14,25	6	-1,52	-0,86	-1,19	0,1965	4,6
3	17,1	22,8	19,95	9	-0,86	-0,20	-0,53	0,3448	8,2
4	22,8	28,5	25,65	9	-0,20	0,45	0,13	0,3956	9,4
5	28,5	34,2	31,35	6	0,45	1,11	0,78	0,2943	7,0
6	34,2	39,9	37,05	3	1,11	1,77	1,44	0,1415	3,3
7	39,9	45,6	42,75	2	1,77	2,42	2,10	0,044	1,0

Для упрощения расчетов значения интеграла вероятности $\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-t^2/2} dt$ и плотности вероятности $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$, используемые в расчетах, приведены в специальных таблицах (3.15, 3.16).

Таблица 3.15

Значения интеграла вероятности $\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-t^2/2} dt$

t	Ф(t)	t	Ф(t)	t	Ф(t)	t	Ф(t)
0.00	0.00000	1.00	0.68269	2.00	0.95450	3.00	0.99730
0.01	0.00798	1.01	0.68750	2.01	0.95557	3.01	0.99739
0.02	0.01596	1.02	0.69227	2.02	0.95662	3.02	0.99747
0.03	0.02393	1.03	0.69699	2.03	0.95764	3.03	0.99755
0.04	0.03191	1.04	0.70166	2.04	0.95865	3.04	0.99763
0.05	0.03988	1.05	0.70628	2.05	0.95964	3.05	0.99771
0.06	0.04784	1.06	0.71086	2.06	0.96060	3.06	0.99779
0.07	0.05581	1.07	0.71538	2.07	0.96155	3.07	0.99786
0.08	0.06376	1.08	0.71986	2.08	0.96247	3.08	0.99793
0.09	0.07171	1.09	0.72429	2.09	0.96338	3.09	0.99800
0.10	0.07966	1.10	0.72867	2.10	0.96427	3.10	0.99806
0.11	0.08759	1.11	0.73300	2.11	0.96514	3.11	0.99813
0.12	0.09552	1.12	0.73729	2.12	0.96599	3.12	0.99819
0.13	0.10348	1.13	0.74152	2.13	0.96683	3.13	0.99825
0.14	0.11134	1.14	0.74571	2.14	0.96765	3.14	0.99831
0.15	0.11924	1.15	0.74986	2.15	0.96844	3.15	0.99837
0.16	0.12712	1.16	0.75395	2.16	0.96923	3.16	0.99842
0.17	0.13499	1.17	0.75800	2.17	0.96999	3.17	0.99848
0.18	0.14285	1.18	0.76200	2.18	0.97074	3.18	0.99853
0.19	0.15069	1.19	0.76595	2.19	0.97148	3.19	0.99858
0.20	0.15852	1.20	0.76986	2.20	0.97219	3.20	0.99863
0.21	0.16633	1.21	0.77372	2.21	0.97289	3.21	0.99867
0.22	0.17413	1.22	0.77754	2.22	0.97358	3.22	0.99872
0.23	0.18191	1.23	0.78130	2.23	0.97425	3.23	0.99876
0.24	0.18967	1.24	0.78502	2.24	0.97491	3.24	0.99880
0.25	0.19741	1.25	0.78870	2.25	0.97555	3.25	0.99855
0.26	0.20514	1.26	0.79233	2.26	0.97618	3.26	0.99889
0.27	0.21284	1.27	0.79592	2.27	0.97679	3.27	0.99892

Продолжение таблицы 3.15

t	Φ(t)	t	Φ(t)	t	Φ(t)	t	Φ(t)
0.28	0.22052	1.28	0.79945	2.28	0.97739	3.28	0.99896
0.29	0.22818	1.29	0.80295	2.29	0.97798	3.29	0.99900
0.30	0.23582	1.30	0.80640	2.30	0.97855	3.30	0.99903
0.31	0.24344	1.31	0.80980	2.31	0.97911	3.31	0.99907
0.32	0.25103	1.32	0.81316	2.32	0.97966	3.32	0.99910
0.33	0.25860	1.33	0.81648	2.33	0.98019	3.33	0.99913
0.34	0.26614	1.34	0.81975	2.34	0.98072	3.34	0.99916
0.35	0.27366	1.35	0.82298	2.35	0.98123	3.35	0.99919
0.36	0.28115	1.36	0.82617	2.36	0.98172	3.36	0.99922
0.37	0.28862	1.37	0.82931	2.37	0.98221	3.37	0.99925
0.38	0.29605	1.38	0.83241	2.38	0.98269	3.38	0.99928
0.39	0.30346	1.39	0.83547	2.39	0.98315	3.39	0.99930
0.40	0.31084	1.40	0.83849	2.40	0.98360	3.40	0.99933
0.41	0.31819	1.41	0.84146	2.41	0.98405	3.41	0.99935
0.42	0.32552	1.42	0.84439	2.42	0.98448	3.42	0.99937
0.43	0.33280	1.43	0.84728	2.43	0.98490	3.43	0.99940
0.44	0.34006	1.44	0.85013	2.44	0.98531	3.44	0.99942
0.45	0.34729	1.45	0.85294	2.45	0.98571	3.45	0.99944
0.46	0.35448	1.46	0.85571	2.46	0.98611	3.46	0.99946
0.47	0.36164	1.47	0.85844	2.47	0.98649	3.47	0.99948
0.48	0.36877	1.48	0.86113	2.48	0.98686	3.48	0.99950
0.49	0.37587	1.49	0.86378	2.49	0.98723	3.49	0.99952
0.50	0.38292	1.50	0.86639	2.50	0.98758	3.50	0.99953
0.51	0.38995	1.51	0.86696	2.51	0.98793	3.51	0.99955
0.52	0.39694	1.52	0.87149	2.52	0.98826	3.52	0.99957
0.53	0.40389	1.53	0.87398	2.53	0.98859	3.53	0.99958
0.54	0.41080	1.54	0.87644	2.54	0.98891	3.54	0.99960
0.55	0.41768	1.55	0.87886	2.55	0.98923	3.55	0.99961
0.56	0.42452	1.56	0.88124	2.56	0.98953	3.56	0.99963
0.57	0.43132	1.57	0.88358	2.57	0.98983	3.57	0.99964
0.58	0.43809	1.58	0.88589	2.58	0.99012	3.58	0.99966
0.59	0.44481	1.59	0.88817	2.59	0.99040	3.59	0.99967
0.60	0.45149	1.60	0.89040	2.60	0.99068	3.60	0.99968
0.61	0.45814	1.61	0.89260	2.61	0.99095	3.61	0.99969
0.62	0.46474	1.62	0.89477	2.62	0.99121	3.62	0.99971
0.63	0.47131	1.63	0.89690	2.63	0.99146	3.63	0.99972

Продолжение таблицы 3.15

t	Φ(t)	t	Φ(t)	t	Φ(t)	t	Φ(t)
0.64	0.47783	1.64	0.89899	2.64	0.99171	3.64	0.99973
0.65	0.48431	1.65	0.90106	2.65	0.99195	3.65	0.99974
0.66	0.49075	1.66	0.90309	2.66	0.99219	3.66	0.99975
0.67	0.49714	1.67	0.90508	2.67	0.99241	3.67	0.99976
0.68	0.50350	1.68	0.90704	2.68	0.99263	3.68	0.99977
0.69	0.50981	1.69	0.90897	2.69	0.99285	3.69	0.99978
0.70	0.51607	1.70	0.91087	2.70	0.99307	3.70	0.99978
0.71	0.52230	1.71	0.91273	2.71	0.99327	3.71	0.99979
0.72	0.52848	1.72	0.91457	2.72	0.99347	3.72	0.99980
0.73	0.53461	1.73	0.91637	2.73	0.99367	3.73	0.99981
0.74	0.54070	1.74	0.91814	2.74	0.99386	3.74	0.99982
0.75	0.54675	1.75	0.91988	2.75	0.99404	3.75	0.99982
0.76	0.55275	1.76	0.92159	2.76	0.99422	3.76	0.99983
0.77	0.55870	1.77	0.92327	2.77	0.99439	3.77	0.99984
0.78	0.56461	1.78	0.92492	2.78	0.99456	3.78	0.99984
0.79	0.57047	1.79	0.92655	2.79	0.99473	3.79	0.99985
0.80	0.57629	1.80	0.92814	2.80	0.99489	3.80	0.99986
0.81	0.58206	1.81	0.92970	2.81	0.99505	3.81	0.99986
0.82	0.58778	1.82	0.93124	2.82	0.99520	3.82	0.99987
0.83	0.59346	1.83	0.93275	2.83	0.99535	3.83	0.99987
0.84	0.59909	1.84	0.93423	2.84	0.99549	3.84	0.99988
0.85	0.60468	1.85	0.93569	2.85	0.99563	3.85	0.99988
0.86	0.61021	1.86	0.93711	2.86	0.99576	3.86	0.99989
0.87	0.61570	1.87	0.93852	2.87	0.99590	3.87	0.99989
0.88	0.62114	1.88	0.93989	2.88	0.99602	3.88	0.99990
0.89	0.62653	1.89	0.94124	2.89	0.99615	3.89	0.99990
0.90	0.63188	1.90	0.94257	2.90	0.99627	3.90	0.99990
0.91	0.63718	1.91	0.94387	2.91	0.99639	3.91	0.99991
0.92	0.64243	1.92	0.94514	2.92	0.99650	3.92	0.99991
0.93	0.64763	1.93	0.94639	2.93	0.99661	3.93	0.99992
0.94	0.65278	1.94	0.94762	2.94	0.99672	3.94	0.99992
0.95	0.65789	1.95	0.94882	2.95	0.99682	3.95	0.99992
0.96	0.66294	1.96	0.95000	2.96	0.99692	3.96	0.99992
0.97	0.66795	1.97	0.95116	2.97	0.99702	3.97	0.99993
0.98	0.67291	1.98	0.95230	2.98	0.99712	3.98	0.99993
0.99	0.67783	1.99	0.95341	2.99	0.99721	3.99	0.99993

Таблица 3.16

Значения плотности вероятности $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$

t	0	1	2	3	4	5	6	7	8	9
0	0,398942	0,398922	0,398862	0,398763	0,398623	0,398444	0,398225	0,397966	0,397668	0,39733
0,1	0,396953	0,396536	0,39608	0,395585	0,395052	0,394479	0,393868	0,393219	0,392531	0,391806
0,2	0,391043	0,390242	0,389404	0,388529	0,387617	0,386668	0,385683	0,384663	0,383606	0,382515
0,3	0,381388	0,380226	0,379031	0,377801	0,376537	0,37524	0,373911	0,372548	0,371154	0,369728
0,4	0,36827	0,366782	0,365263	0,363714	0,362135	0,360527	0,35889	0,357225	0,355533	0,353812
0,5	0,352065	0,350292	0,348493	0,346668	0,344818	0,342944	0,341046	0,339124	0,33718	0,335213
0,6	0,333225	0,331215	0,329184	0,327133	0,325062	0,322972	0,320864	0,318737	0,316593	0,314432
0,7	0,312254	0,31006	0,307851	0,305627	0,303389	0,301137	0,298872	0,296595	0,294305	0,292004
0,8	0,289692	0,287369	0,285036	0,282694	0,280344	0,277985	0,275618	0,273244	0,270864	0,268477
0,9	0,266085	0,263688	0,261286	0,258881	0,256471	0,254059	0,251644	0,249228	0,246809	0,24439
1	0,241971	0,239551	0,237132	0,234714	0,232297	0,229882	0,22747	0,22506	0,222653	0,220251
1,1	0,217852	0,215458	0,213069	0,210686	0,208308	0,205936	0,203571	0,201214	0,198863	0,19652
1,2	0,194186	0,19186	0,189543	0,187235	0,184937	0,182649	0,180371	0,178104	0,175847	0,173602
1,3	0,171369	0,169147	0,166937	0,16474	0,162555	0,160383	0,158225	0,15608	0,153948	0,151831
1,4	0,149727	0,147639	0,145564	0,143505	0,14146	0,139431	0,137417	0,135418	0,133435	0,131468
1,5	0,129518	0,127583	0,125665	0,123763	0,121878	0,120009	0,118157	0,116323	0,114505	0,112704
1,6	0,110921	0,109155	0,107406	0,105675	0,103961	0,102265	0,100586	0,098925	0,097282	0,095657
1,7	0,094049	0,092459	0,090887	0,089333	0,087796	0,086277	0,084776	0,083293	0,081828	0,08038
1,8	0,07895	0,077538	0,076143	0,074766	0,073407	0,072065	0,07074	0,069433	0,068144	0,066871

Продовження табл. 3.16

t	0	1	2	3	4	5	6	7	8	9
1,9	0,065616	0,064378	0,063157	0,061952	0,060765	0,059595	0,058441	0,057304	0,056183	0,055079
2	0,053991	0,052919	0,051864	0,050824	0,0498	0,048792	0,0478	0,046823	0,045861	0,044915
2,1	0,043984	0,043067	0,042166	0,04128	0,040408	0,03955	0,038707	0,037878	0,037063	0,036262
2,2	0,035475	0,034701	0,033941	0,033194	0,03246	0,03174	0,031032	0,030337	0,029655	0,028985
2,3	0,028327	0,027682	0,027048	0,026426	0,025817	0,025218	0,024631	0,024056	0,023491	0,022937
2,4	0,022395	0,021862	0,021341	0,020829	0,020328	0,019837	0,019356	0,018885	0,018423	0,017971
2,5	0,017528	0,017095	0,01667	0,016254	0,015848	0,015449	0,01506	0,014678	0,014305	0,01394
2,6	0,013583	0,013234	0,012892	0,012558	0,012232	0,011912	0,0116	0,011295	0,010997	0,010706
2,7	0,010421	0,010143	0,009871	0,009606	0,009347	0,009094	0,008846	0,008605	0,00837	0,00814
2,8	0,007915	0,007697	0,007483	0,007274	0,007071	0,006873	0,006679	0,006491	0,006307	0,006127
2,9	0,005953	0,005782	0,005616	0,005454	0,005296	0,005143	0,004993	0,004847	0,004705	0,004567
3	0,004432	0,004301	0,004173	0,004049	0,003928	0,00381	0,003695	0,003584	0,003475	0,00337
3,1	0,003267	0,003167	0,00307	0,002975	0,002884	0,002794	0,002707	0,002623	0,002541	0,002461
3,2	0,002384	0,002309	0,002236	0,002165	0,002096	0,002029	0,001964	0,001901	0,00184	0,00178
3,3	0,001723	0,001667	0,001612	0,00156	0,001508	0,001459	0,001411	0,001364	0,001319	0,001275
3,4	0,001232	0,001191	0,001151	0,001112	0,001075	0,001038	0,001003	0,000969	0,000936	0,000904
3,5	0,000873	0,000843	0,000814	0,000785	0,000758	0,000732	0,000706	0,000681	0,000657	0,000634
3,6	0,000612	0,00059	0,000569	0,000549	0,000529	0,00051	0,000492	0,000474	0,000457	0,000441
3,7	0,000425	0,000409	0,000394	0,00038	0,000366	0,000353	0,00034	0,000327	0,000315	0,000303
3,8	0,000292	0,000281	0,000271	0,00026	0,000251	0,000241	0,000232	0,000223	0,000215	0,000207
3,9	0,000199	0,000191	0,000184	0,000177	0,00017	0,000163	0,000157	0,000151	0,000145	0,000139
4	0,000134	0,000129	0,000124	0,000119	0,000114	0,000109	0,000105	0,000101	0,000097	0,000093

В таблице 3.17 представлен расчет фактического значения критерия χ^2 Пирсона на основе эмпирических и теоретических частот [3]. Как видим, в таблице 3.14 эмпирические интервальные частоты m_i (1 и 6, 3 и 2) и рассчитанные по ним теоретические частоты $\tilde{m}_i$ сгруппированы, так как при использовании критерия χ^2 Пирсона целесообразно, чтобы эмпирическая и теоретическая частоты в каждой градации имели значение не менее 5.

Далее, определено значение $\dfrac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$, которое в сумме дает фактическое значение критерия χ^2 Пирсона.

Таблица 3.17

Расчет фактического значения критерия χ^2 Пирсона

i (k)	m_i	$\tilde{m}_i$	$\dfrac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$
1	1 6 } 7	1,7 4,6 } 6,3	0,08
2	9	8,2	0,08
3	9	9,4	0,02
4	6	7	0,14
5	3	3,3	
6	} 5	} 4,3	0,11
7	2	1	
χ^2			0,43

Теперь по специальной таблице 3.18 определяется теоретическое значение критерия χ^2 Пирсона. Для его нахождения используется 5% уровень значимости и степень свободы $v = k - 3$, где k - количество интервалов.

Далее фактическое и теоретическое значения критерия χ^2 Пирсона χ^2 Пирсона сравниваются между собой. Гипотеза о соответствии эмпирического распределения нормальному закону полностью подтверждается при условии $\chi^2 < \chi^2_{теор}$.

В приведенном примере (табл. 3.14) гипотеза о достоверности данных (о соответствии эмпирического распределения нормальному закону распределения) полностью подтверждается, поскольку при 5%-

ном уровне значимости и степени свободы ν = 5 – 3 = 2 теоретическое значение χ^2 равно 5,99, т.е. $\chi^2_{ф} < \chi^2_{теор.}$

Таблица 3.18

Значения χ^2 Пирсона для различных степеней свободы и уровней значимости

Степень свободы, $v = k - 3$	Уровни значимости		Степень свободы, $v = k - 3$	Уровни значимости	
	0,01 (1%)	0,05 (5%)		0,01 (1%)	0,05 (5%)
1	6.63	3.84	26	45.64	38.89
2	9.21	**5.99**	27	46.96	40.11
3	11.34	7.81	28	48.28	41.34
4	13.28	9.49	29	49.59	42.56
5	15.09	11.07	30	50.89	43.77
6	16.81	12.59	31	52.19	44.99
7	18.48	14.07	32	53.49	46.19
8	20.09	15.51	33	54.78	47.40
9	21.67	16.92	34	56.06	48.60
10	23.21	18.31	35	57.34	49.80
11	24.73	19.68	36	58.62	51.00
12	26.22	21.03	37	59.89	52.19
13	27.69	22.36	38	61.16	53.38
14	29.14	23.68	39	62.43	54.57
15	30.58	25.00	40	63.69	55.76
16	32.00	26.30	41	64.95	56.94
17	33.41	27.59	42	66.21	58.12
18	34.81	28.87	43	67.46	59.30
19	36.19	30.14	44	68.71	60.48
20	37.57	31.41	45	69.96	61.66
21	38.93	32.67	46	71.20	62.83
22	40.29	33.92	47	72.44	64.00
23	41.64	35.17	48	73.68	65.17
24	42.98	36.42	49	74.92	66.34
25	44.31	37.65	50	76.15	67.50

3.4. ПРЕОБРАЗОВАНИЕ ДАННЫХ В СЛУЧАЕ ИХ НЕСООТВЕТСТВИЯ НОРМАЛЬНОМУ ЗАКОНУ РАСПРЕДЕЛЕНИЯ

Любое отклонение данных от нормального закона распределения делает полученные результаты и выводы по ним недостоверными [4].

Если реальная кривая плотности вероятности (полигон) заметно отличается от кривой нормального закона распределения, если фактическое значение χ^2 Пирсона больше или равно табличному (теоретическому) значению: $\chi^2{}_{ф} \geq \chi^2{}_{т}$, то для «нормализации» кривой распределения (или полигона) можно осуществить преобразование исходных данных с помощью некоторой функции (например взять логарифм или возвести в квадрат). Как правило, в случае положительной асимметрии применяется логарифмирование, а в случае отрицательной – возведение в квадрат. Положительная асимметрия наблюдается тогда, когда значение моды не совпадает со значением средней арифметической, а лежит левее нее ($Mo < X_{ср}$). Отрицательная асимметрия указывает на обратную ситуацию ($Mo > X_{ср}$) [4].

Если же после преобразования повторилось несоответствие данных нормальному закону распределения, значит к полученным данным необходимо подбирать другой закон распределения (Пуассона, гамма, биноминальный и т.д.).

РАЗДЕЛ 4. ВЫЯВЛЕНИЕ ЗАВИСИМОСТИ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ ДРУГ ОТ ДРУГА

Выявить зависимость двух **однородных** и **достоверных** выборок друг от друга можно с помощью корреляционно-регрессионного анализа. Этот анализ позволяет определить тесноту и форму связи между ними. Теснота связи проверяется с помощью первой – «корреляционной» части этого анализа; форма связи (вид уравнения) определяется «регрессионной» частью.

Две выборки с одинаковым количеством значений, проверенные на однородность и достоверность, и обязательно соответствующие этим критериям, то есть, являющимися однородными и достоверными, выстраиваются в виде корреляционного ряда:

$$\begin{array}{lcccccc} X & 4 & 6 & 3 & 3 & 1 & 4 \\ У & 2 & 5 & 7 & 4 & 3 & 2 \end{array}$$

Тесноту связи между двумя выборками определяют с помощью коэффициента корреляции (r), который может иметь значение в пределах от -1 до +1:

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{n \cdot S_x \cdot S_y},$$

где n – количество парных значений; S_y, S_x - стандартное (среднее квадратическое) отклонение по выборкам $У$, X [3].

Ниже представлена вспомогательная таблица (табл. 4.1), используемая для расчета коэффициента корреляции.

Таблица 4.1

Вспомогательная таблица для расчета коэффициента корреляции

X_i	$X_i – X_{ср}$	$(X_i – X_{ср})^2$	$У_i$	$У_i – У_{ср}$	$(У_i – У_{ср})^2$	$(X_i – X_{ср})(У_i – У_{ср})$
…	…	…	…	…	…	…
$\sum X_i$	$\sum(X_i – X_{ср})$	$\sum(X_i – X_{ср})^2$	$\sum У$	$\sum(У_i – У_{ср})$	$\sum(У_i – У_{ср})^2$	$\sum(X_i – X_{ср})(Уi – У_{ср})$

Анализ полученного значения коэффициента корреляции r позволяет определить уровень тесноты связи двух выборочных совокупностей ($У$, X). Если коэффициент корреляции $r < 0,3$ – теснота связи слабая. При значении r в пределах от 0,3 до 0,7 – теснота связи средняя. Если $r > 0,7$, имеет значение теснота связи сильная. Значение коэффициента корреляции берется по модулю. На этом заканчивается первая - «корреляционная» часть этого анализа.

Для перехода ко второй - «регрессионной» части необходимо определить статистическую значимость коэффициента корреляции. Она проверяется с помощью критерия Стьюдента. Вначале рассчитывается его фактическое значение, которое затем сравнивается с табличным (теретическим). Для нахождения табличного (теоретического) значения этого критерия используется 5% уровень значимости и степень свободы (n-1), где n – количество парных ($У$, X) значений двух выборок (табл. 1.1).

Если $t_ф < t_т$ = (гипотеза H_0) - коэффициент корреляции статистически незначим, а уравнение регрессии имеет нелинейный вид;

если $t_ф \geq t_т$ = (гипотеза H_1) – коэффициент корреляции статистически значим, а уравнение регрессии имеет линейный вид.

Для расчета фактического значения критерия Стьюдента ($t_ф$) применяются формулы, учитывающие количество n парных ($У$, X) значений двух выборок.

При условии n < 50 фактическое значение критерия Стьюдента определяется по формуле [3]:

$$t_ф = t = \frac{|\hat{z}|}{\sigma_z},$$

где $\hat{z} = \frac{1}{2}\ln\frac{1+r_{xy}}{1-r_{xy}}$; $\sigma_z = \frac{1}{\sqrt{n-3}}$.

При условии n > 50 фактическое значение критерия Стьюдента определяется по формуле:

$$t_ф = t = \frac{|r_{xy}|}{\sigma_r}, \text{ где } \sigma_r = \frac{1-r_{xy}^2}{\sqrt{n-1}}.$$

Далее рассчитанное фактическое значение критерия Стьюдента сравнивается с табличным значением этого критерия ($t_т$), определенного по табл. 1.1 для 5%-ного уровня значимости и степени свободы n-1.

Далее, исходя из полученного результата, принимается либо нулевая либо альтернативная гипотеза:

$t_ф < t_т$ = (гипотеза H_0) – коэффициент корреляции статистически незначим, а уравнение регрессии имеет нелинейный вид;

$t_ф \geq t_т$ = (гипотеза H_1) – коэффициент корреляции статистически значим, а уравнение регрессии имеет линейный вид.

Если принимается гипотеза H_1 ($t_ф \geq t_т$), коэффициент корреляции признается статистически значимым, а уравнение регрессии приобретает вид:

$$\overline{y}(x) = ax + b,$$

где a, b – коэффициенты уравнения, определяемые по следующим формулам:

$$a = r_{xy}\frac{S_y}{S_x},$$

$$b = \overline{y} - a\overline{x},$$

где S_y, S_x - стандартные отклонения по выборкам $У$, X; r_{xy} - коэффициент корреляции.

После расчета коэффициентов a и b линейного уравнения проверяется их статистическая значимость. Проверка статистической значимости осуществляется с помощью критерия Стьюдента. Вначале рассчитывается его фактическое значение, которое затем сравнивается с табличным (теретическим). Для нахождения табличного значения этого критерия используется 5%-ный уровень значимости и степень свободы n-1, где n – количество парных ($У$, X) значений двух выборок (табл. 1.1).

Для расчета фактических значений критерия Стьюдента для коэффициентов a и b используются следующие формулы [3]:

$$t_ф = |a|/\sigma_a,$$

$$t_ф = |b|/\sigma_b,$$

где σ_a и σ_b рассчитываются с помощью следующих формул:

$$\sigma_a = \frac{S_y}{S_x\sqrt{n}}, \quad \sigma_b = \frac{S_y}{\sqrt{n}}\left[1+\frac{1}{C_{V_x}^2}\right]^{1/2}, \text{ где } C_{V_x} = \frac{S_x}{|\overline{x}|}.$$

Окончательное принятие уравнения регрессии линейного вида между двумя выборками (X, $У$) возможно только при таких условиях:

$t_ф$ (r_{xy}) $\geq t_т$ = (гипотеза H_1);

$t_ф$ (a) $\geq t_т$ = (гипотеза H_1);

$t_ф$ (b) $\geq t_т$ = (гипотеза H_1).

Если хоть одно из этих условий не выдерживается, то между выборками Х и У существует форма связи нелинейного вида.

Поиск регрессионных уравнений нелинейного вида является более сложным. Для упрощения можно воспользоваться программой Excel. В Excel предусмотрен простой способ выведения уравнений регрессии от одной переменной (рис. 4.1). Для этого необходимо [5]:

1. Занести значения X и $У$ в соседние столбцы, например, в А и В.
2. Нажать мастер диаграмм и выбрать тип графика «Точечная».
3. В следуещем окне построения графика сделать соответствующие подписи: общего названия, оси X и оси $У$ и определить густоту линий координатной сетки.
4. Выбрать положение графика в Excel и нажать «Готово».

5. Выделить график левой клавишей мыши. При этом вместо меню «Данные» появится меню «Диаграмма».
6. В меню «Диаграмма» выбрать команду «Добавить линию тренда». Тренд – это то же самое, что и уравнение регрессии. Появится окно «Линия тренда», в котором необходимо выбрать тип регрессии.
7. Нажать на клавишу «Параметры» и назначить: тип уравнения, показ уравнения на диаграмме и величину достоверности аппроксимации R^2 (коэффициента детерминации, характеризующего вероятность приближения реальной зависимости к регрессионной).
8. В результате Excel выведет уравнение регрессии и построит его график. Целесообразно построить несколько трендов и выбрать за основу прогнозирования такой из них, для которого значение коэффициента детерминации R^2 максимальное.
9. Для прогнозирования необходимо: а) выделить график; б) через меню «Диаграмма» попасть в окно «Линия тренда»; в) в окошке «Прогноз» этого окна назначить соответствующее значение аргумента и нажать «ОК».

Таким образом можно вывести любое из уравнений регрессии, предусмотренных окном «Линия тренда»: логарифмическое, полиноминальное, степенное, экспоненциальное (примечание: если среди значений встречаются отрицательные числа, то логарифмическое, степенное, и экспоненциальное уравнения регрессии выбирать нельзя).

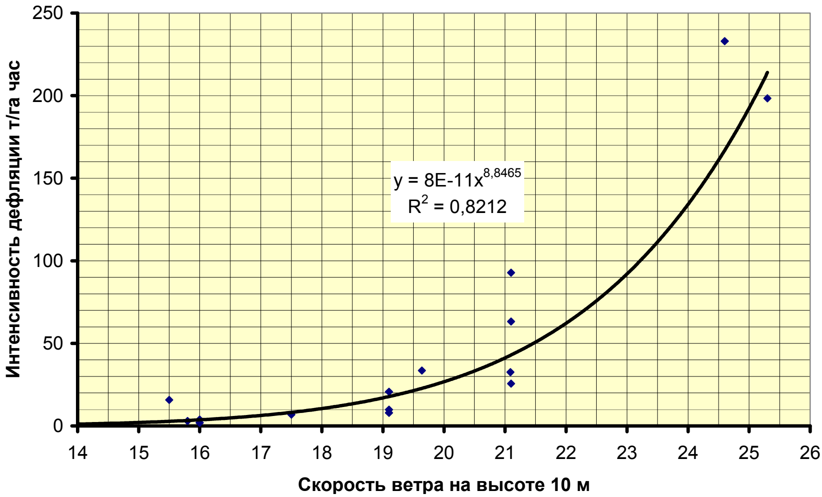


Рис. 4.1 Заключительный этап выведения уравнения регрессии от одной переменной с помощью Excel

Величина R^2 называется коэффициентом детерминации. Если из этого коэффициента извлечь корень, можно получить корреляционное отношение η_{xy} («эта»), характеризующее тесноту связи нелинейной зависимости двух выборок X и $У$. При линейной зависимости такую тесноту связи характеризует коэффициент корреляции r_{xy}.

Самым простым способом выявления зависимости двух проверенных на однородность и достоверность выборок друг от друга является применение Excel без предварительных расчетов «вручную» коэффициента корреляции и его статистической значимости. Для занесенных в соседние столбики значений X и $У$ в Excel строится несколько трендов и выбираются те уравнение и график, для которых значение R^2 (коэффициент детерминации) - максимален. Полученное уравнение характеризует форму связи двух выборок, а корень, извлеченный из R^2, дает значение коеффициента корреляции или корреляционного отношения (при криволинейной зависимости), и характеризует тесноту связи.

РАЗДЕЛ 5. СРАВНЕНИЕ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ МЕЖДУ СОБОЙ

Две выборочные совокупности можно сравнивать между собой с помощью параметрических и непараметрических критериев. Параметрические критерии используются в том случае, если обе выборки соответствуют нормальному закону распределения. К таким критериям относится критерий Фишера-Снедекора [3].

Если две выборочные совокупности не соответствуют или не проверялись на соответствие нормальному закону распределения, или хотя бы одна выборка не соответствует указанному закону, используются непараметрические критерии, например, критерий Вилкоксона [3].

В данном учебном пособии рассматривается только параметрический критерий Фишера-Снедекора, так как мы исходим из условия, что экспериментальные данные обязательно должны проверяться исследователем на однородность и достоверность и быть, соответственно, однородными и достоверными.

5.1. СРАВНЕНИЕ ДВУХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ МЕЖДУ СОБОЙ С ПОМОЩЬЮ КРИТЕРИЯ ФИШЕРА-СНЕДЕКОРА [3]

Расчет критерия Фишера-Снедекора состоит из двух частей. В первой части рассчитывается $F_{ф}$ путем сравнения дисперсий двух выборок S^2_{x}, S^2_{y}:

$$F_{ф} = \frac{S_x^2}{S_y^2}\ .$$

В числителе всегда располагается выборка с большей дисперсией по модулю.

Далее фактическое значение критерия Фишера ($F_{ф}$) сравнивается с $F_{т}$ (табличным). Для нахождения этого критерия (табл. 5.1) используется 5%-ный уровень значимости и степени свободы

$$k_1 = n - 1,$$
$$k_2 = m - 1,$$

где k_1 - число степеней свободы большей дисперсии, k_2 - число степеней свободы меньшей дисперсии

Таблица 5.1

Значения F-критерия Фишера
при уровне значимости $\alpha = 0,05$ [2]

k_2 \ k_1	1	2	3	4	5	6	8	12	24	∞
1	161,5	199,5	215,7	224,6	230,2	233,9	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65

Продолжение таблицы 5.1

k_2 \ k_1	1	2	3	4	5	6	8	12	24	∞
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
100	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1

Далее фактическое и теоретическое значение критерия Фишера сравниваются между собой:

если $F_ф \geq F_т$ имеет место гипотеза H'_1 – выборки существенно различаются, то есть они не однородны;

если $F_ф < F_т$ (гипотеза H'_0)- выборки существенно не различаются, то есть они однородны.

Однако окончательно эти гипотезы утверждаются после расчета второй части, т.е. критерия Стьюдента. При использовании этого критерия вначале рассчитывается его фактическое значение ($t_ф$):

$$t_ф = \frac{|\overline{x} - \overline{y}|}{\sqrt{\frac{S_x^2(n-1) + S_y^2(m-1)}{n+m-2}\left(\frac{1}{n} + \frac{1}{m}\right)}}$$

Теперь по таблице 1.1 определяется теоретическое значение данного критерия.

Для его нахождения используется 5%-ный уровень значимости и степень свободы $v = n + m - 2$, где n и m - объемы выборок.

Далее фактическое значение ($t_ф$) сравнивается с $t_т$ - табличным (теретическим) значением:

если $t_ф < t_т$ (гипотеза H''_0) - выборки существенно не различаются, то есть они однородны;

если $t_ф \geq t_т$ (гипотеза H''_1) – выборки существенно различаются, то есть они не однородны.

Только при условии, что в двух частях расчета будет принята гипотеза H_0 ($F_ф < F_т = H'_0$ и $t_ф < t_т = H''_0$), две выборочные совокупности не будут существенно различаться, то есть будут однородны.

Ниже (в табл. 5.2) приведен пример заполнения заключительной таблицы по проверке данных на однородность с помощью критерия Фишера-Снедекора.

Таблица 5.2

Оценка существенности различий

Сравниваемые варианты	Критерий Фишера (F)	Критерий Стьюдента (t)	Сущест-венность различий
B_1 – контроль	1,29>1 ($H_1^{'}$)	7,86>1,96($H_1^{''}$)	существ.
B_2 – контроль	1,76>1,53 ($H_1^{'}$)	7,68>1,96($H_1^{''}$)	существ.
B_3 – контроль	1,23>1,22 ($H_1^{'}$)	22,95>1,96($H_1^{''}$)	существ.

5.2. СРАВНЕНИЕ ТРЕХ (И БОЛЕЕ) ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ МЕЖДУ СОБОЙ С ПОМОЩЬЮ КРИТЕРИЯ КРУСКАЛА-УОЛЛИСА

Для сравнения трех и более выборок между собой используется критерий Крускала-Уоллиса [3]. При использовании этого критерия вначале рассчитывается его фактическое значение:

$$Q = \frac{12}{N(N+1)} \sum_{\gamma=1}^{a} \frac{1}{n_\gamma} \left[\sum_{i \in S_\gamma} R_i \right]^2 - 3(N+1).$$

Для нахождения табличного значения этого критерия (по значению χ^2 Пирсона) используется 5% уровень значимости и степень свободы (n-1), где n – количество сравниваемых выборок.

Далее фактическое значение сравнивается с табличным (теретическим) значением:

$Q < \chi^2_{т}$. (гипотеза H_0) - выборки существенно не различаются, то есть они однородны;

$Q \geq \chi^2_{т}$. (гипотеза H_1) - выборки существенно различаются, то есть они не однородны.

Ниже (табл. 5.3, 5.4) приведен пример проверки трех выборочных совокупностей на однородность с помощью критерия Крускала-Уоллиса.

Таблица 5.3

Исходные данные трех, сравниваемых между собой, выборочных совокупностей

№ п/п	Выборочные совокупности		
	X	Y	Z
1	16	17	16
2	17	30	17
3	18	37	19
4	20	39	32
5	21	40	33
6	28	44,5	36
7	30	30	36
8	35	23	38
9	40	36	40
10	40	38	43
11	40	40	44
12	42	24	45
13	52	30	45
14	55	29	46
15	60,4	43	49
16	67	47	60
17	70	42	70
18	70	38	85
19	71,4	22	85
20	73	36	59
21	98	31	9
22	98	33	34
23	100,8	35	36

В табл. 5.4 приведена полная объединенная выборка, на основе которой для каждой величины определяется ранг, равный порядковому номеру. Если одинаковые значения принадлежат разным выборкам, то им присваивается один и тот же средний ранг.

Таблица 5.4

Объединенный ранжированный ряд данных

Характеристика объединенного ряда											
N	R_i	$i\in S_\gamma$	N	R_i	$i\in S_\gamma$	N	R_i	$i\in S_\gamma$	N	R_i	$i\in S_\gamma$
1	1	9z	18	17,5	30y	35	34	38y	52	52	47y
2	2	16x	19	17,5	30y	36	36	39y	53	53	49z
3	3	16z	20	20	31y	37	39,5	40x	54	54	52x
4	5	17x	21	21	32z	38	39,5	40x	55	55	55x
5	5	17z	22	22,5	33z	39	39,5	40x	56	56	59z
6	5	17y	23	22,5	33y	40	39,5	40z	57	57	60z
7	7	18x	24	24	34z	41	39,5	40y	58	58	60,4x
8	8	19z	25	25,5	35x	42	39,5	40y	59	59	67x
9	9	20x	26	25,5	35y	43	43,5	42x	60	61	70x
10	10	21x	27	29	36z	44	43,5	42y	61	61	70x
11	11	22y	28	29	36z	45	45,5	43z	62	61	70z
12	12	23y	29	29	36z	46	45,5	43y	63	63	71,4x
13	13	24y	30	29	36y	47	47	44z	64	64	73x
14	14	28x	31	29	36y	48	48	44,5y	65	65,5	85z
15	15	29y	32	32	37y	49	49,5	45z	66	65,5	85z
16	17,5	30x	33	34	38z	50	49,5	45z	67	67,5	98x
17	17,5	30y	34	34	38y	51	51	46z	68	67,5	98x

Далее определяется сумма рангов для каждой выборки:

$$\frac{1}{n_x}\left[\sum_{i\in S_x} R_i\right]^2 = \frac{1}{23}[845,5]^2 = 31081,3$$

$$\frac{1}{n_y}\left[\sum_{i\in S_y} R_i\right]^2 = \frac{1}{23}[931,0]^2 = 37685,3$$

$$\frac{1}{n_z}\left[\sum_{i\in S_z} R_i\right]^2 = \frac{1}{23}[638,5]^2 = 17725,3$$

На основании суммы рангов рассчитывается фактическое значение критерия Крускала-Уоллиса:

$$Q = \frac{12}{69\cdot 70}\cdot 86491,9 - 3\cdot 70 = 4,88$$

Согласно табличным данным, на уровне значимости α = 0,05 и степени свободы $\nu = 3 - 1 = 2$, $\chi^2_{\text{т}}(\alpha, \nu) = 5,99$ (табл. 3.15). Таким образом, $Q < \chi^2_{\text{т}}$ и гипотеза об однородности трех выборочных совокупностей подтверждается.

СПИСОК ЛИТЕРАТУРЫ

1. Доспехов Б. А. Методика полевого опыта (с основами статистической обработки результатов исследований). Изд. 4-е, перераб. и доп. – М.: Колос, 1979. – 416 с.
2. Гмурман В.Е. Руководство к решению задач по теории вероятности и математической статистике: Учеб. пособие – 12-е изд. Перераб. – М.: Высшее образование, 2006г. – 476с.
3. Школьний Е. П., Гончарова Л. Д., Миротворська Н. К. Методи обробки та аналізу гідрометеорологічної інформації (збірник задач) / Навчальн. посібник. – К.: Міносвіти і науки України, 2000. – 419 с.
4. Самнер Г. Математика для географов – М.: Прогресс, 1981.– 296 с.
5. Куценко М.В. Вступ до географічних інформаційних систем та моделювання стану довкілля: Навч. Посібник.– Харків: Екограф, 2008. – 204 с.

Н а у к о в е   в и д а н н я

ЗУБОВА Лілія Григорівна
доктор технічних наук, професор

ОСНОВИ МАТЕМАТИЧНОЇ ОБРОБКИ ЕКСПЕРИМЕНТАЛЬНЫХ ДАНИХ

(російською мовою)

В авторській редакції

Підписано до друку 01.10.2013
Формат $60\times84^{1}/_{16}$ Папір типограф. Гарнітура Times.
Друк офсетний. Умов. друк. арк. 3,49. Обл.-вид. арк. 4,58.
Тираж 50 прим. Вид. № 1025. Замов № 1091. Ціна договірна.

Видавництво «НОУЛІДЖ»
Свідоцтво про реєстрацію серія ДК №2884 від 26.06.2007 р.
91051, м. Луганськ, кв. Якіра, 3/316,
Тел. (050) 475-35-13, e-mail: nickvnu@gmail.com