

Matematické metody v ekonomii

Мусрепов С.Б., Мусрепов Н.С., Баймурзин А.С.

НАО Северо-Казахстанский Университет им. М. Козыбаева

УСТАНОВЛЕНИЕ СВЯЗИ МЕЖДУ ОБЪЕКТАМИ ДЕЙСТВИТЕЛЬНОСТИ МЕТОДОМ ЛИНЕЙНОЙ РЕГРЕССИИ.

Введение

В данной работе, рассмотрим один из методов регрессионного анализа, который позволит установить связь между разными параметрами, характер и силу их взаимосвязи. Данный метод позволяет также устанавливать причины происходящих явлений, и помогает предсказывать будущие процессы.

Исходя из этого, в данной работе рассмотрим парную линейную регрессию.

Простая (парная) линейная регрессия, представляет собой модель, где независимая переменная y рассматривается как функция одной независимой переменной x , в явном виде:

$$y = a + b \cdot x,$$

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

Материалы и методы исследования

Для рассмотрения и использования парной линейной регрессии, необходимо провести ряд последовательных работ.

Первое, в данной работе рассмотрим линейную связь регрессионного анализа между валовым сбором зерна (в центнерах) и посевной площадью зерновых и зернобобовых (в гектарах) в Северо- Казахстанской области за 2024 год.

Валовой сбор зерновых (в весе после доработки, центнеров)					Посевная площадь зерновых и зернобобовых-всего (гектаров)				
Все категории хозяйств					Все категории хозяйств				
	2024г					2024г			
Итого по области	61 006 248				Итого по области	3 335 553			
Петропавловск	1 642				Петропавловск	75			
Айыртауский	5 063 921				Айыртауский	289 324			
Акжарский	3 506 465				Акжарский	224 625			
М.Жумабаева	6 312 333				М.Жумабаева	333 072			
Есильский	4 054 861				Есильский	211 547			
Жамбылский	4 241 100				Жамбылский	231 611			
Кызылжарский	3 216 482				Кызылжарский	151 489			
Мамлютский	2 611 229				Мамлютский	130 528			
Шал акына	4 002 577				Шал акына	219 034			
Аккайынский	3 778 125				Аккайынский	169 748			
Тайыншинский	8 871 619				Тайыншинский	440 687			
Тимирязевский	4 065 619				Тимирязевский	215 877			
Уалихановский	1 954 777				Уалихановский	145 215			
Г.Мусрепова	9 325 498				Г.Мусрепова	572 721			

Приступаем к составлению исходной матрицы элементов системы линейных уравнений:

	1	2	3	4	5	6	7	8
	Итого по об.	Петропавловск	Айыртауский	Акжарский	М.Жумабаева	Есильский	Жамбылский	Кызылжарский
X	3 335 553	75	289 324	224 625	333 072	211 547	231 611	151 489
Y	61 006 248	1 642	5 063 921	3 506 465	6 312 333	4 054 861	4 241 100	3 216 482

	9	10	11	12	13	14	15
	Мамлютский	Шал акына	Аккайынский	Тайыншинский	Тимирязевский	Уалихановский	Г.Мусрепова
	130 528	219 034	169 748	440 687	215 877	145 215	572 721
	2 611 229	4 002 577	3 778 125	8 871 619	4 065 619	1 954 777	9 325 498

где x- посевная площадь(га); y- валовой сбор зерновых(центнерах).

1

Следующим этапом приступаем к определению дисперсии случайной величины σ_x^2 меры разброса данной случайной величины, то есть её отклонения от математического ожидания по следующим формулам:

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2; \quad \sigma_y^2 = \overline{y^2} - \bar{y}^2$$

Для определения дисперсии случайной величины вычисляют средние значения x и y:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N}; \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{N}, \quad \text{где } N - \text{число экспериментов.}$$

Дисперсия случайной величины σ_x^2 - эта мера разброса данной случайной величины и ее отклонения от математического ожидания и определяется следующими формулами и принимают значения:

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = 18247913783; \quad \sigma_y^2 = \overline{y^2} - \bar{y}^2 = 5,7038E+12$$

Далее приступаем к нахождению параметров a и b уравнения линейной регрессии $\hat{y} = a + b \cdot x$:

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = \frac{1,35382E+12 - 1,03821E+12}{18247913783} = 17,29554527$$

$$a = \bar{y} - b \bar{x} = 4357589,143 - 17,29554527 \cdot 238253,7857 = 236860,0063$$

Получаем уравнение регрессии: $\hat{y} = 236860,0063 + 17,29554527 \cdot x$

Следующий шаг линейной регрессии является проверка эффективности и состоятельности уравнения регрессии.

Она определяется с помощью линейного коэффициента парной корреляции:

$$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{\sigma_x^2 \cdot \sigma_y^2}} = \frac{1,35382E+12 - 238253,8 \cdot 4357589}{\sqrt{18247913783 \cdot 5,70378E+12}} = 0,978271944$$

Определяем интервал, в котором находится полученный коэффициент парной корреляции по шкале 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)

В данном случае, полученный коэффициент находится в интервале от (0,9) до (1.0), это означает, что линейная связь между параметрами прямая и весьма высокая.

Далее определяем стандартную ошибку остаточной компоненты следующей формулой:

$$S_E = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \hat{y}_i)^2} = \sqrt{\frac{1,37247E+15}{14-1}} =$$

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

$$\delta_a = S_E * \sqrt{\frac{\sum_{i=1}^N x_i^2}{N * \sum_{i=1}^N (x_i - \bar{x})^2}} = \quad \delta_b = \frac{S_E^2}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2}} = \quad \begin{matrix} 314921,9; \\ 208876919,4 \end{matrix}$$

2

Для обоих параметров находим критерий Стьюдента – t следующими формулами:

$$t_a = a/\delta_a = 0,75212297; \quad t_b = b/$$

Производим сравнение полученных значений t – критерия с критическими по таблице Стьюдента при $P = 0,05$ и $f = 13 (f = N - 1)$ (число степеней свободы), получаем значение

$$t_{\text{таб}} = 2,160$$

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
∞	1,645	1,960	2,326	2,576	3,090	3,291

В результате сравнения полученных значений критерий t_a и t_b и критерия Стьюдента по данной таблице, можно сделать вывод о том, что коэффициент a выгоден в небольшом % уровне значимости, а коэффициент b ненадежен на данном уровне значимости.

После определения t – критерия, приступаем к расчету предельной ошибки которое указывает на возможное расхождение средних и максимум ошибок, при заданной вероятности ее появления – Δ :

$$\Delta_a = t_{\text{таб}} * \mu_a = 680231,304; \quad \Delta_b = t_{\text{таб}} * \mu_b = 451174145,9;$$

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

Следующим этапом необходимо установить, что найденные коэффициенты регрессии являются эффективными и состоятельными, т.е. $a=236860,0063$ и $b=17,29554527$.

Для этого рассмотрим понятия доверительные интервалы и рассчитаем границы его предела:

$$\begin{aligned}
 (a - \Delta_a) &< a < (a + \Delta_a) \\
 -443371 &< a < 917091,31 \\
 (b - \Delta_b) &< b < (b + \Delta_b) \\
 17,29555 &< b < 236860,006
 \end{aligned}$$

Результаты исследования и заключение

В результате были установлены доверительные границы для коэффициентов α и b .

Следует также отметить следующие существующие правила:

Во-первых, чем уже доверительный интервал, тем точнее становится оценка параметра;

во-вторых, если доверительный интервал включает нулевое значение, то оцениваемый параметр статистический равен нулю;

в третьих, если границы доверительного интервала имеют разные знаки, т.е. в эти границы попадает значение нуля, то оцениваемый параметр принимается нулевым.

Таким образом, исходя из этих правил, в частности, одна из границ доверительного интервала имеет разные знаки, т.е. попадает значение нуля, то наш оцениваемый параметр принимается нулевым ($\alpha = 0$).

Далее определяем среднюю ошибку аппроксимации следующей формулой и получаем значения:

$$\bar{A} = \frac{1}{14} * \sum_{i=1}^n \left| \frac{y_i - \hat{y}_{x_i}}{y_i} \right| * 100\% = 14,2\%$$
 , которое также показывает о некачественном подборе значения исходных данных уравнения регрессии, так как принимает значение выше 10%. А допустимая ошибка аппроксимации должна составлять 8%-10 % и она показывает качество подбора модели по отношению к исходным данным.

Далее, учитывая все выше перечисленные недостатки, следует отметить, также и о том, что полученные оценки коэффициентов регрессии α и b показывают неэффективность и состоятельность, а само уравнение $Y = 236860,0063 + 17,29554527 * X$ не может использоваться для моделирования и прогнозирования динамики процесса.

Список используемой литературы.

1. Парная линейная регрессия. [Электрон.ресурс] - URL: http://lifeprog.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.

PEDAGOGICKÉ VĚDY

Metodické základy vzdělávacího procesu

Юсупова Г.К., Бекмагамбетова Р.К. РАЗВИТИЕ НАВЫКОВ САМОРЕГУЛИРУЕМОГО ОБУЧЕНИЯ БУДУЩИХ ПЕДАГОГОВ ДОШКОЛЬНЫХ ОРГАНИЗАЦИЙ ЧЕРЕЗ ИСПОЛЬЗОВАНИЕ ЦИФРОВЫХ ИНСТРУМЕНТОВ В ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ	71
Досбенбетова А.Ш., Нарбутаева А.М., АКТУАЛЬНОСТЬ ПЕДАГОГИЧЕСКОГО НАСЛЕДИЯ АЛЬ-ФАРАБИ В ПРОЦЕССЕ ПОДГОТОВКИ БУДУЩЕГО УЧИТЕЛЯ	74

PSYCHOLOGIE A SOCIOLOGIE

Klinická psychologie

Пронич К. А., Григоркевич В.Б. ВЛИЯНИЕ СОВРЕМЕННОЙ ЖУРНАЛИСТИКИ НА РАЗВИТИЕ ЛИЧНОСТИ И ФОРМИРОВАНИЕ МЕДИАГРАМОТНОСТИ МОЛОДЁЖИ	82
CONTENTS.....	86

314650

314599

314651

314631

314608

314647

314616

314646

314597

314649

314593