

Министерство образования и науки Украины
Севастопольский национальный технический университет



«АНАЛИЗ МОДЕЛИ МНОЖЕСТВЕННОЙ ЛИНЕЙНОЙ РЕГРЕССИИ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к проведению практического занятия №1

по дисциплине «Эконометрия»

для студентов специальности
6.050201 - «Менеджмент организаций»
дневной формы обучения

Севастополь
2008



УДК 658

«Анализ модели множественной линейной регрессии»

методические указания к проведению практического занятия №1 по дисциплине «Эконометрия» для студентов специальности 8.050201 – «Менеджмент организаций» дневной формы обучения / Сост. А.В. Цуканов, Т.А. Кокодей. – Севастополь: Изд-во СевНТУ, 2008г. – 24 с.

Целью методических указаний является закрепление теоретических знаний и получение практических навыков в области многофакторного регрессионного анализа. Методические указания предназначены для студентов специальности «Менеджмент организаций» дневной формы обучения.

Методические указания утверждены на заседании кафедры менеджмента и экономико-математических методов, (протокол № 5 от 23.01.2008 г.).

Допущено учебно-методическим центром СевНТУ в качестве методических указаний.

Рецензент:

Фисун С.Н., канд. техн. наук, доцент кафедры «Кибернетика и вычислительная техника».

СОДЕРЖАНИЕ

1. Цель работы.....	4
2. Теоретические основы и примеры расчётов: линейная модель множественной регрессии (1МНК).....	4
2.1. Оценка параметров модели.....	5
2.2. Проверка коэффициентов на значимость.....	6
2.3. Проверка адекватности уравнения множественной регрессии в целом...	8
2.4. Предпосылки метода наименьших квадратов.....	9
3. Индивидуальные расчётно-практические задания.....	14
4. Содержание отчета о практическом занятии.....	15
Библиографический список.....	16
Приложение А. (справочное) Вспомогательные сведения из высшей математики	17
Приложение Б. (справочное) Статистические таблицы.....	20

1. ЦЕЛЬ РАБОТЫ

Целью данной работы является получение практических навыков оценивания и проверки адекватности линейной модели множественной регрессии.

2. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ И ПРИМЕРЫ РАСЧЁТОВ: ЛИНЕЙНАЯ МОДЕЛЬ МНОЖЕСТВЕННОЙ РЕГРЕССИИ (1МНК)

Термин «регрессия» (термин был впервые использован в работе Пирсона - Pearson, 1908) в переводе означает «отступление», «возврат» и представляет собой количественную меру изменения (уменьшения или увеличения) одного (результативного) признака в зависимости от изменения на определенную величину других коррелирующих с ним (факторных) признаков.

Простая линейная регрессия представляет собой регрессию между двумя переменными y и x , т.е. модель вида $y = f(x) + u = \alpha_0 + \alpha_1 \cdot x + u$, где y — результативный признак; x - признак-фактор, u — случайная ошибка.

В результате построения уравнения регрессии (прямая на рисунке 1) получаются не точные значение (точки y_i), а модельные значения $\tilde{y}_x$, отличающиеся от точных на некоторую величину u_i (ошибку), рисунок 1.

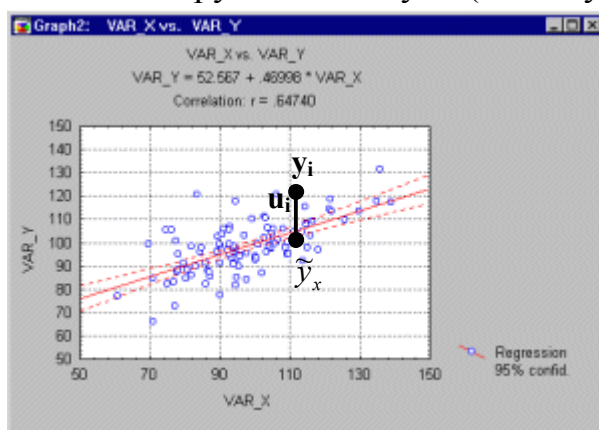


Рисунок 1 – График функции однофакторной линейной регрессии

Множественная регрессия представляет собой регрессию результативного признака с двумя и большим числом факторов, т. е. модель вида $y = f(x_1, x_2, \dots, x_n) + u$.

В уравнении регрессии корреляционная по сути связь признаков представляется в виде функциональной связи, выраженной соответствующей функцией (1).

$$y = \tilde{y}_x + u = \alpha_0 + \alpha_1 \cdot x_1 + \dots + \alpha_n x_n + u, \quad (1)$$

где y — фактическое значение результативного признака;

$\tilde{y}_x$ - модельное значение результативного признака;

x_i - признак-фактор;

α_i — параметр регрессионной модели;

u — случайная величина (ошибка, возмущение, остаток), характеризующая отклонения реального значения результативного признака от теоретического (предсказанные значения “ y ” минус наблюдаемые). Она включает влияние не учтенных в модели факторов, случайных ошибок и особенностей измерения.

Например, специалисты по кадрам обычно используют процедуры множественной регрессии для определения заработной платы, адекватной выполненной работе. Можно определить некоторое количество факторов или параметров, таких, как “размер ответственности” (x_1) или “число подчиненных” (x_2), которые, как ожидается, оказывают влияние на заработную плату y (стоимость работы). Кадровый аналитик затем проводит исследование размеров окладов среди сравнимых компаний на рынке, записывая размер жалования и соответствующие характеристики (т.е. значения параметров) по различным должностям. Эта информация может быть использована при анализе с помощью множественной регрессии для построения регрессионного уравнения в следующем виде

$$y = 0.5 + 3,5 \cdot x_1 + 7,3 \cdot x_2 + u,$$

где y — заработная плата, тыс.грн.;
 x_1 — размер ответственности;
 x_2 — число подчинённых, человек.

Как только эта линия регрессии определена, аналитик оказывается в состоянии построить график ожидаемой оплаты труда и реальных обязательств компании по выплате жалования. Таким образом, аналитик может определить, какие должности недооценены (лежат ниже линии регрессии), какие оплачиваются слишком высоко (лежат выше линии регрессии), а какие оплачены адекватно.

2.1. Оценка параметров модели

Классический подход к оцениванию параметров линейной модели основан на обычном или одношаговом методе наименьших квадратов (МНК).

Этот метод позволяет получить такие оценки параметров, при которых сумма квадратов отклонений фактических значений результативного признака y_i от расчетных (теоретических) $\tilde{y}_i$ минимальна (2).

$$\sum_i (y_i - \tilde{y}_i)^2 \rightarrow \min. \quad (2)$$

Чтобы найти минимум функции (5), надо вычислить производные по каждому из параметров и приравнять их к нулю, т.к. равенство нулю производной — необходимое условие экстремума. В результате получается система уравнений (в матричном виде представлена формулой 3, решение которой и позволяет получить оценки параметров регрессии, формула (1).

$$\alpha = [X^T \cdot X]^{-1} \cdot X^T \cdot Y. \quad (3)$$

Пример 1. Составить линейную модель множественной регрессии y от x_1 и x_2 по данным таблицы 1.



Таблица 1 – Исходные данные

Номер предприятия \ Переменная	1	2	3	4
$x_1, (\%)$	1	1	-1	-1
$x_2, (\%)$	1	-1	1	-1
$y, (\text{тыс. грн.})$	0,5	5	7	12

Этап 1. Составим матрицы X (первый столбец всегда состоит из «1», поскольку в модели присутствует α_0 , второй и третий столбцы – значения x_1 и x_2) и Y (значения y)

$$X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{pmatrix} \quad Y = \begin{pmatrix} 0,5 \\ 5 \\ 7 \\ 12 \end{pmatrix}$$

$$\text{Этап 2. } X^T \cdot X = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

$$\text{Этап 3. } (X^T \cdot X)^{-1} = \begin{pmatrix} 0,25 & 0 & 0 \\ 0 & 0,25 & 0 \\ 0 & 0 & 0,25 \end{pmatrix}$$

$$\text{Этап 4. } X^T \cdot Y = \begin{pmatrix} 24,5 \\ -13,5 \\ -9,5 \end{pmatrix}$$

$$\text{Этап 5. } \alpha = [X^T \cdot X]^{-1} \cdot X^T \cdot Y = \begin{pmatrix} 6,125 \\ -3,375 \\ -2,375 \end{pmatrix}$$

Получаем модель: $y = 26,125 - 3,375x_1 - 2,375x_2$.

2.2. Проверка коэффициентов на значимость

Значимость отдельных коэффициентов уравнения множественной регрессии оценивается с помощью Т-критерия Стьюдента по формуле (4).

$$t_{p_{\alpha_i}} = \frac{|\alpha_i|}{\sqrt{\sigma_{\alpha_i}^2}}, \quad (4)$$

где α_i - оценка i -го коэффициента модели;

$\sigma_{\alpha_i}^2$ - оценка дисперсии параметра α_i - i -й элемент главной диагонали дисперсионно-ковариационной матрицы S_{α_i} , формула (5).

$$S_{\alpha_i} = \sigma_u^2 \cdot [X^T \cdot X]^{-1} = \frac{RSS}{n-k} \cdot [X^T \cdot X]^{-1} = \frac{Y^T \cdot Y - \alpha^T \cdot X^T \cdot Y}{n-k} \cdot [X^T \cdot X]^{-1}, \quad (5)$$

где σ_u^2 - оценка дисперсии ошибки u ;

RSS - сумма квадратов ошибок.

Полученное по формуле (4) значение расчётного критерия Стьюдента $t_{p_{\alpha_i}}$ для каждого коэффициента $\alpha_0, \alpha_1, \dots, \alpha_i$ сравнивается с табличным (приложение Б) при уровне значимости $\alpha = 0,05$ или $0,01$. Если фактическое значение $t_{p_{\alpha_i}}$ превышает табличное, то коэффициент (параметр) модели значим ($\alpha_i \neq 0$) с вероятностью $(1-\alpha) \cdot 100 = 95\%$ или 99% . Следует принимать число степеней свободы $n-k$ (n - число наблюдений, k - число коэффициентов модели).

Порядок работы при проверке значимости коэффициента по t-статистике

1. Формулируем нулевую гипотезу H_0 о не значимости параметра $\alpha_i = 0$, и альтернативную H_1 – о его значимости.

2. Выбираем уровень значимости α (1% или 5%) – максимально допустимой вероятности ошибочного принятия гипотезы H_1 . Вычисляем число степеней свободы $(n-1)$.

3. По таблицам распределения Стьюдента определяем критическое значение критерия (приложение Б)

4. Если расчётный t-критерий по абсолютной величине больше табличного, то коэффициент является значимым при выбранном уровне значимости α . В противном случае коэффициент незначим (на данном уровне α)

Пример 2. Для уравнения регрессии, полученного в Примере 1., вычислить значение Т-критерия Стьюдента и определить статистическую значимость каждого коэффициента модели $\alpha_0, \alpha_1, \alpha_2$ при уровне значимости 5%.

1. Рассчитаем сумму квадратов ошибок:

$$RSS = Y^T \cdot Y - \alpha^T \cdot X^T \cdot Y = 218,25 - 218,1875 = 0,0625.$$

Тогда σ_u^2 - оценка дисперсии ошибки u : $\sigma_u^2 = 0,0625 / (4-3) = 0,0625$.

2. Рассчитаем дисперсионно-ковариационную матрицу

$$\frac{RSS}{n-k} \cdot [X^T \cdot X]^{-1} = (0,0625/1) * \begin{pmatrix} 0,25 & 0 & 0 \\ 0 & 0,25 & 0 \\ 0 & 0 & 0,25 \end{pmatrix} =$$

$$= \begin{pmatrix} 0,015625 & 0 & 0 \\ 0 & 0,015625 & 0 \\ 0 & 0 & 0,015625 \end{pmatrix}.$$

Тогда соответствующие дисперсии коэффициентов

$$\sigma_{\alpha_0}^2 = \sigma_{\alpha_1}^2 = \sigma_{\alpha_2}^2 = 0,015625.$$

Табличное значение (двусторонний) $t_{\text{табл}}(0.05; (4-3)) = 12.706$.

А расчётные значения t-критериев.

$$t_{p_{\alpha_0}} \frac{|6,125|}{\sqrt{0,015625}} = 49.$$

$$t_{p_{\alpha_1}} \frac{|-3,375|}{\sqrt{0,015625}} = 27.$$

$$t_{p_{\alpha_2}} \frac{|-2,375|}{\sqrt{0,015625}} = 19.$$

Каждый из коэффициентов значим $\alpha_i \neq 0$, т.к. больше табличного значения 12,7 при уровне значимости 5%.

2.3. Проверка адекватности уравнения множественной регрессии в целом

Значимость уравнения множественной регрессии в целом ($\alpha_1 \neq \alpha_2 \neq 0$) оценивается с помощью F -критерия Фишера:

$$F_p = \frac{R^2}{1-R^2} \cdot \frac{n-k}{k-1}, \quad (6)$$

где R^2 - коэффициент детерминации - индикатор степени подгонки модели к фактическим данным - или часть вариации (дисперсии) зависимой переменной y , которая объясняется уравнением регрессии. Чем меньше разброс значений остатков около линии регрессии по отношению к общему разбросу значений, тем лучше прогноз. Если имеется R-квадрат равный 0.4, то 40% от исходной изменчивости или дисперсии могут быть объяснены, а 60% остаточной изменчивости остаются необъясненными моделью. R-квадрат определяется при помощи формулы (7)

$$R^2 = \frac{\mathbf{C}^T \cdot \mathbf{X}^T \cdot \mathbf{Y} - n \cdot \bar{Y}^2}{\mathbf{Y}^T \cdot \mathbf{Y} - n \cdot \bar{Y}^2}, \quad (7)$$

n - число наблюдений;

k - число коэффициентов факторов;

$\bar{Y}$ - среднее значение зависимой переменной.

Полученное по формуле (6) значение F сравнивается с табличным (приложение Б) при уровне значимости $\alpha = 0,05$ или $0,01$. Если фактическое значение F -критерия Фишера превышает табличное $F_p > F_{табл}$, то уравнение статистически значимо, т.е. модель адекватна ($\alpha_1 \neq \alpha_2 \neq 0$) с вероятностью $(1 - \alpha) * 100 = 95\%$ или 99% . Следует принимать число степеней свободы $k-1$ (по горизонтали), $n-k$ (по вертикали).

Порядок работы при проверке значимости уравнения по F-критерию

1. Формулируем нулевую гипотезу H_0 о неадекватности модели в целом ($\alpha_0 = \alpha_1 = \alpha_2 = 0$) и альтернативную H_1 – о его адекватности.

2. Выбираем уровень значимости α (1% или 5%). Вычисляем число степеней свободы: $k-1$, $n-k$.

3. По таблицам F-распределения Фишера определяем критическое значение критерия (приложение Б).

4. Если F-расчётное больше F-табличного, то модель является в целом адекватной при выбранном уровне значимости α . В противном случае – не адекватной.

Пример 3. Для уравнения, полученного в Примере 1., вычислить значение F -критерия Фишера и определить статистическую значимость уравнения (адекватность модели) при уровне значимости 5%.

$$1. \mathcal{E}^T \cdot X^T \cdot Y - n \cdot \bar{Y}^2 = 218,1875 - 150,0625 = 68,125.$$

$$2. Y^T \cdot Y - n \cdot \bar{Y}^2 = 218,25 - 150,0625 = 68,1875.$$

$$3. R^2 = 68,125 / 68,1875 = 0,999.$$

$$4. F_p = 545.$$

Сравним с $F_{табл}(0,05; (3-1); (4-3)) = 199.5$ для уровня значимости 5% и числа степеней свободы 2 и 1 (приложение Б).

Вывод: модель в целом адекватна, т.к. $F_p > F_{табл}$.

Предпосылки метода наименьших квадратов

Регрессионный анализ функции (1), основанный на обычном или одношаговом методе наименьших квадратов (МНК), должен удовлетворять четырем условиям Гаусса—Маркова:

1. Математическое ожидание случайной составляющей (ошибки), $M(u_i)$ в любом наблюдении должно быть равно нулю.

2. Дисперсия случайной составляющей σ_u^2 должна быть постоянна для всех наблюдений.

3. Отсутствие систематической связи между значениями случайной составляющей u_i в любых двух наблюдениях.

4. Случайная составляющая должна быть распределена независимо от переменных x_i , которые не являются стохастическими, т.е. не имеют случайной составляющей.

После того как проведена оценка параметров модели, рассчитывая разности фактических и теоретических значений $y - \tilde{y}_x$ можно получить оценки случайной составляющей (остатков или ошибок) u . В задачу регрессионного анализа входит не только построение самой модели, но и исследование остаточных величин.

Необходимость этого объясняется тем, что при использовании МНК предполагалось, что остатки представляют собой независимые случайные величины (условие №3, 4) и их среднее значение равно 0 (условие №1); они имеют одинаковую (постоянную) дисперсию (условие №2).

Таким образом, исследование остатков предполагают проверку наличия перечисленных предпосылок МНК:

Случайных характер остатков

Для проверки строится график зависимости остатков u_i от теоретических значений результативного признака $\tilde{y}_x$. Если на графике получена горизонтальная полоса, то остатки представляют собой случайные величины и МНК оправдан, а теоретические значения хорошо аппроксимируют фактические значения y . Пример случайности остатков приведен на рисунке 2.

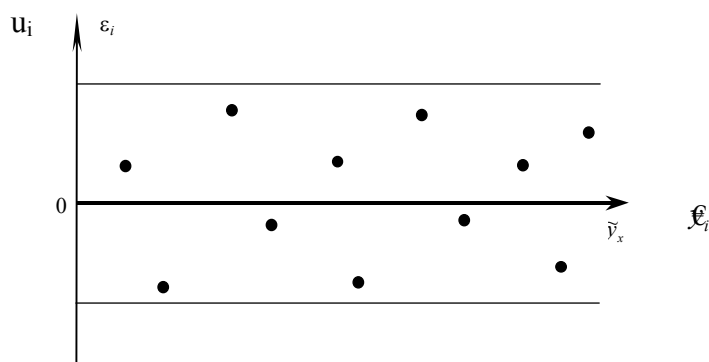


Рисунок 2 - Зависимость u_i от y

Возможны различные случаи зависимости остатков от теоретических значений $\tilde{y}_x$, рисунок 3.

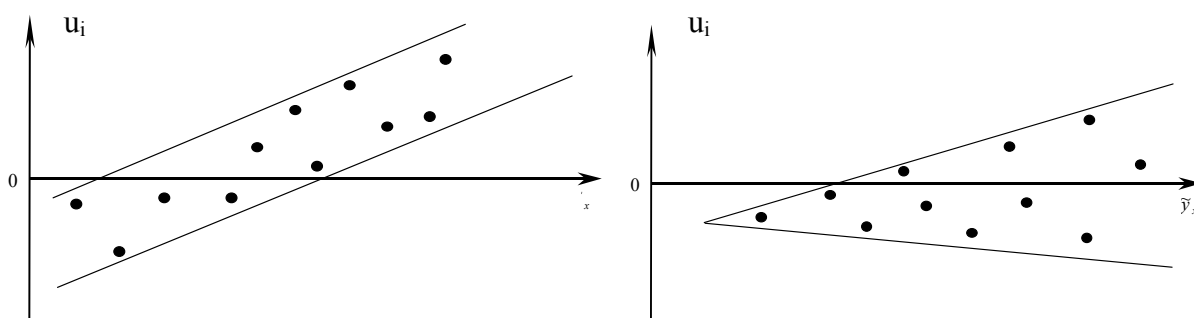


Рисунок 3 - Зависимость u_i от $\tilde{y}_x$

Нулевая средняя величина остатков, не зависящая от x_i (№1, №4)

Эта предпосылка означает, что математическое ожидание случайной составляющей (ошибки), $M(u_i)$ в любом наблюдении должно быть равно нулю, т.е. $\sum_{i=1}^n (\tilde{y}_x - y_i) = 0$. Это условие выполнимо для линейных моделей. Для определения независимости величины остатков от x_i , как и в случае определения независимости u_i от y , строится график u_i от x_i . Если остатки на графике расположены в виде горизонтальной полосы, то они независимы от значений x_i . Если же зависимость присутствует, то модель является неадекватной.

Гомоскедастичность (№2)

Гомоскедастичность остатков означает, что дисперсия каждого отклонения одинакова для всех значений x . Если это условие не соблюдается, то имеет место гетероскедастичность. Наличие гетероскедастичности можно наглядно видеть из поля корреляции, рисунок 4.

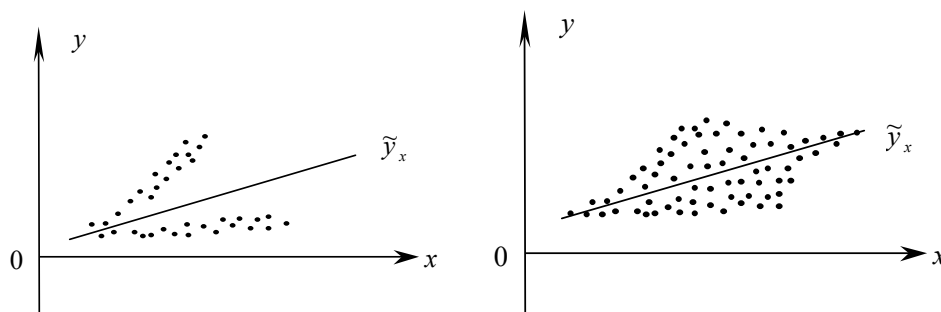


Рисунок 4 - Поле корреляции

Т.к. дисперсия характеризует отклонение, то из рисунков видно, что в первом случае дисперсия остатков растет по мере увеличения x , а во втором – дисперсия остатков достигает максимальной величины при средних значениях величины x и уменьшается при минимальных и максимальных значениях x . Наличие гетероскедастичности будет сказываться на уменьшении эффективности оценок параметров уравнения регрессии. Наличие гомоскедастичности или гетероскедастичности можно определять также по графику зависимости остатков от теоретических значений y .

Отсутствие автокорреляции остатков (№3)

Под автокорреляцией остатков понимают зависимость распределения значений остатков u_i друг от друга. Автокорреляция остатков означает наличие корреляции между остатками текущих и предыдущих

(последующих) наблюдений. Оценить эту зависимость можно вычислив коэффициент корреляции между этими остатками по формуле (8).

$$r_{u_i u_j} = \frac{\overline{u_i u_j} - \bar{u}_i \cdot \bar{u}_j}{\sigma_{u_i} \sigma_{u_j}}. \quad (8)$$

Если этот коэффициент окажется существенно отличным от нуля, то остатки автокоррелированы.

Пример 4.

По четырем предприятиям региона (таблица 2) изучается зависимость выработки продукции на одного работника y (тыс. грн.) от ввода в действие новых основных фондов x_2 (% от стоимости фондов на конец года) и от удельного веса рабочих высокой квалификации в общей численности рабочих x_1 (%). Составлена модель множественной регрессии $y = 2,38 + 3,38x_1 + 2,15x_2$.

Проверить выполнение предпосылок 1МНК для полученной модели.

Таблица 2 – Анализ остатков (ошибок) уравнения регрессии

Предприятие Показатель	№1	№2	№3	№4
x_1 , (%)	1	2	3	5
x_2 , (%)	0	1	3	4
y , (тыс. грн.)	6	11	19	28
$\tilde{y}_x$, (тыс. грн.)	5,79	11,31	19,07	27,87
u_i , (тыс. грн.)	0,21	-0,31	-0,07	0,13

Вычисляем теоретические значения по уравнению регрессии, а остатки по формуле $u = y - \tilde{y}_x$ и записываем в таблицу 2.

Теперь для проверки случайного характера остатков построим график их зависимости от теоретических значений $\tilde{y}_x$, рисунок 5.

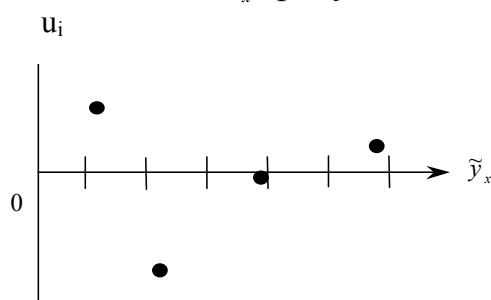


Рисунок 5 - Зависимость u_i от $\tilde{y}_x$

Хотя по четырем точкам судить трудно, но в целом можно сделать вывод, что остатки распределены случайно. Из этого же рисунка можно сделать вывод о гомоскедастичности остатков, т. к. дисперсия каждого отклонения одинакова для всех значений x .

Вычислим величину суммарного отклонения:

$$\sum_{k=1}^n (y - \tilde{y}_x) = 0,21 - 0,31 - 0,07 + 0,13 = -0,04.$$

По малости этой величины можно сделать вывод о практически нулевой средней величине остатков.

Коэффициент автокорреляции остатков находим по следующим рядам данных:

u_i , (тыс. грн.)	-0,31	-0,07	0,13
u_{i-1} , (тыс. грн.)	0,21	-0,31	-0,07

$$\bar{u}_i = \frac{-0,31 - 0,07 + 0,13}{3} = -0,083$$

$$\bar{u}_{i-1} = \frac{0,21 - 0,31 - 0,07}{3} = -0,057$$

$$\overline{u_i u_j} = \frac{(-0,31) \cdot 0,21 + (-0,07) \cdot (-0,31) + 0,13 \cdot (-0,07)}{3} = -0,018$$

$$\begin{aligned} \sigma_{u_i} &= \sqrt{\frac{\sum_{k=1}^3 (u_{ik} - \bar{u}_i)^2}{3}} = \\ &= \sqrt{\frac{(-0,31 + 0,083)^2 + (-0,07 + 0,083)^2 + (0,13 + 0,083)^2}{3}} = 0,18 \end{aligned}$$

$$\begin{aligned} \sigma_{u_{i-1}} &= \sqrt{\frac{\sum_{k=1}^3 (u_{i-1k} - \bar{u}_{i-1})^2}{3}} = \\ &= \sqrt{\frac{(0,21 + 0,057)^2 + (-0,31 + 0,057)^2 + (-0,07 + 0,057)^2}{3}} = 0,212 \end{aligned}$$

Отсюда находим

$$r_{u_i u_j} = \frac{\overline{u_i u_j} - \bar{u}_i \cdot \bar{u}_j}{\sigma_{u_i} \sigma_{u_j}} = \frac{-0,018 - (-0,083) \cdot (-0,057)}{0,18 \cdot 0,212} = -0,348$$

Коэффициент корреляции незначителен и его можно считать приемлемым. Таким образом установлено, что были все предпосылки к тому, чтобы применять 1МНК и линейное уравнение регрессии к исходным данным.

3. ИНДИВИДУАЛЬНЫЕ РАСЧЁТНО-ПРАКТИЧЕСКИЕ ЗАДАНИЯ

Задание 1.

Даны матрицы $A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 4 & 1 \\ 1 & -4 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$, $C = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$ и число $\alpha = 2$. Найти $A^T B + \alpha C$, (приложение А).

Задание 2. Задана квадратная матрица 3-го порядка. Вычислите определитель $\det A$ и обратную матрицу A^{-1} .

$$A = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 3 & 1 \\ 3 & 2 & 0 \end{bmatrix}$$

Задание 3.

По исходным данным таблицы 3 в соответствии с вариантом выполнить:

1. Оценить параметры $\alpha_0, \alpha_1, \alpha_2$ модели $y = y + u = \alpha_0 + \alpha_1 \cdot x_1 + \alpha_2 x_2 + u$
2. Проверить коэффициенты $\alpha_0, \alpha_1, \alpha_2$ на значимость
3. Оценить адекватность (значимость) модели в целом.
4. Проверить были ли все предпосылки к тому, чтобы применять МНК и линейное уравнение регрессии к исходным данным.

Таблица 3 – Исходные данные

Номер предприятия	1	2	3	4
x_1	1	1	-1	-1
x_2	1	-1	1	-1
y	6	11	19	29
Вариант				
1	6	10	19	29
2	5	11	19	29
3	6	11	16	29
4	8	11	19	29
5	6	11	19	26
6	6	12	19	29
7	10	11	19	29
8	8	13	19	30
9	5	10	19	29
10	6	11	18	31
11	6	11	19	31
12	6	11	17	25
13	6	11	15	26
14	6	11	20	28
15	6	13	18	29
16	6	16	20	29
17	6	11	19	29
18	5	12	19	29
19	9	11	19	29
20	6	11	15	29
21	6	11	18	29
22	6	11	19	30
23	4	11	19	29
24	6	15	19	29
25	6	10	19	29
26	6	11	13	29

4. СОДЕРЖАНИЕ ОТЧЕТА О ПРАКТИЧЕСКОМ ЗАНЯТИИ

- 1) Название и цель работы.
- 2) Условия индивидуальных заданий.
- 3) Ход решения индивидуальных заданий.
- 3) Выводы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Себер Дж. Линейный регрессионный анализ/ Дж. Себер. - М.: Мир, 1980-200с.
2. Демиденко Е. З. Линейная и нелинейная регрессия/ Е. З.Демиденко.- М.: Финансы и статистика, 1981 -320с.
1. Maddala G. S. Introduction to Econometrics, 3rd Edition, Wiley, 2001. (1st ed. Macmillan, 1988)
2. Greene, W H, Econometric analysis Hardcover - 4 edition (July 28, 1999) Prentice Hall; Paperback 3 edition (April 2000) Prentice Hall
3. Judge, G G et al. The theory and practice of econometrics; 2nd ed Wiley, 1985

ПРИЛОЖЕНИЕ А (справочное)

Вспомогательные сведения из высшей математики

Понятие матрицы

Матрицей называется прямоугольная таблица из чисел, содержащая некоторое количество m строк и некоторое количество n столбцов. Числа m и n называются порядками матрицы. В случае, если $m = n$, матрица называется квадратной, а число $m = n$ - ее порядком.

Умножение матрицы на число

Произведением матрицы A_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$) на вещественное число k называется матрица C_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$), элементы которой равны $c_{ij} = k \cdot a_{ij}$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$). (1.3)

Пример 1. Дана матрица A , найти матрицу $5A$

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 2 & 0 \\ -1 & -3 \end{bmatrix}.$$

Сложение матриц

Суммой двух матриц, например: A и B , имеющих одинаковое количество строк и столбцов, m и n , называется матрица C_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$), элементы которой равны $c_{ij} = a_{ij} + b_{ij}$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$). $C = A + B$.

Пример 2. Найти алгебраическую сумму матриц:

$$\begin{pmatrix} 2 & -4 \\ 3 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ -2 & 5 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ -2 & 7 \end{pmatrix}.$$

Умножение матриц

Матрицы можно умножить, если число столбцов первой матрицы равно числу строк второй.

$C_{mk} = A_{mn} \cdot B_{nk}$; $c_{ij} = a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + \dots + a_{in} \cdot b_{nj}$; ($i = 1, 2, \dots, m$), ($j = 1, 2, \dots, k$).

Пример 3.

$$C_{23} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 7 \\ 11 & 10 & 17 \end{pmatrix}$$

$$c_{11} = 1 \cdot 1 + 2 \cdot 2 = 5; \quad c_{12} = 1 \cdot 2 + 2 \cdot 1 = 4; \quad c_{13} = 1 \cdot 3 + 2 \cdot 2 = 7;$$

$$c_{21} = 3 \cdot 1 + 4 \cdot 2 = 11; \quad c_{22} = 3 \cdot 2 + 4 \cdot 1 = 10; \quad c_{23} = 3 \cdot 3 + 4 \cdot 2 = 17.$$

Свойства произведения:

$$A \cdot B \cdot C = (A \cdot B) \cdot C = A \cdot (B \cdot C);$$

$$A \cdot B \cdot B \cdot A;$$

$$A(B + C) = AB + AC \text{ - умножение слева;}$$

$$(B + C)A = BA + CA \text{ - умножение справа;}$$

$$0A = A0 = 0; \quad EA = AE = A;$$

E -единичная матрица («1» по диагонали, «0» остальные элементы)

Вычисление определителя матрицы

1. Определителем матрицы 1×1 , состоящей из одного числа, будем считать само это число.

2. Определитель матрицы 2×2 вычисляется по формуле (A.1)

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (\text{A.1})$$

3. Определитель матрицы 3×3 вычисляется (например, разложением по первой строке)

по формуле (A.2)

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \quad (\text{A.2})$$

Можно использовать разложение по любой строке или столбцу.

Дополнительным минором M_{ij} элемента матрицы n -го порядка a_{ij} ($i, j = 1, \dots, n$) называется определитель матрицы $n-1$ -го порядка, полученной из матрицы A вычеркиванием i -ой строки и j -го столбца, на пересечении которых стоит данный элемент.

Алгебраическим дополнением A_{ij} элемента матрицы n -го порядка a_{ij} ($i, j = 1, \dots, n$) называется число $(-1)^{i+j} M_{ij}$, где M_{ij} — дополнительный минор.

Тогда определителем (или детерминантом) матрицы A называется число, которое вычисляется как $\det A = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n}$

Пример 4. Вычислить определитель матрицы A :

$\det A =$

$$\begin{vmatrix} 1 & 2 & 1 \\ 0 & -2 & 3 \\ 3 & 1 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} -2 & 3 \\ 1 & 1 \end{vmatrix} - 2 \cdot \begin{vmatrix} 0 & 3 \\ 3 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & -2 \\ 3 & 1 \end{vmatrix} = (-2 \cdot 1 - 1 \cdot 3) - 2(0 \cdot 1 - 3 \cdot 3) + (0 \cdot 1 + 3 \cdot 2) =$$

$$= -5 + 18 + 6 = 19.$$

Транспонирование.

Если в матрице A_{mn} поменять местами строки и столбцы, получим транспонированную матрицу A_{nm}^T .

Пример 5.

$$A_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 0 & 2 \end{pmatrix}, \quad A_{32}^T = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 3 & 2 \end{pmatrix}.$$

Свойства транспонирования: $(A^T)^T = A$; $(A + B)^T = A^T + B^T$; $(A \ B \ C)^T = C^T B^T A^T$.

Нахождение обратной матрицы

Квадратная матрица называется неособенной, если ее определитель не равен нулю. Всякая неособенная матрица имеет обратную. Это такая матрица A^{-1} , которая, будучи умножена на исходную A слева или справа, дает единичную

$$A^{-1}A = AA^{-1} = E.$$

Свойства обратной матрицы:

ПРИЛОЖЕНИЕ Б
(справочное)
Статистические таблицы

Таблица Б.1 - t-распределение: критические значения t

Число степеней свободы	Тесты	Уровень значимости					
		10%	5,0%	2%	1,0%	0,2%	0,10%
	Двусторонний	10%	5,0%	2%	1,0%	0,2%	0,10%
	Односторонний	5%	2,5%	1%	0,5%	0,1%	0,05%
1		6,314	12,706	31,821	63,657	318,31	636,62
2		2,920	4,303	6,965	9,925	22,327	31,598
3		2,353	3,182	4,541	5,841	10,214	12,924
4		2,132	2,776	3,747	4,604	7,173	8,610
5		2,015	2,571	3,365	4,032	5,893	6,869
6		1,943	2,447	3,143	3,707	5,208	5,959
7		1,895	2,365	2,998	3,499	4,785	5,408
8		1,860	2,306	2,896	3,355	4,501	5,041
9		1,833	2,262	2,821	3,250	4,297	4,781
10		1,812	2,228	2,764	3,169	4,144	4,587
11		1,796	2,201	2,718	3,106	4,025	4,437
12		1,782	2,179	2,681	3,055	3,930	4,318
13		1,771	2,160	2,650	3,012	3,852	4,221
14		1,761	2,145	2,624	2,977	3,787	4,140
15		1,753	2,131	2,602	2,947	3,733	4,073
16		1,746	2,120	2,583	2,921	3,686	4,015
17		1,740	2,110	2,567	2,898	3,646	3,965
18		1,734	2,101	2,552	2,878	3,610	3,922
19		1,729	2,093	2,539	2,861	3,579	3,883
20		1,725	2,086	2,528	2,845	3,552	3,850
21		1,721	2,080	2,518	2,831	3,527	3,819
22		1,717	2,074	2,508	2,819	3,505	3,792
23		1,714	2,069	2,500	2,807	3,485	3,767
24		1,711	2,064	2,492	2,797	3,467	3,745
25		1,708	2,060	2,485	2,787	3,450	3,725
26		1,706	2,056	2,479	2,779	3,435	3,707
27		1,703	2,052	2,473	2,771	3,421	3,690
28		1,701	2,048	2,467	2,763	3,408	3,674
29		1,699	2,045	2,462	2,756	3,396	3,659
30		1,697	2,042	2,457	2,750	3,385	3,646
40		1,684	2,021	2,423	2,704	3,307	3,551
60		1,671	2,000	2,390	2,660	3,232	3,460
120		1,658	1,980	2,358	2,617	3,160	3,373
Г		1,645	1,960	2,326	2,576	3,090	3,291

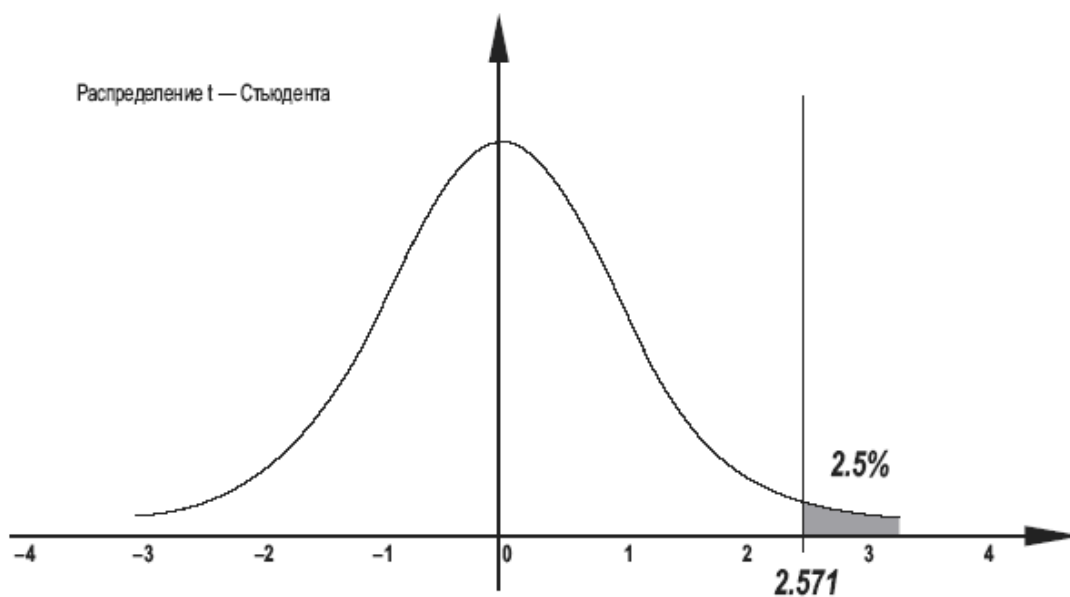


Рисунок Б.1 – График плотности распределения для $k=5$

Таблица Б.2 - F-распределение: Критические значения F с ν_1 и ν_2 степенями свободы. Уровень значимости 5%

$\nu_1 \backslash \nu_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	Γ
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
Γ	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00



Таблица Б.3 - F-распределение: Критические значения F с v_1 и v_2 степенями свободы. Уровень значимости 1%

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	Γ
1	4052	4999,5	5403	5625	5764	5859	5928	5981	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49	99,50
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,20	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
Γ	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00



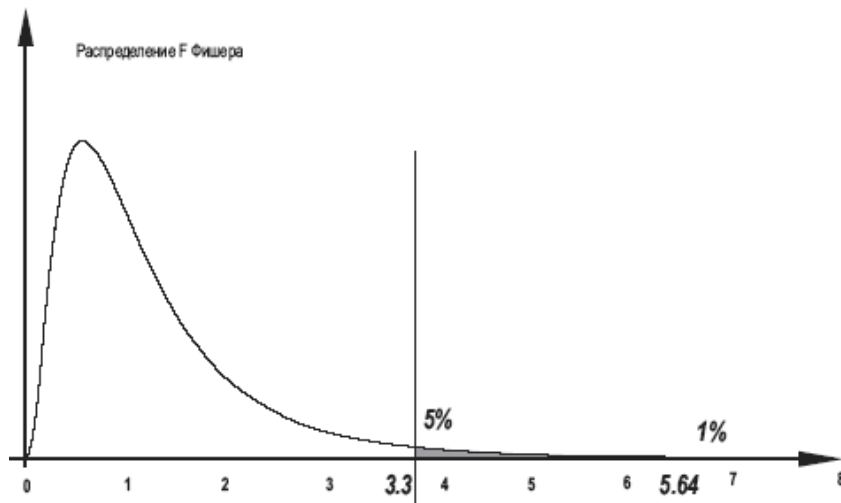


Рисунок Б.2 – График плотности распределения для $k_1=5$, $k_2=10$

Заказ № _____ от « _____ » _____ 2008. Тираж _____ экз.

Изд-во СевНТУ