Факультет «Инженерный бизнес и менеджмент»
МГТУ им. Н.Э. Баумана

Домашнее задание по курсу лекций «Статистика». Часть 2 2016 год

Работа выполняется студентами на листах формата А4.

Первый лист – титульный. На нем указывается тема работы, учебная группа, номер варианта студента, а также Ф.И.О. студента и преподавателя, принимающего работу.

ВАЖНО! В работе во всех задачах необходимо приводить **расчетные формулы** и отражать **ход расчетов.**

В завершение каждого пункта задачи необходимо привести выводы по результатам проведенных расчетов.

Домашнее задание содержит двадцать восемь вариантов. Выбор варианта определяется порядковым номером студента в журнале учета успеваемости.

Исходные данные, необходимые для выполнения работ, даны в Таблице №3.

Задача 3. «Временные ряды. Метод наименьших квадратов (МНК)»

Исходные данные (выбранный вариант по Таблице №3) – набор n-пар чисел (t_k , x_k), k = 1,2,…,n, где t_k – независимая переменная (например, время), а $x(t_k)$ – зависимая (например, индекс инфляции). Предполагается, что переменные связаны зависимостью

$$x(t_k) = a\, t_k + b + e_k\,,\ k = 1,2,\ldots,n,$$

где a и b – параметры, не известные статистику и подлежащие оцениванию, а e_k – погрешности, искажающие зависимость.

Таблица для внесения исходных данных.

t_k						
x_k						

1. Методом наименьших квадратов оцените параметры a и b линейной зависимости. Выпишите восстановленную зависимость.

Например, по результатам вычислений оценки параметров a и b равны:

$a^* = 3{,}14;$ $\quad b^* = 9{,}03$, тогда восстановленная зависимость будет выглядеть: $x^*(t) = 3{,}14 \cdot t + 9{,}03$

---

2. Вычислите восстановленные значения зависимой переменной, сравните их с исходными значениями (найдите разности) и проверьте условие точности вычислений (при отсутствии ошибок в вычислениях сумма исходных значений должна равняться сумме восстановленных, или сумма попарных разностей будет равна нулю: $\sum_{k=1}^{n}(x^*(t_k) - x(t_k)) = 0;\ k = 1,2,\ldots,n$).

3. Найдите остаточную сумму квадратов и оцените дисперсию погрешностей: SS; $(\sigma^2)^*$.

4. Выпишите точечный прогноз, а также верхнюю и нижнюю доверительные границы для него (для доверительной вероятности 0,95).

Например: $x^*(t) = 3{,}14t + 9{,}03 \pm 1{,}96 \cdot 1{,}49\sqrt{\frac{1}{6} + \frac{(t - 5{,}67)^2}{63{,}1}}$

---

5. Рассчитайте прогнозное значение и доверительные границы для него для момента $t =$ __ (см. вариант по Таблице № 3 – столбец «$t_{прогн.}$»).

Пример: точечный прогноз будет выглядеть для t = 12:

$$x^*(12) = 3{,}14 \cdot 12 + 9{,}03 \pm 1{,}96 \cdot 1{,}49\sqrt{\frac{1}{6} + \frac{(12 - 5{,}67)^2}{63{,}1}} = 46{,}71 \pm 2{,}615.$$

---

6. Как изменятся результаты, если доверительная вероятность будет увеличена? А если она будет уменьшена?

Таблица № 3. Исходные данные к Задаче №3.

№ вар.	Обозна-чения	Значения						$t_{прогн.}$
1.	t_k	1	2	4	5	7	10	12
	x_k	11	17	20	26	33	42	
2.	t_k	3	4	6	7	8	10	12
	x_k	10	12	16	22	26	30	
3.	t_k	1	4	6	8	9	10	12
	x_k	12	20	28	33	35	40	
4.	t_k	2	4	5	6	8	9	12
	x_k	17	24	27	28	33	35	
5.	t_k	2	4	5	7	8	10	12
	x_k	11	15	21	26	28	32	
6.	t_k	2	3	5	6	8	10	12
	x_k	10	12	18	25	30	42	
7.	t_k	1	3	5	6	7	10	12
	x_k	12	18	26	28	32	42	
8.	t_k	1	3	4	5	7	10	12
	x_k	12	16	20	26	34	40	
9.	t_k	2	3	5	6	7	10	12
	x_k	17	18	26	28	32	42	
10.	t_k	3	4	6	7	8	10	12
	x_k	10	15	19	24	31	36	
11.	t_k	3	5	6	7	9	10	12
	x_k	11	17	22	28	32	38	
12.	t_k	3	4	5	8	9	10	12
	x_k	11	13	17	23	27	32	
13.	t_k	10	25	30	40	55	70	100
	x_k	90	130	180	205	195	210	
14.	t_k	30	40	45	55	70	80	100
	x_k	95	128	125	138	145	170	

№ вар.	Обозна-чения	Значения						$t_{прогн.}$
15.	t_k	25	30	35	50	70	75	100
	x_k	100	90	130	140	180	210	
16.	t_k	25	38	45	60	72	77	100
	x_k	90	95	115	133	142	160	
17.	t_k	3	5	7	8	9	10	12
	x_k	19	26	32	33	35	42	
18.	t_k	15	30	40	55	60	80	100
	x_k	100	130	135	140	150	170	
19.	t_k	5	25	30	35	55	70	100
	x_k	75	100	90	130	200	180	
20.	t_k	20	28	35	50	62	75	100
	x_k	83	95	119	127	140	165	
21.	t_k	15	27	40	55	68	75	100
	x_k	92	100	108	125	135	160	
22.	t_k	20	27	40	45	60	80	100
	x_k	110	125	140	150	155	180	
23.	t_k	15	28	40	58	60	80	100
	x_k	100	120	125	140	160	165	
24.	t_k	18	25	40	50	67	75	100
	x_k	78	85	102	115	140	162	
25.	t_k	10	12	21	35	50	70	100
	x_k	128	147	168	190	210	222	
26.	t_k	20	27	50	60	75	80	100
	x_k	100	95	115	125	150	157	
27.	t_k	5	30	35	50	70	75	100
	x_k	75	90	130	140	180	210	
28.	t_k	25	30	35	50	55	70	100
	x_k	100	90	130	140	200	180	

Основные расчетные формулы

1) Оценки параметров линейного тренда a и b:

$$a^* = \frac{\overline{xt} - \bar{x} \cdot \bar{t}}{\overline{t^2} - (\bar{t})^2}$$

$$b^* = \bar{x} - a \cdot \bar{t}$$

2) Критерий оценки точности приближения функции по МНК:

$$\sum_{k=1}^{n} (x^*(t_k) - x(t_k)) = 0, \qquad k = 1,2,\ldots,n$$

3) Остаточная сумма квадратов:

$$SS = \sum_{k=1}^{n} (x^*(t_k) - x(t_k))^2 = \sum SS_k = \sum_{k=1}^{n} ((a^* - a) \cdot t_k + (b^* - b) - e_k)^2$$

4) Состоятельная оценка остаточной дисперсии:

$$(\sigma^2)^* = \frac{SS}{n}$$

5) Доверительные границы для прогностической функции:

$$x_{верх}(t) = a^* t + b^* + \delta(t),$$

$$x_{нижн}(t) = a^* t + b^* - \delta(t),$$

где погрешность $\delta(t)$ имеет вид:

$$\delta(t) = U(p) \cdot \sigma^* \cdot \sqrt{\frac{1}{n} + \frac{(t - \bar{t})^2}{\sum_{k=1}^{n} (t_k - \bar{t})^2}}, \quad \sigma^* = \sqrt{\frac{SS}{n}},$$

где t – момент времени для расчета прогнозного значения,

$\bar{t}$ – среднее значение показателя времени в исходном ряду,

σ^* – корень из оценки остаточной дисперсии,

а при доверительной вероятности $\gamma = 0{,}95$: $U(p) = 1{,}96$.