# Лабораторная работа 14.

# Тема: Корреляционно-регрессионный анализ в Excel с использованием настройки Пакет анализа

Для проведения **корреляционно-регрессионного анализа** в первую очередь необходимо построить **матрицу коэффициентов парной корреляции** для оценки степени влияния факторов на зависимую переменную и друг на друга. Для построения матрицы коэффициентов парной корреляции необходимо выбирать команду меню *Данные/Анализ данных/Корреляция*.

Проведение множественного корреляционно-регрессионного анализа.

## **1.Теоретические аспекты корреляционного анализа.**

Изменение любого экономического показателя зависит от большого числа факторов, но из них лишь некоторые оказывают существенное воздействие на исследуемый показатель. Доля влияния остальных факторов столь незначительна, что их игнорирование не может привести к существенным отклонениям исследуемого объекта.

В большинстве случаев между экономическими явлениями не существует строгой функциональной взаимосвязи, поэтому в экономике говорят не о функциональных, а о корреляционных или статистических зависимостях.

Нахождение, оценка и анализ таких зависимостей и оценка их параметров являются одним из разделов эконометрики.

Эконометрика - это наука, в которой на базе реальных статистических данных строятся, анализируются и совершенствуются модели реальных экономических явлений.

При рассмотрении взаимосвязей выделяют одну из величин как независимую, а другие как зависимые. При рассмотрении зависимости двух случайных величин говорят о парной регрессии. Зависимость нескольких переменных называют множественной регрессией.

В регрессионных моделях зависимая (объясняемая) переменная *Y* может быть представлена в виде функции *f(X1, X2, …, Xn),* где *X1, X2, …, Xn* - независимые (объясняющие) переменные или факторы. В зависимости от вида функции *f(X1, X2, …, Xn)* модели делятся на линейные и нелинейные. В зависимости от количества включенных в модель факторов *X* модели делятся на однофакторные и многофакторные.

Основными этапами построения регрессионной модели являются:

* Построение системы показателей (факторов). Сбор и предварительный анализ исходных данных. Построение матрицы коэффициентов парной корреляции.
* Выбор вида модели и численная оценка ее параметров.
* Проверка качества модели.
* Оценка влияния отдельных факторов на основе модели.
* Прогнозирование на основе модели регрессии.

Выбор факторов, влияющих на исследуемый показатель, производится на основании качественного и количественного анализа исследуемых явлений.

Исключение части факторов осуществляется на основе анализа парных коэффициентов корреляции и оценкой их значимости. Коэффициент парной корреляции определяется по формуле:



где ****– среднее значение факторного признака,

**** – среднее значение результативного признака.

Значение коэффициентов парной корреляции лежит в интервале от *‑1*до*+1*. Его положительное значение свидетельствует о прямой связи, отрицательное - об обратной, т.е. когда растет одна переменная, другая уменьшается. Связь считается достаточно сильной, если коэффициент корреляции по абсолютной величине превышает *0,7* и слабой, если меньше *0,4*.

Для оценки значимости коэффициента корреляции применяется *t* - критерий Стьюдента. при этом фактическое значение этого критерия *(tнабл)*



сравнивается с критическим значением *tкр* которое берется из таблицы значений *t* с учетом заданного уровня значимости *(α = 0.05)* и числа степеней свободы *(n - 2)*.

Если *tнабл > tкр,* то полученное значение коэффициента парной корреляции признается значимым.

Одним из условий регрессионной модели является предположение о функциональной независимости объясняющих переменных. связь между факторами называется мультиколлинеарностью, которая делает вычисление параметров модели либо невозможным, либо затрудняет содержательную интерпретацию параметров модели. Считают явление мультиколлинеарности в исходных данных установленным, если коэффициент парной корреляции между двумя переменными больше *0.8*. Чтобы избавиться от мультиколлинеарности, в модель включают лишь один из функционально связанных между собой факторов, причем тот который в большей степени связан с зависимой переменной.

## **2 Математическая постановка задачи.**

Переменные:

*Х0*- валовой внутренний продукт, млрд. руб.

*X]* - объем промышленной продукции, млрд. руб.

*Х2*- инвестиции в основной капитал, млрд. руб.

*Х3*- розничный товарооборот, млрд. руб.

*Х4*- объем платных услуг населению, млрд. руб.

*Х5*- доходы консолидированного бюджета, млрд. руб.

*Хб*- расходы консолидированного бюджета, млрд. руб.

*Х7* - общая численность официально зарегистрированных безработных, тыс. чел.

*Х8* - номинальная начисленная среднемесячная заработная плата, тыс. руб.

*Х9* - денежные доходы населения, млрд. руб.

*Х10* - денежные расходы и сбережения населения, млрд. руб.

|  |  |  |
| --- | --- | --- |
| Вариант | Зависимая переменная | Независимые переменные |
| 1 | *X0* | *X1 – X10* |
| 2 | *X1* | *X0, X2 - X10* |
| 3 | *X2* | *X0, X1, X3 - X10* |
| 4 | *X3* | *X0 - X2, X4 - X10* |
| 5 | *X4* | *X0 – X3, X5 - X10* |
| 6 | *X5* | *X0 – X4, X6 - X10* |
| 7 | *X6* | *X0 – X5, X7 - X10* |
| 8 | *X7* | *X0 – X6, X8 - X10* |
| 9 | *X8* | *X0 – X7, X9 - X10* |
| 10 | *X9* | *X0 – X8, X10* |
| 11 | *X10* | *X0 – X9* |
| 12 | *X0* | *X1 – X10* |
| 13 | *X1* | *X0, X2 - X10* |
| 14 | *X2* | *X0, X1, X3 - X10* |
| 15 | *X3* | *X0 - X2, X4 - X10* |

Даны данные для показателей *Х0*-*Х10*. Исследовать их зависимость. Для этого:

1.Построить матрицу коэффициентов парной корреляции, проанализировать ее, сделать вывод о необходимости включения в модель данных факторов.

2. Рассчитать параметры линейной и экспоненциальной моделей. Для расчета параметров линейной модели использовать функцию *ЛИНЕЙН* и инструмент *Регрессия* надстройки *Пакет анализа*, для расчета параметров экспоненциальной - функцию *ЛГРФПРИБЛ*. Для линейной и экспоненциальной моделей рассмотреть случаи, когда аргумент *Константа* в функциях *ЛИНЕИН* и *ЛГФРФПРИБЛ* имеет значение *ИСТИНА* и *ЛОЖЬ*.

3. Сделать выводы: 1) о значимости коэффициентов, входящих в модель; 2) об адекватности модели фактическим данным;

4. На основе проведенного анализа определить вид модели, наиболее точно описывающей фактические данные;

5. Рассчитать прогнозные значения, используя выбранную модель. Найти отклонение фактических данных от расчетных. Сделать вывод;

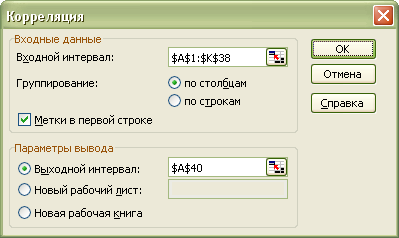
6. Построить график, отражающий фактические и расчетные данные.

**Исходные данные к заданию 1**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **X0** | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** | **X10** |
| 4044,3 | 4356,4 | 605,1 | 1626,7 | 500,2 | 2063,2 | 1604 | 84,7 | 398,2 | 2612,8 | 2651 |
| 4201,3 | 4376,5 | 620,1 | 1602,5 | 512,7 | 2143,7 | 1869,2 | 84,7 | 400,4 | 2554,4 | 2497,4 |
| 4861,2 | 5012 | 914 | 1880,7 | 562 | 2447,7 | 2320,2 | 82,2 | 427,6 | 2783,4 | 2785 |
| 4560 | 4906,7 | 862,1 | 1982,7 | 556,1 | 2406,4 | 2496,5 | 80,3 | 440,6 | 2922,6 | 2837,7 |
| 4886,3 | 4854,9 | 958,4 | 2037 | 535,4 | 2592,9 | 2277,3 | 77,4 | 449,9 | 2842,8 | 2839,9 |
| 5788,8 | 4926,3 | 1488,9 | 2193,9 | 582,7 | 2698 | 2834,4 | 84,4 | 476,4 | 3339,9 | 3274,4 |
| 5539,7 | 4835,1 | 1231,5 | 2152,1 | 533,1 | 2529,7 | 2563,8 | 73,1 | 487,9 | 3252,5 | 3188,2 |
| 6431,4 | 5254,4 | 1429,6 | 2227 | 532,9 | 2644,9 | 2434,2 | 73 | 497,4 | 3164 | 3224 |
| 6681,5 | 5588,4 | 1679,5 | 2344,4 | 550,7 | 2793,7 | 2616,7 | 72,5 | 485,2 | 3503,3 | 3444,4 |
| 5836 | 5416,7 | 1326,2 | 2341,7 | 567,6 | 2669,2 | 2565,9 | 71,1 | 501,5 | 3309,6 | 3318,1 |
| 5881,2 | 5477,2 | 1456,8 | 2211,9 | 625,9 | 2845 | 2904,5 | 69,3 | 503,3 | 3350,4 | 3336,6 |
| 6355,4 | 5503,9 | 2523,6 | 2629,8 | 716,9 | 2990,5 | 4770,1 | 67,9 | 562,2 | 4172,2 | 4072,7 |
| 4995,9 | 5842,5 | 846 | 2017,5 | 633,6 | 2659,8 | 1982,3 | 72,2 | 519 | 3299,1 | 3328,2 |
| 5175,2 | 5984,6 | 923,8 | 2009,4 | 635,5 | 2636,6 | 2517,9 | 73,4 | 525,4 | 3335,3 | 3292,8 |
| 5971,6 | 6446,3 | 1173,3 | 2260 | 679,5 | 2943,1 | 3048,8 | 72,8 | 558,5 | 3619,7 | 3608,4 |
| 5568,1 | 6082,5 | 1156,7 | 2400,1 | 622,6 | 2890,9 | 2984,8 | 71,8 | 561,6 | 3836,6 | 3732,5 |
| 6025,1 | 6301,7 | 1450,2 | 2508,1 | 635,3 | 3051,5 | 2788,9 | 69,3 | 579,3 | 3361,1 | 3376,4 |
| 7025,8 | 6603,9 | 1845,2 | 2684,1 | 680,9 | 3249 | 3344 | 66,5 | 604 | 4203,3 | 4082 |
| 6782,7 | 6593,6 | 1566,4 | 2736,6 | 670 | 3052,6 | 3026,7 | 62,3 | 612,3 | 3961,9 | 3932,6 |
| 7775,5 | 7003,6 | 1729,7 | 2824,5 | 678,6 | 3349,7 | 2894,7 | 60 | 623,5 | 4016,9 | 3999,9 |
| 7993,4 | 6823,4 | 1987,3 | 2880,2 | 684,4 | 3456,3 | 3094,8 | 56,4 | 606,4 | 4247,3 | 4192,3 |
| 7169,8 | 6610,8 | 1902,7 | 2812,9 | 788,2 | 3731,2 | 3119,8 | 53,2 | 618,4 | 4146,8 | 4186,5 |
| 7155,5 | 6482,3 | 1839,1 | 2704,2 | 765,1 | 3517,8 | 3327,3 | 52,6 | 611,5 | 4277,5 | 4255,5 |
| 7628,4 | 6491,8 | 3953,7 | 3224,2 | 833,5 | 3823,1 | 4507 | 52 | 668 | 6379,6 | 6297 |
| 6194,3 | 6319,8 | 1351,2 | 2584,7 | 795,3 | 3482,9 | 2321,8 | 53,5 | 617,2 | 4148,2 | 4283,1 |
| 6352,4 | 6607,3 ' | 1185,3 | 2466,7 | 770,1 | 3347,6 | 2941 | 53 | 614,3 | 4180,2 | 4152,6 |
| 7220,6 | 7068,7 | 1715,5 | 2928,3 | 815,7 | 3585,4 | 3284,1 | 51,7 | 659,4 | 4601,5 | 4584,6 |
| 6804,6 | 6895,9 | 1536,4 | 3036,4 | 758,7 | 3678,3 | 3856,4 | 50,2 | 661,9 | 4800,2 | 4687,9 |
| 7325,9 | 7459,9 | 1823,1 | 3021,1 | 777,8 | 3801,6 | 3647,7 | 47,7 | 686,9 | 4242,8 | 4284,6 |
| 8336,5 | 7647,9 | 2452,1 | 3237,6 | 837,3 | 4002,1 | 4038,2 | 46,3 | 710,2 | 5270,7 | 5144,8 |
| 8236,2 | 7660,3 | 2076,6 | 3247,1 | 820,4 | 3990,3 | 4067,5 | 46,7 | 732 | 4788 | 4769,1 |
| 9214,2 | 8158,4 | 2129,2 | 3436,9 | 829,1 | 4212 | 3588,1 | 48,6 | 737,3 | 4984,9 | 4984,9 |
| 9721,8 | 7857,1 | 2502,7 | 3472,8 | 820,8 | 4154,2 | 3781,3 | 46,5 | 713,4 | 5239 | 5198,5 |
| 8686,4 | 8336,9 | 2238,7 | 3504,1 | 872,2 | 4322,7 | 4369,4 | 45,5 | 738 | 4993,5 | 5050,6 |
| 8615,6 | 8589,3 | 2417,6 | 3357,1 | 916 | 4623,1 | 4506,1 | 45,2 | 736,4 | 5327,6 | 5300,1 |
| 9378,7 | 8902,3 | 3838,4 | 4034,7 | 974,8 | 4817,9 | 7101,1 | 44,1 | 795,4 | 6410,2 | 6293,5 |
| 7860,4 | 9516,9 | 1468,6 | 3450,4 | 938,5 | 4632 | 2747,2 | 49,6 | 756,3 | 5257 | 5272,2 |

## **3. Проведение корреляционного анализа средствами MS Excel .**

Для построения матрицы коэффициентов парной корреляции необходимо выбирать команду меню *Сервис/Анализ данных/Корреляция.* Откроется следующее диалоговое окно:



Далее следует нажать кнопку *OK*. После этого будет создана матрица коэффициентов парной корреляции:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Y | X0 | X1 | X2 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
| Y | 1 |  |  |  |  |  |  |  |  |  |  |
| X0 | 0.954 | 1 |  |  |  |  |  |  |  |  |  |
| X1 | 0.926 | 0.880 | 1 |  |  |  |  |  |  |  |  |
| X2 | 0.820 | 0.795 | 0.609 | 1 |  |  |  |  |  |  |  |
| X4 | 0.917 | 0.835 | 0.926 | 0.716 | 1 |  |  |  |  |  |  |
| X5 | 0.969 | 0.918 | 0.960 | 0.740 | 0.965 | 1 |  |  |  |  |  |
| X6 | 0.817 | 0.739 | 0.660 | 0.888 | 0.747 | 0.751 | 1 |  |  |  |  |
| X7 | -0.920 | -0.885 | -0.891 | -0.689 | -0.928 | -0.939 | -0.680 | 1 |  |  |  |
| X8 | 0.968 | 0.914 | 0.963 | 0.721 | 0.954 | 0.973 | 0.746 | -0.953 | 1 |  |  |
| X9 | 0.943 | 0.874 | 0.852 | 0.881 | 0.923 | 0.924 | 0.833 | -0.885 | 0.921 | 1 |  |
| X10 | 0.947 | 0.879 | 0.863 | 0.873 | 0.934 | 0.934 | 0.819 | -0.898 | 0.930 | 0.998 | 1 |

Анализ матрицы коэффициентов парной корреляции показывает, что существенное влияние на зависимую переменную оказывают все факторы.Для исключения явления мультиколлинеарности все факторы кроме *X2* и *X5* следует исключить из модели.