ВСЕРОССИЙСКИЙ ЗАОЧНЫЙ ФИНАНСОВО-ЭКОНОМИЧЕСКИЙ ИНСТИТУТ

КАФЕДРА СТАТИСТИКИ

О Т Ч Е Т

о результатах выполнения

компьютерной лабораторной работы

Автоматизированный априорный анализ статистической совокупности в среде MS Excel

Вариант № 62

Выполнил:

Проверил:

Брянск 2009 г.

**Постановка задачи**

При проведении статистического наблюдения за деятельностью предприятий корпорации получены выборочные данные о среднегодовой стоимости основных производственных фондов и выпуске продукции за год по 32-м предприятиям, выпускающим однотипную продукцию (выборка 10%-ная, механическая).

В статистическом исследовании эти предприятия выступают как ***единицы выборочной совокупности***. ***Генеральную совокупность*** образуют ***все предприятия корпорации***. Анализируемые признаки предприятий – *Среднегодовая стоимость основных производственных фондов* и *Выпуск продукции* – изучаемые признаки единиц совокупности.

Для автоматизации статистических расчетов используются средства электронных таблиц процессора Excel.

Выборочные данные представлены на ***Листе 1*** ***Рабочего файла*** в табл.1 (ячейки B4:C35):

|  |  |  |
| --- | --- | --- |
| Номер предприятия | Среднегодовая стоимость основных производственных фондов, млн.руб. | Выпуск продукции, млн. руб. |
| 1 | 3446,00 | 3296,00 |
| 2 | 4054,00 | 3616,00 |
| 3 | 4182,00 | 4032,00 |
| 4 | 4406,00 | 4480,00 |
| 5 | 2870,00 | 2240,00 |
| 6 | 4630,00 | 3840,00 |
| 7 | 4758,00 | 5184,00 |
| 8 | 3574,00 | 3520,00 |
| 9 | 4374,00 | 4128,00 |
| 10 | 5046,00 | 5152,00 |
| 11 | 1910,00 | 4800,00 |
| 12 | 5526,00 | 5440,00 |
| 13 | 4214,00 | 4288,00 |
| 14 | 4630,00 | 4672,00 |
| 15 | 5302,00 | 5664,00 |
| 16 | 6070,00 | 6080,00 |
| 17 | 4534,00 | 4096,00 |
| 18 | 5014,00 | 4864,00 |
| 19 | 3990,00 | 3040,00 |
| 20 | 5078,00 | 4160,00 |
| 21 | 5654,00 | 5600,00 |
| 22 | 3894,00 | 3168,00 |
| 23 | 3094,00 | 2976,00 |
| 24 | 5174,00 | 4768,00 |
| 25 | 4630,00 | 4160,00 |
| 26 | 4310,00 | 3936,00 |
| 27 | 3350,00 | 2560,00 |
| 28 | 4502,00 | 4000,00 |
| 29 | 5206,00 | 4384,00 |
| 30 | 6070,00 | 1600,00 |
| 31 | 4950,00 | 4160,00 |
| 32 | 3638,00 | 3712,00 |

В процессе исследования совокупности необходимо решить ряд задач***.***

I. Статистический анализ ***выборочной совокупности***

1. Выявить наличие среди исходных данных резко выделяющихся значений признаков (аномалий в данных) и исключить их из выборки.
2. Рассчитать обобщающие статистические показатели совокупности по изучаемым признакам: среднюю арифметическую (), моду (**Мо**), медиану (**Ме**), размах вариации (**R**), дисперсию(), среднее квадратическое отклонение (), коэффициент вариации (**V*σ***).

1. На основе рассчитанных показателей в предположении, что распределения единиц по обоим признакам близки к нормальному, оценить:

а) степень колеблемости значений признаков в совокупности;

б) степень однородности совокупности по изучаемым признакам;

в) количество попаданий индивидуальных значений признаков в диапазоны (), (), ()..

1. Сравнить распределения единиц совокупности по двум изучаемым признакам на основе анализа:

а) колеблемости признаков;

б) однородности единиц;

в) надежности (типичности) средних значений признаков.

1. Построить интервальный вариационный ряд и гистограмму распределения единиц совокупности по признаку *Среднегодовая стоимость основных производственных фондов* и установить характер (тип) этого распределения.

II. Статистический анализ ***генеральной совокупности***

1. Рассчитать генеральную дисперсию **,** генеральное среднее квадратическое отклонение и ожидаемый размах вариации признаков **RN**. Сопоставить значения генеральной и выборочной дисперсий.

1. Для изучаемых признаков рассчитать:

а) среднюю ошибку выборки;

б) предельные ошибки выборки для уровней надежности P=0,683, P=0,954 и границы, в которых будут находиться средние значения признака в генеральной совокупности при заданных уровнях надежности.

1. Рассчитать коэффициенты асимметрии **As** и эксцесса **Ek**. На основе полученных оценок охарактеризовать особенности формы распределения единиц генеральной совокупности по каждому из изучаемых признаков.

III. Экономическая интерпретация результатов статистического исследования предприятий

В этой части исследования необходимо ответить на ряд вопросов.

1. Типичны ли образующие выборку предприятия по значениям изучаемых экономических показателей?
2. Каковы наиболее характерные для предприятий значения показателей среднегодовой стоимости основных фондов и выпуска продукции?
3. Насколько сильны различия в экономических характеристиках предприятий выборочной совокупности? Можно ли утверждать, что выборка сформирована из предприятий с достаточно близкими значениями по каждому из показателей?
4. Какова структура предприятий выборочной совокупности по среднегодовой стоимости основных фондов? Каков удельный вес предприятий с наибольшими, наименьшими и типичными значениями данного показатели? Какие именно это предприятия?
5. Носит ли распределение предприятий по группам закономерный характер и какие предприятия (с более высокой или более низкой стоимостью основных фондов) преобладают в совокупности?
6. Каковы ожидаемые средние величины среднегодовой стоимости основных фондов и выпуска продукции на предприятиях корпорации в целом? Какое максимальное расхождение в значениях каждого показателя можно ожидать?

**2. Выводы по результатам выполнения лабораторной работы[[1]](#footnote-1)**

I. I. Статистический анализ ***выборочной совокупности***

**Задача 1**

**Вывод:**

Количество аномальных единиц наблюдения (табл.2) равно 2, номера предприятий 11; 30.

**Задача 2**

Рассчитанные выборочные показатели представлены в двух таблицах — табл.3 и табл.5. На основе этих таблиц формируется ***единая таблица*** (табл.8) значений выборочных показателей, перечисленных в условии Задачи 2.

Таблица 8

Описательные статистики выборочной совокупности

|  |  |
| --- | --- |
| Обобщающие статистические показатели совокупности по изучаемым признакам | Признаки |
| *Среднегодовая стоимость* *основных производственных* *фондов*  | *Выпуск продукции*  |
| Средняя арифметическая (), *млн. руб.* | *4470,00* | *4173,87* |
| Мода (Мо),  *млн. руб.* | *4630,00* | *4160,00* |
| Медиана (Ме),  *млн. руб.* | *4518,00* | *4144,00* |
| Размах вариации (R),  *млн. руб.* | *3200,00* | *3840,00* |
| Дисперсия () | *579106,13* | *824093,58* |
| Среднее квадратическое отклонение (), *млн. руб.* | *760,99* | *907,79* |
| Коэффициент вариации (**V*σ***), *%* | *17,02* | *21,75* |

**Задача 3**

3а). ***Степень колеблемости признака*** определяется по значению коэффициента вариации **V*σ***в соответствии с оценочной шкалой колеблемости признака:

**0%<V*σ*40%**  - колеблемость незначительная;

**40%< V*σ*60%**  - колеблемость средняя (умеренная);

**V*σ*>60%** - колеблемость значительная.

**Вывод:**

Для признака Среднегодовая стоимость основных производственных фондов показатель V =17,02%. Так как значение показателя лежит в диапазоне 0%<V 40% оценочной шкалы, следовательно, колеблемость незначительная.

Для признака Выпуск продукции показатель V =21,75%. Так как значение показателя лежит в диапазоне 0%<V 40% оценочной шкалы, следовательно, колеблемость незначительная.

3б). ***Степень однородности совокупности*** по изучаемому признакудля нормального и близких к нормальному распределений устанавливается по значению коэффициента вариации **V*σ*.** Если **V*σ*33%**, то по данному признаку расхождения между значениями признака невелико. Если при этом единицы наблюдения относятся к одному определенному типу, то изучаемая совокупность однородна.

**Вывод:**

Для признака Среднегодовая стоимость основных производственных фондов показатель следовательно, по данному признаку выборочная совокупность однородна.

Для признака Выпуск продукции показатель , следовательно, по данному признаку выборочная совокупность однородна

3в). Для оценки количества попаданий индивидуальных значений признаков ***xi*** в тот или иной диапазон отклонения от средней , а также для выявления структуры рассеяния значений ***xi*** по 3-м диапазонам формируется табл.9 (с конкретными числовыми значениями границ диапазонов).


## Таблица 9

Распределение значений признака по диапазонам рассеяния признака относительно



|  |  |  |  |
| --- | --- | --- | --- |
|  | Границы диапазонов, млн. руб. | Количество значений xi, находящихся в диапазоне | Процентное соотношение рассеяния значений xi по диапазонам, % |
|  | Первый признак | Второй признак | Первый признак | Второй признак | Первый признак | Второй признак |
| А | 1 | 2 | 3 | 4 | 5 | 6 |
|  | [3709,01; 5230,99] | [3266,07; 5081,66] | 20 | 19 | 66,66 | 63,33 |
|  | [2948,02; 5991,98] | [2358,27; 5989,46] | 28 | 28 | 93,33 | 93,33 |
|  | [2187,03; 6752,97] | [1450,48; 6897,25] | 30 | 30 | 100,00 | 100,00 |

На основе данных табл.9 структура рассеяния значений признака по трем диапазонам (графы 5 и 6) сопоставляется со структурой рассеяния ***по правилу «трех сигм»***, справедливому для нормальных и близких к нему распределений:

**68,3%** значений располагаются в диапазоне (),

**95,4%** значений располагаются в диапазоне (),

**99,7%** значений располагаются в диапазоне ().

Если полученная в табл. 9 структура рассеяния ***хi*** по 3-м диапазонам незначительно расходится с правилом «трех сигм», можно предположить, что распределение единиц совокупности по данному признаку близко к нормальному.

***Расхождение с правилом «трех сигм»*** ***может быть существенным***. Например, менее 60% значений ***хi*** попадают в центральный диапазон () или значительно более 5% значения ***хi*** выходит за диапазон (). В этих случаях распределение нельзя считать близким к нормальному.

**Вывод**:

Сравнение данных графы 5 табл.9 с правилом «трех сигм» показывает на их незначительное (существенное) расхождение, следовательно, распределение единиц совокупности по признаку Среднегодовая стоимость основных производственных фондов можно (нельзя) считать близким к нормальному.

Сравнение данных графы 6 табл.9 с правилом «трех сигм» показывает на незначительное (существенное) расхождение, следовательно, распределение единиц совокупности по признаку Выпуск продукции можно (нельзя) считать близким к нормальному.

**Задача 4**

Для ответа на вопросы 4а) – 4в) необходимо воспользоваться табл.8 и ***сравнить*** величины показателей для ***двух признаков***.

Для сравнения степени колеблемости значений изучаемых признаков, степени однородности совокупности по этим признакам, надежности их средних значений используются коэффициенты вариации **V*σ*** признаков.

**Вывод:**

Так как V для первого признака больше (меньше), чем V для второго признака, то колеблемость значений первого признака больше (меньше) колеблемости значений второго признака, совокупность более однородна по первому (второму) признаку, среднее значение первого признака является более (менее) надежным, чем у второго признака.

**Задача 5**

Интервальный вариационный ряд распределения единиц совокупности по признаку *Среднегодовая стоимость основных производственных фондов* представлен в табл.7, а его гистограмма и кумулята – на рис.2.

Возможность отнесения распределения признака «*Среднегодовая стоимость основных производственных фондов*» к семейству нормальных распределений устанавливается путем анализа формы гистограммы распределения. Анализируются количество вершин в гистограмме, ее асимметричность и выраженность «хвостов», т.е. частоты появления в распределении значений, выходящих за диапазон **().**

1. При анализе формы гистограммы прежде всего следует оценить распределение вариантов признака по интервалам (группам). Если на гистограмме четко прослеживаются два-три «горба» частот вариантов, это говорит о том, что значения признака концентрируются сразу в нескольких интервалах, что не соответствует нормальному закону распределения.

Если гистограмма имеет ***одновершинную*** форму, есть основания предполагать, что выборочная совокупность может иметь характер распределения, близкий к нормальному.

2. Для дальнейшего анализа формы распределения используются описательные параметры выборки – показатели центра распределения (, **Mo**, **Me**) и вариации (). Совокупность этих показателей позволяет дать качественную оценку близости эмпирических данных к нормальной форме распределения.

Нормальное распределение является *симметричным*, и для него выполняются соотношения:

=Mo=Me

***Нарушение*** этих соотношений свидетельствует о ***наличии асимметрии*** распределения. Распределения с небольшой или умеренной асимметрией в большинстве случаев относятся к нормальному типу.

3. Для анализа длины «хвостов» распределения используется правило «трех сигм». Согласно этому правилу в нормальном и близким к нему распределениях крайние значения признака (близкие к ***хmin*** и ***хmax***) встречаются много реже (5-7 % всех случаев), чем лежащие в диапазоне (). ***Следовательно, по проценту выхода значений признака за пределы диапазона () можно судить о соответствии длины «хвостов» распределения нормальному закону***.

**Вывод:**

1. Гистограмма является одновершинной (многовершинной).

2. Распределение приблизительно симметрично (существенно асимметрично), так как параметры , Mo, Me отличаются незначительно (значительно):

= 4470,00, Mo=4630,00, Me=4518,00.

3. “Хвосты” распределения не очень длинны (являются длинными), т.к. согласно графе 5 табл.9 6,67% вариантов лежат за пределами интервала ( )=(2948,02; 5991,98) млн. руб.

Следовательно, на основании п.п. 1,2,3, можно (нельзя) сделать заключение о близости изучаемого распределения к нормальному.

II. Статистический анализ ***генеральной совокупности***

**Задача 1**

Рассчитанные в табл.3 генеральные показатели представлены в табл.10.

##### Таблица 10

Описательные статистики генеральной совокупности

|  |  |
| --- | --- |
| Обобщающие статистические показатели совокупности по изучаемым признакам | Признаки |
| *Среднегодовая стоимость* *основных производственных* *фондов* | *Выпуск продукции* |
| Стандартное отклонение *, млн. руб.* | 774,00 | 923,32 |
| Дисперсия  | 599075,31 | 852510,60 |
| Асимметричность **As** | -0,15 | 0,04 |
| Эксцесс **Ek** | -0,34 | -0,21 |

Для нормального распределения справедливо равенство

**RN=6*σ*N.**

В условиях близости распределения единиц генеральной совокупности к нормальному это соотношение используется ***для прогнозной оценки размаха вариации признака в генеральной совокупности.***

Ожидаемый размах вариации признаков RN:

- для первого признака RN =4644,00,

- для второго признака RN =5539,92.

Соотношение между генеральной и выборочной дисперсиями:

- для первого признака 1,03, т.е. расхождение между дисперсиями незначительное (значительное);

-для второго признака 1,03, т.е. расхождение между дисперсиями незначительное (значительное).

**Задача** **2**

Применение выборочного метода наблюдения связано с измерением степени достоверности статистических характеристик генеральной совокупности, полученных по результатам выборочного наблюдения. Достоверность генеральных параметров зависит от ***репрезентативности выборки***, т.е. от того, насколько полно и адекватно представлены в выборке статистические свойства генеральной совокупности.

Как правило, статистические характеристики выборочной и генеральной совокупностей не совпадают, а отклоняются на некоторую величину ***ε***, которую называют ***ошибкой выборки*** ***(ошибкой репрезентативности)***. Ошибка выборки – это разность между значением показателя, который был получен по выборке, и генеральным значением этого показателя. Например, разность

= |-|

определяет ошибку репрезентативности для средней величины признака.

Так как ошибки выборки всегда случайны, вычисляют среднюю и предельную ошибки выборки.

1. Для среднего значения признака ***средняя ошибка выборки***  (ее называют также **стандартной ошибкой**) выражает среднее квадратическое отклонение ***σ*** выборочной средней от математического ожидания **M[]** генеральной средней .

Для изучаемых признаков средние ошибки выборки даны в табл. 3:

- для признака *Среднегодовая стоимость основных производственных фондов*

=141,31,

- для признака *Выпуск продукции*

**=**168,57.

2. ***Предельная ошибка выборки***  определяет границы, в пределах которых лежит генеральная средняя . Эти границы задают так называемый ***доверительный интервал генеральной средней***  – случайную область значений, которая с вероятностью **P**, близкой к 1, ***гарантированно содержит*** значение генеральной средней. Эту вероятность называют ***доверительной вероятностью*** или ***уровнем надежности***.

Для уровней надежности **P=0,954**; **P=0,683** оценки предельных ошибок выборки даны в табл. 3 и табл. 4.

Для генеральной средней предельные значения и доверительные интервалы определяются выражениями:

,

Предельные ошибки выборки и ожидаемые границы для генеральных средних представлены в табл. 11.

Таблица 11

Предельные ошибки выборки и ожидаемые границы для генеральных средних

|  |  |  |  |
| --- | --- | --- | --- |
| Доверительнаявероятность**Р** | Коэффииентдоверия**t** | Предельные ошибки выборки, млн. руб. | Ожидаемые границы для средних , млн. руб. |
| для первогопризнака | для второгопризнака | для первогопризнака | для второгопризнака |
| 0,683 | 1 | 143,88 | 171,64 | 4326,124613,88 | 4002,224345,51 |
| 0,954 | 2 | 294,61 | 351,44 | 4175,394764,61 | 3822,424525,31 |

**Вывод:**

Увеличение уровня надежности ведет к расширению (сужению) ожидаемых границ для генеральных средних.

**Задача 3**

Рассчитанныев табл.3значения коэффициентов асимметрии **As** и эксцесса **Ek** даны в табл.10.

***1.Показатель асимметрии As*** оценивает смещение ряда распределения влево или вправо по отношению к оси симметрии нормального распределения.

Если асимметрия правосторонняя (**As**>0) то ***правая часть эмпирической кривой оказывается длиннее левой***, т.е. имеет место неравенство >**Me>Mo,** что означает ***преимущественное появление в распределении более высоких значений признака*** (среднее значение больше серединного **Me** и модального **Mo**).

Если асимметрия левосторонняя (**As**<0), то ***левая часть эмпирической кривой оказывается длиннее правой*** и выполняется неравенство <**Me<Mo,** означающее, что в ***распределении чаще встречаются более низкие значения признака*** (среднее значение меньше серединного **Me** и модального **Mo**).

Чем больше величина |**As**|, тем более асимметрично распределение. Оценочная шкала асимметрии:

|**As**| 0,25 - асимметрия незначительная;

0,25<|**As**|0,5 - асимметрия заметная (умеренная);

|**As**|>0,5 - асимметрия существенная.

**Вывод:**

Для признака Среднегодовая стоимость основных производственных фондов наблюдается незначительная (заметная, существенная) левосторонняя (правосторонняя) асимметрия. Следовательно, в распределении преобладают более низкие значения признака.

Для признака Выпуск продукции наблюдается незначительная (заметная, существенная) левосторонняя (правосторонняя) асимметрия. Следовательно, в распределении преобладают более высокие значения признака.

***2.Показатель эксцесса*** **Ek** характеризует крутизну кривой распределения - ее заостренность или пологость по сравнению с нормальной кривой.

Как правило, ***коэффициент эксцесса вычисляется только для симметричных или близких к ним распределений***.

Если **Ek**>0, то вершина кривой распределения располагается выше вершины нормальной кривой, а форма кривой является более островершинной, чем нормальная. Это говорит о скоплении значений признака в центральной зоне ряда распределения, т.е. ***о преимущественном появлении в данных значений, близких к средней величине***.

Если **Ek**<0, то вершина кривой распределения лежит ниже вершины нормальной кривой, а форма кривой более пологая по сравнению с нормальной. Это означает, что значения признака не концентрируются в центральной части ряда, а рассеяны по всему диапазону от ***xmax*** до ***xmin***.

Для нормального распределения **Ek**=0. Чем больше абсолютная величина |**Ek**|, тем существеннее распределение отличается от нормального.

***При*** **незначительном *отклонении* Ek *от нуля форма кривой эмпирического распределения*** **незначительно** ***отличается от формы нормального распределения.***

**Вывод:**

1. Так как для признака Среднегодовая стоимость основных производственных фондов Ek>0 (Ek<0), то кривая распределения является более островершинной (пологовершинной) по сравнению с нормальной кривой. При этом Ek незначительно (значительно) отличается от нуля (Ek=|0,34|) Следовательно, по данному признаку форма кривой эмпирического распределения значительно (незначительно) отличается от формы нормального распределения.

2.Так как для признака Выпуск продукции Ek>0 (Ek<0), то кривая распределения является более островершинной (пологовершинной) по сравнению с нормальной кривой. При этом Ek незначительно (значительно) отличается от нуля (Ek=|0,21|). Следовательно, по данному признаку форма кривой эмпирического распределения значительно (незначительно) отличается от формы нормального распределения.

**III. Экономическая интерпретация результатов статистического исследования предприятий[[2]](#footnote-2)**

1. *Типичны ли образующие выборку предприятия по значениям изучаемых экономических показателей?*

*Предприятия с резко выделяющимися значениями показателей приведены в табл.2. После их исключения из выборки, оставшиеся 30 предприятий являются типичными (нетипичными) по значениям изучаемых экономических показателей.*

1. *Каковы наиболее характерные для предприятий значения показателей среднегодовой стоимости основных производственных фондов и выпуска продукции?*

Ответ на вопрос следует из анализа данных табл.9, где приведен диапазон значений признака ***()***, содержащий наиболее характерные для предприятий значения показателей.

Для среднегодовой стоимости основных производственных фондов наиболее характерные значения данного показателя находятся в пределах от 3709,01 млн. руб. до 5230,99 млн. руб. и составляют 66,66% от численности совокупности.

 Для выпуска продукции наиболее характерные значения данного показа-теля находятся в пределах от 3266,07 млн. руб. до 5081,66 млн. руб. и составляют 63,33% от численности совокупности.

1. *Насколько сильны различия в экономических характеристиках предприятий выборочной совокупности? Можно ли утверждать, что выборка сформирована из предприятий с достаточно близкими значениями по каждому из показателей?*

Ответы на вопросы следуют из значения коэффициента вариации (табл.8), характеризующего степень однородности совокупности (см. вывод к задаче 3б). Максимальное расхождение в значениях показателей определяется размахом вариации **Rn.** (табл.8).

Для среднегодовой стоимости основных производственных фондов различия в значениях показателя значительны (незначительны). Максимальное расхождение в значениях данного показателя 3200,00 млн. руб.

1. *Какова структура предприятий выборочной совокупности по среднегодовой стоимости основных производственных фондов? Каков удельный вес предприятий с наибольшими, наименьшими и типичными значениями данного показатели? Какие именно это предприятия?*

Структура предприятий представлена в табл.7 Рабочего файла.

Предприятия с наиболее типичными значениями показателя входят в интервал от 3709,01 млн. руб. до 5230,99 млн. руб. Их удельный вес 66,66%. Это предприятия №№ 22, 19, 2, 3, 13, 26, 9, 4, 28, 17, 6, 14, 25, 7, 31, 18, 10, 20, 24, 29.

Предприятия с наибольшими значениями показателя входят в интервал от 5430,00 млн. руб. до 6070,00 млн. руб. Их удельный вес 100,00 %. Это предприятия №№ 12, 21, 16.

Предприятия с наименьшими значениями показателя входят в интервал от 2870,00 млн. руб. до 3510,00 млн. руб. Их удельный вес 13,33%. Это предприятия №№ 5, 23, 27, 1.

1. *Носит ли распределение предприятий по группам закономерный характер и какие предприятия (с более высокой или более низкой стоимостью основных фондов) преобладают в совокупности?*

Ответ на вопрос следует из вывода к задаче 5 и значения коэффициента асимметрии (табл.8).

Распределение предприятий на группы по среднегодовой стоимости основных производственных фондов носит закономерный характер, близкий к нормальному (незакономерный характер). В совокупности преобладают предприятия с более высокой (низкой) стоимостью основных фондов.

1. *Каковы ожидаемые средние величины среднегодовой стоимости основных фондов и выпуска продукции на предприятиях корпорации в целом? Какое максимальное расхождение в значениях каждого показателя можно ожидать?*

Ответ на первый вопрос следует из данных табл.11. Максимальное расхождение в значениях показателя определяется величиной размаха вариации **RN.**

По корпорации в целом ожидаемые с вероятностью 0,954 средние величины показателей находятся в интервалах:

для среднегодовой стоимости основных производственных фондов – от 4175,39 млн. руб. до 4764,61 млн. руб.;

для выпуска продукции - от 3822,42 млн. руб. до 4525,31 млн. руб.;

Максимальные расхождения в значениях показателей:

для среднегодовой стоимости основных производственных фондов -3200,00 млн. руб.;

для выпуска продукции - 3840,00 млн. руб.

**ПРИЛОЖЕНИЕ**

**Результативные таблицы и графики**

|  |  |  |
| --- | --- | --- |
|  |  | Таблица 2 |
| Аномальные единицы наблюдения |
| Номер предприятия | Среднегодовая стоимость основных производственных фондов, млн.руб. | Выпуск продукции, млн. руб. |
| 11 | 1910,00 | 4800,00 |
| 30 | 6070,00 | 1600,00 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | Таблица 3 |
| Описательные статистики |
| По столбцу "Среднегодовая стоимость основных производственных фондов, млн.руб." | По столбцу "Выпуск продукции, млн.руб" |
| *Столбец1* |  | *Столбец2* |  |
|  |  |  |  |
| Среднее | 4470 | Среднее | 4173,866667 |
| Стандартная ошибка | 141,3123385 | Стандартная ошибка | 168,5734857 |
| Медиана | 4518 | Медиана | 4144 |
| Мода | 4630 | Мода | 4160 |
| Стандартное отклонение | 774,00 | Стандартное отклонение | 923,3150071 |
| Дисперсия выборки | 599075,3103 | Дисперсия выборки | 852510,6023 |
| Эксцесс | -0,34 | Эксцесс | -0,21 |
| Асимметричность | -0,152503649 | Асимметричность | 0,042954448 |
| Интервал | 3200 | Интервал | 3840 |
| Минимум | 2870 | Минимум | 2240 |
| Максимум | 6070 | Максимум | 6080 |
| Сумма | 134100 | Сумма | 125216 |
| Счет | 30 | Счет | 30 |
| Уровень надежности(95,4%) | 294,6096545 | Уровень надежности(95,4%) | 351,4440204 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | Таблица 4 |
| Предельные ошибки выборки |
| По столбцу "Среднегодовая стоимость основных производственных фондов, млн.руб." | По столбцу "Выпуск продукции, млн.руб" |
| *Столбец1* |  | *Столбец2* |  |
|  |  |  |  |
| Уровень надежности(68,3%) | 143,8849956 | Уровень надежности(68,3%) | 171,6424447 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | Таблица 5 |
|  | Выборочные показатели вариации  |  |
| По столбцу "Среднегодовая стоимость основных производственных фондов, млн.руб." | По столбцу "Выпуск продукции, млн.руб" |
| Стандартное отклонение | 760,9902321 | Стандартное отклонение | 907,7960025 |
| Дисперсия  | 579106,1333 | Дисперсия  | 824093,5822 |
| Коэффициент вариации, % | 17,02438998 | Коэффициент вариации, % | 21,74952089 |
|  |  |  |  |

|  |  |
| --- | --- |
|  | Таблица 6 |
| *Карман* | *Частота* |
|  | 1 |
| 3510 | 3 |
| 4150 | 5 |
| 4790 | 11 |
| 5430 | 7 |
| 6070 | 3 |
|   |   |

|  |  |  |
| --- | --- | --- |
|  |  | Таблица 7 |
| Интервальный ряд распределения предприятий по стоимости основных производственных фондов |
| *Группа предприятий по стоимости основных фондов* | *Число предприятий в группе* | *Накопленная частость группы.%* |
| 2870-3510 | 4 | 13,33% |
| 3510-4150 | 5 | 30,00% |
| 4150-4790 | 11 | 66,67% |
| 4790-5430 | 7 | 90,00% |
| 5430-6070 | 3 | 100,00% |
|   |   |   |
| Итого | 30 |  |

ВСЕРОССИЙСКИЙ ЗАОЧНЫЙ ФИНАНСОВО-ЭКОНОМИЧЕСКИЙ ИНСТИТУТ

КАФЕДРА СТАТИСТИКИ

О Т Ч Е Т

о результатах выполнения

компьютерной лабораторной работы

Автоматизированный корреляционно-регрессионный анализ взаимосвязи статистических данных в среде MS Excel

Вариант № 62

Выполнил:

Проверил:

Брянск 2009 г.

# 1. Постановка задачи статистического исследования

Корреляционно-регрессионный анализ взаимосвязи признаков является составной частью проводимого статистического исследования деятельности 30-ти предприятий и частично использует результаты ЛР-1.

В ЛР-2 изучается взаимосвязь между факторным признаком *Среднегодовая стоимость основных производственных фондов* (признак **Х**) и результативным признаком *Выпуск продукции* (признак **Y**), значениями которых являются исходные данные ЛР-1 после исключения из них аномальных наблюдений.

|  |  |  |
| --- | --- | --- |
| Номер предприятия | Среднегодовая стоимость основных производственных фондов, млн.руб. | Выпуск продукции, млн. руб. |
| 1 | 3446,00 | 3296,00 |
| 2 | 4054,00 | 3616,00 |
| 3 | 4182,00 | 4032,00 |
| 4 | 4406,00 | 4480,00 |
| 5 | 2870,00 | 2240,00 |
| 6 | 4630,00 | 3840,00 |
| 7 | 4758,00 | 5184,00 |
| 8 | 3574,00 | 3520,00 |
| 9 | 4374,00 | 4128,00 |
| 10 | 5046,00 | 5152,00 |
| 12 | 5526,00 | 5440,00 |
| 13 | 4214,00 | 4288,00 |
| 14 | 4630,00 | 4672,00 |
| 15 | 5302,00 | 5664,00 |
| 16 | 6070,00 | 6080,00 |
| 17 | 4534,00 | 4096,00 |
| 18 | 5014,00 | 4864,00 |
| 19 | 3990,00 | 3040,00 |
| 20 | 5078,00 | 4160,00 |
| 21 | 5654,00 | 5600,00 |
| 22 | 3894,00 | 3168,00 |
| 23 | 3094,00 | 2976,00 |
| 24 | 5174,00 | 4768,00 |
| 25 | 4630,00 | 4160,00 |
| 26 | 4310,00 | 3936,00 |
| 27 | 3350,00 | 2560,00 |
| 28 | 4502,00 | 4000,00 |
| 29 | 5206,00 | 4384,00 |
| 31 | 4950,00 | 4160,00 |
| 32 | 3638,00 | 3712,00 |

В процессе статистического исследования необходимо решить ряд задач.

1. Установить наличие ***статистической связи*** между факторным признаком **Х** и результативным признаком **Y** графическим методом.
2. Установить наличие ***корреляционной связи*** между признаками **Х** и **Y** методом аналитической группировки.
3. Оценить тесноту связи признаков **Х** и **Y** на основе эмпирического корреляционного отношения ***η***.
4. Построить однофакторную линейную регрессионную модель связи признаков **Х** и **Y**, используя инструмент **Регрессия** надстройки **Пакет анализа,** и оценить тесноту связи признаков **Х** и **Y** на основе линейного коэффициента корреляции ***r***.
5. Определить адекватность и практическую пригодность построенной линейной регрессионной модели, оценив:

а) значимость и доверительные интервалы коэффициентов ***а0, а1***;

б) индекс детерминации **R2** и его значимость;

в) точность регрессионной модели.

1. Дать экономическую интерпретацию:

а) коэффициента регрессии ***а1***;

б) коэффициента эластичности ***К*Э**;

в) остаточных величин ***εi***.

1. Найти наиболее адекватное нелинейное уравнение регрессии с помощью средств инструмента **Мастер диаграмм**.

**2. Выводы по результатам выполнения лабораторной работы[[3]](#footnote-3)**

**Задача 1**

Установление наличия ***статистической связи*** между факторным признаком **Х** и результативным признаком **Y** графическим методом.

Статистическая связь является разновидностью стохастической (случайной) связи, при которой с изменением факторного признака **X** ***закономерным*** образом изменяется какой–либо из обобщающих статистических показателей распределения результативного признака **Y**.

**Вывод:**

Точечный график связи признаков (диаграмма рассеяния, полученная в ЛР-1 после удаления аномальных наблюдений) позволяет сделать вывод, что имеет (не имеет) место статистическая связь. Предположительный вид связи – линейная (нелинейная) прямая (обратная).

**Задача 2**

Установление наличия ***корреляционной связи*** между признаками **Х** и **Y** методом аналитической группировки.

Корреляционная связь – важнейший частный случай стохастической статистической связи, когда под воздействием вариации факторного признака **Х** ***закономерно*** изменяются от группы к группе ***средние групповые значения*** результативного признака **Y** (усредняются результативные значения , полученные под воздействием фактора ). Для выявления наличия корреляционной связи используется метод аналитической группировки.

**Вывод:**

Результаты выполнения аналитической группировки предприятий по факторному признаку Среднегодовая стоимость основных производственных фондов даны в табл. 2.2 Рабочего файла, которая показывает, что с увеличением значений факторного признака Х закономерно (незакономерно) увеличиваются (уменьшаются) средние групповые значения результативного признака . Следовательно, между признаками Х и Y существует корреляционная связь.

**Задача 3**

Оценка тесноты связи признаков **Х** и **Y** на основе эмпирического корреляционного отношения.

Для анализа тесноты связи между факторным и результативным признаками рассчитывается показатель ***η*** – эмпирическое корреляционное отношение, задаваемое формулой

,

где и - соответственно межгрупповая и общая дисперсии результативного признака **Y** - *Выпуск продукции* (индекс ***х*** дисперсии означает, что оценивается мера влияния признака **Х** на **Y)***.*

Для качественной оценки тесноты связи на основе показателя эмпирического корреляционного отношения служит шкала Чэддока:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Значение **η** | 0,1 – 0,3 | 0,3 – 0,5 | 0,5 – 0,7 | 0,7 – 0,9 | 0,9 – 0,99 |
| Сила связи | Слабая | Умеренная | Заметная | Тесная | Весьма тесная |

Результаты выполненных расчетов представлены в табл. 2.4 Рабочего файла.

**Вывод:**

Значение коэффициента η =0,9028, что в соответствии с оценочной шкалой Чэддока говорит о весьма тесной степени связи изучаемых признаков.

**Задача 4**

Построение однофакторной линейной регрессионной модели связи изучаемых признаков с помощью инструмента **Регрессия** надстройки **Пакет анализа** и оценка тесноты связи на основе линейного коэффициента корреляции **r**.

**4.1.** Построение регрессионной модели заключается в нахождении аналитического выражения связи между факторным признаком **X** и результативным признаком **Y.**

Инструмент **Регрессия** на основе исходных данных (***xi , yi***), производит расчет параметров ***а0***и ***а1*** уравнения однофакторной линейной регрессии , а также вычисление ряда показателей, необходимых для проверки адекватности построенного уравнения исходным (фактическим) данным.

***Примечание****. В результате работы инструмента* ***Регрессия*** *получены четыре результативные таблицы (начиная с заданной ячейки* ***А75****).* ***Эти таблицы выводятся в Рабочий файл без нумерации, поэтому необходимо присвоить им номера табл.2.5 – табл.2.8 в соответствии с их порядком****.*

**Вывод:**

Рассчитанные в табл.2.7 (ячейки В91 и В92) коэффициенты а0 и а1 позволяют построить линейную регрессионную модель связи изучаемых признаков в виде уравнения -695,5510+1,0894х.

**4.2.** В случае линейности функции связи для оценки ***тесноты связи*** признаков **X** и **Y,** устанавливаемой по построенной модели, используется ***линейный коэффициент корреляции r***.

Значение коэффициента корреляции **r** приводится в табл.2.5 в ячейке **В78** (термин "*Множественный* **R**").

**Вывод:**

Значение коэффициента корреляции r =0,9132, что в соответствии с оценочной шкалой Чэддока говорит о весьма тесной степени связи изучаемых признаков.

**Задача 5**

Анализ адекватности и практической пригодности построенной линейной регрессионной модели.

Анализ адекватности регрессионной модели преследует цель оценить, насколько построенная теоретическая модель взаимосвязи признаков отражает фактическую зависимость между этими признаками, и тем самым оценить практическую пригодность синтезированной модели связи.

Оценка соответствия построенной регрессионной модели исходным (фактическим) значениям признаков **X** и **Y** выполняется в 4 этапа:

1. оценка статистической значимости коэффициентов уравнения ***а0, а1*** и определение их доверительных интервалов для заданного уровня надежности;
2. определение практической пригодности построенной модели на основе оценок линейного коэффициента корреляции ***r*** и индекса детерминации **R2**;
3. проверка значимости уравнения регрессии в целом по **F**-критерию Фишера;
4. оценка погрешности регрессионной модели**.**

**5.1** Оценка статистической значимости коэффициентов уравнения и определение их доверительных интервалов

Так как коэффициенты уравнения ***а0 , а1*** рассчитывались, исходя из значений признаков только для 30-ти пар (***xi , yi***), то полученные значения коэффициентов являются лишь ***приближенными оценками*** фактических параметров связи ***а0 , а1***. Поэтому необходимо:

* + 1. проверить значения коэффициентов на неслучайность (т.е. узнать, насколько они типичны для всей генеральной совокупности предприятий отрасли);
		2. определить (с заданной доверительной вероятностью **0,95** и **0,683**) пределы, в которых могут находиться значения ***а0, а1*** для генеральной совокупности предприятий.

Для анализа коэффициентов ***а0, а1*** линейного уравнения регрессии используется табл.2.7, в которой:

 – значения коэффициентов ***а0, а1*** приведены в ячейках **В91** и **В92** соответственно;

 – рассчитанный уровень значимости коэффициентов уравнения приведен в ячейках **Е91** и **Е92**;

 – доверительные интервалы коэффициентов с уровнем надежности **Р=0,95** и **Р=0,683** указаны в диапазоне ячеек **F91:I92**.

**5.1.1** Определение значимости коэффициентов уравнения

Уровень значимости – это величина ***α***=1–**Р**, где **Р** – заданный уровень надежности (доверительная вероятность).

Режим работы инструмента **Регрессия** использует по умолчанию уровень надежности **Р=0,95.** Для этого уровня надежности **у**ровень значимости равен ***α*** = 1 – 0,95 = **0,05.** Этот уровень значимости считается ***заданным***.

В инструменте **Регрессия** надстройки **Пакет анализа** для каждого из коэффициентов **а0** и ***а1*** вычисляется **у**ровень его значимости ***αр*,** который указан в результативной таблице (табл.**2.7** термин "Р-*значение*"). Если ***рассчитанный*** для коэффициентов ***а0, а1*** уровень значимости ***αр*,** меньше ***заданного*** уровня значимости ***α*= 0,05**, то этот коэффициент признается ***неслучайным*** (т.е. типичным для генеральной совокупности), в противном случае – ***случайным***.

***Примечание****. В случае, если признается случайным свободный член* ***а0****, то уравнение регрессии целесообразно построить заново без свободного члена* ***а0.*** *В этом случае в диалоговом окне* **Регрессия** *необходимо задать те же самые параметры за исключением лишь того, что следует активизировать флажок* ***Константа-ноль*** *(это означает, что модель будет строиться при условии* ***а0=0****). В лабораторной работе такой шаг не предусмотрен.*

*Если незначимым (случайным) является коэффициент регрессии* ***а1,*** *то взаимосвязь между признаками* ***X*** *и* ***Y*** *в принципе не может аппроксимироваться линейной моделью.*

**Вывод:**

Для свободного члена а0 уравнения регрессии рассчитанный уровень значимости есть αр =0,1061. Так как он меньше (больше) заданного уровня значимости α=0,05, то коэффициент а0 признается типичным (случайным).

Для коэффициента регрессии а1 рассчитанный уровень значимости есть αр =0 . Так как он меньше (больше) заданного уровня значимости α=0,05, то коэффициент а1 признается типичным (случайным).

**5.1.2** Зависимость доверительных интервалов коэффициентов уравнения от заданного уровня надежности

Доверительные интервалы коэффициентов ***а0, а1*** построенного уравнения регрессии при уровнях надежности **Р=0,95** и **Р**=**0,683** представлены в табл.2.7, на основе которой формируется табл.2.9.

Таблица 2.9

**Границы доверительных интервалов коэффициентов уравнения**

|  |  |
| --- | --- |
| Коэффициенты | Границы доверительных интервалов |
| Для уровня надежности **Р=0,95** | Для уровня надежности **Р=0,683** |
| нижняя | верхняя | нижняя | верхняя |
| ***а0*** | -1548,8999 | 157,7979 | -1119,9924 | -271,1096 |
| ***а1*** | 0,9012 | 1,2776 | 0,9957 | 1,1830 |

**Вывод:**

В генеральной совокупности предприятий значение коэффициента а0 следует ожидать с надежностью Р=0,95 в пределах -1548,8999 а0 157,7979, значение коэффициента а1 в пределах 0,9012 а1 1,2776. Уменьшение уровня надежности ведет к расширению (сужению) доверительных интервалов коэффициентов уравнения.

* 1. ***Определение практической пригодности построенной регрессионной модели.***

Практическую пригодность построенной модели можно охарактеризовать по величине линейного коэффициента корреляции ***r***:

* + близость к единице свидетельствует о хорошей аппроксимации исходных (фактических) данных с помощью построенной линейной функции связи ;

* + близость к нулю означает, что связь между фактическими данными **Х** и **Y** нельзя аппроксимировать как построенной, так и любой другой линейной моделью, и, следовательно, для моделирования связи следует использовать какую-либо подходящую нелинейную модель.

Пригодность построенной регрессионной модели для практического использования можно оценить и по величине ***индекса детерминации*** **R2**, показывающего, ***какая часть общей вариации признака* Y *объясняется в построенной модели вариацией фактора*** **X.**

В основе такой оценки лежит равенство **R = *r***(имеющее место для линейных моделей связи), а также шкала Чэддока, устанавливающая качественную характеристику тесноты связи в зависимости от величины ***r***.

Согласно шкале Чэддока ***высокая степень тесноты связи*** признаков достигается лишь при >0,7, т.е. при >0,7. Для индекса детерминации **R2** это означает выполнение неравенства **R2*>*0,5.**

При недостаточно тесной связи признаков **X**, **Y** (слабой, умеренной, заметной) имеет место неравенство 0,7, а следовательно, и неравенство **.**

С учетом вышесказанного, практическая пригодность построенной модели связи оценивается по величине **R2** следующим образом:

* + неравенство **R2 >0,5** позволяет считать, что построенная модель пригодна для практического применения, т.к. в ней достигается высокая степень тесноты связи признаков **X** и **Y**, при которой более 50% вариации признака **Y** объясняется влиянием фактора **Х**;
	+ неравенство означает, что построенная модель связи практического значения не имеет ввиду недостаточной тесноты связи между признаками **X** и **Y**, при которойменее 50% вариации признака **Y** объясняется влиянием фактора **Х**, и, следовательно, фактор **Х** влияет на вариацию **Y** в значительно меньшей степени, чем другие (неучтенные в модели) факторы.

Значение индекса детерминации **R2** приводится в табл.2.5 в ячейке **В79** (термин "**R -** *квадрат*").

**Вывод:**

Значение линейного коэффициента корреляции r и значение индекса детерминации R2 согласно табл. 2.5 равны: r =0,9132, R2 =0,8339. Поскольку и , то построенная линейная регрессионная модель связи пригодна (не пригодна) для практического использования.

* 1. ***Общая оценка адекватности регрессионной модели по F-критерию Фишера***

Адекватность построенной регрессионной модели фактическим данным (***xi, yi***) устанавливается по критерию Р.Фишера, оценивающему статистическую значимость (неслучайность) индекса детерминации **R2**.

Рассчитанная для уравнения регрессии оценка значимости **R2** приведена в табл.2.6 в ячейке **F86** (термин "*Значимость* ***F***"). Если она меньше заданного уровня значимости ***α***=**0,05**, то величина **R2** признается неслучайной и, следовательно, построенное уравнение регрессии может быть использовано как модель связи между признаками **Х** и **Y** для ***генеральной совокупности*** предприятий отрасли.

**Вывод:**

Рассчитанный уровень значимости αр индекса детерминации R2 есть αр=0. Так как он меньше(больше) заданного уровня значимости α=0,05, то значение R2 признается типичным (случайным) и модель связи между признаками Х и Y -695,5510+1,0894х применима (неприменима) для генеральной совокупности предприятий отрасли в целом.

* 1. ***Оценка погрешности регрессионной модели***

Погрешность регрессионной модели можно оценить по величине стандартной ошибки построенного линейного уравнения регрессии . Величина ошибки оценивается как среднее квадратическое отклонение по совокупности отклонений исходных (фактических) значений ***yi*** признака **Y** от его теоретических значений , рассчитанных по построенной модели.

Погрешность регрессионной модели выражается в процентах и рассчитывается как величина **.100.**

В адекватных моделях погрешность не должна превышать 12%-15%.

Значение приводится в выходной таблице "*Регрессионная статистика*" (табл.2.5) в ячейке **В81** (термин "*Стандартная ошибка*"), значение – в таблице описательных статистик (ЛР-1, ***Лист 1***, табл.3, столбец 2).

**Вывод:**

Погрешность линейной регрессионной модели составляет .100= .100=9,1749%, что подтверждает (не подтверждает) адекватность построенной модели -695,5510+1,0894х.

**Задача 6**

Дать экономическую интерпретацию:

1) коэффициента регрессии ***а1***;

3) остаточных величин ***i***.

2) коэффициента эластичности ***К*Э**;

**6.1** Экономическая интерпретация коэффициента регрессии а1

В случае линейного уравнения регрессии **=*a0+a1x*** величина коэффициента регрессии ***a1*** показывает, на сколько ***в среднем*** (в абсолютном выражении) изменяется значение результативного признака **Y** при изменении фактора **Х** на единицу его измерения. Знак при ***a1*** показывает направление этого изменения.

**Вывод:**

Коэффициент регрессии а1 =1,0894 показывает, что при увеличении факторного признака Среднегодовая стоимость основных производственных фондов на 1 млн руб. значение результативного признака Выпуск продукции увеличивается (уменьшается) в среднем на 1,0894 млн руб.

**6.2** Экономическая интерпретация коэффициента эластичности

С целью расширения возможностей экономического анализа явления используется коэффициент эластичности , который измеряется в процентах и показывает, ***на сколько процентов*** изменяется ***в среднем*** результативный признак при изменении факторного признака на 1%.

Средние значения и приведены в таблице описательных статистик (ЛР-1, ***Лист 1***, табл.3).

Расчет коэффициента эластичности:

= =1,1667%

**Вывод:**

Значение коэффициента эластичности Кэ=1,1667% показывает, что при увеличении факторного признака Среднегодовая стоимость основных производственных фондов на 1% значение результативного признака Выпуск продукции увеличивается (уменьшается) в среднем на 1,1667 %.

**6.3** Экономическая интерпретация остаточных величин εi

Каждый их остатков характеризует отклонение фактического значения ***yi*** от теоретического значения , рассчитанного по построенной регрессионной модели и определяющего, какого среднего значения следует ожидать, когда фактор **Х** принимает значение ***xi***.

Анализируя остатки, можно сделать ряд практических выводов, касающихся выпуска продукции на рассматриваемых предприятиях отрасли.

Значения остатков ***i*** (таблица остатков из диапазона **А98:С128**) имеют как положительные, так и отрицательные отклонения от ожидаемого в среднем объема выпуска продукции (которые в итоге уравновешиваются, т.е.).

Экономический интерес представляют ***наибольшие расхождения*** между фактическим объемом выпускаемой продукции ***yi*** и ожидаемым усредненным объемом .

**Вывод:**

Согласно таблице остатков максимальное превышение ожидаемого среднего объема выпускаемой продукции имеют три предприятия - с номерами 20, 19, 29 а максимальные отрицательные отклонения - три предприятия с номерами 7, 15, 32. Именно эти шесть предприятий подлежат дальнейшему экономическому анализу для выяснения причин наибольших отклонений объема выпускаемой ими продукции от ожидаемого среднего объема и выявления резервов роста производства.

**Задача 7**

Нахождение наиболее адекватного ***нелинейного*** уравнения регрессии с помощью средств инструмента **Мастер диаграмм.**

Уравнения регрессии и их графики построены для 3-х видов нелинейной зависимости между признаками и представлены на диаграмме 2.1 ***Рабочего файла***.

Уравнения регрессии и соответствующие им индексы детерминации **R2** приведены в табл.2.10 (при заполнении данной таблицы ***коэффициенты уравнений необходимо указывать не в компьютерном формате, а в общепринятой десятичной форме чисел***).

## Таблица 2.10

## Регрессионные модели связи

|  |  |  |
| --- | --- | --- |
| Вид уравнения | Уравнение регрессии | Индексдетерминации **R2** |
| Полином 2-го порядка | 5Е-05*х*2+0,6*х*+201,7 | 0,8353 |
| Полином 3-го порядка | 8Е-08*х*3-0,001*х*2+5,1*х*-5982,3 | 0,8381 |
| Степенная функция | 0,2*х*1,1788 | 0,8371 |

Выбор наиболее адекватного уравнения регрессии определяется максимальным значением индекса детерминации **R2**: ***чем ближе значение R2 к единице, тем более точно регрессионная модель соответствует фактическим данным***.

**Вывод:**

Максимальное значение индекса детерминации R2 =0,8381. Следовательно, наиболее адекватное исходным данным нелинейное уравнение регрессии имеет вид 8Е-08х3-0,001х2+5,1х-5982,3.

**ПРИЛОЖЕНИЕ**

**Результативные таблицы и графики**

|  |  |  |
| --- | --- | --- |
|  |  | Таблица 2.1 |
| Исходные данные |
| Номер предприятия | Среднегодовая стоимость основных производственных фондов, млн.руб. | Выпуск продукции, млн. руб. |
| 5 | 2870,00 | 2240,00 |
| 23 | 3094,00 | 2976,00 |
| 27 | 3350,00 | 2560,00 |
| 1 | 3446,00 | 3296,00 |
| 8 | 3574,00 | 3520,00 |
| 32 | 3638,00 | 3712,00 |
| 22 | 3894,00 | 3168,00 |
| 19 | 3990,00 | 3040,00 |
| 2 | 4054,00 | 3616,00 |
| 3 | 4182,00 | 4032,00 |
| 13 | 4214,00 | 4288,00 |
| 26 | 4310,00 | 3936,00 |
| 9 | 4374,00 | 4128,00 |
| 4 | 4406,00 | 4480,00 |
| 28 | 4502,00 | 4000,00 |
| 17 | 4534,00 | 4096,00 |
| 6 | 4630,00 | 3840,00 |
| 14 | 4630,00 | 4672,00 |
| 25 | 4630,00 | 4160,00 |
| 7 | 4758,00 | 5184,00 |
| 31 | 4950,00 | 4160,00 |
| 18 | 5014,00 | 4864,00 |
| 10 | 5046,00 | 5152,00 |
| 20 | 5078,00 | 4160,00 |
| 24 | 5174,00 | 4768,00 |
| 29 | 5206,00 | 4384,00 |
| 15 | 5302,00 | 5664,00 |
| 12 | 5526,00 | 5440,00 |
| 21 | 5654,00 | 5600,00 |
| 16 | 6070,00 | 6080,00 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  | Таблица 2.2 |
| Зависимость выпуска продукции от среднегодовой стоимости основных фондов  |
| Номер группы | Группы предприятий по стоимости основеных фондов | Число предприятий  | Выпуск продукции |
| Всего | В среднем на одно предприятие |
| 1 | 2870-3510 | 4 | 11072,00 | 2768,00 |
| 2 | 3510-4150 | 5 | 17056,00 | 3411,20 |
| 3 | 4150-4790 | 11 | 46816,00 | 4256,00 |
| 4 | 4790-5430 | 7 | 33152,00 | 4736,00 |
| 5 | 5430-6070 | 3 | 17120,00 | 5706,67 |
| Итого |   | 30 | 125216,00 | 4173,87 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | Таблица 2.3 |
| Показатели внутригрупповой вариации |
| Номер группы | Группы предприятий по стоимости основеных фондов | Число предприятий  | Внутригрупповая дисперсия |
| 1 | 2870-3510 | 4 | 161024,00 |
| 2 | 3510-4150 | 5 | 68239,36 |
| 3 | 4150-4790 | 11 | 138891,64 |
| 4 | 4790-5430 | 7 | 262729,14 |
| 5 | 5430-6070 | 3 | 73955,56 |
| Итого |   | 30 |   |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | Таблица 2.4 |
| Показатели дисперсии и эмпирического корреляционного отношения |
| Общая дисперсия | Средняя из внутригрупповых дисперсия | Межгрупповая дисперсия | Эмпирическое корреляционное отношение |
| 824093,5822 | 152469,0489 | 671624,5333 | 0,902765617 |

**Выходные таблицы**

|  |  |
| --- | --- |
|  | Таблица 2.5 |
| ВЫВОД ИТОГОВ |  |
|  |  |
| *Регрессионная статистика* |
| Множественный R | 0,9132 |
| R-квадрат | 0,833912798 |
| Нормированный R-квадрат | 0,827981112 |
| Стандартная ошибка | 382,9463742 |
| Наблюдения | 30 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | Таблица 2.6 |
| Дисперсионный анализ |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Значимость F* |
| Регрессия | 1 | 20616665,55 | 20616665,55 | 140,5861384 | 1,97601E-12 |
| Остаток | 28 | 4106141,913 | 146647,9255 |  |  |
| Итого | 29 | 24722807,47 |   |   |   |

Таблица 2.7

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Коэффициенты* | *Стандартная ошибка* | *t-статистика* | *P-Значение* |
| Y-пересечение | -695,5510 | 416,5909893 | -1,669625628 | 0,106137752 |
| Переменная X 1 | 1,0894 | 0,09187519 | 11,85690257 | 1,97601E-12 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Нижние 95%* | *Верхние 95%* | *Нижние 68,3%* | *Верхние 68,3%* |
| -1548,899908 | 157,7979239 | -1119,992383 | -271,1096012 |
| 0,901157173 | 1,277553188 | 0,995748659 | 1,182961703 |
|  |  | Таблица 2.8 |
| ВЫВОД ОСТАТКА |  |
|  |  |  |
| *Наблюдение* | *Предсказанное Y* | *Остатки* |
| 1 | 2430,898377 | -190,8983771 |
| 2 | 2674,913938 | 301,0860623 |
| 3 | 2953,788864 | -393,788864 |
| 4 | 3058,366961 | 237,6330386 |
| 5 | 3197,804425 | 322,1955755 |
| 6 | 3267,523156 | 444,4768439 |
| 7 | 3546,398082 | -378,3980824 |
| 8 | 3650,97618 | -610,9761798 |
| 9 | 3720,694911 | -104,6949114 |
| 10 | 3860,132375 | 171,8676254 |
| 11 | 3894,99174 | 393,0082597 |
| 12 | 3999,569838 | -63,56983771 |
| 13 | 4069,288569 | 58,7114307 |
| 14 | 4104,147935 | 375,8520649 |
| 15 | 4208,726032 | -208,7260325 |
| 16 | 4243,585398 | -147,5853982 |
| 17 | 4348,163496 | -508,1634956 |
| 18 | 4348,163496 | 323,8365044 |
| 19 | 4348,163496 | -188,1634956 |
| 20 | 4487,600959 | 696,3990412 |
| 21 | 4696,757154 | -536,7571535 |
| 22 | 4766,475885 | 97,5241149 |
| 23 | 4801,335251 | 350,6647491 |
| 24 | 4836,194617 | -676,1946167 |
| 25 | 4940,772714 | -172,7727141 |
| 26 | 4975,63208 | -591,6320798 |
| 27 | 5080,210177 | 583,7898228 |
| 28 | 5324,225738 | 115,7742622 |
| 29 | 5463,663201 | 136,3367991 |
| 30 | 5916,834956 | 163,1650438 |

**ПРИЛОЖЕНИЕ**

**Результативные таблицы и графики**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  | Таблица 3.1 |
| Исходные данные |
| **Годы** | **Выпуск продукции, млн. руб.** |  | **Месяцы** | **Выпуск продукции, млн. руб.** |
| **1** | **12320,00** |  | **январь** | **1105,00** |
| **2** | **12560,00** |  | **февраль** | **1171,00** |
| **3** | **12950,00** |  | **март** | **1230,00** |
| **4** | **12830,00** |  | **апрель** | **1200,00** |
| **5** | **13065,00** |  | **май** | **1260,00** |
| **6** | **15237,00** |  | **июнь** | **1240,00** |
|  |  |  | **июль** | **1296,00** |
|  |  |  | **август** | **1271,00** |
|  |  |  | **сентябрь** | **1350,00** |
|  |  |  | **октябрь** | **1371,00** |
|  |  |  | **ноябрь** | **1383,00** |
|  |  |  | **декабрь** | **1360,00** |
|  |  |  | **Итого** | **15237,00** |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  | Таблица 3.2 |
| Показатели динамики выпуска продукции |
| **Годы** | **Выпуск продукции, млн. руб.** | **Абсолютный прирост, млн. руб.** | **Темп роста, %** | **Темп прироста, %** | **Абсолютное значение 1% прироста** |
| **цепной** | **базисный** | **цепной** | **базисный** | **цепной** | **базисный** |
| **1-й** | 12320,00 |   |   |   |   |   |   |   |
| **2-й** | 12560,00 | 240,00 | 240,00 | 101,9 | 101,9 | 1,9 | 1,9 | 123,2 |
| **3-й** | 12950,00 | 390,00 | 630,00 | 103,1 | 105,1 | 3,1 | 5,1 | 125,6 |
| **4-й** | 12830,00 | -120,00 | 510,00 | 99,1 | 104,1 | -0,9 | 4,1 | 129,5 |
| **5-й** | 13065,00 | 235,00 | 745,00 | 101,8 | 106,0 | 1,8 | 6,0 | 128,3 |
| **6-й** | 15237,00 | 2 172,00 | 2 917,00 | 116,6 | 123,7 | 16,6 | 23,7 | 130,65 |

Таблица 3.3

Средние показатели ряда динамики

|  |  |  |
| --- | --- | --- |
|

|  |
| --- |
| Средний уровень ряда динамики, млн. руб., |

 | 13160,33 |
|

|  |
| --- |
| Средний абсолютный прирост***,*** млн. руб., |

 | 583,40 |
|

|  |
| --- |
| Средний темп роста, %, |

 | 104,3 |
|

|  |
| --- |
| Средний темп прироста, %, |

 | 4,3 |

Таблица 3.4

Прогноз выпуска продукции на 7-ой год

|  |  |
| --- | --- |
| По среднему абсолютному приросту, млн. руб., | 15820,40 |
|

|  |
| --- |
| По среднему темпу роста, %, |

 | 15892,19 |

|  |  |  |
| --- | --- | --- |
|  |  | Таблица 3.5 |
| Выпуск продукции за 6-ой год |
| **Месяцы** | **Выпуск продукции, млн. руб.** | **Скользящее среднее** |
| **январь** | **1105,00** |  |
| **февраль** | **1171,00** | **1168,67** |
| **март** | **1230,00** | **1200,33** |
| **апрель** | **1200,00** | **1230,00** |
| **май** | **1260,00** | **1233,33** |
| **июнь** | **1240,00** | **1265,33** |
| **июль** | **1296,00** | **1269,00** |
| **август** | **1271,00** | **1305,67** |
| **сентябрь** | **1350,00** | **1330,67** |
| **октябрь** | **1371,00** | **1368,00** |
| **ноябрь** | **1383,00** | **1371,33** |
| **декабрь** | **1360,00** |  |

1. Все статистические показатели необходимо представить в таблицах ***с точностью до 2-х знаков*** после запятой. Таблицы и пробелы в формулировках выводов заполнять ***вручную***. В выводах при выборе альтернативного варианта ответа ***ненужный вариант вычеркнуть***. [↑](#footnote-ref-1)
2. Выводы должны раскрывать ***экономический смысл*** результатов проведенного статистического анализа совокупности предприятий, поэтому ответы на поставленные вопросы задач 1-6, должны носить ***экономический характер*** со ссылками на результаты анализа статистических свойств совокупности (п. 1-5 для выборочной совокупности и п. 1-3 для генеральной совокупности). В Методических указаниях к лабораторной работе №1 (стр.7-9) разъяснено, на основании каких статистических показателей делаются соответствующие экономические выводы. [↑](#footnote-ref-2)
3. Все статистические показатели необходимо представить в таблицах с точностью до 4-х знаков после запятой. Таблицы и пробелы в формулировках выводов заполнять ***вручную***. В выводах при выборе альтернативного варианта ответа ***ненужный вариант*** ***вычеркивается***. [↑](#footnote-ref-3)