

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Локтионова Оксана Геннадьевна

Должность: проректор по учебной работе

Дата подписания: 16.09.2023 19:32:55

Уникальный программный ключ:

0b817ca911e6668abb13a5d436f39e51a11a11c577e947d6a4951ff4564088

1

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ

Проректор по учебной работе

« 14 » _____ О.Г. Локтионова, _____ г.



ПРОВЕРКА НОРМАЛЬНОСТИ РАСПРЕДЕЛЕНИЯ СОГЛАСНО КРИТЕРИЮ ПИРСОНА

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Проверка нормальности распределения согласно критерию Пирсона : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2017. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1. Цель работы: Получить практические и теоретические навыки по проверки нормальности распределения результатов измерения при большом числе измерений $n > 50$.

2. Задание: Произвести более 50 измерений заданной физической величины. Проверить гипотезу о нормальности распределения полученных результатов наблюдений согласно критерию Пирсона (критерию χ^2).

3. Краткие теоретические положение

Проверка нормальности распределения согласно критерию χ^2 сводится к следующему.

1. *Данные наблюдений группируют по интервалам.*

Весь диапазон полученных результатов наблюдений $X_{max} - X_{min}$ разделяют на r интервалов шириной $\Delta X_i = (X_{max} - X_{min})/r$, ($i=1, 2, \dots, r$) и подсчитывают частоты, m_i , равные числу результатов, лежащих в каждом i -м интервале, т. е. меньших или равных его правой и больших левой границы. Число r интервалов выбирается в зависимости от числа наблюдений согласно следующим рекомендациям табл.1:

Таблица 1

n	r
40—100	7—9
100—500	8—12
500—1000	10—16
1000—10000	12—22

Отношения

$$P_i^* = \frac{m_i}{n} \quad (1)$$

где n – общее число наблюдений, называются **частотами** и представляют собой статистические оценки вероятностей попадания результата наблюдений в i -й интервал. Распределение частостей по интервалам образует статистическое распределение результатов наблюдений.

Если разделить частоту на длину интервала, то получим величины

$$p_i^* = \frac{1}{\Delta X_i} P_i^* = \frac{m_i}{n \Delta X_i} \quad (2)$$

являющиеся оценками средней плотности распределения в интервале ΔX_i .

Если в некоторые интервалы попадает меньше пяти наблюдений, то такие интервалы объединяют с соседними. При этом число степеней свободы k уменьшается на единицу.

2. Вычисляют среднее арифметическое \bar{X}

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

и точечную оценку среднего квадратического отклонения результата наблюдений s_X

$$s_X = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2},$$

которые принимают в качестве параметров теоретического нормального распределения с плотностью $p_X(X)$.

3. Для каждого интервала находят вероятности попадания в

них результатов наблюдений приближенно как произведение плотности теоретического распределения в середине интервала на его длину:

$$P_i \approx p_X \left(\frac{X_i + X_{i+1}}{2} \right) \Delta X_i.$$

4. Для каждого интервала вычисляют величину меры расхождения χ_i^2 ($i = 1, 2, \dots, r$)

$$\chi_i^2 = \frac{(m_i - nP_i)^2}{nP_i}$$

и суммируют их по всем i , в результате чего получают меру расхождения χ_k^2

$$\chi_k^2 = \sum_{i=1}^r \frac{n}{P_i} (P_i^* - P_i)^2 = \sum_{i=1}^r \frac{(m_i - nP_i)^2}{nP_i} = \sum_{i=1}^r \chi_i^2.$$

5. Определяют число степеней свободы $k=r-3$ и, задаваясь уровнем значимости $q=1-\alpha$, находят по табл. 5 приложения значения $\chi_{k; \frac{1}{2}q}^2$

и $\chi_{k; 1-\frac{1}{2}q}^2$. Если $\chi_{k; \frac{1}{2}q}^2 < \chi_k^2 \leq \chi_{k; 1-\frac{1}{2}q}^2$, то распределение результатов наблюдений считают нормальным.

4. Порядок выполнения работы

1. На основании исходных данных, в соответствии со своим вариантом (табл. 6 приложения), заполнить табл. 2 приложения.

2. Рассчитать параметры данного распределения, *среднее ариф-*

метическое \bar{X} и точечную оценку среднего квадратического отклонения результата наблюдений s_x .

3. Заполнить таблицу 3 приложения расчетными значениями. Плотность нормированного распределения $p(t_i)$ выбрать из табл. 4 приложения по рассчитанным значениям нормированное отклонение от среднего арифметического t_i .

4. Рассчитать меру расхождения χ_k^2

5. Задаваясь уровнем значимости $q=0,1$ и рассчитанной величиной степени свободы k найти из табл. 5 приложения значения $\chi_{k; \frac{1}{2}q}^2$ и

$$\chi_{k; 1-\frac{1}{2}q}^2.$$

6. Сделать вывод о нормальности распределения данных результатов наблюдений.

Список литературы

1. Бурдун Г.Д., Макаров Б.Н. Основы метрологии. Учебное пособие для вузов. Издание третье, переработанное – М.: Изд-во стандартов, 1985, 256 с. ил.
2. ГОСТ 8.207 «ГСИ. Прямые измерения с многократными наблюдениями. Методы обработки результатов наблюдений.»

ПРИЛОЖЕНИЯ

Таблица 2.

Результаты измерений

Номер измерения	Результат измерения
1	
·	
·	
n	

Таблица 3.

Результаты расчетов

Номер интервала i	Средина интервала X_i	Частота m_i	Отклонение от среднего $X_i - \bar{X}$	Нормированное отклонение от среднего арифметического $t_i = \frac{X_i - \bar{X}}{S_X}$	Плотность нормированного распределения $p(t_i)$	Плотность в серединах интервалов $p(X_i) = \frac{p(t_i)}{S_X}$	Теоретическая частота $nP_i = n\Delta X_i \cdot p(X_i)$	Отклонение χ^2_i
1								
·								
·								
r								

$n =$ $\Delta X =$ $\bar{X} =$ $s_X =$

$k =$ $\chi_k^2 =$

Дифференциальная функция нормированного нормального распределения

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	t
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973	0,0
0,1	3970	3965	3861	3956	3951	3945	3939	3932	3925	3918	0,1
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825	0,2
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697	0,3
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538	0,4
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352	0,5
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144	0,6
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920	0,7
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685	0,8
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444	0,9
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203	1,0
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965	1,1
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736	1,2
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518	1,3
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315	1,4
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127	1,5
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957	1,6
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804	1,7
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669	1,8
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551	1,9
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449	2,0
2,1	0440	0431	0422	0413	0404	0396	0388	0379	0371	0363	2,1
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290	2,2
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229	2,3
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180	2,4
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139	2,5
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107	2,6
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081	2,7
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061	2,8
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046	2,9
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034	3,0
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025	3,1
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018	3,2
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013	3,3
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009	3,4
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006	3,5
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004	3,6
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003	3,7
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002	3,8
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001	3,9

Таблица 5

Интегральная функция χ^2 - распределения Пирсона.Значения $\chi^2_{k;P}$ для различных значений k и P

k	P												
	0,01	0,02	0,05	0,1	0,2	0,3	0,5	0,7	0,8	0,9	0,95	0,98	0,99
1	0,000157	0,000628	0,00393	0,0158	0,0642	0,148	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,02	0,04	0,103	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	7,824	9,21
3	0,115	0,185	0,352	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,345
4	0,297	0,429	0,711	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	0,554	0,752	1,145	1,61	2,343	3	4,351	6,064	7,289	9,236	11,07	13,388	15,086
6	0,872	1,134	1,635	2,204	3,07	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	1,239	1,564	2,167	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,622	18,475
8	1,647	2,032	2,733	3,49	4,594	5,527	7,344	9,524	11,03	13,362	15,507	18,168	20,09
9	2,088	2,532	3,325	4,168	5,38	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	2,558	3,059	3,94	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	3,053	3,609	4,575	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	3,571	4,178	5,226	6,304	7,807	9,034	11,34	14,011	15,812	18,549	21,026	24,054	26,217
13	4,107	4,765	5,892	7,041	8,634	9,926	12,34	15,119	16,985	19,812	22,362	25,471	27,688
14	4,66	5,368	6,571	7,79	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141
15	5,229	5,985	7,261	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	5,812	6,614	7,962	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	29,633	32
17	6,408	7,255	8,672	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	7,015	7,906	9,39	10,865	12,857	14,44	17,338	20,601	22,76	25,989	28,869	32,346	34,805
19	7,633	8,567	10,117	11,651	13,716	15,352	18,338	21,689	23,9	27,204	30,144	33,687	36,191
20	8,26	9,237	10,851	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,41	35,02	37,566
21	8,897	9,915	11,591	13,24	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	9,542	10,6	12,338	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,659	40,289
23	10,196	11,293	13,091	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	10,856	11,992	13,848	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,27	42,98
25	11,524	12,697	14,611	16,473	18,94	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	12,198	13,409	15,379	17,292	19,82	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	12,878	14,125	16,151	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	44,14	46,963
29	14,256	15,574	17,708	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	49,588

k	P												
	0,01	0,02	0,05	0,1	0,2	0,3	0,5	0,7	0,8	0,9	0,95	0,98	0,99
30	14,953	16,306	18,493	20,599	23,364	25,508	29,336	33,53	36,25	40,256	43,773	47,962	50,892
31	15,655	17,042	19,281	21,434	24,255	26,44	30,336	34,598	37,359	41,422	44,985	49,226	52,191
32	16,362	17,783	20,072	22,271	25,148	27,373	31,336	35,665	38,466	42,585	46,194	50,487	53,486
33	17,073	18,527	20,867	23,11	26,042	28,307	32,336	36,731	39,572	43,745	47,4	51,743	54,775
34	17,789	19,275	21,664	23,952	26,938	29,242	33,336	37,795	40,676	44,903	48,602	52,995	56,061
35	18,509	20,027	22,465	24,797	27,836	30,178	34,336	38,859	41,778	46,059	49,802	54,244	57,342
36	19,233	20,783	23,269	25,643	28,735	31,115	35,336	39,922	42,879	47,212	50,998	55,489	58,619
37	19,96	21,542	24,075	26,492	29,635	32,053	36,336	40,984	43,978	48,363	52,192	56,73	59,893
38	20,691	22,304	24,884	27,343	30,537	32,992	37,335	42,045	45,076	49,513	53,384	57,969	61,162
39	21,426	23,069	25,695	28,196	31,441	33,932	38,335	43,105	46,173	50,66	54,572	59,204	62,428
40	22,164	23,838	26,509	29,051	32,345	34,872	39,335	44,165	47,269	51,805	55,758	60,436	63,691
41	22,906	24,609	27,326	29,907	33,251	35,813	40,335	45,224	48,363	52,949	56,942	61,665	64,95
42	23,65	25,383	28,144	30,765	34,157	36,755	41,335	46,282	49,456	54,09	58,124	62,892	66,206
43	24,398	26,159	28,965	31,625	35,065	37,698	42,335	47,339	50,548	55,23	59,304	64,116	67,459
44	25,148	26,939	29,787	32,487	35,974	38,641	43,335	48,396	51,639	56,369	60,481	65,337	68,71
45	25,901	27,72	30,612	33,35	36,884	39,585	44,335	49,452	52,729	57,505	61,656	66,555	69,957
46	26,657	28,504	31,439	34,215	37,795	40,529	45,335	50,507	53,818	58,641	62,83	67,771	71,201
47	27,416	29,291	32,268	35,081	38,708	41,474	46,335	51,562	54,906	59,774	64,001	68,985	72,443
48	28,177	30,08	33,098	35,949	39,621	42,42	47,335	52,616	55,993	60,907	65,171	70,197	73,683
49	28,941	30,871	33,93	36,818	40,534	43,366	48,335	53,67	57,079	62,038	66,339	71,406	74,919
50	29,707	31,664	34,764	37,689	41,449	44,313	49,335	54,723	58,164	63,167	67,505	72,613	76,154
51	30,475	32,459	35,6	38,56	42,365	45,261	50,335	55,775	59,248	64,295	68,669	73,818	77,386
52	31,246	33,256	36,437	39,433	43,281	46,209	51,335	56,827	60,332	65,422	69,832	75,021	78,616
53	32,019	34,055	37,276	40,308	44,199	47,157	52,335	57,879	61,414	66,548	70,993	76,223	79,843
54	32,793	34,856	38,116	41,183	45,117	48,106	53,335	58,93	62,496	67,673	72,153	77,422	81,069
55	33,571	35,659	38,958	42,06	46,036	49,055	54,335	59,98	63,577	68,796	73,311	78,619	82,292
56	34,35	36,464	39,801	42,937	46,955	50,005	55,335	61,031	64,658	69,919	74,468	79,815	83,514
57	35,131	37,27	40,646	43,816	47,876	50,956	56,335	62,08	65,737	71,04	75,624	81,009	84,733
58	35,914	38,078	41,492	44,696	48,797	51,906	57,335	63,129	66,816	72,16	76,778	82,201	85,95
59	36,698	38,888	42,339	45,577	49,718	52,858	58,335	64,178	67,894	73,279	77,93	83,391	87,166
60	37,485	39,699	43,188	46,459	50,641	53,809	59,335	65,226	68,972	74,397	79,082	84,58	88,379

Таблица 6

№	№ варианта									
	1	2	3	4	5	6	7	8	9	10
1	201,576	86,114	19,046	158,772	80,915	181,724	135,073	62,720	43,996	40,701
2	201,561	86,127	19,000	158,755	80,917	181,722	135,086	62,721	43,975	40,690
3	201,558	86,117	19,055	158,807	80,956	181,725	135,082	62,707	44,003	40,734
4	201,570	86,126	19,001	158,805	80,956	181,713	135,073	62,731	43,990	40,729
5	201,562	86,134	19,015	158,779	80,948	181,718	135,080	62,711	43,998	40,703
6	201,555	86,079	19,033	158,802	80,949	181,724	135,088	62,737	43,996	40,691
7	201,588	86,171	18,979	158,792	80,944	181,718	135,083	62,701	43,971	40,714
8	201,566	86,106	19,037	158,791	80,948	181,721	135,071	62,735	43,994	40,697
9	201,544	86,064	19,050	158,806	80,974	181,724	135,085	62,712	43,992	40,739
10	201,570	86,127	19,016	158,789	80,941	181,726	135,082	62,723	43,993	40,742
11	201,559	86,121	19,080	158,829	80,943	181,723	135,091	62,699	43,989	40,723
12	201,571	86,123	19,001	158,789	80,934	181,725	135,091	62,715	43,999	40,746
13	201,557	86,124	19,072	158,756	80,934	181,734	135,082	62,705	43,999	40,721
14	201,561	86,119	19,024	158,777	80,931	181,726	135,088	62,731	43,991	40,718
15	201,569	86,140	19,042	158,768	80,926	181,723	135,087	62,734	43,983	40,720
16	201,564	86,138	19,018	158,791	80,949	181,726	135,085	62,721	43,976	40,726
17	201,570	86,132	19,012	158,810	80,958	181,721	135,076	62,715	43,995	40,749
18	201,564	86,099	19,013	158,802	80,968	181,724	135,089	62,694	43,985	40,664
19	201,578	86,132	19,002	158,782	80,954	181,720	135,083	62,726	43,974	40,746
20	201,577	86,114	19,040	158,821	80,952	181,718	135,085	62,702	44,004	40,712
21	201,582	86,131	19,034	158,773	80,933	181,721	135,088	62,729	43,984	40,742
22	201,570	86,105	18,996	158,761	80,916	181,722	135,083	62,725	43,988	40,733
23	201,564	86,124	19,027	158,807	80,968	181,720	135,083	62,715	44,007	40,708
24	201,603	86,115	19,036	158,785	80,955	181,721	135,080	62,715	43,994	40,736
25	201,556	86,120	19,033	158,815	80,966	181,721	135,075	62,749	43,987	40,706
26	201,560	86,106	19,040	158,771	80,926	181,722	135,084	62,735	44,001	40,735
27	201,575	86,145	19,035	158,793	80,955	181,716	135,077	62,705	43,993	40,692
28	201,546	86,112	19,033	158,782	80,923	181,726	135,089	62,691	43,979	40,707
29	201,578	86,140	19,017	158,767	80,950	181,730	135,073	62,747	43,989	40,698
30	201,546	86,073	19,054	158,784	80,936	181,720	135,084	62,737	43,980	40,706
31	201,567	86,116	19,008	158,798	80,956	181,724	135,089	62,710	44,007	40,681
32	201,574	86,130	19,007	158,825	80,943	181,720	135,079	62,718	43,992	40,736
33	201,580	86,135	18,980	158,796	80,935	181,720	135,089	62,735	43,997	40,751
34	201,568	86,115	19,015	158,775	80,916	181,718	135,074	62,722	43,990	40,713
35	201,563	86,121	19,023	158,785	80,924	181,726	135,078	62,730	43,982	40,711
36	201,561	86,100	19,032	158,796	80,963	181,716	135,076	62,752	43,981	40,683
37	201,586	86,127	19,041	158,740	80,923	181,733	135,076	62,728	43,987	40,719
38	201,562	86,132	19,030	158,784	80,929	181,725	135,085	62,712	44,004	40,772
39	201,590	86,124	19,029	158,784	80,923	181,726	135,079	62,730	43,985	40,733

№ варианта										
№	1	2	3	4	5	6	7	8	9	10
40	201,540	86,095	19,014	158,798	80,945	181,718	135,075	62,712	43,981	40,725
41	201,565	86,118	19,054	158,800	80,947	181,722	135,091	62,716	44,008	40,729
42	201,557	86,098	19,035	158,788	80,958	181,719	135,085	62,714	44,015	40,678
43	201,565	86,130	19,037	158,808	80,934	181,714	135,081	62,703	44,004	40,731
44	201,580	86,119	18,987	158,768	80,951	181,722	135,069	62,766	43,964	40,732
45	201,554	86,128	19,032	158,803	80,954	181,719	135,089	62,714	44,010	40,702
46	201,574	86,142	18,997	158,792	80,948	181,725	135,088	62,721	44,007	40,720
47	201,562	86,133	19,045	158,766	80,923	181,727	135,084	62,724	43,997	40,704
48	201,557	86,112	19,028	158,810	80,978	181,719	135,073	62,712	43,985	40,719
49	201,561	86,094	19,007	158,819	80,982	181,711	135,080	62,731	43,991	40,745
50	201,551	86,100	19,010	158,784	80,943	181,722	135,080	62,733	44,002	40,715

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
1	139,752	125,290	118,606	56,888	189,731	119,778	125,777	189,498	73,219	42,442
2	139,751	125,266	118,602	56,931	189,690	119,714	125,788	189,513	73,237	42,435
3	139,749	125,234	118,612	56,881	189,620	119,730	125,803	189,511	73,252	42,431
4	139,749	125,261	118,611	56,901	189,673	119,683	125,799	189,554	73,218	42,439
5	139,751	125,258	118,602	56,894	189,648	119,703	125,803	189,461	73,251	42,451
6	139,750	125,251	118,609	56,950	189,630	119,730	125,718	189,545	73,207	42,439
7	139,751	125,264	118,608	56,893	189,684	119,714	125,785	189,539	73,214	42,443
8	139,751	125,302	118,607	56,902	189,645	119,694	125,823	189,503	73,205	42,449
9	139,753	125,291	118,599	56,907	189,639	119,732	125,751	189,539	73,207	42,451
10	139,752	125,280	118,598	56,886	189,626	119,741	125,785	189,529	73,246	42,437
11	139,750	125,261	118,604	56,904	189,575	119,651	125,743	189,481	73,205	42,444
12	139,750	125,278	118,608	56,903	189,694	119,712	125,773	189,592	73,236	42,449
13	139,753	125,297	118,599	56,898	189,612	119,696	125,733	189,488	73,236	42,453
14	139,751	125,297	118,606	56,944	189,569	119,704	125,774	189,552	73,227	42,449
15	139,750	125,268	118,598	56,932	189,660	119,708	125,745	189,606	73,239	42,434
16	139,749	125,291	118,608	56,934	189,665	119,756	125,781	189,550	73,208	42,438
17	139,748	125,274	118,602	56,929	189,656	119,701	125,778	189,540	73,206	42,425
18	139,748	125,235	118,602	56,945	189,594	119,659	125,730	189,541	73,208	42,447
19	139,754	125,260	118,609	56,933	189,674	119,679	125,801	189,550	73,267	42,441
20	139,749	125,260	118,602	56,907	189,628	119,731	125,745	189,568	73,236	42,444
21	139,747	125,275	118,603	56,911	189,670	119,734	125,789	189,522	73,261	42,450
22	139,750	125,276	118,606	56,889	189,690	119,688	125,764	189,510	73,187	42,451
23	139,750	125,270	118,603	56,942	189,638	119,682	125,748	189,548	73,221	42,425
24	139,749	125,270	118,598	56,939	189,600	119,738	125,760	189,488	73,216	42,446
25	139,749	125,253	118,601	56,933	189,618	119,712	125,742	189,581	73,250	42,446

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
26	139,745	125,260	118,605	56,910	189,687	119,691	125,785	189,493	73,224	42,428
27	139,751	125,267	118,611	56,910	189,665	119,706	125,759	189,598	73,257	42,439
28	139,748	125,236	118,605	56,914	189,620	119,746	125,749	189,538	73,209	42,439
29	139,751	125,291	118,600	56,943	189,634	119,723	125,738	189,546	73,247	42,445
30	139,750	125,250	118,607	56,899	189,582	119,740	125,769	189,555	73,223	42,432
31	139,750	125,277	118,598	56,934	189,589	119,700	125,804	189,533	73,234	42,435
32	139,748	125,312	118,598	56,881	189,664	119,672	125,817	189,548	73,237	42,429
33	139,752	125,264	118,600	56,883	189,592	119,708	125,819	189,519	73,226	42,442
34	139,751	125,267	118,599	56,934	189,725	119,677	125,735	189,502	73,227	42,441
35	139,749	125,282	118,601	56,887	189,590	119,732	125,703	189,523	73,250	42,448
36	139,751	125,292	118,599	56,920	189,679	119,759	125,736	189,577	73,209	42,427
37	139,747	125,256	118,610	56,876	189,657	119,688	125,759	189,489	73,239	42,445
38	139,749	125,280	118,609	56,896	189,637	119,717	125,763	189,510	73,236	42,424
39	139,750	125,243	118,603	56,888	189,599	119,718	125,755	189,526	73,250	42,432
40	139,748	125,267	118,599	56,928	189,633	119,700	125,785	189,522	73,198	42,449
41	139,751	125,260	118,608	56,867	189,660	119,703	125,784	189,506	73,194	42,440
42	139,748	125,255	118,606	56,932	189,587	119,734	125,732	189,604	73,244	42,446
43	139,752	125,299	118,593	56,909	189,602	119,747	125,759	189,499	73,250	42,429
44	139,751	125,276	118,604	56,947	189,671	119,686	125,835	189,562	73,234	42,441
45	139,752	125,247	118,613	56,940	189,615	119,697	125,743	189,567	73,221	42,450
46	139,748	125,278	118,607	56,933	189,655	119,746	125,785	189,515	73,216	42,423
47	139,749	125,281	118,596	56,941	189,592	119,734	125,759	189,580	73,207	42,448
48	139,750	125,255	118,610	56,900	189,614	119,720	125,787	189,494	73,214	42,441
49	139,749	125,287	118,605	56,923	189,660	119,693	125,707	189,547	73,257	42,432
50	139,747	125,310	118,605	56,937	189,668	119,710	125,817	189,512	73,249	42,440

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Ложионова
« 14 » _____ г.



РАСЧЕТ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ ДЛЯ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ И СРЕДНЕГО КВАДРАТИЧЕСКОГО ОТКЛОНЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ С НОРМАЛЬНЫМ РАСПРЕДЕЛЕНИЕМ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Расчет доверительных интервалов для математического ожидания и среднего квадратического отклонения случайной величины с нормальным распределением : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: Научиться рассчитывать доверительные интервалы для математического ожидания и среднего квадратического отклонения случайной величины с нормальным распределением.

2 Задание: Рассчитывать доверительные интервалы для математического ожидания и среднего квадратического отклонения случайной величины в соответствии со своим вариантом (прил. В).

3 Краткие теоретические сведения

Как известно, выборочное среднее и выборочная дисперсия являются случайными величинами, оценками математического ожидания и дисперсии, причем их совпадение с теоретическими характеристиками имеет нулевую вероятность.

Иногда бывает удобно указать интервал, внутри которого недоступная для непосредственного измерения характеристика попадает с достаточно большой, близкой к единице, вероятностью. Такой интервал называют **доверительным интервалом**.

3.1 Расчет доверительного интервала для оценки математического ожидания при неизвестной дисперсии

Пусть результаты измерения величины X имеют нормальное распределение, но значение дисперсии неизвестно. Тогда доверительный интервал для математического ожидания величины X определим, как

$$P\{\bar{X} - t_p s_{\bar{X}} \leq X < \bar{X} + t_p s_{\bar{X}}\} = 2 \int_0^{t_p} S(t; k) dt,$$

где \bar{X} - среднее арифметическое результатов измерения

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

где X_i - i -й результат измерения;

n - число измерений;

$s_{\bar{X}}$ - средне квадратическое отклонение среднего арифметического результатов измерения

$$s_{\bar{X}} = \frac{s_X}{\sqrt{n}},$$

где s_X - среднее квадратическое отклонением результатов измерений

$$s_X = +\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2};$$

t_P – коэффициент Стьюдента, который определяется при заданной доверительной вероятности P ;

$S(t;k)$ - плотность распределения Стьюдента;

k – степеней свободы $k=n-1$.

Значение коэффициента Стьюдента выбирают из справочных таблиц (прил. А).

После вычислений, записывают итог измерения, как

$$X = \bar{X} \pm t_P s_{\bar{X}}; P = \dots \%$$

Пример: При измерении диаметра вала были получены следующие результаты, мм

18,977	18,911	18,999	18,986
18,958	18,968	18,988	18,998
18,927	18,967	18,982	18,966
18,956	18,984	18,953	18,937
18,981	18,954	18,939	18,985

Для расчета среднего, среднеквадратического отклонения, заполним следующую таблицу

№	X_i	X_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
1.	18,977	77	11	121
2.	18,958	58	-8	64
3.	18,927	27	-39	1521
4.	18,956	56	-10	100
5.	18,981	81	15	225
6.	18,911	11	-55	3025
7.	18,968	68	2	4
8.	18,967	67	1	1
9.	18,984	84	18	324

№	X_i	X_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
10.	18,954	54	-12	144
11.	18,999	99	33	1089
12.	18,988	88	22	484
13.	18,982	82	16	256
14.	18,953	53	-13	169
15.	18,939	39	-27	729
16.	18,986	86	20	400
17.	18,998	98	32	1024
18.	18,966	66	0	0
19.	18,937	37	-29	841
20.	18,985	85	19	361
		$\Sigma=1316$		$\Sigma=10882$

Далее получим

$$\bar{X} = 1316/20 = 66 = 18,966 \text{ мм};$$

$$s_X = \sqrt{10882/19} = 24 = 0,024 \text{ мм};$$

$$s_{\bar{X}} = \frac{s_X}{\sqrt{n}} = \frac{0,024}{\sqrt{20}} = 0,0054 \text{ мм}.$$

При заданной величине доверительной вероятности (например $P=0,99$) и числе степеней свободы $k=20-1=19$, по таблице прил. А определяем значение коэффициента Стьюдента

$$t_{99\%} = 2,861.$$

Далее рассчитаем доверительную границу

$$t_{99\%} s_{\bar{X}} = 2,861 \cdot 0,0054 \approx 0,015 \text{ мм}.$$

Конечный результат измерения детали запишем, как

$$X = (18,966 \pm 0,015) \text{ мм}; P=99\%.$$

3.2 Доверительный интервал для оценки среднего квадратического отклонения нормального распределения

Доверительный интервал для оценки среднего квадратического отклонения нормального распределения определяется из следующего соотношения

$$P\{s_{X_1} > \sigma_X \geq s_{X_2}\} = 1 - q,$$

где s_{X_1}, s_{X_2} - границы доверительного интервала

$$s_{X_1} = \frac{\sqrt{n-1}s_X}{\chi_{k, \frac{1}{2}q}}, s_{X_2} = \frac{\sqrt{n-1}s_X}{\chi_{k, 1-\frac{1}{2}q}},$$

где $\chi_{k, \frac{1}{2}q}, \chi_{k, 1-\frac{1}{2}q}$ - коэффициенты определяемые из справочных таблиц

распределения Пирсона (хи-квадрат);

q – уровень значимости (вероятность выхода дисперсии за границы доверительного интервала)

$$q=1-P,$$

где P - уровень доверительной вероятности.

Пример: Для предыдущего примера рассчитаем доверительный интервал для среднего квадратического отклонения.

По заданному уровню доверительной вероятности $P=90\%$ ($q=1-0,9=0,1$) и числе степеней свободы $k=19$ в таблице прил. Б находим значения

$$\chi_{19;0,05}^2 = 7,633; \chi_{19;0,05} = \sqrt{7,633} = 2,763;$$

$$\chi_{19;0,95}^2 = 31,144; \chi_{19;0,95} = \sqrt{31,144} = 5,49.$$

Рассчитаем границы доверительного интервала

$$s_{X_1} = \frac{\sqrt{n-1}s_X}{\chi_{19;0,05}} = \frac{\sqrt{20-1} \cdot 0,024}{2,763} = 0,0378 \text{ мм},$$

$$s_{X_2} = \frac{\sqrt{n-1}s_X}{\chi_{19;0,95}} = \frac{\sqrt{20-1} \cdot 0,024}{5,49} = 0,0190 \text{ мм}.$$

Полученные результаты говорят о том, что значение среднего квадратического отклонения результатов измерения диаметров вала с вероятностью 90% лежит в интервале 0,0190-0,0378 мм.

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение В), студент изучает теоретические сведения согласно пункту 3. Далее выполняет действия аналогичные действиям в

рассмотренных примерах с учетом имеющихся особенностей задания.

В отчёте должны найти отражение следующие пункты:

- название практической работы;
- цель работы;
- индивидуальное задание для выполнения практической работы;
- краткие теоретические сведения;
- результаты выполнения работы: таблица с исходными, результаты расчета;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое доверительный интервал?
2. Для чего используются доверительные интервалы?
3. Как рассчитать границы доверительного интервала для оценки математического ожидания при неизвестной дисперсии?
4. Как рассчитать границы доверительного интервала для оценки среднего квадратического отклонения нормального распределения?

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
Распределение Стьюдента (t_p)

k	P											
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,98	0,99
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750
300	0,126	0,254	0,386	0,525	0,675	0,843	1,038	1,284	1,650	1,968	2,339	2,592

ПРИЛОЖЕНИЕ Б

Распределение хи-квадрат ($\chi^2_{k,P}$)

k	P													
	0,01	0,02	0,05	0,1	0,2	0,3	0,4	0,5	0,7	0,8	0,9	0,95	0,98	0,99
1	0,000157	0,000628	0,00393	0,0158	0,0642	0,148	0,275	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,0201	0,0404	0,103	0,211	0,446	0,713	1,022	1,386	2,408	3,219	4,605	5,991	7,824	9,210
3	0,115	0,185	0,352	0,584	1,005	1,424	1,869	2,366	3,665	4,642	6,251	7,815	9,837	11,345
4	0,297	0,429	0,711	1,064	1,649	2,195	2,753	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	0,554	0,752	1,145	1,610	2,343	3,000	3,655	4,351	6,064	7,289	9,236	11,070	13,388	15,086
6	0,872	1,134	1,635	2,204	3,070	3,828	4,570	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	1,239	1,564	2,167	2,833	3,822	4,671	5,493	6,346	8,383	9,803	12,017	14,067	16,622	18,475
8	1,646	2,032	2,733	3,490	4,594	5,527	6,423	7,344	9,524	11,030	13,362	15,507	18,168	20,090
9	2,088	2,532	3,325	4,168	5,380	6,393	7,357	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	2,558	3,059	3,940	4,865	6,179	7,267	8,295	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	3,053	3,609	4,575	5,578	6,989	8,148	9,237	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	3,571	4,178	5,226	6,304	7,807	9,034	10,182	11,340	14,011	15,812	18,549	21,026	24,054	26,217
13	4,107	4,765	5,892	7,042	8,634	9,926	11,129	12,340	15,119	16,985	19,812	22,362	25,472	27,688
14	4,660	5,368	6,571	7,790	9,467	10,821	12,078	13,339	16,222	18,151	21,064	23,685	26,873	29,141
15	5,229	5,985	7,261	8,547	10,307	11,721	13,030	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	5,812	6,614	7,962	9,312	11,152	12,624	13,983	15,338	18,418	20,465	23,542	26,296	29,633	32,000
17	6,408	7,255	8,672	10,085	12,002	13,531	14,937	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	7,015	7,906	9,390	10,865	12,857	14,440	15,893	17,338	20,601	22,760	25,989	28,869	32,346	34,805
19	7,633	8,567	10,117	11,651	13,716	15,352	16,850	18,338	21,689	23,900	27,204	30,144	33,687	36,191
20	8,260	9,237	10,851	12,443	14,578	16,266	17,809	19,337	22,775	25,038	28,412	31,410	35,020	37,566
21	8,897	9,915	11,591	13,240	15,445	17,182	18,768	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	9,542	10,600	12,338	14,041	16,314	18,101	19,729	21,337	24,939	27,301	30,813	33,924	37,659	40,289
23	10,196	11,293	13,091	14,848	17,187	19,021	20,690	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	10,856	11,992	13,848	15,659	18,062	19,943	21,652	23,337	27,096	29,553	33,196	36,415	40,270	42,980
25	11,524	12,697	14,611	16,473	18,940	20,867	22,616	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	12,198	13,409	15,379	17,292	19,820	21,792	23,579	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	12,879	14,125	16,151	18,114	20,703	22,719	24,544	26,336	30,319	32,912	36,741	40,113	44,140	46,963
28	13,565	14,847	16,928	18,939	21,588	23,647	25,509	27,336	31,391	34,027	37,916	41,337	45,419	48,278
29	14,256	15,574	17,708	19,768	22,475	24,577	26,475	28,336	32,461	35,139	39,087	42,557	46,693	49,588
30	14,953	16,306	18,493	20,599	23,364	25,508	27,442	29,336	33,530	36,250	40,256	43,773	47,962	50,892

ПРИЛОЖЕНИЕ В
Исходные данные к практической работе

№ варианта	n	P	№ варианта	n	P
21.	15	0,9	22.	15	0,95
23.	16	0,95	24.	16	0,98
25.	17	0,98	26.	17	0,8
27.	18	0,8	28.	18	0,7
29.	19	0,7	30.	19	0,9
31.	20	0,9	32.	20	0,95
33.	19	0,95	34.	19	0,98
35.	18	0,98	36.	18	0,8
37.	17	0,8	38.	17	0,7
39.	16	0,7	40.	16	0,9

№ варианта										
№	1	2	3	4	5	6	7	8	9	10
1	28,855	43,026	31,981	58,657	93,441	6,005	7,002	13,808	50,663	10,034
2	28,916	43,073	32,039	58,632	93,467	6,003	7,001	13,838	50,663	10,002
3	28,859	42,976	32,002	58,673	93,439	5,994	6,999	13,808	50,615	9,997
4	28,863	43,029	31,975	58,653	93,462	5,998	6,998	13,818	50,616	9,976
5	28,915	43,057	32,063	58,634	93,446	5,987	7,000	13,818	50,657	10,020
6	28,853	42,982	31,964	58,684	93,469	5,986	6,995	13,814	50,650	9,979
7	28,934	43,040	32,050	58,651	93,451	6,012	7,006	13,847	50,612	9,984
8	28,955	43,068	31,967	58,688	93,464	6,005	6,995	13,831	50,619	9,974
9	28,902	43,037	31,951	58,668	93,436	6,008	7,002	13,799	50,630	9,975
10	28,960	43,074	32,021	58,657	93,428	5,998	7,003	13,850	50,625	9,981
11	28,961	43,018	32,051	58,655	93,439	6,000	7,004	13,800	50,641	9,922
12	28,895	43,006	31,977	58,682	93,442	6,005	6,998	13,809	50,661	10,000
13	28,919	43,046	31,948	58,615	93,423	6,005	7,001	13,812	50,664	9,955
14	28,938	43,059	31,990	58,682	93,461	5,999	6,996	13,829	50,657	10,029
15	28,956	43,024	32,027	58,638	93,446	5,996	6,998	13,796	50,653	9,960
16	28,960	43,006	32,026	58,646	93,459	5,997	7,001	13,819	50,653	9,938
17	28,867	43,047	32,018	58,649	93,477	6,003	6,997	13,856	50,610	9,964
18	28,905	43,075	32,009	58,654	93,469	6,011	6,998	13,842	50,660	10,056
19	28,889	42,980	31,919	58,639	93,454	6,002	6,998	13,799	50,645	10,005
20	28,882	43,029	32,052	58,673	93,464	6,008	7,001	13,848	50,644	10,016

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
1	10,958	41,593	50,389	14,044	80,489	15,983	17,001	95,096	19,000	19,992
2	11,048	41,594	50,369	13,985	80,489	15,972	16,984	95,076	18,949	20,002
3	11,008	41,572	50,389	13,972	80,483	15,973	16,996	95,172	18,952	19,995
4	11,024	41,569	50,370	13,994	80,489	16,035	16,966	95,116	19,036	19,997
5	11,004	41,554	50,371	14,006	80,487	15,960	16,991	95,180	18,907	19,994
6	11,000	41,593	50,369	13,976	80,482	15,967	16,972	95,110	19,014	19,992
7	10,987	41,564	50,378	13,996	80,487	15,952	16,978	95,154	18,966	20,002
8	10,973	41,561	50,377	13,981	80,486	15,991	17,003	95,105	18,992	20,004
9	11,009	41,551	50,371	14,023	80,482	15,967	16,996	95,143	18,972	19,998
10	10,987	41,584	50,371	14,025	80,482	15,964	17,018	95,105	19,010	19,986
11	11,013	41,569	50,364	14,011	80,482	15,997	16,989	95,135	18,913	19,996
12	11,073	41,592	50,366	13,967	80,488	15,948	17,006	95,169	19,005	19,999
13	11,019	41,591	50,362	13,988	80,489	15,976	17,006	95,169	18,977	20,003
14	11,010	41,589	50,365	14,035	80,483	15,968	16,972	95,155	19,028	20,008
15	11,014	41,583	50,364	14,046	80,484	15,919	16,987	95,080	18,994	20,004
16	10,996	41,581	50,375	14,011	80,488	15,968	17,006	95,183	19,012	20,005
17	11,032	41,588	50,361	13,999	80,484	15,960	17,031	95,170	18,994	19,994
18	10,927	41,569	50,379	14,010	80,483	15,943	16,986	95,114	19,024	19,994
19	11,042	41,567	50,373	13,981	80,484	16,062	16,998	95,149	19,036	20,001
20	10,993	41,580	50,355	14,024	80,484	16,050	16,983	95,133	18,973	20,008

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Ложионова
« 14 » _____ г.



ПОСТРОЕНИЕ ГИСТОГРАММЫ, ПОЛИГОНА ЧАСТОТ И ПОЛИГОНА НАКОПЛЕННЫХ ЧАСТОТ РАСПРЕДЕЛЕНИЯ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЯ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Построение гистограммы, полигона частот и полигона накопленных частот распределения результатов измерения : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: Научиться выполнять группировку данных и оформлять результаты обработки эмпирической выборки графически (в виде полигона, гистограммы, эмпирической функции распределения).

2 Задание: Построить полигон, гистограмму и эмпирическую функцию распределения результатов измерения диаметра шейки вала (см. п. 8) и оформить отчет о выполненной работе.

3 Краткие теоретические сведения

3.1. Построение эмпирических распределений. Оценка плотности вероятностей и функции распределения

Просматривая результаты наблюдений (измерений) трудно заметить какую-либо закономерность в их изменении. Выявить такие закономерности позволяют статистические методы.

Предварительная статистическая обработка опытных данных начинается обычно с того, что их располагают в порядке возрастания (точнее, неубывания). Упорядоченная таким образом выборка называется *вариационным рядом*, а сама процедура упорядочения – *ранжированием* (или сортировкой) опытных данных.

Наглядной формой графического представления эмпирических данных является *гистограмма* и *полигон*. Построение гистограммы (полигона) позволяет обоснованно выдвигать гипотезы о законе распределения обрабатываемых экспериментальных данных. Зная закон распределения наблюдаемой случайной величины можно решать многие практические задачи, связанные с обработкой результатов измерений, контролем качества продукции, оценкой эффективности и стабильности технологических процессов.

При непрерывном распределении случайной величины X эмпирическая плотность распределения вероятностей может быть изображена в виде *гистограммы* или *полигона относительных частот*. Для этого должна быть выполнена *группировка* значений выборки, которая состоит в следующем.

Весь интервал $[x_{\min}, x_{\max}]$, в котором заключены элементы выборки, разбивается на ряд частичных непересекающихся интервалов (*разрядов*) $(a_i, b_i]$ длины h и подсчитывается число элементов выборки n_i , попавших в i -й интервал $i = 1, 2, \dots, m$. Параллельно вычис-

ляется и *относительная частота* $w_i = n_i / n$. При графическом изображении гистограммы и полигона каждый интервал удобнее представлять не двумя границами a_i и b_i , а одним значением $\bar{x}_i = a_i + h/2$ - серединой интервала.

Гистограммой частот называется ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы длины h , а высоты равны отношению n_i/h (*плотность частоты*). Площадь частичного i -го прямоугольника равна $h(n_i/h) = n_i$ - числу выборочных элементов, попавших в i -й интервал. Площадь гистограммы частот равна объему выборки n .

Полигон частот – это ломаная линия, получающаяся при соединении точек с координатами $(\bar{x}_i, n_i/h)$, т.е. соединяются середины верхних сторон прямоугольников гистограммы.

Аналогично определяется *гистограмма* и *полигон относительных частот*.

Для *гистограммы относительных частот* площадь частичного i -го прямоугольника равна $h(w_i/h) = w_i$ - относительной частоте элементов, попавших в i -й интервал. Площадь гистограммы относительных частот равна сумме всех относительных частот, т. е. *единице*. При достаточно большом объеме выборки высоты построенных прямоугольников дают приближенные значения плотности распределения $f(x)$ в средних точках \bar{x}_i интервалов (a_i, b_i) .

Полигон относительных частот – это ломаная линия, соединяющая точки $(\bar{x}_1, w_1/h), (\bar{x}_2, w_2/h), \dots, (\bar{x}_m, w_m/h)$, где \bar{x}_i - середины интервалов группирования, w_i - соответствующие им относительные частоты.

Полигон накопленных частот – это график ломаной линии, соединяющей точки

$$(b_1, n_1), (b_2, n_1 + n_2), \dots, \left(b_k, \sum_{i=1}^k n_i \right), \dots, \left(b_m, \sum_{i=1}^m n_i \right),$$

с абсциссами, равными правым границам интервалов группирования, и ординатами, равными накопленным частотам.

Аналогично, *полигон накопленных относительных частот* – это ломаная линия, получающаяся соединением точек с координатами

$$(b_1, w_1), (b_2, w_1 + w_2), \dots, \left(b_k, \sum_{i=1}^k w_i\right), \dots, \left(b_m, \sum_{i=1}^m w_i = 1\right).$$

Эмпирической (выборочной) функцией распределения называется функция $\widehat{F}(x)$, определяющая для каждого значения x относительную частоту события $X < x$. Строится она так. Для данного числа x подсчитывается число элементов вариационного ряда $x_{(i)}$, меньших x и делится на объем выборки:

$$\widehat{F}_n(x) = \begin{cases} 0, & x < x_{\min} \equiv x_{(1)} \\ \frac{k}{n}, & x_{(k)} < x < x_{(k+1)}, \quad k = 1, 2, \dots, n-1, \\ 1, & x > x_{\max} \equiv x_{(n)} \end{cases}$$

Используя функцию единичного скачка

$$H(u) = \begin{cases} 0, & u < 0 \\ 1, & u \geq 0 \end{cases},$$

формулу для эмпирической функции распределения можно записать более компактно

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n H(x - x_{(i)}).$$

В математической статистике доказывается, что гистограмма и полигон относительных частот являются *состоятельными оценками плотности распределения*, а полигон накопленных относительных частот (или эмпирическая функция распределения) – *состоятельной оценкой истинной функции распределения* генеральной совокупности. При этом, чем больше объем выборки, тем мельче можно взять интервалы разбиения h и тем точнее гистограмма и эмпирическая функция распределения будут аппроксимировать соответствующие теоретические распределения.

Число интервалов группирования эмпирических данных m существенно влияет на вид гистограммы. Четких рекомендаций на этот счет не существует, ясно лишь то, что это число не должно быть «очень малым» или «очень большим». Для выборок среднего объема (несколько сотен элементов) число интервалов обычно выбирают в диапазоне 8 – 12. Согласно рекомендации ВНИИ Метрологии им.

Д.И. Менделеева в зависимости от объема выборки n число интервалов группирования m следует выбирать согласно таблице

n	40 – 100	100 – 500	500 – 1000	1000 – 10000
m	7 – 9	8 – 12	10 – 16	12 – 22

Часто в литературе по математической статистике для выбора оптимального значения m рекомендуется формула

$$m = 1 + 3.32 \lg(n) = 1 + \log_2(n) = 1 + 1.44 \ln(n),$$

записанная с использованием десятичного, двоичного или натурального логарифма; или формула

$$m = 5 \lg(n),$$

где n - объем выборки. Существуют и иные рекомендации. Однако все рекомендуемые формулы не имеют строгого обоснования и дают лишь ориентировочное количество интервалов m , которое при желании может быть изменено в ту или другую сторону.

При определении диапазона возможных значений параметра m можно ориентироваться на формулы

$$m_{\min} = 0.55n^{0.4}, \quad m_{\max} = 1.25n^{0.4},$$

дающие оценки минимального и максимального значения данного параметра. В качестве оптимального значения m рекомендуется выбирать целое число из интервала $[m_{\min}, m_{\max}]$. Обычно в этот интервал попадает несколько целых чисел; после построения гистограмм для каждого из этих чисел, «оптимальное» значение параметра m определяется визуально. Для симметричных распределений в качестве значения m рекомендуется выбирать нечетное целое число из интервала $[m_{\min}, m_{\max}]$.

После определения m вычисляется длина интервалов группирования

$$h = \frac{x_{\max} - x_{\min}}{m},$$

и производятся все дальнейшие вычисления, необходимые для построения гистограммы, полигона и эмпирической функции распределения.

Пример: При обработке валика на токарном станке необходимо выдержать размер $\varnothing 22,27_{-0,44}$.

Из текущей продукции станка берется выборка из деталей, изготовленных при неизменной настройке, объемом $n=90$ шт. Результаты измерений занесены в таблицу 1.

Таблица 1 – Результаты измерения диаметра деталей

22,02	21,97	22,03	22,05	22,01	22,04	21,98	22,04	21,99	21,93	22,05	21,98	21,91
21,99	22,05	21,99	21,84	21,81	22,06	22,00	22,04	21,92	22,07	21,92	22,0	22,0
21,98	22,06	22,03	22,05	21,98	22,06	22,04	21,88	22,08	21,98	22,04	22,0	22,01
21,92	21,96	21,99	21,99	21,94	21,90	21,93	21,94	22,07	22,09	21,82	21,92	21,99
21,94	22,06	22,02	21,99	22,15	22,00	21,88	21,97	21,99	22,13	21,88	22,03	21,96
21,89	21,97	21,93	21,95	21,98	22,01	22,05	22,04	22,09	21,87	22,09	22,01	22,07
21,89	22,0	21,96	22,06	21,93	22,02	21,95	22,06	22,03	22,05	22,03	21,89	

Находим наибольшее и наименьшее значения

$$X_{max} = 22,15; X_{min} = 21,81.$$

Определяем размах

$$R = 22,15 - 21,81 = 0,34.$$

Принимаем количество интервалов равное 9 ($m=9$). Определяем ширину интервала

$$h = 0,34/9 = 0,038.$$

Для удобства построения выбираем ширину интервала равную 0,04 ($h=0,04$).

Границы интервалов устанавливаем следующими: левая граница первого интервала принимается 21,80 (меньше X_{min}), правая отстоит на ширину интервала и составляет 21,84. Последующие границы 21,88; 21,92; 21,96 и т.д.

Правая граница последнего интервала 22,16, что больше наибольшего из имеющихся значений.

Определяем частоту попадания размеров в установленные интервалы (n_i).

Результаты подсчетов сводим в таблицу. Накопленная частота определяется путем прибавления каждого последующего значения частоты к сумме предыдущих значений.

Таблица 3.2 – Бланк для подсчета частоты

№	Интервалы	Середина интервала	Штриховые отметки частоты	Частота	Накопленная частота
---	-----------	--------------------	---------------------------	---------	---------------------

1	21,80-21,84	21,82		3	3
2	21,84-21,88	21,86		4	7
3	21,88-21,92	21,90		9	16
4	21,92-21,96	21,94		12	28
5	21,96-22,00	21,98		24	52
6	22,00-22,04	22,02		17	69
7	22,04-22,08	22,06		16	85
8	22,08-22,12	22,10		3	88
9	22,12-22,16	22,14		2	90

Строим гистограмму и полигон частот распределения (рисунок 1).

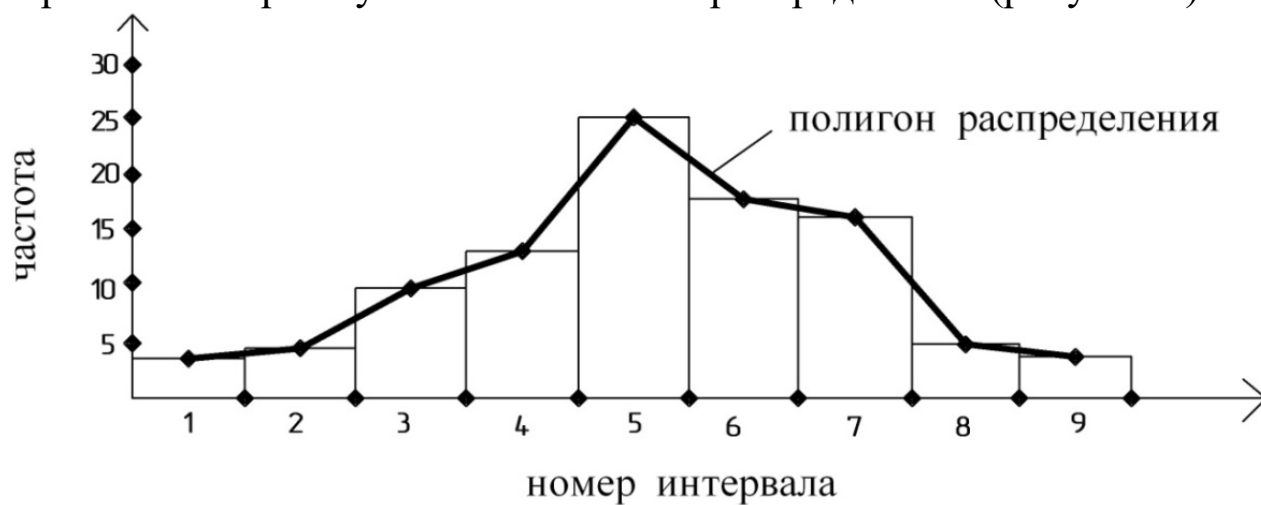


Рис. 1 – Гистограмма и полигон частот распределения значений диаметра вала

Далее строим полигон накопленных частот (рис. 2)



Рис. 2 - Полигон накопленных частот распределения значений диаметра вала

3.2 Анализ формы гистограммы

Гистограмма может иметь различную форму, по которой можно судить об условиях и результатах исследуемого процесса. При анализе гистограммы можно выяснить в удовлетворительном ли состоянии находится партия изделий и технологический процесс, а также предложить мероприятия по приведению процесса в стабильное состояние.

На рисунке 3 приведены типичные виды гистограмм, которыми можно воспользоваться как образцами при анализе процессов.

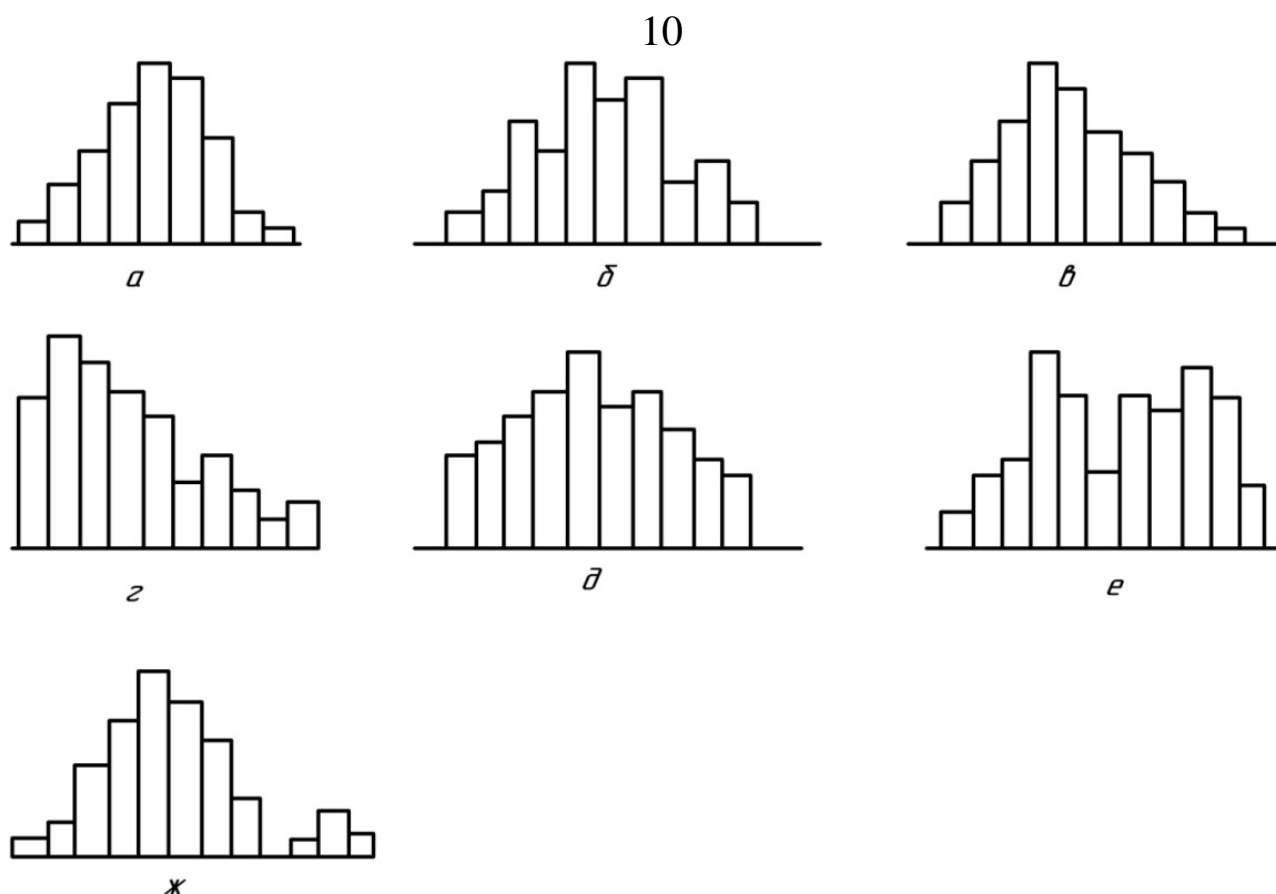


Рис. 3 – Виды гистограмм

Обычный тип, симметричный (рисунок 3, а). Гистограмма с таким распределением встречается чаще всего. Она указывает на стабильность процесса.

Гребенка (мультимодальный тип) (рисунок 3, б). Интервалы через один имеют более низкие частоты. Такая форма встречается, когда число единичных наблюдений, попадающих в интервал, колеблется, или когда действует определенное правило округления данных.

Положительно (отрицательно) скошенное распределение (рисунок 3, в). Среднее значение гистограммы локализуется слева (справа) от центра размаха. Частоты довольно резко спадают при движении влево (вправо) и, наоборот, медленно вправо (влево). Такая асимметричная форма встречается, когда невозможно получить значение ниже определенного, например, для диаметра деталей.

Распределение с обрывом слева (справа) (рисунок 3, г). Это одна из тех форм, которые часто встречаются при 100%-м контроле изделий из-за плохой воспроизводимости процесса, а также, когда,

например, отобраны и исключены из выборки все изделия с параметрами ниже (или выше) контрольного норматива.

Плато (равномерное и прямоугольное распределение) (рисунок 3, д). Такая гистограмма получается в случаях, когда объединяются несколько распределений, в которых средние значения различаются незначительно. Например, детали обрабатываются на станках-дублерах. Анализ такой гистограммы целесообразно проводить, используя метод расслоения.

Двухпиковое (бимодальное) распределение (рисунок 3, е). Такая форма встречается, когда смешиваются два распределения с далеко отстоящими средними значениями, например, в случае наличия разницы между двумя видами материалов, двумя операторами, разными типоразмерами станков. В этом случае можно провести расслоение по двум видам факторов, исследовать причины различия и принять соответствующие меры для его устранения.

Распределение с изолированным пиком (рисунок 3, ж). Рядом с распределением обычного типа появляется маленький изолированный пик. Эта форма гистограммы появляется при наличии малых включений данных из другого распределения или грубых ошибок при измерениях.

По виду гистограммы можно приблизительно судить о стабильности и устойчивости процесса. Однако для более точного решения этих задач необходимо знать закон распределения случайной величины.

На практике эмпирические распределения строятся в виде гистограммы или полигона, т.е. ломаной линии, соединяющей середины частот в каждом интервале (рисунок 1). Эмпирические распределения значений случайных величин могут подчиняться различным законам. Зная закон распределения, можно заранее предсказать, что те или иные значения случайной величины могут появиться с той или иной вероятностью. Знание законов распределения бывает необходимым для принятия определенных решений по анализу и управлению процессами. Закон распределения может быть представлен в интегральной или дифференциальной форме, которая является более наглядной и поэтому более предпочтительной.

Среди различных законов наиболее широко применяется закон нормального распределения (закон Гаусса). Этому закону подчиня-

ются случайные величины, на которые оказывают влияние многочисленные, примерно равные по силе воздействия, факторы. Этому закону подчиняются размеры деталей, обрабатываемых на настроенных станках; результаты многократных измерений, при отсутствии систематических погрешностей и многие другие величины. Нормальное распределение свидетельствует об устойчивости технологического процесса.

Для того чтобы убедиться, что распределение подчиняется закону Гаусса, необходимо провести проверку гипотезы нормальности распределения.

Если гипотеза нормальности распределения не подтверждается, то оценивание процесса ведут на основе того закона распределения, который был установлен, либо, что бывает чаще, получают приближенную оценку зоны рассеивания.

При подтверждении гипотезы о соответствии эмпирического распределения нормальному закону становится возможным исследование воспроизводимости процесса, т.е. определяется неизменность основных параметров процесса: среднего арифметического значения и среднего квадратического отклонения. После чего можно провести статистический анализ точности технологического процесса.

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение), студент изучает теоретические сведения согласно пункту 3. Далее выполняет действия аналогичные действиям в рассмотренном примере (пункт 3.1) с **учетом имеющихся особенностей задания.**

В отчёте должны найти отражение следующие пункты:

- название практической работы;
- цель работы;
- индивидуальное задание для выполнения практической работы;
- краткие теоретические сведения;
- результаты выполнения работы: таблица с исходными, таблица с параметрами распределения;

- подробные выводы по работе.

Контрольные вопросы

1. Что графически представляет собой гистограмма распределения?
2. Как строится полигон распределения?
3. Как определяется количество и ширина интервалов при построении гистограммы?
4. О чем свидетельствует форма гистограммы?

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
Исходные данные к практической работе

№ варианта										
№	1	2	3	4	5	6	7	8	9	10
1	28,855	43,026	31,981	58,657	93,441	6,005	7,002	13,808	50,663	10,034
2	28,916	43,073	32,039	58,632	93,467	6,003	7,001	13,838	50,663	10,002
3	28,859	42,976	32,002	58,673	93,439	5,994	6,999	13,808	50,615	9,997
4	28,863	43,029	31,975	58,653	93,462	5,998	6,998	13,818	50,616	9,976
5	28,915	43,057	32,063	58,634	93,446	5,987	7,000	13,818	50,657	10,020
6	28,853	42,982	31,964	58,684	93,469	5,986	6,995	13,814	50,650	9,979
7	28,934	43,040	32,050	58,651	93,451	6,012	7,006	13,847	50,612	9,984
8	28,955	43,068	31,967	58,688	93,464	6,005	6,995	13,831	50,619	9,974
9	28,902	43,037	31,951	58,668	93,436	6,008	7,002	13,799	50,630	9,975
10	28,960	43,074	32,021	58,657	93,428	5,998	7,003	13,850	50,625	9,981
11	28,961	43,018	32,051	58,655	93,439	6,000	7,004	13,800	50,641	9,922
12	28,895	43,006	31,977	58,682	93,442	6,005	6,998	13,809	50,661	10,000
13	28,919	43,046	31,948	58,615	93,423	6,005	7,001	13,812	50,664	9,955
14	28,938	43,059	31,990	58,682	93,461	5,999	6,996	13,829	50,657	10,029
15	28,956	43,024	32,027	58,638	93,446	5,996	6,998	13,796	50,653	9,960
16	28,960	43,006	32,026	58,646	93,459	5,997	7,001	13,819	50,653	9,938
17	28,867	43,047	32,018	58,649	93,477	6,003	6,997	13,856	50,610	9,964
18	28,905	43,075	32,009	58,654	93,469	6,011	6,998	13,842	50,660	10,056
19	28,889	42,980	31,919	58,639	93,454	6,002	6,998	13,799	50,645	10,005
20	28,882	43,029	32,052	58,673	93,464	6,008	7,001	13,848	50,644	10,016
21	28,874	43,038	32,011	58,669	93,450	6,004	7,002	13,787	50,616	10,033
22	28,917	42,985	32,011	58,645	93,474	6,006	6,999	13,826	50,626	9,944
23	28,915	43,050	32,025	58,606	93,453	6,008	6,999	13,837	50,663	9,994
24	28,900	43,044	31,954	58,639	93,461	6,003	7,004	13,819	50,649	9,976
25	28,946	43,068	32,010	58,634	93,445	5,991	6,998	13,825	50,635	9,982
26	28,911	42,983	31,963	58,691	93,432	5,996	6,997	13,826	50,661	10,029
27	28,907	43,021	32,004	58,646	93,433	6,000	7,001	13,835	50,620	9,956
28	28,870	42,998	31,985	58,644	93,475	5,999	6,998	13,812	50,620	10,013
29	28,890	43,027	32,028	58,605	93,462	5,997	6,998	13,844	50,656	9,994
30	28,904	43,006	31,954	58,659	93,441	5,999	7,001	13,843	50,621	9,895
31	28,952	43,006	32,013	58,669	93,427	6,004	7,000	13,847	50,658	9,984
32	28,961	43,060	32,040	58,653	93,455	5,981	7,002	13,844	50,617	10,012
33	28,892	42,984	32,039	58,690	93,459	6,001	6,998	13,831	50,664	9,974
34	28,893	42,978	32,000	58,626	93,432	5,996	7,000	13,811	50,617	10,002
35	28,912	43,025	32,071	58,626	93,466	6,006	7,003	13,823	50,636	9,972

№ варианта										
№	1	2	3	4	5	6	7	8	9	10
36	28,916	42,981	31,982	58,670	93,456	5,991	7,003	13,828	50,619	9,983
37	28,928	43,059	31,983	58,649	93,449	6,008	7,000	13,787	50,630	9,963
38	28,915	43,067	32,022	58,671	93,433	6,004	7,001	13,789	50,641	9,989
39	28,925	43,017	32,030	58,606	93,479	5,999	6,996	13,844	50,619	10,034
40	28,863	43,075	31,926	58,636	93,440	5,994	7,001	13,848	50,635	9,963
41	28,893	43,010	32,008	58,667	93,451	6,008	7,002	13,847	50,663	9,961
42	28,853	43,040	31,974	58,610	93,436	6,003	6,990	13,812	50,654	9,983
43	28,889	42,987	31,974	58,644	93,476	6,004	7,002	13,818	50,625	9,979
44	28,867	43,071	31,983	58,669	93,453	6,009	6,999	13,790	50,648	9,976
45	28,933	43,041	32,019	58,663	93,429	6,003	7,005	13,841	50,617	9,981
46	28,852	43,040	31,981	58,623	93,451	6,003	6,996	13,797	50,639	10,049
47	28,913	43,054	32,084	58,607	93,459	6,002	6,999	13,844	50,631	9,938
48	28,868	43,075	31,995	58,627	93,444	5,999	6,999	13,796	50,640	10,002
49	28,904	43,075	32,008	58,633	93,454	5,999	6,999	13,813	50,634	10,024
50	28,890	43,033	32,018	58,680	93,448	5,995	6,998	13,835	50,615	9,992
51	28,890	42,973	32,024	58,644	93,442	6,007	7,004	13,849	50,631	10,020
52	28,854	43,068	31,945	58,630	93,459	6,001	7,005	13,799	50,663	9,965
53	28,944	42,988	32,020	58,648	93,442	6,007	6,998	13,817	50,630	10,058
54	28,930	43,029	32,027	58,653	93,441	5,997	7,002	13,803	50,613	9,996
55	28,943	42,973	31,959	58,614	93,461	6,002	7,000	13,797	50,629	9,985
56	28,942	43,066	32,003	58,668	93,468	5,987	6,999	13,789	50,644	9,981
57	28,882	43,043	32,059	58,619	93,477	6,003	7,002	13,859	50,630	9,960
58	28,924	43,014	32,037	58,694	93,452	6,012	6,997	13,827	50,623	10,036
59	28,898	43,043	32,031	58,627	93,432	5,996	6,999	13,822	50,630	10,021
60	28,935	42,982	31,999	58,654	93,424	5,995	6,994	13,838	50,612	10,031

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
1	10,958	41,593	50,389	14,044	80,489	15,983	17,001	95,096	19,000	19,992
2	11,048	41,594	50,369	13,985	80,489	15,972	16,984	95,076	18,949	20,002
3	11,008	41,572	50,389	13,972	80,483	15,973	16,996	95,172	18,952	19,995
4	11,024	41,569	50,370	13,994	80,489	16,035	16,966	95,116	19,036	19,997
5	11,004	41,554	50,371	14,006	80,487	15,960	16,991	95,180	18,907	19,994
6	11,000	41,593	50,369	13,976	80,482	15,967	16,972	95,110	19,014	19,992
7	10,987	41,564	50,378	13,996	80,487	15,952	16,978	95,154	18,966	20,002
8	10,973	41,561	50,377	13,981	80,486	15,991	17,003	95,105	18,992	20,004
9	11,009	41,551	50,371	14,023	80,482	15,967	16,996	95,143	18,972	19,998
10	10,987	41,584	50,371	14,025	80,482	15,964	17,018	95,105	19,010	19,986
11	11,013	41,569	50,364	14,011	80,482	15,997	16,989	95,135	18,913	19,996

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
12	11,073	41,592	50,366	13,967	80,488	15,948	17,006	95,169	19,005	19,999
13	11,019	41,591	50,362	13,988	80,489	15,976	17,006	95,169	18,977	20,003
14	11,010	41,589	50,365	14,035	80,483	15,968	16,972	95,155	19,028	20,008
15	11,014	41,583	50,364	14,046	80,484	15,919	16,987	95,080	18,994	20,004
16	10,996	41,581	50,375	14,011	80,488	15,968	17,006	95,183	19,012	20,005
17	11,032	41,588	50,361	13,999	80,484	15,960	17,031	95,170	18,994	19,994
18	10,927	41,569	50,379	14,010	80,483	15,943	16,986	95,114	19,024	19,994
19	11,042	41,567	50,373	13,981	80,484	16,062	16,998	95,149	19,036	20,001
20	10,993	41,580	50,355	14,024	80,484	16,050	16,983	95,133	18,973	20,008
21	11,065	41,583	50,390	13,982	80,484	16,047	17,013	95,125	18,981	20,005
22	11,050	41,570	50,363	13,985	80,486	15,874	17,021	95,086	19,011	19,997
23	10,997	41,568	50,389	13,956	80,481	15,973	17,009	95,083	19,002	19,993
24	11,050	41,585	50,388	14,014	80,484	16,005	16,992	95,170	18,981	19,996
25	11,048	41,574	50,374	13,990	80,486	15,976	17,023	95,143	19,009	19,992
26	11,065	41,567	50,354	14,031	80,489	16,011	17,030	95,078	18,975	19,997
27	10,994	41,590	50,360	14,040	80,484	16,063	16,986	95,151	18,993	19,995
28	10,931	41,591	50,390	13,999	80,486	15,978	16,972	95,079	19,007	20,000
29	10,997	41,555	50,382	13,993	80,486	16,030	17,033	95,068	18,913	20,009
30	10,931	41,576	50,364	13,982	80,488	16,008	16,978	95,095	18,948	20,003
31	10,946	41,574	50,374	13,940	80,484	15,969	17,001	95,066	18,995	20,013
32	10,976	41,571	50,374	13,996	80,482	16,013	16,987	95,065	18,967	20,008
33	11,029	41,569	50,353	13,977	80,486	15,996	17,027	95,154	19,064	19,990
34	10,971	41,588	50,367	13,986	80,485	15,988	17,006	95,161	18,981	19,998
35	10,996	41,545	50,388	13,995	80,483	15,988	17,005	95,129	18,960	20,002
36	10,970	41,595	50,394	13,980	80,483	15,974	16,994	95,106	18,984	19,997
37	11,065	41,593	50,365	14,033	80,485	15,993	17,002	95,173	19,068	19,996
38	11,031	41,586	50,392	13,989	80,484	16,061	16,976	95,132	18,984	19,990
39	11,024	41,556	50,397	14,036	80,484	16,029	17,003	95,087	19,056	20,002
40	10,972	41,549	50,353	13,993	80,485	15,983	16,980	95,133	19,064	20,010
41	11,005	41,546	50,354	14,007	80,489	15,974	17,004	95,170	18,976	19,999
42	11,064	41,569	50,387	14,031	80,488	16,072	16,980	95,063	18,947	20,006
43	10,968	41,559	50,390	13,965	80,484	16,039	16,990	95,125	18,990	19,983
44	11,000	41,548	50,359	14,015	80,485	15,975	16,999	95,174	18,981	20,007
45	10,997	41,551	50,362	14,003	80,484	15,929	16,998	95,111	18,974	20,005
46	10,959	41,559	50,373	14,017	80,487	15,915	16,966	95,174	18,982	19,991
47	11,018	41,594	50,380	13,950	80,487	16,001	16,996	95,160	19,041	19,994
48	11,016	41,568	50,368	13,990	80,481	15,956	16,995	95,163	18,973	20,008
49	10,942	41,560	50,394	13,963	80,484	16,013	17,003	95,109	19,025	20,008
50	10,975	41,577	50,359	13,978	80,486	16,035	16,974	95,073	18,986	20,010

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
51	11,015	41,566	50,395	14,017	80,481	16,021	17,017	95,172	19,042	20,003
52	11,030	41,591	50,367	14,007	80,487	15,941	17,024	95,129	19,114	20,007
53	11,036	41,588	50,392	14,011	80,482	16,042	17,016	95,106	18,939	19,995
54	11,018	41,565	50,366	14,046	80,484	15,893	17,013	95,079	18,975	19,999
55	10,997	41,548	50,391	13,976	80,483	15,874	16,994	95,125	18,939	20,006
56	10,958	41,582	50,383	13,946	80,485	16,014	17,011	95,181	19,029	19,992
57	10,981	41,590	50,384	14,017	80,486	16,034	16,998	95,172	19,016	19,999
58	11,060	41,546	50,362	14,021	80,484	15,968	16,986	95,173	18,997	20,007
59	11,038	41,557	50,360	14,036	80,484	15,924	17,022	95,149	19,015	19,991
60	11,007	41,567	50,375	13,993	80,481	15,952	17,021	95,176	19,030	19,999

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Жукмонова
« 14 » _____ г.



РАСЧЕТ ОЦЕНОЧНЫХ ХАРАКТЕРИСТИК РАСПРЕДЕЛЕНИЯ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЯ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Расчет оценочных характеристик распределения результатов измерения: методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: Получить сведения об основных характеристиках распределения результатов измерения и практические навыки их расчета.

2 Задание: Произвести расчет основных характеристик распределения результатов измерения размера изделия (см. п. 8) и оформить отчет о выполненной работе.

3 Краткие теоретические сведения

3.1 Основные характеристики распределения

3.1.1 Среднее арифметическое \bar{X} результатов наблюдений, которое принимая за оценку истинного значения измеряемой величины

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

где X_i - i -й результат измерения;

n – число измерений.

3.1.2 Дисперсия результатов наблюдений D_X , является характеристикой их рассеивания относительно математического ожидания:

$$D_X = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Дисперсия имеет размерность квадрата измеряемой величины, поэтому она не совсем удобна в качестве характеристики рассеивания (рис. 1). Для этого используется положительное значение корня квадратного из дисперсии, называемое средним квадратическим отклонением результатов наблюдений:

$$s_X = +\sqrt{D_X}$$

или

$$s_X = +\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

3.1.3 Оценку $s_{\bar{X}}$ среднего квадратического отклонения среднего арифметического:

$$s_{\bar{X}} = \frac{s_X}{\sqrt{n}}.$$

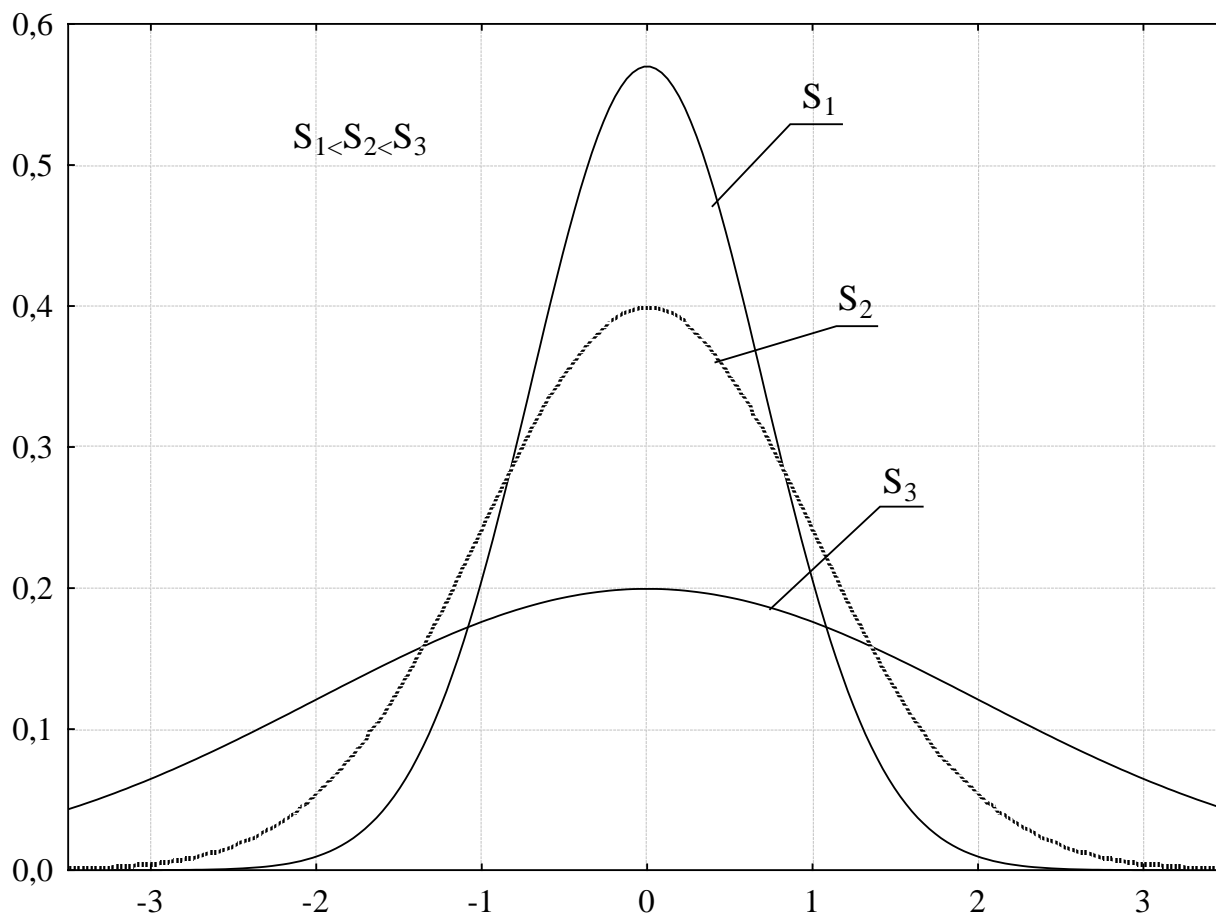


Рисунок 1 – Графики плотности нормального распределения для различных значений s_x

3.1.4 Мода. Мода выборки - это значение, наиболее часто встречающееся в выборке.

Пример: Для набора значений: 10, 11, 12, 11, 14, 13, 11, 12 мода равна

$$M_o=11.$$

3.1.5 Медиана. Медиана выборки - это значение, которое разбивает выборку на две равные части. Половина наблюдений лежит ниже медианы, и половина наблюдений лежит выше медианы. Медиана вычисляется следующим образом. Изучаемая выборка упорядочивается в порядке возрастания. Получаемая последовательность a_k , где $k=1, \dots, 2 \cdot m+1$ называется вариационным рядом или порядковыми статистиками. Если число наблюдений нечетно, то медиана оценивается как: a_{m+1} . Если число наблюдений четно, то медиана оценивается как: $(a_m + a_{m+1})/2$.

Пример: Для рассмотренного выше примера, построим вариационный ряд (в порядке возрастания): 10, 11, 11, 11, 12, 12, 13, 14. Т.к. количество членов ряда равно 8, то медиана будет равна

$$Me = (11 + 12) / 2 = 11,5.$$

3.1.6 Асимметрия. Асимметрия или коэффициент асимметрии As является мерой несимметричности распределения.

$$As = \frac{n}{(n-1)(n-2)s_X^3} \sum_{i=1}^n (X_i - \bar{X})^3$$

Если этот коэффициент значительно отличается от 0, распределение является несимметричным (рис. 2), т.е. если вы разобьете такое распределение пополам в точке среднего (или медианы), то распределения значений с двух сторон от этой центральной точки будут неодинаковыми (т.е. несимметричными). Такое распределение можно назвать "скошенным".

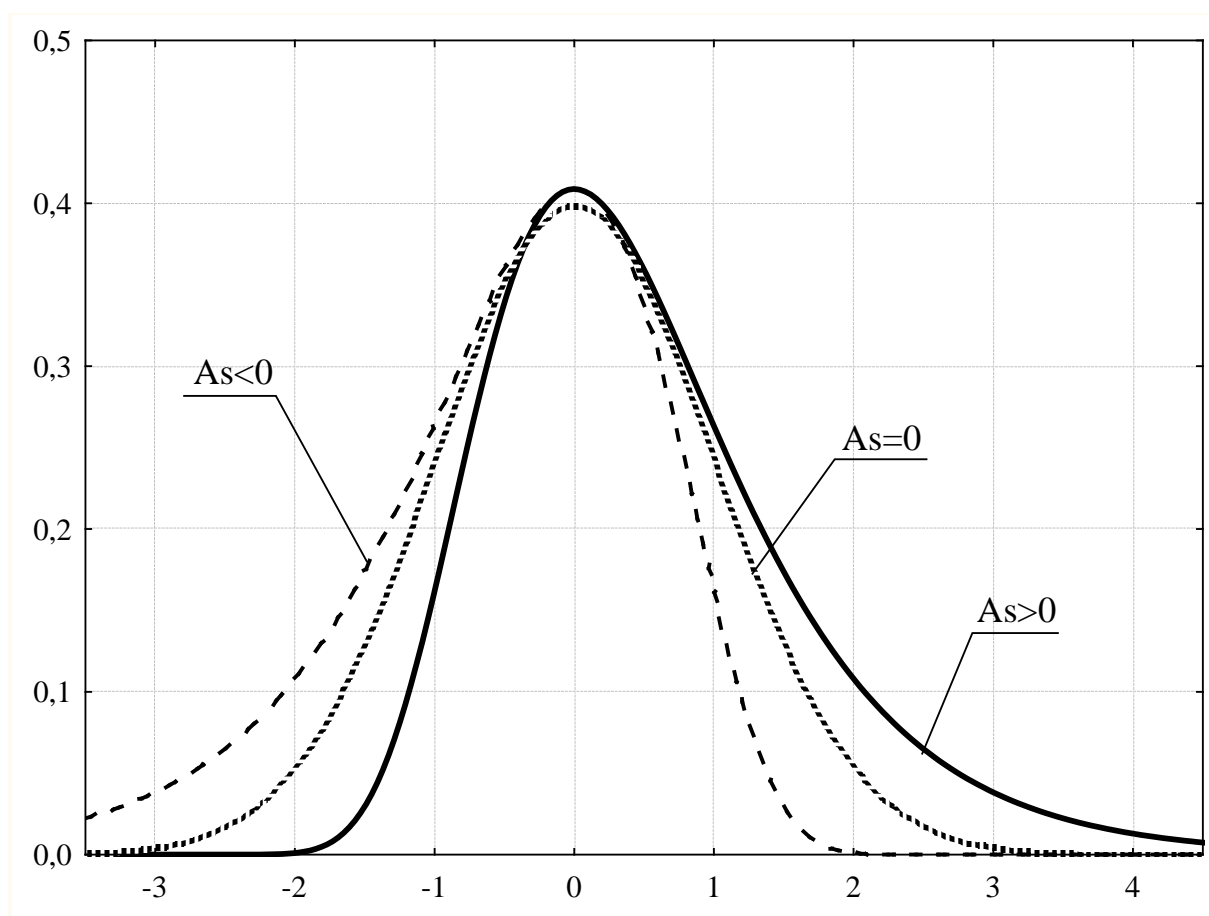


Рисунок 2 – Случаи с положительной и отрицательной асимметрией

Плотность нормального распределения симметрична относительно среднего.

Принято считать, что асимметрия выше 0,5 (независимо от знака) считается значительной, а меньше 0,25 не значительной.

3.1.7 Эксцесс. Эксцесс или точнее, коэффициент эксцесса E_k измеряет "пикообразность" распределения (рис. 3). Если эксцесс значительно отличается от 0, то функция плотности либо имеет более закругленный, либо более острый пик, чем пик плотности нормального распределения. Функция плотности нормального распределения имеет эксцесс равный 0. Оценка эксцесса (выборочный эксцесс) вычисляется по формуле:

$$E_k = \frac{n(n+1)}{(n-1)(n-2)(n-3)s_X^4} \sum_{i=1}^n (X_i - \bar{X})^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

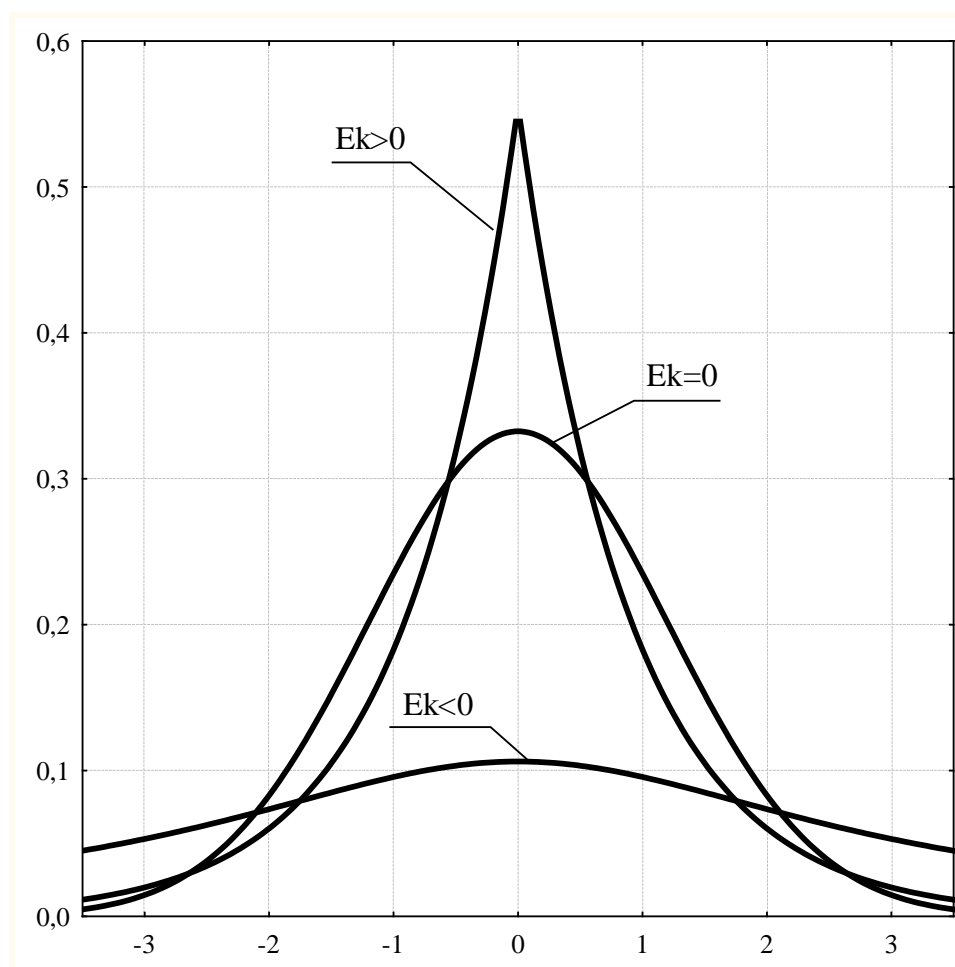


Рисунок 3 – Случаи с положительным и отрицательным эксцессом

4 Расчет основных характеристик распределения

Расчет рассмотрим на примере для 20 результатов

18,977	18,911	18,999	18,986
18,958	18,968	18,988	18,998
18,927	18,967	18,982	18,966
18,956	18,984	18,953	18,937
18,981	18,954	18,939	18,985

Для расчета среднего, среднеквадратического отклонения, асимметрии и эксцесса заполним следующую таблицу

№	X_i	X_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})^3$	$(X_i - \bar{X})^4$
1.	18,977	77	11	121	1331	14641
2.	18,958	58	-8	64	-512	4096
3.	18,927	27	-39	1521	-59319	2313441
4.	18,956	56	-10	100	-1000	10000
5.	18,981	81	15	225	3375	50625
6.	18,911	11	-55	3025	-166375	9150625
7.	18,968	68	2	4	8	16
8.	18,967	67	1	1	1	1
9.	18,984	84	18	324	5832	104976
10.	18,954	54	-12	144	-1728	20736
11.	18,999	99	33	1089	35937	1185921
12.	18,988	88	22	484	10648	234256
13.	18,982	82	16	256	4096	65536
14.	18,953	53	-13	169	-2197	28561
15.	18,939	39	-27	729	-19683	531441
16.	18,986	86	20	400	8000	160000
17.	18,998	98	32	1024	32768	1048576
18.	18,966	66	0	0	0	0
19.	18,937	37	-29	841	-24389	707281
20.	18,985	85	19	361	6859	130321
		$\Sigma=1316$		$\Sigma=10882$	$\Sigma=-166348$	$\Sigma=15761050$

Далее получим

$$\bar{X} = 1316/20 = 65,8 = 18,966;$$

$$s_x = \sqrt{10882/19} = 24 = 0,024;$$

$$As = \frac{20}{19 \cdot 18 \cdot 24^3} (-166348) = -0,7;$$

$$Ek = \frac{20 \cdot 21}{19 \cdot 18 \cdot 17 \cdot 24^4} 15761050 - \frac{3 \cdot 19^2}{18 \cdot 17} = -0,10748$$

5 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение), студент изучает теоретические сведения согласно пункту 3. Далее выполняет на компьютере действия аналогичные действиям в рассмотренном примере (пункт 4) **с учетом имеющихся особенностей задания.**

В отчёте по работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы: таблица с исходными, таблица с параметрами распределения;
- подробные выводы по работе.

Контрольные вопросы

1. Назовите основные параметры распределения.
2. Что характеризуют основные параметры распределения?
3. В каком случае дисперсия одного распределения больше другого?
4. В каком случае эксцесс одного распределения больше другого?
5. В каком случае асимметрия одного распределения больше другого?

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
Исходные данные к лабораторной работе

№ варианта										
№	1	2	3	4	5	6	7	8	9	10
1	201,576	86,114	19,046	158,772	80,915	181,724	135,073	62,720	43,996	40,701
2	201,561	86,127	19,000	158,755	80,917	181,722	135,086	62,721	43,975	40,690
3	201,558	86,117	19,055	158,807	80,956	181,725	135,082	62,707	44,003	40,734
4	201,570	86,126	19,001	158,805	80,956	181,713	135,073	62,731	43,990	40,729
5	201,562	86,134	19,015	158,779	80,948	181,718	135,080	62,711	43,998	40,703
6	201,555	86,079	19,033	158,802	80,949	181,724	135,088	62,737	43,996	40,691
7	201,588	86,171	18,979	158,792	80,944	181,718	135,083	62,701	43,971	40,714
8	201,566	86,106	19,037	158,791	80,948	181,721	135,071	62,735	43,994	40,697
9	201,544	86,064	19,050	158,806	80,974	181,724	135,085	62,712	43,992	40,739
10	201,570	86,127	19,016	158,789	80,941	181,726	135,082	62,723	43,993	40,742
11	201,559	86,121	19,080	158,829	80,943	181,723	135,091	62,699	43,989	40,723
12	201,571	86,123	19,001	158,789	80,934	181,725	135,091	62,715	43,999	40,746
13	201,557	86,124	19,072	158,756	80,934	181,734	135,082	62,705	43,999	40,721
14	201,561	86,119	19,024	158,777	80,931	181,726	135,088	62,731	43,991	40,718
15	201,569	86,140	19,042	158,768	80,926	181,723	135,087	62,734	43,983	40,720
16	201,564	86,138	19,018	158,791	80,949	181,726	135,085	62,721	43,976	40,726
17	201,570	86,132	19,012	158,810	80,958	181,721	135,076	62,715	43,995	40,749
18	201,564	86,099	19,013	158,802	80,968	181,724	135,089	62,694	43,985	40,664
19	201,578	86,132	19,002	158,782	80,954	181,720	135,083	62,726	43,974	40,746
20	201,577	86,114	19,040	158,821	80,952	181,718	135,085	62,702	44,004	40,712

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
1	139,752	125,290	118,606	56,888	189,731	119,778	125,777	189,498	73,219	42,442
2	139,751	125,266	118,602	56,931	189,690	119,714	125,788	189,513	73,237	42,435
3	139,749	125,234	118,612	56,881	189,620	119,730	125,803	189,511	73,252	42,431
4	139,749	125,261	118,611	56,901	189,673	119,683	125,799	189,554	73,218	42,439
5	139,751	125,258	118,602	56,894	189,648	119,703	125,803	189,461	73,251	42,451
6	139,750	125,251	118,609	56,950	189,630	119,730	125,718	189,545	73,207	42,439
7	139,751	125,264	118,608	56,893	189,684	119,714	125,785	189,539	73,214	42,443
8	139,751	125,302	118,607	56,902	189,645	119,694	125,823	189,503	73,205	42,449
9	139,753	125,291	118,599	56,907	189,639	119,732	125,751	189,539	73,207	42,451
10	139,752	125,280	118,598	56,886	189,626	119,741	125,785	189,529	73,246	42,437
11	139,750	125,261	118,604	56,904	189,575	119,651	125,743	189,481	73,205	42,444
12	139,750	125,278	118,608	56,903	189,694	119,712	125,773	189,592	73,236	42,449
13	139,753	125,297	118,599	56,898	189,612	119,696	125,733	189,488	73,236	42,453

№ варианта										
№	11	12	13	14	15	16	17	18	19	20
14	139,751	125,297	118,606	56,944	189,569	119,704	125,774	189,552	73,227	42,449
15	139,750	125,268	118,598	56,932	189,660	119,708	125,745	189,606	73,239	42,434
16	139,749	125,291	118,608	56,934	189,665	119,756	125,781	189,550	73,208	42,438
17	139,748	125,274	118,602	56,929	189,656	119,701	125,778	189,540	73,206	42,425
18	139,748	125,235	118,602	56,945	189,594	119,659	125,730	189,541	73,208	42,447
19	139,754	125,260	118,609	56,933	189,674	119,679	125,801	189,550	73,267	42,441
20	139,749	125,260	118,602	56,907	189,628	119,731	125,745	189,568	73,236	42,444

№ варианта										
№	21	22	23	24	25	26	27	28	29	30
1	46,918	196,268	63,567	56,574	158,499	99,586	179,138	65,290	53,947	18,977
2	46,914	196,298	63,551	56,582	158,498	99,612	179,227	65,304	53,927	18,958
3	46,917	196,293	63,557	56,529	158,520	99,581	179,175	65,308	53,948	18,927
4	46,917	196,332	63,580	56,558	158,504	99,595	179,127	65,289	53,893	18,956
5	46,920	196,256	63,560	56,564	158,509	99,579	179,219	65,316	53,936	18,981
6	46,917	196,321	63,567	56,529	158,524	99,632	179,236	65,318	53,930	18,911
7	46,917	196,309	63,573	56,573	158,473	99,622	179,187	65,302	53,918	18,968
8	46,913	196,298	63,562	56,565	158,490	99,616	179,233	65,306	53,942	18,967
9	46,916	196,322	63,569	56,595	158,515	99,589	179,238	65,314	53,949	18,984
10	46,916	196,329	63,547	56,639	158,527	99,604	179,240	65,309	53,935	18,954
11	46,913	196,289	63,563	56,491	158,458	99,606	179,192	65,315	53,923	18,999
12	46,915	196,296	63,559	56,614	158,556	99,630	179,162	65,306	53,920	18,988
13	46,915	196,298	63,573	56,584	158,488	99,584	179,253	65,324	53,940	18,982
14	46,919	196,216	63,561	56,565	158,504	99,618	179,228	65,307	53,927	18,953
15	46,918	196,320	63,561	56,585	158,554	99,599	179,121	65,311	53,926	18,939
16	46,915	196,340	63,566	56,504	158,460	99,604	179,178	65,297	53,932	18,986
17	46,913	196,302	63,552	56,577	158,514	99,591	179,151	65,312	53,941	18,998
18	46,913	196,311	63,549	56,505	158,442	99,622	179,306	65,320	53,931	18,966
19	46,914	196,314	63,542	56,549	158,461	99,605	179,202	65,302	53,949	18,937
20	46,916	196,248	63,566	56,515	158,497	99,634	179,235	65,319	53,923	18,985

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Жолмонова
« 14 » _____ 2023 г.



НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ ИЗУЧЕНИЯ СВЯЗИ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Непараметрические методы изучения связи : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: исследовать зависимости между результатами наблюдений непараметрическими методами изучения связи.

2 Задание: Рассчитать коэффициенты корреляции знаков, Спирмена, Кендалла, конкордации и установить значимость этих коэффициентов.

3 Краткие теоретические сведения

Корреляционно-регрессионный метод применим только к количественным признакам. Однако задача измерения связи ставится перед статистикой и по отношению к таким признакам, как пол, образование, занятие, семейное состояние человека, отрасль, форма собственности предприятия, т. е. признакам, не имеющим количественного выражения. Для измерения связи между качественными (атрибутивными) признаками в статистике широко используются: коэффициент сопряженности А. А. Чупрова, коэффициенты ассоциации, контингенции, а также коэффициенты ранговой корреляции Спирмена и Кендалла.

3. 1. Коэффициент корреляции знаков (коэффициент Фехнера)

Простейшим непараметрическим показателем тесноты связи между двумя признаками x и y является коэффициент Фехнера. В основе его расчета лежит принцип сопоставления не абсолютных значений признаков x и y , а их отклонений от среднего уровня. Применение коэффициента Фехнера в практических расчетах основано на предположении, что отклонения эмпирических значений признака x_i от его средней величины $x_i - \bar{X}$ носят случайный характер и должны случайным образом сочетаться с отклонениями эмпирических значений признака y от его среднего уровня $y_i - \bar{Y}$. Соотношение пар совпадений или несовпадений знаков отклонений $x_i - \bar{X}$ и $y_i - \bar{Y}$ позволяет судить о наличии и степени тесноты связи между x и y . Коэффициент Фехнера (K_{Φ}) определяется по формуле следующего вида:

$$K_{\Phi} = \frac{C - H}{C + H}, \quad (1)$$

где C – число совпадений знаков отклонений; H – число несовпадений знаков отклонений. Коэффициент Фехнера может принимать как положительные, так и отрицательные значения в пределах от (-1) до $(+1)$, т. е. $-1 \leq K_{\Phi} \leq +1$.

При $K_{\Phi} = \pm 1$ связь между признаками x и y функциональная.

При $K_{\Phi} = 0$ связь отсутствует.

Промежуточные значения коэффициента Фехнера характеризуют степень тесноты связи между двумя признаками. Знак коэффициента Фехнера свидетельствует о направлении связи между двумя признаками: если $K_{\Phi} \in [-1; 0]$, то связь обратная, т. е. с увеличением или снижением x снижается или увеличивается y ; если $K_{\Phi} \in [0; +1]$, то связь прямая, т. е. с увеличением или снижением x увеличивается или снижается y .

При значении $K_{\Phi} > 0,6$ делается вывод о наличии сильной прямой (обратной) зависимости между признаками.

Пример. Рассчитаем коэффициент Фехнера по данным о деятельности аудиторско-консультационных фирм Москвы в 2001 г., построив для этого табл. 1.

$$K_{\Phi} = \frac{C - H}{C + H} = \frac{8 - 2}{8 + 2} = 0,6.$$

Таблица 1 - Расчетные данные для определения K_{Φ}

№ п/п	Совокупная выручка, млн. руб., y	Общая численность профессионалов, чел., x	Знаки отклонений		
			$x_i - \bar{X}$	$y_i - \bar{Y}$	C
1	2,62	23	- 1,081	- 30,6	1
2	3,04	32	- 0,661	- 21,6	1
3	3,15	50	- 0,551	- 3,6	1
4	3,83	53	+ 0,129	- 0,6	
5	3,58	55	- 0,121	+ 1,4	
6	4,08	58	+ 0,379	+ 4,4	1
7	4,09	59	+ 0,389	+ 5,4	1
8	4,20	62	+ 0,499	+ 8,4	1
9	4,18	69	+ 0,479	+ 15,4	1
10	4,24	75	+ 0,539	+ 21,4	1
Итого	37,01	536			8
Средняя	3,701	53,6			

Таким образом, связь между совокупной выручкой и численностью профессионалов аудиторско-консультационных фирм Москвы прямая и сильная.

Недостатком коэффициента Фехнера, что значительно сокращает возможности его практической реализации, является равенство весов различных по абсолютной величине отклонений фактических значений признаков от их среднего уровня. Кроме того неточность этого показателя проявляется еще и в том, что он учитывает только знаки отклонений, а не числовые значения отклонений.

3.2. Ранговые коэффициенты связи (коэффициенты Спирмена, Кендалла, конкордации)

В ряде случаев, когда факторный признак может иметь не количественное выражение, а атрибутивное (т. е. не имеющее количественного выражения), то в качестве условных обозначений значений признаков и оценки связей между ними также используются ранги и ранговые коэффициенты связи.

Например, на объем реализации товаров оказывают влияние такие факторы, как уровень образования, мода и др., но которые нельзя выразить каким-либо числом.

Ранжирование – это процедура упорядочения объектов изучения, которая выполняется на основе предпочтения.

Ранг – это порядковый номер значений признака, расположенных в порядке возрастания или убывания их величин. Если значения признака имеют одинаковую количественную оценку, то ранг всех этих значений принимается равным средней арифметической от соответствующих номеров мест, которые их определяют. Данные ранги называются *связными*.

Пример. Проранжируем предприятия автомобильной промышленности одного из регионов по величине балансовой прибыли. Наиболее предпочтительному предприятию, величина балансовой прибыли которого наибольшая, присваивается ранг «1»; затем в порядке уменьшения величины балансовой прибыли были проранжированы все рассматриваемые предприятия автомобильной промышленности (табл. 2).

Таблица 2 - Балансовая прибыль предприятий автомобильной промышленности одного из регионов в 1998 г.

Предприятие	Балансовая прибыль, млн. руб.	Ранжирование (ранги)
1	10	6,5
2	12	4
3	10	6,5
4	12	4
5	12	4
6	15	2
7	17	1

Среди непараметрических методов оценки тесноты связи наибольшее значение имеют ранговые коэффициенты ρ – Спирмена и τ – Кендалла. Эти коэффициенты могут быть использованы для определения тесноты связи, как между количественными, так и между качественными признаками при условии, если их значения упорядочить или проранжировать по степени убывания или возрастания признака.

Так, например, можно при помощи одной группы экспертов проранжировать кандидатов на занятие какой-либо должности по степени профессиональной подготовленности, а другую группу экспертов просить проранжировать тех же кандидатов по личностным и этическим качествам, а затем измерить связь между рангами.

3.2.1. Коэффициент корреляции рангов Спирмена

Коэффициент корреляции рангов (коэффициент Спирмена) рассчитывается по формуле (для случая, когда нет связанных рангов)

$$\rho_{x/y} = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)},$$

где d_i^2 – квадрат разности рангов $d_i^2 = (R_x - R_y)^2$, R_x – ранг качественных показателей, R_y – ранг качественных показателей, n – число наблюдений (число пар рангов). Коэффициент Спирмена принимает

любые значения в интервале $[-1; 1]$. Если ранги по обоим признакам совпадают, то $d_i^2=0$, $\rho=1$ и, следовательно, связь полная прямая. Если $\rho=-1$, связь полная обратная, при $\rho=0$ связь между признаками отсутствует. Значимость коэффициента корреляции рангов Спирмена проверяется на основе t -критерия Стьюдента. Расчетное значение критерия определяется по формуле:

$$t_P = \rho_{x/y} \sqrt{\frac{n-2}{1-\rho_{x/y}^2}}.$$

Значение коэффициента корреляции считается статистически существенным, если $t_P > t_{кр}(P; k=n-2)$. Значение $t_{кр}(P=1-q; k=n-2)$ выбирается из справочных таблиц (приложение А).

Пример. Вычислим коэффициент корреляции рангов по данным о стоимости основных фондов и выпуске продукции (млн. руб.). Соответствующие расчеты приведены в табл. 3.

Таблица 3. Расчет коэффициента Спирмена

Наименование фирмы	Стоимость основных фондов x , млн. руб.	Выпуск продукции y , млн.руб.	Знак отклонения от средней арифметической		Ранги			
			$x_i - \bar{X}$	$y_i - \bar{Y}$	по x	по y	Разность рангов	Квадрат разности рангов
«Светлана»	6,0	2,4	-	-	1,0	1,0	0	0
«XXI век»	8,0	4,0	-	-	2,0	3,5	- 1,5	2,25
«Золотой век»	9,0	3,6	-	-	3,0	2,0	+ 1,0	1,00
«Перспектива»	10,0	4,0	-	-	4,5	3,5	+ 1,0	1,00
«Бест»	10,0	4,5	-	-	4,5	5,0	- 0,5	0,25
«Карен»	11,0	4,6	+	-	6,0	6,0	0	0
«Элита»	12,0	5,6	+	+	7,0	8,0	- 1,0	1,00
«Интерстиль»	13,0	6,5	+	+	8,0	9,0	- 1,0	1,00
«Гейзер»	14,0	7,0	+	+	9,0	10,0	- 1,0	1,00
«Олимп»	15,0	5,0	+	+	10,0	7,0	+ 3,0	9,00
Итого	108,0	47,2						16,5
Средняя	10,8	4,72						

Ранги стоимости основных фондов для фирм «Перспектива» и «Бест» определяются как средняя из $\bar{X} = (4,0 + 5,0)/2 = 4,5$, потому что стоимости основных фондов в 10 млн. руб. соответствуют ранги 4 и 5 (это так называемые связанные ранги). Аналогично ранги выпуска продукции для фирмы «XXI век» и «Золотой век» определяют

ся как средняя из $\bar{X} = (3,0 + 4,0)/2 = 3,5$. Итак, коэффициент Спирмена

$$\rho_{x/y} = 1 - \frac{6 \cdot 16,5}{10(10^2 - 1)} = 1 - 0,1 = 0,9.$$

Полученный ранговый коэффициент корреляции свидетельствует о наличии прямой тесной связи между величиной основных фондов и выпуском продукции. Ранговый коэффициент корреляции более точный по сравнению с коэффициентом корреляции знаков, потому что он учитывает не только знаки отклонений, но и место величины признака в данном ряду.

3.2.2. Коэффициент корреляции рангов Кендалла τ

Ранговый коэффициент корреляции Кендалла τ может также использоваться для измерения взаимосвязи между качественными и количественными признаками, характеризующими однородные объекты, ранжированные по одному принципу. Расчет рангового коэффициента Кендалла осуществляется по формуле:

$$\tau_{xy} = 1 - \frac{2S}{n(n-1)},$$

где n – число наблюдений; S – сумма разностей между числом последовательностей и числом инверсий по второму признаку.

Расчет данного коэффициента выполняется в следующей последовательности:

- 1) значения x ранжируются в порядке возрастания или убывания;
- 2) значения y располагаются в порядке, соответствующем значениям x ;
- 3) для каждого ранга y определяется число следующих за ним значений рангов, превышающих его величину. Суммируя таким образом числа, определяют величину P как меру соответствия последовательностей рангов по x и y и учитывают со знаком (+);
- 4) для каждого ранга определяется число следующих за ним рангов, меньших его величины. Суммарная величина обозначается через Q и фиксируется со знаком (–);

5) определяется сумма баллов по всем членам ряда. Этот коэффициент также изменяется в пределах $-1 < \tau < 1$. Он дает несколько более строгую оценку связи, нежели коэффициент Спирмена. Как правило, коэффициент Кендалла меньше коэффициента Спирмена, соотношение между ними:

$$r_s \approx \frac{3}{2} \tau.$$

Это соотношение выполняется *при большом числе наблюдений, $n > 30$* , и слабых либо умеренно тесных связях. Тогда $S = P - Q$. Можно показать, что $P + Q = -n(n-1)$, так что τ может быть представлен как

$$\tau = \frac{P - Q}{P + Q}.$$

Пример (корреляции рангов Кендалла). В приведенном выше примере:

$$P = 9 + 6 + 7 + 6 + 5 + 4 + 2 + 1 + 0 + 0 = 40.$$

$$Q = 0 + (-1) + 0 + 0 + 0 + 0 + 1 + 1 + 1 + 0 = -4.$$

Таким образом,

$$\tau_{xy} = 1 - \frac{2(40 - 4)}{10(10 - 1)} = \frac{72}{90} = 0,8,$$

что также свидетельствует о наличии тесной связи между рассматриваемыми признаками. Хотя $\tau < \rho_{x/y}$ ($\rho_{x/y} = 0,9$), но поскольку связь тесная, соотношение между этими двумя коэффициентами не вполне соответствует упомянутому: коэффициент Спирмена в нашем примере превосходит τ не в 1,5 раза, а на 125 %. Если в изучаемой совокупности есть связанные ранги, то расчеты необходимо проводить по следующей формуле:

$$\tau = \frac{S}{\sqrt{\left[\frac{n(n-1)}{2} - U_x \right] \left[\frac{n(n-1)}{2} - U_y \right]}},$$

где $U_x = \frac{\sum t_x(t_x - 1)}{2}$; $U_y = \frac{\sum t_y(t_y - 1)}{2}$;

t_x, t_y – число связанных рангов в ряду X и Y соответственно.

Рассмотрим расчет коэффициента корреляции рангов Кендалла для случая наличия связанных рангов:

$$P = 9 + 6 + 7 + 6 + 5 + 4 + 2 + 1 + 0 + 0 = 40.$$

10

$$Q = 0 + (-1) + 0 + 0 + 0 + 0 + (-1) + (-1) + (-1) + 0 = -4.$$

$$U_x = (2(2 - 1))/2 = 1. \quad U_y = (2(2 - 1))/2 = 1.$$

$$\tau = \frac{40 - 4}{\sqrt{\left[\frac{10(10-1)}{2} - 1\right] \left[\frac{10(10-1)}{2} - 1\right]}} = 0,82,$$

что свидетельствует о существенной связи между номинальной стоимостью основных фондов и выпуском продукции. Связь между признаками можно признать статистически значимой, если значения коэффициентов ранговой корреляции Спирмена и Кендалла больше 0,5.

3.2.3. Множественный коэффициент ранговой корреляции (коэффициент конкордации W)

Для определения тесноты связи между произвольным числом ранжированных признаков применяется *множественный коэффициент ранговой корреляции (коэффициент конкордации) (W)*, который вычисляется по формуле:

$$W = \frac{12S}{m^2(n^3 - n)},$$

где m – количество факторов; n – число наблюдений; S – отклонение суммы квадратов рангов от средней квадратов рангов.

Пример. Коэффициент конкордации. Определим тесноту связи между уставным капиталом, числом выставленных акций и числом занятых на предприятиях, выставивших акции на аукционы в 2001 г. (табл. 4):

$$S = 2863 - \frac{165^2}{10} = 2863 - 2722,5 = 140,5,$$

$$W = \frac{12S}{m^2(n^3 - n)} = \frac{12 \cdot 140,5}{(91000 - 10)} = 0,018.$$

Таблица 4 -Расчет коэффициента конкордации

Номер	Уставный	Число вы-	Число заня-	R_x	R_y	R_z	Сумма	Квадраты
-------	----------	-----------	-------------	-------	-------	-------	-------	----------

предприятия	капитал, тыс. руб., x	ставлен- ных акций, y	тых на пред- приятия, z				строк	сумм
1	29540	856	119	9	7	1	17	289
2	16050	930	125	1	9	2	12	144
3	41020	1563	132	10	10	3	23	529
4	23500	682	141	6	5	4	15	225
5	26250	616	150	7	3	5	15	225
6	17950	495	165	4	2	6	12	144
7	28130	815	178	8	6	7	21	441
8	17510	858	181	3	8	8	19	361
9	17000	467	201	2	1	9	12	144
10	22640	661	204	5	4	10	19	361
Итого	–	–	–	–	–	–	165	2863

Значимость коэффициента конкордации проверяется на основе χ^2 -критерия Пирсона:

$$\chi_p^2 = \frac{12S}{m \cdot n(n-1)}.$$

Для нашего примера:

$$\chi_p^2 = \frac{12 \cdot 140,5}{3 \cdot 10(10-1)} = 6,24.$$

Расчетное значение $\chi_p^2 = 6,24$ меньше $\chi_{кр}^2 = 16,919$ ($P=1-q=0,05$, $k=n-1=9$), что подтверждает незначимость коэффициента конкордации и свидетельствует о слабой связи между рассматриваемыми признаками. В случае наличия связанных рангов коэффициент конкордации определяется по формуле:

$$W = \frac{S}{1/12 m^2 (n^3 - n) - m \sum_{j=1}^m T_j},$$

где $T_j = 1/12 m \sum_{j=1}^m (t_j^3 - t_j)$; t_j – количество связанных рангов по отдель-

ным показателям. Проверка значимости осуществляется по формуле:

$$\chi_p^2 = \frac{S}{1/12 m \cdot n(n-1) - \frac{1}{n-1} \sum_{j=1}^m T_j}.$$

Коэффициент конкордации принимает любые значения в интервале $[-1; 1]$.

Пример. Коэффициент конкордации (в случае наличия связанных рангов). По данным предприятий нефтеперерабатывающей промышленности определим зависимость прибыли от реализации, от среднегодовой стоимости основных производственных фондов и объема валовой продукции (табл. 5):

$$S = 1229,5 - \frac{84^2}{7} = 1229,5 - 1008 = 221,5$$

$$T_x = \frac{1}{12} [(2^3 - 2) + (2^3 - 2)] = 1;$$

$$T_y = \frac{1}{12} [(2^3 - 2) + (2^3 - 2)] = 1;$$

$$T_z = \frac{1}{12} [(2^3 - 2) + (2^3 - 2)] = 1;$$

$$\sum T_j = T_x + T_y + T_z = 1 + 1 + 1 = 3;$$

$$W = \frac{S}{1/12 m^2 (n^3 - n) - m \sum_{j=1}^m T_j} = \frac{221,5}{\frac{1}{12} 3^2 (7^3 - 7) - 3 \cdot 3} = \frac{221,5}{243,0} = 0,91.$$

Таблица 5 - Расчет коэффициента конкордации (в случае наличия связанных рангов)

Номер предприятия	Прибыль от реализации, млн. руб. x	Объем валовой продукции, млрд. руб. y	Среднегодовая стоимость ОПФ, млрд. руб. z	R_x	R_y	R_z	Сумма строк	Квадраты сумм
1	40	1,7	0,27	1,5	1	1,5	4	16
2	75	3,2	0,55	3	5	4	12	144
3	82	2,9	0,97	4,5	3,5	5,5	13,5	182,25
4	40	1,8	0,27	1,5	2	1,5	5	25
5	106	11,8	0,98	6	6,5	7	19,5	380,25
6	82	2,9	0,35	4,5	3,5	3	11	121
7	109	11,8	0,97	7	6,5	5,5	19	361
Итого	-	-	-	28	28	28	84	1229,5

Расчетное значение χ^2 -критерия Пирсона для проверки значимости коэффициента конкордации по данным нашего примера составило:

$$\chi_p^2 = \frac{221,5}{1/12 \cdot 3 \cdot 7(7-1) - \frac{1}{7-1} \cdot 3} = 22,15.$$

Расчетное значение $\chi_p^2 = 22,15$ больше $\chi_{кр}^2 = 12,592$, ($P=1-q=0,05$; $k=n-1=6$), что подтверждает значимость коэффициента конкордации и свидетельствует о сильной связи между рассматриваемыми признаками. Преимуществом ранговых коэффициентов корреляции Спирмена, Кендалла и конкордации является то, что с их помощью можно измерять и оценивать связи как между количественными, так и между атрибутивными признаками, которые поддаются ранжированию.

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение В), студент изучает теоретические сведения согласно пункту 3. Далее выполняет расчеты аналогичные в рассмотренных примерах с учетом имеющихся особенностей задания.

В отчёте по практической работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое корреляция?
2. Как рассчитывается коэффициент корреляции знаков.
3. Как рассчитывается коэффициент Спирмена.
4. Как рассчитывается коэффициент Кендалла.
5. Как рассчитывается коэффициент конкордации.

6. Как установить значимость коэффициента корреляции знаков.
7. Как установить значимость коэффициента Спирмена.
8. Как установить значимость коэффициента Кендалла.
9. Как установить значимость коэффициента конкордации.

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
Распределение Стьюдента (t_p)

k	P											
	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,99
1	0,079	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	63,657
2	0,071	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	9,925
3	0,068	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	5,841
4	0,067	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	4,604
5	0,066	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	4,032
6	0,065	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,707
7	0,065	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	3,499
8	0,065	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	3,355
9	0,064	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	3,250
10	0,064	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	3,169
11	0,064	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	3,106
12	0,064	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	3,055
13	0,064	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	3,012
14	0,064	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,977
15	0,064	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,947
16	0,064	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,921
17	0,064	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,898
18	0,064	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,878
19	0,064	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,861
20	0,063	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,845
21	0,063	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,831
22	0,063	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,819
23	0,063	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,807
24	0,063	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,797
25	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,787
26	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,779
27	0,063	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,771
28	0,063	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,763
29	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,756
30	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,750
300	0,063	0,126	0,254	0,386	0,525	0,675	0,843	1,038	1,284	1,650	1,968	2,592

ПРИЛОЖЕНИЕ Б

Значения $\chi_{k;P}^2$ для различных значений k и P

k	P												
	0,01	0,02	0,05	0,1	0,2	0,3	0,5	0,7	0,8	0,9	0,95	0,98	0,99
1	0,000157	0,000628	0,00393	0,0158	0,0642	0,148	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,02	0,04	0,103	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	7,824	9,21
3	0,115	0,185	0,352	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,345
4	0,297	0,429	0,711	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	0,554	0,752	1,145	1,61	2,343	3	4,351	6,064	7,289	9,236	11,07	13,388	15,086
6	0,872	1,134	1,635	2,204	3,07	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	1,239	1,564	2,167	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,622	18,475
8	1,647	2,032	2,733	3,49	4,594	5,527	7,344	9,524	11,03	13,362	15,507	18,168	20,09
9	2,088	2,532	3,325	4,168	5,38	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	2,558	3,059	3,94	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	3,053	3,609	4,575	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	3,571	4,178	5,226	6,304	7,807	9,034	11,34	14,011	15,812	18,549	21,026	24,054	26,217
13	4,107	4,765	5,892	7,041	8,634	9,926	12,34	15,119	16,985	19,812	22,362	25,471	27,688
14	4,66	5,368	6,571	7,79	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141
15	5,229	5,985	7,261	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	5,812	6,614	7,962	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	29,633	32
17	6,408	7,255	8,672	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	7,015	7,906	9,39	10,865	12,857	14,44	17,338	20,601	22,76	25,989	28,869	32,346	34,805
19	7,633	8,567	10,117	11,651	13,716	15,352	18,338	21,689	23,9	27,204	30,144	33,687	36,191
20	8,26	9,237	10,851	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,41	35,02	37,566
21	8,897	9,915	11,591	13,24	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	9,542	10,6	12,338	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,659	40,289
23	10,196	11,293	13,091	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	10,856	11,992	13,848	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,27	42,98
25	11,524	12,697	14,611	16,473	18,94	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	12,198	13,409	15,379	17,292	19,82	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	12,878	14,125	16,151	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	44,14	46,963
29	14,256	15,574	17,708	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	49,588

<i>k</i>	P												
	0,01	0,02	0,05	0,1	0,2	0,3	0,5	0,7	0,8	0,9	0,95	0,98	0,99
30	14,953	16,306	18,493	20,599	23,364	25,508	29,336	33,53	36,25	40,256	43,773	47,962	50,892
31	15,655	17,042	19,281	21,434	24,255	26,44	30,336	34,598	37,359	41,422	44,985	49,226	52,191
32	16,362	17,783	20,072	22,271	25,148	27,373	31,336	35,665	38,466	42,585	46,194	50,487	53,486
33	17,073	18,527	20,867	23,11	26,042	28,307	32,336	36,731	39,572	43,745	47,4	51,743	54,775
34	17,789	19,275	21,664	23,952	26,938	29,242	33,336	37,795	40,676	44,903	48,602	52,995	56,061
35	18,509	20,027	22,465	24,797	27,836	30,178	34,336	38,859	41,778	46,059	49,802	54,244	57,342
36	19,233	20,783	23,269	25,643	28,735	31,115	35,336	39,922	42,879	47,212	50,998	55,489	58,619
37	19,96	21,542	24,075	26,492	29,635	32,053	36,336	40,984	43,978	48,363	52,192	56,73	59,893
38	20,691	22,304	24,884	27,343	30,537	32,992	37,335	42,045	45,076	49,513	53,384	57,969	61,162
39	21,426	23,069	25,695	28,196	31,441	33,932	38,335	43,105	46,173	50,66	54,572	59,204	62,428
40	22,164	23,838	26,509	29,051	32,345	34,872	39,335	44,165	47,269	51,805	55,758	60,436	63,691
41	22,906	24,609	27,326	29,907	33,251	35,813	40,335	45,224	48,363	52,949	56,942	61,665	64,95
42	23,65	25,383	28,144	30,765	34,157	36,755	41,335	46,282	49,456	54,09	58,124	62,892	66,206
43	24,398	26,159	28,965	31,625	35,065	37,698	42,335	47,339	50,548	55,23	59,304	64,116	67,459
44	25,148	26,939	29,787	32,487	35,974	38,641	43,335	48,396	51,639	56,369	60,481	65,337	68,71
45	25,901	27,72	30,612	33,35	36,884	39,585	44,335	49,452	52,729	57,505	61,656	66,555	69,957
46	26,657	28,504	31,439	34,215	37,795	40,529	45,335	50,507	53,818	58,641	62,83	67,771	71,201
47	27,416	29,291	32,268	35,081	38,708	41,474	46,335	51,562	54,906	59,774	64,001	68,985	72,443
48	28,177	30,08	33,098	35,949	39,621	42,42	47,335	52,616	55,993	60,907	65,171	70,197	73,683
49	28,941	30,871	33,93	36,818	40,534	43,366	48,335	53,67	57,079	62,038	66,339	71,406	74,919
50	29,707	31,664	34,764	37,689	41,449	44,313	49,335	54,723	58,164	63,167	67,505	72,613	76,154
51	30,475	32,459	35,6	38,56	42,365	45,261	50,335	55,775	59,248	64,295	68,669	73,818	77,386
52	31,246	33,256	36,437	39,433	43,281	46,209	51,335	56,827	60,332	65,422	69,832	75,021	78,616
53	32,019	34,055	37,276	40,308	44,199	47,157	52,335	57,879	61,414	66,548	70,993	76,223	79,843
54	32,793	34,856	38,116	41,183	45,117	48,106	53,335	58,93	62,496	67,673	72,153	77,422	81,069
55	33,571	35,659	38,958	42,06	46,036	49,055	54,335	59,98	63,577	68,796	73,311	78,619	82,292
56	34,35	36,464	39,801	42,937	46,955	50,005	55,335	61,031	64,658	69,919	74,468	79,815	83,514
57	35,131	37,27	40,646	43,816	47,876	50,956	56,335	62,08	65,737	71,04	75,624	81,009	84,733
58	35,914	38,078	41,492	44,696	48,797	51,906	57,335	63,129	66,816	72,16	76,778	82,201	85,95

ПРИЛОЖЕНИЕ В

Исходные данные к работе

Таблица 1

	№ варианта														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>q</i>	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05

Таблица 2

Вар. 1		Вар. 2		Вар. 3		Вар. 4		Вар. 5		Вар. 6		Вар. 7		Вар. 8	
<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>
83	4,0	65	2,0	80	1,0	94	10,0	27	2,0	100	4,0	64	3,0	57	2,0
28	5,0	77	10,0	74	2,0	25	5,0	36	10,0	54	8,0	63	7,0	85	1,0
24	3,0	53	6,0	91	4,0	49	2,0	12	7,0	85	9,0	32	3,0	61	9,0
58	3,0	18	2,0	26	2,0	78	10,0	70	5,0	42	7,0	43	8,0	94	7,0
92	10,0	58	4,0	90	1,0	77	10,0	59	7,0	99	9,0	14	10,0	40	5,0
98	1,0	33	1,0	71	10,0	49	5,0	38	6,0	25	10,0	11	8,0	87	8,0
69	8,0	71	1,0	17	9,0	72	2,0	94	9,0	85	10,0	93	5,0	78	5,0
10	6,0	100	10,0	80	9,0	87	7,0	43	4,0	61	10,0	72	6,0	40	10,0
92	2,0	24	4,0	94	5,0	22	6,0	80	2,0	11	8,0	12	4,0	32	7,0
83	10,0	10	2,0	94	6,0	34	6,0	72	8,0	75	9,0	88	7,0	99	5,0
94	4,0	24	1,0	87	1,0	57	10,0	91	3,0	88	8,0	21	6,0	80	3,0
65	8,0	40	3,0	13	2,0	17	3,0	26	8,0	37	5,0	99	10,0	23	10,0
10	7,0	88	5,0	82	2,0	18	3,0	10	8,0	57	4,0	44	1,0	50	8,0
53	10,0	29	10,0	94	1,0	29	6,0	68	5,0	44	5,0	10	6,0	69	5,0
55	10,0	81	8,0	73	6,0	72	4,0	100	3,0	25	3,0	61	6,0	35	8,0

Вар. 9		Вар. 10		Вар. 11		Вар. 12		Вар. 13		Вар. 14		Вар. 15	
<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>	<i>X</i>	<i>Y</i>
36	10,0	45	10,0	67	5,0	59	10,0	64	7,0	82	6,0	61	9,0
92	4,0	77	6,0	91	6,0	41	10,0	66	7,0	78	1,0	26	6,0
37	8,0	75	4,0	30	1,0	30	9,0	93	6,0	93	3,0	94	4,0
47	6,0	17	5,0	61	4,0	67	9,0	74	6,0	47	5,0	45	7,0
53	8,0	63	3,0	75	6,0	34	5,0	73	10,0	61	4,0	20	9,0
48	4,0	38	4,0	37	1,0	56	8,0	71	2,0	39	5,0	55	6,0
81	7,0	10	7,0	59	9,0	40	9,0	10	5,0	61	2,0	88	8,0
22	2,0	63	1,0	38	6,0	67	4,0	33	2,0	69	6,0	18	3,0
76	9,0	52	2,0	52	5,0	30	2,0	80	6,0	74	3,0	70	6,0
55	10,0	40	2,0	63	5,0	33	10,0	28	3,0	14	10,0	16	8,0
56	2,0	84	7,0	92	9,0	11	3,0	92	2,0	97	2,0	56	6,0
52	8,0	39	1,0	53	5,0	98	4,0	67	10,0	76	1,0	87	10,0
94	3,0	11	9,0	95	2,0	15	9,0	49	7,0	85	9,0	92	2,0
42	2,0	11	10,0	30	9,0	12	2,0	83	4,0	61	4,0	42	7,0
44	5,0	72	9,0	57	8,0	68	4,0	37	3,0	25	1,0	32	9,0

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Жолтимонова
« 14 » _____ 20__ г.



РЕГРЕССИОННЫЙ АНАЛИЗ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Регрессионный анализ : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: исследовать зависимости между результатами наблюдений и определить их характер.

2 Задание: построить диаграммы рассеивания, рассчитать коэффициенты корреляции для результатов наблюдения трёх случайных величин, а установить значимость этих коэффициентов.

3 Краткие теоретические сведения

3.1 Линейная регрессия

Регрессионный анализ позволяет приближенно определить форму связи между результативным и факторными признаками, а также решить вопрос о том, значима ли эта связь. Вид функции, с помощью которой приближенно выражается форма связи, выбирают заранее, исходя из содержательных соображений или визуального анализа данных. Математическое решение задачи основано на методе наименьших квадратов, с помощью которого определяют степень приближения полученной функции к опытным данным.

В вычислительном аспекте метод наименьших квадратов сводится к составлению и решению системы так называемых **нормальных уравнений**. Исходным этапом для этого является подбор вида функции, отображающей статистическую связь.

Тип функции в каждом конкретном случае можно подобрать путем прикидки на графике исходных данных подходящей, т. е. достаточно хорошо приближающей эти данные, линии, которая может быть записана в виде

$$Y = a_0 + a_1 X, \quad (1)$$

где Y – результативный признак или зависимая переменная; X – факторный признак или независимая переменная; a_0 и a_1 – параметры уравнения, которые могут быть найдены методом наименьших квадратов.

Для нахождения искомых параметров нужно составить систему уравнений, которая в данном случае будет иметь вид

$$\begin{cases} na_0 + a_1 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i \\ a_0 \sum_{i=1}^n X_i + a_1 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i \end{cases} \quad (2)$$

Полученная система может быть решена методом Гаусса. Искомые параметры системы из двух нормальных уравнений можно вычислить и непосредственно с помощью последовательного использования нижеприведенных формул:

$$a_1 = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}, \quad a_0 = \bar{Y} - a_1 \bar{X}, \quad (3)$$

где Y_i , – i -е значение результативного признака; X_i – i -е значение факторного признака; \bar{Y} и \bar{X} – средние арифметические результативного и факторного признаков соответственно; n – число значений признака Y_i , или, что то же самое, число значений признака X_i .

Уравнение регрессии не только определяет форму анализируемой связи, но и показывает, в какой степени изменение одного признака сопровождается изменением другого признака.

Коэффициент при X (a_1), называемый **коэффициентом регрессии**, показывает, на какую величину в среднем изменяется результативный признак Y при изменении факторного признака X на единицу.

3.2 Нелинейная регрессия

Не всегда связь между признаками может быть достаточно хорошо представлена линейной функцией. Иногда для описания существующей связи более пригодными, а порой и единственно возможными являются более сложные нелинейные функции. Ограничимся рассмотрением наиболее простых из них.

Одним из простейших видов нелинейной зависимости является парабола, которая в общем виде может быть представлена функцией

$$Y = a_0 + a_1 X + a_2 X^2. \quad (4)$$

Неизвестные параметры a_0 , a_1 , a_2 находятся в результате решения следующей системы уравнений:

$$\begin{cases} na_0 + a_1 \sum_{i=1}^n X_i + a_2 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n Y_i, \\ a_0 \sum_{i=1}^n X_i + a_1 \sum_{i=1}^n X_i^2 + a_2 \sum_{i=1}^n X_i^3 = \sum_{i=1}^n Y_i X_i, \\ a_0 \sum_{i=1}^n X_i^2 + a_1 \sum_{i=1}^n X_i^3 + a_2 \sum_{i=1}^n X_i^4 = \sum_{i=1}^n Y_i X_i^2. \end{cases} \quad (5)$$

Систему уравнений (5) можно представить в виде:

$$\begin{cases} g_{00}a_0 + g_{01}a_1 + g_{02}a_2 = h_0; \\ g_{10}a_0 + g_{11}a_1 + g_{12}a_2 = h_1; \\ g_{20}a_0 + g_{21}a_1 + g_{22}a_2 = h_2; \end{cases} \quad (6)$$

где

$$\begin{aligned} g_{00} &= n; \quad g_{01} = \sum_{i=1}^n X_i; \quad g_{02} = \sum_{i=1}^n X_i^2; \quad h_0 = \sum_{i=1}^n Y_i; \\ g_{10} &= \sum_{i=1}^n X_i; \quad g_{11} = \sum_{i=1}^n X_i^2; \quad g_{12} = \sum_{i=1}^n X_i^3; \quad h_1 = \sum_{i=1}^n X_i Y_i; \\ g_{20} &= \sum_{i=1}^n X_i^2; \quad g_{21} = \sum_{i=1}^n X_i^3; \quad g_{22} = \sum_{i=1}^n X_i^4; \quad h_2 = \sum_{i=1}^n X_i^2 Y_i \end{aligned} \quad (7)$$

Решение системы (6) также может быть найдено с помощью правила Крамера:

$$a_0 = \frac{\Delta_0}{\Delta}, \quad a_1 = \frac{\Delta_1}{\Delta}, \quad a_2 = \frac{\Delta_2}{\Delta}, \quad (8)$$

где Δ – определитель системы (6), составленный из коэффициентов при неизвестных:

$$\begin{aligned} \Delta &= \begin{vmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{vmatrix} = \\ &= g_{00}g_{11}g_{22} + g_{10}g_{21}g_{02} + g_{01}g_{12}g_{20} - \\ &\quad - g_{20}g_{11}g_{02} - g_{10}g_{01}g_{22} - g_{21}g_{12}g_{00}. \end{aligned} \quad (9)$$

Определители Δ_0 , Δ_1 и Δ_2 получаются из определителя Δ путем замены столбца коэффициентов при соответствующих переменных столбцом правых частей системы:

$$\Delta_0 = \begin{vmatrix} h_0 & g_{01} & g_{02} \\ h_1 & g_{11} & g_{12} \\ h_2 & g_{21} & g_{22} \end{vmatrix} = \begin{aligned} &= h_0 g_{11} g_{22} + h_1 g_{21} g_{02} + g_{01} g_{12} h_2 - \\ &\quad - h_2 g_{11} g_{02} - h_1 g_{01} g_{22} - g_{21} g_{12} h_0, \end{aligned} \quad (10)$$

$$\Delta_1 = \begin{vmatrix} g_{00} & h_0 & g_{02} \\ g_{10} & h_1 & g_{12} \\ g_{20} & h_2 & g_{22} \end{vmatrix} = \begin{aligned} &= g_{00} h_1 g_{22} + g_{10} h_2 g_{02} + h_0 g_{12} g_{20} - \\ &\quad - g_{20} h_1 g_{02} - g_{10} h_0 g_{22} - h_2 g_{12} g_{00}, \end{aligned} \quad (11)$$

$$\Delta_2 = \begin{vmatrix} g_{00} & g_{01} & h_0 \\ g_{10} & g_{11} & h_1 \\ g_{20} & g_{21} & h_2 \end{vmatrix} = \begin{aligned} &= g_{00} g_{11} h_2 + g_{10} g_{21} h_0 + g_{01} h_1 g_{20} - \\ &\quad - g_{20} g_{11} h_0 - g_{10} g_{01} h_2 - g_{21} h_1 g_{00}. \end{aligned} \quad (12)$$

Дает ли преимущества описание связи с помощью параболы по сравнению с описанием, построенным по гипотезе линейности? Ответ на этот вопрос можно получить, рассчитав последовательный F-критерий.

На практике для изучения связей используются полиномы более высоких порядков (3-го и 4-го порядков). Составление системы, ее решение, а также решение вопроса о полезности повышения порядка функции для этих случаев аналогичны описанным. При этом никаких принципиально новых моментов не возникает, но существенно увеличивается объем расчетов.

Кроме класса парабол для анализа нелинейных связей можно применять и другие виды функций. Для расчета неизвестных параметров этих функций рекомендуется использовать метод наименьших квадратов, как наиболее мощный и широко применяемый.

Однако метод наименьших квадратов не универсален, поскольку он может использоваться только при условии, что выбранные для выравнивания функции линейны по отношению к своим параметрам. Не все функции удовлетворяют этому условию, но большинство применяемых на практике с помощью специальных преобразований могут быть приведены к стандартной форме функции с линейными параметрами.

3.3 Коэффициент детерминации

При регрессионном анализе важным является вопрос - в какой степени величина X определяет величину Y ? На этот вопрос можно ответить, рассчитав, какая часть вариации результативного признака Y может быть объяснена влиянием факторного признака X .

Рассмотрим отношение

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}. \quad (15)$$

Оно показывает долю разброса, учитываемого регрессией, в общем разбросе результативного признака и носит название **коэффициента детерминации**. Этот показатель, равный отношению факторной вариации к полной вариации признака, позволяет судить о том, насколько «удачно» выбран вид функции. Проведя расчеты, основанные на одних и тех же исходных данных, для нескольких типов функций, мы можем из них выбрать такую, которая дает наибольшее значение R^2 и, следовательно, в большей степени, чем другие функции, объясняет вариацию результативного признака. Действительно, при расчете R^2 для одних и тех же данных, но разных функций знаменатель выражения (15) остается неизменным, а числитель показывает ту часть вариации результативного признака, которая учитывается выбранной функцией. Чем больше R^2 , т. е. чем больше числитель, тем больше изменение факторного признака объясняет изменение результативного признака и тем, следовательно, лучше уравнение регрессии, лучше выбор функции.

Коэффициент детерминации всегда находится в пределах интервала $[0;1]$. Если значение R^2 близко к единице, это означает, что построенная модель объясняет почти всю изменчивость соответствующих переменных. И наоборот, значение R^2 близкое к нулю, означает плохое качество построенной модели.

Коэффициент детерминации R^2 показывает, на сколько процентов $R^2 \cdot 100\%$ найденная функция регрессии описывает связь между исходными значениями Y и X . На рис. 1 показана $(\hat{Y}_i - \bar{Y})$ – объясненная регрессионной моделью вариация и $(Y_i - \bar{Y})$ – общая вариация. Соответственно, величина $(1 - R^2) \cdot 100\%$ показывает, сколько процентов вариации параметра Y обусловлены факторами, не включенными в регрессионную модель.

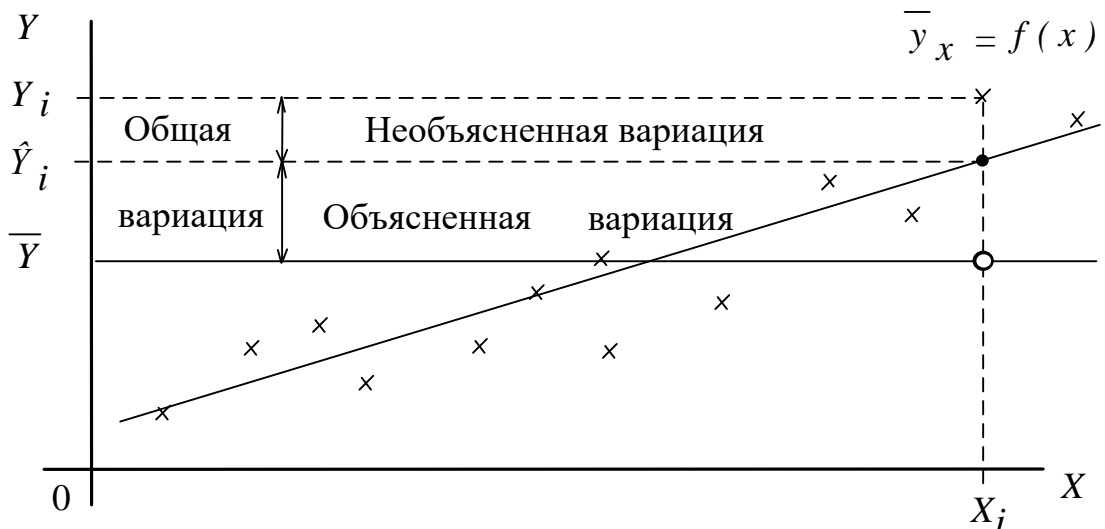


Рис. 1. Графическая интерпретация коэффициента детерминации (линейная регрессия)

При высоком значении коэффициента детерминации $R^2 \geq 75\%$ можно делать прогноз $Y^* = f(X^*)$ для конкретного значения X^* в пределах диапазона исходных данных. При прогнозах значений, не входящих в диапазон исходных данных, справедливость полученной модели гарантировать нельзя. Это объясняется тем, что может проявиться влияние новых факторов, которые модель не учитывает.

3.4 Критерий значимости регрессии

Критерий значимости регрессии позволяет дать ответы на вопросы, существует ли связь? Значимо ли уравнение регрессии, используемое для отображения предполагаемой связи?

Мерой значимости линии регрессии может служить следующее соотношение

$$F[m, n - m - 1] = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m}$$

где n – число значений X или, что то же самое, значений Y ; m – число факторных признаков (независимых переменных, в нашем случае $m=1$).

Действительно, связь тем больше, чем значительнее мера рассеяния признака, обусловленная регрессией, превосходит меру рассеяния отклонений фактических значений от выравненных.

Соотношение (13) позволяет решить вопрос о значимости регрессии. Регрессия значима, т. е. между признаками существует связь, если для данного уровня значимости вычисленное значение $F[m, n-m-1]$ превышает критическое значение $F_{кр}[m, n-m-1]$, стоящее на пересечении m -го столбца и $[n-m-1]$ -й строки таблицы Фишера (Приложение А).

3.5 Пример расчета

Для 20 парных наблюдений X и Y (табл. 1) построить уравнение линейной регрессии и нелинейной (полином второй степени), сделать вывод о значимости полученных уравнений (при $q=0,05$) и о том, какое уравнение лучше описывает взаимосвязь между параметрами X и Y .

Таблица 1.

№ наб-я	1	2	3	4	5	6	7	8	9	10
X_i	1,338	1,803	0,844	0,933	0,604	1,649	0,543	0,071	1,349	0,919
Y_i	9,392	13,124	5,932	6,52	4,492	11,945	4,191	1,844	9,477	6,336
№ наб-я	1	2	3	4	5	6	7	8	9	10
X_i	0,794	1,909	0,257	1,817	1,219	1,062	0,046	1,674	0,839	0,026
Y_i	5,577	14,199	2,717	13,332	8,512	7,406	1,763	12,117	6,02	1,687

Для построения линейной регрессии заполним следующую таблицу (табл. 2).

Таблица 2

№	X_i	Y_i	X_i^2	$X_i Y_i$	\hat{Y}_i	$Y_i - \hat{Y}_i$	$(Y_i - \hat{Y}_i)^2$	$Y_i - \bar{Y}$	$(Y_i - \bar{Y})^2$
1	1,338	9,392	1,7902	12,566	9,664	-0,272	0,074	2,063	4,255
2	1,803	13,124	3,2508	23,663	12,738	0,386	0,149	5,795	33,580
3	0,844	5,932	0,7123	5,007	6,398	-0,466	0,217	-1,397	1,952
4	0,933	6,52	0,8705	6,083	6,987	-0,467	0,218	-0,809	0,655
5	0,604	4,492	0,3648	2,713	4,812	-0,320	0,102	-2,837	8,049
6	1,649	11,945	2,7192	19,697	11,720	0,225	0,051	4,616	21,306
7	0,543	4,191	0,2948	2,276	4,408	-0,217	0,047	-3,138	9,848
8	0,071	1,844	0,0050	0,131	1,288	0,556	0,309	-5,485	30,087
9	1,349	9,477	1,8198	12,784	9,737	-0,260	0,068	2,148	4,613
10	0,919	6,336	0,8446	5,823	6,894	-0,558	0,312	-0,993	0,986
11	0,794	5,577	0,6304	4,428	6,068	-0,491	0,241	-1,752	3,070
12	1,909	14,199	3,6443	27,106	13,439	0,760	0,577	6,870	47,195
13	0,257	2,717	0,0660	0,698	2,518	0,199	0,040	-4,612	21,272
14	1,817	13,332	3,3015	24,224	12,831	0,501	0,251	6,003	36,034
15	1,219	8,512	1,4860	10,376	8,877	-0,365	0,134	1,183	1,399
16	1,062	7,406	1,1278	7,865	7,840	-0,434	0,188	0,077	0,006
17	0,046	1,763	0,0021	0,081	1,123	0,640	0,410	-5,566	30,982
18	1,674	12,117	2,8023	20,284	11,885	0,232	0,054	4,788	22,924
19	0,839	6,02	0,7039	5,051	6,365	-0,345	0,119	-1,309	1,714
20	0,026	1,687	0,0007	0,044	0,990	0,697	0,485	-5,642	31,834
Σ	19,696	146,58	26,437	190,9			4,045		311,762
	$\bar{X} = 0,985$	$\bar{Y} = 7,329$							

Далее рассчитаем коэффициенты линейного уравнения

$$a_1 = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} = \frac{20 \cdot 190,9 - 19,696 \cdot 146,58}{20 \cdot 26,437 - 19,696^2} = 6,61;$$

$$a_0 = \bar{Y} - a_1 \bar{X} = 7,329 - 6,61 \cdot 0,985 = 0,82.$$

Построим диаграмму рассеивания экспериментальных данных и линию полученного уравнения регрессии $Y=0,82+6,61X$ (рис. 2). На рисунке хорошо заметны отклонения экспериментальных точек от линии регрессии.

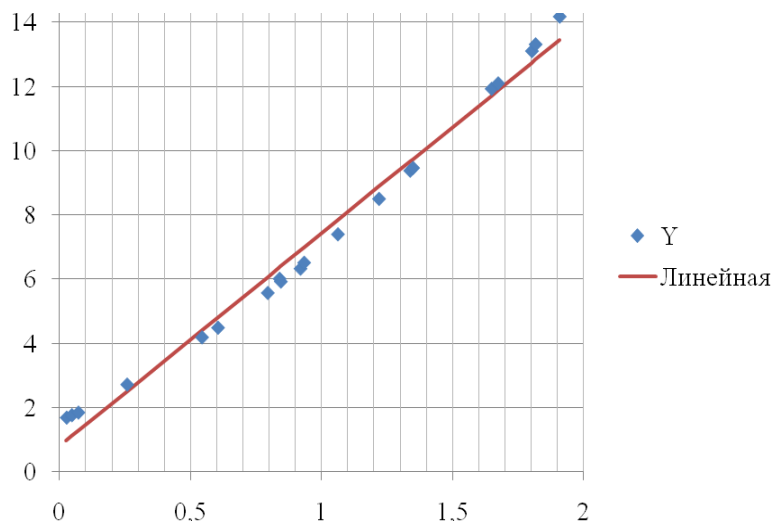


Рис. 2 - Диаграмма рассеивания экспериментальных данных и линия регрессии $Y=0,82+6,61X$

Выполним оценку значимости линейной модели, рассчитаем коэффициент детерминации

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = 1 - \frac{4,045}{311,762} = 0,987$$

и значение F -критерия

$$F[1,18] = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} = \frac{0,987}{1 - 0,987} \cdot \frac{20 - 1 - 1}{1} \approx 1369$$

Из таблицы приложения А при $q=0,05$ и $k_1=1$ и $k_2=18$ выбираем критическое значение F -критерия

$$F_{кр}[1,18] = 4,41.$$

Т.к. $F[1,18] = 1369 > F_{кр}[1,18] = 4,41$, то полученное уравнение можно признать значимым.

Далее выполним построения уравнения регрессии в виде полинома второй степени, для этого заполним следующую таблицу (табл. 3).

Таблица 3

№	X_i	Y_i	X_i^2	X_i^3	X_i^4	$X_i^2 Y$	\hat{Y}_i	$Y_i - \hat{Y}_i$	$(Y_i - \hat{Y}_i)^2$
1	1,338	9,392	1,7902	2,395	3,205	16,814	9,387	0,005	0,00002
2	1,803	13,124	3,2508	5,861	10,568	42,664	13,217	-0,093	0,00862
3	0,844	5,932	0,7123	0,601	0,507	4,226	5,948	-0,016	0,00025
4	0,933	6,52	0,8705	0,812	0,758	5,676	6,520	$3,6 \cdot 10^{-4}$	$1,3 \cdot 10^{-7}$
5	0,604	4,492	0,3648	0,220	0,133	1,639	4,510	-0,018	0,00034
6	1,649	11,945	2,7192	4,484	7,394	32,481	11,885	0,060	0,00359
7	0,543	4,191	0,2948	0,160	0,087	1,236	4,169	0,022	0,00047
8	0,071	1,844	0,0050	$3,6 \cdot 10^{-4}$	$2,5 \cdot 10^{-5}$	0,009	1,865	-0,021	0,00042
9	1,349	9,477	1,8198	2,455	3,312	17,246	9,471	0,006	0,00003
10	0,919	6,336	0,8446	0,776	0,713	5,351	6,428	-0,092	0,00852
11	0,794	5,577	0,6304	0,501	0,397	3,516	5,636	-0,059	0,00344
12	1,909	14,199	3,6443	6,957	13,281	51,745	14,170	0,029	0,00084
13	0,257	2,717	0,0660	0,017	0,004	0,179	2,702	0,015	0,00022
14	1,817	13,332	3,3015	5,999	10,900	44,015	13,341	-0,009	0,00008
15	1,219	8,512	1,4860	1,811	2,208	12,649	8,500	0,012	0,00015
16	1,062	7,406	1,1278	1,198	1,272	8,353	7,386	0,020	0,00041
17	0,046	1,763	0,0021	$9,7 \cdot 10^{-5}$	$4,5 \cdot 10^{-6}$	0,004	1,759	0,004	0,00002
18	1,674	12,117	2,8023	4,691	7,853	33,955	12,097	0,020	0,00040
19	0,839	6,02	0,7039	0,591	0,496	4,238	5,916	0,104	0,01077
20	0,026	1,687	0,0007	$1,8 \cdot 10^{-5}$	$4,6 \cdot 10^{-7}$	0,001	1,676	0,011	0,00013
Σ	19,696	146,58	26,437	39,530	63,088	285,996			0,039

Рассчитаем значения коэффициентов системы уравнений (6)

$$g_{00} = 20; g_{01} = 19,696; g_{02} = 26,437; h_0 = 146,58;$$

$$g_{10} = 19,696; g_{11} = 26,437; g_{12} = 39,530; h_1 = 190,9;$$

$$g_{20} = 26,437; g_{21} = 39,530; g_{22} = 63,088; h_2 = 285,996,$$

на основании которых рассчитаем значения определителей (10)-(12)

$$\Delta = \begin{vmatrix} 20 & 19,696 & 26,437 \\ 19,696 & 26,437 & 39,53 \\ 26,437 & 39,53 & 63,088 \end{vmatrix} = 320,58;$$

$$\Delta_0 = \begin{vmatrix} 146,58 & 19,696 & 26,437 \\ 190,9 & 26,437 & 39,53 \\ 285,996 & 39,53 & 63,088 \end{vmatrix} = 502,99;$$

$$\Delta_1 = \begin{vmatrix} 20 & 146,58 & 26,437 \\ 19,696 & 190,9 & 39,53 \\ 26,437 & 285,996 & 63,088 \end{vmatrix} = 1304,3;$$

$$\Delta_2 = \begin{vmatrix} 20 & 19,696 & 146,58 \\ 19,696 & 26,437 & 190,9 \\ 26,437 & 39,53 & 285,996 \end{vmatrix} = 425,26.$$

По полученным значениям рассчитываем значения коэффициентов уравнения (4)

$$a_0 = \frac{\Delta_0}{\Delta} = \frac{502,99}{320,58} = 1,569;$$

$$a_1 = \frac{\Delta_1}{\Delta} = \frac{1304,3}{320,58} = 4,069;$$

$$a_2 = \frac{\Delta_2}{\Delta} = \frac{425,26}{320,58} = 1,327.$$

Строим диаграмму рассеивания экспериментальных данных и линию полученного уравнения регрессии $Y = 1,569 + 4,069X + 1,327X^2$ (рис. 3). Следует отметить, что экспериментальные точки достаточно плотно лежат на линии регрессии.

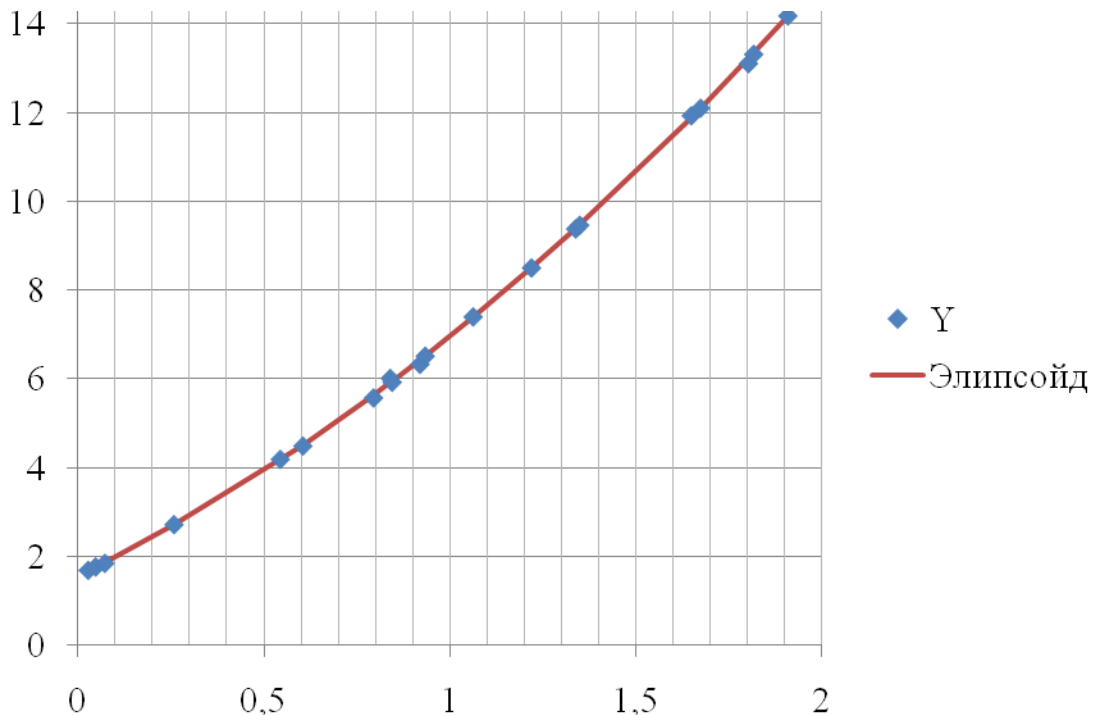


Рис. 3 - Диаграмма рассеивания экспериментальных данных и линия регрессии $Y=1,569+4,069X +1,327X^2$

Выполним оценку значимости нелинейной модели, рассчитаем коэффициент детерминации

$$R^2 = 1 - \frac{0,039}{311,762} = 0,999$$

и значение F -критерия

$$F[1,18] = \frac{0,999}{1-0,999} \cdot \frac{18}{1} \approx 144\,985.$$

Т.к. $F[1,18] = 144\,985 > F_{кр}[1,18] = 4,41$, то полученное уравнение можно также признать значимым.

Сравнивая два уравнения, можно сделать вывод, что взаимосвязь между X и Y лучше описывается уравнением $Y=1,569+4,069X +1,327X^2$.

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение Б), студент изучает теоретические сведения согласно пункту 3. Далее выполняет расчеты аналогичные в рассмотренном примере **с учетом имеющихся особенностей задания.**

В отчёте по практической работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое линейная регрессия?
2. Что такое многофакторная регрессия?
3. Что такое нелинейная регрессия?
4. Что рассчитывается при построении регрессионного уравнения?
5. Что такое критерий значимости регрессии?
6. Что такое коэффициент детерминации?
7. Какова интерпретация коэффициентов уравнения множественной регрессии?
8. Чем характеризуется точность уравнения регрессии?
9. Как осуществляется выбор «наилучшего» уравнения регрессии?

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
F-распределение (при $q=0,05$)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
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	8,85	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	4,37
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	2,41
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,69	2,51	2,30
13	4,67	3,81	3,41	3,18	3,03	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	2,37	2,20	2,01	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	3,39	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,23	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	2,31	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,55	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
31	4,16	3,30	2,91	2,68	2,52	2,41	2,25	2,08	1,88	1,61
35	4,12	3,27	2,87	2,64	2,49	2,37	2,22	2,04	1,83	1,56
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,20	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,47
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	2,53	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	1,33
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,03	1,85	1,63	1,28
∞	3,84	3,00	2,61	2,37	2,21	2,10	1,94	1,75	1,52	1,03

F-распределение (при $q=0,1$)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
2	8,53	9,00	9,16	9,24	9,29	9,33	9,37	9,41	9,45	9,49
3	5,54	5,46	5,39	5,34	5,31	5,28	5,25	5,22	5,18	5,13
4	4,54	4,32	4,19	4,11	4,05	4,01	3,95	3,90	3,83	3,76
5	4,06	3,78	3,62	3,52	3,45	3,40	3,34	3,27	3,19	3,11
6	3,78	3,46	3,29	3,18	3,11	3,05	2,98	2,90	2,82	2,72
7	3,59	3,26	3,07	2,96	2,88	2,83	2,75	2,67	2,58	2,47
8	3,46	3,11	2,92	2,81	2,73	2,67	2,59	2,50	2,40	2,29
9	3,36	3,01	2,81	2,69	2,61	2,55	2,47	2,38	2,28	2,16
10	3,29	2,92	2,73	2,61	2,52	2,46	2,38	2,28	2,18	2,06
11	3,23	2,86	2,66	2,54	2,45	2,39	2,30	2,21	2,10	1,97
12	3,18	2,81	2,61	2,48	2,39	2,33	2,24	2,15	2,04	1,90
13	3,14	2,76	2,56	2,43	2,35	2,28	2,20	2,10	1,98	1,85
14	3,10	2,73	2,52	2,39	2,31	2,24	2,15	2,05	1,94	1,80
15	3,07	2,70	2,49	2,36	2,27	2,21	2,12	2,02	1,90	1,76
16	3,05	2,67	2,46	2,33	2,24	2,18	2,09	1,99	1,87	1,72
17	3,03	2,64	2,44	2,31	2,22	2,15	2,06	1,96	1,84	1,69
18	3,01	2,62	2,42	2,29	2,20	2,13	2,04	1,93	1,81	1,66
19	2,99	2,61	2,40	2,27	2,18	2,11	2,02	1,91	1,79	1,63
20	2,97	2,59	2,38	2,25	2,16	2,09	2,00	1,89	1,77	1,61
21	2,96	2,57	2,36	2,23	2,14	2,08	1,98	1,87	1,75	1,59
22	2,95	2,56	2,35	2,22	2,13	2,06	1,97	1,86	1,73	1,57
23	2,94	2,55	2,34	2,21	2,11	2,05	1,95	1,84	1,72	1,55
24	2,93	2,54	2,33	2,19	2,10	2,04	1,94	1,83	1,70	1,53
25	2,92	2,53	2,32	2,18	2,09	2,02	1,93	1,82	1,69	1,52
26	2,91	2,52	2,31	2,17	2,08	2,01	1,92	1,81	1,68	1,50
27	2,90	2,51	2,30	2,17	2,07	2,00	1,91	1,80	1,67	1,49
28	2,89	2,50	2,29	2,16	2,06	2,00	1,90	1,79	1,66	1,48
29	2,89	2,50	2,28	2,15	2,06	1,99	1,89	1,78	1,65	1,47
30	2,88	2,49	2,28	2,14	2,05	1,98	1,88	1,77	1,64	1,46
31	2,87	2,48	2,27	2,14	2,04	1,97	1,88	1,77	1,63	1,45
35	2,85	2,46	2,25	2,11	2,02	1,95	1,85	1,74	1,60	1,41
40	2,84	2,44	2,23	2,09	2,00	1,93	1,83	1,71	1,57	1,38
45	2,82	2,42	2,21	2,07	1,98	1,91	1,81	1,70	1,55	1,35
50	2,81	2,41	2,20	2,06	1,97	1,90	1,80	1,68	1,54	1,33
60	2,79	2,39	2,18	2,04	1,95	1,87	1,77	1,66	1,51	1,29
70	2,78	2,38	2,16	2,03	1,93	1,86	1,76	1,64	1,49	1,27
80	2,77	2,37	2,15	2,02	1,92	1,85	1,75	1,63	1,48	1,25
90	2,76	2,36	2,15	2,01	1,91	1,84	1,74	1,62	1,47	1,23
100	2,76	2,36	2,14	2,00	1,91	1,83	1,73	1,61	1,46	1,22
∞	2,71	2,30	2,08	1,95	1,85	1,77	1,67	1,55	1,38	1,03

ПРИЛОЖЕНИЕ Б
Исходные данные к работе

Таблица 1

	№ варианта														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>q</i>	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05

Таблица 2

№	№ варианта								
	1			2			3		
1	178,435	202,092	144,037	194,150	14,380	12,011	113,620	138,372	10,443
2	178,416	201,730	143,846	194,104	14,325	12,009	113,616	138,355	10,668
3	178,370	201,814	143,907	194,116	14,357	12,005	113,614	138,368	10,617
4	178,457	201,802	143,892	194,113	14,317	12,013	113,606	138,362	10,622
5	178,398	201,720	143,924	194,122	14,345	11,987	113,624	138,374	10,685
6	178,444	201,868	143,899	194,116	14,343	12,002	113,585	138,346	10,463
7	178,351	202,009	144,018	194,142	14,354	11,978	113,610	138,371	10,485
8	178,465	201,810	143,978	194,135	14,352	11,994	113,631	138,358	10,483
9	178,466	201,773	143,980	194,129	14,394	12,000	113,641	138,381	10,554
10	178,490	201,816	144,029	194,144	14,375	11,991	113,628	138,364	10,458
11	178,442	201,754	143,880	194,111	14,306	11,995	113,634	138,375	10,465
12	178,510	202,089	144,062	194,153	14,426	11,969	113,707	138,418	10,637
13	178,512	201,997	144,099	194,162	14,408	11,994	113,621	138,376	10,469
14	178,360	201,826	144,041	194,144	14,391	11,984	113,619	138,370	10,428
15	178,417	201,812	144,059	194,155	14,393	11,991	113,650	138,380	10,539

№	№ варианта								
	4			5			6		
1	49,293	142,859	169,932	197,329	23,619	203,924	61,057	34,567	206,450
2	49,353	142,918	169,724	197,099	23,735	203,934	61,220	34,554	206,483
3	49,336	142,788	169,858	197,098	23,761	203,950	61,146	34,588	206,380
4	49,333	142,860	169,732	197,095	23,826	203,944	61,149	34,564	206,372
5	49,340	142,836	169,747	197,119	23,690	203,923	61,269	34,629	206,326
6	49,318	142,894	169,765	197,029	23,763	203,929	61,147	34,588	206,444
7	49,324	142,829	169,923	197,097	23,702	203,928	61,088	34,592	206,411
8	49,316	142,862	169,766	197,196	23,729	203,935	61,235	34,584	206,376
9	49,325	142,843	169,602	196,985	23,723	203,920	61,233	34,602	206,288
10	49,317	142,816	169,630	197,146	23,699	203,938	61,280	34,565	206,391
11	49,321	142,881	169,856	197,242	23,698	203,925	61,152	34,575	206,378
12	49,338	142,897	169,697	197,101	23,714	203,923	61,189	34,544	206,465
13	49,323	142,934	169,603	197,085	23,725	203,927	61,200	34,602	206,391

№ варианта									
№	4			5			6		
14	49,322	142,859	169,691	197,210	23,677	203,930	61,238	34,591	206,372
15	49,322	142,928	169,860	197,128	23,681	203,931	61,241	34,632	206,318

№ варианта									
№	7			8			9		
1	55,995	80,812	206,307	138,182	183,767	44,262	145,114	154,467	191,922
2	56,028	80,817	206,274	138,133	183,792	44,302	145,091	154,464	191,933
3	55,948	80,732	206,265	138,204	183,777	44,260	145,104	154,479	191,945
4	55,986	80,935	206,242	138,148	183,797	44,263	145,113	154,441	191,904
5	56,096	80,891	206,266	138,156	183,806	44,291	145,104	154,429	191,921
6	56,098	80,818	206,255	138,068	183,798	44,280	145,106	154,425	191,908
7	55,984	80,738	206,281	138,194	183,765	44,257	145,113	154,375	191,910
8	55,846	80,690	206,254	138,217	183,775	44,282	145,104	154,423	191,906
9	56,117	80,934	206,316	138,155	183,788	44,294	145,083	154,365	191,887
10	56,057	80,826	206,266	138,129	183,785	44,274	145,118	154,444	191,891
11	56,062	80,776	206,274	138,165	183,776	44,227	145,127	154,468	191,919
12	55,858	80,735	206,264	138,091	183,814	44,208	145,163	154,463	191,929
13	56,076	80,752	206,283	138,141	183,803	44,235	145,136	154,464	191,929
14	56,025	80,861	206,281	138,170	183,798	44,261	145,103	154,415	191,913
15	55,851	80,832	206,248	138,189	183,771	44,246	145,129	154,339	191,901

№ варианта									
№	10			11			12		
1	198,209	43,903	200,697	152,883	155,297	114,758	50,691	195,826	96,043
2	198,198	43,910	200,691	152,882	155,270	114,765	50,612	195,813	95,961
3	198,191	43,937	200,693	152,868	155,253	114,762	50,498	195,780	95,920
4	198,212	43,910	200,697	152,882	155,172	114,743	50,751	195,837	96,074
5	198,203	43,933	200,694	152,912	155,317	114,742	50,615	195,809	95,988
6	198,212	43,902	200,686	152,840	155,303	114,760	50,651	195,818	96,005
7	198,210	43,902	200,696	152,870	155,177	114,750	50,618	195,807	95,969
8	198,211	43,936	200,683	152,889	155,339	114,759	50,577	195,799	95,934
9	198,227	43,898	200,692	152,898	155,270	114,763	50,613	195,809	95,986
10	198,221	43,911	200,683	152,867	155,273	114,761	50,609	195,802	95,958
11	198,208	43,922	200,684	152,876	155,272	114,744	50,637	195,817	95,997
12	198,200	43,898	200,696	152,888	155,270	114,766	50,527	195,795	95,914
13	198,200	43,920	200,691	152,894	155,340	114,750	50,643	195,817	95,998
14	198,209	43,910	200,686	152,874	155,337	114,746	50,524	195,788	95,925
15	198,218	43,875	200,688	152,898	155,359	114,746	50,576	195,791	95,964

№ варианта									
№	13			14			15		
1	206,732	118,321	182,070	190,116	180,982	83,113	117,240	114,728	139,641
2	206,735	118,275	181,949	190,122	181,088	83,153	117,263	114,733	139,626
3	206,773	118,307	182,053	190,117	180,977	83,117	117,242	114,730	139,637
4	206,675	118,325	182,000	190,117	180,978	83,120	117,249	114,738	139,641
5	206,761	118,298	182,005	190,121	181,039	83,145	117,258	114,735	139,636
6	206,727	118,302	181,996	190,129	181,229	83,223	117,180	114,731	139,636
7	206,748	118,315	181,988	190,118	181,054	83,124	117,281	114,738	139,624
8	206,639	118,283	181,984	190,117	181,039	83,138	117,203	114,728	139,645
9	206,622	118,313	182,020	190,117	181,014	83,136	117,330	114,734	139,649
10	206,688	118,298	181,996	190,114	180,985	83,127	117,321	114,733	139,635
11	206,690	118,290	181,955	190,117	181,042	83,173	117,181	114,731	139,631
12	206,661	118,290	181,968	190,123	181,085	83,161	117,306	114,728	139,649
13	206,749	118,297	181,969	190,119	181,041	83,132	117,329	114,724	139,649
14	206,717	118,318	182,034	190,121	181,038	83,151	117,232	114,732	139,633
15	206,695	118,281	181,924	190,118	181,046	83,150	117,246	114,728	139,643

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Жукмонова
« 14 » _____ г.



КОРРЕЛЯЦИОННЫЙ АНАЛИЗ СВЯЗЕЙ МЕЖДУ КОЛИЧЕСТВЕННЫМИ ПРИЗНАКАМИ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Корреляционный анализ связей между количественными признаками : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.

305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: исследовать зависимости между результатами наблюдений и определить их характер.

2 Задание: построить диаграммы рассеивания, рассчитать коэффициенты корреляции для результатов наблюдения трёх случайных величин, а установить значимость этих коэффициентов.

3 Краткие теоретические сведения

В технологической практике исследователь часто сталкивается с необходимостью установления факта существования функциональных или иных зависимостей между экспериментальными данными, нередко такая связь может быть случайной. Различные постановки задач статистического исследования можно классифицировать следующим образом: задачи корреляционного анализа (задачи исследования наличия взаимосвязей между отдельными группами переменных); задачи регрессионного анализа; задачи дисперсионного анализа.

3.1 Понятие о корреляционном анализе

Корреляция (от лат. *Correlatio* – соотношение) – мера связи. Практическая статистика использует это понятие для обозначения группы показателей, описывающих степень сходства в варьировании двух и более переменных. Предварительный анализ зависимости между оценками параметра дают простые коэффициенты корреляций. Коэффициент корреляции является мерой линейной зависимости между анализируемыми переменными. Кроме того, в статистическом анализе используются коэффициенты частной корреляции, которые исключают зависимость остальных переменных, кроме рассматриваемой. Все эти корреляции весьма полезны для первичного анализа взаимозависимости. Корреляционный анализ определяет меру линейности зависимости независимо от положения прямой, то есть наличие некоторой гипотетической функциональной зависимости.

Оценкой коэффициента корреляции двух случайных величин X и Y является величина:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}},$$

где n - объем выборки X и Y ;

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i;$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

Сильная корреляция означает закономерность: чем более отличается от своего среднего значение одной переменной, тем дальше от своего среднего значение другой. Также можно утверждать, что чем ближе точки на графике рассеивания параметров X и Y расположены к прямой линии больше корреляция между X и Y .

Коэффициент может принимать как положительные, так и отрицательные значения. Знак означает направленность изменчивости двух переменных. Если коэффициент корреляции больше нуля, то для данного наблюдения отличия от среднего бывают, как правило, в одну сторону. Если меньше нуля – увеличение первой переменной, как правило, связано с уменьшением второй. Сравнение именно со своими средними позволяет сравнивать переменные разной размерности и природы, например метры с килограммами или секунды с часами. Нас может интересовать, в зависимости от цели исследования, отличие коэффициента корреляции от нуля (есть ли вообще связь между переменными) или отличие двух коэффициентов друг от друга (одинаковы ли связи, например, между переменными A и B , с одной стороны и переменными C и D , с другой).

Нельзя сразу принимать полученный коэффициент корреляции как окончательный результат, если исследование отвечает на сложные вопросы. Если вероятность нуль-гипотезы достаточно мала, то можно сказать только, что между переменными существует какая-то связь, и не более того. Ответы на другие вопросы – предмет дальнейшего анализа, а часто – и специальных опытов.

Коэффициент корреляции находится в пределе $|r| \leq 1$. Абсолютная величина коэффициента означает силу связи (условно принима-

ем: 0 – 0,3 – слабая связь; 0,3 – 0,7 – связь средней силы; 0,7 – 1 – сильная связь).

Знак коэффициента определяет, совпадают ли направления изменчивости переменных. «+»– совпадают (чем больше одно, тем больше другое), «-»– противоположны (чем больше одно, тем меньше другое). На рис. 1. приведены различные случаи корреляционной связи.

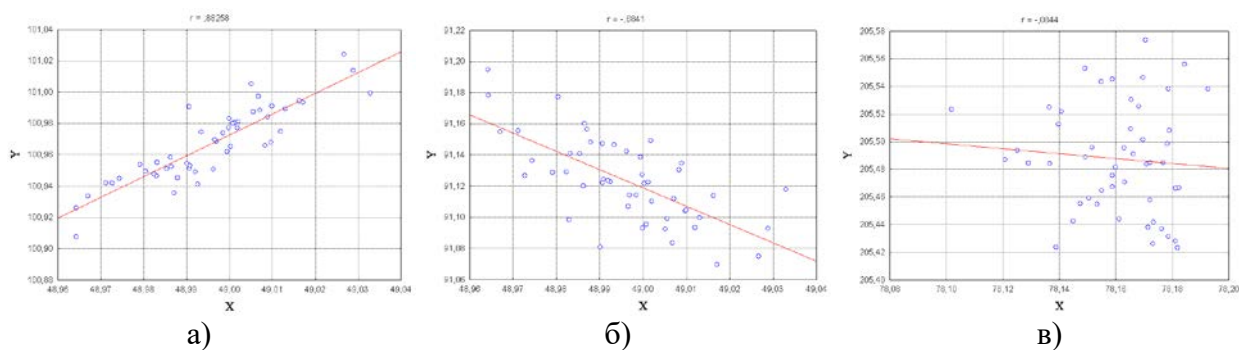


Рис. 1 – Графики рассеивания с различной силой корреляционной связи:
а – сильная, положительная; б – средняя отрицательная;
в – слабая отрицательная

3.2 Проверка значимости коэффициентов корреляции

Для проверки значимости коэффициентов корреляции чаще всего используют распределение Стьюдента и условие

$$\frac{|r_{XY}| \sqrt{n-2}}{\sqrt{1-r_{XY}^2}} < t_{P,k},$$

где $t_{P,k}$ – коэффициент Стьюдента, который берется из справочных таблиц (приложение А) при заданной доверительной вероятности $P=1-q$ и числе степеней свободы $k=n-2$.

Если условие выполняется, то гипотеза об отсутствии корреляционной связи принимается.

3.3. Пример выполнения работы

В таблице представлены результаты 15-ти наблюдений двух случайных величин X и Y .

	Номер наблюдения														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	5,48	5,57	5,49	5,40	5,55	5,41	5,54	5,46	5,33	5,41	5,55	5,53	5,42	5,60	5,35
Y	1,61	1,54	1,56	1,47	1,60	1,51	1,61	1,45	1,43	1,43	1,60	1,56	1,47	1,61	1,48

Для расчета коэффициента корреляции заполним следующую таблицу

№	X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	Y_i	$Y_i - \bar{Y}$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$
1	5,48	0,01	0,0001	1,61	0,08	0,0066	0,0006
2	5,57	0,10	0,0095	1,54	0,01	0,0001	0,0011
3	5,49	0,02	0,0003	1,56	0,03	0,0010	0,0005
4	5,40	-0,07	0,0053	1,47	-0,06	0,0034	0,0043
5	5,55	0,08	0,0060	1,60	0,07	0,0051	0,0055
6	5,41	-0,06	0,0039	1,51	-0,02	0,0003	0,0012
7	5,54	0,07	0,0045	1,61	0,08	0,0066	0,0055
8	5,46	-0,01	0,0002	1,45	-0,08	0,0062	0,0010
9	5,33	-0,14	0,0204	1,43	-0,10	0,0097	0,0141
10	5,41	-0,06	0,0039	1,43	-0,10	0,0097	0,0062
11	5,55	0,08	0,0060	1,60	0,07	0,0051	0,0055
12	5,53	0,06	0,0033	1,56	0,03	0,0010	0,0018
13	5,42	-0,05	0,0028	1,47	-0,06	0,0034	0,0031
14	5,60	0,13	0,0162	1,61	0,08	0,0066	0,0104
15	5,35	-0,12	0,0150	1,48	-0,05	0,0024	0,0060
	$\Sigma=82,09$		$\Sigma=0,0973$	$\Sigma=22,93$		$\Sigma=0,0674$	$\Sigma=0,0667$
	$\bar{X} = 5,47$			$\bar{Y} = 1,53$			

Рассчитаем величину коэффициента корреляции

$$r_{XY} = \frac{0,0667}{\sqrt{0,0973 \cdot 0,0674}} = 0,83.$$

На рис. 2 показана диаграмма рассеивания результатов наблюдения X и Y .

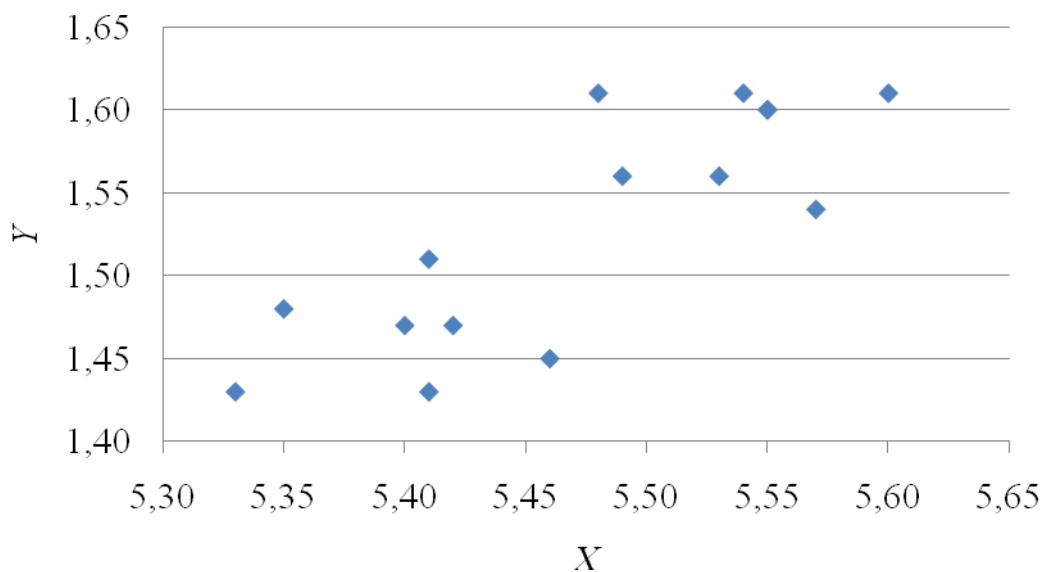


Рис. 2 диаграмма рассеивания результатов наблюдения X и Y

Исходя из рассчитанного коэффициента корреляции и диграммы рассеивания, можно сделать вывод, что между наблюдениями X и Y существует сильная положительная связь.

Далее выполним проверку значимости коэффициентов корреляции при уровне значимости $q=0,05$

$$t = \frac{|r_{XY}| \sqrt{n-2}}{\sqrt{1-r_{XY}^2}} = \frac{|0,83| \sqrt{15-2}}{\sqrt{1-0,83^2}} = 5,23 > t_{0,95;13} = 2,160,$$

где $t_{0,95;13}$ – коэффициент Стьюдента из справочной таблицы приложения А, взятый при заданной доверительной вероятности $P=1-q=1-0,05=0,95$ и числе степеней свободы $k=n-2=15-2=13$.

Т.к. условие выполняется, то можно сделать вывод о значимости коэффициента корреляции между наблюдениями X и Y .

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение Б), студент изучает теоретические сведения согласно пункту 3. Далее выполняет расчеты аналогичные в рассмотренном примере с учетом имеющихся особенностей задания.

В отчёте по практической работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое корреляция?
2. Что такое коэффициент корреляции?
3. Как рассчитывается коэффициент корреляции?
4. Какова область значений коэффициента корреляции?
5. Что означает положительная и отрицательная корреляция?
6. При каком значении коэффициента корреляции связь является сильной, средней и слабой?

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
Распределение Стьюдента (t_p)

k	P											
	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,99
1	0,079	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	63,657
2	0,071	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	9,925
3	0,068	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	5,841
4	0,067	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	4,604
5	0,066	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	4,032
6	0,065	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,707
7	0,065	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	3,499
8	0,065	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	3,355
9	0,064	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	3,250
10	0,064	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	3,169
11	0,064	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	3,106
12	0,064	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	3,055
13	0,064	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	3,012
14	0,064	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,977
15	0,064	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,947
16	0,064	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,921
17	0,064	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,898
18	0,064	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,878
19	0,064	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,861
20	0,063	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,845
21	0,063	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,831
22	0,063	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,819
23	0,063	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,807
24	0,063	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,797
25	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,787
26	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,779
27	0,063	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,771
28	0,063	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,763
29	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,756
30	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,750
300	0,063	0,126	0,254	0,386	0,525	0,675	0,843	1,038	1,284	1,650	1,968	2,592

ПРИЛОЖЕНИЕ Б
Исходные данные к работе

Таблица 1

	№ варианта														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>q</i>	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05

Таблица 2

№	№ варианта								
	1			2			3		
1	178,435	202,092	144,037	194,150	14,380	12,011	113,620	138,372	10,443
2	178,416	201,730	143,846	194,104	14,325	12,009	113,616	138,355	10,668
3	178,370	201,814	143,907	194,116	14,357	12,005	113,614	138,368	10,617
4	178,457	201,802	143,892	194,113	14,317	12,013	113,606	138,362	10,622
5	178,398	201,720	143,924	194,122	14,345	11,987	113,624	138,374	10,685
6	178,444	201,868	143,899	194,116	14,343	12,002	113,585	138,346	10,463
7	178,351	202,009	144,018	194,142	14,354	11,978	113,610	138,371	10,485
8	178,465	201,810	143,978	194,135	14,352	11,994	113,631	138,358	10,483
9	178,466	201,773	143,980	194,129	14,394	12,000	113,641	138,381	10,554
10	178,490	201,816	144,029	194,144	14,375	11,991	113,628	138,364	10,458
11	178,442	201,754	143,880	194,111	14,306	11,995	113,634	138,375	10,465
12	178,510	202,089	144,062	194,153	14,426	11,969	113,707	138,418	10,637
13	178,512	201,997	144,099	194,162	14,408	11,994	113,621	138,376	10,469
14	178,360	201,826	144,041	194,144	14,391	11,984	113,619	138,370	10,428
15	178,417	201,812	144,059	194,155	14,393	11,991	113,650	138,380	10,539

№	№ варианта								
	4			5			6		
1	49,293	142,859	169,932	197,329	23,619	203,924	61,057	34,567	206,450
2	49,353	142,918	169,724	197,099	23,735	203,934	61,220	34,554	206,483
3	49,336	142,788	169,858	197,098	23,761	203,950	61,146	34,588	206,380
4	49,333	142,860	169,732	197,095	23,826	203,944	61,149	34,564	206,372
5	49,340	142,836	169,747	197,119	23,690	203,923	61,269	34,629	206,326
6	49,318	142,894	169,765	197,029	23,763	203,929	61,147	34,588	206,444
7	49,324	142,829	169,923	197,097	23,702	203,928	61,088	34,592	206,411
8	49,316	142,862	169,766	197,196	23,729	203,935	61,235	34,584	206,376
9	49,325	142,843	169,602	196,985	23,723	203,920	61,233	34,602	206,288
10	49,317	142,816	169,630	197,146	23,699	203,938	61,280	34,565	206,391
11	49,321	142,881	169,856	197,242	23,698	203,925	61,152	34,575	206,378
12	49,338	142,897	169,697	197,101	23,714	203,923	61,189	34,544	206,465
13	49,323	142,934	169,603	197,085	23,725	203,927	61,200	34,602	206,391

№ варианта									
№	4			5			6		
14	49,322	142,859	169,691	197,210	23,677	203,930	61,238	34,591	206,372
15	49,322	142,928	169,860	197,128	23,681	203,931	61,241	34,632	206,318

№ варианта									
№	7			8			9		
1	55,995	80,812	206,307	138,182	183,767	44,262	145,114	154,467	191,922
2	56,028	80,817	206,274	138,133	183,792	44,302	145,091	154,464	191,933
3	55,948	80,732	206,265	138,204	183,777	44,260	145,104	154,479	191,945
4	55,986	80,935	206,242	138,148	183,797	44,263	145,113	154,441	191,904
5	56,096	80,891	206,266	138,156	183,806	44,291	145,104	154,429	191,921
6	56,098	80,818	206,255	138,068	183,798	44,280	145,106	154,425	191,908
7	55,984	80,738	206,281	138,194	183,765	44,257	145,113	154,375	191,910
8	55,846	80,690	206,254	138,217	183,775	44,282	145,104	154,423	191,906
9	56,117	80,934	206,316	138,155	183,788	44,294	145,083	154,365	191,887
10	56,057	80,826	206,266	138,129	183,785	44,274	145,118	154,444	191,891
11	56,062	80,776	206,274	138,165	183,776	44,227	145,127	154,468	191,919
12	55,858	80,735	206,264	138,091	183,814	44,208	145,163	154,463	191,929
13	56,076	80,752	206,283	138,141	183,803	44,235	145,136	154,464	191,929
14	56,025	80,861	206,281	138,170	183,798	44,261	145,103	154,415	191,913
15	55,851	80,832	206,248	138,189	183,771	44,246	145,129	154,339	191,901

№ варианта									
№	10			11			12		
1	198,209	43,903	200,697	152,883	155,297	114,758	50,691	195,826	96,043
2	198,198	43,910	200,691	152,882	155,270	114,765	50,612	195,813	95,961
3	198,191	43,937	200,693	152,868	155,253	114,762	50,498	195,780	95,920
4	198,212	43,910	200,697	152,882	155,172	114,743	50,751	195,837	96,074
5	198,203	43,933	200,694	152,912	155,317	114,742	50,615	195,809	95,988
6	198,212	43,902	200,686	152,840	155,303	114,760	50,651	195,818	96,005
7	198,210	43,902	200,696	152,870	155,177	114,750	50,618	195,807	95,969
8	198,211	43,936	200,683	152,889	155,339	114,759	50,577	195,799	95,934
9	198,227	43,898	200,692	152,898	155,270	114,763	50,613	195,809	95,986
10	198,221	43,911	200,683	152,867	155,273	114,761	50,609	195,802	95,958
11	198,208	43,922	200,684	152,876	155,272	114,744	50,637	195,817	95,997
12	198,200	43,898	200,696	152,888	155,270	114,766	50,527	195,795	95,914
13	198,200	43,920	200,691	152,894	155,340	114,750	50,643	195,817	95,998
14	198,209	43,910	200,686	152,874	155,337	114,746	50,524	195,788	95,925
15	198,218	43,875	200,688	152,898	155,359	114,746	50,576	195,791	95,964

№ варианта									
№	13			14			15		
1	206,732	118,321	182,070	190,116	180,982	83,113	117,240	114,728	139,641
2	206,735	118,275	181,949	190,122	181,088	83,153	117,263	114,733	139,626
3	206,773	118,307	182,053	190,117	180,977	83,117	117,242	114,730	139,637
4	206,675	118,325	182,000	190,117	180,978	83,120	117,249	114,738	139,641
5	206,761	118,298	182,005	190,121	181,039	83,145	117,258	114,735	139,636
6	206,727	118,302	181,996	190,129	181,229	83,223	117,180	114,731	139,636
7	206,748	118,315	181,988	190,118	181,054	83,124	117,281	114,738	139,624
8	206,639	118,283	181,984	190,117	181,039	83,138	117,203	114,728	139,645
9	206,622	118,313	182,020	190,117	181,014	83,136	117,330	114,734	139,649
10	206,688	118,298	181,996	190,114	180,985	83,127	117,321	114,733	139,635
11	206,690	118,290	181,955	190,117	181,042	83,173	117,181	114,731	139,631
12	206,661	118,290	181,968	190,123	181,085	83,161	117,306	114,728	139,649
13	206,749	118,297	181,969	190,119	181,041	83,132	117,329	114,724	139,649
14	206,717	118,318	182,034	190,121	181,038	83,151	117,232	114,732	139,633
15	206,695	118,281	181,924	190,118	181,046	83,150	117,246	114,728	139,643

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Жолтимонова
« 14 » _____ 2023 г.



ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Однофакторный дисперсионный анализ : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: Ознакомить с методом однофакторного дисперсионного анализа и овладеть практическими навыками анализа технологических процессов данным методом.

2 Задание:

2.1 По табл. 1 приложения Б в соответствии со своим вариантом выбрать уровень значимости критериев.

2.2 Для исходных данных, приведенных в табл. 2 приложения Б, провести анализа технологических процессов данным методом однофакторного дисперсионного анализа.

2.3. Оформить отчет.

3 Краткие теоретические сведения

Идея дисперсионного анализа, как и сам термин «дисперсия», принадлежит английскому статистику Р. Фишеру. Метод был разработан в 1920-х годах и используется для определения степени влияния на изучаемый показатель некоторых факторов, в том числе и не поддающихся количественному измерению (достаточно, чтобы его можно было измерить хотя бы в шкале наименований). При исследовании зависимостей такого рода одной из наиболее простых является ситуация, когда можно указать один только фактор, который, возможно, влияет на конечный результат, и этот фактор может принимать лишь конечное число значений (уровней). В этом случае реализуется алгоритм однофакторного дисперсионного анализа. Если же на изменение показателя в равной степени могут оказывать влияние несколько факторов, то для установления степени их совместного влияния используется уже алгоритм двухфакторного или многофакторного анализа.

3.1 Однофакторный дисперсионный анализ

Предположим, что анализируется влияние на случайную величину X фактора A , изучаемого на k уровнях (A_1, A_2, \dots, A_k) . На каждом уровне A_i проведены n наблюдений $(x_{i1}, x_{i2}, \dots, x_{in})$, $i = 1, k$ случайной величины X . Следовательно, на всех k уровнях фактора A произведены kn наблюдений.

Поясним суть и последовательность проведения дисперсионного анализа.

Расположим экспериментальные данные в виде таблицы:

Номер наблюдения	Уровни фактора A					
	A_1	A_2	...	A_i	...	A_k
1	x_{11}	x_{21}	...	x_{i1}	...	x_{k1}
2	x_{12}	x_{22}	...	x_{i2}	...	x_{k2}
\vdots	\vdots	\vdots	...	\vdots	...	\vdots
j	x_{1j}	x_{2j}	...	x_{ij}	...	x_{kj}
\vdots	\vdots	\vdots	...	\vdots	...	\vdots
n	x_{1n}	x_{2n}	...	x_{in}	...	x_{kn}
Σ	X_1	X_2	...	X_i	...	X_k

Здесь $X_i = \sum_{j=1}^n x_{ij}, i = \overline{1, k}$.

Рассмотрим оценки различных дисперсий, возникающие при анализе таблицы результатов наблюдений. Для оценки дисперсии, характеризующей изменение данных на уровне A_i (по строкам таблицы), имеем

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 = \frac{1}{n-1} \left[\sum_{j=1}^n x_{ij}^2 - \frac{1}{n} \left(\sum_{j=1}^n x_{ij} \right)^2 \right]. \quad (1)$$

Находим оценку дисперсии, характеризующей рассеяние рассматриваемой случайной величины вне влияния фактора A, по формуле

$$\begin{aligned} s_0^2 &= \frac{1}{k} \sum_{i=1}^k s_i^2 = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 = \\ &= \frac{1}{k(n-1)} \left[\sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - \frac{1}{n} \sum_{i=1}^k \left(\sum_{j=1}^n x_{ij} \right)^2 \right]. \end{aligned} \quad (2)$$

Оценка s_i^2 имеет $(n-1)$ степеней свободы, а оценка s_0^2 – $k(n-1)$ степеней свободы.

Оценка s^2 выборочной дисперсии с использованием всех наблюдений равна

$$s^2 = \frac{1}{kn-1} \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x})^2, \quad (3)$$

где $\bar{x} = \frac{1}{k} \sum_{i=1}^k \bar{x}_i$; $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}$.

Следовательно

$$s^2 = \frac{1}{kn-1} \left[\sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - \frac{1}{kn} \left(\sum_{i=1}^k \sum_{j=1}^n x_{ij} \right)^2 \right]. \quad (4)$$

Оценку дисперсии, характеризующей изменение математических ожиданий m_i , связанное с влиянием фактора A рассчитаем как

$$s_A^2 = \frac{n}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{x})^2. \quad (5)$$

Оценка s_A^2 имеет $(k-1)$ степеней свободы.

Влияние фактора A признается значимым, если значимо отношение s_A^2/s_0^2 . Отношение s_A^2/s_0^2 признается значимым с доверительной вероятностью q , если

$$s_A^2/s_0^2 > F_{q,k-1,k(n-1)}$$

где $F_{q,k-1,k(n-1)}$ – квантиль F распределения с $k-1$ и $k(n-1)$ степенями свободы. Для его нахождения используются таблицы приложения А.

Итоговая схема вычислений следующая. Вычисляем последовательно суммы

$$Q_1 = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2; \quad Q_2 = \frac{1}{n} \sum_{i=1}^k X_i^2; \quad Q_3 = \frac{1}{kn} \left(\sum_{i=1}^k X_i \right)^2.$$

Далее находим

$$s_0^2 = \frac{Q_1 - Q_2}{k(n-1)}; \quad s_A^2 = \frac{Q_2 - Q_3}{k-1}.$$

Сравниваем s_A^2 и s_0^2 и устанавливаем наличие влияния фактора A . Если влияние фактора A не установлено, то всю выборку наблюдений будем считать однородной с общей дисперсией

$$s^2 = \frac{Q_1 - Q_3}{kn-1};$$

Пример. Провести дисперсионный анализ данных, представленных таблицей, при уровне значимости $q = 0,05$.

Номер наблюдения	Уровни фактора A				
	A ₁	A ₂	A ₃	A ₄	A ₅
1	3,2	2,6	2,9	3,6	3,0
2	3,1	3,1	2,6	3,4	3,4
3	3,1	2,7	3,0	3,2	3,2
4	2,8	2,9	3,1	3,3	3,5
5	3,3	2,7	3,0	3,5	2,9
6	3,0	2,8	2,8	3,3	3,1

Для вычисления величин Q_1 , Q_2 и Q_3 заполним следующие таблицы

Номер наблюдения	x_{1j}	x_{2j}	x_{3j}	x_{4j}	x_{5j}^2		
1	3,2	2,6	2,9	3,6	3,0	$Q_2 = \frac{1}{6} \sum_{i=1}^5 X_i^2;$ $Q_3 = \frac{1}{5 \cdot 6} \left(\sum_{i=1}^5 X_i \right)^2.$	
2	3,1	3,1	2,6	3,4	3,4		
3	3,1	2,7	3,0	3,2	3,2		
4	2,8	2,9	3,1	3,3	3,5		
5	3,3	2,7	3,0	3,5	2,9		
6	3,0	2,8	2,8	3,3	3,1		
X_i	18,50	16,80	17,40	20,30	19,10	$\Sigma X_i = 92,10$	$Q_3 = 282,75$
X_i^2	342,25	282,24	302,76	412,09	364,81	$\Sigma X_i^2 = 1704,15$	$Q_2 = 284,03$

Номер наблюдения	x_{1j}^2	x_{2j}^2	x_{3j}^2	x_{4j}^2	x_{5j}^2		
1	10,24	6,76	8,41	12,96	9,00	$Q_1 = \sum_{i=1}^5 \sum_{j=1}^6 x_{ij}^2;$	
2	9,61	9,61	6,76	11,56	11,56		
3	9,61	7,29	9,00	10,24	10,24		
4	7,84	8,41	9,61	10,89	12,25		
5	10,89	7,29	9,00	12,25	8,41		
6	9,00	7,84	7,84	10,89	9,61		
$\sum_{j=1}^6 x_{ij}^2$	57,19	47,20	50,62	68,79	61,07	$Q_1 = 284,87$	

Далее вычисляем дисперсии

$$s_0^2 = \frac{284,87 - 284,03}{5(6-1)} = 0,034; \quad s_A^2 = \frac{284,03 - 284,35}{5-1} = 0,319.$$

$$\frac{s_A^2}{s_0^2} = \frac{0,319}{0,034} = 9,45.$$

Из таблиц для $k_1 = k - 1 = 4$ и $k_2 = k(n - 1) = 25$ находим $F_{0,05;4;25} = 2,76$.

Т.к. $s_A^2/s_0^2 = 9,45 > F_{0,05;4;25} = 2,76$ влияние фактора А на поведение наблюдаемой случайной величины следует признать значимым.

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение Б), студент изучает теоретические сведения согласно пункту 3. Далее выполняет расчеты аналогичные в рассмотренном примере **с учетом имеющихся особенностей задания.**

В отчёте по практической работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое дисперсионный анализ?
2. Можно ли с помощью дисперсионного анализа построить математическую модель объекта?
3. Какие гипотезы проверяются в дисперсионном анализе?
4. Основные предпосылки при решении задач с помощью дисперсионного анализа.
5. Основная идея однофакторного дисперсионного анализа.

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.

2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
F-распределение (при $q=0,05$)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
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	8,85	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	4,37
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	2,41
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,69	2,51	2,30
13	4,67	3,81	3,41	3,18	3,03	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	2,37	2,20	2,01	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	3,39	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,23	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	2,31	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,55	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
31	4,16	3,30	2,91	2,68	2,52	2,41	2,25	2,08	1,88	1,61
35	4,12	3,27	2,87	2,64	2,49	2,37	2,22	2,04	1,83	1,56
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,20	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,47
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	2,53	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	1,33
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,03	1,85	1,63	1,28
∞	3,84	3,00	2,61	2,37	2,21	2,10	1,94	1,75	1,52	1,03

F-распределение (при $q=0,1$)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
2	8,53	9,00	9,16	9,24	9,29	9,33	9,37	9,41	9,45	9,49
3	5,54	5,46	5,39	5,34	5,31	5,28	5,25	5,22	5,18	5,13
4	4,54	4,32	4,19	4,11	4,05	4,01	3,95	3,90	3,83	3,76
5	4,06	3,78	3,62	3,52	3,45	3,40	3,34	3,27	3,19	3,11
6	3,78	3,46	3,29	3,18	3,11	3,05	2,98	2,90	2,82	2,72
7	3,59	3,26	3,07	2,96	2,88	2,83	2,75	2,67	2,58	2,47
8	3,46	3,11	2,92	2,81	2,73	2,67	2,59	2,50	2,40	2,29
9	3,36	3,01	2,81	2,69	2,61	2,55	2,47	2,38	2,28	2,16
10	3,29	2,92	2,73	2,61	2,52	2,46	2,38	2,28	2,18	2,06
11	3,23	2,86	2,66	2,54	2,45	2,39	2,30	2,21	2,10	1,97
12	3,18	2,81	2,61	2,48	2,39	2,33	2,24	2,15	2,04	1,90
13	3,14	2,76	2,56	2,43	2,35	2,28	2,20	2,10	1,98	1,85
14	3,10	2,73	2,52	2,39	2,31	2,24	2,15	2,05	1,94	1,80
15	3,07	2,70	2,49	2,36	2,27	2,21	2,12	2,02	1,90	1,76
16	3,05	2,67	2,46	2,33	2,24	2,18	2,09	1,99	1,87	1,72
17	3,03	2,64	2,44	2,31	2,22	2,15	2,06	1,96	1,84	1,69
18	3,01	2,62	2,42	2,29	2,20	2,13	2,04	1,93	1,81	1,66
19	2,99	2,61	2,40	2,27	2,18	2,11	2,02	1,91	1,79	1,63
20	2,97	2,59	2,38	2,25	2,16	2,09	2,00	1,89	1,77	1,61
21	2,96	2,57	2,36	2,23	2,14	2,08	1,98	1,87	1,75	1,59
22	2,95	2,56	2,35	2,22	2,13	2,06	1,97	1,86	1,73	1,57
23	2,94	2,55	2,34	2,21	2,11	2,05	1,95	1,84	1,72	1,55
24	2,93	2,54	2,33	2,19	2,10	2,04	1,94	1,83	1,70	1,53
25	2,92	2,53	2,32	2,18	2,09	2,02	1,93	1,82	1,69	1,52
26	2,91	2,52	2,31	2,17	2,08	2,01	1,92	1,81	1,68	1,50
27	2,90	2,51	2,30	2,17	2,07	2,00	1,91	1,80	1,67	1,49
28	2,89	2,50	2,29	2,16	2,06	2,00	1,90	1,79	1,66	1,48
29	2,89	2,50	2,28	2,15	2,06	1,99	1,89	1,78	1,65	1,47
30	2,88	2,49	2,28	2,14	2,05	1,98	1,88	1,77	1,64	1,46
31	2,87	2,48	2,27	2,14	2,04	1,97	1,88	1,77	1,63	1,45
35	2,85	2,46	2,25	2,11	2,02	1,95	1,85	1,74	1,60	1,41
40	2,84	2,44	2,23	2,09	2,00	1,93	1,83	1,71	1,57	1,38
45	2,82	2,42	2,21	2,07	1,98	1,91	1,81	1,70	1,55	1,35
50	2,81	2,41	2,20	2,06	1,97	1,90	1,80	1,68	1,54	1,33
60	2,79	2,39	2,18	2,04	1,95	1,87	1,77	1,66	1,51	1,29
70	2,78	2,38	2,16	2,03	1,93	1,86	1,76	1,64	1,49	1,27
80	2,77	2,37	2,15	2,02	1,92	1,85	1,75	1,63	1,48	1,25
90	2,76	2,36	2,15	2,01	1,91	1,84	1,74	1,62	1,47	1,23
100	2,76	2,36	2,14	2,00	1,91	1,83	1,73	1,61	1,46	1,22
∞	2,71	2,30	2,08	1,95	1,85	1,77	1,67	1,55	1,38	1,03

ПРИЛОЖЕНИЕ Б
Исходные данные к работе

Таблица 1

	№ варианта														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>q</i>	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05

Таблица 2

Номер наблюдения	Уровни фактора А				
	A_1	A_2	A_3	A_4	A_5
Вариант 1					
1	25,19	25,45	24,90	24,62	25,29
2	25,03	25,38	24,99	24,67	25,15
3	25,13	24,61	25,48	25,29	25,49
4	25,15	25,39	24,86	24,99	24,64
5	24,61	24,62	24,57	24,62	24,51
6	25,17	24,74	25,02	24,76	25,33
7	25,23	25,43	25,41	25,06	25,03
8	25,32	25,08	25,37	25,14	25,29
9	24,58	25,25	24,67	25,22	24,77
10	25,22	24,99	24,84	25,25	24,52
Вариант 2					
1	32,52	32,63	33,34	32,80	33,20
2	33,36	32,93	32,90	33,21	33,28
3	32,76	32,59	33,07	33,33	32,77
4	33,47	32,51	32,64	32,87	32,95
5	32,93	32,52	33,49	32,89	32,99
6	32,76	33,38	33,45	33,07	33,03
7	32,55	33,24	33,45	32,94	32,87
8	33,45	33,23	32,91	33,33	32,76
9	32,74	32,61	33,40	33,00	33,20
10	33,41	33,24	33,05	33,30	33,23
Вариант 3					
1	23,13	23,06	22,54	22,97	22,59
2	22,95	22,98	22,66	22,77	22,80
3	22,71	23,26	23,41	23,39	23,35
4	23,50	22,56	23,27	22,64	23,41
5	22,72	23,20	23,41	23,40	22,87
6	23,37	23,42	23,13	22,99	22,62
7	22,93	23,30	22,58	23,00	22,79
8	22,59	22,80	23,42	22,82	22,78

Номер наблюдения	Уровни фактора А				
	A ₁	A ₂	A ₃	A ₄	A ₅
9	22,71	23,43	23,19	23,35	23,25
10	23,19	22,60	22,75	22,77	22,91
Вариант 4					
1	41,62	42,04	42,04	41,96	42,26
2	41,93	41,83	42,08	42,42	41,92
3	41,96	42,34	41,61	41,70	42,40
4	42,32	41,97	42,14	42,10	42,27
5	41,65	42,24	41,91	42,24	41,59
6	41,55	42,30	41,93	41,86	41,84
7	42,44	42,35	41,77	42,20	41,51
8	42,32	41,92	41,55	41,94	41,71
9	41,62	42,48	41,64	41,90	42,24
10	41,63	42,46	41,83	41,54	42,44
Вариант 5					
1	64,32	63,79	64,01	64,29	64,19
2	63,74	64,36	63,89	64,10	64,37
3	64,19	64,27	64,46	63,98	64,35
4	64,42	63,59	63,82	63,90	64,02
5	64,34	63,93	63,56	64,09	64,26
6	64,35	63,77	63,66	64,21	64,16
7	64,48	63,59	64,23	64,06	63,57
8	63,99	64,28	63,97	63,61	63,80
9	64,34	63,62	63,80	64,01	64,44
10	64,43	63,93	63,98	63,62	63,72
Вариант 6					
1	44,28	44,49	44,00	44,29	43,60
2	43,87	43,67	44,09	44,43	44,06
3	43,95	44,43	44,30	44,03	43,67
4	44,06	44,18	43,95	43,53	43,91
5	44,41	43,67	43,64	43,57	43,55
6	44,16	44,27	44,37	43,84	44,16
7	44,16	44,33	44,00	43,79	44,09
8	44,07	43,64	44,14	43,86	43,58
9	43,96	44,03	44,19	43,95	44,24
10	44,26	43,72	44,15	43,86	43,79
Вариант 7					
1	25,98	25,89	26,36	25,77	25,94
2	26,04	25,99	26,05	26,15	26,19
3	25,65	25,57	25,56	26,18	25,66

Номер наблюдения	Уровни фактора А				
	A ₁	A ₂	A ₃	A ₄	A ₅
4	26,07	25,66	25,75	25,55	26,24
5	25,80	25,81	26,38	26,01	25,60
6	26,23	26,50	25,60	26,01	26,22
7	25,76	26,45	25,75	26,11	26,20
8	25,70	25,92	25,77	26,46	25,98
9	25,83	26,26	26,36	26,06	26,07
10	26,10	26,07	25,95	25,65	25,63
Вариант 8					
1	14,23	13,54	14,47	13,88	13,77
2	13,71	13,92	13,63	14,31	13,93
3	14,41	13,72	14,32	14,20	13,56
4	13,58	13,98	13,80	13,72	13,68
5	14,07	13,85	14,15	13,86	14,17
6	13,88	14,35	13,57	13,83	14,16
7	13,77	14,12	14,13	14,17	13,64
8	14,21	13,76	13,92	14,03	13,76
9	13,77	14,21	13,89	13,77	13,79
10	14,39	13,53	13,82	14,26	13,64
Вариант 9					
1	74,28	74,31	74,34	74,37	73,60
2	74,27	73,55	74,39	74,07	74,43
3	73,99	74,18	73,76	73,97	73,83
4	74,05	73,77	73,89	74,39	73,85
5	74,35	73,85	74,37	73,82	74,29
6	73,94	74,42	73,58	73,65	74,40
7	73,84	73,84	73,73	73,91	73,70
8	73,50	74,22	74,28	74,19	73,99
9	74,08	74,46	74,36	73,77	73,56
10	74,23	74,11	74,07	73,57	73,95
Вариант 11					
1	61,20	60,87	60,86	60,61	61,12
2	60,55	61,46	60,76	60,92	61,46
3	60,78	60,95	60,75	61,13	61,00
4	61,35	61,43	60,51	61,03	60,51
5	61,35	61,34	61,21	60,63	61,42
6	60,78	60,60	61,08	61,48	61,26
7	61,40	60,99	60,77	60,58	61,19
8	60,81	60,92	60,83	61,29	60,75
9	61,22	60,90	60,57	60,70	60,60

Номер наблюдения	Уровни фактора А				
	A ₁	A ₂	A ₃	A ₄	A ₅
10	61,18	61,44	61,08	60,59	61,01
Вариант 12					
1	57,36	57,32	57,32	56,51	57,26
2	56,74	57,19	56,50	56,57	57,01
3	56,65	56,68	56,53	57,44	57,04
4	56,76	57,33	57,35	56,99	57,02
5	57,25	57,06	56,98	56,72	57,18
6	57,42	57,37	57,45	56,55	56,53
7	56,67	57,09	57,11	57,27	56,60
8	56,58	56,54	56,69	57,20	56,82
9	57,18	56,58	57,12	56,93	56,89
10	56,99	56,68	57,26	56,58	56,84
Вариант 13					
1	78,74	79,19	79,32	79,04	79,31
2	78,97	79,22	78,61	79,48	78,54
3	78,86	79,44	79,03	78,55	78,64
4	78,95	79,02	78,86	79,11	79,45
5	78,87	78,78	79,24	79,33	78,98
6	78,89	78,94	78,84	79,32	79,47
7	79,50	78,68	78,67	78,67	78,62
8	78,97	79,43	78,77	79,33	78,75
9	79,17	79,03	79,02	79,17	78,65
10	78,61	79,43	78,77	78,95	79,05
Вариант 14					
1	78,81	79,35	79,31	79,49	79,00
2	79,42	79,15	78,89	78,94	78,88
3	78,53	79,27	78,65	79,32	78,55
4	79,22	79,37	79,26	79,05	79,22
5	78,80	79,44	78,84	78,94	79,29
6	78,64	78,79	78,62	79,15	78,58
7	78,63	78,69	78,71	78,86	79,37
8	78,79	78,52	79,44	79,24	78,78
9	78,75	78,91	78,95	79,37	79,18
10	78,96	79,11	78,97	79,19	79,28
Вариант 15					
1	69,25	68,87	68,65	69,16	68,64
2	69,42	69,44	69,29	68,86	68,63
3	69,46	69,32	69,26	68,64	68,66
4	69,41	68,87	68,81	69,46	69,45

Номер наблюдения	Уровни фактора А				
	A_1	A_2	A_3	A_4	A_5
5	69,19	69,28	68,90	69,11	68,55
6	69,19	68,54	68,59	68,60	68,69
7	69,37	68,98	69,44	68,60	68,65
8	69,28	68,76	68,50	69,48	68,94
9	69,03	69,36	68,97	69,43	69,07
10	69,44	68,56	69,14	69,36	69,39

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра машиностроительных технологий и оборудования

УТВЕРЖДАЮ

Проректор по учебной работе

О.Г. Жолмонова

« 14 »

г.



ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Методические указания к выполнению лабораторной
и самостоятельной работы для студентов
по направлению подготовки 15.04.01

Курск 2023

УДК 519.6

Составитель Куц В.В.

Рецензент

Кандидат технических наук, доцент А.Н. Гречухин

Проверка статистических гипотез : методические указания к выполнению лабораторной и самостоятельной работы для студентов по направлению подготовки 15.04.01 / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: В.В. Куц; ЮЗГУ. Курск, 2023. 11 с.: ил. 3.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ. л. 0,87 . Уч.-изд. л. 0,79.

Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: Получить сведения и практические навыки проверки статистических гипотез.

2 Задание:

2.1 По табл. 1 приложения Г в соответствии со своим вариантом выбрать уровень значимости критериев.

2.2 Для исходных данных, приведенных в табл. 2 приложения Г, провести проверку статистических гипотез о равенстве средних значений и дисперсий двух нормально распределенных выборок.

2.3. Для исходных данных, приведенных в табл. 3 приложения Г, провести проверку статистических гипотез о равенстве ряда дисперсий нормально распределенных выборок.

2.4. Оформить отчет.

3 Краткие теоретические сведения

3.1. Задачи статистической проверки гипотез

Статистическая проверка гипотез, т. е. предположений, относящихся к эмпирическим распределениям изучаемых случайных величин, играет важную роль в статистических исследованиях.

Если эмпирическая кривая распределения большой выборки по своему внешнему виду приближается к какому-либо теоретическому закону распределения, то возникает вопрос, можно ли данную выборку рассматривать как выборку из генеральной совокупности, имеющей распределение именно по этому закону. Решение этого вопроса имеет важное значение для исследователя, так как знание закона распределения изучаемой величины позволяет извлечь из экспериментов дополнительную информацию. Если производится две серии испытаний с фактором A и без него и в результате получаются разные значения средних и дисперсий изучаемой переменной величины, то возникает вопрос, является ли это различие в средних и дисперсиях влиянием фактора A или оно носит чисто случайный характер.

Решение перечисленных и им подобных задач в математической статистике производится путем постановки и проверки так называемой «нулевой гипотезы». При этом под «нулевой гипотезой» подразумевается допущение об отсутствии интересующего нас различия

между выборками или их статистическими характеристиками. Например, нас интересует, можно ли по полученному распределению в большой выборке из генеральной совокупности считать, что последняя имеет нормальное распределение. Для того чтобы прийти к вполне определенному заключению, хотя бы и вероятностного характера, мы делаем гипотетическое допущение, что распределение выборки несущественно отличается от нормального и, следовательно, на основании закона больших чисел можно считать, что и генеральная совокупность имеет нормальное распределение. Другими словами, мы выдвигаем «нулевую гипотезу» об отсутствии различия между эмпирическим распределением и теоретическим нормальным или гипотезу о том, что данная выборка взята из нормальной совокупности. Теперь надо проверить эту гипотезу и в результате проверки либо отбросить ее, либо принять.

Для проверки гипотез в математической статистике пользуются рядом критериев, которые называют в этом случае *критериями согласия*. Для того чтобы принять или забраковать гипотезу при помощи этих критериев, установлены уровни значимости их. представляет собой достаточно малое значение вероятности, отвечающее событиям, которые в данной обстановке исследования можно считать практически невозможными. Обычно принимают пяти- двух- или однопроцентный уровень значимости. В технике чаще всего принимают пятипроцентный уровень значимости.

Уровень значимости называют также *доверительным уровнем вероятности*, который соответственно может быть принят равным $q = 0,05$; $0,02$ или $0,01$; иногда принимают $q = 0,001$. Эти уровни доверительной вероятности соответствуют классификации явлений на редкие ($q = 0,05$), очень редкие ($q = 0,01$) и чрезвычайно редкие ($q = 0,001$). Выбирая тот или иной уровень значимости критерия или уровень доверительной вероятности q , мы тем самым устанавливаем и область допустимых его значений, которая выражается вероятностью $P = 1 - q$.

С уменьшением уровня значимости расширяется область допустимых значений критерия и вместе с тем теряется его чувствительность, так как повышается вероятность принять гипотезу даже в тех случаях, когда эта гипотеза неверна. Но вместе с тем выбор доста-

точно малого уровня значимости гарантирует от возможности неправильно забраковать верную гипотезу.

Статистические приемы проверки гипотез не обладают полной определенностью. Если используемый критерий попадает в область допустимых значений, то нельзя еще сделать вывода о правильности гипотезы, а можно лишь заключить, что наблюдаемое значение критерия не противоречит этой гипотезе, что можно признать допустимость гипотезы до тех пор, пока более обстоятельные исследования, с помощью более точных критериев или при увеличенном числе наблюдений не подтвердят это или не приведут к противоположному заключению. Поэтому статистическими методами нельзя пользоваться формально, а необходимо их сочетать с анализом физической сущности изучаемого явления. Когда гипотеза, основанная на теоретическом анализе физической сущности явления, подтверждается также статистическими приемами, то достоверность ее можно считать достаточно надежной.

3.2. Проверка гипотезы равенства двух выборочных средних

Предположим, что из одной и той же генеральной совокупности взяты две выборки, которые для величины x дают средние \bar{X}_1 и \bar{X}_2 , отличные одна от другой. Требуется узнать, случайно или не случайно они отличаются друг от друга. Этот вопрос имеет важное значение при проведении опытов. Если расхождение между \bar{X}_1 и \bar{X}_2 будет существенно, то это может указать на ошибки в опытах или в методике их выполнения, тогда как случайность их расхождения указывает на отсутствие таких ошибок.

Подобный вопрос возникает и при исследовании влияния различных факторов на изучаемый признак. Если опыты с фактором A и без него дали отличные друг от друга \bar{X}_1 и \bar{X}_2 , то при случайном отличии значений их очевидно, что фактор A не влияет на исследуемый признак и, наоборот, влияет при существенном расхождении между \bar{X}_1 и \bar{X}_2 . Наконец, может возникнуть на практике и такой вопрос: принадлежат ли две выборки одной и той же генеральной совокупности. И этот вопрос можно разрешить, сравнивая выборочные средние

\bar{X}_1 и \bar{X}_2 и оценивая их расхождение. Если выборки взяты из одной и той же генеральной совокупности, то расхождение между \bar{X}_1 и \bar{X}_2 будет случайно, и, наоборот, оно будет существенно, когда выборки не будут принадлежать одной и той же совокупности.

Рассмотрим случай, когда выборки берутся из нормальной генеральной совокупности, тогда оценка расхождения двух выборочных средних производится при помощи критерия t Стюдента. Если выборки берутся из нормальной совокупности и при $n < 25$, то величина

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}},$$

где n_1 и n_2 - объем выборок; s_1^2 и s_2^2 - дисперсии выборок, подчинена распределению Стюдента и может быть оценена при помощи таблицы вероятностей $P(|t| \geq t_T)$ (см. приложение А).

При оценке полученного значения t по таблице приложения А необходимо принимать $k = n_1 + n_2 - 2$. Таблица $P(|t| \geq t_T)$ дает вероятность случайных значений t , которые численно не менее наблюдаемого значения t_T . Если эта вероятность будет очень мала (практически, когда $P \leq q$), то наша нулевая гипотеза о несущественном, случайном расхождении между выборочными средними должна быть забракована. Если же вероятность $P(|t| \geq t_T)$ будет достаточно велика (практически, когда $P > q$), то гипотеза однородности выборочных средних может быть принята.

При $n > 25$, t критерий вычисляют по формуле

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

Пример 1. С автомата, обрабатывающего втулки $D = 20^{+0,2}$ мм, было взято в разное время две выборки по 5 шт. каждая. Результаты измерения диаметров втулок приведены в табл. 1 .

Таблица 1- Результаты измерения диаметра втулок

№ выборки	№ детали					\bar{X}	s^2
	1	2	3	4	5		
Выборка 1	20,05	20,08	20,1	20,1	20,09	20,084	0,0004
Выборка 2	20,10	20,15	20,05	20,08	20,10	20,096	0,0013

Распределение диаметров втулок предполагается нормальным. Поскольку выборки взяты из продукции одного и того же станка, можно предполагать, что $\sigma_1^2 = \sigma_2^2$. Пусть теперь наша гипотеза заключается в том, что генеральные средние в моменты взятия выборок были равны между собой, т. е. $\bar{X}_1 = \bar{X}_2$. Другими словами, настройка станка в момент взятия пробы № 1 и № 2 не изменилась.

В результате вычислений средних арифметических и дисперсий выборок получено

$$\bar{X}_1 = 20,084; \bar{X}_2 = 20,096; s_1^2 = 0,0004; s_2^2 = 0,0013.$$

Определим t :

$$t = \frac{|20,084 - 20,096|}{\sqrt{5 \cdot 0,0004 + 5 \cdot 0,0013}} \sqrt{\frac{5 \cdot 5(5 + 5 - 2)}{5 + 5}} = 0,58.$$

Из таблицы приложения А находим, что при $k=5+5-2=8$ вероятность $P(|t| \geq t_\tau) = 0,58$. Принимая $q=0,05$, то вероятность не мала, она значительно больше доверительного уровня $q=0,05$, поэтому наша гипотеза может быть принята.

3.3 Проверка гипотезы равенства двух выборочных дисперсий

Пусть имеются две выборки из нормальной совокупности. Объем каждой выборки равен n_1 и n_2 . Дисперсии этих выборок соответственно равны s_1^2 и s_2^2 . Можно ли считать при наличии некоторых различий между величинами s_1^2 и s_2^2 , что данные выборки принадлежат одной и той же генеральной совокупности? Или можно поставить вопрос так: произведено два опыта, из которых один опыт про-

изводился с фактором A , а другой - без него. Каждый опыт повторялся n раз. В результате обработки статистических данных получено, что дисперсия признака x в опытах с фактором A равна величине s_A^2 , а без него - s_0^2 . Оказывает ли существенное влияние исследуемый фактор A на признак x ? Для ответа на поставленные вопросы необходимо произвести сравнение дисперсий и оценить, является ли существенным их различие. Сравнение дисперсий производится по их отношению:

$$F = \frac{s_1^2}{s_2^2}.$$

В числителе всегда ставится наибольшее значение из двух наблюдаемых дисперсий. Отношение F , когда выборки берутся из нормальной генеральной совокупности, подчиняется обратному F -распределению вероятностей (распределению Фишера). Табличные значения обратного F -распределения вероятностей приведены в таблицах приложения Б (в зависимости от q (0,05 и 0,1), $k_1=n_1-1$ и $k_2=n_2-1$).

Для проверки нашей гипотезы необходимо вычислить наблюдаемое значение F , а затем определить $k_1=n_1-1$ и $k_2=n_2-1$, и найти для этих значений и величины q табличное значение F_T .

Если окажется, что $F \geq F_T$, то наша гипотеза должна быть забракована. Если же окажется, что $F < F_T$ то гипотеза принимается.

Пример 3. С двух автоматов, обрабатывающих одинаковые детали, взято две выборки $n_1=n_2=10$. При этом оказалось, что $s_1^2 = 400$ мкм² и $s_2^2 = 325$ мкм². Ранее было установлено, что рассеивание размеров деталей, обработанных на автоматах, следует нормальному закону распределения.

Можно ли считать, что оба станка обеспечивают одинаковую точность обработки? Предположим, что оба станка дают одинаковую точность и наблюдаемое расхождение между дисперсиями случайно. Для проверки нашей нулевой гипотезы определим критерий F :

$$F = \frac{400}{325} = 1,23.$$

По таблице приложения Б для $q = 0,05$ при $k_1 = k_2 = 9$ находим $F_T = 3,23$, следовательно, $F < F_T$. Поэтому надо считать нашу гипотезу верной, а наблюдаемое различие в значениях дисперсий выборок случайным.

3.4 Проверка гипотезы равенства ряда дисперсий

Пусть имеется m выборок не равных объемов (n_1, n_2, \dots, n_m) взятых из одной или m генеральных совокупностей, имеющих нормальные распределения. При этом дисперсии этих совокупностей имеют одинаковые значения, т. е. $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2$, а математические ожидания могут быть и не равны друг другу.

Дисперсии выборок $s_1^2, s_2^2, \dots, s_m^2$ несколько отличаются друг от друга по величине. Требуется проверить гипотезу о том, что это различие дисперсий выборок носит случайный характер, и, следовательно, дисперсии генеральных совокупностей σ_i^2 , из которых взяты выборки, равны между собой, т. е. $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2$.

Данная гипотеза может быть проверена по критерию Бартлетта, в соответствии с которым случайная величина

$$Q = \frac{2,3026 \left[(N - m) \lg s^2 - \sum_{i=1}^m (n_i - 1) \lg s_i^2 \right]}{1 + \frac{1}{3(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{N - m} \right)},$$

где

$$N = \sum_{i=1}^m n_i ;$$

$$s^2 = \frac{\sum_{i=1}^m s_i^2 (n_i - 1)}{N - m} ,$$

имеет распределение близкое к χ^2 с $k = m - 1$ степенями свободы.

Задаваясь доверительным уровнем вероятности, например, $q = 0,05$ и пользуясь таблицей приложения В, определяют верхний критический предел χ^2 . Если $Q < \chi^2$ то гипотеза принимается, если $Q > \chi^2$, то гипотеза отвергается.

Для вычисления Q рекомендуется составлять вспомогательную табл. 2.

Таблица 2 - Вспомогательная таблица для вычисления

№ вы- борки	s_i^2	n_i	$n_i - 1$	$s_i^2(n_i - 1)$	$\frac{1}{n_i - 1}$	$\lg s_i^2$	$(n_i - 1)\lg s_i^2$
1							
m							
		$\sum n_i$	$\sum (n_i - 1)$	$\sum s_i^2(n_i - 1)$	$\sum \frac{1}{n_i - 1}$		$\sum (n_i - 1)\lg s_i^2$

Если объемы выборок равны, т. е. $n_1 = n_2 = \dots = n_m = n$, то формула для Q принимает вид

$$Q = \frac{2,3026m(n-1) \left[\lg s^2 - \frac{1}{m} \sum_{i=1}^m \lg s_i^2 \right]}{1 + \frac{m+1}{3m(n-1)}}$$

где s^2 вычисляется по формуле

$$s^2 = \frac{1}{m} \sum_{i=1}^m s_i^2$$

Однако при равном объеме выборок проверку гипотезы однородности дисперсий проще производить упрощенным приемом, основанным на вычислении критерия Кохрена G :

$$G = \frac{s_{i \max}^2}{\sum_{i=1}^m s_i^2}$$

Критические значения G_T для уровней значимости 0,05 и 0,1 в зависимости от объема выборок n и числа выборок m приведены в приложении В.

Если найденное по данным выборок G меньше табличного G_T ($G < G_T$), то гипотеза однородности дисперсий генеральных совокупностей, из которых были взяты выборки, принимается. Если $G > G_T$, то гипотеза отвергается.

Пример 4. С четырёх автоматов, настроенных на обработку одних и тех же деталей, взято по одной текущей выборке объема $n_1 = n_2 = n_3 = 10$. Дисперсии выборок имеют следующие значения: $s_1^2 = 100$ мкм², $s_2^2 = 300$ мкм², $s_3^2 = 200$ мкм², $s_4^2 = 400$ мкм². Требуется установить, одинакова ли точность автоматов, т. е. одинаково ли рассеивание случайных погрешностей обработки на этих автоматах, если предварительными исследованиями установлено, что это рассеивание подчиняется закону нормального распределения.

Для решения поставленной задачи необходимо проверить гипотезу однородности выборочных дисперсий. Проверку этой гипотезы произведем при помощи критерия Q Бартлета и критерия G .

Для вычисления критерия Q нужно вычислить s^2 , $\lg s^2$ и $\sum \lg s_i^2$. Получаем следующее

$$s^2 = \frac{100 + 300 + 200 + 400}{4} = 250;$$

$$\lg 100 = 2; \lg 300 = 2,48; \lg 200 = 2,30; \lg 400 = 2,60;$$

$$\sum \lg s_i^2 = 2 + 2,48 + 2,30 + 2,6 = 9,38$$

$$\lg s^2 = \lg 250 = 2,398;$$

$$Q = \frac{2,3026 \cdot 4(10-1) \left[2,398 - \frac{9,38}{4} \right]}{1 + \frac{4+1}{3 \cdot 4(10-1)}}$$

По таблице приложения Б для $k=3$ и доверительной вероятности $P = 0,05$ $\chi^2 = 7,8$. Так как $Q < \chi^2$, то гипотеза однородности диспер-

сии принимается, т. е. наблюдаемые значения s_i^2 отличаются друг от друга случайно. Это подтверждается и критерием G :

$$G = \frac{400}{100 + 300 + 200 + 400} = 0,4$$

По таблице приложения В для доверительной вероятности $q = 0,05$ $G = 0,5$. Так как $G < G_T$, то гипотеза подтверждается.

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение Г), студент изучает теоретические сведения согласно пункту 3. Далее выполняет расчеты аналогичные рассмотренным примерам с учетом имеющихся особенностей задания.

В отчёте по практической работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое «нулевая гипотеза» при проверке статистических гипотез?
2. Что такое критериями согласия при проверке статистических гипотез?
3. Что такое уровень значимости статистических критериев?
4. Как влияет уровень значимости на область допустимых значений критерия и его чувствительность?
5. Проверка гипотезы равенства двух выборочных средних при $n < 25$.

6. Проверка гипотезы равенства двух выборочных средних при $n > 25$.
7. Проверка гипотезы равенства двух выборочных дисперсий.
8. Проверка гипотезы равенства ряда дисперсий по критерию Бартлета для выборок не равных объемов.
9. Проверка гипотезы равенства ряда дисперсий по критерию Бартлета для выборок равных объемов.
10. Проверка гипотезы равенства ряда дисперсий по критерию Кохрена

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
Распределение Стьюдента (t_p)

k	P											
	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,99
1	0,079	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	63,657
2	0,071	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	9,925
3	0,068	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	5,841
4	0,067	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	4,604
5	0,066	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	4,032
6	0,065	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,707
7	0,065	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	3,499
8	0,065	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	3,355
9	0,064	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	3,250
10	0,064	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	3,169
11	0,064	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	3,106
12	0,064	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	3,055
13	0,064	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	3,012
14	0,064	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,977
15	0,064	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,947
16	0,064	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,921
17	0,064	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,898
18	0,064	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,878
19	0,064	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,861
20	0,063	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,845
21	0,063	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,831
22	0,063	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,819
23	0,063	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,807
24	0,063	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,797
25	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,787
26	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,779
27	0,063	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,771
28	0,063	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,763
29	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,756
30	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,750
300	0,063	0,126	0,254	0,386	0,525	0,675	0,843	1,038	1,284	1,650	1,968	2,592

ПРИЛОЖЕНИЕ Б
F-распределение (при $q=0,05$)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
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	8,85	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	4,37
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	2,41
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,69	2,51	2,30
13	4,67	3,81	3,41	3,18	3,03	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	2,37	2,20	2,01	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	3,39	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,23	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	2,31	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,55	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
31	4,16	3,30	2,91	2,68	2,52	2,41	2,25	2,08	1,88	1,61
35	4,12	3,27	2,87	2,64	2,49	2,37	2,22	2,04	1,83	1,56
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,20	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,47
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	2,53	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	1,33
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,03	1,85	1,63	1,28
∞	3,84	3,00	2,61	2,37	2,21	2,10	1,94	1,75	1,52	1,03

F-распределение (при $q=0,1$)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
2	8,53	9,00	9,16	9,24	9,29	9,33	9,37	9,41	9,45	9,49
3	5,54	5,46	5,39	5,34	5,31	5,28	5,25	5,22	5,18	5,13
4	4,54	4,32	4,19	4,11	4,05	4,01	3,95	3,90	3,83	3,76
5	4,06	3,78	3,62	3,52	3,45	3,40	3,34	3,27	3,19	3,11
6	3,78	3,46	3,29	3,18	3,11	3,05	2,98	2,90	2,82	2,72
7	3,59	3,26	3,07	2,96	2,88	2,83	2,75	2,67	2,58	2,47
8	3,46	3,11	2,92	2,81	2,73	2,67	2,59	2,50	2,40	2,29
9	3,36	3,01	2,81	2,69	2,61	2,55	2,47	2,38	2,28	2,16
10	3,29	2,92	2,73	2,61	2,52	2,46	2,38	2,28	2,18	2,06
11	3,23	2,86	2,66	2,54	2,45	2,39	2,30	2,21	2,10	1,97
12	3,18	2,81	2,61	2,48	2,39	2,33	2,24	2,15	2,04	1,90
13	3,14	2,76	2,56	2,43	2,35	2,28	2,20	2,10	1,98	1,85
14	3,10	2,73	2,52	2,39	2,31	2,24	2,15	2,05	1,94	1,80
15	3,07	2,70	2,49	2,36	2,27	2,21	2,12	2,02	1,90	1,76
16	3,05	2,67	2,46	2,33	2,24	2,18	2,09	1,99	1,87	1,72
17	3,03	2,64	2,44	2,31	2,22	2,15	2,06	1,96	1,84	1,69
18	3,01	2,62	2,42	2,29	2,20	2,13	2,04	1,93	1,81	1,66
19	2,99	2,61	2,40	2,27	2,18	2,11	2,02	1,91	1,79	1,63
20	2,97	2,59	2,38	2,25	2,16	2,09	2,00	1,89	1,77	1,61
21	2,96	2,57	2,36	2,23	2,14	2,08	1,98	1,87	1,75	1,59
22	2,95	2,56	2,35	2,22	2,13	2,06	1,97	1,86	1,73	1,57
23	2,94	2,55	2,34	2,21	2,11	2,05	1,95	1,84	1,72	1,55
24	2,93	2,54	2,33	2,19	2,10	2,04	1,94	1,83	1,70	1,53
25	2,92	2,53	2,32	2,18	2,09	2,02	1,93	1,82	1,69	1,52
26	2,91	2,52	2,31	2,17	2,08	2,01	1,92	1,81	1,68	1,50
27	2,90	2,51	2,30	2,17	2,07	2,00	1,91	1,80	1,67	1,49
28	2,89	2,50	2,29	2,16	2,06	2,00	1,90	1,79	1,66	1,48
29	2,89	2,50	2,28	2,15	2,06	1,99	1,89	1,78	1,65	1,47
30	2,88	2,49	2,28	2,14	2,05	1,98	1,88	1,77	1,64	1,46
31	2,87	2,48	2,27	2,14	2,04	1,97	1,88	1,77	1,63	1,45
35	2,85	2,46	2,25	2,11	2,02	1,95	1,85	1,74	1,60	1,41
40	2,84	2,44	2,23	2,09	2,00	1,93	1,83	1,71	1,57	1,38
45	2,82	2,42	2,21	2,07	1,98	1,91	1,81	1,70	1,55	1,35
50	2,81	2,41	2,20	2,06	1,97	1,90	1,80	1,68	1,54	1,33
60	2,79	2,39	2,18	2,04	1,95	1,87	1,77	1,66	1,51	1,29
70	2,78	2,38	2,16	2,03	1,93	1,86	1,76	1,64	1,49	1,27
80	2,77	2,37	2,15	2,02	1,92	1,85	1,75	1,63	1,48	1,25
90	2,76	2,36	2,15	2,01	1,91	1,84	1,74	1,62	1,47	1,23
100	2,76	2,36	2,14	2,00	1,91	1,83	1,73	1,61	1,46	1,22
∞	2,71	2,30	2,08	1,95	1,85	1,77	1,67	1,55	1,38	1,03

ПРИЛОЖЕНИЕ В
Критические значения G
 при $q=0,05$

m	n-1										
	2	3	4	5	6	7	8	9	10	16	32
2	0,9750	0,9392	0,9057	0,8772	0,8534	0,8332	0,8159	0,8010	0,7880	0,7341	0,6694
3	0,8709	0,7977	0,7457	0,7070	0,6770	0,6531	0,6333	0,6167	0,6025	0,5466	0,4835
4	0,7679	0,6839	0,6287	0,5894	0,5598	0,5365	0,5175	0,5018	0,4884	0,4365	0,3797
5	0,6838	0,5981	0,5440	0,5063	0,4783	0,4564	0,4387	0,4241	0,4118	0,3645	0,3134
6	0,6161	0,5321	0,4803	0,4447	0,4184	0,3980	0,3817	0,3682	0,3568	0,3136	0,2674
7	0,5612	0,4800	0,4307	0,3972	0,3726	0,3536	0,3384	0,3259	0,3154	0,2756	0,2334
8	0,5157	0,4377	0,3910	0,3594	0,3362	0,3185	0,3043	0,2927	0,2829	0,2461	0,2073
9	0,4775	0,4027	0,3584	0,3285	0,3067	0,2901	0,2768	0,2659	0,2568	0,2226	0,1867
10	0,4450	0,3733	0,3311	0,3028	0,2823	0,2666	0,2541	0,2439	0,2353	0,2033	0,1698
12	0,3924	0,3264	0,2880	0,2624	0,2440	0,2299	0,2187	0,2096	0,2020	0,1735	0,1441
15	0,3346	0,2758	0,2419	0,2195	0,2034	0,1912	0,1815	0,1737	0,1671	0,1427	0,1177
20	0,2705	0,2205	0,1921	0,1735	0,1602	0,1502	0,1422	0,1358	0,1305	0,1106	0,0905
24	0,2354	0,1908	0,1656	0,1491	0,1374	0,1286	0,1216	0,1160	0,1113	0,0940	0,0765
30	0,1979	0,1593	0,1377	0,1236	0,1137	0,1061	0,1003	0,0955	0,0915	0,0769	0,0622
40	0,1575	0,1258	0,1082	0,0968	0,0887	0,0827	0,0779	0,0741	0,0709	0,0593	0,0477
60	0,1132	0,0895	0,0765	0,0682	0,0623	0,0579	0,0544	0,0517	0,0494	0,0410	0,0326
120	0,0633	0,0494	0,0419	0,0371	0,0337	0,0312	0,0292	0,0277	0,0264	0,0217	0,0170

при $q=0,1$

m	n-1										
	2	3	4	5	6	7	8	9	10	16	32
2	0,9500	0,9027	0,8646	0,8347	0,8107	0,7911	0,7747	0,7607	0,7486	0,7000	0,6434
3	0,8174	0,7430	0,6934	0,6578	0,6307	0,6092	0,5918	0,5772	0,5647	0,5161	0,4619
4	0,7076	0,6286	0,5787	0,5438	0,5178	0,4974	0,4809	0,4672	0,4557	0,4110	0,3621
5	0,6239	0,5462	0,4983	0,4654	0,4409	0,4220	0,4067	0,3941	0,3835	0,3427	0,2988
6	0,5591	0,4842	0,4389	0,4079	0,3852	0,3676	0,3535	0,3418	0,3321	0,2947	0,2548
7	0,5074	0,4359	0,3931	0,3640	0,3427	0,3263	0,3132	0,3025	0,2934	0,2590	0,2225
8	0,4653	0,3970	0,3565	0,3292	0,3092	0,2939	0,2817	0,2716	0,2632	0,2314	0,1977
9	0,4302	0,3651	0,3267	0,3009	0,2821	0,2677	0,2562	0,2469	0,2390	0,2093	0,1780
10	0,4005	0,3383	0,3018	0,2774	0,2596	0,2461	0,2353	0,2264	0,2190	0,1912	0,1620
12	0,3529	0,2958	0,2626	0,2405	0,2245	0,2123	0,2026	0,1947	0,1881	0,1633	0,1375
15	0,3009	0,2500	0,2207	0,2013	0,1874	0,1767	0,1684	0,1615	0,1558	0,1345	0,1124
20	0,2434	0,2001	0,1756	0,1594	0,1478	0,1390	0,1321	0,1265	0,1218	0,1044	0,0865
24	0,2120	0,1734	0,1515	0,1372	0,1269	0,1192	0,1131	0,1081	0,1040	0,0888	0,0732
30	0,1785	0,1450	0,1262	0,1139	0,1051	0,0985	0,0933	0,0891	0,0856	0,0727	0,0596
40	0,1424	0,1147	0,0993	0,0893	0,0822	0,0769	0,0727	0,0693	0,0665	0,0561	0,0457
60	0,1028	0,0820	0,0705	0,0631	0,0579	0,0540	0,0509	0,0484	0,0464	0,0389	0,0313
120	0,0578	0,0455	0,0388	0,0345	0,0315	0,0292	0,0275	0,0260	0,0249	0,0206	0,0164

ПРИЛОЖЕНИЕ Г
Исходные данные к работе

Таблица 1

	№ варианта														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>q</i>	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05

Таблица 2

№ вар-та	№ измерения									
	1	2	3	4	5	6	7	8	9	10
1	10,17	9,36	10,22	10,07	10,11	9,87	10,47	8,98	9,40	9,38
	10,46	9,85	10,71	10,74	10,35	10,03	10,52	9,50	10,14	10,51
2	30,39	30,15	30,86	30,16	29,93	30,07	29,24	29,99	28,75	30,15
	29,97	30,21	30,36	29,90	30,84	30,13	30,41	29,92	29,90	30,10
3	49,99	49,95	49,96	49,94	49,91	49,86	50,03	50,08	49,95	49,99
	50,20	50,13	50,03	50,14	50,13	50,03	50,16	50,03	50,03	50,07
4	70,46	70,03	70,00	70,30	69,75	70,47	70,02	69,72	70,58	68,59
	70,18	70,06	70,11	70,66	70,03	70,95	70,09	69,36	70,20	69,38
5	89,97	90,02	90,02	89,83	89,99	90,06	89,99	90,01	90,04	90,01
	90,05	89,90	89,95	89,96	89,91	89,99	89,93	90,06	89,96	89,99
6	110,03	109,56	110,34	110,43	109,62	110,37	109,99	110,17	110,25	110,20
	110,10	109,94	110,04	110,21	110,65	110,56	110,11	110,29	110,60	109,70
7	130,10	129,98	129,95	130,01	130,13	130,14	129,99	129,98	130,00	129,96
	130,07	130,08	130,08	130,05	130,04	129,99	130,02	129,98	130,00	130,04
8	149,74	150,00	150,24	150,04	149,98	150,28	150,00	149,94	149,71	150,16
	149,95	149,91	150,50	149,80	149,95	150,23	150,07	150,01	150,02	150,24
9	170,34	170,12	169,86	170,17	169,87	169,90	169,94	169,94	170,18	169,85
	170,04	169,99	169,89	170,05	169,71	170,12	170,01	169,98	170,00	169,96
10	189,85	189,88	190,29	189,38	190,37	189,82	189,37	189,62	189,41	189,18
	189,99	190,91	189,74	190,04	189,95	189,43	189,95	190,37	189,88	189,60
11	210,40	210,83	209,76	209,82	209,95	210,12	209,85	210,87	210,11	210,21
	209,89	210,14	209,94	210,14	209,32	210,45	210,05	210,79	210,51	209,54
12	230,49	230,13	230,08	230,06	230,01	229,54	229,68	230,19	229,63	230,14
	230,36	230,06	230,75	230,02	230,25	230,14	230,07	229,94	229,50	230,20
13	250,21	249,94	250,11	249,87	249,84	250,18	250,16	250,19	249,93	250,28
	249,96	250,31	250,15	249,63	249,86	250,44	250,78	250,15	250,22	250,24
14	269,91	269,48	269,64	270,03	270,09	269,85	269,65	269,81	269,72	269,73
	269,77	270,25	270,24	270,29	269,96	269,91	270,42	270,01	270,10	269,96
15	290,16	289,82	289,66	290,09	290,17	289,47	290,19	290,45	290,18	289,95
	290,28	290,13	290,00	289,99	290,09	290,40	289,94	290,55	289,87	290,03

Таблица 3

№ вар-та		№ выборки				
		1	2	3	4	5
1	n_i	24	24	24	24	24
	s_i^2	430	150	170	400	400
2	n_i	16	19	43	49	28
	s_i^2	250	360	150	360	500
3	n_i	50	50	50	50	50
	s_i^2	130	370	230	240	320
4	n_i	20	31	46	29	39
	s_i^2	190	410	180	310	280
5	n_i	18	18	18	18	18
	s_i^2	370	420	370	460	280
6	n_i	36	31	29	32	24
	s_i^2	420	470	240	170	420
7	n_i	24	24	24	24	24
	s_i^2	250	410	310	220	340
8	n_i	35	3	7	44	47
	s_i^2	400	410	130	130	470
9	n_i	17	17	17	17	17
	s_i^2	310	290	110	180	110
10	n_i	1	3	43	44	41
	s_i^2	190	450	220	130	120
11	n_i	45	45	45	45	45
	s_i^2	490	460	130	390	420
12	n_i	27	36	39	0	45
	s_i^2	130	120	450	120	180
13	n_i	19	19	19	19	19
	s_i^2	300	290	230	380	230
14	n_i	22	35	40	17	25
	s_i^2	450	260	350	370	350
15	n_i	38	38	38	38	38
	s_i^2	420	480	280	270	200