

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

**ИНСТИТУТ ТЕХНОЛОГИЙ (ФИЛИАЛ) ФЕДЕРАЛЬНОГО
ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО ОБРАЗОВАТЕЛЬНОГО**

УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ

**«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ» В Г. ВОЛГОДОНСКЕ РОСТОВСКОЙ
ОБЛАСТИ**

(Институт технологий (филиал) ДГТУ в г. Волгодонске)

Кафедра «Технический сервис и информационные технологии»

Теория систем и системный анализ

Практикум

Для обучающихся по направлению 09.03.02

Профиль Информационные системы

г. Волгодонск
2021

Практическая работа №1

По 12 территориям России за 200X г. известны значения двух признаков (табл. 2.1).

Таблица 2.1

Район	Средняя заработная плата работающего, тыс. руб., x	Доля денежных средств, направляемых на прирост сбережений во вкладах, %, y
Респ. Карелия	1	2,01
Архангельская обл.	2	2,83
Вологодская обл.	3	3,46
Новгородская обл.	4	3,99
Ростовская обл.	5	4,47
Орловская обл.	6	4,89
Ярославская обл.	7	5,29
Калужская обл.	8	5,65
Ленинградская обл.	9	6,01
Мурманская обл.	10	6,32
Респ. Коми	11	6,63
Московская обл.	12	6,92

Задание

- Для характеристики зависимости y от x рассчитать параметры следующих функций: 1) линейной; 2) степенной; 3) показательной; 4) равноугольной гиперболы.
- Оценить каждую модель через среднюю ошибку аппроксимации, показатель корреляции и F-критерий Фишера. Выбрать лучшую модель.

Решение

- Для расчета параметров a и b линейной регрессии $y = a + bx$ составим следующую таблицу (табл. 2.2).

Таблица 2.2

Район	x	y	$x * y$	x^2	y^2	\hat{y}_x	A_i
Респ. Карелия	1	2,01	2,01	1	4,0401	2,521154	25,4305
Архангельская обл.	2	2,83	5,66	4	8,0089	2,948671	4,19333
Вологодская обл.	3	3,46	10,38	9	11,9716	3,376189	2,42229
Новгородская обл.	4	3,99	15,96	16	15,9201	3,803706	4,66902
Ростовская обл.	5	4,47	22,35	25	19,9809	4,231224	5,34175
Орловская обл.	6	4,89	29,34	36	23,9121	4,658741	4,72922
Ярославская обл.	7	5,29	37,03	49	27,9841	5,086259	3,85144
Калужская обл.	8	5,65	45,2	64	31,9225	5,513776	2,41104
Ленинградская обл.	9	6,01	54,09	81	36,1201	5,941294	1,1432

Мурманская обл.	10	6,32	63,2	100	39,9424	6,368811	0,77233
Респ. Коми	11	6,63	72,93	121	43,9569	6,796329	2,50873
Московская обл.	12	6,92	83,04	144	47,8864	7,223846	4,39084
ИТОГО	78	58,47	441,19	650	311,646		61,8637
Среднее	6,5	4,8725	36,7658	54,167	25,9705		5,15531
Дисперсия	11,9167	2,2293					
Ср.кв.отклонение	3,4521	1,4931					

Напомним, что средние значения рассчитываются по формулам $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$,

$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, $\overline{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i$, $\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$ и т.д., где $n=12$ (число наблюдений в рассматриваемой задаче). Дисперсия определяется по формулам $\sigma_x^2 = \overline{x^2} - \bar{x}^2$, $\sigma_y^2 = \overline{y^2} - \bar{y}^2$, а среднеквадратическое отклонение есть корень квадратный из дисперсии.

Рассчитаем параметры a и b линейной регрессии $y = a + bx$.

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = \frac{36,7658 - 6,5 \times 4,8725}{11,9167} = 0,4275,$$

$$a = \bar{y} - b \cdot \bar{x} = 4,8725 - 0,4275 \times 6,5 = 2,0936.$$

Уравнение линейной регрессии имеет вид $\hat{y}_x = 2,0936 + 0,4275x$.

Рассчитаем линейный коэффициент парной корреляции:

$$r_{xy} = b \frac{\sigma_x}{\sigma_y} = 0,4275 \frac{3,4521}{1,4931} = 0,9884.$$

Связь сильная (так как $r_{xy} > 0,7$), прямая (так как $r_{xy} > 0$).

Коэффициент детерминации $r_{xy}^2 = (0,9884)^2 = 0,9770$, значит, вариация результата на 97,7 % объясняется вариацией фактора x , а на 2,3 % – вариацией неучтенных в модели признаков. Подставляя в уравнение регрессии фактические значения x , определим теоретические (расчетные) значения \hat{y}_x . Найдем величину средней ошибки аппроксимации:

$$\bar{A} = \frac{1}{n} \sum A_i = \frac{1}{n} \sum \left| \frac{y - \hat{y}_x}{y} \right| \times 100 \% = 5,15531 \% \approx 5,2 \%.$$

В среднем, расчетные значения от фактических отклоняются на 5,2 %, это говорит об удовлетворительном качестве модели, поскольку средняя ошибка аппроксимации не превышает 8–10 %.

Оценим теперь статистическую надежность полученной модели с помощью F -критерия Фишера. Выдвигаем гипотезу H_0 о статистической незначимости уравнения регрессии и показателя тесноты связи:

$$F_{\text{факт}} = \frac{r_{xy}^2}{1 - r_{xy}^2} (n - 2) = \frac{0,997}{1 - 0,997} (12 - 2) = 3323,34.$$

Определяем по таблице значений F -критерия Фишера $F_{\text{табл}}(0,05; k_1 = 1; k_2 = 10) = 4,96$. Так как $F_{\text{табл}} < F_{\text{факт}}$, то гипотеза о случайной природе оцениваемых характеристик отклоняется и признается их статистическая значимость и надежность.

2. Построению степенной модели $y = ax^b$ предшествует процедура линеаризации переменных. Проведем линеаризацию путем логарифмирования обеих частей уравнения:

$$\lg y = \lg a + b \lg x;$$

$$Y = C + bX, \text{ где } Y = \lg y, C = \lg a, X = \lg x.$$

Для расчетов используем данные таблицы 2.3.

Таблица 2.3

Район	X	Y	X^2	XY	\hat{y}_x	$y - \hat{y}_x$	A_i	$(y - \hat{y}_x)^2$
Респ. Карелия	0	0,303	0	0	2,0044	0,0056	0,2778	0,0000311745
Архангельская обл.	0,301	0,452	0,091	0,136	2,8318	-0,0018	0,0639	0,0000032712
Вологодская обл.	0,477	0,539	0,228	0,257	3,4662	-0,0062	0,179	0,0000383477
Новгородская обл.	0,602	0,601	0,362	0,362	4,0007	-0,0107	0,2691	0,0001152464
Ростовская обл.	0,699	0,65	0,489	0,455	4,4715	-0,0015	0,0336	0,0000022573
Орловская обл.	0,778	0,689	0,606	0,536	4,897	-0,007	0,1428	0,0000487626
Ярославская обл.	0,845	0,723	0,714	0,611	5,2882	0,0018	0,0348	0,0000033965
Калужская обл.	0,903	0,752	0,816	0,679	5,6522	-0,0022	0,0385	0,0000047380
Ленинградская обл.	0,954	0,779	0,911	0,743	5,994	0,016	0,2661	0,0002557166
Мурманская обл.	1	0,801	1	0,801	6,3173	0,0027	0,0432	0,0000074573
Респ. Коми	1,041	0,822	1,084	0,856	6,6247	0,0053	0,0801	0,0000282301
Московская обл.	1,079	0,84	1,165	0,907	6,9184	0,0016	0,0234	0,0000026192
ИТОГО	8,68	7,951	7,464	6,343			1,4523	0,0005412173
Среднее	0,723	0,663	0,622	0,529			0,121	0,0000451014
Дисперсия	0,099	0,025						

Рассчитаем C и b :

$$b = \frac{\overline{XY} - \bar{X} \cdot \bar{Y}}{\sigma_X^2} = \frac{0,529 - 0,723 \cdot 0,663}{0,099} = 0,499,$$

$$C = \bar{Y} - b \cdot \bar{X} = 0,663 - 0,499 \times 0,723 = 0,302.$$

Получим линейное уравнение $y = 0,302 + 0,499X$.

Выполнив его потенцирование, получим:

$$\hat{y} = 10^{0,302} \cdot x^{0,499} = 2,004x^{0,499}.$$

Подставляя в уравнение регрессии фактические значения x , определим теоретические (расчетные) значения \hat{y}_x . По ним рассчитаем показатели: тесноты связи – индекс корреляции ρ_{xy} и среднюю ошибку аппроксимации:

$$\rho_{xy} = \sqrt{1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2}} = \sqrt{1 - \frac{\sum(y - \hat{y}_x)^2 / n}{y^2 - \bar{y}^2}} = \sqrt{1 - \frac{0,0000451014}{2,2293}} = 0,99999.$$

Связь сильная. Коэффициент детерминации $\rho_{xy}^2 = 0,99998$, значит, вариация результата на 99,998 % объясняется вариацией фактора x , а на 0,002 % – вариацией не учтенных в модели признаков.

Величина средней ошибки аппроксимации $\bar{A} = 0,121\%$ говорит об удовлетворительном качестве модели, поскольку \bar{A} не превышает 8–10 %.

Оценим теперь статистическую надежность полученной модели с помощью F -критерия Фишера. Выдвигаем гипотезу H_0 о статистической незначимости уравнения регрессии и показателя тесноты связи.

$$F_{\text{факт}} = \frac{\rho_{xy}^2}{1 - \rho_{xy}^2} (n - 2) = \frac{0,99998}{1 - 0,99998} (12 - 2) = 499\,990.$$

$$F_{\text{табл}}(0,05; k_1 = 1; k_2 = 10) = 4,96.$$

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

3. Построению показательной модели $y = ab^x$ предшествует процедура линеаризации переменных. Проведем линеаризацию путем логарифмирования обеих частей уравнения:

$$\lg y = \lg a + x \lg b;$$

$$Y = C + Bx, \text{ где } Y = \lg y, C = \lg A, B = \lg b.$$

Для расчетов используем данные таблицы 2.4.

Таблица 2.4

Район	x	Y	Yx	\hat{y}_x	$y - \hat{y}_x$	A_i	$(y - \hat{y}_x)^2$
Респ. Карелия	1	0,303	0,303196	2,668038	-0,65804	32,73823	0,433015
Архангельская обл.	2	0,452	0,903573	2,945627	-0,11563	4,085775	0,01337
Вологодская обл.	3	0,539	1,617228	3,252097	0,207903	6,008746	0,043223
Новгородская обл.	4	0,601	2,403892	3,590453	0,399547	10,0137	0,159638
Ростовская обл.	5	0,65	3,251538	3,964012	0,505988	11,31963	0,256023
Орловская обл.	6	0,689	4,135853	4,376438	0,513562	10,5023	0,263746
Ярославская обл.	7	0,723	5,06419	4,831772	0,458228	8,662149	0,209973
Калужская обл.	8	0,752	6,016388	5,334481	0,315519	5,584403	0,099552
Ленинградская обл.	9	0,779	7,00987	5,889493	0,120507	2,005106	0,014522
Мурманская обл.	10	0,801	8,007171	6,50225	-0,18225	2,883699	0,033215
Респ. Коми	11	0,822	9,036649	7,178759	-0,54876	8,276909	0,301137
Московская. обл.	12	0,84	10,08127	7,925654	-1,00565	14,53257	1,01134
ИТОГО	78	7,951	57,83082			116,6132	2,838753
Среднее	6,5	0,663	4,819235			9,717769	0,236563
Дисперсия	11,9167	0,025					

Рассчитаем C и B :

$$b = \frac{\overline{xY} - \bar{x} \cdot \bar{Y}}{\sigma_x^2} = \frac{4,819235 - 6,5 \cdot 0,663}{11,9167} = 0,04299,$$

$$C = \bar{Y} - b \cdot \bar{x} = 0,663 - 0,04299 \times 6,5 = 0,38321.$$

Получим линейное уравнение $y = 0,38321 + 0,04299x$.

Выполнив его потенцирование, получим:

$$\hat{y} = 10^{0,38321} \cdot 10^{0,04299x} = 2,416609 \cdot 1,104042^x.$$

Подставляя в уравнение регрессии фактические значения x , определим теоретические (расчетные) значения \hat{y}_x . По ним рассчитаем показатели: тесноты связи – индекс корреляции ρ_{xy} и среднюю ошибку аппроксимации:

$$\rho_{xy} = \sqrt{1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2}} = \sqrt{1 - \frac{\sum(y - \hat{y}_x)^2/n}{y^2 - \bar{y}^2}} = \sqrt{1 - \frac{0,236563}{2,2293}} = 0,945454.$$

Связь сильная. Коэффициент детерминации $\rho_{xy}^2 = 0,89$, значит, вариация результата на 89 % объясняется вариацией фактора x , а на 11 % – вариацией неучтенных в модели признаков.

Величина средней ошибки аппроксимации $\bar{A} = 9,717769\%$ говорит о не очень хорошем, но все же удовлетворительном качестве модели, поскольку \bar{A} близко к 10 %.

Оценим теперь статистическую надежность полученной модели с помощью F -критерия Фишера. Выдвигаем гипотезу H_0 о статистической незначимости уравнения регрессии и показателя тесноты связи.

$$F_{\text{факт}} = \frac{\rho_{xy}^2}{1 - \rho_{xy}^2} (n - 2) = \frac{0,89}{1 - 0,89} (12 - 2) = 80,91.$$

$$F_{\text{табл}}(0,05; k_1 = 1; k_2 = 10) = 4,96.$$

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

4. Уравнение равносторонней гиперболы $y = a + \frac{b}{x}$ линеаризуется при замене

$$z = \frac{1}{x}. \text{ Тогда } y = a + bz.$$

Для расчетов используем данные таблицы 2.5.

Таблица 2.5

Район	y	z	z^2	yz	\hat{y}_x	$y - \hat{y}_x$	A_i	$(y - \hat{y}_x)^2$
Респ. Карелия	2,01	1	1	2,01	1,085765	0,924235	45,98182	0,85421
Арханг. обл.	2,83	0,5	0,25	1,415	3,639541	-0,80954	28,60571	0,655357
Вологод. обл.	3,46	0,33	0,111111	1,153333	4,4908	-1,0308	29,79191	1,062549
Новгород. обл.	3,99	0,25	0,0625	0,9975	4,916429	-0,92643	23,21878	0,858272
Рост. обл.	4,47	0,2	0,04	0,894	5,171807	-0,70181	15,70038	0,492533
Орлов. обл.	4,89	0,17	0,027778	0,815	5,342059	-0,45206	9,244557	0,204357

Ярослав. обл.	5,29	0,14	0,020408	0,755714	5,463667	-0,17367	3,282934	0,03016
Калужск. обл.	5,65	0,13	0,015625	0,70625	5,554874	0,095126	1,683655	0,009049
Ленингр. обл.	6,01	0,11	0,012346	0,667778	5,625812	0,384188	6,392484	0,147601
Мурман. обл.	6,32	0,1	0,01	0,632	5,682562	0,637438	10,08604	0,406327
Респ. Коми	6,63	0,090909	0,008264	0,602727	5,728995	0,901005	13,58983	0,811811
Москов. обл.	6,92	0,083333	0,006944	0,576667	5,767688	1,152312	16,6519	1,327823
Среднее	4,8725	0,258601	0,130415	0,935497			17,01917	0,571671
Дисперсия	2,229252	0,06354						

Рассчитаем параметры a и b .

$$b = \frac{\bar{zy} - \bar{z} \cdot \bar{y}}{\sigma_z^2} = \frac{11,22597 - 0,258601 \cdot 4,8725}{0,06354} = -5,10755,$$

$$a = \bar{y} - b \cdot \bar{z} = 4,8725 + 5,10755 \times 0,258601 = 6,193318.$$

Получено уравнение $\hat{y} = 6,193318 - \frac{5,10755}{x}$.

Подставляя в уравнение регрессии фактические значения x , определим теоретические (расчетные) значения \hat{y}_x . По ним рассчитаем показатели: тесноты связи – индекс корреляции ρ_{xy} и среднюю ошибку аппроксимации:

$$\rho_{xy} = \sqrt{1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2}} = \sqrt{1 - \frac{\sum(y - \hat{y}_x)^2 / n}{y^2 - \bar{y}^2}} = \sqrt{1 - \frac{0,571671}{2,2293}} = 0,862299.$$

Связь сильная. Коэффициент детерминации $\rho_{xy}^2 = 0,74$, значит, вариация результата на 74 % объясняется – вариацией фактора x , а на 26 % – вариацией неучтенных в модели признаков.

Величина средней ошибки аппроксимации $\bar{A} = 17,01917\%$ говорит о неудовлетворительном качестве модели, поскольку \bar{A} превышает 10 %.

Оценим теперь статистическую надежность полученной модели с помощью F -критерия Фишера. Выдвигаем гипотезу H_0 о статистической незначимости уравнения регрессии и показателя тесноты связи.

$$F_{\text{факт}} = \frac{\rho_{xy}^2}{1 - \rho_{xy}^2} (n - 2) = \frac{0,74}{1 - 0,74} (12 - 2) = 28,46.$$

$$F_{\text{табл}}(0,05; k_1 = 1; k_2 = 10) = 4,96.$$

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

Проведенные исследования показывают, что рассматриваемую в задаче зависимость лучше всего описывает степенная кривая, поскольку для этой модели показатель корреляции оказался больше, а ошибка аппроксимации меньше, чем для других построенных функций, при этом уравнение степенной регрессии статистически надёжно.

ЛАБОРАТОРНАЯ РАБОТА №2

1. Один хороший студент имеет возможность повторить материал в ночь перед экзаменом. Но накануне он столкнулся с небольшой проблемой. Перед завтрашним экзаменом его сокурсники организовали вечеринку, в которой он не хочет участвовать. Студент имеет альтернативы:

a_1 – участвовать в вечеринке всю ночь,

a_2 – половину ночи участвовать в вечеринке, а половину – учиться,

a_3 – учиться всю ночь.

Экзамен может быть легким (s_1), средним (s_2), трудным (s_3) в зависимости от настроения преподавателя. Ожидаемые баллы, полученные на экзамене студентом:

	s_1	s_2	s_3
a_1	85	60	40
a_2	92	85	81
a_3	100	88	82

Какой выбор должен сделать студент (использовать все четыре критерия)?

Пусть студент заинтересован в оценке, которую он получит на экзамене: 90, 80, 70 и 60 баллов. Если он получит оценку ниже 60 баллов, то считается, что экзамен не сдан. Изменит ли такое отношение к оценкам выбор студента?

ЛАБОРАТОРНАЯ РАБОТА №3

По 14 территориям Уральского и Западно-Сибирского района по данным за ноябрь 2000 г. изучалась зависимость потребительских расходов на душу населения (тыс.руб., y) от средней заработной платы и выплат социального характера (тыс.руб., x). Было получено следующее уравнение линейной регрессии $\hat{y}_x = 0,2491x + 192,71$, $r_{xy}^2 = 0,8556$. Известны следующие данные $\sum (y - \hat{y}_x)^2 = 46993,97$, $\overline{x^2} = 1815993$, $\bar{x} = 1222,86$.

Задание.

1. Оценить статистическую значимость параметров регрессии и корреляции.
2. Выполнить прогноз потребительских расходов на душу населения, если значение средней заработной платы и выплат социального характера увеличится на 6% от своего среднего уровня.
3. Определить доверительный интервал прогноза для уровня значимости $\alpha = 0,05$, рассчитать ошибку прогноза.

Решение.

1. Оценку статистической значимости параметров регрессии проведем с помощью t-статистики Стьюдента и путем расчета доверительного интервала каждого из показателей.

Выдвигаем гипотезу H_0 о статистически незначимом отличии от нуля показателей a, b и r_{xy} .

По таблице определяем значение $t_{табл}$ для уровня значимости $\alpha = 0,05$ и числа степеней свободы $df = n - 2 = 14 - 2 = 12$ $t_{табл} = 2,1788$.

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = 320613,7576.$$

Определим случайные ошибки показателей a, b и r_{xy} .

$$m_a = \sqrt{\frac{\sum (y - \hat{y}_x)^2 \overline{x^2}}{(n-2)n\sigma_x^2}} = \sqrt{\frac{46993,97 \cdot 1815993}{12 \cdot 14 \cdot 320613,7576}} = 39,80.$$

$$m_b = \sqrt{\frac{\sum (y - \hat{y}_x)^2}{(n-2)n\sigma_x^2}} = \sqrt{\frac{46993,97}{12 \cdot 14 \cdot 320613,7576}} = 0,03.$$

$$m_r = \sqrt{\frac{1 - r_{xy}^2}{n-2}} = \sqrt{\frac{1 - 0,8556}{14-2}} = 0,12.$$

Тогда фактические значения t-статистики составят:

$$t_a = \frac{a}{m_a} = \frac{192,71}{39,80} = 4,84, \quad t_b = \frac{b}{m_b} = \frac{0,2491}{0,03} = 8,3, \quad t_r = \frac{r_{xy}}{m_r} = \frac{0,925}{0,12} = 7,7.$$

Фактические значения t-статистики превосходят табличные значения $t_a = 4,84 > t_{табл} = 2,1788$, $t_b = 8,3 > t_{табл} = 2,1788$, $t_r = 7,7 > t_{табл} = 2,1788$, поэтому гипотеза H_0 отклоняется, т.е. a, b и r_{xy} не случайно отличаются от нуля, а статистически значимы.

Рассчитаем доверительные интервалы для a, b . Для этого определим предельную ошибку каждого из показателей:

$$\Delta_a = t_{табл} \cdot m_a = 2,1788 \cdot 39,8 = 86,48, \quad \Delta_b = t_{табл} \cdot m_b = 2,1788 \cdot 0,03 = 0,65.$$

Доверительные интервалы:

$$\begin{aligned} \gamma_a &= a \pm \Delta_a = 192,71 \pm 86,48, \\ \gamma_{a_{\min}} &= 192,71 - 86,48 = 106,23, \quad \gamma_{a_{\max}} = 192,71 + 86,48 = 279,19; \\ \gamma_b &= b \pm \Delta_b = 0,2491 \pm 0,03, \\ \gamma_{b_{\min}} &= 0,2491 - 0,03 = 0,2191, \quad \gamma_{b_{\max}} = 0,2491 + 0,03 = 0,2791. \end{aligned}$$

Анализ верхней и нижней границ доверительных интервалов приводит к выводу о том, что с вероятностью $p = 1 - \alpha = 0,95$ параметры a и b , находясь в указанных границах, не принимают нулевых значений, т.е. не являются статистически незначимыми и существенно отличны от нуля.

2. Полученные оценки уравнения регрессии позволяют использовать его для прогноза. Если прогнозное значение средней заработной платы и выплат социального характера составит $x_p = \bar{x} \cdot 1,06 = 1222,86 \cdot 1,06 = 1296,23$ тыс. руб., тогда прогнозное значение потребительских расходов на душу населения составит $\hat{y}_p = 0,2491 \cdot 1296,23 + 192,71 = 515,60$ тыс. руб.

3. Ошибка прогноза составит:

$$m_{\hat{y}_p} = \sqrt{\frac{\sum (y - \hat{y})_x^2}{n - 2}} \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{n\sigma_x^2}} = \sqrt{\frac{46993,97}{14 - 2}} \sqrt{1 + \frac{1}{14} + \frac{(1296,23 - 1222,86)^2}{14 \cdot 320613,7576}} = 64,81 \text{ тыс. руб.}$$

Предельная ошибка прогноза, которая в 95% случаев не будет превышена, составит:

$$\Delta_{\hat{y}_p} = t_{табл} \cdot m_{\hat{y}_p} = 2,1788 \cdot 64,81 = 141,21 \text{ тыс. руб.}$$

Доверительный интервал прогноза:

$$\begin{aligned} \gamma_{\hat{y}_p} &= \hat{y}_p \pm \Delta_{\hat{y}_p} = 515,60 \pm 141,21. \\ \gamma_{\hat{y}_{p_{\min}}} &= 515,60 - 141,21 = 374,39, \quad \gamma_{\hat{y}_{p_{\max}}} = 515,60 + 141,21 = 656,81. \end{aligned}$$

Выполненный прогноз потребительских расходов на душу населения оказался надежным ($p = 1 - \alpha = 0,95$), но неточным, так как диапазон верхней и нижней границ доверительного интервала составляет 1,75 раза:

$$D_\gamma = \frac{\gamma_{\hat{y}_{p_{\max}}}}{\gamma_{\hat{y}_{p_{\min}}}} = \frac{656,81}{374,39} = 1,75.$$