

1. Линейная парная регрессия

Краткая теоретическая справка

Регрессия [regression] – зависимость среднего значения какой-либо случайной величины от некоторой другой величины (парная регрессия) или нескольких величин (множественная регрессия).

Уравнение линейной парной регрессии имеет вид: $\hat{y} = a + bx$.

Для оценки параметров a , b методом наименьших квадратов (МНК) необходимо решить систему нормальных уравнений:

$$\begin{cases} na + b \sum x = \sum y, \\ a \sum x + b \sum x^2 = \sum yx. \end{cases} \quad (1.1)$$

Можно воспользоваться готовыми формулами решения системы:

$$b = \frac{\text{cov}(x; y)}{\sigma_x^2} = \frac{\overline{y \cdot x} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2}, \quad a = \bar{y} - b \cdot \bar{x}, \quad (1.2)$$

где $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ – среднее значение фактора X;

$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ – среднее значение результирующей переменной Y;

$\overline{y \cdot x} = \frac{1}{n} \sum_{i=1}^n y_i \cdot x_i$ – среднее значение произведения переменных X и Y;

$\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$ – среднее значение квадрата переменной X;

$\text{cov}(x; y) = \overline{x \cdot y} - \bar{x} \cdot \bar{y}$ – ковариация переменных X и Y;

$\sigma_x^2 = D_x = \overline{x^2} - \bar{x}^2$ – дисперсия переменной X.

Коэффициент регрессии b показывает, на сколько единиц в среднем по совокупности изменится результирующая переменная Y, если факторная переменная X увеличится на одну единицу.

Для оценки тесноты линейной связи между переменными используют линейный коэффициент парной корреляции:

$$r_{xy} = b \frac{\sigma_x}{\sigma_y} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}, \quad (1.3)$$

где $\sigma_x = \sqrt{D_x} = \sqrt{\overline{x^2} - \bar{x}^2}$ – среднеквадратическое отклонение (СКО) переменной X;

$\sigma_y = \sqrt{D_y} = \sqrt{y^2 - \bar{y}^2}$ – среднее квадратическое отклонение (СКО) переменной Y.

Можно считать, что:

1) если $r_{xy} > 0$, то имеется *прямая* линейная связь между переменными X и Y;

2) если $r_{xy} < 0$, то имеется *обратная* линейная связь между переменными X и Y;

3) если $r_{xy} \approx 0$ ($|r_{xy}| < 0,1$), то линейная связь между переменными X и Y *отсутствует*.

Качественная оценка тесноты связи величин X и Y может быть выявлена на основе шкалы Чеддока:

Тестона связи	Значение коэффициента корреляции
Слабая	0,1-0,3
Умеренная	0,3-0,5
Заметная	0,5-0,7
Высокая	0,7-0,9
Весьма высокая	0,9-0,99

Для оценки качества уравнения регрессии использую *коэффициент детерминации* R^2 .

Коэффициент детерминации характеризует долю дисперсии, объясняемую регрессией, в общей дисперсии результативного признака (квадрат коэффициента корреляции):

$$R^2 = r_{xy}^2. \quad (1.4)$$

Коэффициент детерминации показывает, какую часть вариации (изменения) результативной переменной Y объясняет вариация (изменение) фактора X. Чем ближе R^2 к единице, тем лучше регрессионная модель.

Оценка статистической значимости уравнения регрессии в целом осуществляется с помощью *F-критерия Фишера*. Проверяется гипотеза H_0 о статистической незначимости уравнения регрессии. Для этого рассчитывается фактическое значение критерия по формуле:

$$F_{\text{факт}} = \frac{\sum(\hat{y} - \bar{y}) / m}{\sum(y - \hat{y})^2 / (n - m - 1)}, \quad (1.5)$$

где n – число единиц совокупности;

m – число параметров при переменных x .

Если применяется линейное уравнение регрессии, то расчет $F_{\text{факт}}$ упрощается:

$$F_{\text{факт}} = \frac{R^2}{1 - R^2} \cdot (n - 2). \quad (1.6)$$

$F_{\text{табл}}$ – это максимально возможное значение критерия, которое могло сформироваться под влиянием случайных факторов при данных степенях свободы и уровне значимости α . Уровень значимости α – вероятность отвергнуть правильную гипотезу при условии, что она верна. Имеются таблицы критических (табличных) значений F-критерия: $F(\alpha; k_1; k_2)$, где $k_1 = m$, $k_2 = n - m - 1$. Для линейного уравнения парной регрессии с уровнем значимости $\alpha = 0,05$ необходимо в таблице значений (приложение №4) найти значение $F(0,05; 1; n - 2)$.

Если $F_{\text{табл}} < F_{\text{факт}}$, то гипотеза H_0 о случайной природе оцениваемых характеристик отклоняется и признается их статистическая значимость и надежность.

Для оценки статистической значимости коэффициентов регрессии и коэффициента корреляции рассчитывается t-критерий Стьюдента. Выдвигается гипотеза H_0 о случайной природе показателей, т.е. о незначимом их отличии от нуля. Наблюдаемые значения t-критерия рассчитываются по формулам:

$$t_b = \frac{b}{m_b}, t_a = \frac{a}{m_a}, t_r = \frac{|r_{xy}|}{m_r}, \quad (1.7)$$

где m_b, m_a, m_r – случайные ошибки параметров линейной регрессии и коэффициента корреляции.

Для линейной парной регрессии выполняется равенство $t_b = t_r = \sqrt{F}$, поэтому проверки гипотез о значимости коэффициента регрессии при факторе и коэффициента корреляции равносильны проверке гипотезы о статистической значимости уравнения регрессии в целом.

Вообще, случайные ошибки рассчитываются по формулам:

$$m_r = \sqrt{\frac{1 - r_{xy}^2}{n - 2}}, m_b = \frac{S_{\text{ост}}}{\sigma_x \sqrt{n}}, m_a = S_{\text{ост}} \frac{\sqrt{\sum x^2}}{n \sigma_x}. \quad (1.8)$$

где $S_{\text{ост}}^2$ – остаточная дисперсия на одну степень свободы:

$$S_{\text{ост}} = \sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}}. \quad (1.9)$$

Табличное (критическое) значение t-статистики находят по таблицам распределения t-Стьюдента при уровне значимости $\alpha = 0,05$ и числе степеней свободы $k = n - 2$ (приложение №5). Если $t_{\text{табл}} < t_{\text{факт}}$, то H_0 отклоняется, т.е. коэффициенты регрессии не случайно отличаются от нуля и сформировались под влиянием систематически действующего фактора.

Решение типового примера

Имеются статистические данные (приложение 1, ст. 0, ст. 31) о численности населения (оценка на конец 2008 года; сотни тысяч человек) (фактор X) и об обороте розничной торговли в этих же регионах (в фактически действовавших ценах, млрд. руб.) (результативный признак Y) в некоторых регионах Российской Федерации.

№	Регион	X	Y	№	Регион	X	Y
1	Республика Адыгея	4,4	25,6	9	Ставропольский край	27,1	203,6
2	Республика Дагестан	27,1	217,3	10	Астраханская область	10,1	77,7
3	Кабардино-Балкарская Республика	8,9	48	11	Волгоградская область	26	186,1
4	Республика Калмыкия	2,8	7,4	12	Ростовская область	42,4	423,4
5	Карачаево-Черкесская Республика	4,3	22,6	13	Ивановская область	10,7	57,9
6	Республика Северная Осетия - Алания	7	37,4	14	Калужская область	10	80,7
7	Чеченская Республика	12,4	25,2	15	Костромская область	6,9	38,1
8	Краснодарский край	51,4	500,7				

Требуется:

1. Построить график зависимости между переменными, по которому необходимо подобрать модель уравнения регрессии.

2. Рассчитать параметры уравнения регрессии $y = a + bx$ методом наименьших квадратов.

3. Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

4. Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием F-критерия Фишера при уровне значимости $\alpha = 0,05$.

5. Оценить значимость коэффициентов регрессии и корреляции по t-критерию Стьюдента при уровне значимости $\alpha = 0,05$.

6. Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1,3 от его среднего уровня по совокупности.

Решение:

1. Построим график зависимости оборота розничной торговли от численности населения в среде MS Excel с помощью точечной диаграммы (рис. 1.1). По оси абсцисс откладываем значения фактора X (численность населения), по оси ординат – значения результативного признака Y (оборот розничной торговли). Визуальный анализ диаграммы показывает, что можно использовать линейную регрессионную модель.

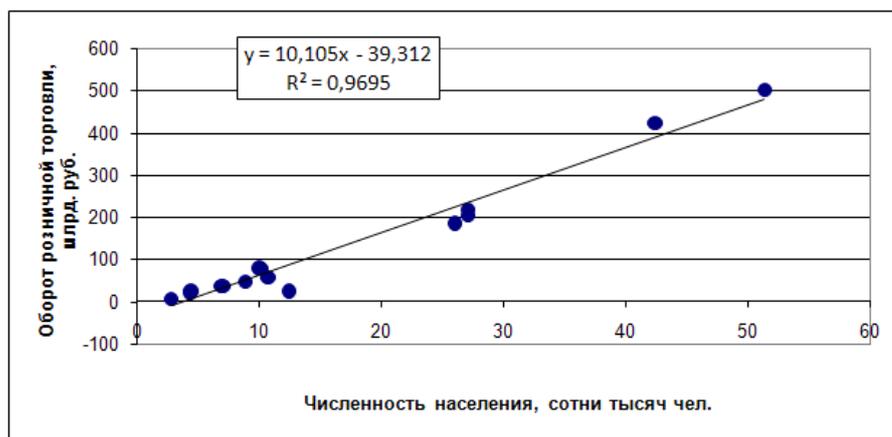


Рис. 1.1. Зависимость оборота розничной торговли от численности населения по регионам РФ.

2. Составим расчетную таблицу, заполним столбцы 1-6.

Таблица 1.1

№	x	y	x^2	y^2	xy	\hat{y}	$y - \hat{y}$	$(y - \hat{y})^2$
1	2	3	4	5	6	7	8	9
1.	4,4	25,6	19,36	655,36	112,64	5,15	20,45	418,22
2.	27,1	217,3	734,41	47219,29	5888,83	234,53	-17,23	296,89
3.	8,9	48,0	79,21	2304,00	427,20	50,62	-2,62	6,87
4.	2,8	7,4	7,84	54,76	20,72	-11,02	18,42	339,23
5.	4,3	22,6	18,49	510,76	97,18	4,14	18,46	340,81
6.	7,0	37,4	49,00	1398,76	261,80	31,42	5,98	35,73
7.	12,4	25,2	153,76	635,04	312,48	85,99	-60,79	3695,26
8.	51,4	500,7	2641,96	250700,49	25735,98	480,08	20,62	425,21
9.	27,1	203,6	734,41	41452,96	5517,56	234,53	-30,93	956,70
10.	10,1	77,7	102,01	6037,29	784,77	62,75	14,95	223,58
11.	26,0	186,1	676,00	34633,21	4838,60	223,42	-37,32	1392,42
12.	42,4	423,4	1797,76	179267,56	17952,16	389,14	34,26	1174,07
13.	10,7	57,9	114,49	3352,41	619,53	68,81	-10,91	119,04
14.	10,0	80,7	100,00	6512,49	807,00	61,74	18,96	359,60
15.	6,9	38,1	47,61	1451,61	262,89	30,41	7,69	59,11
Сумма	251,5	1951,7	7276,31	576185,99	63639,34			9842,74
Среднее	16,77	130,11	485,09	38412,40	4242,62			
Дисперсия	203,97	21482,92						
СКО	14,28	146,57						

Найдем дисперсию и среднее квадратическое отклонение переменных:

$$\text{Дисперсия: } D_x = \overline{x^2} - \bar{x}^2 = 485,09 - 16,77^2 = 203,97;$$

$$D_y = \overline{y^2} - \bar{y}^2 = 38412,4 - 130,11^2 = 21482,92.$$

Среднее квадратическое отклонение (СКО):

$$\sigma_x = \sqrt{D_x} = \sqrt{203,97} = 14,28;$$

$$\sigma_y = \sqrt{D_y} = \sqrt{21482,92} = 146,57.$$

Найдем параметры уравнения регрессии:

$$b = \frac{\text{cov}(x; y)}{\sigma_x^2} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{D_x} = \frac{4242,62 - 16,77 \cdot 130,11}{203,97} = 10,1;$$

$$a = \bar{y} - b \cdot \bar{x} = 130,11 - 10,1 \cdot 16,77 = -39,31.$$

Таким образом, уравнение регрессии имеет вид: $\hat{y} = 10,1x - 39,31$.

Вывод: с увеличением численности населения на 1 единицу (т.е. на 100 000 человек) оборот розничной торговли увеличивается в среднем на 10 100 млн. рублей.

3. Оценим тесноту связи между переменными с помощью показателей корреляции и детерминации.

Линейный коэффициент корреляции:

$$r_{xy} = b \frac{\sigma_x}{\sigma_y} = 10,1 \cdot \frac{14,28}{146,57} = 0,98.$$

Вывод: линейная связь между оборотом розничной торговли и численностью населения по регионам весьма высокая, т.к. $r_{xy} = 0,98 \geq 0,9$.

Коэффициент детерминации:

$$R^2 = r_{xy}^2 = 0,98^2 = 0,97.$$

Вывод: Коэффициент детерминации показывает, что 97% различий в обороте розничной торговли по регионам (y) объясняется вариацией численности населения в этих регионах (x), а 3% другими, неучтенными факторами.

4. Охарактеризуем статистическую надежность результатов регрессионного анализа с использованием F-критерия Фишера при уровне значимости $\alpha = 0,05$.

Рассчитаем фактическое значение критерия:

$$F_{\text{факт}} = \frac{R^2}{1 - R^2} \cdot (n - 2) = \frac{0,97}{1 - 0,97} \cdot (15 - 2) = 412,61.$$

Найдем табличное значение критерия при уровне значимости $\alpha = 0,05$ и степенях свободы: $k_1 = m = 1$, $k_2 = n - m - 1 = 15 - 1 - 1 = 13$:

$$F_{\text{табл}}(\alpha = 0,05; k_1 = 1; k_2 = 13) = 4,67 \text{ (см. приложение №4).}$$

Т.к. $F_{\text{табл}} = 4,67 < 412,61 = F_{\text{факт}}$, то уравнение регрессии является статистически значимым, надежным.

5. Оценим значимость коэффициентов регрессии и корреляции по t-критерию Стьюдента при уровне значимости $\alpha = 0,05$.

Поскольку сделан вывод о статистической значимости уравнения регрессии в целом по критерию Фишера, то статистическими значимо отличны от нуля и коэффициент корреляции r_{xy} и коэффициент регрессии b при факторе.

Продолжим заполнять расчетную таблицу 1.1, заполним столбцы 7-9.

Для заполнения столбца 7, необходимо рассчитать теоретические значения переменной y соответствующие имеющимся значениям фактора x с помощью полученного уравнения регрессии $y = 10,1x - 39,31$.

$$\text{Так, для } x = 4,4: \hat{y} = 10,1 \cdot 4,4 - 39,31 = 5,15.$$

Для $x = 27,1$: $\hat{y} = 10,1 \cdot 27,1 - 39,31 = 234,53$ и т.д.

Столбец 8 заполняется как разность соответствующих значений столбца 2 и столбца 7, т.е. рассчитывается по формуле $y - \hat{y}$.

Так, для $i = 1$: $y - \hat{y} = 25,6 - 5,15 = 20,45$;

для $i = 2$: $y - \hat{y} = 217,3 - 234,53 = -17,23$ и т.д.

Далее заполним 9 столбец расчетной таблицы по формуле $(y - \hat{y})^2$ и сумму этого столбца.

Теперь можно рассчитать фактическое (наблюдаемое) значение t-критерия для свободного члена a в уравнении регрессии. Найдем остаточную дисперсию на одну степень свободы (в числителе дроби под корнем сумма значений 9 столбца расчетной таблицы 1.1):

$$S_{ост} = \sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}} = \sqrt{\frac{9842,74}{15 - 2}} = 27,52.$$

Случайная ошибка свободного члена уравнения регрессии:

$$m_a = S_{ост} \frac{\sqrt{\sum x^2}}{n \sigma_x} = 27,52 \cdot \frac{\sqrt{7276,31}}{15 \cdot 14,28} = 10,96.$$

Значение t-статистики для параметра a :

$$t_a = \frac{|a|}{m_a} = \frac{39,31}{10,96} = 3,59.$$

Табличное значение t-критерия для уровня значимости $\alpha = 0,05$ и числа степеней свободы $k = n - 2 = 13$ равно 2,1604 (см. приложение 5). Поскольку $t_{табл} = 2,1604 < 3,59 = t_{факт}$, то значение свободного члена уравнения регрессии статически значимо отличается от нуля.

7. Определим прогнозное значение результативного признака, если возможное значение факторного признака составит 1,3 от его среднего уровня по совокупности.

Прогнозное значение результативного признака определяется путем подстановки в уравнение регрессии возможного значения факторного признака $x_p = \bar{x} \cdot 1,3 = 16,77 \cdot 1,3 = 21,8$.

Прогнозное значение оборота розничной торговли составит:

$$y = 10,1 \cdot 21,8 - 39,31 = 180,94.$$

Значит, при численности населения в регионе в 2,18 млн. человек возможный оборот розничной торговли в регионе составит 180,94 млрд. руб.¹

¹ Для прогнозного значения результативного признака корректнее рассчитать доверительные интервалы. См. Елисеева И.И. Практикум по эконометрике. – М.: Финансы и статистика, 2005. – С. 9.

Задание для самостоятельной работы

Имеются статистические данные (приложение №1) по некоторым регионам Российской Федерации. Для факторной переменной X (численность населения, столбец 0) и результативной переменной Y (столбец N , где N – номер варианта) требуется:

1. Рассчитать параметры уравнения регрессии $\hat{y} = a + bx$ методом наименьших квадратов.

2. Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

3. Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием F -критерия Фишера при уровне значимости $\alpha = 0,05$.

4. Оценить значимость коэффициентов регрессии и корреляции по t -критерию Стьюдента при уровне значимости $\alpha = 0,05$.

5. Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1,3 от его среднего уровня по совокупности.

Приложение 1

Статистические данные к теме «Линейная парная регрессия»

Все данные по регионам представлены за 2008 год²

№	Регион	0	1	2	3	4	5	6	7	8	9	10
1	Республика Адыгея	4,4	1,6	16	10,1	117	28,3	11	124	1,7	94	8,1
2	Республика Дагестан	27,1	5,7	173	7,6	513	3603,9	45	625	10,8	3109	24,6
3	Кабардино-Балкарская Республика	8,9	2,2	74	9,1	178	44,9	15	58	3,8	454	11,7
4	Республика Калмыкия	2,8	1,7	24	9,1	67	21,9	6	106	1,4	368	13,5
5	Карачаево-Черкесская Республика	4,3	2,0	32	9,4	112	17,1	8	111	1,6	44	6,4
6	Республика Северная Осетия - Алания	7,0	2,2	36	9,2	203	23,8	18	214	4,9	230	11,3
7	Чеченская Республика	12,4	1,7	168	11,8	276	0,2	8	138	3,0	397	9,0
8	Краснодарский край	51,4	13,9	126	13,2	1408	381,7	109	1469	22,0	3253	128,2
9	Ставропольский край	27,1	9,8	108	11,1	688	165,9	57	806	12,1	3091	55,2
10	Астраханская область	10,1	4,6	40	12,3	236	92,2	20	261	6,7	873	18,9
11	Волгоградская область	26,0	10,5	108	12,0	713	192,0	55	777	13,2	823	53,3
12	Ростовская область	42,4	17,2	147	12,5	1220	235,6	89	1256	16,2	2320	98,1
13	Ивановская область	10,7	4,3	29	10,2	324	126,2	25	379	5,5	185	28,9
14	Калужская область	10,0	4,9	26	14,1	291	110,0	24	265	4,0	135	26,3
15	Костромская область	6,9	4,8	19	11,5	205	90,1	16	365	2,5	1671	18,1

Расшифровка столбцов:

- 0 – Численность населения (оценка на конец года; сотни тысяч человек);
- 1 – Численность работников органов местного самоуправления (на конец года; тысяч человек);
- 2 – Численность безработных (тысяч человек);
- 3 – Среднемесячная номинальная начисленная заработная плата работников организаций (тысяч рублей);
- 4 – Численность пенсионеров (тысяч человек);
- 5 – Розничная продажа водки и ликероводочных изделий (десятки тысяч декалитров);
- 6 – Жилищный фонд (общая площадь жилых помещений; миллионов квадратных метров);
- 7 – Число дошкольных образовательных учреждений;
- 8 – Численность врачей всех специальностей (на конец года; тысяч человек);
- 9 – Использование свежей воды (миллионов кубических метров);
- 10 – Число предприятий и организаций (на конец года; тысяч).

² Источник: <http://www.gks.ru>

Продолжение таблицы

№	Регион	11	12	13	14	15	16	17	18	19	20	21
1	Республика Адыгея	16,1	9,6	21,2	11,9	3,1	6,7	54,3	41,6	33,2	21,8	80,8
2	Республика Дагестан	205,3	40,3	27,5	11,6	21,8	37,5	28,2	88,7	333,5	31,3	870,9
3	Кабардино-Балкарская Республика	27,4	19,5	29,7	19,6	12,5	16,6	67,2	35,1	192,0	6,9	296,3
4	Республика Калмыкия	5,5	9,7	29,2	22,5	0,9	1,5	43,7	10,6	10,0	0,3	18,0
5	Карачаево-Черкесская Республика	3,3	12,4	11,3	6,5	13,5	2,0	14,2	21,2	167,2	2,4	51,6
6	Республика Северная Осетия - Алания	8,9	14,2	16,8	12,2	7,5	3,1	52,1	10,1	130,4	6,6	32,8
7	Чеченская Республика	1,5	8,5	18,4	11,9	3,0	5,0	18,2	0,2	18,8	3,7	21,4
8	Краснодарский край	339,6	185,3	368,9	224,1	52,9	57,0	1163,4	318,0	512,0	81,0	613,8
9	Ставропольский край	161,0	76,4	294,5	225,5	25,9	18,1	841,3	183,4	258,8	35,7	199,5
10	Астраханская область	78,7	15,2	7,2	2,4	8,7	18,0	4,0	38,6	151,7	10,4	479,4
11	Волгоградская область	127,0	70,7	316,4	214,9	32,4	29,8	518,3	124,2	361,3	18,2	713,9
12	Ростовская область	209,3	115,8	455,2	286,6	36,0	38,4	887,1	213,7	380,2	75,0	552,9
13	Ивановская область	67,2	8,9	23,9	7,3	10,9	3,4	10,1	47,8	108,0	13,4	87,7
14	Калужская область	59,7	18,4	34,1	9,8	23,6	5,2	19,1	67,5	323,1	20,6	105,7
15	Костромская область	31,0	12,2	25,9	7,5	10,9	3,7	8,3	30,1	160,2	8,1	105,2

Расшифровка столбцов:

- 11 – Производство хлеба и хлебобулочных изделий (тысяч тонн);
12 – Продукция сельского хозяйства (в хозяйствах всех категорий; в фактически действовавших ценах; млрд. рублей);
13 – Посевные площади всех сельскохозяйственных культур (в хозяйствах всех категорий; десятки тысяч гектаров);
14 – Посевные площади зерновых и зернобобовых культур (в хозяйствах всех категорий; десятки тысяч гектаров);
15 – Посевные площади картофеля (в хозяйствах всех категорий; тысяч гектаров);
16 – Посевные площади овощей (в хозяйствах всех категорий; тысяч гектаров);
17 – Валовой сбор зерна (в весе после доработки) (в хозяйствах всех категорий; десятки тысяч тонн);
18 – Розничная продажа коньяков (тысяч декалитров);
19 – Валовой сбор картофеля (в хозяйствах всех категорий; тысяч тонн);
20 – Ввоз бензина автомобильного (десятки тысяч тонн);
21 – Валовой сбор овощей (в хозяйствах всех категорий; тысяч тонн).

Продолжение таблицы

№	Регион	22	23	24	25	26	27	28	29	30	31
1	Республика Адыгея	6,3	15,4	4,9	10,4	5,3	6,5	0,8	0,5	70	25,6
2	Республика Дагестан	52,3	96,7	90,6	52,6	43,8	45,7	9,1	2,7	523	217,3
3	Кабардино-Балкарская Республика	2,4	71,0	22,9	31,2	10,8	6,5	2,5	1,3	253	48,0
4	Республика Калмыкия	0,9	0,5	33,2	16,0	2,6	2,6	0,6	0,1	91	7,4
5	Карачаево-Черкесская Республика	0,9	6,7	20,5	22,1	6,1	5,9	0,9	2,2	121	22,6
6	Республика Северная Осетия - Алания	2,0	11,4	12,9	18,6	12,5	11,4	2,0	2,0	166	37,4
7	Чеченская Республика	22,1	11,4	23,4	25,9	6,6	28,0	0,6	2,2	173	25,2
8	Краснодарский край	77,4	256,1	68,7	136,8	137,9	182,3	39,4	32,9	1053	500,7
9	Ставропольский край	14,2	41,1	38,7	61,1	57,2	31,1	10,7	10,6	401	203,6
10	Астраханская область	37,2	11,2	23,4	15,3	18,5	28,8	3,8	8,0	268	77,7
11	Волгоградская область	28,6	143,6	31,7	47,9	71,0	55,7	8,2	17,6	373	186,1
12	Ростовская область	38,6	102,2	58,9	99,6	88,1	92,9	20,1	27,4	1336	423,4
13	Ивановская область	1,6	15,0	8,8	18,0	16,8	12,1	1,7	2,1	257	57,9
14	Калужская область	1,5	17,2	12,8	22,6	21,0	23,5	6,3	2,1	989	80,7
15	Костромская область	1,0	10,6	8,3	15,7	10,1	9,7	1,5	2,2	17	38,1

Расшифровка столбцов:

22 – Ввоз пиломатериалов (тысяч кубических метров);

23 – Валовой сбор плодов и ягод (в хозяйствах всех категорий; тысяч тонн);

24 – поголовье крупного рогатого скота (в хозяйствах всех категорий; на конец года; десятки тысяч голов);

25 – Производство молока (в хозяйствах всех категорий; десятки тысяч тонн);

26 – Объем платных услуг населению (миллиардов рублей);

27 – Объем работ, выполненных по виду экономической деятельности «Строительство» (в фактически действовавших ценах; млрд. рублей);

28 – Ввод в действие жилых домов (десятки тысяч квадратных метров общей площади);

29 – Отправление грузов железнодорожным транспортом общего пользования (миллионов тонн);

30 – Численность исследователей с учеными степенями (человек);

31 – Оборот розничной торговли (в фактически действовавших ценах; млрд. рублей).

Приложение 4

Таблица значений F-критерия Фишера при уровне значимости $\alpha = 0,05$

$k_1 \backslash k_2$	1	2	3	4	5	6	8	12	24	∞
1	161,5	199,5	215,7	224,6	230,2	233,9	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26

Приложение 5

Критические значения t -критерия Стьюдента при уровне значимости 0,10, 0,05, 0,01 (двухсторонний)

Число степеней свободы d.f.	α			Число степеней свободы d.f.	α		
	00,10	0,05	0,01		00,10	0,05	0,01
1	6,3138	12,706	63,657	18	1,7341	2,1009	2,8784
2	2,9200	4,3027	9,9248	19	1,7291	2,0930	2,8609
3	2,3534	3,1825	5,8409	20	1,7247	2,0860	2,8453
4	2,1318	2,7764	4,5041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,4995	24	1,7109	2,0639	2,7969
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7530	2,1315	2,9467	60	1,6707	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6577	1,9799	2,6174
17	1,7396	2,1098	2,8982	∞	1,6449	1,9600	2,5758