Лабораторная работа №1

Тема: "Построение регрессионных моделей"

Цель работы:

Изучить основные понятия регрессионных моделей, ознакомиться с методикой построения регрессионных моделей.

Основные понятия:

Решение задач с применением теории планирования эксперимента (ТПЭ) предусматривает использование информации об изучаемом процессе для выбора общей последовательности управления экспериментами, которая уточняется после очередного этапа проведения исследований на основе вновь полученных сведений. Тем самым достигается возможность рационального управления экспериментами при неполном первоначальном знании характеристик исследуемого объекта. Целесообразность применения тем выше, чем сложнее исследуемая система.

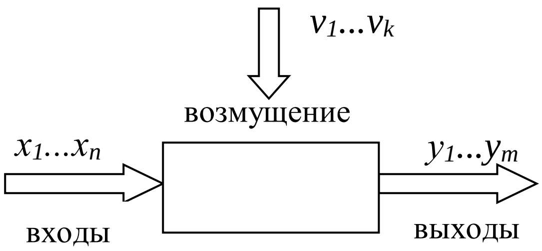

Рис.1 "черный ящик"

где $x_1 \ldots x_n$ – управляемые независимые параметры (факторы); $y_1 \ldots y_m$ – функции выхода (отклика), представляют собой реакции системы на воздействие факторов $x_1 \ldots x_n$; $v_1 \ldots v_k$ – случайные возмущающие воздействия.

Исследуемый объект (реальный объект, модель объекта) рассматривается как "черный ящик", представленный на рис.1. «Черный ящик» представляет собой сложную гомоморфную (неполную) модель кибернетической системы. Он только тогда является удовлетворительной моделью системы, когда содержит такое количество информации, которое отражает разнообразие системы.

Переменные $x_1...x_n$ принято называть факторами. Им можно сопоставить геометрическое понятие *факторного пространства* – пространство, координатные оси которого соответствуют значениям факторов.

Важнейшей задачей методов обработки, полученной в ходе эксперимента информации, является задача построения математической модели изучаемого явления, процесса, объекта. Под экспериментом понимается совокупность операций, совершаемых над объектом исследования с целью получения информации об его свойствах.

Хорошо спланированный эксперимент обеспечивает оптимальную обработку результатов, и, следовательно, возможность четких статистических выводов.

Цель планирования эксперимента – нахождение таких условий и правил проведения опытов, при которых удается получить надежную и достоверную информацию об объекте с наименьшей затратой труда, а также представить эту информацию в компактной и удобной форме с количественной оценкой точности.

Зависимость между случайными величинами называется *регрессией*. Она понимается как зависимость между математическими ожиданиями этих величин.

Построение регрессионных моделей – это многоступенчатый, итерационный процесс. Первая построенная модель в процессе статистического анализа, может оказаться не адекватной данным. Диагностика регрессионных моделей позволяет обнаружить несоответствие модели данным и наметить пути для дальнейшего улучшения построенной модели.

Существует два вида уравнения регрессии:

1. Простая (парная) регрессия, представляет собой модель, где среднее значение зависимой (объясняемой) переменной y рассматривается как функция одной независимой (объясняющей) переменной x.

В неявном виде парная регрессия – это модель вида:

$$y = f(x)$$

В явном виде:

$$y=a+b*x \qquad (1)$$

где x – зависимая переменная, y – независимая переменная, a – свободный член уравнения регрессии, b – коэффициент уравнения регрессии.

2. Множественная (зависит от множества факторов), представляет собой модель, где среднее значение зависимой (объясняемой) переменной y рассматривается как функция нескольких независимых (объясняющих) переменных $x_1, x_2, \ldots x_n$.

В неявном виде множественная регрессия – это модель вида:

$$y = f(x_1, x_2, x_3 \ldots x_n)$$

В явном виде:

$$y =a+b_1x1+b_2x_2+\ldots+b_nx_n$$

где $x_1, x_2, x_3 \ldots x_n$ – зависимые переменные, y – независимая переменная, a – свободный член уравнения регрессии, $b_{1\ldots}b_n$ – коэффициенты уравнения регрессии.

Большинство вероятностно-статистических моделей можно свести к парной регрессии, поэтому данная регрессия получила широкое распространение.

Построение регрессионных моделей может потребовать построение следующих моделей парной регрессии:

- Линейная регрессия: $y= a+b*x+ \varepsilon$
- Равносторонняя гипербола: $y = a + \frac{b}{x} + \varepsilon$
- Степенная регрессия: $y = a*x^b * \varepsilon$
- Показательная регрессия: $y = a+b^x * \varepsilon$
- Экспоненциальная регрессия: $y =e^{(a+b*x)}*\varepsilon$

где b – коэффициент, a – свободный член уравнения регрессии, ε – ошибка.

В случае линейной регрессии ошибка ε равна нулю.

В данной лабораторной работе рассматривается парная линейная регрессия.

Для того что бы рассчитать парную регрессию, необходимо провести ряд расчетов.

Дисперсия случайной величины σ_{x}^2 – мера разброса данной случайной величины, то есть её отклонения от математического ожидания.

$$\sigma_x^2 = \overline{x^2} - \overline{x}^2 \tag{2}$$

$$\sigma_y^2 = \overline{y^2} - \overline{y}^2 \tag{3}$$

Для нахождения дисперсии случайной величины понадобятся средние значения $\overline{x}$ и $\overline{y}$, которые рассчитываются следующим образом.

$$\overline{x} = \frac{\sum_1^n x}{N} \tag{4}$$

$$\overline{y} = \frac{\sum_1^n y}{N} \tag{5}$$

где N – число экспериментов

Для решения уравнения парной регрессии необходимо рассчитать свободный член a и коэффициент b.

$$b = \frac{\overline{xy} - \overline{x} * \overline{y}}{\sigma_x^2} \tag{6}$$

$$\overline{xy} = \frac{\sum_1^n x_i * y_i}{N} \tag{7}$$

$$a = \overline{y} - b\overline{x} \tag{8}$$

где i=1…N

Степень тесноты парной линейной зависимости определяет линейный коэффициент корреляции r_{xy}.

$$r_{xy} = \frac{\overline{\text{xy}} - \overline{\text{x}} * \overline{\text{y}}}{\sqrt{\sigma_{\text{x}}^2 * \sigma_{\text{y}}^2}} \tag{9}$$

Иногда показателю тесноты связи можно дать качественную оценку. Для оценки связи линейного коэффициента корреляции используется шкала Чеддока, представленная в приложении 1.

В дальнейшем, чтобы рассчитать границы доверительного интервала, необходимо знать стандартную ошибку остаточной компоненты – S_E.

$$S_{\text{E}} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \tag{10}$$

где N – число экспериментов, i – порядковый номер эксперимента, $i=1...N$, $\hat{y}_i$ рассчитывается, как

$$\hat{y}_i = a + bx_i \tag{11}$$

Прогноз, полученный подстановкой в уравнение регрессии ожидаемого значения фактора, называют точечным прогнозом. Вероятность точной реализации такого прогноза крайне мала. Необходимо сопроводить его значением *средней квадратичной ошибкой прогноза* или *доверительным интервалом прогноза*.

Найдем средние квадратичные (стандартные) ошибки оценивания коэффициента b и свободного члена a уравнения регрессии:

$$\delta_a = S_E * \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N * \sum_{i=1}^{n}(x_i - \overline{x})^2}} \tag{12}$$

$$\delta_b = \frac{(S_E)^2}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2}} \tag{13}$$

где δ_a – средняя квадратичная ошибка оценивания свободного члена a, δ_b – средняя квадратичная ошибка оценивания коэффициента b.

Для вычисления *доверительных границ прогноза* нужно умножить найденную среднюю квадратичную ошибку прогноза на t-критерий Стьюдента для заданной доверительной вероятности (обычно задают 95% или p=0,05). При малых объемах выборки интервал получается довольно широким.

t-критерий Стьюдента – общее название для класса методов статистической проверки гипотез (статистических критериев), основанных на распределении Стьюдента. Наиболее частые случаи применения t-критерия связаны с проверкой равенства средних значений в двух выборках. Для применения данного критерия необходимо, чтобы исходные данные имели нормальное распределение.

Требование нормальности распределения данных является необходимым для точного t-теста. Однако, даже при других распределениях данных возможно использование t-статистики. Во многих случаях эта статистика асимптотически имеет стандартное нормальное распределение, поэтому можно использовать квантили этого распределения. Однако, часто даже в этом случае используют квантили не стандартного нормального распределения, а соответствующего распределения Стьюдента, как в точном t-тесте. Асимптотически они эквивалентны, однако на малых выборках доверительные интервалы распределения Стьюдента шире и надежнее.

Рассчитаем значение критерия Стьюдента для свободного члена a и коэффициета b.

$$t_a = \frac{a}{\delta_a}\ ; \quad t_b = \frac{b}{\delta_b} \qquad (14,15)$$

где t_a – критерий Стьюдента для свободного члена a, t_b – критерий Стьюдента для коэффициента b.

С помощью таблицы, представленной в приложении 2, t-критерия Стьюдента находится критическое значение ($t_{табл}$) критерия для заданной степени свободы.

Для того, чтобы определить подходят ли модели свободный член a и коэффициент b необходимо сравнить критическое и рассчитанное значения критерия.

- Если рассчитанное (t_a и t_b) значение t-критерия Стьюдента *равно или больше критического* ($t_{табл}$), найденного по таблице, делаем вывод о статистической значимости различий между сравниваемыми величинами.

$$t_a \geq t_{табл}$$

- Если значение рассчитанного (t_a и t_b) t-критерия Стьюдента *меньше* критического ($t_{табл}$) значит различия сравниваемых величин статистически не значимы.

$$t_a \leq t_{табл}$$

После определения t-критерия Стьюдента, рассчитываем предельную ошибку – максимально возможное расхождение средних и максимум ошибок при заданной вероятности ее появления – Δ:

$$\Delta_a = t_{табл} * \mu_a\ ; \qquad \Delta_b = t_{табл} * \mu_b \qquad (16,17)$$

где Δ_a – предельная ошибка свободного члена a, Δ_b – предельная ошибка коэффициента b.

Для того чтобы определить являются коэффициенты регрессии эффективными и состоятельными, необходимо рассмотреть доверительные интервалы, для этого рассчитаем границы доверительного интервала:

$$(a - \Delta a) < a < (a + \Delta a) \qquad (18)$$

$$(b - \Delta b) < b < (b + \Delta b) \qquad (19)$$

Таким образом, полученные оценки коэффициента регрессии b и свободного члена a уравнения регрессии являются эффективными и состоятельными, а само уравнение $\hat{y}=a+bx$ будет использоваться для моделирования и прогнозирования динамики.

Задание на лабораторную работу:

1. Выполнить расчет в Excel.

2. Для нахождения коэффициентов уравнения регрессии выполнить все табличные расчеты.
3. Составить уравнение регрессии.
4. Выяснить, являются ли полученные оценки коэффициентов уравнения регрессии эффективными и состоятельными.
5. Построить на одном графике корреляционное поле, чтобы сделать предварительный вывод является ли регрессия линейная, а так же построить линию тренда.
6. Получить выводы о проделанной работе.
7. Ответить на контрольные вопросы в отчете.
8. Оформить отчет.

Варианты:

I

n	1	2	3	4	5	6	7	8	9	10	11	12
X	451	163	513	112	123	561	312	123	456	156	126	131
У	145	123	235	145	120	160	100	240	135	123	230	125

II

n	1	2	3	4	5	6	7	8	9	10	11	12
X	52	110	170	141	150	160	200	230	240	260	270	300
У	100	90	130	31	60	39	58	70	80	150	120	130

III

n	1	2	3	4	5	6	7	8	9	10	11	12
X	120	150	140	156	216	156	153	324	102	233	200	123
У	354	325	205	385	524	562	655	244	241	145	254	145

IV

n	1	2	3	4	5	6	7	8	9	10	11	12
X	26	48	65	23	150	123	264	156	152	154	520	415

У	221	153	155	102	156	264	435	156	203	325	456	163

V

n	1	2	3	4	5	6	7	8	9	10	11	12
X	15	26	65	320	652	156	896	123	16	263	459	213
У	125	163	162	263	563	23	463	126	133	152	213	159

VI

n	1	2	3	4	5	6	7	8	9	10	11	12
X	651	165	849	156	156	132	456	251	489	265	56	256
У	156	859	156	165	49	566	145	156	566	456	56	156

VII

n	1	2	3	4	5	6	7	8	9	10	11	12
X	156	165	415	894	465	163	496	131	566	561	562	354
У	526	414	561	561	456	321	648	654	321	456	464	456

VIII

n	1	2	3	4	5	6	7	8	9	10	11	12
X	463	135	563	456	131	452	496	131	566	561	562	354
У	526	414	561	561	456	321	156	123	542	563	451	456

IX

n	1	2	3	4	5	6	7	8	9	10	11	12
X	123	465	789	744	456	452	453	131	566	879	562	789
У	465	123	123	416	435	746	156	325	748	563	464	456

X

n	1	2	3	4	5	6	7	8	9	10	11	12
X	165	156	498	3,15	496	132	981	135	158	864	456	894
У	465	135	496	159	791	168	489	163	89	213	489	564

Пример выполнения работы:

Для выполнения работы потребуется программа Microsoft Office Excel. На основании выбранного преподавателем варианта необходимо выполнить расчеты.

n	1	2	3	4	5	6	7	8	9	10	11	12
X	100	90	150	31	60	39	40	70	80	150	120	130
У	131	110	170	141	150	160	200	230	240	260	270	300

Промежуточные расчеты для уравнения линейной парной регрессии

Для того чтобы получить параметры уравнения линейной парной регрессии, понадобятся промежуточные расчеты (таблица 1), где будут рассчитаны:

Произведение: *(x*y)*

Квадрат числа: x^2; y^2

Разность: $(x-\bar{x})$, $(y-\bar{y})$, где $\bar{x} = \frac{\sum_1^{12} x}{12}$ (см. формулы 4,5)

Квадрат разности: $(x-\bar{x})^2$; $(y-\bar{y})^2$

Сумма: $\sum x$; $\sum y$; $\sum(x*y)$; $\sum x^2$; $\sum y^2$; $\sum(x-\bar{x})$; $\sum(x-\bar{x})^2$;

Также необходимо посчитать средние значения: $\overline{xy}$; $\overline{x^2}$; $\overline{y^2}$

Таблица 1

Промежуточные расчеты уравнения

№	x	$y_{эксп}$	$(x*y)$	x^2	y^2	$(x-\bar{x})$	$(x-\bar{x})^2$
1	100	131	13100	10000	17161	11,7	136,1
2	90	110	9900	8100	12100	1,7	2,78
3	150	170	25500	22500	28900	61,7	3802,78
4	31	141	4371	961	19881	-57,3	3287,1
5	60	150	9000	3600	22500	-28,3	802,78
6	39	160	6240	1521	25600	-49,3	2433,78
7	40	200	8000	1600	40000	-48,3	2336,1
8	70	230	16100	4900	52900	-18,3	336,1
9	80	240	19200	6400	57600	-8,3	6469,4

Продолжение таблицы 1

10	150	260	39000	22500	67600	61,7	3802,78
11	120	270	32400	14400	72900	31,7	1002,78
12	130	300	39000	16900	90000	41,7	1736,1
Сумма:	1060	2362	221811	113382	507142	$0,7*10^{13}$	19748,7
Средние:	$\bar{x}$	$\bar{y}$	$\overline{xy}$	$\overline{x^2}$	$\overline{y^2}$		
	88,3	196,83	18484,25	9448,5	42261,83		

Произведение управляемого независимого параметра *x* на функцию выхода *y*:

$(x*y)_1$ = 100*131 = 13100

...

$(x*y)_{12}$ =130*300 = 39000

Квадрат числа независимого параметра *x* на функцию выхода *y*:

$x^2{}_1 = 100^2 = 10000$

...

$x^2{}_{12} = 130^2 = 16900$

$y^2{}_1 = 131^2 = 17161$

...

$y^2{}_{12} = 300^2 = 90000$

Разность управляемого независимого параметра *x* и среднего значения управляемого независимого параметра $\bar{x}$:

$(x_1 - \bar{x})$ = 100-88,3 = 11,7

...

$(x_{12} - \bar{x})$ = 130-88,3 =41,6

Сумма управляемых независимых параметров *x*:

$\sum x$ = (100+90+ ... +120+130) = 1060

...

$\sum y^2$ = (17161+ ... +90000) = 507142

Средние от суммы:

$$\bar{x} = \frac{1060}{12} = 88,3$$

...

$$\overline{y^2} = \frac{507142}{12} = 42261{,}83$$

Значения из заполненной таблицы, удобно использовать при дальнейших расчетах.

Параметры уравнения линейной парной регрессии

Чтобы расчитать коэффициенты для уравнения регрессии найдем дисперсию, см. формулы (2), (3):

$$\sigma_x^2 = \overline{x^2} - \overline{x}^2 = 9448{,}5 - (88{,}3)^2 = 1645{,}72$$

$$\sigma_y^2 = \overline{y^2} - \overline{y}^2 = 42261{,}83 - (196{,}83)^2 = 3518{,}47$$

Параметры a и b уравнения линейной регрессии $\hat{y}=a+bx$ рассчитываются по формулам, см. формулы (6), (8):

$$b = \frac{\overline{xy} - \overline{x} * \overline{y}}{\sigma_x^2} = \frac{18484{,}25 - 88{,}3 * 196{,}83}{1645{,}72} = 0{,}67$$

$$a = \overline{y} - b\overline{x} = 196{,}83 - 0{,}7 * 88{,}3 = 137{,}94$$

Полученные коэффициенты подставляются в уравнение регрессии, см. формулу (1):

$$\hat{y} = 137{,}94 + 0{,}7x$$

Таблица промежуточных расчетов для уравнения регрессии

Прежде чем найти доверительный интервал, для удобства, следует рассчитать таблицу промежуточных расчетов (таблица 2):

Таблица 2

Промежуточные расчеты для проверки эффективности и самостоятельности уравнения

№	***x***	***y*_эксп**	$\hat{y}_{рас}$	$(y - \hat{y})$	$(y - \hat{y})^2$

1	100	131	204,6	-73.61	5418,76
2	90	110	197,94	-87.94	7734,25
3	150	170	237,95	-67,95	4617,25
4	31	141	158,61	-17,61	309,96
5	60	150	177,94	-27,94	780,74
6	39	160	163,94	-3,94	15,52
7	40	200	164,61	35,39	1252,7
8	70	230	184,61	45,39	2060,31
9	80	240	191,28	48,72	2373,93
10	150	260	237,95	22,05	486,19
11	120	270	217,95	52,1	2709,47
12	130	300	224,62	75,4	5682,89
Сумма	1060	2362	2362	$-0,14*10^{12}$	33441,96

Табличные расчеты:

Уравнение регрессии для каждого x:

$\hat{y}_1 = 137{,}94 + 0{,}7x$ = 137,94 + 0.7*100 = 204,6

...

$\hat{y}_{12}$ = 137,94 + 0,7*130 = 224,62

Вычитание (y – ŷ):

$(y - y\hat{})_1$ = 131- 204,6 = -73,61

...

$(y - \hat{y})_{12}$ = 300 – 224,62= 75,4

Возведение в квадрат $(y - \hat{y})^2$:

$(y - \hat{y})^2{}_1 = (-73{,}61)^2$ = 5 418,76

...

$(y - \hat{y})^2{}_{12} = (75{,}4)^2$ = 5682,89

Поле корреляции

Чтобы построить поле корреляции в программе Microsoft Office Excel 13, следует выбрать: «Вставка» - «Рекомендуемые диаграммы» - «Точечная». Диаграмма строится по данным x и y (рис.2. Поле корреляции):

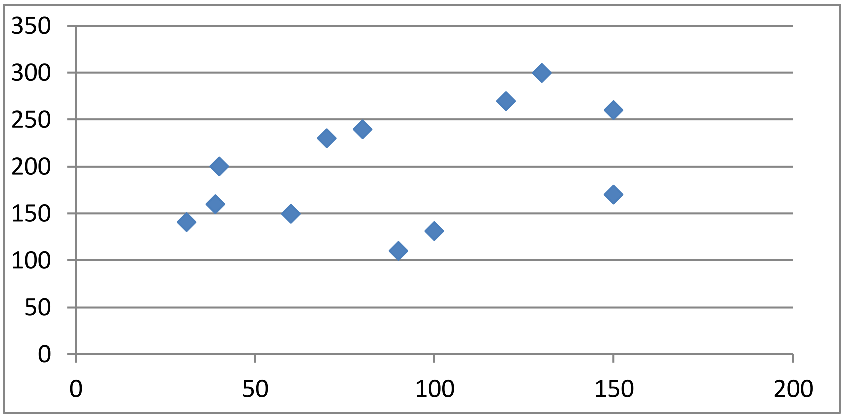

Рис. 2. Поле корреляции

Линия тренда

После расчета коэффициентов регрессии можно строить линию тренда.

$$\hat{y} = 137{,}94 + 0{,}7x$$

Линия тренда строится на имеющемся поле корреляции. Для построения нам нужно выбрать: Кликнуть в любом месте диаграммы и затем нажмите иконку с символом **плюс** (+) рядом с диаграммой, чтобы открыть меню «**Элементы диаграммы**» (Chart elements). Другой вариант: нажмите кнопку «**Добавить элемент диаграммы**» (Add Chart Elements), которая находится в разделе «**Макеты диаграмм**» (Chart Layouts) на вкладке «**Конструктор**» (Design).

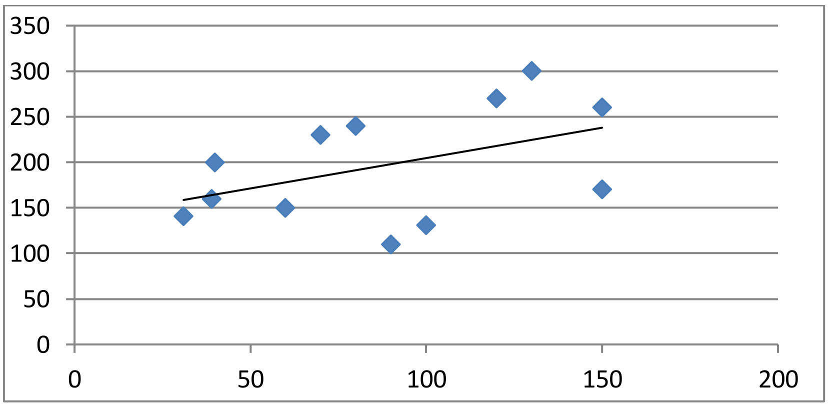

Рис.3. Линия тренда

Расчеты для проверки эффективности и состоятельности уравнения регрессии

Значение линейного коэффициента парной корреляции для определения связи рассчитывается по формуле (9):

$$r_{xy} = \frac{\overline{xy} - \overline{x} * \overline{y}}{\sqrt{\sigma_x^2 * \sigma_y^2}} = \frac{18484{,}25 - 88{,}3 * 196{,}83}{\sqrt{1645{,}72 * 3518{,}47}} = 0{,}456$$

Интервал, в котором находится рассчитанный коэффициент определяется по шкале Чеддока.

Коэффициент находится в интервале от 0,3 до 0,5, связь умеренная, прямая.

Далее рассчитывается стандартная ошибка остаточной компоненты, формулы (10):

$$S_{\text{Е}} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \widehat{y_i})^2} = \sqrt{\frac{33\,441{,}96}{12-1}} = 55{,}14$$

Найдем средние квадратичные (стандартные) ошибки оценивания коэффициента b и свободного члена a уравнения регрессии, формулы (12) и (13) :

$$\delta_a = S_E * \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N * \sum_{i=1}^{N}(x_i - \overline{x})^2}} = 55{,}14 * \sqrt{\frac{113382}{12 * 19748{,}67}} = 38{,}14$$

$$\delta_b = \frac{S_E^2}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2}} = \frac{55{,}14^2}{\sqrt{19748{,}67}} = 21{,}63$$

Найдем t – критерий Стьюдента для обоих параметров, смотреть формулы (14) и (15):

$$t_a = \frac{a}{\mu_a} = \frac{137{,}94}{38{,}14} = 3{,}62$$

$$t_b = \frac{b}{\mu_b} = \frac{0{,}67}{21{,}63} = 0{,}03$$

Сравниваем полученное значение t-критерия Стьюдента с критическим при p=0,05 и f = 11 (f=N-1) (число степеней свободы) значением, указанным в таблице: $t_{табл}$ = 2,20, можно сказать, что с вероятностью 95% коэффициент a надёжен, коэффициент b ненадёжен при данном уровне значимости.

Так как рассчитанное значение критерия $t_a = 3{,}62$ больше критического, делаем вывод о том, что наблюдаемые различия статистически значимы, коэффициент a надёжен. Значение критерия $t_b = 0{,}03$ меньше табличного, значит различия сравниваемых величин статистически не значимы, коэффициент b ненадёжен при данном уровне значимости

Для расчета доверительного интервала определяем предельную ошибку Δ, формулы (16) и (17):

$$\Delta_a = t_{табл} * \mu_a = 2{,}20 * 3{,}62 = 7{,}964$$
$$\Delta_b = t_{табл} * \mu_b = 2{,}20 * 0{,}031 = 0{,}068$$

Доверительные интервалы для коэффициентов регрессии, рассчитываются по формулам (18), (19):

$$(a - \Delta a) < a < (a + \Delta a)$$
$$130{,}06 < a < 145{,}82$$
$$(b - \Delta b) < b < (b + \Delta b)$$
$$0{,}602 < b < 0{,}738$$

Если границы интервала имеют разные знаки, т.е. в эти границы попадает ноль, то оцениваемый параметр принимается нулевым.

В отношении (18), (19) справедливы общие правила: чем уже доверительный интервал, тем точнее оценка параметра; если доверительный интервал включает нулевое значение, то оцениваемый параметр статистически незначим (равен нулю).

Таким образом, из-за ненадёжности, полученные оценки коэффициента регрессии b не являются эффективными и состоятельными, а само уравнение y = 137,94 + 0,67x не может использоваться для моделирования и прогнозирования динамики. Это обусловлено большой ошибкой уравнения регрессии.

Содержание оформления отчета:

1. Титульный лист, содержащий информацию о студенте (группа, ФИО, номер варианта).
2. Содержание отчета.
3. Выполненная работа, формулы, расчеты.
4. График.
5. Ответы на контрольные вопросы.
6. Выводы по лабораторной работе.

Контрольные вопросы:

1. Дайте определение понятию эксперимент.
2. Что называется планированием эксперимента? Факторное пространство-это....?
3. Перечислите виды регрессии.
4. Дайте определение понятию «черный ящик»
5. Дайте определение регрессии.
6. Какова цель отбора особенности измерения переменных?
7. Дайте определение дисперсии случайной величины.

Литература:

1. Парная линейная регрессия. [Электрон.ресурс] - URL: http://life-prog.ru/1_30099_tema-parnaya-lineynaya-regressiya.html, 30.05.2016 г.

2. Построение регрессионной модели системы двух случайных величин. [электронный ресурс] - URL: http://www.studfiles.ru/preview/5443443/page:3/, 30.05.2016 г.
3. Астахова Л.Г. Математическая теория планирования эксперимента: учебное пособие [Текст]: / Астахова Л.Г.;- Владикавказ, 2013.- 96 с.
4. Бойко Н.Г. Теория и методы инженерного эксперимента: курс лекций. [Текст]:/ Бойко Н.Г. [и др.]; - Донецк: ЛО-ГОС, 2009.- 155с.

Ответы на вопросы:

1. Эксперимент – это совокупность операций совершаемых над объектом исследования с целью получения информации об его свойствах.
2. Нахождение таких условий и правил проведения опытов, при которых удается получить надежную и достоверную информацию об объекте с наименьшей затратой труда.

   Пространство – координатные оси которого соответствуют значениям факторов.
3. Виды регрессии: парная y = f(x); множественная y = f(x1, x2, x3 … xn).
4. Черный ящик – это сложная гомоморфная (неполная) модель кибернетической системы, в которой соблюдается разнообразие.
5. Регрессия – зависимость между случайными величинами.
6. Цель этого отбора – уменьшить ошибки.
7. Это мера разброса данной случайной величины, то есть её отклонения от математического ожидания.

Приложение 1

Шкала Чеддока

	Значение коэффициента корреляции при наличии	
	прямой связи	обратной связи
Слабая	0,1–0,3	(-0,3)–(-0,1)
Умеренная	0,3–0,5	(-0,5)–(-0,3)
Заметная	0,5–0,7	(-0,7)–(-0,5)
Высокая	0,7–0,9	(-0,7)–(-0,9)
Весьма высокая	0,9–1	(-1)–(-0,9)

Приложение 2

t-критерий Стьюдента

Число степеней свободы	Уровень					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,314	12,71	31,82	63,66	318,3	636,6
2	2,920	4,303	6,965	9,925	22,33	31,60
3	2,353	3,182	4,541	5,841	10,21	12,92
4	2,132	2,776	3,747	4,604	7,173	8,610
5	2,015	2,571	3,365	4,032	5,893	6,869
6	1,943	2,447	3,143	3,707	5,208	5,959
7	1,895	2,365	2,998	3,499	4,785	5,408
8	1,860	2,306	2,896	3,355	4,501	5,041
9	1,833	2,262	2,821	3,250	4,297	4,781
10	1,812	2,228	2,764	3,169	4,144	4,587
11	1,796	2,201	2,718	3,106	4,025	4,437
12	1,782	2,179	2,681	3,055	3,930	4,318
13	1,771	2,160	2,650	3,012	3,852	4,221
14	1,761	2,145	2,624	2,977	3,787	4,140
15	1,753	2,131	2,602	2,947	3,733	4,073
16	1,746	2,120	2,583	2,921	3,686	4,015
17	1,740	2,110	2,567	2,898	3,646	3,965
18	1,734	2,101	2,552	2,878	3,610	3,922
19	1,729	2,093	2,539	2,861	3,579	3,883
20	1,725	2,086	2,528	2,845	3,552	3,850
21	1,721	2,080	2,518	2,831	3,527	3,819
22	1,717	2,074	2,508	2,819	3,505	3,792
23	1,714	2,069	2,500	2,807	3,485	3,767
24	1,711	2,064	2,492	2.797	3,467	3,745
25	1,708	2,060	2,485	2,787	3,450	3,725
26	1,706	2,056	2,479	2,779	3,435	3,707
27	1,703	2,052	2,473	2,771	3,421	3,690
28	1,701	2,048	2,467	2,763	3,408	3,674
29	1,699	2,045	2,462	2,756	3,396	3,659
30	1,697	2,042	2,457	2,750	3,385	3,646
00	1,645	1,960	2,326	2,576	3,090	3,291