МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ЛУГАНСКОЙ НАРОДНОЙ РЕСПУБЛИКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
ЛУГАНСКОЙ НАРОДНОЙ РЕСПУБЛИКИ
«ЛУГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ ВЛАДИМИРА ДАЛЯ»

Стахановский инженерно-педагогический институт менеджмента
Кафедра социально-экономических и педагогических дисциплин

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

для проведения практических занятий по дисциплине

«СТАТИСТИКА»

для студентов направления подготовки
44.03.04 Профессиональное обучение (по отраслям)

Луганск 2021

УДК 311

Рекомендовано к изданию Учебно-методическим советом
ГОУ ВО ЛНР «ЛГУ им. В.ДАЛЯ»
(протокол № от . .2021 г.)

Методические указания для проведения практических занятий по дисциплине **«Статистика»** для студентов направления подготовки **44.03.04 Профессиональное обучение (по отраслям)**. / Сост.: Д.С. Тимошенко. – **Стаханов**: ГОУ ВО ЛНР «ЛГУ им. В.Даля», 2021. – 71 с.

Методические указания содержат 10 тем, описание которых сопровождается теоретическими сведениями. К каждой теме приведены краткие методические указания, в которых приводится методика расчета основных показателей.

Предназначен для студентов профилей «Экономика и управление», «Управление персоналом».

Составитель: ст. преподаватель Тимошенко Д.С.

Ответственный за выпуск: доц. Карчевская Н.В.

Рецензент: доц. Карчевский В.П.

Часть 1. Теоретическая статистика

Тема 1. Абсолютные и относительные статистические величины

Методические указания по теме

Задача 1. Расход топлива на производственные нужды предприятия характеризуется в отчетном периоде следующими данными:

Вид топлива	Теплотворная способность, МДж/кГ	Расход, т	
		по плану	фактически
Дизельное топливо	41,9	1000	1050
Мазут	40,1	750	730
Уголь	26,4	500	555

Определить общее количество потребленного условного топлива (1 т.у.т. = 29,3 МДж/кГ) по плану и фактически, а также процент выполнения плана по общему расходу топлива.

Решение. Учитывая стандартную теплотворную способность 29,3 МДж/кГ, определяем количество потребленного условного топлива каждого вида по плану (X'_{1i}) и фактически (X_{1i}):

- дизельное топливо: $X'_{1дт}$ = 41,9/29,3*1000 = 1430,034 т.у.т.
  дизельное топливо: $X_{1дт}$ = 41,9/29,3*1050 = 1501,536 т.у.т.;
- мазут: $X'_{1м}$ = 40,1/29,3*750 = 1026,451 т.у.т.
  мазут: $X_{1м}$ = 40,1/29,3*730 = 999,078 т.у.т.;
- уголь: $X'_{1у}$ = 26,4/29,3*500 = 450,512 т.у.т.
  уголь: $X_{1у}$ = 26,4/29,3*555 = 500,068 т.у.т.

Суммируя количество потребленного условного топлива каждого вида, получим общее количество потребленного условного топлива:

- по плану $X'_1 = \sum X'_{1i}$ = 2906,997 т.у.т.;
- фактически $X_1 = \sum X_{1i}$ = 3000,682 т.у.т.

Для определения процента выполнения плана необходимо рассчитать *индекс выполнения плана*, то есть отношение значений по факту и плану отчетного периода:

$$i_{ВП} = \frac{X_1}{X'_1}, \qquad (1)$$

Применяя формулу (1), имеем: $i_{ВП}$ = 3000,682/2906,997 = 1,032, то есть план по общему расходу топлива перевыполнен на 3,2%.

Задача 2. Рассчитать индекс и темп изменения, если в марте произведено продукции 130 тонн, а в феврале 100 тонн.

Решение. *Индекс изменения (динамики)* характеризует изменение какого-либо явления во времени. Он представляет собой отношение значений одной и той же абсолютной величины в разные периоды времени. Данный индекс определяется по формуле (2):

$$i_Д = \frac{X_1}{X_0}, \qquad (2)$$

где подиндексы означают: 1 — отчетный или анализируемый период, 0 — прошлый или базисный период.

Критериальным значением индекса динамики (темпа роста) служит единица, то есть если $i_Д$>1, то имеет место рост явления во времени; если $i_Д$=1 – стабильность; если $i_Д$<1 – наблюдается спад явления. Применяя формулу (2), имеем: $i_Д$= 130/100 = 1,3 (или 130%) > 1 – рост объема произведенной продукции.

Темп изменения (прироста) определяется по формуле (3):

$$T = i_Д - 1. \qquad (3)$$

Применяя формулу (3), имеем: *T* = 1,3 – 1 = 0,3 (или 30%), то есть объем произведенной продукции вырос в марте по сравнению с февралем на 30%.

Задача 3. Рассчитать индексы планового задания, выполнения плана и динамики, если выпуск продукции в отчетном году составил 100 млн. рублей, на следующий год планировалось 140 млн. рублей, а фактически получено 112 млн. рублей.

Решение. *Индекс планового задания* – это отношение значений одной и той же абсолютной величины по плану анализируемого периода и по факту базисного. Он определяется по формуле (4):

$$i_{ПЗ} = \frac{X'_1}{X_0}, \qquad (4)$$

где X'_1 — план анализируемого периода; X_0 — факт базисного периода.

Применяя формулу (4) имеем: $i_{ПЗ}$= 140/100 = 1,4 (или 140%), то есть на следующий год планировалось выпустить продукции в размере 140% от объема предыдущего года.

Индекс выполнения плана определим, применяя формулу (1): $i_{ВП}$ = 112/140 = 0,8 (или 80%), то есть план по увеличению выпуска продукции выполнили лишь на 80% или недовыполнили на 20%.

Индекс динамики можно определить по формуле (2) или перемножая индексы планового задания и выполнения плана, то есть $i_Д = \frac{X_1}{X_0} = i_{ПЗ} i_{ВП}$ = 1,12.

Задача 4. Суммарные денежные доходы россиян в 2005 г. составили 13522,5 млрд. руб., из которых 8766,7 млрд. руб. составила оплата труда, 1748,4 млрд. руб. – социальные выплаты, 1541,7 млрд. руб. – доход от предпринимательской деятельности, 1201,5 млрд. руб. – доходы от собственности, остальное – прочие доходы. Рассчитать относительные величины структуры и координации, приняв за основу оплату труда. Построить секторную (круговую) диаграмму структуры доходов.

Решение. *Индекс структуры (доля*) – это отношение какой-либо части величины (совокупности) ко всему ее значению. Он определяется по формуле (5):

$$i_{CT} = d = \frac{f}{\sum f} \qquad (5)$$

Применяя формулу (5) и округляя значения до 3-х знаков после запятой, имеем:

- доля оплаты труда d_{OT} = 8766,7/13522,5 = 0,648 или 64,8%;
- доля социальных выплат d_{CB} =1748,4/13522,5 = 0,129 или 12,9%;
- доля доходов от предпринимательской деятельности $d_{ПД}$ =1541,7/13522,5 = 0,114 или 11,4%;
- доля доходов от собственности $d_{ДС}$ =1201,5/13522,5 = 0,089 или 8,9%.

Долю прочих доходов найдем, используя формулу (6), согласно которой сумма всех долей равна единице:

$$\sum d = 1. \qquad (6)$$

Таким образом, доля прочих доходов $d_{проч}$ = 1 – 0,648 – 0,129 – 0,114 – 0,089 = 0,020 или 2,0%.

Для иллюстрации структуры (составных частей) доходов построим секторную диаграмму (рис.1):

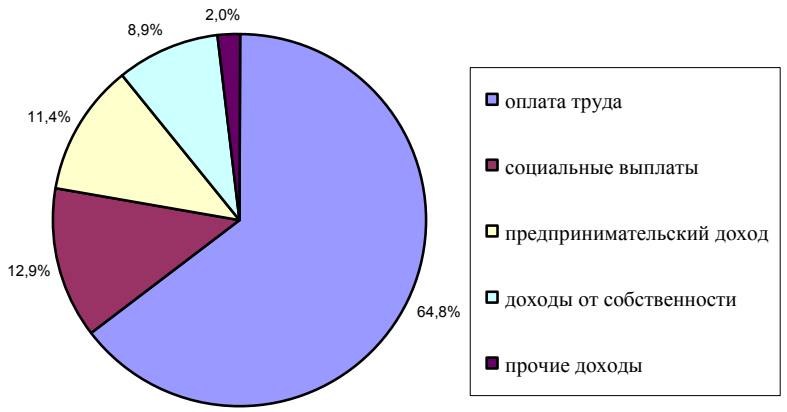


Рис.1. Структура денежных доходов населения РФ в 2005 году.

Таким образом, очевидно, что наибольшую долю в суммарных денежных доходах составляет оплата труда (64,8%), на 2-м месте – социальные выплаты

(12,9%), затем следуют предпринимательский доход (11,4%), доходы от собственности (8,9%), а прочие доходы составляют лишь 2%.

Индекс координации – это отношение какой-либо части величины к другой ее части, принятой за основу (базу сравнения). Он определяется по формуле (7):

$$i_K = \frac{f}{f_б}. \qquad (7)$$

Применяя формулу (7) и принимая за основу оплату труда, имеем:

- индекс координации социальных выплат $i_{K_{СВ}}$ = 1748,4/8766,7 ≈ 0,129/0,648 = 0,199;
- индекс координации предпринимательского дохода $i_{K_{ПД}}$=1541,7/8766,7 ≈ 0,114/0,648 = 0,176;
- индекс координации доходов от собственности $i_{K_{ДС}}$ = 1201,5/8766,7 ≈ 0,089/0,648 = 0,137;
- индекс координации прочих доходов $i_{K_{проч}}$ ≈ 0,02/0,648 = 0,031.

Таким образом, социальные выплаты составляют 19,9% от оплаты труда, предпринимательский доход – 17,6%, доходы от собственности – 13,7%, а прочие доходы – 3,1%.

Задача 5. Запасы воды в озере Байкал составляют 23000 км^3, а в Ладожском озере 911 км^3. Рассчитать относительные величины сравнения запасов воды этих озер.

Решение. *Индекс сравнения* – это отношение значений одной и той же величины в одном периоде или моменте времени, но для разных объектов или территорий. Он определяется по формуле (8):

$$i_C = \frac{X_A}{X_Б}, \qquad (8)$$

где *А, Б* — признаки сравниваемых объектов или территорий.

Применяя формулу (8) и принимая за объекты *А* и *Б*, соответственно, озера Байкал и Ладожское, найдем индекс сравнения: i_C = 23000/911 = 25,25, то есть запасов воды в озере Байкал в 25,25 раза больше, чем в Ладожском озере.

Меняя базу сравнения, найдем индекс сравнения Ладожского озера с Байкалом по той же формуле: i_C = 911/23000 = 0,0396 или 3,96%, то есть запасы воды в Ладожском озере составляют 3,96% запасов воды в озере Байкал.

Задача 6. Рассчитать относительную величину интенсивности валового внутреннего продукта (ВВП) в сумме 1416,1 млрд. $ на душу населения в России в 2004 году при численности населения в 144,2 млн. человек.

Решение. *Показатель интенсивности* – это отношение значений двух разнородных абсолютных величин для одного периода времени и одной территории или объекта. Он определяется по формуле (9):

$$i_{ИН} = \frac{X}{Y}. \qquad (9)$$

Применяя формулу (9) имеем: $i_{ИН}$ = 1416,1/0,1442 = 9820,39 $/чел в год.

Контрольные задания по теме

Вариант 1. Определить общее производство моющих средств в условных тоннах (условная жирность 40%) по плану и фактически, а также процент выполнения плана по следующим данным:

Вид продукта	Жирность, %	Физическая масса, т	
		по плану	фактически
Мыло хозяйственное	60	500	600
Мыло туалетное	80	1000	1500
Стиральный порошок	10	50000	40000

Вариант 2. По данным о численности жителей двух крупнейших городов России (тыс. чел) определить индексы сравнения и динамики.

Город / Год	2004	2005
Москва	10391	10407
Санкт-Петербург	4624	4600

Вариант 3.

1. По плану на 2005 год намечалось увеличение товарооборота на 3%. В 2005 году плановое задание перевыполнили на 600 млн. руб. или на 2,5%. Определить фактический прирост товарооборота (в млн. руб.) в 2005 году по сравнению с 2004 годом.

1. По данным о товарообороте из предыдущей задачи, состоящего из реализации собственной продукции и продажи покупных товаров, определить относительные величины координации и структуры собственной и покупной продукции в 2004 и 2005 годах, если известно, что доля собственной продукции в 2004 году составила 65%, а в 2005 году она увеличилась на 10%.

Вариант 4. Жилищный фонд и численность населения России следующие (на начало года):

Год	2002	2003	2004	2005
Весь жилищный фонд, млн. м^2	2853	2885	2917	2949

Численность населения, млн. чел.	145,6	145,0	144,2	143,5

Охарактеризовать изменение обеспеченности населения жилой площадью с помощью относительных величин динамики и координации.

Вариант 5.

1. В России в 2004 численность женщин составила 77144,3 тыс. чел, а мужчин – 67023,9 тыс. чел. Рассчитать относительные величины структуры и координации.

2. По плану объем продукции в отчетном году должен возрасти по сравнению с прошлым годом на 2,5%. План выпуска продукции перевыполнен на 3,0%. Определить фактический выпуск продукции в отчетном году, если известно, что объем продукции в прошлом году составил 25,3 млн. руб.

Вариант 6. Определить общий объем фактически выпущенной продукции по следующим данным по трем филиалам предприятия, выпускающих однородную продукцию:

Номер филиала	Планируемый объем выпуска продукции, млн. руб.	Выполнение намеченного плана, %
1	500	104
2	750	92
3	250	116

Вариант 7. По промышленному предприятию за отчетный год имеются следующие данные о выпуске продукции:

Наименование продукции	План на I квартал, тыс. т	Фактический выпуск, тыс. т			Отпускная цена за 1 т, у.е.
		январь	февраль	март	
Сталь арматурная	335	110	115	108	1700
Прокат листовой	255	75	90	100	2080

Определить процент выполнения квартального плана: 1) по выпуску каждого вида продукции; 2) в целом по выпуску всей продукции.

Вариант 8. Определить процент выполнения плана по продажам условных школьных тетрадей (1 у.ш.т. – 12 листов) по каждому виду тетрадей и в целом по магазину по следующим данным:

Вид тетради	Цена, руб./шт.	Объем продаж, тыс. шт.	
		по плану	фактически
Тетрадь общая 90 листов	20	50	40
Тетрадь общая 48 листов	13	200	350
Тетрадь общая 16 листов	9	700	500

Вариант 9. В России на начало 2005 года численность населения составила 144,2 млн. чел., в течение года: родилось 1,46 млн. чел., умерло – 2,3 млн. чел., мигрировало из других государств 2,09 млн. чел., мигрировало за границу – 1,98 млн. чел. Охарактеризовать изменение численности населения в 2005 году с помощью относительных величин.

Вариант 10. Определить общий объем фактически выпущенной условной консервной продукции (1 у.к.б. = 0,33 л) по следующим данным:

Вид продукции	Планируемый объем выпуска продукции, тыс. шт.	Выполнение плана, %
Томатная паста 1 л	500	85
Томатная паста 0,5 л	750	104
Томатная паста 0,2 л	250	130

Тема 2. Средние величины и показатели вариации

Методические указания по теме

Задача 1. Имеются следующие данные о возрастном составе студентов группы заочного отделения ВУЗа (лет): 19; 19; 19; 20; 20; 20; 20; 20; 20; 20; 20; 20; 21; 21; 21; 22; 23; 23; 24; 25; 25; 25; 26; 27; 29.

Для анализа распределения студентов по возрасту требуется: 1) построить интервальный ряд распределения и его график; 2) рассчитать модальный, медианный и средний возраст, установить его типичность с помощью коэффициентов вариации; 3) проверить распределение на нормальность с помощью коэффициентов асимметрии и эксцесса.

Решение. Для построения интервального ряда из дискретного используется формула Стерджесса, с помощью которой определяется оптимальное количество интервалов (n):

$$n = 1 + 3{,}322 \lg N, \qquad (10)$$

где N – число величин в дискретном ряде.

В нашей задаче $n = 1 + 3{,}322 lg 25 = 1 + 3{,}322*1{,}398 = 5{,}64$. Так как число интервалов не может быть дробным, то округлим его до ближайшего целого числа, т.е. до 6.

После определения оптимального количества интервалов определяем размах интервала по формуле:

$$h = H / n, \qquad (11)$$

где H – размах вариации, определяемый по формуле (12).

$$H = X_{мах} - X_{min}, \qquad (12)$$

где $X_{мах}$ и X_{min} — максимальное и минимальное значения в совокупности.

В нашей задаче $h = (29 – 19)/6 = 1{,}67$.

Интервальная группировка данных приведена в первом столбце таблицы 1, которая содержит также алгоритм и промежуточные расчеты.

Таблица 1. Вспомогательные расчеты для решения задачи

X_i , лет	f_i	$X_И$	$X_И f_i$	$X_И - \overline{X}$	$\lvert X_И - \overline{X} \rvert f_i$	$(X_И - \overline{X})^2$	$(X_И - \overline{X})^2 f_i$	$(X_И - \overline{X})^3 f_i$	$(X_И - \overline{X})^4 f_i$
до 20,67	12	19,833	237,996	-2,134	25,602	4,552	54,623	-116,539	248,638
20,67-22,33	4	21,5	86,000	-0,467	1,866	0,218	0,871	-0,406	0,189
22,33-24	3	23,167	69,501	1,200	3,601	1,441	4,323	5,190	6,231
24-25,67	3	24,833	74,499	2,866	8,599	8,217	24,650	70,659	202,543
25,67-27,33	2	26,5	53,000	4,533	9,067	20,552	41,105	186,348	844,806
более 27,33	1	28,167	28,167	6,200	6,200	38,446	38,446	238,383	1478,091
Итого	25	—	549,163	—	54,937	—	164,018	383,636	2780,498

На основе этой группировки строится график распределения возраста студентов (рис.2).

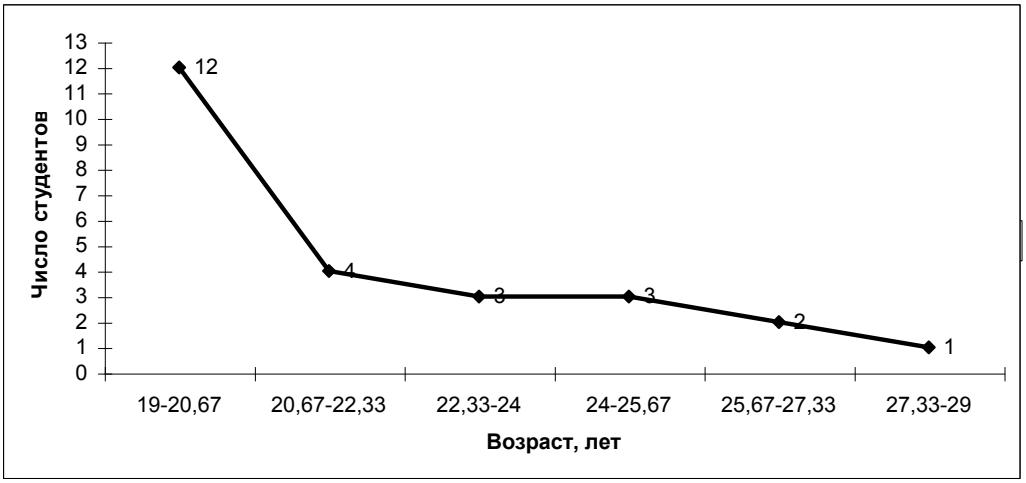


Рис.2. График распределения возраста студентов.

Мода – это наиболее часто повторяющееся значение признака. Для интервального ряда с равными интервалами величина моды определяется по формуле (13):

$$Mo = X_{Mo} + h\frac{f_{Mo} - f_{Mo-1}}{2f_{Mo} - f_{Mo-1} - f_{Mo+1}}, \qquad (13)$$

где X_{Mo} – нижнее значение модального интервала; f_{Mo} – число наблюдений или объем взвешивающего признака (вес признака) в модальном интервале; f_{Mo-1} – то же для интервала, предшествующего модальному; f_{Mo+1} – то же для интервала, следующего за модальным; h – величина интервала изменения признака в группах.

В нашей задаче чаще всего повторяется (12 раз) первый интервал возраста (до 20,67), значит, это и есть модальный интервал. Используя формулу (13), определяем точное значение модального возраста:

Mo = 19 + 1,667*(12-0)/(2*12-4-0) = 20 (лет).

Медиана – это такое значение признака, которое приходится на середину ранжированного ряда. Таким образом, в ранжированном ряду распределения одна половина ряда имеет значения признака больше медианы, другая – меньше медианы. Для интервального ряда с равными интервалами величина медианы определяется так:

$$Me = X_{Me} + h\frac{0{,}5\sum f - f'_{Me-1}}{f_{Me}}, \qquad (14)$$

где X_{Me} – нижняя граница медианного интервала; h – его величина (размах); f'_{Me-1} – сумма наблюдений (или объема взвешивающего признака), накопленная

до начала медианного интервала; f_{Me} – число наблюдений или объем взвешивающего признака в медианном интервале.

В нашей задаче второй интервал возраста (от 20,67 до 22,33) является медианным, так как на него приходится середина ряда распределения возраста. Используя формулу (14), определяем точное значение медианного возраста:

Me = 20,67 + 1,667*(12,5-12)/4 = 20,878 (года).

Средняя величина – это обобщающий показатель совокупности, характеризующий уровень изучаемого явления или процесса. Средние величины могут быть *простыми* и *взвешенными.* Простая средняя рассчитывается при наличии двух и более статистических величин, расположенных в произвольном (несгруппированном) порядке, по общей формуле (15). Взвешенная средняя величина рассчитывается по сгруппированным статистическим величинам с использованием общей формулы (16).

$$\overline{X} = \sqrt[m]{\frac{\sum X_i^m}{N}}; \qquad (15) \qquad\qquad \overline{X} = \sqrt[m]{\frac{\sum X_i^m f_i}{\sum f_i}}. \qquad (16)$$

При этом обозначено: X_i – значения отдельных статистических величин или середин группировочных интервалов; *m* - показатель степени, от значения которого зависят *виды* средних величин. Используя формулы (15) и (16) при разных показателях степени *m*, получаем частные формулы каждого вида (см. таблицу 2).

Таблица 2. Виды степенных средних и их применение

m	*Название средней*	*Формула расчета средней*		*Когда применяется*
		простая	*взвешенная*	
1	Арифметическая	$\overline{X}_{ар} = \frac{\sum X_i}{N}$ (17)	$\overline{X}_{ар} = \frac{\sum X_i f_i}{\sum f_i}$ (18)	Чаще всего, кроме тех случаев, когда должны применяться другие виды средних
–1	Гармоническая	$\overline{X}_{ГМ} = \frac{N}{\sum \frac{1}{X_i}}$ (19)	$\overline{X}_{ГМ} = \frac{\sum f_i}{\sum \frac{f_i}{X_i}}$ (20)	Для осреднения величин с дробной размерностью при наличии дополнительных данных по числителю дробной размерности
0	Геометрическая	$\overline{X}_{геом} = \sqrt[N]{\prod_{i=1}^{N} X_i}$ (21)	$\overline{X}_{геом} = \sqrt[N]{\prod_{i=1}^{N} X_i^{f_i}}$ (22)	Для осреднения цепных индексов динамики
2	Квадратическая	$\overline{X}_{кв} = \sqrt{\frac{\sum X_i^2}{N}}$ (23)	$\overline{X}_{кв} = \sqrt{\frac{\sum X_i^2 f_i}{\sum f_i}}$ (24)	Для осреднения вариации признака (расчет средних отклонений)
3	Кубическая	$\overline{X}_{куб} = \sqrt[3]{\frac{\sum X_i^3}{N}}$ (25)	$\overline{X}_{куб} = \sqrt[3]{\frac{\sum X_i^3 f_i}{\sum f_i}}$ (26)	Для расчета индексов нищеты населения
1	Хронологическая	$\overline{X}_{ХР} = \frac{\frac{X_1 + X_N}{2} + \sum_{2}^{N-1} X_i}{N-1}$ (27)	$\overline{X}_{ХР} \frac{\sum (X_i + X_{i+1}) f_i}{2\sum f_i}$ (28)	Для осреднения моментных статистических величин

Выбор вида формулы средней величины зависит от содержания осредняемого признака и конкретных данных, по которым ее приходится вычислять. Показатель степени m в общей формуле средней величины оказывает существенное влияние на значение средней величины: по мере увеличения степени возрастает и средняя величина (правило мажорантности средних величин), то есть $\overline{X}_{ГМ} < \overline{X}_{геом} < \overline{X}_{ар} < \overline{X}_{КВ} < \overline{X}_{куб}$. Так, если $m \to +\infty$, то $\overline{X} \to X_{\max}$, а если $m \to -\infty$, то $\overline{X} \to X_{\min}$.

В нашей задаче, применяя формулу (18) и подставляя вместо X_i середины интервалов возраста $X_И$, определяем средний возраст студентов: $\overline{X}_{ар}$ = 549,163/25 = 21,967 (года). Теперь осталось определить типичность или нетипичность найденной средней величины. Это осуществляется с помощью расчета показателей вариации. Чем ближе они к нулю, тем типичнее найденная средняя величина для изучаемой статистической совокупности. При этом критериальным значением коэффициента вариации служит 1/3.

Коэффициенты вариации рассчитываются как отношение среднего отклонения к средней величине. Поскольку среднее отклонение может определяться линейным и квадратическим способами, то соответствующими могут быть и коэффициенты вариации.

Среднее линейное отклонение определяется по формулам (29) и (30):

$$Л = \frac{\sum \left| X_i - \overline{X} \right|}{N} \text{ – простое;} \qquad (29)$$

$$Л = \frac{\sum \left| X_i - \overline{X} \right| f_i}{\sum f_i} \text{ – взвешенное.} \qquad (30)$$

Среднее квадратическое отклонение определяется как корень квадратный из дисперсии, то есть по формуле (31):

$$\sigma = \sqrt{Д}\,. \qquad (31)$$

Дисперсия определяется по формулам (32) или (33):

$$Д = \frac{\sum \left(X_i - \overline{X} \right)^2}{N} \text{ – простая;} \qquad (32)$$

$$Д = \frac{\sum \left(X_i - \overline{X} \right)^2 f_i}{\sum f_i} \text{ – взвешенная.} \qquad (33)$$

В нашей задаче, применяя формулу (30), определим ее числитель и внесем в расчетную таблицу. В итоге получим среднее линейное отклонение: *Л* = 54,937/25 = 2,198 (года). Разделив это значение на средний возраст, получим *линейный коэффициент вариации*: $\lambda = \frac{Л}{\overline{X}}$ = 2,198/21,967 = 0,100. По значению этого коэффициента для рассмотренной группы студентов делаем вывод о типичности среднего возраста, т.к. расчетное значение коэффициента вариации не превышает критериального (0,100 < 0,333).

Применяя формулу (33), получим в итоге дисперсию: *Д* = 164,018/25 = 6,561. Извлечем из этого числа корень и получим в результате среднее квадратическое отклонение: $\sigma = \sqrt{Д}$ = 2,561 (года). Разделив это значение на средний возраст, получим *квадратический коэффициент вариации*: $v = \frac{\sigma}{\overline{X}}$ = 2,561/21,967 = 0,117. По значению этого коэффициента для рассмотренной группы студентов можно сделать вывод о типичности среднего возраста, т.к. расчетное значение коэффициента вариации не превышает критериального (0,117 < 0,333).

В качестве *показателей асимметрии* используются: коэффициент асимметрии – нормированный момент третьего порядка (34) и коэффициент асимметрии Пирсона (35):

$$r_3 = \frac{\mu_3}{\sigma^3}, \qquad (34) \qquad\qquad As = \frac{\overline{X} - Mo}{\sigma}. \qquad (35)$$

Если значение коэффициента асимметрии положительно, то в ряду преобладают варианты, которые больше средней (правосторонняя скошенность), если отрицательно – левосторонняя скошенность. Если коэффициент асимметрии равен 0, то вариационный ряд симметричен.

В нашей задаче $\mu_3 = \frac{\sum\left(X_i - \overline{X}\right)^3 f_i}{\sum f_i}$ =383,636/25 = 15,345; σ^3 =2,561^3= 16,797; r_3=15,345/16,797 = 0,914 > 0, значит, распределение студентов по росту с правосторонней асимметрией. Это подтверждает и значение коэффициента асимметрии Пирсона: *As* = (21,967-20)/2,561 = 0,768.

Для характеристики *крутизны распределения* используется центральный момент 4-го порядка:

$$\mu_4 = \frac{\sum\left(X_i - \overline{X}\right)^4 f_i}{\sum f_i}. \qquad (36)$$

Для образования безразмерной характеристики определяется нормированный момент 4-го порядка $r_4 = \frac{\mu_4}{\sigma^4}$, который и характеризует крутизну (заостренность) графика распределения. При измерении асимметрии эталоном служит нормальное (симметричное) распределение, для которого r_4=3. Поэтому для оценки крутизны данного распределения в сравнении с нормальным вычисляется эксцесс распределения (37):

$$Ex = \frac{\mu_4}{\sigma^4} - 3. \qquad (37)$$

Для приближенного определения эксцесса может быть использована формула Линдберга (38):

$$Ex = d_{\sigma/2} - 0{,}3829\,, \qquad (38)$$

где $d_{\sigma/2}$– доля количества вариант, лежащих в интервале, равном половине σ (в ту и другую сторону от средней величины).

В нашей задаче числитель центрального момента 4-го порядка рассчитан в последнем столбце расчетной таблицы. В итоге по формуле (37) имеем: Ex = (2780,498/25)/$2{,}561^4$–3 = 111,220/43,017–3 = -0,415. Так как Ex<0, то распределение низковершинное. Это подтверждает и приблизительный расчет по формуле (38): в интервале $21{,}967 \pm 0{,}5*2{,}561$, то есть от 20,687 до 23,248 находится примерно 21,4% студентов. Таким образом, Ex = 0,214 – 0,3829 = –0,169.

Контрольные задания по теме

По имеющимся в следующей таблице данным по группе из 20 студентов заочного отделения необходимо:

1) построить интервальный ряд распределения признака и его график;

2) рассчитать модальное, медианное и среднее значение, установить его типичность с помощью коэффициентов вариации;

3) проверить распределение на нормальность с помощью коэффициентов асимметрии и эксцесса.

№ п/п	**Вариант**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	Рост, см	Вес, кг	Доход, у.е./мес.	IQ (тест Айзенка)	Тет-радь, листов	Воз-раст, лет	Соот-ношение «рост/вес»	Стаж работы, мес.	Кол-во друзей, чел.	Время решения контрольной, час.
1	159	45	430	95	24	20	3,533	26	5	8,5
2	160	61	640	115	32	25	2,623	63	7	6,2
3	161	56	610	111	24	28	2,875	94	10	6,8
4	162	48	330	97	24	19	3,375	16	4	12,0
5	162	54	420	105	60	23	3,000	49	2	7,5
6	164	58	290	98	16	20	2,828	14	6	10,0
7	166	51	480	109	90	26	3,255	78	9	7,2
8	169	62	610	120	24	19	2,726	10	5	4,2
9	170	70	840	122	48	30	2,429	130	10	3,5
10	170	72	330	92	24	20	2,361	20	3	9,5
11	171	73	560	110	16	28	2,342	86	8	7,8

№ п/п	Вариант									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	Рост, см	Вес, кг	Доход, у.е./мес.	IQ (тест Айзенка)	Тет-радь, листов	Воз-раст, лет	Соот-ношение «рост/вес»	Стаж работы, мес.	Кол-во друзей, чел.	Время решения контрольной, час.
12	171	64	450	102	48	21	2,672	29	4	8,0
13	172	73	350	108	32	26	2,356	75	7	6,0
14	174	68	310	100	48	21	2,559	22	4	4,8
15	176	81	380	104	64	20	2,173	32	1	8,6
16	176	84	340	104	48	19	2,095	21	5	10,0
17	178	76	660	128	90	27	2,342	96	8	4,5
18	181	90	450	106	48	26	2,011	70	9	12,5
19	183	68	540	105	32	23	2,691	59	6	10,5
20	192	95	750	117	60	27	2,021	98	4	6,5

Тема 3. Выборочное наблюдение

Методические указания по теме

Задача 1. На предприятии в порядке случайной бесповторной выборки было опрошено 100 рабочих из 1000 и получены следующие данные об их доходе за месяц:

Доход, у.е.	до 300	300-500	500-700	700-1000	более 1000
Число рабочих	8	28	44	17	3

С вероятностью 0,950 определить:

1) среднемесячный размер дохода работников данного предприятия;

2) долю рабочих предприятия, имеющих месячный доход более 700 у.е.;

3) необходимую численность выборки при определении среднемесячного дохода работников предприятия, чтобы не ошибиться более чем на 50 у.е.;

4) необходимую численность выборки при определении доли рабочих с размером месячного дохода более 700 у.е., чтобы при этом не ошибиться более чем на 5%.

Решение. Выборочный метод (выборка) используется, когда применение сплошного наблюдения физически невозможно из-за огромного массива данных или экономической нецелесообразности. Учитывая, что на основе выборочного обследования нельзя точно оценить изучаемый параметр (например, среднее значение – $\overline{X}$ или долю какого-то признака – p) генеральной совокупности, необходимо найти пределы, в которых он находится. Для этого необходимо определить изучаемый параметр по данным выборки (выборочную среднюю – $\tilde{X}$ и/или выборочную долю – w) и его дисперсию ($Д_в$). Для этого построим вспомогательную таблицу 3.

Таблица 3. Вспомогательные расчеты для решения задачи

X_i	f_i	$X_И$	$X_И f_i$	$(X_И - \tilde{X})^2$	$(X_И - \tilde{X})^2 f_i$
до 300	8	200	1600	137641	1101128
300 - 500	28	400	11200	29241	818748
500 - 700	44	600	26400	841	37004
700 - 1000	17	850	14450	77841	1323297
более 1000	3	1150	3450	335241	1005723
Итого	100		57100		4285900

По формуле (18) получим средний доход в выборке: $\tilde{X}$ = 57100/100 = 571 (у.е.). Применив формулу (33) и рассчитав ее числитель в последнем столбце таблицы, получим дисперсию среднего выборочного дохода: $Д_в$ = 4285900/100 = 42859.

Затем необходимо определить *предельную ошибку выборки* по формуле (39)[1]:

$$\Delta = t\mu, \qquad (39)$$

где t – *коэффициент доверия*, зависящий от вероятности, с которой определяется предельная ошибка выборки; μ – *средняя ошибка выборки*, определяемая для повторной выборки по формуле (40), а для бесповторной – по формуле (41):

$$\mu = \sqrt{\frac{Д_в}{n}}, \qquad (40) \qquad\qquad \mu = \sqrt{\frac{Д_в}{n}\left(1-\frac{n}{N}\right)}, \qquad (41)$$

где n – численность выборки; N – численность генеральной совокупности.

В нашей задаче выборка бесповторная, значит, применяя формулу (41), получим среднюю ошибку выборки при определении среднего возраста в генеральной совокупности: $\mu = \sqrt{\frac{42859}{100}\left(1-\frac{100}{1000}\right)} = 19{,}640$ (у.е.).

Для определения средней ошибки выборки при определении доли рабочих с доходами более 700 у.е. в генеральной совокупности необходимо определить дисперсию этой доли. Дисперсия доли альтернативного признака w (признак, который может принимать только два взаимоисключающих значения – например, больше или меньше определенного значения) определяется по формуле (42):

$$Д_в^{альт} = w(1-w). \qquad (42)$$

В нашей задаче долю альтернативного признака (рабочие с доходами более 700 у.е.) найдем как отношение числа таких рабочих к общему числу рабочих в выборке: w = 20/100 = 0,2 или 20%. Теперь определим дисперсию этой доли по формуле (42): $Д_в^{альт}$=0,2*(1-0,2) = 0,16. Теперь можно рассчитать среднюю ошибку выборки по формуле (41): $\mu = \sqrt{\frac{0{,}16}{100}\left(1-\frac{100}{1000}\right)} = 0{,}038$ или 3,8%.

Значения вероятности β и коэффициента доверия t имеются в математических таблицах нормального закона распределения вероятностей (если в выборке более 30 единиц), из которых в статистике широко применяются сочетания, приведенные в таблице 4:

Таблица 4. Значения интеграла вероятностей Лапласа

β	0,683	0,866	0,950	0,954	0,988	0,997
t	1	1,5	1,96	2	2,5	3

В нашей задаче β = 0,950, значит t = 1,96 (то есть предельная ошибка выборки в 1,96 раза больше средней). Предельная ошибка выборки по формуле (39) будет равна: Δ = 1,96*19,64 = 38,494 (у.е.) при определении среднего

[1] Если в выборке более 30 единиц генеральной совокупности

дохода; Δ = 1,96*0,038 = 0,075 или 7,5% при определении доли рабочих с доходами более 700 у.е.

После расчета предельной ошибки находят *доверительный интервал* обобщающей характеристики генеральной совокупности по формуле (43) – для средней величины и по формуле (44) – для доли альтернативного признака:

$$(\tilde{X} - \Delta) \le \overline{X} \le (\tilde{X} + \Delta) \qquad (43) \qquad\qquad (w - \Delta) \le p \le (w + \Delta) \qquad (44)$$

В нашей задаче по формуле (43): 571-38,494$\le \overline{X} \le$571+38,494 или 532,506 у.е. $\le \overline{X} \le$ 609,494 у.е., то есть средний доход всех рабочих предприятия с вероятностью 95% будет лежать в пределах от 532,5 до 609,5 у.е.

Аналогично определяем доверительный интервал для доли по формуле (44): 0,2-0,075$\le p \le$0,2+0,075 или 0,125$\le p \le$0,275, то есть доля рабочих с доходами более 700 у.е. на всем предприятии с вероятностью 95% будет лежать в пределах от 12,5% до 27,5%.

При разработке программы выборочного наблюдения очень часто задается конкретное значение предельной ошибки (Δ) и уровень вероятности (β). Неизвестной остается минимальная численность выборки (n), обеспечивающая заданную точность. Ее можно получить, если подставить формулу (40) или (41) в формулу (39) и выразить из них n. В результате получатся формулы для вычисления необходимой численности повторной (45) и бесповторной (46) выборок.

$$n_{повт} = \frac{Д_в t^2}{\Delta^2}; \qquad (45) \qquad\qquad n_{б/повт} = \frac{Д_в t^2}{\Delta^2 + Д_в t^2 / N}. \qquad (46)$$

В нашей задаче выборка бесповторная, значит, воспользуемся формулой (46), в которую подставим уже рассчитанные дисперсии среднего выборочного дохода рабочих ($Д_в$ = 42859) и доли рабочих с доходами более 700 у.е. ($Д_в$ = 0,16):

$$n_{б/повт} = \frac{42859 * 1{,}96^2}{50^2 + 42859 * 1{,}96^2 / 1000} = 62 \text{ (чел.)}, \qquad n_{б/повт} = \frac{0{,}16 * 1{,}96^2}{0{,}05^2 + 0{,}16 * 1{,}96^2 / 1000} = 197 \text{ (чел.)}.$$

Таким образом, необходимо включить в выборку не менее 62 рабочих при определении среднего месячного дохода работников предприятия, чтобы не ошибиться более чем на 50 у.е., и не менее 197 рабочих при определении доли рабочих с размером месячного дохода более 700 у.е., чтобы при этом не ошибиться более чем на 5%.

Контрольные задания по теме

Для изучения вкладов населения в коммерческом банке города была проведена 5%-я случайная выборка лицевых счетов, в результате которой получено следующее распределение клиентов по размеру вкладов:

Размер вклада, у.е.	Число вкладчиков, чел.									
	Вариант									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
до 5000	10	80	100	50	60	30	90	20	70	40
5 000 – 15 000	40	60	150	30	40	110	75	65	90	80
15 000 – 30 000	25	35	70	90	120	90	130	140	60	95
30 000 – 50 000	30	45	40	5	80	30	60	75	20	115
свыше 50 000	15	10	30	25	50	15	25	5	10	5

С вероятностью 0,954 определить:

1) средний размер вклада во всем банке;

2) долю вкладчиков во всем банке с размером вклада свыше 15000 у.е.;

3) необходимую численность выборки при определении среднего размера вклада, чтобы не ошибиться более чем на 500 у.е.;

4) необходимую численность выборки при определении доли вкладчиков во всем банке с размером вклада свыше 30 000 у.е., чтобы не ошибиться более чем на 10%.

Тема 4. Ряды динамики

Методические указания по теме

Задача 1. Смертность от болезней системы кровообращения в России за период 1995-2004 гг. характеризуется следующим рядом динамики.

Год	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Умершие, тыс. чел.	1163,5	1113,7	1100,3	1094,1	1187,8	1231,4	1253,1	1308,1	1330,5	1287,7

Вычислить: абсолютные, относительные, средние изменения и их темпы базисным и цепным способами. Проверить ряд на наличие в нем линейного тренда, на основе которого рассчитать интервальный прогноз на 2005 год с вероятностью 95%.

Решение. Любое изменение уровней ряда динамики определяется *базисным* (сравнение с первым уровнем) и *цепным* (сравнение с предыдущим уровнем) способами. Оно может быть *абсолютным* (разность уровней ряда) и *относительным* (соотношение уровней).

Базисное абсолютное изменение представляет собой разность конкретного и первого уровней ряда (47), а *цепное абсолютное изменение* представляет собой разность конкретного и предыдущего уровней ряда (48).

$$\Delta Y^{Б} = Y_i - Y_1. \qquad (47) \qquad\qquad \Delta Y^{Ц} = Y_i - Y_{i-1}. \qquad (48)$$

По знаку абсолютного изменения делается вывод о характере развития явления: при $\Delta Y > 0$ — рост, при $\Delta Y < 0$ — спад, при $\Delta Y = 0$ — стабильность.

В нашей задаче эти изменения определены в 3-м и 4-м столбцах таблицы 5. Для проверки правильности расчетов применяется правило, согласно которому сумма цепных абсолютных изменений равняется последнему базисному. В нашей задаче это правило выполняется: $\sum \Delta Y^{Ц}$ =124,2 и $\Delta Y^{Б}_{2004}$=124,2.

Базисное относительное изменение представляет собой соотношение конкретного и первого уровней ряда (49), а *цепное относительное изменение* представляет собой соотношение конкретного и предыдущего уровней ряда (50).

$$i^{Б} = Y_i / Y_1. \qquad (49) \qquad\qquad i^{Ц} = Y_i / Y_{i-1}. \qquad (50)$$

Относительные изменения уровней — это по существу индексы динамики, критериальным значением которых служит 1. Если они больше ее, имеет место рост явления, меньше ее — спад, а при равенстве единице наблюдается стабильность явления.

В нашей задаче эти изменения определены в 5-м и 6-м столбцах таблицы 5.

Вычитая единицу из относительных изменений, получают *темп изменения* уровней, критериальным значением которого служит 0. При положительном темпе изменения имеет место рост явления, при отрицательном — спад, а при нулевом темпе изменения наблюдается стабильность явления. В нашей задаче темпы изменения определены в 7-м и 9-м столбцах таблицы 5, а в 8-м и 10-м сделан вывод о характере развития изучаемого явления. Для проверки правильности расчетов применяется правило, согласно которому произведение цепных относительных изменений равняется последнему базисному. В нашей задаче это правило выполняется: $\prod i^{Ц}$ =1,107 и $i^{Б}_{2004}$=1,107.

Таблица 5. Вспомогательные расчеты для решения задачи

Год	Y	$\Delta Y^{Б}$	$\Delta Y^{Ц}$	$i^{Б}$	$i^{Ц}$	$T^{Б}$	хар-р	$T^{Ц}$	хар-р
1995	1163,5								
1996	1113,7	-49,8	-49,8	0,96	0,96	-0,04	спад	-0,04	спад
1997	1100,3	-63,2	-13,4	0,95	0,99	-0,05	спад	-0,01	спад
1998	1094,1	-69,4	-6,2	0,94	0,99	-0,06	спад	-0,01	спад
1999	1187,8	24,3	93,7	1,02	1,09	0,021	рост	0,086	рост
2000	1231,4	67,9	43,6	1,06	1,04	0,058	рост	0,037	рост
2001	1253,1	89,6	21,7	1,08	1,02	0,077	рост	0,018	рост
2002	1308,1	145	55	1,12	1,04	0,124	рост	0,044	рост
2003	1330,5	167	22,4	1,14	1,02	0,144	рост	0,017	рост
2004	1287,7	124	-42,8	1,11	0,97	0,107	рост	-0,03	спад
Итого	12070		124		1,11				

Обобщенной характеристикой ряда динамики является *средний уровень ряда* $\bar{y}$. Способ расчета $\bar{y}$ зависит от того, моментный ряд или интервальный (см. рис.3):

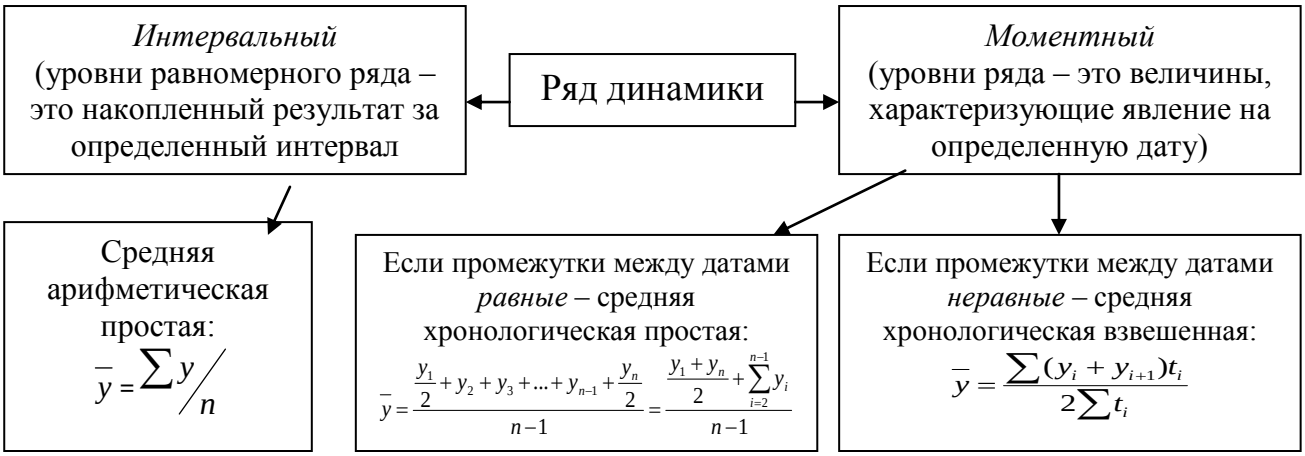


Рис.3. Методы расчета среднего уровня ряда динамики.

В нашей задаче ряд динамики интервальный, значит, применяем формулу средней арифметической простой (17): $\bar{y}$ = 12070,2 / 10 = 1207,02 (тыс. чел.). То есть за период 1995-2004 в России в среднем за год от болезней системы кровообращения умирало 1207,02 тыс. чел.

Кроме среднего уровня в рядах динамики рассчитываются и другие средние показатели – среднее изменение уровней ряда (базисным и цепным способами), средний темп изменения.

Базисное среднее абсолютное изменение – это частное от деления последнего базисного абсолютного изменения на количество изменений уровней (51). *Цепное среднее абсолютное изменение* уровней ряда – это частное от деления суммы всех цепных абсолютных изменений на количество изменений (52).

$$\Delta\bar{Y}^{\boldsymbol{Б}} = \frac{\Delta Y_n^{Б}}{n-1}. \qquad (51) \qquad\qquad \Delta\bar{Y}^{\boldsymbol{Ц}} = \frac{\sum \Delta Y^{Ц}}{n-1}. \qquad (52)$$

По знаку средних абсолютных изменений также судят о характере изменения явления в среднем: рост, спад или стабильность. Из правила контроля базисных и цепных абсолютных изменений следует, что базисное и цепное среднее изменение должны быть равными. В нашей задаче $\Delta\bar{Y}$ = 124,2/9 = 13,8, то есть ежегодно в среднем смертность от болезней системы кровообращения растет на 13,8 тыс. чел.

Наряду со средним абсолютным изменением рассчитывается и среднее относительное. *Базисное среднее относительное изменение* определяется по формуле (53), а *цепное среднее относительное изменение* – по формуле (54):

$$\bar{i}^{\boldsymbol{Б}} = \sqrt[n-1]{i_n^{Б}} = \sqrt[n-1]{Y_n / Y_1}. \qquad (53) \qquad\qquad \bar{i}^{\boldsymbol{Ц}} = \sqrt[n-1]{\prod i_n^{Ц}}. \qquad (54)$$

Естественно, базисное и цепное среднее относительное изменения должны быть одинаковыми и сравнением их с критериальным значением 1 делается вывод о характере изменения явления в среднем: рост, спад или стабильность. В нашей задаче $\bar{i} = \sqrt[9]{1{,}107}$ = 1,0114, то есть ежегодно в среднем смертность от болезней системы кровообращения растет в 1,0114 раза.

Вычитанием 1 из среднего относительного изменения образуется соответствующий *средний темп изменения*, по знаку которого также можно судить о характере изменения изучаемого явления, отраженного данным рядом динамики. В нашей задаче $\bar{T}$ = 1,0114 – 1 = 0,0114, то есть ежегодно в среднем смертность от болезней системы кровообращения растет на 1,14%.

Проверка ряда динамики на наличие в нем *тренда* (тенденции развития ряда) возможна несколькими способами (метод средних, Фостера и Стюарта, Валлиса и Мура и пр.), но наиболее простым является графическая модель, где на графике по оси абсцисс откладывается время, а по оси ординат – уровни ряда. Соединив полученные точки линиями, в большинстве случаев можно

выявить тренд визуально. Тренд может представлять собой прямую линию, параболу, гиперболу и т.п. В итоге приходим к трендовой модели вида:

$$Y_t = \hat{y}_t + \varepsilon_t, \qquad (55)$$

где $\hat{y}_t$ – математическая функция развития; ε_t – случайное или циклическое отклонение от функции; t – время в виде номера периода (уровня ряда). Цель такого метода – выбор теоретической зависимости $\hat{y}_t$ в качестве одной из функций:

$\hat{y}_t = a_0 + a_1 t$ – прямая линия; $\hat{y}_t = a_0 + \frac{a_1}{t}$ – гипербола; $\hat{y}_t = a_0 + a_1 t + a_2 t^2$ – парабола;

$\hat{y}_t = a_0 t^{a_1}$ – степенная; $\hat{y}_t = a_0 + \sum_{k=1}^{m}(a_k \cos kt + b_k \sin kt)$ – ряд Фурье.

Для выявления *тренда* (тенденции развития ряда) в нашей задаче построим график *Y(t)* (рис.4):

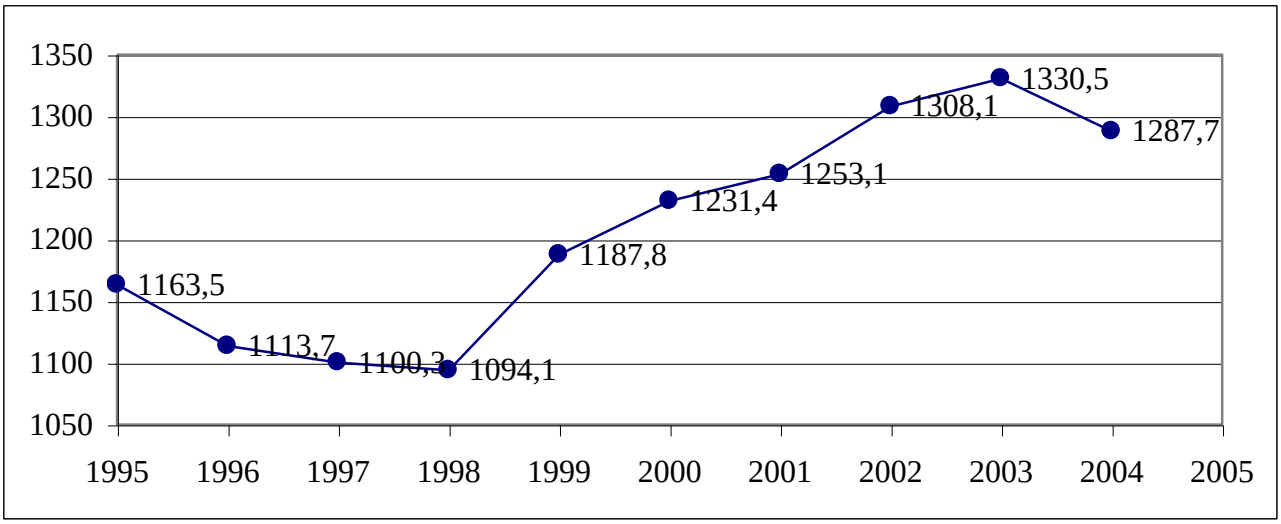


Рис.4. График динамики смертности от болезней системы кровообращения в РФ.

Из данного графика видно, что есть все основания принять уравнение тренда в виде линейной функции.

Определение параметров a_n в этих функциях может вестись несколькими способами, но самые незначительные отклонения аналитических (теоретических) уровней ($\hat{y}_t$ – читается как «игрек, выравненный по *t*») от фактических (Y_t) дает *метод наименьших квадратов – МНК*. При этом методе учитываются все эмпирические уровни и должна обеспечиваться минимальная сумма квадратов отклонений эмпирических значений уровней Y_t от теоретических уровней $\hat{y}_t$:

$$\sum(\hat{y}_t - y)^2 \to \min. \qquad (56)$$

В нашей задаче при выравнивании по прямой вида $\hat{y}_t = a_0 + a_1 t$ параметры a_0 и a_1 отыскиваются по МНК следующим образом. В формуле (55) вместо $\hat{y}_t$ записываем его конкретное выражение $a_0 + a_1 t$. Тогда $S = \sum (a_0 + a_1 t - y)^2 \to \min$. Дальнейшее решение сводится к задаче на экстремум, т.е. к определению того, при каком значении a_0 и a_1 функция двух переменных S может достигнуть минимума. Как известно, для этого надо найти частные производные S по a_0 и a_1, приравнять их к нулю и после элементарных преобразований решить систему двух уравнений с двумя неизвестными.

В соответствии с вышеизложенным найдем частные производные:

$$\begin{cases} \dfrac{\partial S}{\partial a_0} = 2\sum (a_0 + a_1 t - y) = 0 \\ \dfrac{\partial S}{\partial a_1} = 2\sum (a_0 + a_1 t - y)t = 0 \end{cases}$$

Сократив каждое уравнение на 2, раскрыв скобки и перенеся члены с y в правую сторону, а остальные – оставив в левой, получим систему нормальных уравнений:

$$\begin{cases} na_0 + a_1 \sum t = \sum y \\ a_0 \sum t + a_1 \sum t^2 = \sum yt \end{cases} \qquad (57)$$

где n – количество уровней ряда; t – порядковый номер в условном обозначении периода или момента времени; y – уровни эмпирического ряда.

Эта система и, соответственно, расчет параметров a_0 и a_1 упрощаются, если отсчет времени ведется от середины ряда. Например, при *нечетном* числе уровней серединная точка (год, месяц) принимается за нуль. Тогда предшествующие периоды обозначаются соответственно –1, –2, –3 и т.д., а следующие за средним (центральным) – соответственно 1, 2, 3 и т.д. При *четном* числе уровней два серединных момента (периода) времени обозначают –1 и +1, а все последующие и предыдущие, соответственно, через два интервала: ± 3, ± 5, ± 7 и т.д.

При таком порядке отсчета времени (от середины ряда) $\sum t = 0$, поэтому, система нормальных уравнений упрощается до следующих двух уравнений, каждое из которых решается самостоятельно:

$$\begin{cases} na_0 = \sum y \Rightarrow a_0 = \dfrac{\sum y}{n} \\ a_1 \sum t^2 = \sum yt \Rightarrow a_1 = \dfrac{\sum yt}{\sum t^2} \end{cases} \qquad (58)$$

Как видим, при такой нумерации периодов параметр a_0 представляет собой средний уровень ряда. Определим по формуле (58) параметры уравнения прямой, для чего исходные данные и все расчеты необходимых сумм представим в таблице 6.

Из таблицы получаем, что a_0 = 12070,2/10 = 1207,02 и a_1 = 4195/330 = 12,7121. Отсюда искомое уравнение тренда $\hat{y}_t = 1207{,}02 + 12{,}7121t$. В 6-м столбце таблицы 6 приведены трендовые уровни, рассчитанные по этому уравнению. Для иллюстрации построим график эмпирических (маркеры-кружочки) и трендовых уровней (рис.5).

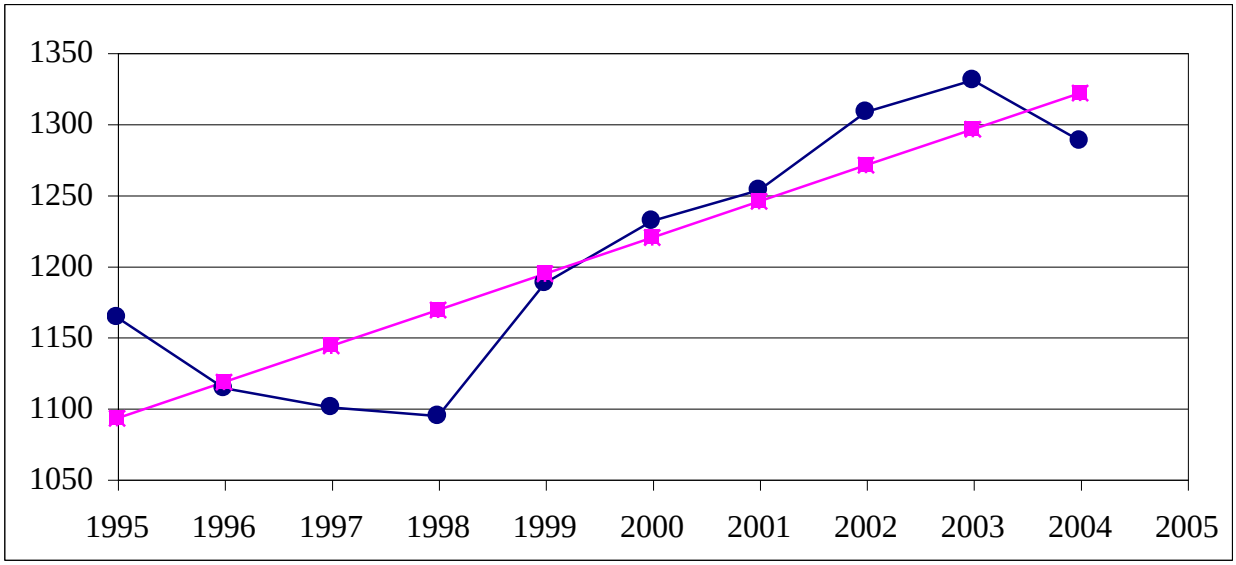


Рис.5. График эмпирических и трендовых уровней смертности от болезней системы кровообращения в РФ.

По полученной модели для каждого периода (каждой даты) определяются теоретические уровни тренда ($\hat{y}_t$) и оценивается *надежность (адекватность) выбранной модели тренда*. Оценку надежности проводят с помощью критерия Фишера, сравнивая его расчетное значение F_p с теоретическими значениями F_T (приложение 1). При этом расчетный критерий Фишера определяется по формуле:

$$F_P = \frac{(n-k)Д_A}{(k-1)Д_o}, \qquad (59)$$

где k – число параметров (членов) выбранного уравнения тренда; $Д_A$ – аналитическая дисперсия, определяемая по формуле (61); $Д_o$ – остаточная дисперсия (62), определяемая как разность фактической дисперсии $Д_Ф$ (60) и аналитической дисперсии:

$$Д_Ф = \frac{\sum(y-\bar{y})^2}{n}; \qquad (60)$$

$$Д_A = \frac{\sum(\hat{y}_t-\bar{y})^2}{n}; \qquad (61)$$

$$Д_O = \frac{\sum(\hat{y}_t-y)^2}{n} = Д_Ф - Д_A. \qquad (62)$$

Сравнение расчетного и теоретического значений критерия Фишера ведется обычно при уровне значимости α с учетом степеней свободы $\nu_1 = k-1$ и

$\nu_2 = n - k$. Уровень значимости α связан с вероятностью β следующей формулой $\beta = 1 - \alpha$. При условии $F_p > F_T$ считается, что выбранная математическая модель ряда динамики адекватно отражает обнаруженный в нем тренд.

Таблица 6. Вспомогательные расчеты для решения задачи

Год	y	t	t^2	yt	$\hat{y}_t$	$(y - \hat{y}_t)^2$	$(\hat{y}_t - \bar{y})^2$	$(y - \bar{y})^2$
1995	1163,5	-9	81	-10471,5	1092,611	5025,263	13089,44	1893,9904
1996	1113,7	-7	49	-7795,9	1118,035	18,79354	7918,3033	8708,6224
1997	1100,3	-5	25	-5501,5	1143,459	1862,733	4039,9506	11389,1584
1998	1094,1	-3	9	-3282,3	1168,884	5592,592	1454,3822	12750,9264
1999	1187,8	-1	1	-1187,8	1194,308	42,35249	161,59803	369,4084
2000	1231,4	1	1	1231,4	1219,732	136,1394	161,59803	594,3844
2001	1253,1	3	9	3759,3	1245,156	63,10136	1454,3822	2123,3664
2002	1308,1	5	25	6540,5	1270,581	1407,705	4039,9506	10217,1664
2003	1330,5	7	49	9313,5	1296,005	1189,915	7918,3033	15247,3104
2004	1287,7	9	81	11589,3	1321,429	1137,652	13089,44	6509,2624
Итого	12070,2	0	330	4195	12070,2	16476,25	53327,348	69803,596

Проверим тренд в нашей задаче на адекватность по формуле (59), для чего в 7-м столбце таблицы 6 рассчитан числитель остаточной дисперсии, а в 8-м столбце – числитель аналитической дисперсии. В формуле (59) можно использовать их числители, так как оба они делятся на число уровней n (n сократятся): F_P = 53327,348*8/(16476,25*1) = 25,893 > F_T, значит, модель адекватна и ее можно использовать для прогнозирования (F_T= 5,32 находим по приложению 1 в 1-ом столбце [$\nu_1 = k - 1 = 1$] и 8-й строке [$\nu_2 = n - k = 8$]).

При составлении прогнозов уровней социально-экономических явлений обычно оперируют не точечной, а интервальной оценкой, рассчитывая так называемые *доверительные интервалы прогноза*. Границы интервалов определяются по формуле (63):

$$\hat{y}_t \pm t_\alpha \sigma_{\hat{y}}, \qquad (63)$$

где $\hat{y}_t$ – точечный прогноз, рассчитанный по модели тренда; t_α – *коэффициент доверия по распределению Стьюдента* при уровне значимости α и числе степеней свободы $\nu = n - 1$ (приложение 2); $\sigma_{\hat{y}}$ – *ошибка аппроксимации*, определяемая по формуле (64):

$$\sigma_{\hat{y}} = \sqrt{\frac{\sum (y - \hat{y}_t)^2}{n - k}}, \qquad (64)$$

где y и $\hat{y}_t$ – соответственно фактические и теоретические (трендовые) значения уровней ряда динамики; n – число уровней ряда; k – число параметров (членов) в уравнении тренда.

Определим доверительный интервал в нашей задаче на 2005 год с уровнем значимости α = (1–0,95) = 0,05. Для этого найдем ошибку

аппроксимации по формуле (64): $\sigma_{\hat{y}} = \sqrt{16476{,}25/(10-2)} =$ 45,38. Коэффициент доверия по распределению Стьюдента t_α = 2,2622 при ν = 10 – 1=9.

Прогноз на 2005 с вероятностью 95% осуществим по формуле (63):

Y_{2005}=(1207,02+12,7121*11) $\pm$ 2,2622*45,38 или 1244,19<Y_{2005}<1449,51 (тыс.чел.).

Контрольные задания по теме

По статистическим данным по России за 2000 – 2005 гг. вычислить: абсолютные, относительные, средние изменения и их темпы базисным и цепным способами. Проверить ряд на наличие в нем линейного тренда, на основе которого рассчитать интервальный прогноз на 2006 год с вероятностью 95%.

Год	**Вариант**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	Валовой сбор сахарной свеклы, млн.т.	Валовой сбор картофеля, млн.т.	Число заключенных браков, тыс.	Число построенных жилых домов, млн.м^2	Поголовье крупного рогатого скота, млн.голов (на конец года)	Производство мяса, млн.т.	Производство яиц, млрд.шт.	Численность населения, тыс.чел. (на начало года)	Среднегодовая численность занятых в экономике, тыс.чел.	Доля расходов на оплату ЖКХ в бюджете домохозяйств, %
2000	14,1	34	897,3	30,3	16,5	4,4	34,1	146890	64327	4,6
2001	14,6	35	1001,6	31,7	15,8	4,5	35,2	146304	64710	5,2
2002	15,7	32,9	1019,8	33,8	15,0	4,7	36,3	145649	65359	6,2
2003	19,4	36,7	1091,8	36,4	13,5	4,9	36,5	144964	65666	7,2
2004	21,8	35,9	979,7	41,0	12,1	5,0	35,8	144168	66407	7,7
2005	21,4	37,3	1066,4	43,6	11,1	4,9	36,8	143474	66939	8,3

Тема 5. Индексы

Методические указания по теме

Задача 1. Имеются следующие данные о продажах торговой точкой двух видов товара:

Товар	Цена за кг, руб.		Объем продаж, тыс. кг	
	Январь	Февраль	Январь	Февраль
Апельсины	20	18	100	160
Бананы	22	25	150	120

Определить: 1) индивидуальные индексы цен, физического объема и выручки; 2) общие индексы цен, физического объема и выручки; 3) абсолютное изменение выручки за счет изменений цен, структурного сдвига и объемов продаж (для каждого фактора в отдельности) по всей продукции и по каждому товару в отдельности. По итогам расчетов сделать аргументированные выводы.

Решение. В основе решения задачи лежит формула (65):

$$Q = pq, \qquad (65)$$

где p – цена товара, q – физический объем (количество), Q – выручка (товарооборот).

Применив формулу (65) к нашей задаче, рассчитаем выручку по каждому товару в январе (Q_{0j}) и феврале (Q_{1j}) в таблице 7.

Таблица 7. Расчет выручки и ее изменения по каждому товару

Товар j	Январь Q_{0j}	Февраль Q_{1j}	Изменение выручки $\Delta Q_j = Q_{1j} - Q_{0j}$
Апельсины	20*100 = 2000	18*160 = 2880	880
Бананы	22*150 = 3300	25*120 = 3000	-300
Итого	5300	5880	580

Из таблицы видно, что абсолютное изменение общей выручки составило: $\Delta\sum Q = \sum Q_1 - \sum Q_0$ = 5880-5300 = 580 тыс. руб., то есть она выросла на 580 тыс. руб.

Общий индекс изменения выручки равняется:

$I_Q = \sum Q_1 / \sum Q_0$ = 5880/5300 = 1,1094, то есть выручка от продажи фруктов увеличилась в 1,1094 раза или на 10,94% в феврале по сравнению с январем.

Определим индивидуальные индексы цен (i_p), физического объема (i_q), выручки (i_Q) и доли товара (i_d) по формуле (2), используя в качестве X_i цены (p), физический объем (q), выручки (Q) и доли товара ($d=q/\sum q$) каждого вида фруктов соответственно. Результаты расчетов представим в таблице 8.

Таблица 8. Расчет индивидуальных индексов

Индивидуальный индекс	апельсины	бананы
количества i_q	160/100 = 1,6	120/150 = 0,8
отпускных цен i_p	18/20 = 0,9	25/22 = 1,136
выручки i_Q	2880/2000=1,44	3000/3300=0,909
доли товара i_d	(160/280)/(100/250) = 1,429	(120/280)/(150/250) = 0,714

Правильность выполненных расчетов проверяется следующим образом:

1) общее изменение выручки должно равняться сумме ее частных (по каждому товару в отдельности) изменений: $\Delta\sum Q$ = 880+(-300) = 580 (тыс. руб.);

2) произведение факторных индивидуальных индексов по периодам должно равняться соответствующему индивидуальному индексу выручки: $i_{Q\text{А}}$=1,6*0,9 =1,44; $i_{Q\text{Б}}$= 0,8*1,136 = 0,909.

Из таблицы видно, что в феврале по сравнению с январем:

– количество проданных апельсинов увеличилось в 1,6 раза или на 60%, а бананов – уменьшилось в 0,8 раза или на 20%;

– цена апельсинов понизилась в 0,9 раза или на 10%, а бананов – повысилась в 1,136 раза или на 13,6%;

– выручка по апельсинам выросла в 1,44 раза или на 44%, а по бананам – снизилась в 0,909 раза или на 9,1%;

– доля проданных апельсинов увеличилась в 1,429 раза или на 42,9%, а бананов – уменьшилась в 0,714 раза или на 28,6%.

Агрегатный общий индекс физического объема Ласпейреса определяется по формуле (66):

$$I_q^{Л} = \frac{\sum q_1 p_0}{\sum q_0 p_0}. \qquad (66)$$

В нашей задаче $I_q^{Л} = \frac{160*20+120*22}{100*20+150*22}$ = 5840/5300 = 1,10189, то есть количество проданных фруктов в базисных (январских) ценах выросло в 1,10189 раза или на 10,189% в феврале по сравнению с январем.

Агрегатный общий индекс цен Пааше рассчитывается по формуле (67):

$$I_p^{П} = \frac{\sum q_1 p_1}{\sum q_1 p_0}. \qquad (67)$$

В нашей задаче $I_p^{П} = \frac{160*18+120*25}{5840}$ = 5880/5840 = 1,00685, то есть цена проданных фруктов при объемах продаж отчетного (февральского) периода выросла в 1,00685 раза или на 0,685% в феврале по сравнению с январем.

Контроль осуществляется по формуле: $I_Q = I_q^{Л}\ I_p^{П}$ = 1,10189*1,00685 = 1,1094.

Агрегатный общий индекс цен Ласпейреса вычисляется по формуле (68):

$$I_p^{Л} = \frac{\sum p_1 q_0}{\sum p_0 q_0}. \qquad (68)$$

В нашей задаче $I_p^{Л} = \frac{18*100+25*150}{5300}$ = 5550/5300 = 1,04717, то есть цена проданных фруктов при объемах продаж базисного (январского) периода выросла в 1,04717 раза или на 4,717% в феврале по сравнению с январем.

Агрегатный общий количественный индекс Пааше рассчитывается по формуле (69):

$$I_q^{П} = \frac{\sum p_1 q_1}{\sum p_1 q_0}. \qquad (69)$$

В нашей задаче $I_q^{П}$ = 5880/5550 =1,05946, то есть количество проданных фруктов в отчетных (февральских) ценах выросло в 1,05946 раза или на 5,946% в феврале по сравнению с январем.

Контроль осуществляется по формуле: $I_Q = I_p^{Л}\ I_q^{П}$ = 1,04717*1,05946 =1,1094.

Средняя геометрическая величина определяется из индексов Ласпейреса и Пааше (по методике Фишера) по формуле (70) для количества товаров и по формуле (71) – для цен:

$$I_q^{Ф} = \sqrt{I_q^{Л} I_q^{П}}; \qquad (70) \qquad\qquad I_p^{Ф} = \sqrt{I_p^{Л} I_p^{П}}. \qquad (71)$$

В нашей задаче $I_q^{Ф} = \sqrt{1{,}10189*1{,}05946}$ =1,0805, то есть в среднем количество проданных фруктов выросло в 1,0805 раза или на 8,05%; $I_p^{Ф} = \sqrt{1{,}04717*1{,}00685}$ =1,0268, то есть в среднем цена проданных фруктов выросла в 1,0268 раза или на 2,68%.

Далее выполняется факторный анализ общей выручки. В его основе лежит следующая трехфакторная мультипликативная модель выручки:

$$I_Q = I_q^{'}\ I_d\ I_p^{П}, \qquad (72)$$

где $I_q^{'} = \frac{\sum q_1}{\sum q_0}$, I_d – индекс структурных сдвигов, показывающий как изменилась выручка под влиянием фактора изменения долей проданных фруктов в отчетном периоде по сравнению с базисным периодом. Он определяется по формуле (73):

$$I_d = \frac{\sum \frac{q_1}{\sum q_1} p_0}{\sum \frac{q_0}{\sum q_0} p_0} = \frac{\sum d_1 p_0}{\sum d_0 p_0}. \qquad (73)$$

В нашей задаче $I_d = \frac{160/280*20+120/280*22}{100/250*20+150/250*22}$ = 0,9838, то есть структурный сдвиг должен был уменьшить отчетную выручку в базисных ценах в 0,9838 раза или на 1,62%.

Тогда изменение выручки за счет изменения общего количества фруктов определяется по формуле (74):

$$\Delta\sum Q_q = (I_q^{'} - 1)\sum Q_0 \,. \qquad (74)$$

В нашей задаче $\Delta\sum Q_q$ = (1,12-1)*5300 = 636 (тыс. руб.), то есть изменение количества проданных фруктов увеличило выручку на 636 тыс. руб.

Изменение общей выручки за счет структурных сдвигов находится по формуле (75):

$$\Delta\sum Q_d = I_q^{'}(I_d - 1)\sum Q_0 \,. \qquad (75)$$

В нашей задаче $\Delta\sum Q_d$ = 1,12*(0,9838-1)*5300 = –96 (тыс. руб.), то есть структурный сдвиг в количестве проданных фруктов уменьшил выручку на 96 тыс. руб.

Изменение общей выручки за счет изменения отпускных цен рассчитывается по формуле (76):

$$\Delta\sum Q_p = I_q^{'} I_d (I_p^{П} - 1)\sum Q_0 \,. \qquad (76)$$

В нашей задаче $\Delta\sum Q_p$ =1,12*0,9838*(1,00685-1)*5300 = 40 (тыс. руб.), то есть изменение цен на фрукты увеличило выручку на 40 тыс. руб.

Контроль правильности расчетов производится по формуле (77), согласно которой общее изменение выручки равно сумме ее изменений за счет каждого фактора в отдельности.

$$\Delta\sum Q = \sum Q_1 - \sum Q_0 = \Delta\sum Q_q + \Delta\sum Q_d + \Delta\sum Q_p \,. \qquad (77)$$

В нашей задаче $\Delta\sum Q$ = 636 + (–96) + 40 = 580 тыс. руб.

Результаты факторного анализа общей выручки заносятся в последнюю строку факторной таблицы 9.

Таблица 9. Результаты факторного анализа выручки

Товар *j*	Изменение выручки, тыс. руб.	В том числе за счет		
		количества продукта	структурных сдвигов	отпускных цен
А	880	240	960	–320
Б	-300	396	–1056	360
Итого	580	636	–96	40

Наконец, ведется факторный анализ изменения частной (по каждому *j*-му товару в отдельности) выручки на основе следующей трехфакторной мультипликативной модели:

$$Q_{1j} = I_q^{'} \, i_{dj} \, i_{pj} \, Q_{0j} \, . \qquad (78)$$

Тогда изменение частной выручки за счет каждого из 3-х факторов (количество, структурный сдвиг и цена) по *j*-му виду товара определяется соответственно по формулам (79) – (81).

$$\Delta Q_{qj} = (I_q^{'} - 1) Q_{0j} \, ; \qquad (79)$$

$$\Delta Q_{dj} = I_q^{'} (i_{dj} - 1) Q_{0j} \, ; \qquad (80)$$

$$\Delta Q_{pj} = I_q^{'} i_{dj} (i_{pj} - 1) Q_{0j} \, . \qquad (81)$$

Так, по апельсинам изменение выручки за счет первого фактора (изменения общего количества проданных фруктов) по формуле (79) равно:

ΔQ_{qA}=(1,12-1)*2000 = 240 (тыс. руб.).

Аналогично по бананам: $\Delta Q_{qБ}$ = (1,12-1)*3300 = 396 (тыс. руб.)

Контроль правильности расчетов:

$\sum \Delta Q_{qj} = \Delta \sum Q_q$, то есть 240 + 396 = 636 (тыс. руб.).

Так, по апельсинам изменение выручки за счет второго фактора (структурных сдвигов в количестве проданных фруктов) по формуле (80) равно:

ΔQ_{dA}=1,12*(1,429-1)*2000 = 960 (тыс. руб.).

Аналогично по бананам: $\Delta Q_{dБ}$=1,12*(0,714-1)*3300 = –1056 (тыс. руб.).

Контроль правильности расчетов:

$\sum \Delta Q_{dj} = \Delta \sum Q_d$, то есть 960 + (–1056) = –96 (тыс. руб.).

И, наконец, по апельсинам изменение выручки за счет 3-го фактора (изменения отпускной цены) по формуле (81) равно:

ΔQ_{pA}=1,12*1,429*(0,9-1)*2000 = –320 (тыс. руб.).

Аналогично по бананам: $\Delta Q_{pБ}$=1,12*0,714*(1,136-1)*3300 = 360 (тыс. руб.).

Контроль правильности расчетов:

$\sum \Delta Q_{pj} = \Delta \sum Q_p$, то есть (–320) + 360= 40 (тыс. руб.)

Результаты факторного анализа частной выручки также заносятся в таблицу 9, в которой все числа оказались взаимно согласованными.

Контрольные задания по теме

Имеются следующие данные о продажах минимаркетом 3-х видов товаров (А, В и С):

Товар	Цена за единицу продукта, руб.		Объем продаж, тыс. штук	
	1 квартал	2 квартал	1 квартал	2 квартал
1 вариант				
А	102	105	205	195
В	56	51	380	423
С	26	30	510	490
2 вариант				
А	112	109	202	260
В	51	48	365	420
С	22	26	477	316
3 вариант				
А	99	103	198	182
В	55	59	370	361
С	20	18	502	456
4 вариант				
А	99	109	188	182
В	55	56	380	385
С	20	21	508	444
5 вариант				
А	120	110	170	220
В	60	58	350	390
С	19	20	550	490

Товар	Цена за единицу продукта, руб.		Объем продаж, тыс. штук	
	1 квартал	2 квартал	1 квартал	2 квартал
6 вариант				
А	130	125	138	198
В	50	56	339	264
С	20	21	613	511
7 вариант				
А	107	110	220	189
В	46	44	490	550
С	18	20	720	680
8 вариант				
А	95	98	264	197
В	48	50	360	294
С	26	25	448	640
9 вариант				
А	89	92	360	294
В	58	56	410	482
С	24	25	558	593
10 вариант				
А	120	125	150	108
В	44	46	513	461
С	16	19	891	550

Определить:

1. Индивидуальные индексы цен, физического объема и товарооборота;
2. Общие индексы цен, физического объема и товарооборота;
3. Абсолютные приросты товарооборота за счет изменений цен, структурного сдвига и объемов продаж (для каждого фактора в отдельности) по всей продукции и по каждому товару в отдельности.

По итогам расчетов сделать аргументированные выводы.

Тема 6. Статистическое изучение взаимосвязей

Методические указания по теме

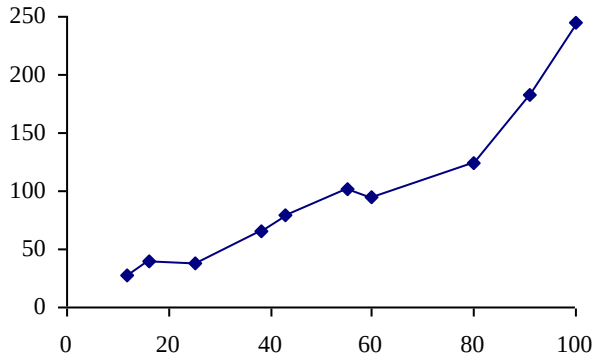
Задача 1. По условным данным таблицы 10 о стоимости основных фондов x и валовом выпуске продукции y (в порядке возрастания стоимости основных фондов) выявить наличие и характер корреляционной связи между признаками x и y.

Таблица 10. Стоимость основных фондов и валовой выпуск по 10 однотипным предприятиям

Предприятия i	Основные производственные фонды, млн. руб. x_i	Валовой выпуск продукции, млн. руб. y_i	$x_i - \bar{x}$	$y_i - \bar{y}$
1	12	28	–	–
2	16	40	–	–
3	25	38	–	–
4	38	65	–	–
5	43	80	–	–
6	55	101	+	+
7	60	95	+	–
8	80	125	+	+
9	91	183	+	+
10	100	245	+	+
Итого	520	1000		

Решение. Для выявления наличия и характера корреляционной связи между двумя признаками в статистике используется **ряд методов**.

1. *Графический метод,* когда корреляционную зависимость для наглядности можно изобразить графически. Для этого, имея n взаимосвязанных пар значений x и y и пользуясь прямоугольной системой координат, каждую такую пару изображают в виде точки на плоскости с координатами x и y. Соединяя последовательно нанесенные точки, получают ломаную линию, именуемую *эмпирической линией регрессии* (см. рисунок справа). Анализируя эту линию, визуально можно определить характер



зависимости между признаками *x* и *y*. В нашей задаче эта линия похожа на восходящую прямую, что позволяет выдвинуть гипотезу о наличии прямой зависимости между величиной основных фондов и валовым выпуском продукции.

2. *Рассмотрение параллельных данных* (значений *x* и *y* в каждой из *n* единиц). Единицы наблюдения располагают по возрастанию значений факторного признака *x* и затем сравнивают с ним (визуально) поведение результативного признака *y*. В нашей задаче в большинстве случаев по мере увеличения значений *x* увеличиваются и значения *y* (за несколькими исключениями – 2 и 3, 6 и 7 предприятия), поэтому, можно говорить о прямой связи между *x* и *y* (этот вывод подтверждает и эмпирическая линия регрессии). Теперь необходимо ее измерить, для чего рассчитывают несколько коэффициентов.

3. *Коэффициент корреляции знаков (Фехнера)* – простейший показатель тесноты связи, основанный на сравнении поведения отклонений индивидуальных значений каждого признака (*x* и *y*) от своей средней величины. При этом во внимание принимаются не величины отклонений ($x_i - \bar{x}$) и ($y_i - \bar{y}$), а их знаки («+» или «–»). Определив знаки отклонений от средней величины в каждом ряду, рассматривают все пары знаков и подсчитывают число их совпадений (*C*) и несовпадений (*H*). Тогда коэффициент Фехнера рассчитывается как отношение разности чисел пар совпадений и несовпадений знаков к их сумме, т.е. к общему числу наблюдаемых единиц:

$$K_{\Phi} = \frac{\sum C - \sum H}{\sum C + \sum H}. \qquad (82)$$

Очевидно, что если знаки всех отклонений по каждому признаку совпадут, то $K_{\Phi}=1$, что характеризует наличие прямой связи. Если все знаки не совпадут, то $K_{\Phi}=-1$ (обратная связь). Если же $\sum C=\sum H$, то $K_{\Phi}=0$. Итак, как и любой показатель тесноты связи, коэффициент Фехнера может принимать значения от 0 до ± 1. Однако, если $K_{\Phi}=1$, то это ни в коей мере нельзя воспринимать как свидетельство функциональной зависимости между *x* и *y*.

В нашей задаче $\bar{x} = \frac{\sum x}{n} = \frac{520}{10} = 52$; $\bar{y} = \frac{\sum y}{n} = \frac{1000}{10} = 100$.

В двух последних столбцах таблицы 10 приведены знаки отклонений каждого *x* и *y* от своей средней величины. Число совпадений знаков – 9, а несовпадений – 1. Отсюда $K_{\Phi}=\frac{9-1}{9+1}=0{,}8$. Обычно такое значение показателя тесноты связи характеризует сильную зависимость, однако, следует иметь в виду, что поскольку K_{Φ} зависит только от знаков и не учитывает величину самих отклонений *x* и *y* от их средних величин, то он практически характеризует не столько тесноту связи, сколько ее наличие и направление.

4. *Линейный коэффициент корреляции* применяется в случае линейной зависимости между двумя количественными признаками x и y. В отличие от $К_Ф$ в линейном коэффициенте корреляции учитываются не только знаки отклонений от средних величин, но и значения самих отклонений, выраженные для сопоставимости в единицах среднего квадратического отклонения t:

$$t_x = \frac{x - \bar{x}}{\sigma_x} \quad \text{и} \quad t_y = \frac{y - \bar{y}}{\sigma_y}.$$

Линейный коэффициент корреляции r представляет собой среднюю величину из произведений нормированных отклонений для x и y:

$$r = \frac{\Sigma\left(\frac{x-\bar{x}}{\sigma x}\right)\left(\frac{y-\bar{y}}{\sigma y}\right)}{n}, \quad (83) \qquad \text{или} \qquad r = \frac{\sum(x-\bar{x})(y-\bar{y})}{n\sigma_x\sigma_y}. \quad (84)$$

Числитель формулы (84), деленный на n, т.е. $\frac{\Sigma(x-\bar{x})(y-\bar{y})}{n} = \overline{(x-\bar{x})(y-\bar{y})}$, представляет собой среднее произведение отклонений значений двух признаков от их средних значений, именуемое *ковариацией*. Поэтому можно сказать, что линейный коэффициент корреляции представляет собой частное от деления ковариации между x и y на произведение их средних квадратических отклонений. Путем несложных математических преобразований можно получить и другие модификации формулы линейного коэффициента корреляции, например:

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x\sigma_y}. \quad (85)$$

Линейный коэффициент корреляции может принимать значения от –1 до +1, причем знак определяется в ходе решения. Например, если $\overline{xy} > \bar{x}\bar{y}$, то r по формуле (85) будет положительным, что характеризует прямую зависимость между x и y, в противном случае (r<0) – обратную связь. Если $\overline{xy} = \bar{x}\bar{y}$, то r=0, что означает отсутствие линейной зависимости между x и y, а при r=1 – функциональная зависимость между x и y. Следовательно, всякое промежуточное значение r от 0 до 1 характеризует степень приближения корреляционной связи между x и y к функциональной. Таким образом, коэффициент корреляции при линейной зависимости служит как мерой тесноты связи, так и показателем, характеризующим степень приближения корреляционной зависимости между x и y к линейной. Поэтому близость значения r к 0 в одних случаях может означать отсутствие связи между x и y, а в других свидетельствовать о том, что зависимость не линейная.

В нашей задаче для расчета r построим вспомогательную таблицу 11.

Таблица 11. Вспомогательные расчеты линейного коэффициента корреляции

i	x_i	y_i	$(x-\bar{x})^2$	$(y-\bar{y})^2$	t_x	t_y	$t_x\, t_y$	$\overline{(x-\bar{x})(y-\bar{y})}$	$\overline{xy}$
1	12	28	1600	5184	-1,36526	-1,10032	1,502223	288	33,6
2	16	40	1296	3600	-1,22873	-0,91693	1,126667	216	64
3	25	38	729	3844	-0,92155	-0,9475	0,873167	167,4	95
4	38	65	196	1225	-0,47784	-0,53488	0,255587	49	247
5	43	80	81	400	-0,30718	-0,30564	0,093889	18	344
6	55	101	9	1	0,102394	0,015282	0,001565	0,3	555,5
7	60	95	64	25	0,273052	-0,07641	-0,02086	-4	570
8	80	125	784	625	0,955681	0,382056	0,365124	70	1000
9	91	183	1521	6889	1,331128	1,268425	1,688436	323,7	1665,3
10	100	245	2304	21025	1,638311	2,215924	3,630373	696	2450
Итого	520	1000	8584	42818			9,516166	1824,4	7024,4

В нашей задаче: $\sigma_x = \sqrt{8584/10} = 29{,}299$; $\sigma_y = \sqrt{42818/10} = 65{,}436$. Тогда по формуле (83) r = 9,516166/10 = 0,9516. Аналогичный результат получаем по формуле (84): r = 1824,4/(29,299*65,436) = 0,9516 или по формуле (85): r = (7024,4 – 52*100) / (29,299*65,436) = 0,9516, то есть связь между величиной основных фондов и валовым выпуском продукции очень близка к функциональной.

Проверка коэффициента корреляции на значимость (существенность). Интерпретируя значение коэффициента корреляции, следует иметь в виду, что он рассчитан для ограниченного числа наблюдений и подвержен случайным колебаниям, как и сами значения x и y, на основе которых он рассчитан. Другими словами, как любой выборочный показатель, он содержит случайную ошибку и не всегда однозначно отражает действительно реальную связь между изучаемыми показателями. Для того, чтобы оценить существенность (значимость) самого r и, соответственно, реальность измеряемой связи между x и y, необходимо рассчитать среднюю квадратическую ошибку коэффициента корреляции σ_r. Оценка существенности (значимости) r основана на сопоставлении значения r с его средней квадратической ошибкой: $\frac{|r|}{\sigma_r}$.

Существуют некоторые особенности расчета σ_r в зависимости от числа наблюдений (объема выборки) – n.

1. Если число наблюдений достаточно велико (n>30), то σ_r рассчитывается по формуле (86):

$$\sigma_r = \frac{1-r^2}{\sqrt{n}}. \qquad (86)$$

Обычно, если $\frac{|r|}{\sigma_r}$>3, то r считается значимым (существенным), а связь – реальной. Задавшись определенной вероятностью, можно определить

доверительные пределы (границы) $r = (r \pm t\sigma_r)$, где t – коэффициент доверия, рассчитываемый по интегралу Лапласа (см. таблицу 4).

2. Если число наблюдений небольшое (n<30), то σ_r рассчитывается по формуле (87):

$$\sigma_r = \frac{\sqrt{1-r^2}}{\sqrt{n-2}}, \qquad (87)$$

а значимость r проверяется на основе t-критерия Стьюдента, для чего определяется расчетное значение критерия по формуле (88) и сопоставляется с $t_{ТАБЛ}$.

$$t_{РАСЧ} = \frac{|r|}{\sigma_r} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}. \qquad (88)$$

Табличное значение $t_{ТАБЛ}$ находится по таблице распределения t-критерия Стьюдента (см. приложение 2) при уровне значимости $\alpha=1-\beta$ и числе степеней свободы $v=n-2$. Если $t_{РАСЧ} > t_{ТАБЛ}$, то r считается значимым, а связь между x и y – реальной. В противном случае ($t_{РАСЧ} < t_{ТАБЛ}$) считается, что связь между x и y отсутствует, и значение r, отличное от нуля, получено случайно.

В нашей задаче число наблюдений небольшое, значит, оценивать существенность (значимость) линейного коэффициента корреляции будем по формулам (87) и (88): $\sigma_r = \frac{\sqrt{1-0{,}9516^2}}{\sqrt{10-2}}$ = 0,3073/2,8284 = 0,1086; $t_{РАСЧ} = \frac{|r|}{\sigma_r}$ = 0,9516/0,1086 = 8,7591. При вероятности 95% $t_{табл}$=2,306, а при вероятности 99% $t_{табл}$=3,355, значит, $t_{РАСЧ} > t_{ТАБЛ}$, что дает возможность считать линейный коэффициент корреляции r = 0,9516 значимым.

5. *Подбор уравнения регрессии* представляет собой математическое описание изменения взаимно коррелируемых величин по эмпирическим (фактическим) данным. Уравнение регрессии должно определить, каким будет среднее значение результативного признака y при том или ином значении факторного признака x, если остальные факторы, влияющие на y и не связанные с x, не учитывать, т.е. абстрагироваться от них. Другими словами, уравнение регрессии можно рассматривать как вероятностную гипотетическую функциональную связь величины результативного признака y со значениями факторного признака x.

Уравнение регрессии можно также назвать *теоретической линией регрессии.* Рассчитанные по уравнению регрессии значения результативного признака называются *теоретическими*. Они обычно обозначаются $\bar{y}_x$ (читается: «игрек, выравненный по *x»)* и рассматриваются как функция от x, т.е. $\bar{y}_x = f(x)$. (Иногда для простоты записи вместо $\bar{y}_x$ пишут $\hat{y}$.)

Найти в каждом конкретном случае тип функции, с помощью которой можно наиболее адекватно отразить ту или иную зависимость между

признаками *x* и *y,* — одна из основных задач регрессионного анализа. Выбор теоретической линии регрессии часто обусловлен формой эмпирической линии регрессии; теоретическая линия как бы сглаживает изломы эмпирической линии регрессии. Кроме того, необходимо учитывать природу изучаемых показателей и специфику их взаимосвязей.

Для аналитической связи между *x* и *y* могут использоваться следующие простые виды уравнений:

$\hat{y} = a_0 + a_1 x$ – прямая линия; $\hat{y} = a_0 + a_1 x + a_2 x^2$ – парабола;

$\hat{y} = a_0 + \frac{a_1}{x}$ – гипербола; $\hat{y} = a_0 a_1^{\,x}$ – показательная функция;

$\hat{y} = a_0 + a_1 \lg x$ – логарифмическая функция и др.

Обычно зависимость, выражаемую уравнением прямой, называют *линейной* (или *прямолинейной),* а все остальные — *криволинейными зависимостями*.

Выбрав тип функции, по эмпирическим данным определяют параметры уравнения. При этом отыскиваемые параметры должны быть такими, при которых рассчитанные по уравнению теоретические значения результативного признака $\bar{y}_x$ были бы максимально близки к эмпирическим данным.

Существует несколько методов нахождения параметров уравнения регрессии. Наиболее часто используется *метод наименьших квадратов* (МНК). Его суть заключается в следующем требовании: искомые теоретические значения результативного признака $\bar{y}_x$ должны быть такими, при которых бы обеспечивалась минимальная сумма квадратов их отклонений от эмпирических значений, т.е.

$$S = \sum (y - \bar{y}_x)^2 \to \min .$$

Поставив данное условие, легко определить, при каких значениях a_0, a_1 и т.д. для каждой аналитической кривой эта сумма квадратов отклонений будет минимальной. Данный метод уже использовался нами в методических указаниях к теме 4 «Ряды динамики», поэтому, воспользуемся формулой (57) для нахождения параметров теоретической линии регрессии в нашей задаче, заменив параметр *t* на *x*.

$$\begin{cases} na_0 + a_1 \sum x = \sum y \\ a_0 \sum x + a_1 \sum x^2 = \sum yx \end{cases} \qquad (89)$$

Исходные данные и все расчеты необходимых сумм представим в таблице 12.

Таблица 12. Вспомогательные расчеты для решения задачи

i	x	y	$x*x$	$y*x$	y'	$(y_i-\bar{y})^2$	$(\bar{y}_x-\bar{y})^2$
1	12	28	144	336	15	5184	7225
2	16	40	256	640	23,5	3600	5852,25
3	25	38	625	950	42,625	3844	3291,891
4	38	65	1444	2470	70,25	1225	885,0625
5	43	80	1849	3440	80,875	400	365,7656
6	55	101	3025	5555	106,375	1	40,64063
7	60	95	3600	5700	117	25	289
8	80	125	6400	10000	159,5	625	3540,25
9	91	183	8281	16653	182,875	6889	6868,266
10	100	245	10000	24500	202	21025	10404
Итого	520	1000	35624	70244	1000	42818	38762,125

$$\begin{cases}10a_0+a_1520=1000\\ a_0520+a_135624=70244\end{cases};\quad \begin{cases}a_0=100-52a_1\\ a_0520+a_135624=70244\end{cases};\quad (100-52a_1)520+35624a_1=70244;$$

$$52000-27040a_1+35624a_1=70244;\quad 8584a_1=18244;\quad a_1=2{,}125;\quad a_0=100-52*2{,}125=-10{,}5.$$

Отсюда искомая линия регрессии: $\hat{y}_t=-10{,}5+2{,}125x$. Для иллюстрации построим график эмпирической (маркеры-кружочки) и теоретической (маркеры-квадратики) линий регрессии.

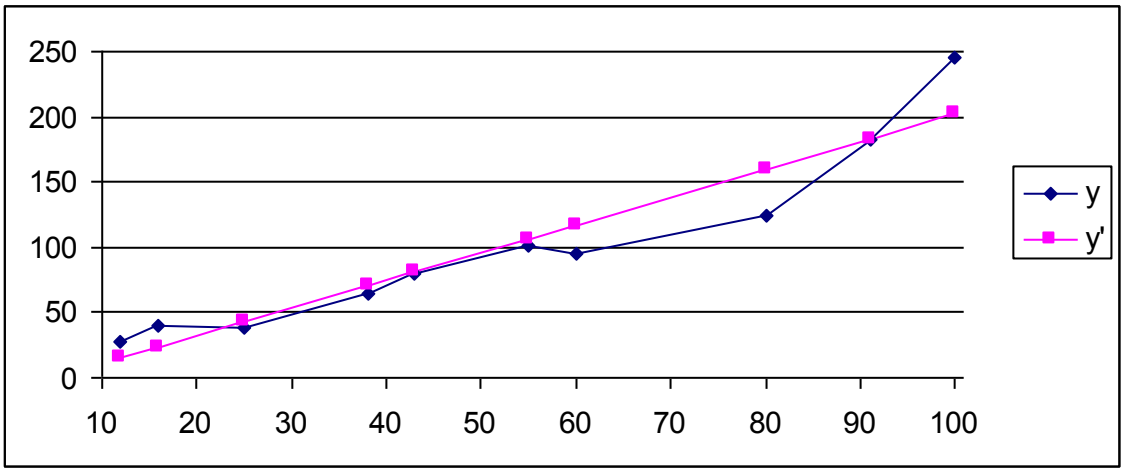


Рис.6. График эмпирической и теоретической линий регрессии.

6. *Теоретическое корреляционное отношение* представляет собой универсальный показатель тесноты связи. Измерить тесноту связи между коррелируемыми величинами – это значит определить, насколько вариация результативного признака обусловлена вариацией факторного признака. Ранее были рассмотрены показатели, с помощью которых можно выявить наличие корреляционной связи между двумя признаками x и y и измерить тесноту этой связи: коэффициент Фехнера и линейный коэффициент корреляции.

Наряду с ними существует универсальный показатель – *корреляционное отношение* (или *коэффициент корреляции по Пирсону*), применимое ко всем случаям корреляционной зависимости независимо от формы этой связи. Следует различать эмпирическое и теоретическое корреляционные отношения. *Эмпирическое корреляционное отношение* рассчитывается на основе правила сложения дисперсий как корень квадратный из отношения межгрупповой дисперсии к общей дисперсии, т.е.

$$\eta_{эмп} = \sqrt{\frac{\delta^2}{\sigma^2}}\,. \qquad (90)$$

Теоретическое корреляционное отношение $\eta_{теор}$ определяется на основе выравненных (теоретических) значений результативного признака $\bar{y}_x$, рассчитанных по уравнению регрессии. $\eta_{теор}$ представляет собой относительную величину, получаемую в результате сравнения среднего квадратического отклонения в ряду теоретических значений результативного признака со средним квадратическим отклонением в ряду эмпирических значений. Если обозначить дисперсию эмпирического ряда игреков через σ_y^2, а теоретического ряда – δ^2, то каждая из них выразится формулами:

$$\sigma_y^2 = \frac{\sum(y_i - \bar{y})^2}{n}, \qquad (91)$$

$$\delta^2 = \frac{\sum(\bar{y}_x - \bar{y})^2}{n}. \qquad (92)$$

Сравнивая вторую дисперсию с первой, получим *теоретический коэффициент детерминации:*

$$\eta^2_{теор} = \frac{\delta^2}{\sigma_y^2} = \frac{\sum(\bar{y}_x - \bar{y})^2}{\sum(y_i - \bar{y})^2}, \qquad (93)$$

который показывает, какую долю в общей дисперсии результативного признака занимает дисперсия, выражающая влияние вариации фактора *x* на вариацию *y*. Извлекая корень квадратный из коэффициента детерминации, получаем *теоретическое корреляционное отношение:*

$$\eta_{теор} = \sqrt{\frac{\delta^2}{\sigma_y^2}} = \sqrt{\frac{\sum(\bar{y}_x - \bar{y})^2}{\sum(y_i - \bar{y})^2}}\,. \qquad (94)$$

Оно может находиться в пределах от 0 до 1. Чем ближе его значение к 1, тем теснее связь между вариацией *y* и *x*. При η<0,3 говорят о малой

зависимости между коррелируемыми величинами, при 0,3<η<0,6 – о средней, при 0,6<η<0,8 – о зависимости выше средней, при η>0,8 – о большой, сильной зависимости. Корреляционное отношение применимо как для парной, так и для множественной корреляции независимо от формы связи. При линейной зависимости $\eta_{теор} \equiv r$.

В нашей задаче расчет необходимых сумм для использования в формуле (93) приведен в последних двух столбцах таблицы 12. Тогда теоретический коэффициент детерминации по формуле (93) равен: $\eta^2_{теор}$ = 38762,125 / 42818 = 0,9053, то есть дисперсия, выражающая влияние вариации фактора *x* на вариацию *y*, составляет 90,53%.

Теоретическое корреляционное отношение по формуле (94) равно: $\eta_{теор} = \sqrt{0{,}9053}$ = 0,9515, что совпадает со значением линейного коэффициента корреляции и, следовательно, можно говорить о большой, сильной зависимости между коррелируемыми величинами.

Контрольные задания по теме

На основе исходных данных контрольных заданий по теме 2 определить наличие и характер корреляционной связи между признаками *x* и *y* 6-ю методами.

При-знак	**Вариант**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
x	Рост	Доход	Возраст	IQ	Доход	Возраст	рост/вес	Стаж	Доход	IQ
y	Вес	Вес	Доход	Доход	Тетрадь	рост/вес	Кол-во друзей	Доход	Кол-во друзей	Время решения

Часть 2. Социально-экономическая статистика

Тема 1. Социально-демографическая статистика

Методические указания по теме

Задача 1. Численность населения города составляла 3000 тыс. чел. на начало года. На конец года она возросла до 3050 тыс. чел. Число родившихся за год составило 35 тыс.чел., число умерших – 15 тыс.чел. Определить:

1) коэффициенты естественного, механического и общего движения населения, установить его тип;

2) перспективную численность населения через 4 года при условии, что коэффициент общего движения населения будет: а) сохраняться на прежнем уровне; б) снижаться ежегодно на 1‰.

Решение. Численность населения (S) в конкретном пункте существенно меняется с течением времени, поэтому, рассчитывается ее *среднегодовое значение* по разным формулам средних величин в зависимости от полноты исходных данных. В нашей задаче среднегодовая численность определяется по формуле средней арифметической простой (17) как полусумма S в начале и конце года: $\bar{S} = (S_н + S_к)/2$ = (3000 + 3050)/2 = 3025 (тыс. чел.).

Изменение численности населения за счет рождений и смертей называется *естественным движением* населения, которое характеризуется рядом коэффициентов: *рождаемости* (95), *смертности* (96) и *естественного движения* (97):

$$К_р = N / \bar{S} * 1000‰, \quad (95)$$

$$К_{см} = M / \bar{S} * 1000‰, \quad (96)$$

$$К_{ЕД} = (N – M)/\bar{S} * 1000‰ = К_р – К_{см}, \quad (97)$$

где N – число родившихся за год, M – число умерших за год.

В нашей задаче по формуле (95) получаем: $К_р$ = 35 / 3025 * 1000‰ = 11,57‰, то есть на каждую 1000 населения приходится 11 младенцев.

По формуле (96) получаем: $К_{см}$ = 15 / 3025*1000‰ = 4,96‰, то есть на каждую 1000 населения приходится 5 умерших.

И, наконец, по формуле (97) получаем: $К_{ЕД} = К_р – К_{см}$,= 11,57‰ – 4,96‰ = 6,61‰, то есть рождаемость превышает смертность на 6,61 промилле (это естественный прирост населения).

Движением населения может происходить и за счет *миграции*, показывающей, куда и откуда, в каком количестве происходит перемещение населения в стране и в международном масштабе. Основными показателями

миграции населения являются: *сальдо миграции* (98) и *коэффициент механического движения населения* (99):

$$\Delta V = V^{+} - V^{-}, \qquad (98)$$

$$K_{МД} = \Delta V / \bar{S} * 1000‰, \qquad (99)$$

где V^{+} и V^{-} – численность, соответственно, прибывшего и выбывшего на постоянное жительство населения.

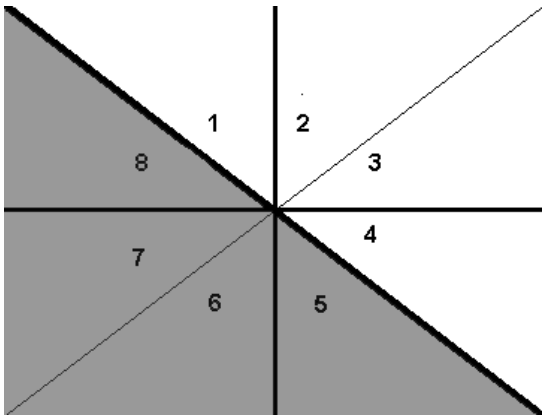
В нашей задаче число прибывших и выбывших неизвестно, но можно найти сальдо миграции, так как данные о миграционном сальдо вместе с данными об естественном движении населения служат основой для расчета коэффициента *общего движения* населения (100):

$$K_{ОД} = K_{ЕД} + K_{МД} = (S_к - S_н)/\bar{S} * 1000‰. \qquad (100)$$

В нашей задаче по формуле (100) получаем: $K_{ОД}$ =(3050 – 3000) / 3025 * 1000‰ = 16,53‰, тогда $K_{МД} = K_{ОД} - K_{ЕД}$ = 16,53‰ – 6,61‰ = 9,92‰, то есть численность прибывших в город больше выбывших на 9,92 промилле (это механический приток или прирост населения).

В социально-демографической статистике различают 8 типов общего движения населения:

1. Естественный прирост превышает механический отток;
2. Естественный прирост превышает механический приток;
3. Механический приток больше естественного прироста;
4. Механический приток больше естественной убыли;
5. Естественная убыль превышает механический приток;
6. Естественная убыль превышает механический отток;
7. Механический отток больше естественной убыли;
8. Механический отток больше естественного прироста.



Их можно изобразить графически (см. рис. 7), если по оси абсцисс отложить $K_{МД}$, а по оси ординат – $K_{ЕД}$.

Типы 1 – 4 (выше жирной диагонали) свидетельствуют о росте населения территории, а типы 5 – 8 (серый цвет) о сокращении ее населения.

В нашей задаче механический приток больше естественного прироста ($|K_{МД}| > |K_{ЕД}|$), значит это 3-й тип общего движения.

Рис. 7. Типы движения населения.

Перспективную численность населения через *t* лет можно определить по формуле (101):

$$S_{н+t} = S_н\left(1+\frac{K_{ОД}}{1000‰_0}\right)^t. \qquad (101)$$

В нашей задаче, считая, что $K_{ОД}$ сохранится на прежнем уровне, определим S через 4 года по формуле (101):

S_4 = 3050*(1+16,53/1000)^4 = 3256,27 (тыс. чел.).

Считая, что $K_{ОД}$ будет ежегодно снижаться на 1‰:

S_1 = 3000*(1+16,53/1000) = 3050 (тыс. чел.);

S_2 = 3050*(1+15,53/1000) = 3097,363 (тыс. чел.);

S_3 = 3097,363*(1+14,53/1000) = 3142,364 (тыс. чел.);

S_4 = 3142,364*(1+13,53/1000) = 3184,877 (тыс. чел.).

Задача 2. Определить среднюю продолжительность предстоящей жизни по следующим данным о числе умирающих из 100 000 человек при переходе от возраста x к x+1 лет (d_x):

Возраст x, лет	0-1	1-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44
d_x	1802	316	184	183	550	749	681	766	1061	1583

Возраст x, лет	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85 и более
d_x	2454	3631	5341	7171	9480	11562	14192	14751	23543

Решение. Исходные данные – это элементы укрупненной в интервалы таблицы смертности. Эти данные показывают, как постепенно с увеличением возраста население выбывает под влиянием смертности. На их основе рассчитываются частные коэффициенты смертности по отдельным возрастным группам. Для определения средней продолжительности предстоящей жизни необходимо рассчитать следующие показатели для каждого возраста x:

1) l_x – число доживающих до возраста x, которое показывает, сколько из 100 000 лиц условного поколения доживают до каждого следующего года (102):

$$l_{x+1} = l_x - d_x; \qquad (102)$$

2) d_x – число умирающих возрасте x;

3) p_x – вероятность дожить до возраста x+1, которая показывает, какая доля из достигших данного возраста x проживает еще полный год и достигнет очередного возраста x+1 (103):

$$p_x = \frac{l_{x+1}}{l_x}; \qquad (103)$$

4) q_x – вероятность умереть в возрасте x (104) (при этом должно выполняться условие: $p_x+q_x=1$):

$$q_x = \frac{d_x}{l_x}; \qquad (104)$$

5) L_x – среднее число живущих в возрастном интервале от x до x+1 года (105):

$$L_x = \frac{l_x + l_{x+1}}{2}; \qquad (105)$$

6) T_x – число человеко-лет, которые проживет население, достигшее x лет до максимального возраста w (106):

$$T_x = \sum_{x}^{w} L_x; \qquad (106)$$

7) E_x – средняя продолжительность предстоящей жизни, которая показывает, сколько лет в среднем предстоит прожить лицам, достигшим возраста x (107):

$$E_x = \frac{T_x}{l_x}; \qquad (107)$$

Показатель (107) является важнейшим обобщающим демографическим показателем, зависящим от социальной ситуации в стране.

В нашей задаче расчет перечисленных показателей приведен в таблице 13. Поскольку возраст задан интервально, то во втором столбце этой таблицы рассчитаны середины интервалов $X_{сер}$ как полусумма нижней и верхней границы интервала, а формула (106) примет следующий вид:

$$T_x = (87{,}5 - X_{сер})*L_x.$$

Таблица 13. Таблица смертности с укрупненными возрастными интервалами в 10 лет

Возраст X, лет	$X_{сер}$	l_x	d_x	p_x	q_x	L_x	T_x	E_x
0-1	0,5	100000	1802	0,982	0,018	99099,0	8621613,0	86,216
1-4	3	98198	316	0,997	0,003	98040,0	8284380,0	84,364
5-9	7,5	97882	184	0,998	0,002	97790,0	7823200,0	79,925
10-14	12,5	97698	183	0,998	0,002	97606,5	7320487,5	74,930
15-19	17,5	97515	550	0,994	0,006	97240,0	6806800,0	69,803
20-24	22,5	96965	749	0,992	0,008	96590,5	6278382,5	64,749
25-29	27,5	96216	681	0,993	0,007	95875,5	5752530,0	59,788
30-34	32,5	95535	766	0,992	0,008	95152,0	5233360,0	54,780
35-39	37,5	94769	1061	0,989	0,011	94238,5	4711925,0	49,720
40-44	42,5	93708	1583	0,983	0,017	92916,5	4181242,5	44,620
45-49	47,5	92125	2454	0,973	0,027	90898,0	3635920,0	39,467
50-54	52,5	89671	3631	0,960	0,040	87855,5	3074942,5	34,291

Возраст X, лет	$X_{сер}$	l_x	d_x	p_x	q_x	L_x	T_x	E_x
55-59	57,5	86040	5341	0,938	0,062	83369,5	2501085,0	29,069
60-64	62,5	80699	7171	0,911	0,089	77113,5	1927837,5	23,889
65-69	67,5	73528	9480	0,871	0,129	68788,0	1375760,0	18,711
70-74	72,5	64048	11562	0,819	0,181	58267,0	874005,0	13,646
75-79	77,5	52486	14192	0,730	0,270	45390,0	453900,0	8,648
80-84	82,5	38294	14751	0,615	0,385	30918,5	154592,5	4,037
85 и более	87,5	23543	23543	0,000	1,000	11771,5	0,0	0,000
Итого		1568920	100000			1518920	79011963	50,361

Таким образом, в итоговой строке таблицы 13 в последнем столбце по формуле (107) определена средняя продолжительность предстоящей жизни в 50,361 года.

Контрольные задания по теме

Задание 1. Имеются следующие условные данные о численности населения города, тыс. чел.:

Вариант	**1**	**2**	**3**	**4**	**5**
Численность в начале года	320	750	105	400	1500
Численность в конце года	315	700	109	440	1475
Число родившихся за год	8,5	15	5,5	15	25
Число умерших за год	12,0	11	6,0	10	37

Определить:

1) коэффициенты естественного, механического и общего движения населения, установить его тип;
2) перспективную численность населения через 5 лет при условии, что коэффициент общего движения населения будет:
   а) сохраняться на прежнем уровне;
   б) ежегодно расти на 1‰.

Задание 2. Определить среднюю продолжительность предстоящей жизни по следующим данным о числе умирающих из 100 000 человек при переходе от возраста x к x+1 лет:

Возраст X, лет	**Вариант**				
	6	**7**	**8**	**9**	**10**
0-10	2300	2170	2139	1856	4020
10-20	735	865	831	395	1980
20-30	1420	1420	1323	1176	5601
30-40	1837	1967	1934	1693	6810
40-50	4017	3887	3919	3973	9820
50-60	8992	9122	9090	9136	11968
60-70	16650	16520	16549	16894	21856
70-80	25755	25885	25856	26099	33295
80-90	38294	38164	38359	38778	4650

Тема 2. Статистика уровня жизни населения

Методические указания по теме

Задача 1. По приведенным в следующей таблице данным (первые три столбца) о распределении населения РФ по ежемесячному среднедушевому доходу (СДД) в 2004 году рассчитать показатели дифференциации доходов (численность населения России в 2004 году составила 144,2 млн. чел.).

№ групп i	Месячный СДД, руб./чел.	Доля населения d_i	Численность населения, млн. чел.	Доход, млн. руб.	Доля доходов q_i	Кумулятивные доли	
						населения d'_i	дохода q'_i
1	до 1000	0,019	2,7398	2054,85	0,00284829	0,019	0,00284829
2	1000-1500	0,043	6,2006	7750,75	0,01074355	0,062	0,01359184
3	1500-2000	0,062	8,9404	15645,7	0,02168699	0,124	0,03527883
4	2000-3000	0,146	21,0532	52633	0,07295623	0,27	0,10823506
5	3000-4000	0,139	20,0438	70153,3	0,09724166	0,409	0,20547671
6	4000-5000	0,118	17,0156	76570,2	0,10613632	0,527	0,31161303
7	5000-7000	0,17	24,514	147084	0,20387767	0,697	0,51549071
8	свыше 7000	0,303	43,6926	349540,8	0,48450929	1	1
	Итого	1	144,2	721432,6	1		

Решение. Сначала определяем абсолютные величины дифференциации. Так, больше всего людей (их доля – 0,303) имели доход свыше 7000 руб./чел. В этом интервале и находится модальный доход, точное значение которого согласно формуле (13)[2] определяется следующим образом:

$$Mo = 7000 + 2000\frac{0,303 - 0,17}{2 * 0,303 - 0,17 - 0} = 7610 \text{ руб./чел.}$$

Доход в интервале 4000-5000 руб./чел. является граничным для половины людей, поэтому согласно формуле (14) значение медианного дохода равно:

$$Me = 4000 + 1000\frac{0,5 - 0,409}{0,118} = 4771,19 \text{ руб./чел.}$$

Затем рассчитываем простейшие относительные величины дифференциации – *децильный* (108) и *фондовый* (109) коэффициенты. Децильный (дециль составляет 10%) коэффициент – это отношение минимального СДД 10% самого богатого населения ($\min СДД_{10\%бог}$) к максимальному СДД 10% самого бедного населения ($\max СДД_{10\%бед}$).

[2] Формулы (13) и (14) должны применяться к распределениям с одинаковым размахом интервалов (СДД), в нашей задаче это условие не соблюдается, а значит, и результаты получены приблизительные

Коэффициент фондов – это отношение среднего СДД 10% самого богатого населения к среднему же СДД 10% самого бедного населения.

$$K_{ДЦ} = \frac{\min СДД_{10\%бог}}{\max СДД_{10\%бед}}, \qquad (108) \qquad\qquad K_{Ф} = \frac{\overline{СДД}_{10\%бог}}{\overline{СДД}_{10\%бед}}. \qquad (109)$$

По исходным данным необходимо отобрать 10% самых бедных людей, т.е. первые три группы (их кумулятивная доля равна 0,124, что ближе всего к необходимым 0,1). Так как первый интервал СДД является открытым, следовательно, представляем его в закрытом виде, используя размах соседнего интервала в размере 500 руб./чел. (т.е. границы 1-й группы составят от 500 до 1000 руб./чел.). Тогда первые три группы самых бедных (12,4%) предстанут в границах 500-2000 с серединой 1250 руб./чел. Если 12,4% бедных имеют размах доходов 1500 руб./чел., то 10% будут иметь размах доходов: 10%*1500/12,4%=1209,68 (руб./чел.). Значит max$СДД_{10\%бед}$ = 500 + 1209,68 = 1709,68 (руб./чел.), а $\overline{СДД}_{10\%бед}$ = 500 + 1209,68 / 2 = 1104,84 (руб./чел.).

Теперь отберем 10 % самых богатых людей – это 8-я группа с доходами от 7000 до 9000 руб./чел. (так как интервал открытый, то применили размах соседнего интервала в размере 2000 руб./чел.), т.е. 30,3% самого богатого населения имеет размах доходов 2000 руб./чел.[3]. Нам нужно отобрать не 30,3%, а 10%, поэтому, решая пропорцию, находим размах доходов 10% самого богатого населения. Он равен 660,07 руб./чел. Отсюда min$СДД_{10\%бог}$ = 9000 - 660,07 = 8339,93 руб./чел., а его среднее значение $\overline{СДД}_{10\%бог}$ = 9000 - 660,07/2 = 8669,97 (руб./чел.).

Таким образом, по формуле (108) децильный коэффициент $K_{ДЦ}$ = 8339,93/1709,68=4,88, а по формуле (109) коэффициент фондов $K_{Ф}$ = 8669,97/1104,84=7,85.

Для расчета более сложных относительных величин дифференциации определим доход и его долю в каждой группе людей, используя середины интервалов СДД и количество людей в группах. Так, доход первой группы составит: 750 руб./чел. * 2,7398 млн. чел. = 2054,85 млн.руб., а его доля равняется 2054,85/721432,6=0,00284829. Аналогично, например, для четвертой группы: 2500*21,0532 = 52633 млн. руб. и 52633/721432,6=0,07295623. Естественно, доли доходов надо определять после суммирования доходов по группам (получается 721432,6 млн. руб.).

Полученные доли людей и доходов вписываются в таблицу, после чего определяются соответствующие кумулятивные доли (нарастающим итогом). Например, кумулятивная доля людей 3-й группы составит 0,019+0,043+0,062=0,1240, а кумулятивная доля их доходов – соответственно 0,00284829+0,01074355+0,02168699=0,03527883. Сумма долей как в обычном, так и в кумулятивном виде должна равняться 1.

---

[3] Очевидно, что данное допущение о максимуме СДД в 9000 руб./чел. является чисто теоретическим, на самом же деле, есть люди, получающие гораздо больше, поэтому полученные показатели являются лишь приблизительными

Кумулятивные доли также вписываются в таблицу, после чего можно определять коэффициенты *локализации* (определяется по формуле Лоренца (110)) и *концентрации* (определяется по формуле Джини (111)) доходов:

$$K_Л = 0{,}5\sum|d_i - q_i|; \qquad (110) \qquad\qquad K_Д = \sum d'_i q'_{i+1} - \sum q'_i d'_{i+1}. \qquad (111)$$

Значения коэффициентов Лоренца и Джини изменяются от 0 до 1. Нулевое их значение свидетельствует об абсолютной равномерности распределения доходов по группам населения. Чем ближе эти коэффициенты к единице, тем в большей мере доходы сосредоточены в отдельной группе населения. Естественно, при этом часть населения оказывается живущей в бедности.

Так, по формуле (110) коэффициент локализации Лоренца равняется:

$$Кл = 0{,}5 * (|0{,}19–0{,}002848| + |0{,}043–0{,}010744| + |0{,}062–0{,}021687| + |0{,}146–0{,}072956| + |0{,}139–0{,}09242| + |0{,}118–0{,}10614| + |0{,}17–0{,}20388| + |0{,}303–0{,}4845|) = 0{,}215.$$

Для наглядности неравномерность распределения доходов изобразим графически в виде кривой Лоренца (рис.8).

По формуле (111) коэффициент концентрации Джини равняется:

$$K_Д = 0{,}019*0{,}013592 + 0{,}062*0{,}03528 + 0{,}124*0{,}108235 + 0{,}27*0{,}2055 + 0{,}409*0{,}3116 + 0{,}527*0{,}5155 + 0{,}697*1 – 0{,}00285*0{,}062 – 0{,}0136*0{,}124 – 0{,}0353*0{,}27 – 0{,}108234*0{,}409 – 0{,}2055*0{,}527 – 0{,}3116*0{,}697 – 0{,}51549*1 = 1{,}168 – 0{,}897 = 0{,}271.$$

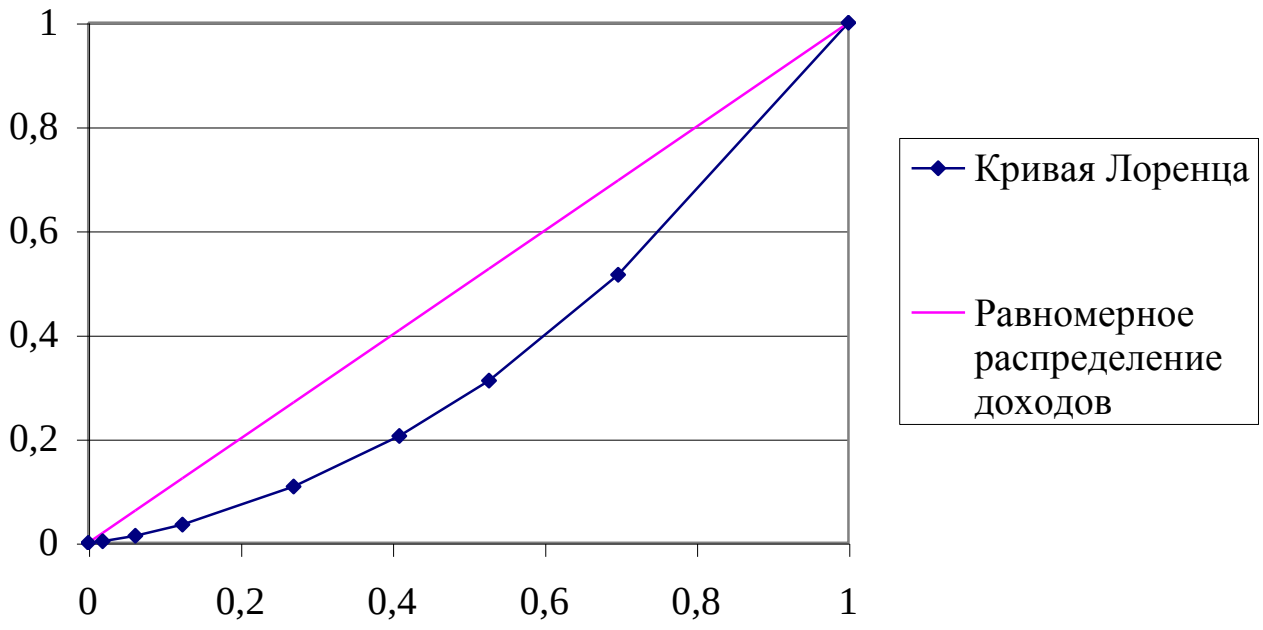


Рис.8. Кривая распределения доходов.

Таким образом, коэффициенты Лоренца и Джини показали, что 0,215–0,271 доходов населения или 21,5–27,1% сосредоточено в руках 10% самых богатых людей, что говорит о неравномерности распределения доходов в России.

Задача 2. Рассчитать индекс развития человеческого потенциала на 2006 год по следующим данным:

Ожидаемая продолжительность жизни, лет	63
Доля грамотных / учащихся	0,82 / 0,65
Паритет покупательной способности валют	1,1
Среднегодовой индекс инфляции	1,03
Среднедушевой ВВП в мес., $/чел	1200

Решение. В качестве обобщающего критерия уровня жизни используется разработанный Программой развития ООН в 1993 г. индекс развития человеческого потенциала (ИРЧП), который базируется на расчете трех индексов и представляет собой простую среднюю арифметическую величину:

$$ИРЧП = \frac{I_{ОБР} + I_{ОЖ} + I_{ВВП}}{3}. \qquad (112)$$

где $I_{ОБР}$ – международный индекс уровня образования, определяемый по формуле (113); $I_{ОЖ}$ – индекс ожидаемой при рождении продолжительности жизни, определяемый по формуле (114); $I_{ВВП}$ – индекс валового внутреннего продукта (ВВП), определяемый по формуле (116).

$$I_{ОБР} = 2/3d_г + 1/3d_у, \qquad (113)$$

где $d_г$ – доля грамотных, $d_у$ – доля учащихся.

$$I_{ОЖ} = \frac{X_0 - X_m}{X_M - X_m}, \qquad (114)$$

где X_m, X_M —минимально и максимально возможная продолжительность жизни, X_0 – ожидаемая при рождении продолжительность жизни, определяемая по формуле (115).

По мировым стандартам X_m = 25 лет, а X_M = 85 лет, значит, для международных сопоставлений надо принимать:

$$X_0 = 85(1- K'_{мс}), \qquad (115)$$

где $K'_{мс}$ – коэффициент младенческой смертности, выраженный в долях единицы.

$$I_{ВВП} = \frac{\overline{ВВП} * ППСВ - ВВП_m}{ВВП_M - ВВП_m}, \qquad (116)$$

где $\overline{ВПП}$ – фактический в стране среднедушевой валовой внутренний продукт; $ВВП_m$ и $ВВП_M$ – минимальный и максимальный размеры среднедушевого ВВП по мировым стандартам; *ППСВ* – паритет покупательной способности валют.

В качестве минимального размера ВВП принято $100 на человека в месяц, а максимальным размером для разумно высокого благосостояния в 1992 г. считалось $5120 на человека в месяц. Максимальный размер на последующие годы корректируется с учетом среднегодового индекса инфляции по формуле (117):

$$ВВП_М = 5120 * i_ф^t, \qquad (117)$$

где t – количество лет с 1992 до расчетного года, а индекс инфляции $i_ф$ можно принять по предыдущему перед расчетным годом.

В нашей задаче индекс образования по формуле (113):

$I_{ОБР}$ = 2/3*0,82 + 1/3*0,65 = 0,763.

Находим индекс ожидаемой при рождении продолжительности жизни по формуле (114): $I_{ОЖ}$ = (63 – 25)/(85 – 25) = 0,633.

Определяем максимальный ВВП по формуле (117):

$ВВП_М = 5120*1{,}03^{14} = 7744{,}459$.

Индекс валового внутреннего продукта находим по формуле (116):

$I_{ВВП}$ = (1200*1,1–100)/(7744,459–100) = 0,160.

ИРЧП по определяем по формуле (112):

ИРЧП = (0,763 + 0,633 + 0,160) / 3 = 0,519.

Контрольные задания по теме

Задание 1. Определить показатели дифференциации доходов населения России по следующим данным.

№ п/п	СДД, руб./чел.	Доли населения, %			
		Вариант			
		1	**2**	**3**	**4**
1	до 1000	20,4	12,5	6,8	3,3
2	1000-1500	19,9	15,0	10,6	6,6
3	1500-2000	16,4	14,4	11,8	8,5
4	2000-3000	20,7	21,7	21,0	17,7
5	3000-4000	10,4	13,4	15,2	15,1
6	4000-5000	5,3	8,2	10,4	11,7
7	5000-7000	4,4	8,2	11,9	15,4
8	более 7000	2,5	6,6	12,3	21,7
	Число жителей, млн.чел. (год)	146,9 (2000)	146,3 (2001)	145,6 (2002)	145,0 (2003)

СДД, руб./чел.	Доли населения, %			
	Вариант			
	6	**7**	**8**	**9**
до 1500	17,3	9,9	6,2	3,2
1500-2500	23	17,5	13,2	8,9
2500-3500	18,1	16,7	14,4	11,5
3500-4500	12,6	13,4	12,8	11,5
4500-6000	11,8	14,3	15	15
6000-8000	8,2	11,4	13,4	14,9
8000-12000	6,1	10,2	13,7	17,3
более 12000	2,9	6,6	11,3	17,7
Число жителей, млн.чел. (год)	145,6 (2002)	145,0 (2003)	144,2 (2004)	143,5 (2005)

Задание 2. Рассчитать индекс развития человеческого потенциала на 2006 год:

Исходные данные	**Вариант 5**	**Вариант 10**
Ожидаемая продолжительность жизни, лет	61	72
Доля грамотных/учащихся	0,78 / 0,45	0,98 / 0,56
Паритет покупательной способности валют	1,5	0,84
Среднегодовой индекс инфляции	1,05	1,02
Среднедушевой ВВП в мес., $/чел	1500	2500

Тема 3. Статистика национального богатства

Методические указания по теме

Задача 1. Имеются следующие данные о динамике балансовой стоимости *основных фондов (Ф)* предприятия:

Дата	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.11	1.12	31.12
Ф, млн. руб.	130	128	120	125	135	124	118	115	119	122	128	125	122

Износ фондов в начале года – 26 млн. руб., норматив отчислений на реновацию (амортизация) - 5%, ликвидационная стоимость - 15% от стоимости выбывших фондов, годовой объем выпущенной продукции - 300 млн. руб., среднесписочная численность персонала - 1000 чел.

1. Определить среднегодовую балансовую стоимость основных фондов.
2. Составить балансы основных фондов по первоначальной полной и остаточной стоимостям.
3. Рассчитать показатели состояния, движения и использования фондов.

Решение. Среднегодовая балансовая стоимость определяется по формуле средней хронологической простой (если временные интервалы равны) или взвешенной (если временные интервалы неравны). В нашей задаче интервалы равные (по 1 месяцу), значит, используем формулу (27):

$$\overline{\Phi} = \frac{\frac{130+122}{2}+128+120+125+135+124+118+115+119+122+128+125}{13-1} = 123{,}75 \text{ (млн.руб.)}$$

Балансы показывают динамику фондов за год. Они строятся по полной *первоначальной* стоимости и по *остаточной* стоимости. Уравнение баланса по полной первоначальной стоимости имеет вид (118):

$$\Phi_к = \Phi_н + П - В, \qquad (118)$$

где $\Phi_к$ и $\Phi_н$ – стоимость фондов на конец и начало года, соответственно; *П* и *В* – стоимость поступивших и выбывших, соответственно, фондов за год.

В нашей задаче по формуле (118) определяем стоимость *Ф* на конец года:

$\Phi_к$ = 130 + (5+10+4+3+6) – (2+8+11+6+3+3+3) =130 + 28 – 36 = 122 (млн. руб.).

В таблице 14 построим баланс основных фондов по полной первоначальной стоимости.

Таблица 14. Баланс основных фондов по полной первоначальной стоимости

Виды основ-ных фондов	Наличие на начало года	Поступило в отчетном году			Выбыло в отчетном году			Наличие на конец года
		Всего	В том числе		Всего	В том числе		
			ввод в действие новых фондов	прочие поступ-ления		ликвидировано основных фондов	прочее выбытие	
Ф	130	28	28	0	36	36	0	122

Схема баланса по полной первоначальной стоимости во многом совпадает с балансом основных фондов по остаточной стоимости (табл. 15). Отличие заключается в том, что в таком балансе помимо учета поступления и выбытия объектов по остаточной (за вычетом износа) стоимости учитывается уменьшение их стоимости за год вследствие износа (*А*), равное сумме начисленной амортизации за год. В основе баланса основных фондов по остаточной стоимости лежит уравнение (119):

$$Ф'_к = Ф'_н + П' - В' - А, \qquad (119)$$

где ' – знак остаточной стоимости.

Остаточная стоимость выбывших фондов за год (*В'*) включает стоимость: проданных (по рыночной стоимости), переданных безвозмездно (по остаточной стоимости) другим предприятиям и ликвидированных (по ликвидационной стоимости) из-за ветхости и износа основных фондов.

В нашей задаче, считая, что продажи и безвозмездной передачи не было, имеем: $B' =$ 0,15*36 = 5,4 (млн. руб.). Тогда, считая все поступившие фонды новыми, по формуле (119) имеем: $Ф'_к$ = (130-26) + 28 – 5,4 – 0,05*130 = 104 + 28 – 5,4 – 6,5 =120,1 (млн. руб.).

Проверка: $Ф_к' = Ф_к - И_к$, где $И_к = И_н + А + В' - В$, поэтому, в нашей задаче $Ф_к'$ = 122 - (26+6,5+5,4-36) = 120,1 (млн. руб.). В таблице 15 построим баланс основных фондов по остаточной стоимости.

Таблица 15. Баланс основных фондов по остаточной стоимости

Виды основ-ных фондов	Наличие на начало года	Поступило в отчетном году			Выбытие и износ за год				Наличие на конец года
		Всего	В том числе		Всего	В том числе			
			ввод в действие новых фондов	прочие поступления		износ основных фондов за год	ликвиди-ровано основных фондов	прочее выбытие	
Ф	104	28	28	0	11,9	6,5	5,4	0	120,1

Используя сведения о наличии фондов по полной и остаточной стоимости, находят их следующие обобщающие показатели:

1) состояния (коэффициенты износа и годности);

2) интенсивности движения (коэффициенты поступления, выбытия, движения и обновления);

3) использования (фондоотдача, фондоемкость и фондовооруженность).

Коэффициент износа определяется на определенную дату (на начало или конец года) как отношение суммы износа основных фондов (*И*) к их полной стоимости (*Ф*), то есть по формуле (120):

$$К_{изн} = И/Ф. \qquad (120)$$

В нашей задаче полная стоимость *Ф* на начало года составляет 130 млн. руб., а износ - 26 млн. руб., следовательно, $К_{изн(н)}$ = 26/130 = 0,20, то есть в начале года 20% фондов были изношенными.

Сумму износа на конец года ($И_к$) можно получить как разность между полной ($Ф_к$) и остаточной стоимостью на эту дату ($Ф'_к$) фондов, то есть по формуле (121):

$$И_к = Ф_к - Ф'_к. \qquad (121)$$

В нашей задаче по формулам (120) и (121) получаем: $К_{изн(к)}$ = (122–120,1) / 122 = 0,016, то есть коэффициент износа уменьшился с 20% в начале года до 1,6% в конце.

Разность между единицей и коэффициентом износа дает величину *коэффициента годности*, отражающего долю неизношенной части основных фондов. Его можно также рассчитать по формуле (122):

$$К_{годн} = Ф'/Ф. \qquad (122)$$

В нашей задаче $К_{годн(н)}$ = 1 – 0,2 = 0,8; $К_{годн(к)}$ = 1 – 0,016 = 0,984 или по формуле (122): $К_{годн(н)}$ = (130-26) / 130 = 0,8; $К_{годн(к)}$ = 120,1 / 122 = 0,984, то есть степень неизношенности фондов увеличилась с 80% в начале года до 98,4% в конце.

К показателям движения основных фондов относят *коэффициенты поступления* (123) и *выбытия* (124):

$$К_п = П / Ф_к; \qquad (123) \qquad\qquad К_в = В / Ф_н. \qquad (124)$$

В нашей задаче $К_п$ = 28 / 122 = 0,230, то есть за год поступило 23% от стоимости всех фондов.

$К_в$ = 36 / 130 = 0,277, то есть за год выбыло 27,7% от стоимости фондов.

Дополнительно можно определить *коэффициенты движения* (125) и *обновления* (126):

$$К_д = (П - В)/ \overline{Ф}; \qquad (125) \qquad\qquad К_о = П / В. \qquad (126)$$

В нашей задаче $К_д$ = (28–36)/123,75 = – 0,065, то есть поступило фондов на 6,5% меньше, чем выбыло.

K_o = 28 / 36 = 0,778, то есть поступило за год 77,8% от стоимости выбывших фондов.

К показателям использования основных фондов относят *фондоотдачу* (127), *фондоемкость* (128) и *фондовооруженность* (129):

$$H = Q / \overline{\Phi}; \quad (127) \qquad h = \overline{\Phi} / Q = 1 / H; \quad (128) \qquad V = \overline{\Phi} / \overline{T}, \quad (129)$$

где Q – объем выпущенной за год продукции; $\overline{T}$ – среднесписочная численность персонала.

В нашей задаче по формуле (127): H = 300 / 123,75 = 2,42, то есть на 1 руб. фондов произведено 2,42 руб. продукции; по формуле (128): h = 123,75 / 300 = 0,42, то есть на 1 руб. произведенной продукции приходится 42 коп. стоимости фондов; по формуле (129): V = 123,75 млн. руб. / 1000 чел. = 123,75 тыс. руб./чел., то есть на 1 работника приходится 123,75 тыс. руб. стоимости фондов.

Задача 2. Имеются следующие данные об использовании *материальных запасов* (*З*) предприятия:

Показатель	Базисный год ($_0$)	Отчетный год ($_1$)
Стоимость *З* в начале года, млн. руб.	30	45
Стоимость *З* в конце года, млн. руб.	12	23
Среднесуточный расход *З*, млн. руб./сут.	0,25	0,3
Годовой объем выпущенной продукции, млн. руб.	80	99

Определить абсолютные и относительные изменения показателей использования *З* предприятия.

Решение. Средний запас $\overline{З}$ предприятия определяется по формуле средней арифметической простой (17) как полусумма стоимости запасов на начало и конец года:

$$\overline{З}_0 = (30 + 12) / 2 = 21 \text{ (млн. руб.)}; \ \overline{З}_1 = (42 + 23) / 2 = 32{,}5 \text{ (млн. руб.)}$$

Для характеристики использования запасов используют следующие относительные показатели: *коэффициент* (скорость) *оборачиваемости* (130), *коэффициент закрепления* (131) и *время обращения* (средняя продолжительность оборота в днях) (132):

$$C = P / \overline{З}, \quad (130) \qquad K_{закр} = \overline{З} / P = 1 / C, \quad (131) \qquad B = 360 * K_{закр}, \quad (132)$$

где P – стоимость реализованной выпущенной продукции за год.

В нашей задаче по формуле (130) определяем: C_0 = 80 / 21 = 3,810 (оборотов в год); C_1 = 100 / 32,5 = 3,077 (оборотов в год), то есть оборачиваемость запасов уменьшилась с 3,81 оборотов в базисном году до 3,077 оборотов в отчетном.

Абсолютное изменение оборачиваемости $\Delta C = C_1 - C_0$ = 3,077 – 3,810 = – 0,733, то есть оборачиваемость уменьшилась на 0,733 раза в год.

Относительное изменение оборачиваемости $i_C = C_1/C_0$ = 3,077/3,81 = 0,808, то есть оборачиваемость запасов уменьшилась в 0,808 раза или на 19,2%.

По формуле (131) получаем: $K_{закр0}$ = 0,2625; $K_{закр1}$ = 0,325, то есть для реализации каждых 100 руб. продукции потребовалось материальных запасов: в базисном периоде – 26,3 руб., а в отчетном периоде – 32,5 руб.

Абсолютное изменение коэффициента закрепления $\Delta K_{закр} = K_{закр1} - K_{закр0}$ = 0,325 – 0,2625 = 0,0625, то есть коэффициент закрепления вырос на 0,0625 в отчетном году по сравнению с базисным.

Относительное изменение коэффициента закрепления $i_{Kзакр} = K_{закр1} / K_{закр0}$ = 0,325 / 0,2625 =1,238, то есть коэффициент закрепления вырос в 1,238 раза или на 23,8% в отчетном году по сравнению с базисным годом.

По формуле (132) получаем: B = 360*0,2625 = 94,5 (дня); B = 360*0,325 = 117 (дней), то есть средняя продолжительность оборота возросла с 94,5 дня в базисном году до 117 дней в отчетном.

Абсолютное изменение времени обращения $\Delta B = B_1 - B_0$ = 117 – 94,5 = 22,5 (дня), то есть время обращения запасов выросло на 22,5 дня в отчетном году по сравнению с базисным.

Относительное изменение времени обращения $i_B = B_1 / B_0$ = 117 / 94,5 = 1,238, то есть время обращения выросло в 1,238 раза или на 23,8% в отчетном году по сравнению с базисным.

Обеспеченность предприятия запасами в днях определяется путем деления размера запаса в начале периода (года) на среднесуточный расход (a), то есть по формуле (133):

$$O_{дн} = З / a. \qquad (133)$$

В нашей задаче по формуле (133) получаем: $O_{дн0}$ = 30 / 0,25 = 120 (дней); $O_{дн1}$ = 45 / 0,3 = 150 (дней), то есть обеспеченность запасами выросла со 120 дней в базисном году до 150 дней в отчетном.

Абсолютное изменение обеспеченности запасами $\Delta O_{дн} = O_{дн1} - O_{дн0}$ = 150 – 120 = 30 (дней), то есть обеспеченность запасами возросла на 30 дней в отчетном году по сравнению с базисным.

Относительное изменение обеспеченности запасами $i_{Одн} = O_{дн1} / O_{дн0}$ = 150 / 120 = 1,25, то есть обеспеченность запасами возросла в 1,25 раза или на 25% в отчетном году по сравнению с базисным.

Контрольные задания

Задание 1. Имеются следующие данные о динамике балансовой стоимости *основных фондов* (*Ф*):

Вариант	Дата	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.11	1.12	31.12
1	Стоимость Ф, млн. руб.	90	85	75	71	81	80	90	77	65	70	75	85	95
2		260	280	300				250	240	220		235	245	255
3		33	31	35	39	29	32	34	36	40	45	42	41	39
4		150	150	148	146	144	154	151	149	155	150	145	140	130
5		400			450			425	420	435	445	475	450	425

Годность основных фондов в начале года составляла 75%, норматив отчислений на реновацию - 10%, ликвидационная стоимость - 5% от стоимости выбывших фондов, годовая выручка - 200 млн. руб., среднесписочная численность производственного персонала - 1500 чел.

1. Определить среднегодовую балансовую стоимость основных фондов.
2. Составить балансы фондов по первоначальной полной и остаточной стоимостям.
3. Рассчитать показатели состояния, движения и использования основных фондов.

Задание 2. Имеются следующие условные данные по предприятию за базисный (БП) и отчетный (ОП) периоды:

Вариант	**6**		**7**		**8**		**9**		**10**	
Период	БП	ОП	БП	ОП	БП	ОП	БП	ОП	БП	ОП
Стоимость запасов в начале года, млн. руб.	15	20	50	40	110	105	1,5	0,9	70	80
Стоимость запасов в конце года, млн. руб.	10	18	20	15	105	95	0,8	0,4	60	50
Среднесуточный расход запасов, млн. руб./сут.	0,14	0,15	0,5	0,75	1,5	1,6	0,05	0,03	2,5	3,5
Годовой объем выпущенной продукции, млн. руб.	45	52	110	85	300	250	2,5	1,5	175	195

Определить абсолютные и относительные изменения показателей использования запасов предприятия, сделать выводы.

Тема 4. Статистика труда

Методические указания по теме

Задача 1. Имеются следующие данные по России на конец 2005 года, млн. человек: численность населения – 142,8, всего занято в экономике – 68,603, численность безработных – 5,2083.

Определить: 1) численность экономически активного населения (ЭАН); 2) коэффициент экономической активности населения; 3) коэффициенты занятости и безработицы.

Решение. *ЭАН (рабочая сила)* – это часть населения, которая предлагает свой труд для производства товаров и услуг. Численность ЭАН определяется как сумма занятых и безработных по формуле (134):

$$P_{ЭАН} = P_{зан} + P_{безр}. \qquad (134)$$

В нашей задаче по формуле (134) получаем: $P_{ЭАН}$ = 68,603 + 5,2083 = 73,8113 (млн. чел.).

Коэффициент экономической активности населения определяется как отношение численности ЭАН к общей численности населения, то есть по формуле (135):

$$K_{ЭАН} = P_{ЭАН} / P. \qquad (135)$$

В нашей задаче по формуле (135) получаем: $K_{ЭАН}$ = 73,8113 / 142,8 = 0,517, то есть численность ЭАН в общей численности населения составляет 51,7%.

Коэффициент занятости населения определяется как отношение численности занятого населения к численности ЭАН, то есть по формуле (136):

$$K_{зан} = P_{зан} / P_{ЭАН}. \qquad (136)$$

В нашей задаче по формуле (136) получаем: $K_{зан}$ = 68,603 / 73,8113 = 0,929, то есть численность занятого населения составляет 92,9% от численности ЭАН.

Коэффициент безработицы определяется как отношение численности безработного населения к численности ЭАН, то есть по формуле (137):

$$K_{безр} = P_{безр} / P_{ЭАН}. \qquad (137)$$

В нашей задаче по формуле (137) получаем: $K_{безр}$ = 5,2083 / 73,8113 = 0,071, то есть численность безработного населения составляет 7,1% от численности ЭАН.

Задача 2. Имеются следующие условные данные по области, тыс. человек:

На начало года:	
– численность трудоспособного населения в трудоспособном возрасте ($T_{ТВнг}$)	500
– численность работающих лиц за пределами трудоспособного возраста ($T_{внеТВнг}$)	15
В течение года:	
– вступило в трудоспособный возраст трудоспособного населения	30
– вовлечено для работы в отраслях экономики лиц пенсионного возраста	3
– прибыло из других областей трудоспособного населения в трудоспособном возрасте	20
– выбыло из состава трудовых ресурсов (в связи с переходом в пенсионный возраст, инвалидность, вследствие смерти и т.д.) трудоспособного населения	10
– выбыло из состава трудовых ресурсов подростков	4
– выбыло трудоспособного населения в трудоспособном возрасте в другие области	5

Определить:

1) численность трудовых ресурсов на начало года ($T_{НГ}$);
2) на конец года:
   а) численность трудоспособного населения в трудоспособном возрасте ($T_{ТВкг}$);
   б) численность работающих лиц, находящихся за пределами трудоспособного возраста ($T_{внеТВкг}$);
   в) численность трудовых ресурсов ($T_{кг}$);
3) среднегодовую численность трудовых ресурсов;
4) коэффициенты естественного, механического и общего движения трудовых ресурсов;
5) перспективную численность трудовых ресурсов на предстоящие 3 года, при условии, что коэффициент общего движения трудовых ресурсов будет ежегодно расти на 1‰.

Решение. Численность трудовых ресурсов (T) определяется как численность трудоспособного населения в трудоспособном возрасте $T_{ТВ}$ и работающих лиц за пределами трудоспособного возраста $T_{внеТВ}$ (лица пенсионного возраста и подростки), то есть по формуле (138):

$$T = T_{ТВ} + T_{внеТВ}. \qquad (138)$$

В нашей задаче по формуле (138) получаем: $T_{нг} = T_{ТВнг} + T_{внеТВнг}$ = 500 + 15 = 515 (тыс. чел.).

В течение года происходит постоянное изменение численности населения трудоспособного возраста по причинам:

1) естественного движения – выбытия за пределы трудоспособного возраста (в пенсионный возраст), а также пополнения при переходе в трудоспособный возраст;

2) механического движения – миграция населения трудоспособного возраста (приезд из других регионов и отъезд в другие регионы).

Таким образом, численность трудоспособного населения в трудоспособном возрасте в конце года можно определить путем корректировки

этой численности в начале года на величину приростов естественного ($\Delta_{ест}$) и механического движения ($\Delta_{мех}$):

$$T_{ТВкг} = T_{ТВнг} + \Delta_{ест} + \Delta_{мех}. \quad (139)$$

В нашей задаче по формуле (139) получаем: $T_{ТВкг}$ = 500 + (30 – 10) + (20 – 5) = 535 (тыс. чел.).

Для определения численности работающих лиц, находящихся за пределами трудоспособного возраста на конец года, необходимо прибавить к этой численности в начале года численность вовлеченных для работы лиц нетрудоспособного возраста ($T_{внеТВвовлеч}$) и вычесть численность выбывших из состава трудовых ресурсов лиц нетрудоспособного возраста ($T_{внеТВвыб}$):

$$T_{внеТВкг} = T_{внеТВнг} + T_{внеТВвовлеч} - T_{внеТВвыб}. \quad (140)$$

В нашей задаче по формуле (140) получаем: $T_{внеТВкг}$ = 15 + 3 – 4 = 14 (тыс. чел.).

Численность трудовых ресурсов на конец года определим по формуле (138): $T_{кг} = T_{ТВкг} + T_{внеТВкг}$ = 535 + 14 = 549 (тыс. чел.).

Среднегодовая численность трудовых ресурсов определяем по формуле средней арифметической простой (17) как полусумма численности трудовых ресурсов на начало и конец года:

$$\overline{T} = \frac{T_{нг} + T_{кг}}{2} = (515 + 549) / 2 = 532 \text{ (тыс. чел.)}.$$

Коэффициенты естественного, механического и общего движения трудовых ресурсов рассчитываются путем деления соответствующих приростов на среднегодовую численность трудовых ресурсов по следующим формулам:

$$K_{ест} = \frac{\Delta_{ест}}{\overline{T}} 1000\%_0, \quad (141)$$

$$K_{мех} = \frac{\Delta_{мех}}{\overline{T}} 1000\%_0, \quad (142)$$

$$K_{общ} = \frac{\Delta_{общ}}{\overline{T}} 1000\%_0 = \frac{\Delta_{ест} + \Delta_{мех}}{\overline{T}} 1000\%_0 = K_{ест} + K_{мех}. \quad (143)$$

В нашей задаче по формуле (141) получаем: $K_{ест}$ = (30 + 3 – 10 – 4) / 532 * 1000‰ = 35,7‰, а по формуле (142): $K_{мех}$ = (20 – 5) / 532 * 1000‰ = 28,2‰, и, наконец, по формуле (143): $K_{общ}$ = (30 + 3 – 10 – 4 + 20 – 5) / 532 * 1000‰ = 35,7‰ + 28,2‰ = 63,9‰.

Перспективную численность трудовых ресурсов на следующий год можно найти по формуле (144):

$$\overline{T}_{след} = \overline{T}_{пред}\left(1 + \frac{К_{общ}}{1000\%_0}\right). \quad (144)$$

В нашей задаче по формуле (144) численность трудовых ресурсов через год составит: $\overline{T}_{след1}$= 535 * (1 + 64,9 / 1000) = 566 (тыс. чел.), а еще через год по той же формуле: $\overline{T}_{след2}$= 566 * (1 + 65,9 / 1000) = 603,299 (тыс. чел.), и, наконец, через три года: $\overline{T}_{след3}$= 603,299 * (1 + 66,9 / 1000) = 643,66 (тыс. чел.).

Задача 3. Заполните годовой баланс рабочего времени предприятия (табл. 16) недостающими данными.

Таблица 16. Баланс рабочего времени за год

Фонды времени	Условные обозначения	Чел.-дней	Использование рабочего времени	Условные обозначения	Чел.-дней
Календарный фонд рабочего времени	КФРВ		Отработанное время	ОВ	
Неявки вследствие праздничных и выходных дней	ПВД	145500	Целодневные простои	ЦП	100
Табельный фонд рабочего времени	ТФРВ		Неявки по уважительным причинам, всего	УП	20000
Неявки в связи с очередными отпусками	ОО	31500	Неявки по неуважительным причинам, всего	НП	5000
Максимально возможный фонд рабочего времени	МВФРВ				
Баланс			Баланс		

Среднесписочная численность работников данного предприятия $\overline{T}$ = 1500 чел., установленная продолжительность рабочего дня в этой отрасли – 8 ч, отработано за год 2625040 чел.-часов.

Определить:

1. Календарный, табельный и максимально возможный фонды рабочего времени;
2. Коэффициенты использования этих фондов времени;
3. Коэффициенты использования рабочего периода и рабочего дня;
4. Интегральный коэффициент использования рабочего времени.

Решение. КФРВ – это сумма человеко-дней, которые отработали бы в течение года все работники предприятия при ежедневной работе. Календарный фонд определяется по формулам (145) и (146):

$$КФРВ = \overline{T}\, d_{год}, \qquad (145)$$

где $d_{год}$ – число календарных дней в году.

$$КФРВ = ОВ + ЦП + (ПВД + ОО) + УП + НП. \qquad (146)$$

В нашей задаче по формуле (145) получаем: *КФРВ* =1500 * 365 = 547500 (чел.-дней), то есть если бы каждый из 1500 работников предприятия работал все календарные дни года, то фонд рабочего времени всех работников за год составил 547500 чел.-дней.

Из формулы (146) можно найти отработанное время: *ОВ* = 547500 – 100 – (145500+31500) – 20000 – 5000 = 345400 (чел.-дней).

ТФРВ – это сумма человеко-дней, которые отработали бы в течение года все работники предприятия, если бы их неявки на работу были связаны только с праздничными и выходными днями. Табельный фонд определяется по формулам (147) или (148):

$$ТФРВ = КФРВ – ПВД, \quad (147)$$

$$ТФРВ = ОВ + ЦП + ОО + УП + НП. \quad (148)$$

В нашей задаче по формуле (147) получаем: *ТФРВ* = 547500 – 145500 = 402000 (чел.-дней) или по формуле (148): *ТФРВ* = 345400 + 100 + 31500 + 20000 + 5000 = 402000 (чел.-дней), то есть если бы каждый работник предприятия не являлся на работу только по причине праздничных и выходных дней, то фонд рабочего времени всех работников за год составил 402000 чел.-дней.

МВФРВ – это сумма человеко-дней, которые отработали бы в течение года все работники предприятия, если бы их неявки на работу были связаны только с реализацией права на отдых (праздничные, выходные дни и очередные отпуска). Максимально возможный фонд определяется по формулам (149) – (151):

$$МВФРВ = КФРВ – (ПВД + ОО), \quad (149)$$

$$МВФРВ = ТФРВ – ОО, \quad (150)$$

$$МВФРВ = ОВ + ЦП + УП + НП. \quad (151)$$

В нашей задаче по формуле (149) получаем: *МВФРВ* = 547500 – (145500 + 31500) = 370500 (чел.-дней). Аналогичный результат получаем по формуле (150): *МВФРВ* = 402000 – 31500 = 370500 (чел.-дней) или по формуле (151): *МВФРВ* = 345400 + 100 + 20000 + 5000 = 370500 (чел.-дней), то есть если бы каждый работник предприятия не являлся на работу только по причине реализации права на отдых, то фонд рабочего времени всех работников за год составил бы 370500 чел.-дней.

Коэффициент использования того или иного фонда времени – это доля отработанного времени в соответствующем фонде. Данные коэффициенты определяются по формулам (152) – (154):

$$К_{испКФРВ} = ОВ / КФРВ, \quad (152)$$

$$К_{испТФРВ} = ОВ / ТФРВ, \quad (153)$$

$$К_{испМВФРВ} = ОВ / МВФРВ. \quad (154)$$

В нашей задаче по формуле (152) получаем: $К_{испКФРВ}$ = 345400 / 547500 = 0,631, то есть из-за неявок по всем причинам и целодневных простоев

фактически отработанный фонд рабочего времени составляет 63,1% от календарного, то есть фонд недоиспользован на 36,9% или на 202100 (345400 – 547500) чел.-дней.

По формуле (153) получаем: $K_{испТФРВ}$ = 345400 / 402000 = 0,859, то есть из-за неявок по всем причинам, кроме выходных и праздничных дней и целодневных простоев фонд недоиспользован на 14,1%.

По формуле (154) получаем: $K_{испМВФРВ}$ = 345400 / 370500 = 0,932, то есть из-за неявок по всем причинам, кроме выходных и праздничных дней и очередных отпусков, а также целодневных простоев фонд недоиспользован на 6,8%.

Коэффициент использования рабочего периода (года) определяется по формуле (155):

$$K_{исп.года} = \frac{d_{факт}}{d_{уст}}, \qquad (155)$$

где $d_{факт}$ – фактическая продолжительность года, $d_{уст}$ – установленная продолжительность года.

Фактическая и установленная продолжительности года определяются по формулам (156) и (157):

$$d_{факт} = \frac{ОВ}{\overline{T}}, \qquad (156) \qquad\qquad d_{уст} = \frac{МВФРВ}{\overline{T}}. \qquad (157)$$

В нашей задаче по формуле (156) получаем: $d_{факт}$ = 345400 / 1500 = 230,2 (дня), то есть из-за неявок по всем причинам и целодневных простоев каждый работник отработал в среднем не 365 календарных дней года, а только 230,2 дня.

По формуле (157) получаем: $d_{уст}$ = 370500 / 1500 = 247 (дней), то есть с учетом неявок вследствие выходных и праздничных дней и очередных отпусков каждый работник должен отработать в среднем 247 дней в году.

И, наконец, по формуле (155) получаем: $K_{исп.года}$ = 230,267 / 247 = 0,9323, то есть из-за неявок по всем причинам, кроме выходных и праздничных дней, а также очередных отпусков, фактическая рабочая продолжительность года недоиспользована на 6,77%, или на 16,8 (230,2 – 247) дня.

Коэффициент использования рабочего дня определяется по формуле (158):

$$K_{исп.дня} = \frac{t_{факт}}{t_{уст}}, \qquad (158)$$

где $t_{факт}$ – фактическая продолжительность дня, определяемая как соотношение количества отработанных человеко-часов и человеко-дней; $t_{уст}$ – установленная продолжительность дня, определяемая нормативами в зависимости от сферы деятельности.

В нашей задаче по формуле (158) получаем: $К_{исп.дня}$ = (2625040/345400) / 8 = 7,6 / 8 = 0,95, то есть продолжительность рабочего дня недоиспользована на 5% или на 0,4 (7,6 – 8) часа.

Интегральный коэффициент использования рабочего времени учитывает использование как рабочего года, так и рабочего дня. Он определяется по формулам (159) – (161):

$$К_{инт} = \frac{ОВ_{(чел-ч)}}{МВФРВ_{(чел-ч)}} = \frac{ОВ_{(чел-ч)}}{МВФРВ_{(чел-дн)} * t_{уст}}, \quad (159)$$

$$К_{инт} = \frac{ОВ_{(ч/чел)}}{МВФРВ_{(ч/чел)}}, \quad (160)$$

$$К_{инт} = К_{исп.года} * К_{исп.дня}, \quad (161)$$

где $ОВ_{чел.-ч.}$ – отработанное время за год в человеко-часах; $МВФРВ_{чел.-ч.}$– максимально возможный фонд рабочего времени в человеко-часах; $МВФРВ_{чел.-дн.}$– максимально возможный фонд рабочего времени в человеко-днях; $МВФРВ_{ч/чел.}$– установленное для работника время на год в часах, $МВФРВ_{ч/чел.} = d_{уст}t_{уст}$; $ОВ_{ч/.чел.}$ – отработанное время за год одним работником.

В нашей задаче по формуле (159) получаем: $К_{инт}$ = 2625040 / (370500*8) = 0,886. Аналогичный результат получаем по формуле (160): $К_{инт}$ = (2625040 / 1500) / (247*8) = 0,886 или по формуле (161): $К_{инт}$ = 0,9323 * 0,95 = 0,886, то есть рабочее время за год недоиспользовано на 11,4%.

Задача 4. Имеются следующие данные по предприятию:

Показатель	Базисный период	Отчетный период
Среднесписочная численность работающих ($\overline{T}$), чел.	3200	3926
Произведенная продукция (выручка – Q), тыс. руб.	42400	53000
Фонд заработной платы (F), тыс. руб.	35200	39072

Определить влияние экстенсивного и интенсивного факторов на изменение: 1) выручки; 2) фонда заработной платы.

Решение. Изменение выручки определим по формуле (2):

i_Q = 53000/ 42400 = 1,25.

Это означает, что в отчетном периоде по сравнению с базисным выручка возросла в 1,25 раза (на 25%) или на 10600 (53000 – 42400) тыс. руб. Данное изменение выручки было получено в результате влияния двух факторов: изменения численности работающих (экстенсивный фактор) и изменения производительности труда (интенсивный фактор). Производительность труда определяется по формуле (162):

$$w = \frac{Q}{\overline{T}}. \quad (162)$$

В нашей задаче по формуле (162) получаем: w_0 = 42400 / 3200 = 13,25 (тыс.руб./чел.), то есть в базисном периоде на одного работающего приходилось 13,25 тыс. руб. продукции; а в отчетном периоде w_1 = 53000 / 3926 = 13,50 (тыс.руб./чел.). Тогда по формуле (2) изменение производительности труда составит: i_w = 13,5 / 13,25 = 1,0189, то есть производительность труда в отчетном периоде по сравнению с базисным возросла примерно в 1,0189 раза (на 1,89%) или на 0,25 (13,5 – 13,25) тыс.руб./чел.

Изменение среднесписочной численности работающих по формуле (2) составило: i_T =3926 / 3200 = 1,2269, то есть среднесписочная численность работающих возросла примерно в 1,2269 раза (на 22,69%) или на 726 (2926 – 3200) чел.

Для проверки правильности расчетов влияния экстенсивного и интенсивного факторов на изменение выручки воспользуемся следующей двухфакторной мультипликативной моделью (163):

$$i_Q = i_T i_w. \qquad (163)$$

В нашей задаче по формуле (163) получаем: 1,25 = 1,2269*1,0189, то есть выручка возросла за счет увеличения среднесписочной численности работающих (экстенсивный фактор) в 1,2269 раза и за счет увеличения производительности труда каждого работающего (интенсивный фактор) в 1,0189 раза. Для определения влияния этих факторов в абсолютных величинах (в рублях) воспользуемся формулами (164) и (165):

$$\Delta Q_T = (i_T - 1)Q_0, \qquad (164)$$
$$\Delta Q_w = i_T(i_w - 1)Q_0. \qquad (165)$$

В нашей задаче по формуле (164) получаем: ΔQ_T = (1,2269–1)*42400 = 9620 (тыс. руб.), то есть вследствие увеличения численности работающих на 22,69% выручка увеличилась на 9620 тыс. руб., что составляет примерно 90,8% от общего изменения (9620 /10600*100%).

По формуле (165) получаем: ΔQ_w = 1,2269*(1,0189–1)*42400 = 980 (тыс. руб.), то есть вследствие увеличения производительности труда работающих на 1,89% выручка увеличилась на 980 тыс. руб., что составляет примерно 9,2% от общего изменения (980/10600*100%).

Для проверки правильности расчетов воспользуемся формулой (166):

$$\Delta Q_T + \Delta Q_w = Q_1 - Q_0. \qquad (166)$$

В нашей задаче по формуле (166) получаем: 9620 + 980 = 53000 – 42400 = 10600.

Изменение фонда заработной платы F определим по формуле (2):

i_F = 39072 / 35200 = 1,11.

Это означает, что в отчетном периоде по сравнению с базисным фонд возрос в 1,11 раза или на 11%. В абсолютном выражении данный прирост составил 3872 (39072 – 35200) тыс. руб. Он был получен в результате влияния двух факторов:

изменения численности работающих (экстенсивный фактор) и изменения уровня заработной платы (интенсивный фактор). Уровень заработной платы определяется по формуле (167):

$$f = \frac{F}{\overline{T}}. \qquad (167)$$

В нашей задаче по формуле (167) получаем: f_0 = 35200 / 3200 = 11 (тыс. руб./чел.), то есть в базисном периоде средний уровень заработной платы составлял 11 тыс. руб./чел.; а в отчетном периоде – f_1 = 39072 / 3926 = 9,952 (тыс.руб./чел.). Тогда по формуле (2) изменение уровня заработной платы составит: i_f = 9,952 / 11 = 0,9047. Это означает, что средний уровень заработной платы в отчетном периоде по сравнению с базисным уменьшился в 0,9047 раза (на 9,53%) или на 1,048 (9,952 – 11) тыс. руб./чел.

Для проверки правильности расчетов влияния экстенсивного и интенсивного факторов на изменение фонда заработной платы воспользуемся следующей двухфакторной мультипликативной моделью (168):

$$i_F = i_T\, i_f. \qquad (168)$$

В нашей задаче по формуле (168) получаем: 1,11 = 1,2269*0,9047, то есть фонд возрос за счет увеличения среднесписочной численности работающих (экстенсивный фактор) в 1,2269 раза и за счет уменьшения уровня заработной платы каждого работающего (интенсивный фактор) в 0,9047 раза. Для определения влияния этих факторов в абсолютных величинах (в рублях) воспользуемся формулами (169) и (170):

$$\Delta F_T = (i_T - 1)F_0, \qquad (169)$$
$$\Delta F_f = i_T(i_f - 1)F_0. \qquad (170)$$

В нашей задаче по формуле (169) получаем: ΔF_T = (1,2269–1)*35200 = 7986,88 (тыс. руб.), а по формуле (170): ΔF_f = 1,2269*(0,9047 – 1)*35200 = –4115,71 (тыс. руб.), то есть вследствие увеличения численности работающих на 22,69% фонд заработной платы должен был увеличиться на 7986,88 тыс. руб., однако, за счет уменьшения уровня заработной платы на 9,53% он уменьшился на 4115,71 тыс. руб.

Для проверки правильности расчетов воспользуемся формулой (171):

$$\Delta F_T + \Delta F_f = F_1 - F_0. \qquad (171)$$

В нашей задаче по формуле (171) получаем: 7986,88 + (–4115,71) = 39072 – 35200 = 3872.

Контрольные задания

Задание 1. Имеются следующие условные данные по стране, млн. человек:

Вариант	**1**	**2**	**3**	**4**	**5**
Численность населения	53	186	42	24	94
Всего занято в экономике	31	105	28	15,5	41
Численность безработных	4	21	2,5	1,8	18

Определить: 1) численность экономически активного населения; 2) коэффициент экономической активности населения; 3) коэффициенты занятости и безработицы.

Задание 2. Имеются следующие условные данные за год по предприятию:

Вариант	**1**	**2**	**3**	**4**	**5**
Неявки вследствие праздничных и выходных дней, тыс. чел.-дн.	100	475	33,6	247,5	9,5
Неявки в связи с очередными отпусками, тыс. чел.-дн.	20,5	95	7	42,5	2,05
Целодневные простои, чел.-дн.	80	700	30	190	5
Неявки по уважительным причинам, всего, тыс. чел.-дн.	25	90	10	44	3
Неявки по неуважительным причинам, всего, тыс. чел.-дн.	4	15	1,5	5,5	0,25
Среднесписочная численность работников, чел	1000	5000	350	2500	100
Всего отработано за год, тыс. чел-часов	1838,64	8941,875	660,946	4649,453	171,157
Установленная продолжительность рабочего дня, часов	8	7,5	7,9	7,7	7

Составьте годовой баланс рабочего времени предприятия и определите:

1. Календарный, табельный и максимально возможный фонды рабочего времени;
2. Коэффициенты использования этих фондов времени;
3. Коэффициенты использования рабочего периода и рабочего дня;
4. Интегральный коэффициент использования рабочего времени.

Задание 3. Имеются следующие условные данные по области, тыс. человек:

Вариант	**6**	**7**	**8**	**9**	**10**
На начало года:					
– численность трудоспособного населения в трудоспособном возрасте	53	186	42	24	94
– численность работающих лиц за пределами трудоспособного возраста	4	21	2,5	1,8	18
В течение года:					
– вступило в трудоспособный возраст трудоспособного населения	3	6	2	1	5
– вовлечено для работы в отраслях экономики лиц пенсионного возраста	1,5	2,5	3	0,6	1,3
– прибыло из других областей трудоспособного населения в	1	2	6	3	5

Вариант	6	7	8	9	10
трудоспособном возрасте					
– выбыло из состава трудовых ресурсов (в связи с переходом в пенсионный возраст, инвалидность, вследствие смерти и т.д.) трудоспособного населения	2	1	5	3	6
– выбыло из состава трудовых ресурсов подростков	2	3	4	1,5	6
– выбыло трудоспособного населения в трудоспособном возрасте в другие области	2,5	3,5	7	4	1,5

Определить:

1) численность трудовых ресурсов на начало года;
2) на конец года:
   а) численность трудоспособного населения в трудоспособном возрасте;
   б) численность работающих лиц, находящихся за пределами трудоспособного возраста;
   в) численность трудовых ресурсов;
3) среднегодовую численность трудовых ресурсов;
4) коэффициенты естественного, механического и общего движения трудовых ресурсов;
5) перспективную численность трудовых ресурсов на предстоящие 3 года, при условии, что коэффициент общего движения трудовых ресурсов будет ежегодно снижаться на 1‰.

Задание 4. Имеются следующие условные данные по предприятию за базисный (БП) и отчетный (ОП) периоды:

Вариант	6		7		8		9		10	
Период	БП	ОП	БП	ОП	БП	ОП	БП	ОП	БП	ОП
Среднесписочная численность работающих, чел.	1000	1500	500	450	2500	2900	5000	4000	200	250
Товарная продукция, млн. руб.	50	60	20	15	100	110	300	200	10	11
Фонд заработной платы, тыс. руб.	40	45	15	14	90	94	250	190	8	10

Определить влияние экстенсивного и интенсивного факторов на изменение: 1) выручки; 2) фонда заработной платы.

Учебное издание

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
для проведения практических занятий по дисциплине
«СТАТИСТИКА»
для студентов направления подготовки
44.03.04 Профессиональное обучение (по отраслям)

С о с т а в и т е л ь:
Дарья Сергеевна Тимошенко

Печатается в авторской редакции.
Компьютерная верстка и оригинал-макет автора.

Подписано в печать _____
Формат 60x84 1/16. Бумага типограф. Гарнитура Times
Печать офсетная. Усл. печ. л. _____. Уч.-изд. л. _____
Тираж 100 экз. Изд. №_____. Заказ №_____. Цена договорная.

Издательство Луганского государственного
университета имени Владимира Даля

Свидетельство о государственной регистрации издательства
МИ-СРГ ИД 000003 от 20 ноября 2015г.

Адрес издательства:91034, г. Луганск, кв. Молодежный, 20а
Телефон: 8 (0642) 41-34-12, **факс:** 8 (0642) 41-31-60
E-mail:uni@snu.edu.ua**http**: www.snu.edu.ua