**Министерство образования и науки Украины**

**Приднепровская государственная академия**

**строительства и архитектуры**

**Кафедра прикладной математики**

**Математическое моделирование и решение задач оптимального управления**

**Тема: Математическое моделирование в управлении**

**Днепропетровск**

**2009**

**Введение**

Экономические процессы характеризует большое число параметров, взаимосвязь и взаимное влияние которых определяют состояние этой, с точки зрения системного анализа, сложной динамической системы и возможности перехода ее в другие состояния. В этой ситуации решения по оптимальному управлению необходимо принимать на основании многомерного статистического анализа стохастической, неполной информации. Всестороннее изучение деятельности предприятий дает возможность фиксировать значения таких показателей, как производительность труда, индекс снижения себестоимости, рентабельность одновременно с показателями, которые могут оказывать влияние на перечисленные результирующие показатели. К таким воздействующим показателям можно отнести трудоемкость единицы продукции, удельный вес рабочих в составе промышленно-производственного персонала, коэффициент сменности оборудования, премии, фондоотдача и другие. Однако следует учесть, что показатели могут быть взаимосвязаны и, следовательно, могут опосредствовано оказывать влияние на результирующие показатели.

Рассмотрение поведения подобных сложных систем, характерной особенностью которых является наличие управления, присуще системному подходу. Основным методом исследования систем в рамках системного подхода является **метод математического моделирования**, базирующийся на использовании средств компьютерной техники. Под математическим моделированием понимают способ исследования различных явлений, процессов путем исследования явлений, имеющих разное физическое содержание, но описываемых одинаковыми математическими соотношениями.

**Математической моделью** реальной системы называется ее описание на каком-либо формальном языке, позволяющее выводить суждения о некоторых чертах поведения этой системы при помощи формальных процедур. Математическая модель может представлять собой функциональные зависимости или графики, уравнения; таблицы или графики, описывающие движение систем и переходы их из одних состояний в другие. Другими словами, математическая модель – это отражение оригинала (системы) в виде функций, уравнений, неравенств, цифр и т.д. Математическая модель – это приближенное описание системы и ее поведения с помощью математической символики. Математическое моделирование – мощный метод познания, а также прогнозирования и управления. Математическое моделирование занимает ведущее место среди других методов исследования, особенно благодаря компьютерной технике, возможности которой позволяют исследование поведения системы осуществлять с помощью машинного эксперимента.

В настоящее время трудно представить себе исследование и прогнозирование экономических явлений, без использования эконометрического моделирования на основе статистических данных, регрессионного анализа и других методов, опирающихся на теорию вероятностей. Экономические законы все более усложняются и, следовательно, в соответствии с законами развития динамических систем должен усиливаться статистический характер законов, их описывающих, который позволяет учитывать влияние случайных факторов.

Таким образом, для выработки оптимального управления сложной системой, каковой является экономический процесс, необходим системный анализ и построение математической модели, которая должна отражать связи между отдельным зависимым параметром и группой влияющих на него показателей, а также связи внутри этой группы, что возможно осуществить методами множественного корреляционного и регрессионного анализа статистических данных. Итак, решение задачи оптимального управления состоит из таких этапов:

* многомерный экономико-статистический анализ показателей производственно–хозяйственной деятельности предприятий;
* составление математической модели задачи оптимизации управления деятельностью предприятий на основании корреляционного и регрессионного анализа статистических данных;
* решение задачи оптимизации, количественное обоснование прогнозируемого результата и рекомендации по его достижению.

Многомерный статистический анализ выполняется средствами надстройки **Excel** «Пакет анализа». Решение нелинейной задачи оптимизации выполняется средствами надстройки **Excel** «Поиск решения».

Для успешного выполнения задания приводятся необходимые сведения из теории вероятностей и математической статистики, и устанавливается связь между параметрами теоретического и статистического распределения изучаемых факторных и результативных признаков.

**I. Статистический анализ в Excel**

**§ 1.1 Очистка информации от засорения**

При статистическом анализе экономической информации принято считать, что экономические показатели подчиняются нормальному закону распределения. Однако на практике это не всегда верно. Наблюдаются отклонения как односторонние, так и двусторонние. Во избежание искажения значений характеристик распределения при обработке информации необходимо очистить ее от засорения случайными отклонениями. Метод выявления аномальных наблюдений и их удаления из совокупности при обработке многомерной статистической информации может привести к отбрасыванию слишком большого количества точек наблюдения. Известны более четко обоснованные методы обнаружения засорения: метод Смирнова–Граббса проверки максимального наблюдения, критерий Граббса для обнаружения одного экстремального наблюдения, критерий исключения нескольких грубых ошибок как обобщение критерия Граббса. Все они применяются к упорядоченной совокупности (вариационному ряду):

 (N25).

Для проверки максимального и минимального значений на наличие грубой ошибки используются критерии

 и  ,

где  ,  , .

При N>25 экстремальные значения могут быть проверены по критерию S:

,

где – стандартное отклонение, определенное для всей выборки ;

 – предполагаемый выброс.

При Sрасч < Sкр гипотеза H0: – выброс отвергается, в противном случае экстремальное значение  считается грубой ошибкой и из дальнейшего рассмотрения исключается. Критические значения критерия S определяются по таблице . При уровне значимости  Sкр так зависит от объема выборки N : значениям N = 30 ; 50 ; 100 ; 1000 соответствуют Sкр = 2,929 ; 3,082 ; 3,283 ; 3,884 .

Парный корреляционный и регрессионный анализ удобно выполнять средствами Excel и надстройки «Пакет анализа» (в меню – Сервис– Анализ данных ).

В данной работе я выполяю статистический анализ совокупности таких показателей производственно-хозяйственной деятельности предприятий (Приложение 1): производительность труда (среднегодовая выработка продукции на одного работника), тыс. грн. **Y1**, премии и вознаграждения на одного работника, % **X8**, среднегодовая численность ППП, чел.**X11**, среднегодовая стоимость основных производственных фондов(ОПФ), тыс. грн. **X12**, среднегодовой фонд заработной платы ППП, тыс. грн. **X13**, непроизводственные расходы, %**X17**.

Выполняю проверки статистических данных на «засорение»:

– копирую все значения показателей на чистый лист;

– упорядочиваю их по возрастанию, выделяю весь столбец без заголовка и нажимаю на панели кнопку сортировки ;

– устанавливаю курсор под последним значением и ввожу функцию Статистическая

– СРЗНАЧ, а затем СТАНДОТКЛ .

– вычисляю значение статистики Sрасч по найденным характеристикам для наибольшего значения, которое нужно подставить в формулу вместо x1 и проверить гипотезу H0 : наибольшее (последнее в столбце) значение – выброс;

– если Sрасч > Sкр (0,05; 50) = 3,082 , это значение является выбросом, и необходимо проверить предыдущее значение , только при этом следует заново определить среднее значение и стандартное отклонение, но уже исключив выброс, как это и выполнено в приведенной таблице;

– проверку на выброс продолжаю до первого значения, для которого гипотеза H0 окажется неправдоподобной, т.е. для которого значение Sрасч окажется меньше Sкр;

– такую же проверку выполняю начиная с наименьшего (первого в столбце) значения, помня о том, что критерий S имеет двустороннюю критическую область, и поэтому следует рассматривать модуль Sрасч.

Такие проверки выполняю для всех показателей. В итоге на новый лист переношу исходные статистические данные, и исключить полностью каждую строку, в которой есть выброс хотя бы одного из показателей. Весь последующий статистический анализ провожу только по очищенным данным. Данные сохраняю в Excel на листе под названием «Очистка от засорения».

**§ 1.2 Проверка закона распределения**

Предварительный анализ статистических данных заключается в проверке соответствия их предположению о нормальном распределении параметров, для чего строю гистограмму и определяю выборочные числовые характеристики. Для построения гистограммы выполняю такую последовательность действий:

размещаю на рабочем листе **Excel** статистические данные наблюдений (без выбросов);

**Сервис – Анализ данных – Гистограмма** (рис.1);

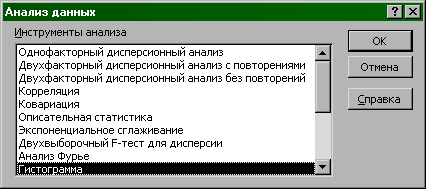


Рис.1.Выбор инструмента анализа.

- в появившемся диалоговом окне **Гистограмма** ввожу в поле **Входные данные** интервал (диапазон) ячеек, содержащий исходные данные, и отмечаю поле **Метки,** т.к., таблица данных имеет заголовки;

- ввожу в поле **Параметры выхода** адрес ячейки, с которой должны размещаться выходные данные (**выходной интервал**) и щелкаю пункт **Вывод графика**;

- OK.

Гистограммы строю для всех признаков статистических данных и сравниваю их с кривой нормального распределения с целью убедиться, что закон распределения каждого признака близок к нормальному, как на приведенной гистограмме.

**Числовые характеристики** для всех признаков оцениваются по выборке с помощью инструмента анализа **Описательная статистика**., вызов которого осуществляется аналогично (см. рис.1 ). В появившемся диалоговом окне **Описательная статистика** ввожу таким же образом **Входные данные** и **Параметры вывода**, только вместо пункта **Вывод графика** отмечаю пункт **Итоговая статистика**.

Результаты применения инструмента **Описательная статистика** к данным наблюдений по результативному признаку **Y1** и выбранным факторным признакамприведены на листе **Excel** под названием «Проверка закона распределения».

Как видно, результаты **Описательной статистики** дают возможность оценить справедливость предположения о нормальном распределении признаков: эксцесс и асимметричность невелики, хотя и отличаются от 0. Нормальный закон распределения факторных признаков подтверждается еще и тем, что значения медианы и моды у них совпадают или близки.

**§ 1.3 Корреляционный анализ**

**Предварительный анализ тесноты взаимосвязи параметров** многомерной модели осуществляю по оценке корреляционной матрицы генеральной совокупности **X** по наблюдениям. Для этого использую инструмент **Анализ данных** в соответствии со следующим алгоритмом:

- размещаю на рабочем листе **Excel** статистические данные в столбцах с соответствующими заголовками (именами переменных);

**- Сервис – Анализ данных – Корреляция**;

- в появившемся диалоговом окне **Корреляция** в соответствующие поля ввожу с помощью мыши входные данные и параметры вывода (см. рис.3 );

- после щелчка мышью по кнопке **OK** на рабочем листе появится матрица, содержащая оценки **парных коэффициентов корреляции**.

Отбираю для дальнейшего анализа пары переменных, имеющие наибольшие значения парных коэффициентов корреляции

( 0,4 ), учитывая, что чем меньше коэффициент rij , тем слабее их связь. Такими парами в приведенном примере (рис.3) являются: Y1–X11; Y1–X12; X11–X13; X12–X13; X17–X8.

Проверяю значимость коэффициентов корреляции на уровне = 0,05. Поскольку объем выборки для всех признаков одинаков и равен 53, критическое значение **r**кр для всех пар будет одинаково и в соответствии с таблицей Фишера–Иейтса **r**кр = **r**табл (0,05;53)< **r**табл(0,05;50) = 0,273. Поскольку для всех коэффициентов выполняется неравенство  > **r**кр , коэффициенты корреляции всех отобранных пар признаков значимо отличаются от нуля, что подтверждает связь между ними.

Дальнейший анализ статистических данных зависит от размерности принимаемой модели. Простейший вариант – двумерная модель. Учитывая, что в приведенном примере Y1 –результирующий признак, определяющий индекс производительность труда, входит в две пары , следует рассмотреть трехмерную модель Y1–X11–X12,

где X11 – среднегодовая численность ППП, а X12 – среднегодовая стоимость основных производственных фондов(ОПФ). В остальных парах следует определить зависимости между X11 и X13, X12 и X13 , X17 и X8. Здесь X5 –удельный вес рабочих в составе промышленно–производственного персонала, X6 –удельный вес покупных изделий, X7 – коэффициент сменности оборудования.

Таким образом, для математической модели задачи выбора оптимального управления деятельностью предприятия с учетом указанных показателей следует установить зависимости: Y2 = F( X4,X8) – целевая функция;

X6 = φ(X4); X8 = φ(X7); X5 = φ(X7) – ограничения.

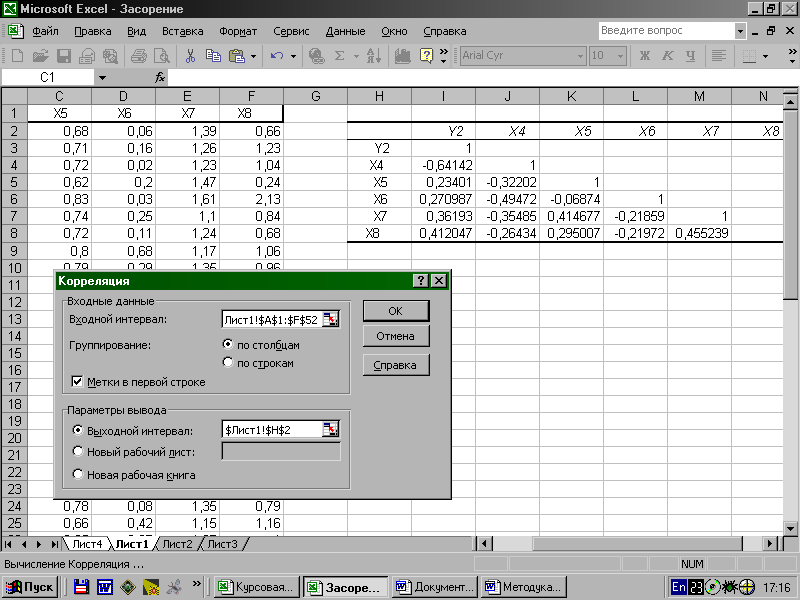


Рис.3.Анализ парной корреляции.

**§1.4 Регрессионный анализ двумерной модели**

В среде **Excel** для двумерного случая линейной регрессии предусмотрено несколько инструментов : статистические функции (КОРРЕЛ, ЛИНЕЙН, ТЕНДЕНЦИЯ и др.) ; инструмент **Регрессия** надстройки **Пакет анализа** ; графические средства при работе с диаграммой – построение **линии тренда**.

С помощью **Пакета анализа** можно получить искомую информацию , следуя такому алгоритму:

- разместить на рабочем листе Excel в двух смежных столбцах с соответствующими заголовками статистические данные по двум признакам, подлежащим исследованию (например, X4 и X6);

**- Сервис – Анализ данных – Регрессия** ;

- в появившемся диалоговом окне **Регрессия** ввести входные данные в поля **Входной интервал Y** (X6)и **Входной интервал X** (X4)и щелкнуть по полю **Метки**, чтобы заголовки не вошли в интервалы данных;

- ввести параметры вывода в поле **Выходной интервал :** адрес левого верхнего угла таблицы результатов или щелкнуть поле **Новый рабочий лист** для вывода на другой лист (см. рис.4);

- для наглядности можно вывести график, щелкнув по полю **График подбора** ;

- **OK**.

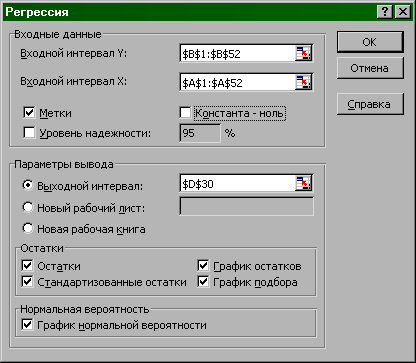


Рис.4.Работа с диалоговым окном Регрессия.

Результат работы инструмента **Регрессия** приведен на рис.5. Итак, выборочное уравнение линейной регрессии X6 на X4 имеет вид:



Выходная таблица содержит коэффициент детерминации **R2 = 0,368802**, что означает, что полученная модель приблизительно на 37% отражает зависимость удельного веса покупных изделий от трудоемкости единицы продукции. Стандартная ошибка (отклонение результата) = 0,118415 означает, что 68% реальных значений результирующего признака x6 находится в диапазоне 0,118415 от линии регрессии. Это следует из того, что условные распределения нормально распределенной генеральной совокупности при фиксировании различных подмножеств компонент являются нормальными.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  |  |  |  |
| ВЫВОД ИТОГОВ | | |  |  |  |  |  |
|  | |  |  |  |  |  |  |
| *Регрессионная статистика* | | |  |  |  |  |  |
| Множественный R | 0,607291 | |  |  |  |  |  |
| R-квадрат | 0,368802 | |  |  |  |  |  |
| Нормированный R-квадрат | 0,35592 | |  |  |  |  |  |
| Стандартная ошибка | 0,118415 | |  |  |  |  |  |
| Наблюдения | 51 | |  |  |  |  |  |
|  |  | |  |  |  |  |  |
| Дисперсионный анализ | | | |  |  |  |  |
|  | *df* | | *SS* | *MS* | *F* | *Значимость F* |  |
| Регрессия | 1 | | 0,401452 | 0,401452 | 28,63014 | 2,3E-06 |  |
| Остаток | 49 | | 0,687078 | 0,014022 |  |  |  |
| Итого | 50 | | 1,088529 |  |  |  |  |
|  |  | |  |  |  |  |  |
|  | *Коэффициенты* | | *Стандартная ошибка* | *t-статистика* | *P-Значение* | *Нижние 95%* | *Верхние 95%* |
| Y-пересечение | 0,557512 | | 0,051111 | 10,90789 | 1,04E-14 | 0,45480 | 0,66022 |
| X4 | -0,85062 | | 0,158973 | -5,35071 | 2,3E-06 | -1,1701 | -0,5312 |

Рис.5. Результаты регрессионного анализа .

В разделе **Дисперсионный анализ** приведены значения таких величин:

**df** – число степеней свободы ; **SS** –сумма квадратов отклонений ; **MS** – дисперсия ; **F** – расчетное значение F–критерия. Поскольку критическое значение критерия Фишера Fкр = 4,03 (m1=1; m2=50;) Fрасч =28,63 > Fкр , и, следовательно с вероятностью  гипотеза об отсутствии связи между рассматриваемыми признаками отвергается. Это означает, что уравнение в целом статистически значимо, т.е. хорошо соответствует данным наблюдений.

Нижняя часть таблицы содержит такие сведения :

**Коэффициенты** – оценки параметров  уравнения регрессии;

**Стандартная ошибка** – стандартные отклонения  ;

**t–статистика –** расчетное значение . Таким образом , можно оценить значимость коэффициентов уравнения регрессии, сравнив расчетное значение t – статистики с критическим значением, найденным по распределению Стьюдента при уровне значимости  и m=50 : tкр =2,009 . Поскольку  > tкр для обоих коэффициентов , то они являются статистически значимыми при уровне доверительной вероятности 0,95.

**Нижние 95%** и **Верхние 95%** определяют нижние и верхние границы доверительных интервалов для коэффициентов уравнения регрессии при . Поскольку доверительные интервалы не содержат **0**, это подтверждает значимость коэффициентов уравнения регрессии.

Для получения линии регрессии и ее уравнения в случае двумерной модели удобным инструментом **Excel** является добавление **линии тренда** к точечной диаграмме, построенной на значениях компонент системы двух заданных случайных величин как результатов наблюдения (см. рис.6).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X4 | X6 | |  | | --- | |  | |  |  |  |  |  |  |
| 0,01 | 0,35 |  |  |  |  |  |  |  |
| 0,02 | 0,42 |  |  |  |  |  |  |  |
| 0,17 | 0,5 |  |  |  |  |  |  |  |
| 0,17 | 0,53 |  |  |  |  |  |  |  |
| 0,18 | 0,68 |  |  |  |  |  |  |  |
| 0,18 | 0,32 |  |  |  |  |  |  |  |
| 0,19 | 0,4 |  |  |  |  |  |  |  |
| 0,22 | 0,54 |  |  |  |  |  |  |  |
| 0,23 | 0,4 |  |  |  |  |  |  |  |
| 0,23 | 0,42 |  |  |  |  |  |  |  |
| 0,23 | 0,47 |  |  |  |  |  |  |  |
| 0,23 | 0,4 |
| 0,24 | 0,56 |
| 0,24 | 0,26 |
| 0,25 | 0,2 |
| 0,25 | 0,33 |
| 0,26 | 0,44 |
| 0,26 | 0,3 |
| 0,26 | 0,27 |
| 0,27 | 0,37 |
| 0,29 | 0,38 |
| 0,29 | 0,34 |
| 0,29 | 0,1 |
| 0,29 | 0,4 |

Рис. 6. Линии тренда.

Алгоритм содержит такие действия:

- разместить на рабочем листе Excel в двух смежных столбцах исходные данные таким образом, чтобы первым был независимый показатель;

**- Вставка – Диаграмма – Точечная** (первый вариант) **– Далее;**

- на закладке **Диапазон данных** ввести диапазон , занимаемый всей таблицей, для чего выделить мышью оба столбца ;

- на закладке **Ряд** ввести в поле **Значения X** диапазон значений независимой величины , а в поле **Значения Y** диапазон значений величины, регрессию которой следует оценить (см.рис.7 );

**Далее –** на закладке **Заголовки** ввести заголовки осей и диаграммы **– Далее –** указать, где разместить диаграмму (на имеющемся листе) – **Готово;**

- откорректировать появившуюся диаграмму, особенно формат осей и надписи, для чего щелкнуть **правой** кнопкой мыши по оси или надписи и в появившемся маленьком диалоговом окне щелкнуть по пункту **Формат оси** (или надписи) ;

- появившемся диалоговом окне **Формат оси** (или надписи ) выбрать нужную закладку и внести необходимые изменения – **OK** ;

- откорректировать полученное корреляционное поле, исключив резко выделяющиеся из общего множества отдельные точки;

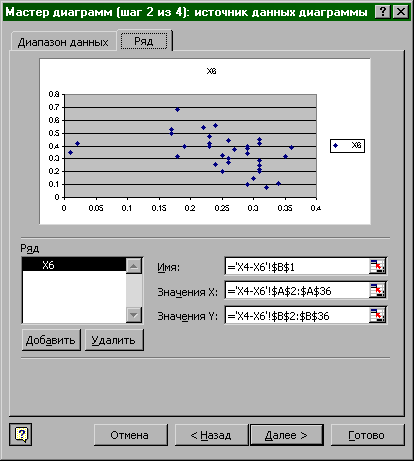


Рис.7. Построение корреляционного поля.

- щелкнуть **правой** кнопкой мыши по любой точке диаграммы и в появившемся диалоговом окне выбрать пункт меню **Добавить линию тренда;**

- в появившемся диалоговом окне на закладке **Тип** выбрать тип зависимости: линейный или полиномиальный (указать порядок приближения);

- щелкнуть по закладке **Параметры** и в появившемся после этого диалоговом окне щелкнуть пункты **показывать уравнение на диаграмме** и **поместить на диаграмму величину достоверности аппроксимации (R^2);**

- записать уравнение регрессии, заменив **y** и **x** на имена результативного и факторного признаков соответственно и оценить значимость полученного уравнения с помощью R^2.

На рис.6 приведены: точечная диаграмма зависимости X6 от X4 и две линии тренда – линейная и нелинейная. Уравнение первой совпадает с уравнением линией регрессии, полученным с помощью инструмента **Регрессия**. Вторая имеет уравнение , т.е. оценку линии регрессии, такого вида:

.

Причем коэффициент детерминации в первом случае равен 0,3688 , а для кубической зависимости R2 = 0,4762 , т.е. предпочтительнее использовать полиномиальную зависимость как лучше согласующуюся со статистическими данными.

Для остальных двух отобранных пар факторных признаков необходимо выполнить такие же действия и получить аналогичные оценки функций регрессии.

**§1.5 Регрессионный анализ трехмерной модели**

Для исследования статистической зависимости одного результирующего признака от двух и более факторных признаков в **Excel** есть две возможности: инструмент **Регрессия** для случая линейной статистической зависимости и непосредственное применение метода наименьших квадратов в случае зависимости любого вида.

Алгоритм применения инструмента **Регрессия** отличается от описанного выше для случая двумерной модели только количеством исходных данных, размещаемых на рабочем листе и соответственно диапазоном входных параметров , вводимом в диалоговом окне **Регрессия** . Выходные данные также отличаются только количеством информации при сохранении их смысла.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Регрессионная статистика* | |  |  |  |  |  |
| Множественный R | 0,762322 |  |  |  |  |  |
| R-квадрат | 0,581135 |  |  |  |  |  |
| Нормированный R-квадрат | 0,563682 |  |  |  |  |  |
| Стандартная ошибка | 50,23613 |  |  |  |  |  |
| Наблюдения | 51 |  |  |  |  |  |
| Дисперсионный анализ | | |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Значимость F* |  |
| Регрессия | 2 | 168064,8 | 84032,39 | 33,2977 | 8,51E-10 |  |
| Остаток | 48 | 121136,1 | 2523,668 |  |  |  |
| Итого | 50 | 289200,9 |  |  |  |  |
|  | *Коэффициенты* | *Стандартная ошибка* | *t-статистика* | *P-Значение* | *Нижние 95%* | *Верхние 95%* |
| Y-пересечение | 225,7848 | 27,41026 | 8,237239 | 9,67E-11 | 170,6728 | 280,8968 |
| X8 | 23,38168 | 10,96783 | 2,131842 | 0,038166 | 1,329382 | 45,43398 |
| X4 | -503,93 | 69,72031 | -7,22788 | 3,29E-09 | -644,112 | -363,748 |

Рис.8. Регрессия Y2 на X4,X8.

На рис.8 приведены результаты применения инструмента **Регрессия** к статистическим данным по признакам X4–X8–Y2 .

Оценка линейной функции регрессии **y2** на **x4,x8** имеет вид:

****

Значение F–критерия Fрасч =33,2977 , что значительно больше Fкр = 3,18 Это означает, что оценка достаточно хорошо согласуется с данными наблюдений. Это подтверждается и достаточно высоким значением коэффициента детерминации R2 = 0,5811351 . Расчетные значения **t** –статистики для свободного члена и коэффициента при **x4** больше **tкр** = 2,009 , что подтверждает их значимость. Для коэффициента при **x8** **tрасч** близко к критическому значению, что ставит под сомнение его значимость.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A** | **B** | **C** | **D** | **E** | **F** | **H** | **I** |
| **1** | **X4** | **X8** | **Y2** | **P(x)** | **ε** | **ε2** | **P2 (x)** | **ε22** |
| **2** | 0,42 | 0,66 | 13,6 | =A$56+B$56\*A2+C$56\*  B2 | =C2-D2 | =E2^2 | =A$59+B$59\*A2+C$59\*B2+D$59\*A2^2+E$59\*B2^2+F$59\*A2\*B2 | =(C2-H2)^2 |
| **3** | 0,51 | 1,23 | 15 | =A$56+B$56\*A3+C$56\*  B3 | =C3-D3 | =E3^2 | =A$59+B$59\*A3+C$59\*B3+D$59\*A3^2+E$59\*B3^2+F$59\*A3\*B3 | =(C3-H3)^2 |
| **4** | 0,38 | 1,04 | 18,1 | =A$56+B$56\*A4+C$56\*  B4 | =C4-D4 | =E4^2 | =A$59+B$59\*A4+C$59\*B4+D$59\*A4^2+E$59\*B4^2+F$59\*A4\*B4 | =(C4-H4)^2 |
| **5** | 0,51 | 0,24 | 21,9 | =A$56+B$56\*A5+C$56\*  B5 | =C5-D5 | =E5^2 | =A$59+B$59\*A5+C$59\*B5+D$59\*A5^2+E$59\*B5^2+F$59\*A5\*B5 | =(C5-H5)^2 |
| **6** | 0,43 | 2,13 | 26,8 | =A$56+B$56\*A6+C$56\*  B6 | =C6-D6 | =E6^2 | =A$59+B$59\*A6+C$59\*B6+D$59\*A6^2+E$59\*B6^2+F$59\*A6\*B6 | =(C6-H6)^2 |
| **7** | 0,43 | 0,84 | 30,1 | =A$56+B$56\*A7+C$56\*  B7 | =C7-D7 | =E7^2 | =A$59+B$59\*A7+C$59\*B7+D$59\*A7^2+E$59\*B7^2+F$59\*A7\*B7 | =(C7-H7)^2 |
| **8** | 0,34 | 0,68 | 32,3 | =A$56+B$56\*A8+C$56\*  B8 | =C8-D8 | =E8^2 | =A$59+B$59\*A8+C$59\*B8+D$59\*A8^2+E$59\*B8^2+F$59\*A8\*B8 | =(C8-H8)^2 |
| **9** | 0,18 | 1,06 | 34,2 | =A$56+B$56\*A9+C$56\*  B9 | =C9-D9 | =E9^2 | =A$59+B$59\*A9+C$59\*B9+D$59\*A9^2+E$59\*B9^2+F$59\*A9\*B9 | =(C9-H9)^2 |

Рис.9. Размещение информации для МНК.

В случае **нелинейной регрессии** специального инструмента в Excel нет, необходимо выполнять действия, предусмотренные методом наименьших квадратов(МНК), используя вычислительные возможности Excel. Расположение исходных данных и формул в таблице Excel приведено на рис.9.

Все формулы вводятся только в верхнюю строку, а затем копируются по всему столбцу. На рис.9 приведены расчеты поиска оценок линейной P(x) и квадратичной P2 (x) функции регрессии. Параметры функции регрессии βj расположены в ячейках A56 ч C56 для линейной зависимости и в ячейках A59 ч F59 для квадратичной зависимости (см. рис.10). Ячейки F53 и I53 содержат значения функций Q – суммы квадратов отклонений.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A** | **B** | **C** | **D** | **E** | **F** | **H** | **I** |
| **50** | 0,02 | 1,14 | 264,8 | =A$56+  B$56\*A50+  C$56\*B50 | =C50-D50 | =E50^2 | =A$59+B$59\*A50+  C$59\*B50+D$59\*A50^2+E$59\*B50^2+F$59\*A50\*B50 | =(C50-H50)^2 |
| **51** | 0,16 | 4,44 | 267,3 | =A$56+  B$56\*A51+  C$56\*B51 | =C51-D51 | =E51^2 | =A$59+B$59\*A51+  C$59\*B51+D$59\*A51^2+E$59\*B51^2+F$59\*A51\*B51 | =(C51-H51)^2 |
| **52** | 0,01 | 1,27 | 355,6 | =A$56+  B$56\*A52+  C$56\*B52 | =C52-D52 | =E52^2 | =A$59+B$59\*A52+  C$59\*B52+D$59\*A52^2+E$59\*B52^2+F$59\*A52\*B52 | =(C52-H52)^2 |
| **53** |  |  |  |  | **Q =** | **=СУММ(F2:**  **F52)** | **Q2 =** | **=СУММ(I2:**  **I52)** |
| **54** |  |  |  |  | **σ =** | **=КОРЕНЬ(F53/51)** | **σ2 =** | **=КОРЕНЬ(I53/51)** |
| **55** | **β0** | **β1** | **β2** |  |  |  |  |  |
| **56** | **225,78481426** | **-503,**  **9302** | **23,381653963** |  |  |  |  |  |
| **57** |  |  |  |  |  |  |  |  |
| **58** | **β0** | **β1** | **β2** | **β3** | **β4** | **β5** |  |  |
| **59** | **247,96413983** | **-930,**  **357130** | **73,537978008** | **1009,39006400157** | **-4,446**  **88827** | **-140,188**  **41146628** |  |  |

Рис.10. Размещение информации для **Поиска решения.**

Значения βj находятся с помощью надстройки Excel **Поиск решения** по такому алгоритму :

* установить курсор на ячейке, содержащей значение функции Q (Q2 ) ;
* **Сервис – Поиск решения ;**
* в появившемся диалоговом окне **Поиск решения** (рис.11) проверить, стоит ли в поле **Установить целевую ячейку** адрес функции Q (Q2), и если нет, то ввести его;
* в поле **Равной** щелкнуть пункт **минимальному значению** ;
* в поле **Изменяя ячейки** ввести диапазон ячеек, которые отведены для значений искомых параметров ;
* щелкнуть по кнопке **Выполнить;**
* если решение найдено, сообщение об этом появится в диалоговом окне, где нужно щелкнуть по пункту **Сохранить найденное решение.** Значения  найдены и находятся в отведенных для них ячейках (рис.10).
* Значение суммы квадратов отклонений найденной оценки функции регрессии от наблюденных значений результирующего признака , т.е. функции Q для линейной регрессии и функции Q2 для квадратичной регрессии , находятся в ячейках F53 и I53, линейная величина отклонений – в ячейке F54 и в ячейке I54.

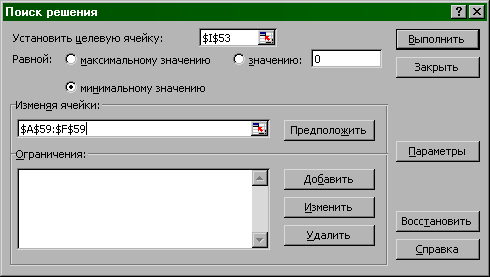


Рис.11. Ввод информации для **Поиска решения.**

Таким образом, коэффициенты линейной функции регрессии P(x) следует считывать из ячеек A56,B56 и С56; коэффициенты нелинейной функции регрессии P2(x) – из ячеек A59  F59. Для рассматриваемого примера линейная функция регрессии совпадает с полученной с помощью инструмента **Регрессия**, а квадратичная

**P2(x) = 247,9641 – 930,3571x4 + 73,538x8 + 1009,39x42 – 4,44689x82 – 140,1884x4x8**

Проверка значимости полученной квадратичной оценки уравнения регрессии выполним так. Определим коэффициент корреляции значений эмпирической функции регрессии и выборочного среднего RyP2(x). Как видно из рис.12 , коэффициент корреляции достаточно большой (0,80921). Выполним еще одну проверку значимости P2(x) с помощью коэффициента детерминации, для чего необходимо вычислить значения Sост, Sфакт .

Размещение нужных формул приведено на рис.12, а промежуточные результаты и значения коэффициента детерминации ниже. Поскольку коэффициент детерминации для случая квадратичной регрессии значительно превосходит коэффициент детерминации для случая линейной регрессии и имеет достаточно большое значение (0,472867), делаем вывод, что квадратичная регрессия достаточно хорошо согласуется со статистическими данными.

Выполним оценку значимости полученного приближения функции в целом с помощью критерия Фишера. Для этого найдем значения критерия Фишера по выборке для рассматриваемых двух видов зависимости (см. рис.12 и 13).

|  |  |  |
| --- | --- | --- |
|  | **R** | **S** |
| **1** | **RyP(x)** | **RyP2(x)** |
| **2** | **=КОРРЕЛ(C2:C52;D2:D52)** | **=КОРРЕЛ(C2:C52;H2:H52)** |
| **3** | |  | | --- | | **Sост** | | |  | | --- | | **Sост** | |
| **4** |  |  |
| **5** | **=F53/48** | **=I53/45** |
| **6**  **7** | |  | | --- | | **Sфакт** | | |  | | --- | | **Sфакт** | |
| **8** |  |  |
| **9** | **=L53/48** | **=N53/45** |
| **10** | **R2** | **R22** |
| **11** | **=1-R5/ (R9 + R5)** | **=1-S5/ (S9 + S5 )** |
| **12** | **Fрасч** | **F2расч** |
| **13** | **=R11\*(51-2-1)/(1-R11)/2** | **=S11\*(51-2-1)/(1-S11)/2** |
| **14** |  |  |
| **15** | **Fкрит =** | **3,205** |

Рис.12.Расчетные формулы

Как видно, расчетное значение F-критерия для квадратичной зависимости значительно превосходит значение Fкрит ,что подтверждает ее значимость. Для линейной зависимости превышение Fрасч не столь велико, что делает снова-таки предпочтительнее квадратичную оценку регрессии **y2** на **x4 и x8 .**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **K** | **L** | **M** | **N** | **O** | **Q** | **R** | **S** |
| |  | | --- | | **1** | |  | ^2 |  | ^2 |  |  | **RyP(x)** | **RyP2(x)** |
| **2** | 66,0145 | 4357,91 | 52,4372 | 2749,66 |  |  | 0,762322 | 0,80921 |
| **3** | 98,0407 | 9611,98 | 63,6085 | 4046,04 |  |  | |  | | --- | | **Sост** | | |  | | --- | | **Sост** | |
| **4** | 36,9723 | 1366,95 | 39,0068 | 1521,53 |  |  |  |  |
| **5** | 121,189 | 14686,7 | 59,1584 | 3499,72 |  |  | 2523,668 | 2218,362 |
| **6** | 36,6828 | 1345,63 | 52,8333 | 2791,36 |  |  | |  | | --- | | **Sфакт** | | |  | | --- | | **Sфакт** | |
| **7** | 66,8451 | 4468,27 | 52,8975 | 2798,14 |  |  |  |  |
| **8** | 25,2325 | 636,678 | 31,6051 | 998,881 |  |  |  |  |
| **9** | -64,2814 | 4132,09 | -63,8871 | 4081,57 |  |  | 3501,349 | 4208,353 |
| **10** | 3,56772 | 12,7286 | 14,147 | 200,138 |  |  | **R2** | **R22** |
| **11** | 43,0760 | 1855,54 | 43,5092 | 1893,05 |  |  | 0,581135 | 0,654822 |
| **12** | -12,1715 | 148,144 | 4,46566 | 19,9421 |  |  | **Fрасч** | **F2расч** |
| **13** | 37,1816 | 1382,47 | 39,3711 | 1550,09 |  |  | 33,29771 | 45,5293 |
| **14** | 68,8203 | 4736,24 | 53,556 | 2868,24 |  |  |  |  |
| **15** | 37,88307 | 1435,127 | 39,90716 | 1592,582 |  |  | **Fкрит** = | 3,205 |

Рис.13.Проверка значимости.

Таким образом, выборочное уравнение регрессии имеет вид :

****

**2. Математическая модель и решение задачи оптимального управления**

Результатом статистического анализа показателей, характеризующих экономический процесс, являются оценки функций регрессии случайных величин (показателей) на одну величину или систему случайных величин. Совокупность всех этих зависимостей является математическим описанием системы и законов перехода ее из одного состояния в другое. Принцип оптимального управления состоит в выборе таких значений показателей, при которых система начинает функционировать наилучшим образом.

Прежде всего, необходимо выбрать критерий оптимальности, т.е. функцию, значение которой должно достичь наибольшего (или наименьшего) из всех возможных в данной ситуации значений. С точки зрения статистического анализа это – один из результативных признаков. Управляемые переменные этой задачи оптимизации – факторные признаки, оказывающие воздействие на результативный признак. Факторные признаки также связаны между собой. Эта связь описывается оценкой функции регрессии одного из факторных признаков на другой факторный признак, полученной в результате регрессионного анализа статистических данных. Выбор таких связанных пар факторных признаков начинается с корреляционного анализа, где отправной точкой является достаточно большой коэффициент парной корреляции. При выборе управляемых переменных задачи следует учесть, что из тесно связанных факторных признаков, особенно с коэффициентом парной корреляции большем 0.5, только один воздействует на результативный признак самостоятельно, а воздействие другого является опосредствованным. Поэтому при выборе математической модели критерия оптимальности учитывается только один из них, а воздействие другого заложено в оценке функции его регрессии на первый фактор.

Оценки функций регрессии факторных признаков (управляемых переменных) друг на друга накладывают ограничения на их возможные значения. Но это не единственные ограничения. Необходимо учесть, что каждый из факторных признаков может принимать значения только в строго ограниченных пределах, которые вытекают из сути самого показателя.

В общем случае математическая модель задачи оптимального управления экономическим процессом, составленная в результате многомерного статистического анализа показателей, содержит:

– целевую функцию

y = f (x1,x2,…,xk) –

функцию регрессии результативного признака Y на факторные признаки X1,X2,…,Xk ;

– ограничения, определяющие область допустимых решений :

xj = φi (xi) , ( i, j = 1,2,…,k) –

функции регрессии факторного признака Xj на факторный признак Xi (i≠J);

 , ( i = 1,2,…,k),

где  и  – нижняя и верхняя границы значений Xi .

Задача оптимизации формулируется следующим образом:

Найти такие значения управляемых переменных, удовлетворяющие всем ограничениям задачи, при которых целевая функция достигает искомого экстремального значения.

В общем случае задача является задачей нелинейного программирования, так как хотя бы одна из функций f (x1,x2,…,xk) или φi (xi) (i = 1,2,…,k) нелинейна относительно управляемых переменных.

Для рассмотренного примера математическая модель имеет вид:

**Y2 =247,9641 – 930,3571\*X4 + 73,538\*X8 + 1009,39\*X4^2 –**

**–4,44689\*X8^2 – 140,1884\*X4\*X8 ––> max**

**X5 = 2,4605\*X7^3 – 10,061\*X7^2+13,815\*X7–5,6226**

**X6 = 18,481\*X4^3– 15,579\*X4^2+2,8223\*X4+0,3562**

**X8 = -86,539\*X7^4+518,28\*X7^3-1141,3\*X7^2+1098,8\*X7-390,07**

**0,2<=X4<=0,5 0,6<=X5<=0,9**

**0 <=X6<=0,7 1 <=X7<= 2 0 <=X8<= 4**

Для решения задачи нелинейной оптимизации следует воспользоваться надстройкой **Excel Поиск решения**. Алгоритм необходимых действий для приведенной математической модели :

1. На рабочем листе Excel расположить исходные данные (см.рис.14).
2. В ячейки A1–E1 записать имена управляемых переменных, в ячейку G1 – имя целевой функции.
3. В ячейки A2 и E2 ввести значения 1, как значения переменных, вошедших в целевую функцию ( при решении нелинейных задач не рекомендуется задавать начальные нулевые значения), значения остальных переменных можно оставлять нулевыми .После окончания поиска решения в ячейках A1–E1 появятся оптимальные значения управляемых переменных, а в ячейке G2 –оптимальное значение целевой функции.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** |
| **1** | **X4** | **X5** | **X6** | **X7** | **X8** |  | **Y2** |  |
| **2** | **0,2** | **0,722559** | **0,445348** | **1,2115** | **1,161317** | **Искомые** | **149,1756** | **ЦФ** |
| **3** | 0,2 | 0,6 | 0 | 1 | 0 | Нижняя гр. |  |  |
| **4** | 0,5 | 0,9 | 0,7 | 2 | 4 | Верхняя гр. |  |  |
| **5** |  | 0,722559 | 0,445348 |  | 1,161316 | Зависимости |  |  |
| **6** |  | **X5=F(X7)** | **X6=F(X4)** |  | **X8=F(X7)** |  |  |  |
| **7** |  |  |  |  |  |  |  |  |

Рис.14.Данные для **Поиска решения.**

1. В ячейки A3–E3 ввести нижние допустимые значения управляемых переменных , в ячейки A4-E4 – верхние.
2. В ячейки B5, C5 , E5 ввести формулы зависимостей, накладывающих ограничения на значения управляемых переменных, в соответствии с математической моделью и адресами (№ ячеек) переменных (рис.15).
3. В ячейку G2 ввести формулу зависимости целевой функции от управляемых переменных (рис.15).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **A** | **B** | **C** | **D** |
| **1** | **X4** | **X5** | **X6** | **X7** |
| **2** | 0,2 | 0,722575104348201 | 0,445348 | 1,21149967733427 |
| **3** | 0,2 | 0,6 | 0 | 1 |
| **4** | 0,5 | 0,9 | 0,7 | 2 |
| **5** |  | =2,4605\*D2^3-10,061\*D2^2+13,815\*D2-5,6226 | =18,481\*A2^3-15,579\*A2^2+2,8223\*A2+0,3562 |  |
|  |  |  |  |  |

|  |  |  |
| --- | --- | --- |
|  | E |  |
| 1 | X8 |  |
| **2** | 1,16131678195123 |  |
| **3** | 0 |  |
| **4** | 4 |  |
| **5** | =-86,539\*D2^4+518,28\*D2^3-1141,3\*D2^2+1098,8\*D2-390,07 |  |
|  |  |  |

|  |  |
| --- | --- |
|  | G |
| 1 | Y2 |
| **2** | =247,9641-930,036\*A2+73,538\*E2+1009,39\*A2^2-4,44689\*E2^2-140,\*A2\*E2 |

1. Вызвать **Сервис – Поиск решения** .

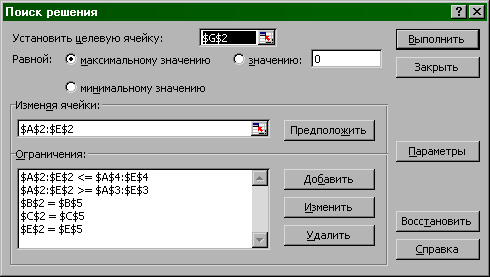


Рис.16. Компьютерная модель задачи.

1. В диалоговом окне ввести необходимые данные (см. рис.16). Для ввода **Ограничений** щелкнуть по кнопке **Добавить** и в появившемся диалоговом окне ввести необходимые ссылки и знаки неравенств, как указано на рис.16.

9. **Выполнить**.

10. Проанализировать полученные результаты и выработать рекомендации по обеспечению оптимального управления.

Как видно из рис.14, оптимальное решение при данных ограничениях и зависимостях свелось к таким результатам.

Максимальное значение индекса снижения себестоимости продукции (**Y2 = 149,1756 )** достигается при таких значениях признаков:

трудоемкость единицы продукции **X4** = 0,5 ;

удельный вес рабочих в составе ППП **X5** = 0,7226 ;

удельный вес покупных изделий **X6** = 0,4453 ;

коэффициент сменности оборудования **X7** = 1,2114 ;

премии и вознаграждения на одного работника **X8** = 1,1613 .

**Литература**

1. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 1997. – 248 с.
2. Дубров А.М. , Мхитарян В.С. , Трошин Л.И. Многомерные статистические методы : Учебник. – М.: Финансы и статистика, 2000. – 352 с.
3. Колемаев В.А. и др. Теория вероятностей и математическая статистика: Учеб. пособие для экон. спец. вузов / В.А.Колемаев, О.В.Староверов, В.Б.Турундаевский ; Под ред. В.А.Колемаева . – М.: Высш. шк. , !991. – 400 с.
4. Исследование операций в экономике : Учебн. пособие для вузов / Н.Ш.Кремер, Б.А. Путко, И.М. Тришин, М.Н.Фридман ; Под ред. Н.Ш.Кремера. – М.: Банки и биржи, ЮНИТИ, 1997. – 407 с.
5. Сивец С.А. Статистические методы в оценке недвижимости и бизнесе. Учебно-практическое пособие по статистике для оценщиков. – Запорожье, 2001. – 320 с.
6. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере / Под ред. Фигурнова В.Э. – М.: ИНФРА , 1998. – 528 с.
7. Акулич И.Л. Математическое программирование в примерах и задачах: Учеб. пособие для студентов эконом. спец. вузов. – М.: Высш. шк. , 1986. – 319 с.
8. Савицкая Г.В. Анализ хозяйственной деятельности предприятия. – Минск: ООО «Новое знание», 2000. – 668 с.
9. Ларсен, Рональд У. Инженерные расчеты в Excel.: Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 544 с.
10. Гурман В.Е. Теория вероятностей и математическая статистика. Учеб. Пособие для втузов. М., « Высш. школа», 1977. – 479 с.
11. Математическая статистика: Учебник / Иванова В.М., Калинина В.Н., Нешумова Л.А. и др. – М.: Высш. школа, 1981. –371 с.
12. Курицкий Б.Я. Поиск оптимальных решений средствами Excel 7.0. – СПб.: BHV – Санки-Петербург, 1997. – 384 с.

**Приложение**

**Показатели производственно-хозяйственной деятельности предприятий**

В таблице приведены следующие показатели производственно-хозяйственной деятельности предприятий

****- производительность труда (среднегодовая выработка продукции на одного работника), тыс. грн.;

**** - индекс снижения себестоимости продукции;

**Y3** - рентабельность, %;

**X4** - трудоемкость единицы продукции;

- удельный вес рабочих в составе промышленно-

производственного персонала (ППП);

- удельный вес покупных изделий;

- коэффициент сменности оборудования;

- премии и вознаграждения на одного работника, %;

- удельный вес потерь от брака, %;

- фондоотдача, грн;

- среднегодовая численность ППП, чел.;

- среднегодовая стоимость основных производственных фондов (ОПФ), тыс. грн.;

- среднегодовой фонд заработной платы ППП, тыс. грн.;

- фондовооруженность труда, грн.;

- оборачиваемость нормируемых оборотных средств, дни;

- оборачиваемость ненормируемых оборотных средств, дни;

- непроизводственные расходы, %.