

Лабораторная работа по статистике
ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ
ФЕДЕРАЦИИ
Сн=24, Нгр=1

Задача 1

Организация статистического наблюдения. Сводка и группировка.

1.1. Определение необходимого и достаточного числа наблюдений.

Условие: Статистическое управление проводит выборочное обследование предприятий малого и среднего бизнеса. Планируется определение и оценка ряда показателей, характеризующих уровень и эффективность деятельности этих предприятий, в том числе общей и средней численности персонала. По предварительным данным известно, что численность работников на них составляет:

$$S_{\max} = (95 + 5 \times C_n) \text{ чел.} = 215 \text{ чел.}, \quad S_{\min} = 40 \text{ чел.}$$

Достоверность P и предельная ошибка Δ оцениваемых показателей должны составить:

$$P = 0,6722 + 0,0105 \times C_n; \quad \Delta = 0,02 + 0,01 \times N_{\text{гр}} .$$

Для наших данных $P=0,9242$; $\Delta = 0,03$.

Определить: Объем выборки n , обеспечивающей заданную достоверность P и предельную ошибку выборки Δ . Расчет провести для случаев, когда объем генеральной совокупности N неизвестен и когда объем N известен и составляет:

$$N = 5000 + 50 \times C_n \times N_{\text{гр}} = 6200,$$

где C_n – номер по списку в журнале группы;

$N_{\text{гр}}$ – номер группы

Решение:

1. Заданы P и Δ (планируется повторная выборка)

$$n_1 = \frac{0.25t^2}{\Delta^2}$$

где t - коэффициент доверия или достоверности, выбираемый по таблице значений интеграла вероятностей Лапласа по заданному значению P (см. практикум, стр. 481).

P	0,6827	0,7887	0,8664	0,9199	0,9545	0,9756	0,9876	0,9940	0,9973
t	1,00	1,25	1,50	1,75	2,00	2,25	2,50	2,75	
	3,00								

Для $P=0,9242$, $t = 1,78$, $n_1 = \frac{0,25 \cdot 1,78^2}{0,03^2} = 880,11$, то есть, для обеспечения заданной достоверности необходимо для обследования выбрать 881 предприятие.

2. Заданы P , Δ , N (планируется бесповторная выборка)

$$n_2 = \frac{0,25Nt^2}{\Delta^2N + 0,25t^2}.$$

Для наших данных получится $n_2 = \frac{0,25 \cdot 6200 \cdot 1,78^2}{0,03^2 \cdot 6200 + 0,25 \cdot 1,78^2} = 770,7$. То есть при бесповторной выборке для обеспечения заданной достоверности нам необходимо отобрать 771 предприятие.

3. Заданы P , Δ , σ (планируется повторная выборка)

$$n_3 = \frac{\sigma^2 t^2}{\Delta^2},$$

где σ^2 - дисперсия, определить которую можно приближённо, исходя из предположения, что показатели численности работников предприятий подчиняются нормальному закону распределения, следующим образом:

$$\sigma \approx \frac{S_{\max} - S_{\min}}{6}.$$

Поскольку предельная ошибка выборки Δ задана в относительном выражении, необходимо и её выразить в абсолютных единицах. Для этого рассчитаем приближенное значение средней численности \bar{S} :

$$\bar{S} \approx \frac{S_{\max} + S_{\min}}{2}.$$

Абсолютное значение предельной ошибки выборки будет равно

$$\Delta_{\text{абс}} \approx \Delta \times \bar{S}.$$

Для наших данных $\sigma \approx \frac{215 - 40}{6} = 29,17$, $\bar{S} \approx \frac{215 + 40}{2} = 127,5$,

$$\Delta_{\text{абс}} \approx 0,03 \times 127,5 = 3,825,$$

$$n_3 = \frac{29,17^2 \cdot 1,78^2}{0,03^2} = 184,22.$$

То есть, чтобы обеспечить заданную надежность нужно отобрать 185 предприятий.

4. Заданы P, Δ, N, σ^2 (планируется бесповторная выборка)

$$n_4 = \frac{\sigma^2 N t^2}{\Delta^2 N + \sigma^2 t^2} \cdot$$

Для наших данных $n_4 = \frac{29,17^2 \cdot 6200 \cdot 1,78^2}{3,825^2 \cdot 6200 + 29,17^2 \cdot 1,78^2} = 178,9$.

То есть, чтобы обеспечить заданную надежность необходимо отобрать 179 предприятий.

Контроль: Значения объема выборки должны подчиняться следующему соотношению:

$$n_1 > n_2 > n_3 > n_4 .$$

Сделать выводы о рекомендуемой выборке, её объеме.

Выводы. Чтобы обеспечить заданный уровень надежности 0,9242 необходимо для обследования отобрать

А) при повторной выборке из генеральной совокупности неизвестного объема 881 предприятие;

Б) при бесповторной выборке из генеральной совокупности объемом 6200 предприятий 771 предприятие;

В) при повторной выборке из генеральной совокупности неизвестного объема при заданной дисперсии 185 предприятий;

Г) при бесповторной выборке из генеральной совокупности объемом 6200 предприятий и данной дисперсии 179 предприятий.

1.2. Сводка и группировка результатов выборочного наблюдения.

Условие: По материалам выборочного обследования получены данные о численности S_i персонала $n = 20$ предприятий (см. таблицу 1.1):

Таблица 1.1 – Результаты выборочного наблюдения

№	1	2	3	4	5	6	7	8	9	1	1	1	1	1	1	1	1	1	2	
п					0					0	1	2	3	4	5	6	7	8	9	0
S_i	9	S_{\max}	8	6	6	7	4	9	7	4	7	6	7	7	5	9	7	6	9	5
чел.	5		5	1	0	0	0	0	5	5	5	0	0	4	5	8	6	5	4	0

Варианты задания: К показателям численности персонала предприятий № 5, 10, 15, 20 (выделены в таблице) прибавить пять номеров C_n по списку в журнале группы:

$$S_i + 5 \times C_n$$

Таблица с нашими данными

№	1	2	3	4	5	6	7	8	9	10	1	1	1	1	15	1	1	1	1	20
---	---	---	---	---	---	---	---	---	---	----	---	---	---	---	----	---	---	---	---	----

<i>n</i>											1	2	3	4		6	7	8	9	
<i>S_i</i>	9	21	8	6	18	7	4	9	7	16	7	6	7	7	17	9	7	6	9	17
чел.	5	5	5	1	0	0	0	0	5	5	5	0	0	4	5	8	6	5	4	0

Содержание задания:

- определить число групп и шаг интервала для интервального ряда распределения;
- построить ряд распределения предприятий по численности персонала с указанием частоты f_i , частости w_i , и их кумулят F_i и W_i соответственно (см. таблицу 1.2);
- построить для данного ряда, используя средства Excel, гистограмму распределения, кумулятивную кривую (см. рис. 1.1 и 1.2).

Число групп k и шаг интервала h рассчитать по формулам:

$$k \approx 1 + \log_2 n ; \quad h \approx \frac{S_{\max} - S_{\min}}{k} .$$

У нас $k \approx 1 + \log_2 20 = 5,32$, то есть разобьем ряд на 5 интервалов.

Длина каждого интервала будет равна $h \approx \frac{215 - 40}{5} = 35$ единиц.

Таблица 1.2 – Интервальный ряд распределения

S_{инт.}	40-75	75-110	110-145	145-180	180-215	Σ
Счёт	57,5	92,5	127,5	162,5	197,5	
Частота f_i	7	8	0	3	2	20
Частость w_i	0,35	0,4	0	0,15	0,1	1
Кумулята F_i	7	15	15	18	20	
Кумулята W_i	0,35	0,75	0,75	0,9	1	

Гистограмма строится по частотам f_i или частостям w_i . Кумулятивная кривая строится по накопленным частотам F_i или частостям W_i

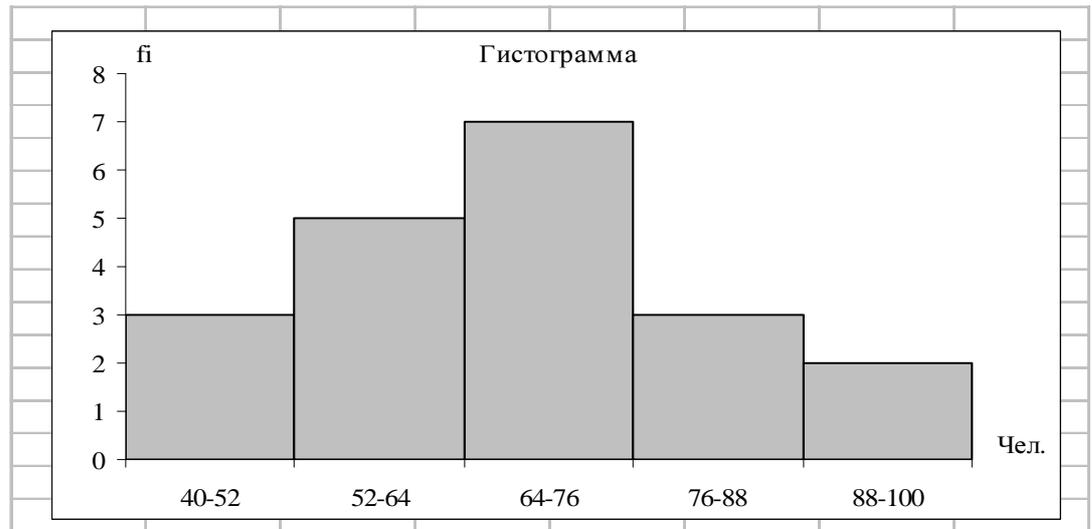


Рис. 1.1. Гистограмма распределения (пример)

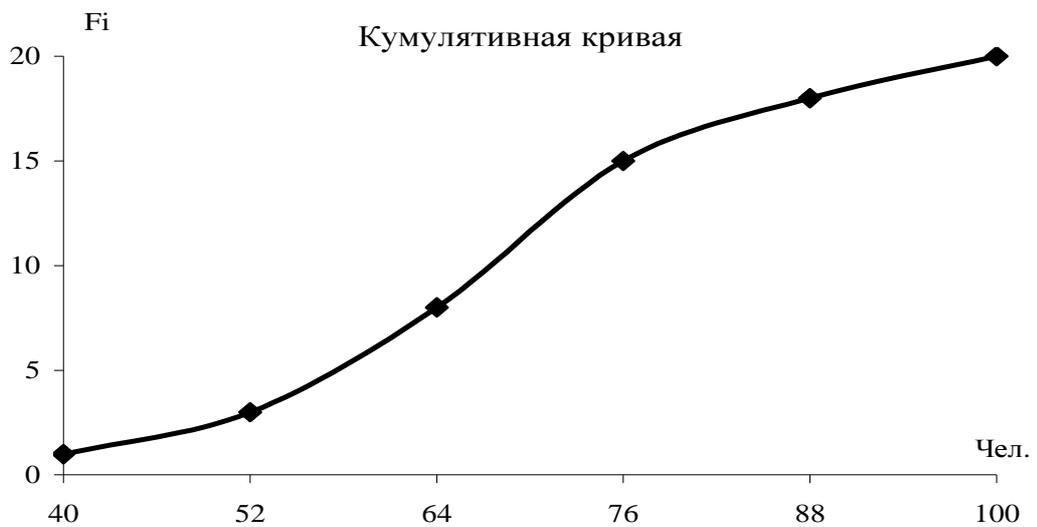
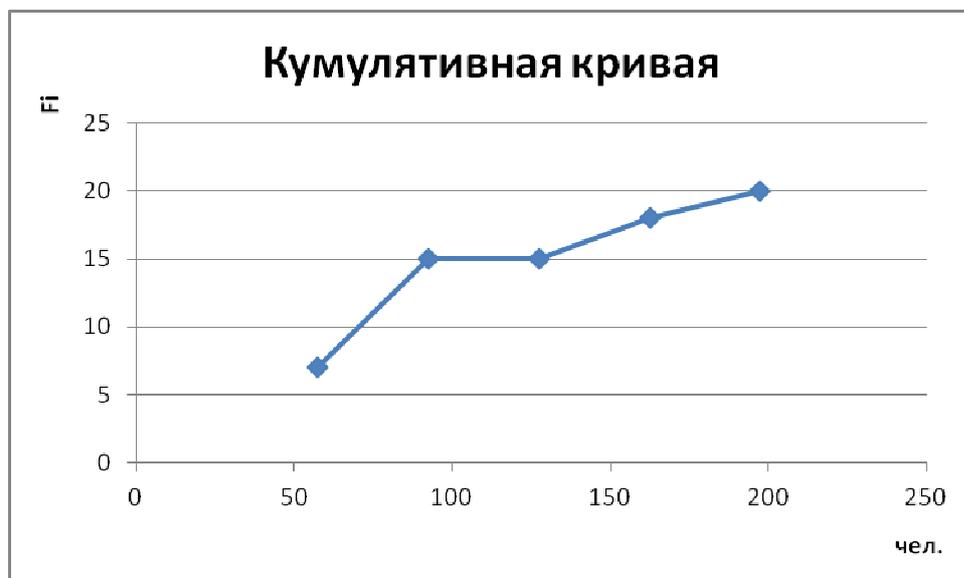
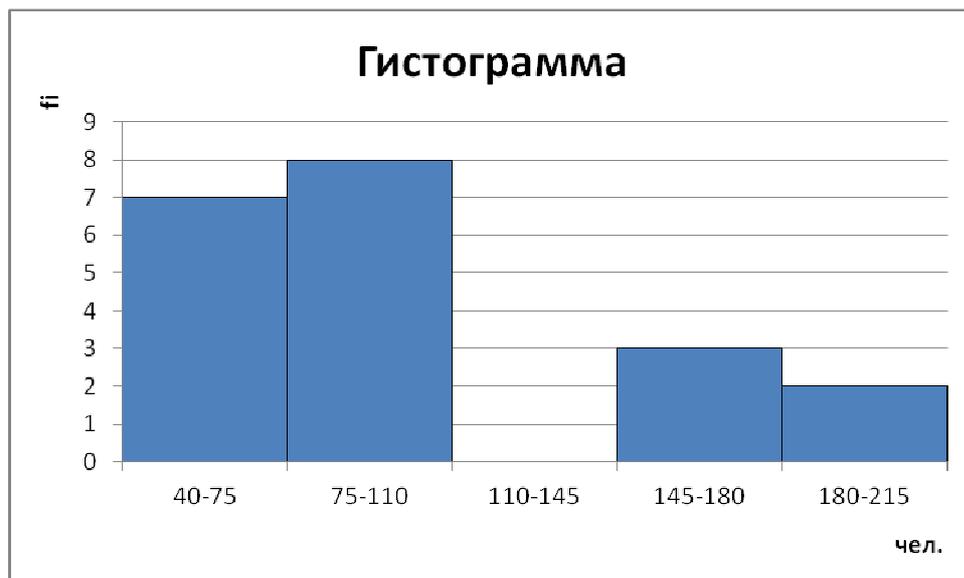


Рис. 1.2. Кумулятивная кривая (пример)

Для наших данных.



Задача 2

Средние величины в статистике

Исходные данные. Численность персонала предприятий и данные из ряда распределения из задачи 1 (табл. 1.1 и 1.2).

Задание: Для оценки выборочной средней \tilde{S} численности персонала предприятий, для оценки свойств средних и для определения предельной ошибки и доверительного интервала генеральной средней \bar{S} :

- вычислить среднюю арифметическую, среднюю квадратическую, среднюю геометрическую, среднюю гармоническую (простую по данным из табл. 1.1 и взвешенную по данным из табл. 1.2).

- по сгруппированным данным (табл. 1.2) вычислить моду (**Mo**) и медиану (**Me**).
- вычислить дисперсию, среднюю μ и предельную Δ ошибки выборочной средней \tilde{S} .

2.1. Вычисление степенных средних:

Средняя арифметическая

- простая (табл. 1.1)
$$\tilde{S}_{\text{ар}} = \frac{\sum S_i}{n} = \frac{2033}{20} = 101,65$$
- взвешенная (табл. 1.2)

$$\tilde{S}_{\text{ар}} = \frac{\sum S_i f_i}{\sum f_i} = \frac{57,5 \cdot 7 + 92,5 \cdot 8 + 162,5 \cdot 3 + 197,5 \cdot 2}{20} = \frac{2025}{20} = 101,25$$

Средняя квадратическая

- простая
$$\tilde{S}_{\text{кв}} = \sqrt{\frac{\sum S_i^2}{n}} = \sqrt{\frac{253613}{20}} = 112,61$$
- взвешенная
$$\tilde{S}_{\text{кв}} = \sqrt{\frac{\sum S_i^2 f_i}{\sum f_i}} = \sqrt{\frac{248825}{20}} = 111,54$$

Средняя геометрическая

- простая
$$\tilde{S}_{\text{геом}} = \sqrt[n]{\prod_{i=1}^n S_i} = 91,97$$
- взвешенная
$$\tilde{S}_{\text{геом}} = \sqrt[n]{\prod_{i=1}^n S_i^{f_i}} = 91,94$$

Средняя гармоническая

- простая
$$\tilde{S}_{\text{гарм}} = \frac{n}{\sum \frac{1}{S_i}} = \frac{20}{0,237} = 84,22$$
- взвешенная
$$\tilde{S}_{\text{гарм}} = \frac{\sum f_i}{\sum \frac{f_i}{S_i}} = \frac{20}{0,237} = 84,45$$

Выводы:

- согласно правилу мажорантности, между средними должно соблюдаться следующее соотношение:

$$\tilde{S}_{\text{кв}} > \tilde{S}_{\text{ар}} > \tilde{S}_{\text{геом}} > \tilde{S}_{\text{гарм}}$$

У нас выполняется для не сгруппированных и сгруппированных данных.

- средняя численность персонала предприятий, вычисленная по данным выборочного наблюдения, равна средней арифметической:

$$\tilde{S} = \tilde{S}_{ap} = 101,65$$

- вычисление других видов средних иллюстрирует свойства степенных средних.

2.2. Вычисление структурных средних

Мода

- аналитическое определение моды (по данным из табл. 1.2)

$$M_o = S_{M_o} + h \times \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})},$$

где: S_{M_o} - левая или нижняя граница модального интервала;

h - шаг или величина интервала;

f_{M_o} - частота модального интервала;

f_{M_o-1} - частота предшествующего интервала;

f_{M_o+1} - частота последующего интервала.

- графическое определение моды (по гистограмме из задания 1, рис. 2.1)

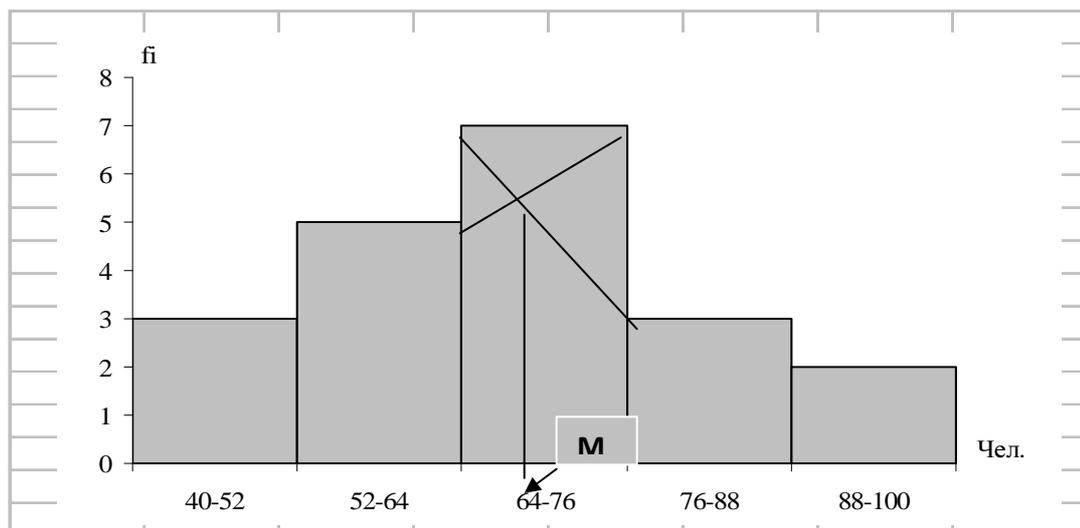
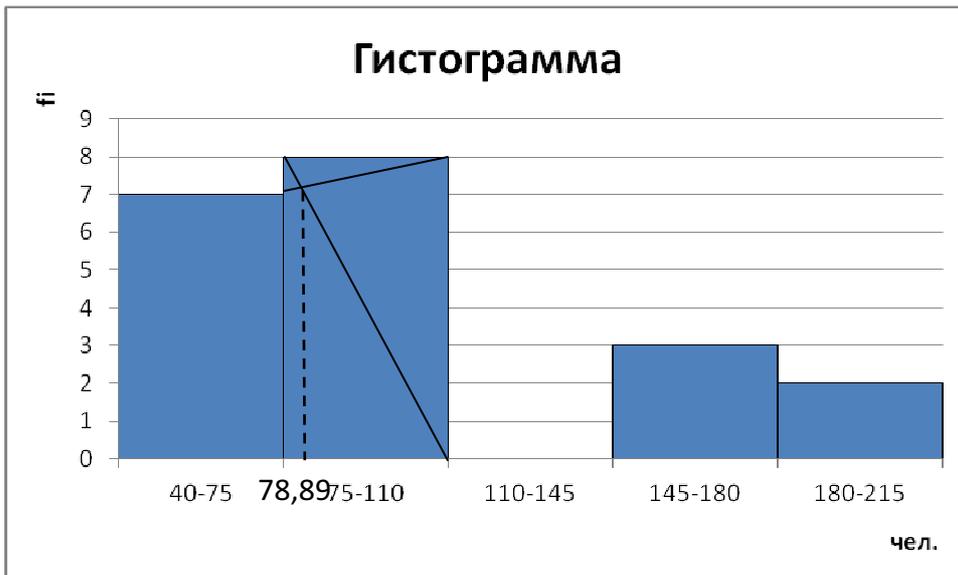


Рис. 2.1. Графическое определение моды (пример)

По нашим данным $M_o = 75 + 35 \times \frac{8 - 7}{(8 - 7) + (8 - 0)} = 78,89$



Медиана

- аналитическое определение медианы (по данным из табл. 1.2)

$$Me = Me = S_{Me} + h \times \frac{N_{Me} - F_{Me-1}}{f_{Me}},$$

где: S_{Me} - левая или нижняя граница медианного интервала (по кумуляте F_i);

h - шаг или величина интервала;

N_{Me} - порядковый номер медианы в ранжированном ряду (

$$N_{Me} = \frac{\sum f_i}{2} = \frac{n}{2});$$

F_{Me-1} - кумулята предшествующего интервала;

f_{Me} - частота медианного интервала;

- графическое определение медианы (по кумулятивной кривой из задания 1, рис. 2.2.)

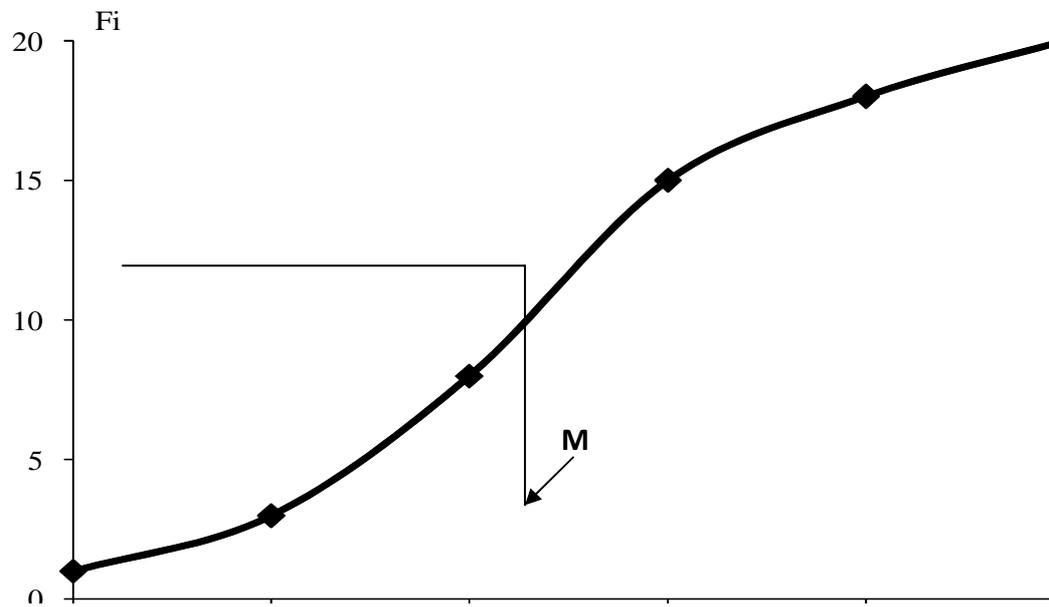


Рис. 2.2. Графическое определение медианы (пример)

По нашим данным $Me = 75 + 35 \times \frac{10 - 7}{8} = 88,125$



2.3. Дисперсия σ^2 , средняя μ и предельная Δ ошибки выборочной средней \tilde{S}

- дисперсия показателей численности персонала предприятий (по средним интервальным значениям численности из табл. 1.2)

$$\sigma^2 = \overline{S^2} - \tilde{S}^2 ,$$

где: \tilde{S}^2 - квадрат средней арифметической;

$\overline{S^2}$ - средняя из квадратов показателей численности персонала.

$$\overline{S^2} = \frac{\sum S_i^2 f_i}{\sum f_i} ;$$

По нашим данным $\overline{S^2} = \frac{248825}{20} = 12441,25;$

$$\sigma^2 = 12441,25 - 101,25^2 = 2189,69$$

- среднее квадратическое отклонение показателей численности персонала предприятий

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

По нашим данным $\sigma = \sqrt{2189,69} = 46,79$

- коэффициент вариации показателей

$$V_\sigma = \frac{\sigma}{\tilde{S}} 100 \%$$

По нашим данным $V_\sigma = \frac{46,79}{101,25} 100 \% = 46,22 \%$

По полученному значению коэффициента вариации **сделать вывод.**

- средняя ошибка выборочной средней

$$\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$$

По нашим данным $\mu = \sqrt{\frac{2189,69}{20} \left(1 - \frac{20}{6200}\right)} = 10,45$

- предельная ошибка выборочной средней

$$\Delta = \mu * t ,$$

где t - коэффициент доверия или достоверности, выбранный или рассчитанный в задаче 1.

По нашим данным $\Delta = 10,45 \cdot 1,78 = 18,59$

Вывод:

По полученному значению выборочной средней \tilde{S} и предельной ошибке этой средней Δ с достоверностью P можно утверждать, что генеральная средняя численность персонала \bar{S} предприятий лежит в интервале:

$$\tilde{S} - \Delta \leq \bar{S} \leq \tilde{S} + \Delta$$

$$101,25 - 18,59 \leq \bar{S} \leq 101,25 + 18,59$$

$$82,66 \leq \bar{S} \leq 119,84$$

Задача 3

Дисперсионный анализ. Общая, внутригрупповая и межгрупповая дисперсии

Условие: При выборочном обследовании установлено, что в общей совокупности из $n = 20$ представлены 8 предприятий малого бизнеса и 12 предприятий среднего бизнеса. Проведена перегруппировка предприятий (по данным из табл. 1.1) по группировочному признаку-фактору. Результаты перегруппировки приведены в таблице 3.1.

Таблица 3.1 – Распределение предприятий по группировочному признаку-фактору

Группы предприятий	Численность персонала S_i								$\sum S$	$\sum S_i^2$
	40	65	61	75	70	74	70	85		
Малые предприятия	40	65	61	75	70	74	70	85	540	37672
Средние предприятия	95	215	180	90	165	75			1493	21594 1
	60	175	98	76	94	170				

Задание:

- установить и назвать группировочный признак-фактор;
- определить общую и внутригрупповые средние показатели численности персонала;
- определить общую, внутригрупповые, среднюю из внутригрупповых и межгрупповую дисперсии средней численности;
- рассчитать эмпирический коэффициент детерминации и эмпирическое корреляционное отношение;
- сделать выводы о влиянии группировочного признака-фактора на вариацию средней численности по группам предприятий.

Предварительно в таблице 3.1 вычислить сумму показателей численности и сумму квадратов показателей численности по каждой группе предприятий.

Алгоритм решения задачи следующий:

Средние показатели численности:

- **средняя численность** для каждой группы предприятий:

для малых предприятий
$$\bar{S}_{м(1)} = \frac{\sum S_{i1}}{n_1};$$

$$\bar{S}_{м(1)} = \frac{540}{8} = 67,5$$

для средних предприятий
$$\bar{S}_{с(2)} = \frac{\sum S_{i2}}{n_2}$$

$$\bar{S}_{с(2)} = \frac{1493}{12} = 124,42.$$

- **общая средняя численность** для всех предприятий:

$$\bar{S}_o = \frac{\sum \bar{S}_j n_j}{\sum n_j} \quad \text{или} \quad \bar{S}_o = \frac{\sum S_i}{n_i}$$

$$\bar{S}_o = \frac{67,5 \cdot 8 + 124,42 \cdot 12}{20} = 101,65$$

Общая, внутригрупповая и межгрупповая дисперсии:

Для вычисления общей и внутригрупповых дисперсий определим **средние из квадратов:**

- для малых предприятий
$$\bar{S}^2_{м(1)} = \frac{\sum S_{i1}^2}{n_1};$$

$$\bar{S}^2_{м(1)} = \frac{37672}{8} = 4709$$

- для средних предприятий
$$\bar{S}^2_{с(2)} = \frac{\sum S_{i2}^2}{n_{21}}$$

$$\bar{S}^2_{с(2)} = \frac{215941}{12} = 17995,08.$$

Внутригрупповые дисперсии:

- для малых предприятий:

$$\sigma_M^2 = \bar{S}^2_M - \bar{S}_M^2;$$

$$\sigma_m^2 = 4709 - 67,5^2 = 152,75;$$

- для средних предприятий:
- $\sigma_c^2 = \overline{S^2}_c - \overline{S}_c^2$

$$\sigma_c^2 = 17995,08 - 124,42^2 = 2515,58.$$

Средние квадратические отклонения:

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

$$\sigma_m = \sqrt{152,75} = 12,36$$

$$\sigma_c = \sqrt{\sigma_c^2}$$

$$\sigma_c = \sqrt{2515,58} = 50,16.$$

Средняя из внутригрупповых дисперсий:

$$\overline{\sigma}^2 = \frac{\sum \sigma_j^2 n_j}{\sum n_j}$$

$$\overline{\sigma}^2 = \frac{152,75 \cdot 8 + 2515,58 \cdot 12}{20} = 1570,45.$$

Межгрупповая дисперсия:

$$\delta^2 = \frac{\sum (\overline{S}_j - \overline{S}_o)^2 n_j}{\sum n_j}$$

$$\delta^2 = \frac{(67,5 - 101,65)^2 \cdot 8 + (124,42 - 101,65)^2 \cdot 12}{20} = 777,48.$$

Общая дисперсия как сумма дисперсий:

$$\sigma_o^2 = \overline{\sigma}^2 + \delta^2$$

$$\sigma_o^2 = 1570,45 + 777,48 = 2347,93.$$

Оценка влияния каждого из факторов на общую дисперсию с помощью эмпирического коэффициента детерминации η^2 :

$$\eta^2 = \frac{\delta^2}{\sigma_o^2};$$

$$\eta^2 = \frac{777,48}{2347,93} = 0,33;$$

эмпирического корреляционного отношения η :

$$\eta = \sqrt{\eta^2} = \sqrt{\frac{\delta^2}{\sigma_o^2}}$$

$$\eta = \sqrt{0,33} = 0,58.$$

По полученным значениям эмпирического коэффициента детерминации η^2 и эмпирического корреляционного отношения η **сделать выводы.**

Выводы. На 33% вариация количества рабочих на предприятиях обусловлена различиями между группами предприятий, на остальные 67% вариация обусловлена другими факторами.
 Эмпирическое корреляционное отношение показывает, что связь умеренная.

Задача 4 Корреляционный анализ

По данным о численности S_i персонала $n = 20$ малых и средних предприятий и уровне фондовооруженности труда R_i (см. табл. 4.1):

- установить факторный и резульативный признаки;
- установить одним из известных Вам способов наличие и характер связи между признаками;
- оценить тесноту связи между ними с помощью непараметрических показателей: коэффициента корреляции рангов **Спирмена**; коэффициента корреляции рангов **Кэндела**; коэффициента корреляции знаков **Фехнера**; коэффициентов сопряженности **Пирсона** и **Чупрова**.

Таблица 4.1. - Численность персонала и уровень фондовооруженности

№_n	1	2	3	4	5	6	7	8	9	1	1	1	1	1	1	1	1	1	2	
					6					0	1	2	3	4	5	6	7	8	9	0
S_i чел.	9	S_{max}	8	6	6	7	4	9	7	4	7	6	7	7	5	9	7	6	9	5
	5		5	1	0	0	0	0	5	5	5	0	0	4	5	8	6	5	4	0
R_i	2	33	2	1	1	1	1	2	2	1	2	1	2	2	1	2	2	2	2	1

тыс.ру б.	8		5	7	5	9	7	6	3	6	4	8	0	3	8	6	6	1	7	7
--------------	---	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Варианты: К показателям фондовооруженности предприятий № 5,10,15,20 прибавить 0,5 номера по списку C_n : $R_i + 0,5 \times C_n$

Наши данные:

№ _n	1	2	3	4	5	6	7	8	9	10	1	1	1	1	15	1	1	1	1	20
S_i чел.	9	21	8	6	18	7	4	9	7	16	7	6	7	7	17	9	7	6	9	17
R_i тыс.ру б.	2	33	2	1	15,	1	1	2	2	16,	2	1	2	2	18,	2	2	2	2	17,
	8		5	7	5	9	7	6	3	5	4	8	0	3	5	6	6	1	7	5

Для вычисления коэффициентов **Спирмена**, **Кэндела** и **Фехнера** провести ранжирование факторных показателей x_i , то есть расположить их в порядке возрастания или убывания. Все вычисления выполнить в таблице 4.2.

Для вычисления коэффициентов **Пирсона** и **Чупрова** провести группировку факторных x_i и результативных y_i показателей в корреляционной таблице 4.3.

Таблица 4.2 – Вычисление показателей тесноты связи

Показатели		Корреляция рангов Спирмена				Корреляция рангов Кэндела		Корреляция знаков Фехнера			
x_i	y_i	R_x	R_y	d	d^2	P_y	Q_y	$x_i \Rightarrow \bar{x}$	$y_i \Rightarrow \bar{y}$	C	H
40	17	1	3,5	-2,5	6,25	16	2	-	-	1	0
60	18	2	6	-4	16	14	4	-	-	1	0
61	17	3	3,5	-0,5	0,25	15	2	-	-	1	0
65	21	4	10	-6	36	10	6	-	-	1	0
70	19	5,5	8	-2,5	6,25	11	4	-	-	1	0
70	20	5,5	9	-3,5	12,25	10	4	-	-	1	0
74	23	7	11,5	-4,5	20,25	9	4	-	+	0	1
75	23	8,5	11,5	-3	9	8	4	-	+	0	1
75	24	8,5	13	-4,5	20,25	7	4	-	+	0	1
76	26	10	16	-6	36	5	5	-	+	0	1
85	25	11	14	-3	9	5	4	-	+	0	1
90	26	12	16	-4	16	4	4	-	+	0	1
94	27	13	18	-5	25	2	5	-	+	0	1
95	28	14	19	-5	25	1	5	-	+	0	1
98	26	15	16	-1	1	1	4	-	+	0	1

165	16,5	16	2	14	196	3	0	+	-	0	1
170	17,5	17	5	12	144	2	1	+	-	0	1
175	18,5	18	7	11	121	1	1	+	-	0	1
180	15,5	19	1	18	324	1	0	+	-	0	1
215	33	20	20	0	0	0	0	+	+	1	0
2033	441				1023,5	125	63			7	13

Коэффициент корреляции рангов Спирмена:

Расчет проводится в следующей последовательности:

Факторные показатели x_i располагаем в порядке их возрастания или убывания, а результивные y_i располагаются в паре с соответствующими значениями x_i .

Результивные показатели y_i ранжируются аналогично в порядке возрастания или убывания их значений.

По разностям $d = R_x - R_y$ между рангами факторных и рангами результивных показателей вычисляем сумму их квадратов $\sum d^2$.

$$\rho_{xy} = 1 - \frac{6\sum d^2}{n^3 - n}$$

Для наших данных $\rho_{xy} = 1 - \frac{6 \cdot 1023,5}{20^3 - 20} = 0,23$

Коэффициент корреляции рангов Кэндела:

$$\tau_{xy} = \frac{2S}{n^2 - n}$$

Для его вычисления также необходимо расположить факторные показатели x_i в порядке возрастания или убывания, а результивные y_i – в паре с соответствующими значениями x_i .

Далее, последовательно рассматриваются значения или ранги только результивного показателя y_i .

По общей сумме числа P_y результивных показателей y_i больших и меньших Q_y по величине и рангу чем последующие вычисляем разность этих сумм:

$$S = \sum P_y - \sum Q_y$$

Для наших данных $S = 125 - 63 = 62$;

$$\tau_{xy} = \frac{2 \cdot 62}{20^2 - 20} = 0,33.$$

Коэффициент корреляции знаков Фехнера:

$$l_{xy} = \frac{n_c - n_n}{n_c + n_n}.$$

Для оценки тесноты связи с помощью этого показателя необходимо вычислить средние значения факторного \bar{X} и результативного \bar{y} показателей:

$$\bar{x} = \frac{\sum x_i}{n}; \quad \bar{y} = \frac{\sum y_i}{n}$$

и последовательно провести сравнение каждого факторного и каждого результативного показателей с их средними значениями. Если показатели больше средних, в таблице ставим знак "+", если они меньше или равны, ставим знак "-".

Далее, отмечая совпадения или несовпадения знаков для факторных и результативных показателей, подсчитываем их число, соответственно $n_c = 7$ и $n_n = 13$.

Для наших данных $\bar{x} = \frac{\sum x_i}{n} = 101,65; \quad \bar{y} = \frac{\sum y_i}{n} = 22,05; \quad n_c = 7; \quad n_n = 13.$

$$l_{xy} = \frac{7 - 13}{7 + 13} = -0,3.$$

Выводы:

По полученным значениям коэффициентов корреляции сделать выводы, о характере и тесноте связи между признаками.

Объясните различия в показателях тесноты связи.

Выводы. Коэффициент ранговой корреляции Спирмена показывает что между признаками слабая прямая зависимость.

Коэффициент ранговой корреляции Кэнделла показывает, что между признаками средняя прямая корреляционная зависимость.

Коэффициент корреляции знаков Фехнера показывает среднюю обратную зависимость.

Коэффициенты корреляции Пирсона и Чупрова.

Провести группировку факторных и результативных показателей по интервальным значениям, рассчитав число групп, шаг интервала для факторных и результативных показателей:

$$k = 1 + \log_2 n; \quad h_x = \frac{X_{\max} - X_{\min}}{k}; \quad h_y = h_y = \frac{Y_{\max} - Y_{\min}}{k}.$$

Для x мы уже рассчитывали число интервалов и их длину. Для y : $k = 5$,

$$h_y = \frac{33 - 15,5}{5} = 3,5.$$

Таблица 4.3 – Корреляционная таблица

		Интервальные значения y_i (число вариантов N_y)					Σn_x
		15,5-19	19-22,5	22,5-26	26-29,5	29,5-33	
Интервальные значения x_i (число вариантов N_x)	40-75	3	3	1			7
	75-110			3	5		8
	110-145						
	145-180	3					3
	180-215	1				1	2
Σn_y		7	3	4	5	1	20

Расчет показателя или критерия взаимной сопряженности Пирсона ϕ^2 провести по формуле:

$$\phi^2 = \frac{\sum \frac{n_{xy}^2}{n_x n_y} - 1}{n_x n_y}.$$

Расчет проводится путем последовательного просмотра таблицы по строкам или по столбцам:

Коэффициент сопряженности Пирсона:

$$K_{\Pi} = \sqrt{\frac{\varphi^2}{1 + \varphi^2}};$$

Коэффициент сопряженности Чупрова:

$$K_{\text{ч}} = \sqrt{\frac{\varphi^2}{\sqrt{(N_x - 1)(N_y - 1)}}}.$$

Для наших данных $\varphi^2 = 2,55 - 1 = 1,55$;

$$K_{\Pi} = \sqrt{\frac{1,55}{1 + 1,55}} = 0,78;$$

$$K_{\text{ч}} = \sqrt{\frac{1,55}{\sqrt{(20-1)(20-1)}}} = 0,29.$$

Выводы: Сделать выводы о тесноте связи. Сопоставить данные показатели с ранее вычисленными показателями тесноты связи.

Выводы. Коэффициент сопряженности Пирсона показывает что между признаками наблюдается тесная прямая зависимость.

Коэффициент сопряженности Чупрова показывает слабую прямую зависимость между признаками.

Задача 5

Показатели динамических рядов

Исходные данные. В процессе выборочного наблюдения получены данные о сумме активов F_i малых и средних предприятий за 10 лет (в млн руб. по состоянию на 01.01 каждого года).

Год	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
F_i	334	588	610	867	917	1162	1232	1510	1562	1898

Варианты. К показателям суммы активов 2005, 2007, 2009, 2011, 2013 добавить пять номеров S_n по списку в журнале группы и два номера группы № гр.:

$$5 \times C_{\text{н}} + 2 \times \text{Нгр} .$$

Наши данные:

Год	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
F_i	456	588	732	867	1039	1162	1354	1510	1684	1898

Определить показатели динамики:

- абсолютный базисный и цепной прирост в млн руб.

$$A_{\text{б}} = y_i - y_0 ; \quad A_{\text{ц}} = y_i - y_{i-1} ;$$

- темпы роста базисные и цепные в %%

$$T_{\text{рб}} = \frac{y_i}{y_0} 100 \% ; \quad T_{\text{рц}} = \frac{y_i}{y_{i-1}} 100 \% .$$

- темпы прироста базисные и цепные в %%

$$T_{\text{пб}} = \frac{A_{\text{б}}}{y_0} 100 \% ; \quad T_{\text{пц}} = \frac{A_{\text{ц}}}{y_{i-1}} 100 \%$$

или

$$T_{\text{пб}} = (T_{\text{рб}} - 100\%) ; \quad T_{\text{пц}} = (T_{\text{рц}} - 100\%) ;$$

- темпы наращивания в %%

$$T_{\text{нар}} = \frac{A_{\text{ц}}}{y_0} 100 \% \quad \text{или} \quad T_{\text{нар}} = (T_{\text{пб}(i)} - T_{\text{пб}(i-1)}) ;$$

- абсолютное значение одного процента прироста в млн руб.

$$A_{\%} = \frac{A_{\text{ц}}}{T_{\text{пц}}} \quad \text{или} \quad A_{\%} = \frac{y_{i-1}}{100\%} .$$

Вычисления провести в таблице 5.1

Таблица 5.1 – Расчет показателей динамики

На 01.01	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Активы F_i	334	588	610	867	917	1162	1232	1510	1562	1898
Ац млн руб.	-	132	144	135	172	123	192	156	174	214
Аб млн руб.	-	132	276	411	583	706	898	1054	1228	1442
Трб %	-	128,9 5	160,5 3	190,1 3	227,8 5	254,8 2	296,9 3	331,1 4	369,3 0	416,2 3
Трц %	-	128,9 5	124,4 9	118,4 4	119,8 4	111,8 4	116,5 2	111,5 2	111,5 2	112,7 1
Тпб %	-	28,95	60,53	90,13	127,8 5	154,8 2	196,9 3	231,1 4	269,3 0	316,2 3
Тпц %	-	28,95	24,49	18,44	19,84	11,84	16,52	11,52	11,52	12,71

%										
Тнар %	-	28,95	31,58	29,61	37,72	26,97	42,11	34,21	38,16	46,93
А% млн руб.	-	4,56	5,88	7,32	8,67	10,39	11,62	13,54	15,1	16,84

Средние показатели динамики:

- средний абсолютный прирост $\bar{A} = \bar{A} = \frac{\sum A_u}{n-1}$ или

$$\bar{A} = \frac{y_n - y_1}{n-1};$$

- средний темп роста $T_p = \sqrt[n-1]{\prod_1^{n-1} T_{pu}}$ или $T_p = \sqrt[n-1]{\frac{y_n}{y_1}}$

- средний темп прироста

$$\bar{T}_u = \bar{T}_p - 100\%;$$

- средняя годовая

сумма

активов

$$\bar{F} = \frac{0.5y_1 + y_2 + y_3 + \dots + 0.5y_n}{n-1}.$$

Для наших данных

$$\bar{A} = \frac{1442}{10-1} = 160,22 \text{ - средний абсолютный прирост;}$$

$$T_p = \sqrt[9]{\frac{1898}{334}} = 1,172 \text{ или } 117,2\% \text{ - средний темп роста;}$$

$$\bar{T}_u = 117,2 - 100\% = 17,2\% \text{ - средний темп прироста;}$$

$$\bar{F} = \frac{10113}{9} = 1123,67 \text{ - средняя годовая сумма активов.}$$

Графики динамического ряда

Построение провести с использованием Excel. Построить три графика – точечный, полигон и столбиковая диаграмма (см. рис. 5.1).

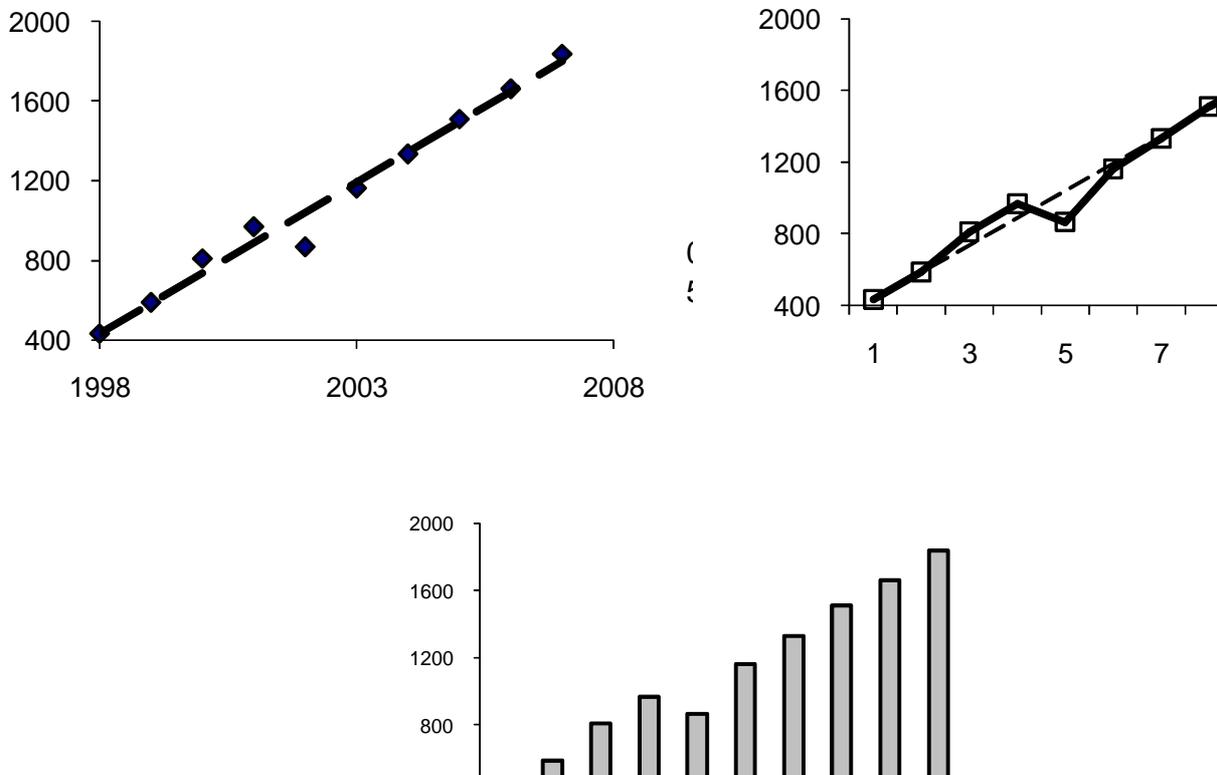


Рис. 5.1. Графики динамического ряда

Для наших данных





Уравнение ряда динамики:

Исходя из предположения, что динамика носит линейный характер, построить линейное уравнение динамики в виде : $Y_t = a_0 + a_1 t$

Параметры уравнения a_0 и a_1 вычислить в таблице 5.2, используя следующие формулы:

$$\bar{y} = \frac{\sum y_i}{n}; \quad \overline{yt} = \frac{\sum y_i t_i}{n}; \quad \sigma_t^2 = \overline{t^2} = \frac{\sum t^2}{n}; \quad a_1 = \frac{\overline{yt}}{\sigma_t^2}; \quad a_0 = \bar{y}$$

Таблица 5.2 – Расчет параметров уравнения динамики

Год	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Σ
t	-9	-7	-5	-3	-1	1	3	5	7	9	
y_i	456	588	732	867	1039	1162	1354	1510	1684	1898	11290
$y_i t_i$	-4104	-4116	-3660	-2601	-1039	1162	4062	7550	11788	17082	26124
t^2	81	49	25	9	1	1	9	25	49	81	330

Лабораторная работа по статистике. Выполнена на www.MatBuro.ru
 ©МатБюро – Решение заданий математики, экономики, программирования
 Сделаем ваши задания на отлично. https://www.matburo.ru/sub_appear.php?p=lst

Y_t	416,5	574,9	733,2	891,5	1049,8	1208,2	1366,5	1524,8	1683,1	1841,5	11290,0
-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	---------

$$\bar{y} = \frac{11290}{10} = 1129 ; \quad \overline{yt} = \frac{26124}{10} = 2612,4 ;$$

$$\sigma_t^2 = \bar{t}^2 = \frac{330}{10n} = 33 ; \quad a_1 = \frac{\overline{yt}}{\sigma_t^2} = \frac{2612,4}{33} = 79,16 ; \quad a_0 = 1129$$

То есть линейное уравнение динамики имеет вид: $Y_t = 1129 + 79,16t$
 Заполним последнюю строку расчетной таблицы.

Прогноз уровня фондовооруженности на 01.01.2016 года провести при $t = 13$.

$Y_{13} = 1129 + 79,16 \cdot 13 = 2158,1$ - прогноз уровня фондовооруженности на 01.01.2016 года.

Задача 6

Индексы количественных и качественных показателей.

Индексы средних взвешенных показателей

Исходные данные. По отчётности обследованных предприятий за два последних квартала получены данные о производстве товаров и услуг (см. табл. 6.1).

Таблица 6.1 – Производство товаров и услуг

Вид товаров и услуг	Число единиц	Цена реализации (тыс.
---------------------	--------------	-----------------------

			руб./ед.)	
	1 кв. q_0	2 кв. q_1	1 кв. p_0	2 кв. p_1
А	2500	2600	3,0	2,5
Б	3000	2900	7,0	7,5
В	3600	3700	5,0	4,5

Варианты заданий. В показатели q_1 по всем видам товаров и услуг добавить двадцать номеров по списку **Сн** и пять номеров группы **Н гр.** В показатели p_1 по всем видам товаров и услуг добавить 0,15 от **Сн** и 0,2 номера группы:

$$q_1 + 20 \times \text{Сн} + 5 \times \text{Н гр}; \quad p_1 + 0,15 \times \text{Сн} + 0,2 \times \text{Н гр}.$$

Вариант задания:

Вид товаров и услуг	Число единиц		Цена реализации (тыс. руб./ед.)	
	1 кв. q_0	2 кв. q_1	1 кв. p_0	2 кв. p_1
А	2500	3085	3	6,3
Б	3000	3385	7	11,3
В	3600	4185	5	8,3

Определить:

- индивидуальные индексы физических объёмов производства и цен реализации продукции;
- сводные (общие) индексы физических объёмов и цен реализации продукции как агрегатные, средние арифметические из индивидуальных, средние гармонические из индивидуальных;
- абсолютное и относительное изменение стоимости продукции;
- влияние количественного и качественного факторов на абсолютное изменение стоимости продукции;
- индексы средней цены переменного, постоянного состава и структурных сдвигов;
- абсолютное влияние изменения средней цены на стоимость товаров и услуг.

Для расчета агрегатных, средних арифметических и средних гармонических индексов провести расчёты индивидуальных индексов и стоимости продукции в таблицах 6.2 и 6.3

Таблица 6.2 – Индивидуальные индексы

Вид товаров и услуг	i_p	i_q
А	2,100	1,234
Б	1,614	1,128
В	1,660	1,163

Таблица 6.3 – Стоимость товаров и услуг

Вид товаров и услуг	p_0q_0	p_1q_1	p_1q_0	p_0q_1
А	7500	19435,5	15750	9255
Б	21000	38250,5	33900	23695
В	18000	34735,5	29880	20925
Σ	46500	92421,5	79530	53875

Сводные индексы физических объёмов:

- агрегатный индекс физических объёмов Ласпейреса

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0};$$

$$I_q = \frac{53875}{46500} = \mathbf{1,159 \text{ или } 115,9\%}.$$

- средний арифметический (из индивидуальных) индекс физических объёмов Ласпейреса

$$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0};$$

$$I_q = \frac{1,234 \cdot 7500 + 1,128 \cdot 21000 + 1,163 \cdot 18000}{46500} = \frac{53875}{46500} = \mathbf{1,159 \text{ или } 115,9\%}$$

- средний гармонический (из индивидуальных) индекс физических объёмов Ласпейреса

$$I_q = \frac{\sum q_1 p_0}{\sum \frac{q_1 p_0}{i_q}}.$$

$$I_q = \frac{\sum q_1 p_0}{\sum \frac{q_1 p_0}{i_q}} = \frac{53875}{\frac{9255}{1,234} + \frac{23695}{1,128} + \frac{20925}{1,163}} = \frac{53875}{46500} = \mathbf{1,159}.$$

Сводные индексы цен:

- агрегатный индекс цен Пааше

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1};$$

$$I_p = \frac{92421,5}{53875} = \mathbf{1,715 \text{ или } 171,5 \%}$$

- средний арифметический (из индивидуальных) индекс цен Пааше

$$I_p = \frac{\sum i_p p_0 q_1}{\sum p_0 q_1};$$

$$I_p = \frac{2,1 \cdot 9255 + 1,614 \cdot 23695 + 1,66 \cdot 20925}{53875} = \frac{92421,5}{53875} = 1,715.$$

- средний гармонический (из индивидуальных) индекс цен Пааше

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}$$

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}} = \frac{92421,5}{\frac{19435,5}{2,1} + \frac{38250,5}{1,614} + \frac{34735,5}{1,66}} = \frac{92421,5}{53875} = 1,715.$$

Сделать выводы.

Выводы. За рассматриваемый период цены на все виды товаров и услуг в среднем повысились на 71,5%, физический объем продаж вырос на 15,9%.

Абсолютное и относительное изменение стоимости товаров и услуг

- абсолютное изменение стоимости

$$\Delta p q = \sum p_1 q_1 - \sum p_0 q_0;$$

$$\Delta p q = 92421,5 - 46500 = \mathbf{45921,5}$$

- относительное изменение стоимости товаров и услуг

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} \quad \text{или} \quad I_{pq} = I_p \times I_q.$$

$$I_{pq} = \frac{92421,5}{46500} = \mathbf{1,988 \text{ или } } I_{pq} = 1,715 \cdot 1,159 = \mathbf{1,988}.$$

Влияние факторов на изменение стоимости продукции

- из-за изменения физических объёмов производства стоимость товаров и услуг увеличилась на

$$\Delta p q_q = \sum q_1 p_0 - \sum q_0 p_0;$$

$$\Delta p q_q = 53875 - 46500 = \mathbf{7375};$$

- из-за изменения цен на продукцию стоимость товаров и услуг увеличилась на:

$$\Delta p q_p = \sum p_1 q_1 - \sum p_0 q_1;$$

$$\Delta p q_p = 92421,5 - 53877,5 = 38544,0$$

- доля количественного фактора в общем изменении стоимости товаров и услуг:

$$\Delta p q_q \% = \frac{\Delta p q_q}{\Delta p q} 100\% ;$$

$$\Delta p q_q \% = \frac{7375}{45921,5} 100\% = \mathbf{16,06\%}.$$

- доля ценового фактора в изменении стоимости товаров и услуг:

$$\Delta p q_p \% = \frac{\Delta p q_p}{\Delta p q} 100\%$$

$$\Delta p q_p \% = \frac{38544,0}{45921,5} 100\% = \mathbf{83,94\%}.$$

Контроль

$$\Delta p q_p + \Delta p q_q = \Delta p q ; 38544,0 + 7375 = 45919,0$$

$$\Delta p q_p \% + \Delta p q_q \% = \mathbf{100,0\%} ; 83,94\% + 16,06\% = 100,0\% .$$

Индексы средних взвешенных показателей

Для расчета необходимы следующие средние цены:

- средняя цена на продукцию во 2-ом квартале

$$\bar{p}_1 = \frac{\sum p_1 q_1}{\sum q_1} ; \bar{p}_1 = \frac{92421,5}{10655} = \mathbf{8,674};$$

- средняя цена на продукцию в 1-ом квартале

$$\bar{p}_0 = \frac{\sum p_0 q_0}{\sum q_0} ; \bar{p}_0 = \frac{46500}{9100} = \mathbf{5,11};$$

- средняя цена на продукцию 2-го квартала в ценах 1-го квартала

$$\bar{p}_{0(1)} = \frac{\sum p_0 q_1}{\sum q_1} ; \bar{p}_{0(1)} = \frac{53875}{10655} = \mathbf{5,92}.$$

Индекс переменного состава (общее изменение средней цены реализации товаров и услуг)

$$I_{nc} = \frac{\bar{p}_1}{\bar{p}_0}; I_{nc} = \frac{8,674}{5,11} = \mathbf{1,697}.$$

Индекс постоянного или фиксированного состава (изменение средней цены из-за изменения цен по каждому виду товаров и услуг)

$$I_{fc} = \frac{\bar{p}_1}{\bar{p}_{0(1)}}; I_{fc} = \frac{8,674}{5,92} = \mathbf{1,465}.$$

Индекс структурных сдвигов (изменение средней цены из-за изменений в объемах и структуре производства товаров и услуг)

$$I_{стр} = \frac{\bar{p}_{0(1)}}{\bar{p}_0}; I_{стр} = \frac{5,92}{5,11} = \mathbf{1,159}$$

Контроль

$$I_{nc} = I_{fc} \times I_{стр}; I_{nc} = 1,465 \times 1,159 = \mathbf{1,697}.$$

Сделать выводы.

Выводы. В среднем цена за рассматриваемый период увеличилась в 1,697 раз, это изменение произошло в 1,465 раз за счет изменения цен по каждому виду товаров и услуг и в 1,159 раз за счет изменений в объемах и структуре производства товаров и услуг.

Оценка абсолютного изменения средней цены товаров и услуг

- общее абсолютное изменение средней цены

$$\Delta \bar{p} = \bar{p}_1 - \bar{p}_0; \Delta \bar{p} = 8,674 - 5,11 = \mathbf{3,564};$$

- абсолютное изменение средней цены из-за изменения уровня цен на товары и услуги

$$\Delta \bar{p}_p = \bar{p}_1 - \bar{p}_{0(1)}; \Delta \bar{p}_p = 8,674 - 5,92 = \mathbf{2,754};$$

- абсолютное изменение средней цены из-за изменения состава и структуры товаров и услуг

$$\Delta \bar{p}_{стр} = \bar{p}_{0(1)} - \bar{p}_0; \Delta \bar{p}_{стр} = 5,92 - 5,11 = \mathbf{0,81}.$$

Сделать выводы.

Выводы. Абсолютное изменение средней цены составило 3,564 тыс. руб. Из-за изменения уровня цен на товары и услуги средняя цена увеличилась на 2,754 тыс. руб. Из-за изменения состава и структуры товаров и услуг средняя цена увеличилась на 0,81 тыс. руб.

Влияние изменения средней цены на выручку от реализации товаров и услуг

- абсолютное изменение стоимости товаров и услуг

$$\Delta pq = \sum p_1 q_1 - \sum p_0 q_0 ; \Delta pq = 92421,5 - 46500 = \mathbf{45921,5};$$

- из-за изменения объёмов производства стоимость товаров и услуг увеличилась на

$$\Delta pq_q = \left(\sum q_1 - \sum q_0 \right) \bar{p}_0 ; \Delta pq_q = (10655 - 9100) \cdot 5,11 = \mathbf{7945,88}$$

- из-за изменения цен на каждый вид продукции стоимость товаров и услуг изменилась на:

$$\Delta pq_p = \left(\bar{p}_1 - \bar{p}_{0(1)} \right) \sum q_1 ; \Delta pq_p = (8,674 - 5,92) \cdot 10655 = \mathbf{29340,39}$$

- из-за изменения средней цены из-за изменения структуры или ассортимента продукции стоимость товаров и услуг изменилась на

$$\Delta pq_{стр} = \left(\bar{p}_{0(1)} - \bar{p}_0 \right) \sum q_1 ; \Delta pq_{стр} = (5,92 - 5,11) \cdot 10655 = \mathbf{8635,234}.$$

Контроль

$$\Delta pq = \Delta pq_q + \Delta pq_p + \Delta pq_{стр} ,$$

$$\Delta pq = 7945,88 + 29340,39 + 8635,234 = \mathbf{45921,5}.$$

Сделать выводы.

Выводы. Абсолютное изменение стоимости товаров и услуг составило 45921,5 тыс. руб. Из-за изменения объёмов производства стоимость увеличилась на 7945,88 тыс. руб. Из-за изменения цен на каждый вид продукции стоимость увеличилась на 29340,39 тыс. руб. И из-за изменения средней цены из-за изменения структуры или ассортимента продукции стоимость товаров и услуг увеличилась на 8635,234 тыс. руб.