

**В.Д. Тудрий, М.А. Верещагин**

**ПРАКТИЧЕСКОЕ РУКОВОДСТВО  
К ПРОИЗВОДСТВУ ПЕРВИЧНОЙ  
СТАТИСТИЧЕСКОЙ ОБРАБОТКИ И  
АНАЛИЗУ МЕТЕОРОЛОГИЧЕСКИХ РЯДОВ**

**Казань – 2009**

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

---

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

*Факультет географии и экологии*

В.Д. Тудрий, М.А. Верещагин

**ПРАКТИЧЕСКОЕ РУКОВОДСТВО  
К ПРОИЗВОДСТВУ ПЕРВИЧНОЙ  
СТАТИСТИЧЕСКОЙ ОБРАБОТКИ И  
АНАЛИЗУ МЕТЕОРОЛОГИЧЕСКИХ РЯДОВ**

*Учебно-методическое пособие*

Казань – 2009

Печатается по решению методической комиссии  
факультета географии и экологии  
Казанского государственного университета  
Утверждено на заседании кафедры метеорологии,  
климатологии и экологии атмосферы  
Протокол № 14 от «4» мая 2009года.

*Рецензент*  
**Зав. кафедрой физической и экономической географии,**  
**д.г.н., проф. В.А. Рубцов**

*Под редакцией*  
**профессора Ю.П. Переведенцева**

**«Практическое руководство к производству первичной статистической обработки и анализу метеорологических рядов».** Учебно-методическое пособие. Для студентов, обучающихся по специальности «Метеорология». В.Д. Тудрий, М.А. Верещагин. Под редакцией Ю.П. Переведенцева – Казань, КГУ, 2009г. – 34с.

Учебно-методическое пособие написано в соответствии с требованиями типовой учебной программы дисциплины «Методы статистической обработки гидрометеорологической информации». Предназначено для выполнения практических работ, следующих за теоретической частью дисциплины. Содержит формулировки заданий, подлежащих выполнению, и практические рекомендации по их решению на компьютере.

## **Содержание**

Предисловие.....	4
Тема 1. Оценка и сравнение статистических характеристик метеорологических рядов.....	5
Тема 2. Тренд-анализ метеорологических рядов.....	16
Тема 3. Корреляционный и регрессионный анализ.....	19
Тема 4. Проверка метеорологических рядов на их однородность.....	22
Тема 5. Анализ внутренней структуры метеорологических рядов.....	24
Список литературы.....	29
Приложение.....	30

## **Предисловие**

Обработка режимно-климатической информации является одной из важнейших составляющих профессиональной подготовки студентов-метеорологов. Её основу составляют методы математической статистики и теории вероятностей. Учебная литература в этой области знаний чрезвычайно разнообразна и далеко не исчерпывается предлагаемым читателю перечнем [1-4]. Для реализации указанных методов с применением ЭВМ к настоящему времени разработано большое количество пакетов стандартных программ, каждый из которых имеет свои специфические особенности.

Предлагаемое учебно-методическое пособие ставит своей целью ознакомить студентов, изучающих курс «Методы статистической обработки гидрометеорологической информации», с практическими приемами использования в учебных целях пакета программ STATGRAPHICS, с полным описанием которого читатель сможет познакомиться, обратившись к работе [2].

# Тема 1

## ОЦЕНКА И СРАВНЕНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК МЕТЕОРОЛОГИЧЕСКИХ РЯДОВ

Любая статистическая обработка метеорологических данных начинается с ввода их в компьютер и контроля подлежащего обработке материала для выявления в нем возможных грубых ошибок. Для этого необходимо вначале

- просмотреть таблицу введенных данных на дисплее,
- вывести на экран монитора графики введенных величин; если ошибка в исходных данных слишком грубая, то на графике будет виден выброс из общего хода изменений переменной.

Затем следует рассчитать основные статистические характеристики введенных рядов

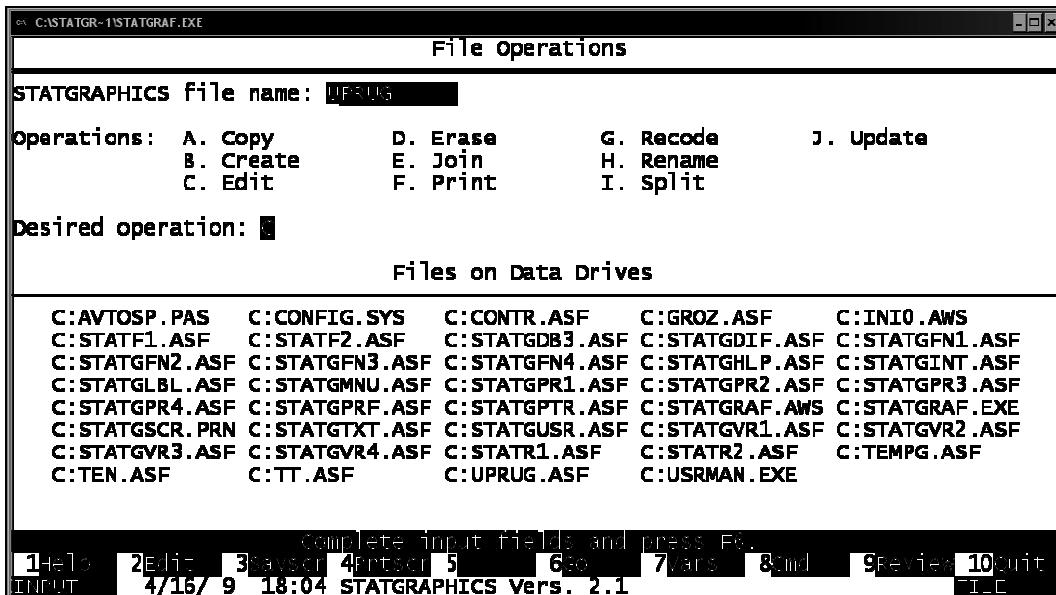
- среднее арифметическое  $\bar{x}$ ,
- дисперсию  $\sigma^2$ ,
- коэффициент вариации (для неотрицательных величин)  $c_v$ ,
- размах колебаний признака в ряду  $R$ ,
- коэффициент асимметрии  $A$ ,
- коэффициент эксцесса  $E$ ,
- моду  $M_0$ ,  $x_{\text{mod}}$ .

Для приобретения практических навыков по расчетам вышеуказанных характеристик предлагается выполнить задание 1.

**ЗАДАНИЕ 1.** Ввести с клавиатуры в оперативное запоминающее устройство два числовых ряда (по указанию преподавателя) и рассчитать с помощью стандартной процедуры пакета STATGRAPHICS основные статистические характеристики этих рядов.

### Рекомендации к выполнению задания.

1. Войти в процедуру "A" (Date Management);
2. Откроется окно с перечнем четырех позиций, войти в позицию 2 (File Operations);
3. Откроется следующее окно



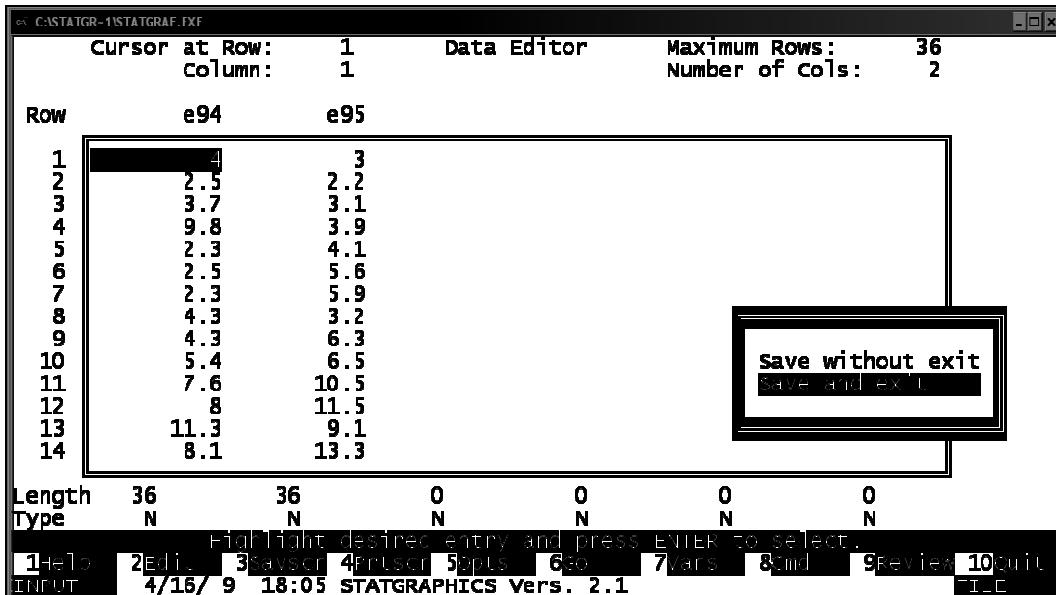
4. В первой строке ввести название файла, в котором будет храниться  $x_i$ ;

5. Выбрать в этом окне операцию "C" (Edit-Редактор), для этого надо вместо "J" ввести "C" в малом окне (Desired operation), нажать клавишу "F6";

6. В этом случае откроется электронная таблица, в которую надо последовательно ввести значения заданной случайной величины  $x_i$ ;

7. После ввода значений  $x_i$  ввести указанную в командной строке команду, нажатием клавиши "F6";

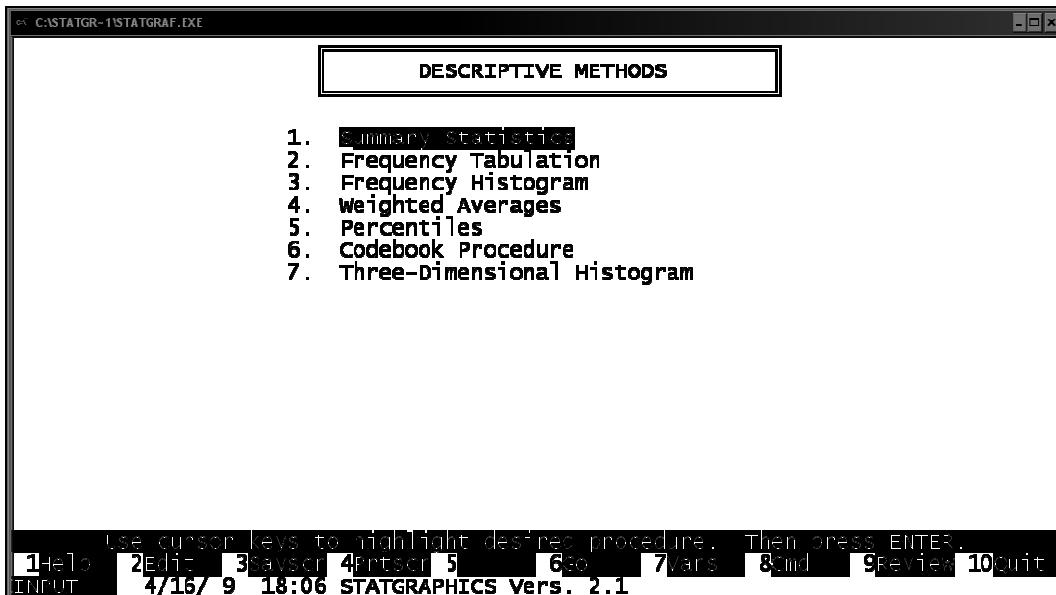
8. Справа на экране появится малое окно;



9. Курсором войти во вторую строку (Save and exit), запомнить значения и выйти из редактора, таблица закроется.

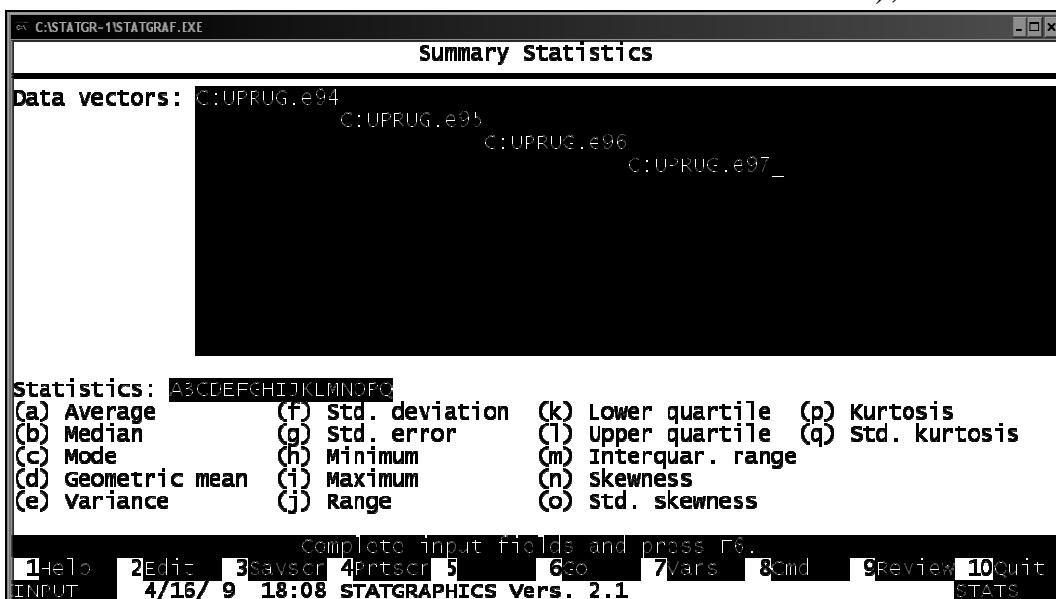
Если надо переименовать переменную, то следует ввести вместо "C" операцию "J" (Update).

10. После формирования переменной в формате Statgraphics приступить к расчетам характеристик. Для этого надо войти в процедуру "F" (Descriptive Methods); откроется список из 7 процедур.



11. Войти в позицию 1 (Summary Statistics);

12. На экране появится перечень основных характеристик (Statistics) и строки Data vectors, в одну из которых надо ввести имя переменной  $x_i$  (здесь можно ввести в 12 строк 12 переменных); (список имеющихся файлов переменных величин можно вызвать нажатием клавиши "F7");



13. Если не надо рассчитывать все характеристики, то в окне Statistics: A B C D ... P Q надо стереть буквы, соответствующие списку характеристик;

14. Нажать клавишу "F6", на экране появятся рассчитанные характеристики распределения случайной величины

Variable:	e94	e95	e96
Sample size	36	36	36
Average	8.675	9.85278	8.62778
Median	8.05	8.3	6.6
Mode	4.3	3.2	2.7
Geometric mean	7.04511	7.71127	6.44927
Variance	26.5236	41.054	35.8981
Standard deviation	5.15011	6.40734	5.9915
Standard error	0.858352	1.06789	0.998583
Minimum	2	1.6	1.1
Maximum	17	22.7	21.2
Range	15	21.1	20.1
Lower quartile	3.95	4	3.25
Upper quartile	13.75	15.2	14.9
Interquartile range	9.8	11.2	11.65
Skewness	0.298545	0.485232	0.49961
Standardized skewness	0.731282	1.18857	1.22379
Kurtosis	-1.39595	-1.16559	-1.15746
Standardized kurtosis	-1.70968	-1.42755	-1.41759

File Edit To quit, click on File or Exit Number:  
**1 Help    2 Edit    3 Layout    4 Options    5 Plotopt    6 Go    7 Stats    8 End    9 Review    10 quit**  
 INPUT    4/16/ 9 18:09 STATGRAPHICS Vers. 2.1

Здесь: в первой строке таблицы – Sample Size – объем выборки  $N$ , во второй Average – среднее значение  $\bar{x}$ , далее медиана –  $M_e$ , мода –  $M_0$ , геометрическое среднее, дисперсия –  $\sigma^2$ , среднеквадратическое отклонение –  $\sigma$ ,  $x_{min}$ ,  $x_{max}$ , размах –  $R$ , нижний квартиль  $x_{0,25}$ , верхний квартиль –  $x_{0,75}$ , межквартильный размах, асимметрия  $A$ , стандартная асимметрия, эксцесс –  $E$ , стандартный эксцесс;

Для оценки величины вариаций разноразмерных величин (например, количества облаков и количества осадков) полезно рассчитать коэффициенты вариации этих величин.

Для оценки изменчивости некоторой переменной в одноразмерном временном ряду можно оценить её среднеквадратическое отклонение, измеряемое в тех же единицах, что и данная случайная величина. Так, если во временном ряду есть пиковые значения, превышающие  $\pm 3\sigma$  ( $\pm 2\sigma$ ), то их с доверительной вероятностью  $P \geq 99\%$  ( $P \geq 95\%$ ), можно считать существенно отличающимися от среднего значения.

Модальное значение случайной величины, или мода ( $x_{mod}$ ), указывает на наиболее часто встречающееся значение случайной величины, ей соответствует модальная вероятность.

Коэффициент асимметрии показывает насколько и в какую сторону смещается модальное значение относительно среднего значения (или относительно нормального распределения случайной величины). Так, если  $x_{mod} < \bar{x}$ , то  $A > 0$  и распределение правостороннее, т.е. имеется большая повторяемость малых значений случайной величины, если же  $x_{mod} > \bar{x}$ , то  $A < 0$  и отмечается большая повторяемость больших значений случайной величины (левостороннее распределение). Асимметрию принято считать

- малой, если  $|A| \leq 0,25$ ;
- умеренной если  $0,25 < |A| \leq 0,5$ ;

- большой, если  $0,5 < |A| \leq 1,5$ ;
- исключительно большой, если  $|A| > 1,5$ .

Коэффициент эксцесса, или эксцесс  $E$ , указывает на степень крутости распределения случайной величины, или на степень его сжатия. Чем больше эксцесс, тем уже распределение и тем меньше среднеквадратическое отклонение  $\sigma$ ; в этом случае значения случайной величины концентрируются в значительной степени вблизи среднего значения. Принято считать эксцесс

- малым, если  $|E| \leq 0,5$ ;
- умеренным, если  $0,5 < |E| \leq 1,0$ ;
- большим, если  $1,0 < |E| \leq 3,0$ ;
- исключительно большим, если  $|E| > 3,0$ .

Если распределение случайной величины круче нормального, то эксцесс положительный  $E > 0$ , если же пологе нормального, то отрицательный  $E < 0$ .

Если имеются два исследуемых ряда (и более) одной размерности (например, температуры воздуха, в  $^{\circ}\text{C}$ ), то сначала необходимо сравнить их средние и дисперсии, оценить достоверность различий этих статистических характеристик. Только после этого можно делать достоверный вывод об отсутствии или наличии существенных отличий между рядами. При этом, обычно, в качестве порогового принимают вероятность ошибки  $a = 0,05$  (доверительная вероятность  $P = 0,95$ ). Для реализации этих рекомендаций необходимо выполнить задание 2.

**ЗАДАНИЕ 2. Оценить достоверность различий средних значений  $\bar{x}_1$  и  $\bar{x}_2$ , и дисперсий  $\sigma_1^2$  и  $\sigma_2^2$  двух выборок.**

#### **Рекомендации к выполнению задания.**

1. Войти в процедуру "G" (Estimation and Testing).
2. Подвести курсор ко второй строке (Two Sample Analysis), нажать клавишу "Enter". На экране высветятся 2 строки для ввода переменной выборки 1 (Sample 1) и переменной выборки 2 (Sample 2).
3. Ввести имя переменных.
4. Нажать клавишу "F6" – высветится экран.

Two-Sample Analysis Results			
Sample Statistics:	C:\UPRUG.e94	C:\UPRUG.e95	Pooled
Number of Obs.	36	36	72
Average	8.675	9.85278	9.26389
Variance	26.5236	41.054	33.7888
Std. Deviation	5.15011	6.40734	5.81281
Median	8.05	8.3	8.1
Conf. Interval For Diff. in Means:	95 Percent		
(Equal Vars.) Sample 1 - Sample 2	-3.91096	1.55541	70 D.F.
(Unequal Vars.) Sample 1 - Sample 2	-3.91318	1.55763	66.9 D.F.
Conf. Interval for Ratio of variances:	95 Percent		
sample 1 ÷ sample 2	0.329441	1.26701	35 35 D.F.
Hypothesis Test for H0: Diff = 0	Computed t statistic = -0.859633		
vs Alt: ne 0	Sig. Level = 0.392926		
at Alpha = 0.05	so do not reject H0.		
Complete input fields and press F6.			
1Hello	2Edit	3Savescr	4Printscr
INPUT	5/ 5/ 9	18:17	STATGRAPHICS Vers. 2.1
10quit	9Review	7Vars	6Go
TWOSAM			

Здесь приводятся значения среднего, дисперсий, среднеквадратического отклонения, медианы для каждой из выборок.

Далее приводятся доверительные интервалы для разности средних значений двух выборок с вероятностью, заданной в малом окне – 95% по умолчанию; при этом рассматриваются 2 варианта – равная дисперсия (Equal Vars.) и неравная дисперсия (Unequal Vars.) двух выборок.

Если доверительный интервал разности средних  $\bar{x}_1 - \bar{x}_2$  включает нулевое значение, то в этом случае с заданной вероятностью ( $\geq 95\%$ ) можно считать, что средние значения  $\bar{x}_1$  и  $\bar{x}_2$  значимо не отличаются; если интервал не включает нулевое значение, то  $\bar{x}_1 \neq \bar{x}_2$ , с вероятностью  $P \geq 95\%$ .

5. Для сравнения дисперсий двух выборок надо напротив строки Conf Interval for Ratio of Variances (Доверительный Интервал для Отношения дисперсий) в малом окне задать значение  $P = 95\%$  доверительной вероятности.

6. Нажать клавишу "F6" – на экране дополнительно выскакивает доверительный интервал для отношения дисперсий  $\sigma_1^2 / \sigma_2^2$ . Если этот доверительный интервал включает значение "1", то в этом случае с вероятностью  $P \geq 95\%$  дисперсии двух выборок одинаковы,  $\sigma_1^2 = \sigma_2^2$ . Если этот интервал не включает "1", то с вероятностью  $P \geq 95\%$   $\sigma_1^2 \neq \sigma_2^2$ .

Для того чтобы установить, как часто встречались те или иные значения случайной величины, рассчитываются абсолютные и относительные частоты (абсолютная и относительная повторяемость случайной величины). Формулы и пример расчетов указанных частот приводятся в [3] (с. 20-21, 135-137). Относительная частота  $P_x$  обнаружения того, что случайная величина  $X$

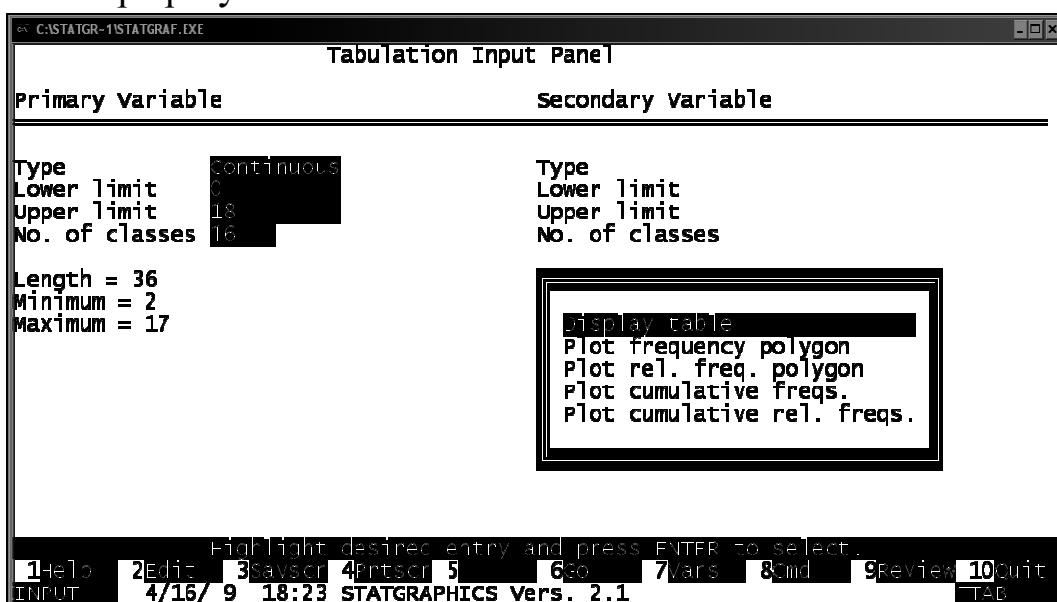
принимает значение  $x$ , при большом объеме выборки  $N$  и малой ширине интервала  $\Delta x$  стремится к плотности вероятности  $f(x)$ .

Распределение случайной величины по частотам представляется графически в виде полигона относительных частот (или гистограммы), который указывает на возможный тип распределения случайной величины (нормальный, экспоненциальный, логарифмический нормальный и пр.).

**ЗАДАНИЕ 3.** Рассчитать абсолютную, относительную, накопленную частоты, построить полигон частот, гистограмму.

**Рекомендации к выполнению задания.**

1. Войти в процедуру "F".
2. Выбрать процедуру 2 (Frequency Tabulation) и нажать клавишу "F6"; на экране появится строка ввода данных;
3. В строке "Data" ввести имя переменной  $x_i$ .
4. Нажать клавишу "F6"; на экране появится запрос ввода параметров, который заполняется программой по умолчанию; здесь компьютер автоматически выбирает число градаций;
5. Нажать клавишу "F6" и в дополнительном окне на экране появится меню форм выбора результатов.



6. Клавишей перемещения курсора выбрать первую строку (Display table – Таблица), нажать клавишу "Enter".

На экране появится таблица результатов.

Frequency Tabulation							
Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below	.00			0	.0000	0	.0000
1	.00	1.13	.563	0	.0000	0	.0000
2	1.13	2.25	1.688	1	.0278	1	.0278
3	2.25	3.38	2.813	5	.1389	6	.1667
4	3.38	4.50	3.938	7	.1944	13	.3611
5	4.50	5.63	5.063	2	.0556	15	.4167
6	5.63	6.75	6.188	0	.0000	15	.4167
7	6.75	7.88	7.313	2	.0556	17	.4722
8	7.88	9.00	8.438	4	.1111	21	.5833
9	9.00	10.13	9.563	1	.0278	22	.6111
10	10.13	11.25	10.688	2	.0556	24	.6667
11	11.25	12.38	11.813	1	.0278	25	.6944
12	12.38	13.50	12.938	2	.0556	27	.7500
13	13.50	14.63	14.063	2	.0556	29	.8056

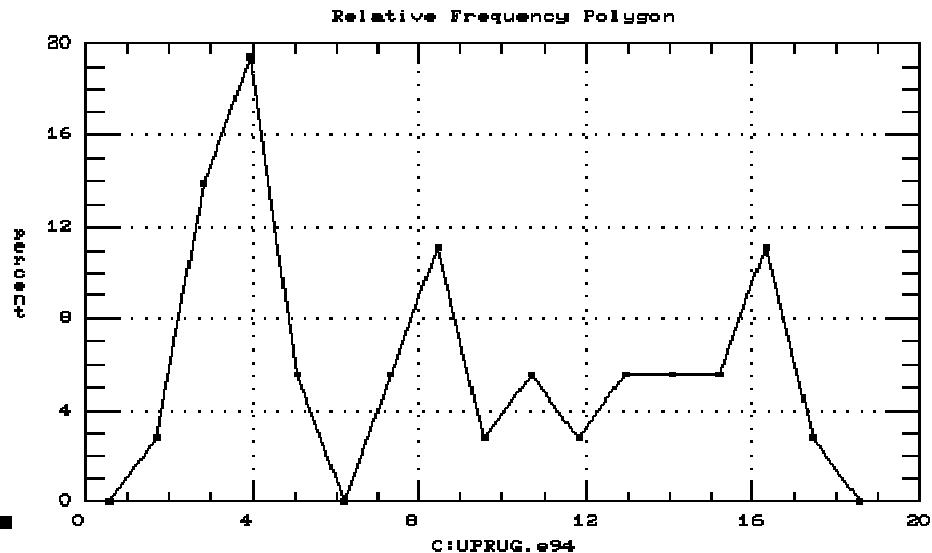
Mean = 8.675 Standard Deviation = 5.15011 Median = 8.05

Use Esc to quit, cursor keys or Page Number:  
**1Hello 2Edit 3Savecr 4Printcr 5Printpt 6Go 7Vars 8End 9Review 10Quit**  
 INPUT **4/16/ 9 18:24 STATGRAPHICS Vers. 2.1**

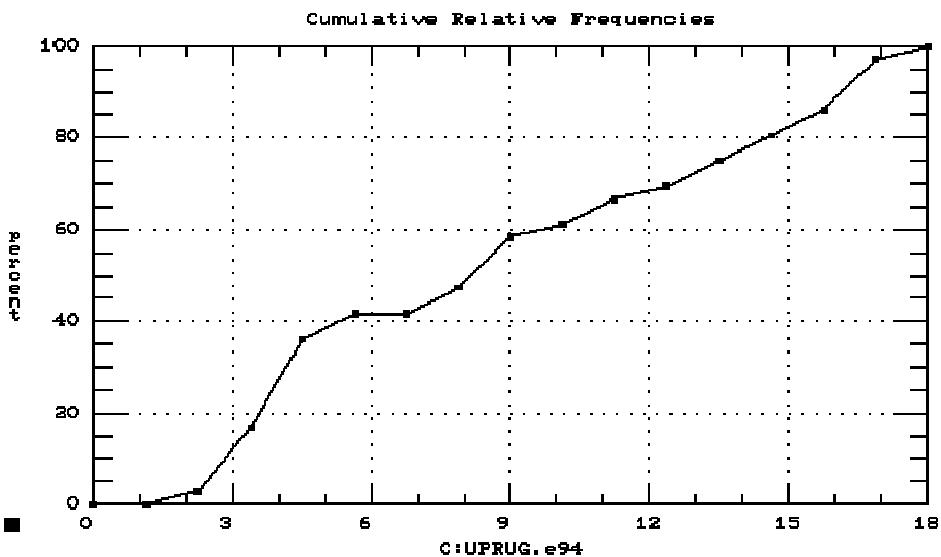
Здесь в столбце Class приводятся номера градаций (интервалов), во втором, третьем и четвертом столбцах приводятся нижняя и верхняя границы класса, градации (интервала), значений случайной величины, среднее значение градации, в пятом и шестом столбцах – абсолютная частота  $a_j$  и относительная частота  $p_j$  в каждой из градаций; в седьмом и восьмом столбцах – накопленная частота и накопленная относительная частота (функция распределения).

7. Нажать клавишу "Esc" (возврат к предыдущему окну), появится дополнительное окно, в котором выбрать Plot rel. freq. polygon.

8. Нажать клавишу "F6", на экране появится график – Полигон относительных частот.



9. Если в дополнительном окне выбрать Plot cumulative relat freqs. и нажать клавишу "F6", то на экране появится график функции распределения  $F(x)$ .



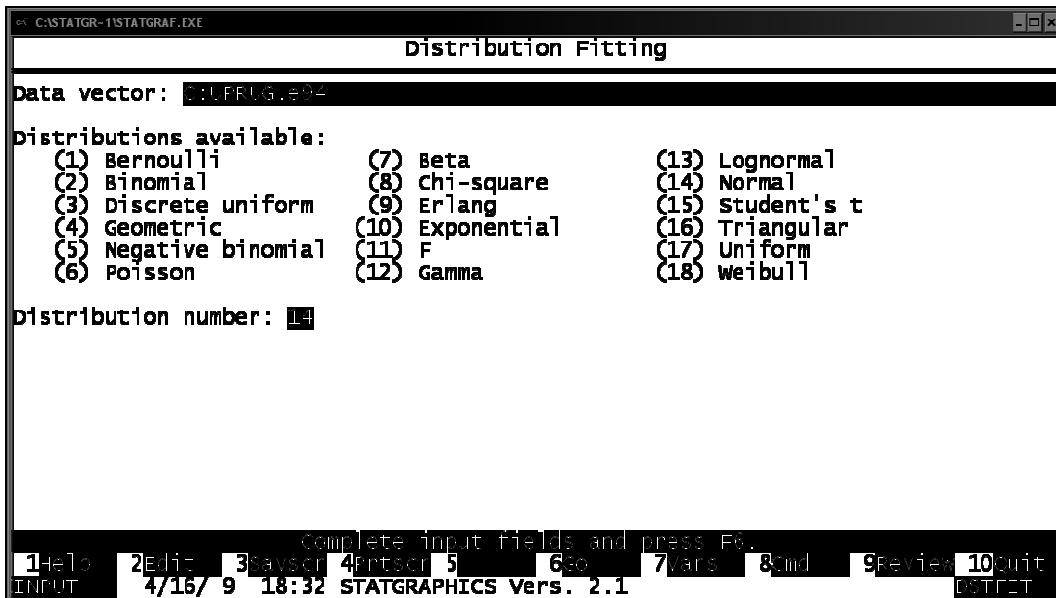
10. Для построения гистограммы распределения случайной величины надо в процедуре "F" главного меню войти в процедуру 3 (Frequency Histogram).

Нередко требуется подобрать к экспериментальному распределению случайной величины соответствующее теоретическое распределение, оценить возможность его аппроксимации с помощью критерия согласия. Для отработки практических навыков в этом направлении необходимо выполнить задание 4.

**ЗАДАНИЕ 4.** Подобрать теоретический закон распределения для описания экспериментального распределения, оценить вероятность аппроксимации с помощью критериев согласия.

#### Рекомендации к выполнению задания.

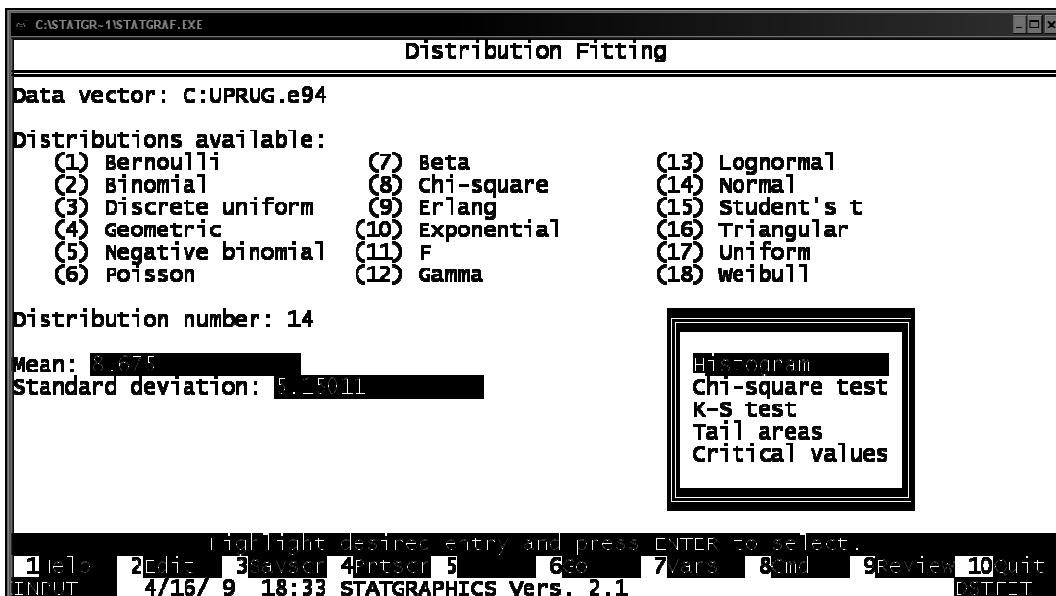
1. Войти в процедуру "H" (Distributions Functions).
2. Откроется подменю из 5 процедур, курсором выбрать 1 позицию (Distributions Fitting), нажать клавишу "Enter".
3. Откроется экран с перечислением 18 законов распределения.
4. В строке Data vector ввести имя случайной переменной  $x_i$ .



5. В малом окне задать номер распределения из списка или по умолчанию оставить номер 14 – Нормальный закон;

6. Нажать клавишу "F6", на экране появится среднее значение (Mean) и среднеквадратическое отклонение (Standard deviation).

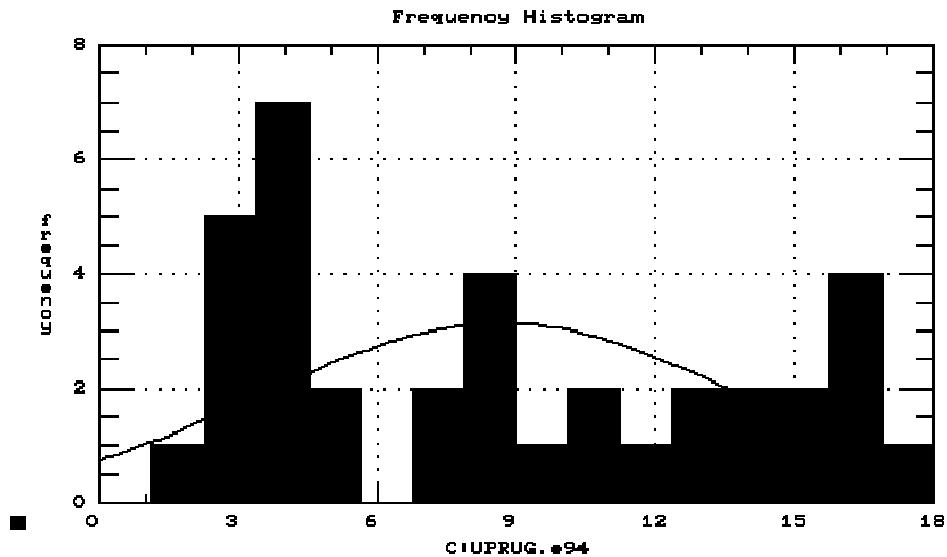
7. Нажать клавишу "F6" – появится малое окно в нижнем правом углу экрана.



8. Курсором выбрать первую строку – Histogram, нажать клавишу "Enter", на экране появится таблица, в которой приводится число интервалов  $x_{max}$ ,  $x_{min}$  и т.д. По умолчанию можно оставить все параметры, выбранные компьютером, без изменения, но можно переписать название графика (Top Title), изменить с помощью пробела в малом окне частоту, например, вместо абсолютной частоты взять относительную (Relative).

9. Нажать клавишу "F6" – на экране высветится график теоретического распределения и гистограмма экспериментального распределения. По виду

графика, по внешней схожести распределений можно принять решение о продолжении анализа.



10. Возвращаемся с помощью клавиши "Esc" в малое окно и выбираем вторую строку – Chi-square test – критерий согласия Хи-квадрат.

11. Нажать клавишу "F6" 2 раза – появится таблица, в которой даются результаты подсчета частот по градациям, а также  $\chi^2$  – Хи-квадрат критерий; под таблицей приводится суммарное значение  $\chi^2$ -критерия и уровень вероятности  $P$  (Sig. level). Чем меньше  $\chi^2$ , тем больше сходство экспериментального и теоретического распределений и тем больше вероятность  $P$ .

Chi-square Test				
Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chisquare
at or below	3.38	6	5	.05305
3.38	6.75	9	7	.39983
6.75	9.00	6	6	.00374
9.00	11.25	3	6	1.49007
above	11.25	12	11	.07172

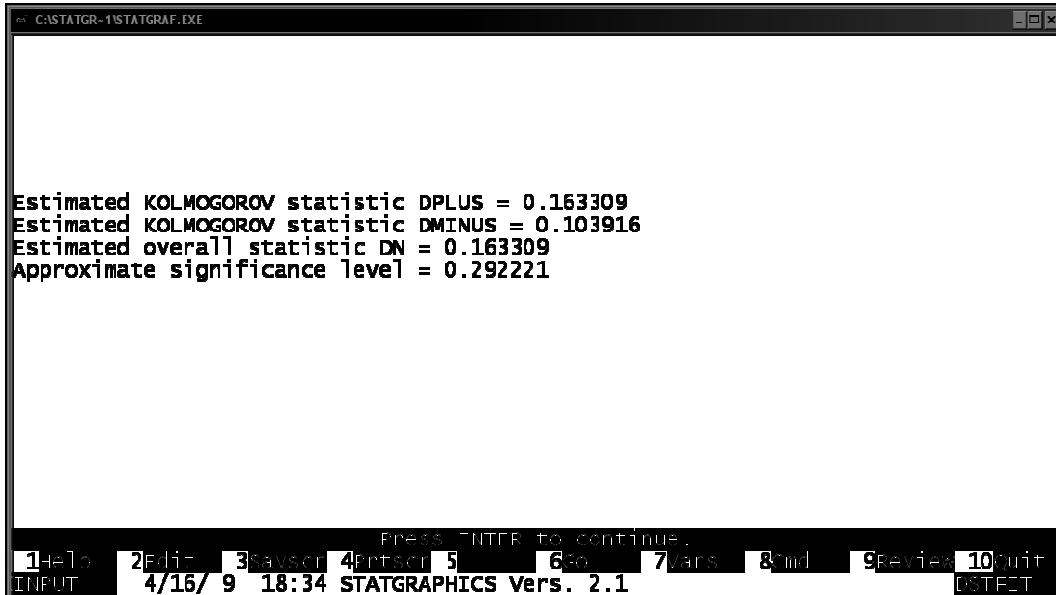
Chisquare = 2.01841 with 2 d.f. sig. level = 0.364508

Use Esc to quit, cursor keys or Page Number:  
 INPUT 1 2 3 4 5 6 7 8 9 10 Page 1.1 of 1.1  
 4/16/ 9 18:34 STATGRAPHICS Vers. 2.1

12. Проверить сходство распределений, применив еще один критерий согласия К-С-Колмогорова-Смирнова. Для этого 2 раза нажать клавишу "Esc" и войти в 3 строку (K-S test).

13. Нажать клавишу "Enter" (или клавишу "F6") и в последней строке распечатки на экране выделяем вероятность доверия  $P$  аппроксимации

выбранным теоретическим законом конкретного экспериментального – Approximate significance level, чем больше  $P$ , тем надежнее аппроксимация.



Если оба критерия согласия указали на высокую вероятность доверия, то данное теоретическое распределение достаточно удовлетворительно описывает экспериментальное распределение.

14. Если надежность аппроксимации небольшая, то надо вернуться в начало меню процедуры " $H$ " (нажать 3 раза клавишу "Esc") и задать в малом окне номер другого теоретического распределения, которое по внешнему виду ближе к экспериментальному. Далее повторить операции 6-13. Наиболее оптимальным считается то теоретического распределение, у которого большая надежность аппроксимации данного экспериментального распределения.

## Тема 2

### ТРЕНД-АНАЛИЗ МЕТЕОРОЛОГИЧЕСКИХ РЯДОВ

При изучении внутренней динамики временных метеорологических рядов часто требуется выявить наличие тенденции в изменениях временного ряда либо в виде линейного, либо параболического тренда.

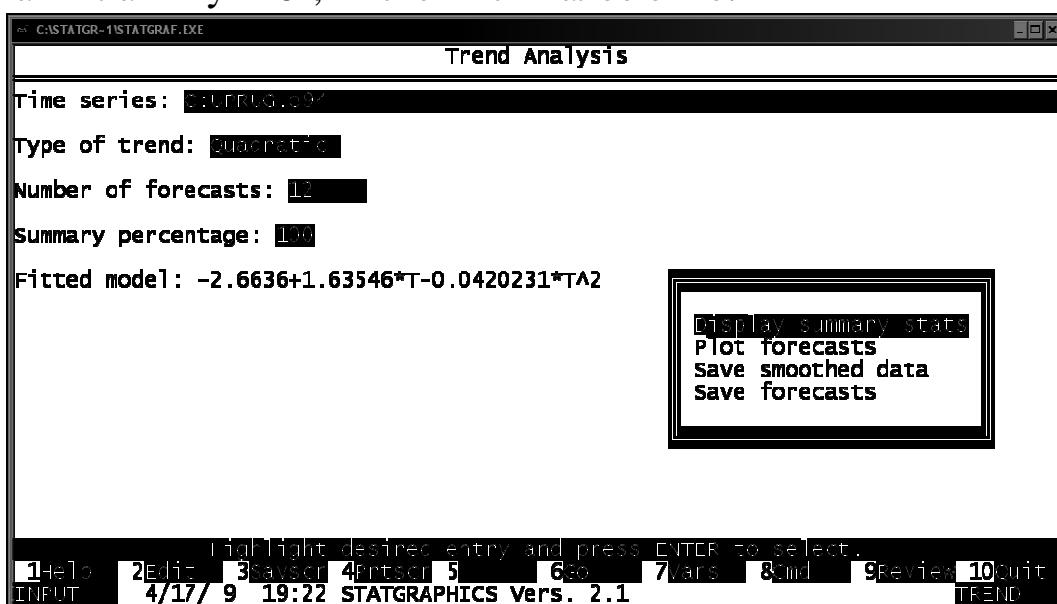
По виду тренда с определенной степенью надежности, можно осуществлять прогнозирование дальнейшего поведения временного ряда.

Для приобретения практических навыков по расчетам параметров тренда необходимо выполнить задание 5.

**ЗАДАНИЕ 5.** Найти параметры уравнения тренда, подобрать оптимальный тип тренда.

**Рекомендации к выполнению задания.**

1. Войти в процедуру "L" (Forecasting – Прогнозирование).
2. Выбрать позицию 4 (Trend Analysis). Нажать клавишу "F6".
3. В строке Time series ввести имя переменной.
4. В строке Type of trend с помощью пробела выбрать тип тренда (линейный, квадратичный, экспоненциальный), выберем Quadratic (квадратичный). В третьей строке можно задать необходимое упреждение для прогноза, по умолчанию можно оставить 12.
5. Нажать клавишу "F6", выскажется малое окно:

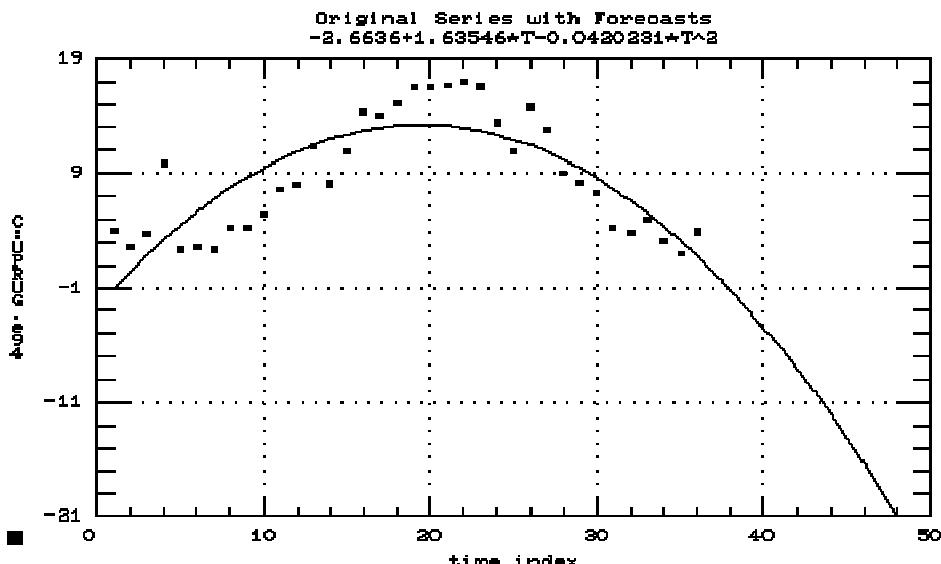


6. Выбрав первую позицию, нажав клавишу "Enter", получим искомое уравнение линейного тренда ( $y = a + bt$ ), значения ошибок – среднюю ошибку  $M.E.$ , среднеквадратическую ошибку –  $M.S.E.$  и т.д., а также прогностические значения случайной величины согласно выбранному уравнению тренда с различными упреждениями (интервалами прогноза)  $\Delta t$ .

		Percent: 100				
Forecast summary		M.E.	M.S.E.	M.A.E.	M.A.P.E.	M.P.E.
7.18381+0.0806049*T	0.00000	25.0857	4.42041	84.0685	-53.9709	
-2.6636+1.63546*T-0.0420231*T^2	0.00000	8.67092	2.58856	46.7796	-15.3321	
EXP(1.73959+0.0114996*T)	1.57951	27.9523	4.48607	69.6643	-26.0162	
EXP(2.10597-1.3249/T)	1.47926	24.8661	4.20083	65.8437	-23.8629	

Use ESC to exit, cursor keys or Page Number:  
**1Hello 2Edit 3Savecr 4Printcr 5Printpt 6Go 7Vars 8End 9Review 10Quit**  
 INPUT 4/16/ 9 18:42 STATGRAPHICS Vers. 2.1 TREN

7. Выбрав в малом окне вторую позицию, на экране получим график тренда, а также экспериментальных и прогностических значений случайной величины:



8. Выбирая, по мере надобности, несколько типов тренда, можно найти уравнение, более эффективно описывающее тенденцию изменения случайной величины во времени. Для этого сравниваются ошибки для 2-3 типов уравнений тренда (*M.E.*, *M.S.E.*, *M.A.E.* и т.д.). Выбирается тот тип, который дает меньшие ошибки.

Кроме того, можно оценить вклад каждого уравнения тренда в полную дисперсию экспериментальных данных. Для этого, вычисляется коэффициент детерминации  $R^2$  в процентах, который указывает, на сколько полно данное уравнение тренда описывает многообразие поведения случайной величины.

Вычисления проводятся с использованием процедуры "K" (Regression Analysis).

При вычислении  $R^2$  сначала необходимо вычислить ординаты тренда для каждого момента времени наблюдения случайной величины, сформировав таким образом новую переменную  $y_t$  и значения экспериментальной случайной переменной  $x_t$ .  $R^2$  вычисляется в процедуре "K". Чем больше оценки  $R^2$ , тем лучше выбранное уравнение описывает тренд случайной величины  $x_t$ .

Другой способ обработки заключается в следующем.

В процедуре "K", позиция 1, вводится зависимая переменная – изучаемая случайная величина; в соответствии с заданием б вводится независимая переменная  $t$  (номер отсчета – 1, 2, 3, 4, 5 и т.д.  $N$  суток или месяцев или лет). Выбирается линейная модель, и проводятся необходимые расчеты значения "R-squared" – это и есть коэффициент детерминации в процентах.

### Тема 3 КОРРЕЛЯЦИОННЫЙ И РЕГРЕССИОННЫЙ АНАЛИЗ

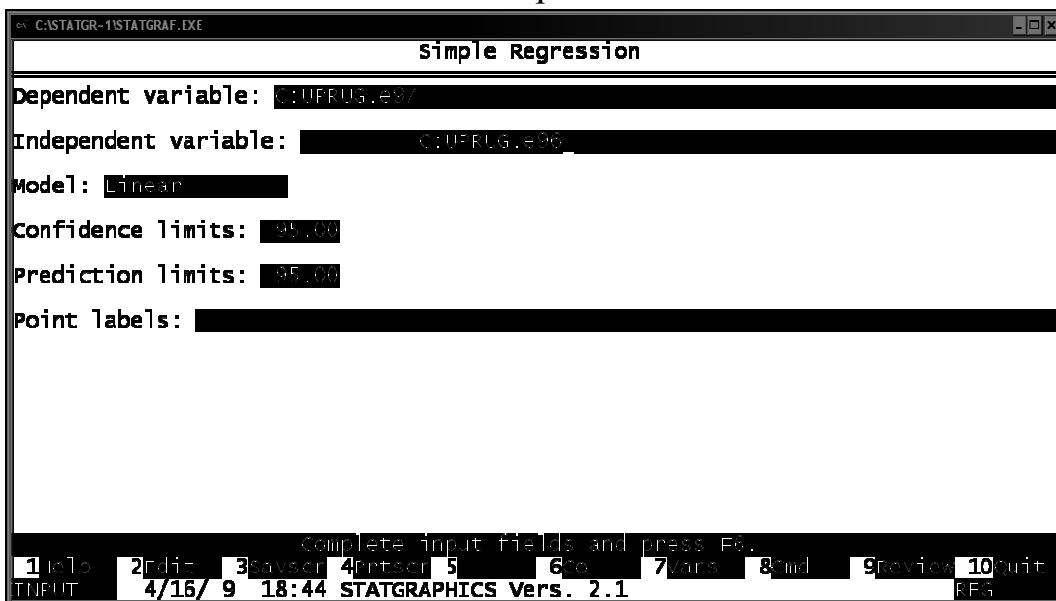
В исследованиях по климату часто возникает необходимость оценки степени взаимосвязи двух взаимодействующих величин (например, температуры поверхности почвы и температуры воздуха). Для этого, в первую очередь, необходимо оценить коэффициент линейной корреляции между признаками, построить уравнение линейной регрессии, связывающее оба признака, оценить их надежность. Для приобретения первичных практических навыков по расчетам коэффициента корреляции и параметров уравнения линейной регрессии необходимо выполнить задание б.

**ЗАДАНИЕ 6.** *Оценить тесноту и вид взаимосвязей двух случайных величин  $y_t$  и  $x_t$ : вычислить коэффициент линейной корреляции  $r(x, y)$ , найти параметры уравнений регрессии, оценить их достоверность и вклад выбранного уравнения регрессии в полные дисперсии  $\sigma^2(y)$ ;  $\sigma^2(x)$ .*

#### **Рекомендации к выполнению задания.**

1. Выбрать процедуру "K" (Regression Analysis), нажать клавишу "Enter".
2. На экране появится шесть позиций.
3. Выбрать процедуру 1 (Simple Regression), нажать клавишу "Enter".
4. На экране появятся строки ввода зависимой переменной  $t$  (Dependent Variable)  $y_t$  и независимой переменной  $x_t$ ; ввести имена соответствующих переменных.

5. В строке Model ввести пробелом тип регрессионной зависимости (линейная, мультипликативная, экспоненциальная, обратная), по умолчанию оставим Linear – Линейная. Остальные строки также оставим без изменений:



6. Нажать клавишу "F6". На экране высветятся результаты вычислений в виде таблицы:

Regression Analysis - Linear model: Y = a+bx					
Dependent variable: C:UPRUG.e97		Independent variable: C:UPRUG.e96			
Parameter	Estimate	Standard Error	T value	Prob. Level	
Intercept	1.07935	0.698716	1.54476	0.131662	
Slope	0.899689	0.0668209	13.4642	3.55271E-15	
Analysis of Variance					
Source	Sum of Squares	Df	Mean Square	F-Ratio	Prob. Level
Model	1017.0070	1	1017.0070	181.2842	.00000
Error	190.74048	34	5.61001		
Total (corr.)	1207.7475	35			
correlation coefficient = 0.917643			R-squared = 84.21 percent		
Stnd. Error of Est. = 2.36855					

Press ENTER to continue.  
 INPUT 4/16/ 9 18:44 STATGRAPHICS Vers. 2.1 REG

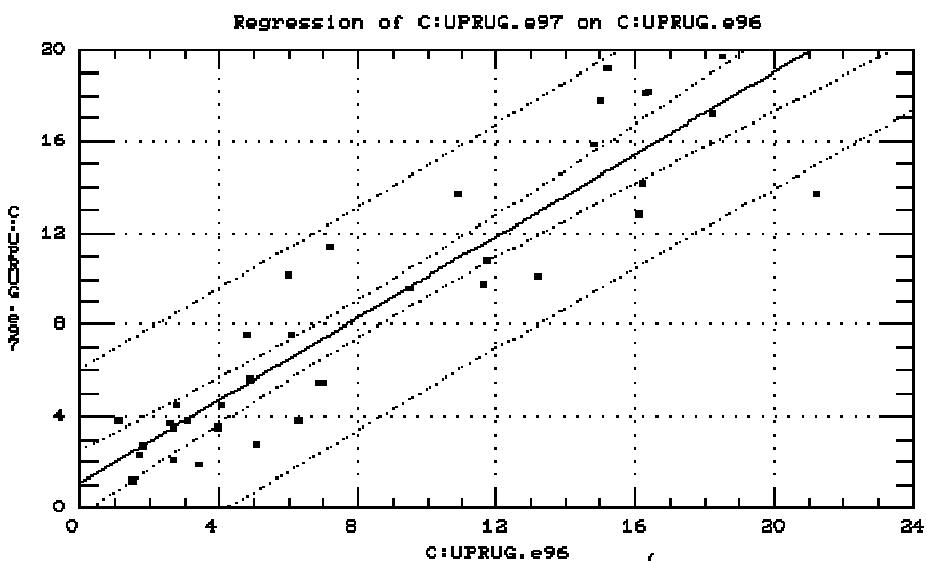
Здесь Intercept (Пересечение) – коэффициент  $a$  в уравнении линейной регрессии  $y = a + bx$ , Slope (Наклон) – коэффициент  $b$ . Так, в приведенном примере, для среднедекадной упругости водяного пара (ст. Казань – университет) за 1997 г. ( $y_t$ ) и за 1996 г. ( $x_t$ ) уравнение регрессии имеет вид  $y_t = 1,08 + 0,9 x_t$ ,  $a \approx 1,08$ ,  $b \approx 0,90$ .

В столбце Prob. Level – уровень вероятности, приводятся значения вероятности ошибки для определения коэффициентов  $a$  и  $b$ , соответственно. Так, здесь вероятность ошибки  $\alpha$  в определении коэффициента  $a$  равна 0,13, для коэффициента  $b$   $\alpha \approx 0$ , т.е. доверительная вероятность определения коэффициента  $a$  есть  $P_a = 1,00 - 0,13 = 0,87$ , для коэффициента  $b$   $P_b = 1,00 -$

$0,00 = 1,00$ , т.е. коэффициенты уравнения регрессии определены достаточно надежно.

В таблице Analysis of Variance приводятся значения сумм квадратов – Sum of Squares для Модели и для ошибки (Error), которые характеризуют их долю в общей дисперсии, в примере, 1017,0 – для модели, 190,7 – для ошибки. Ниже приводится коэффициент корреляции (Correlation Coefficient)  $R = 0,92$  и коэффициент детерминации  $R^2 = 84,2\%$ . То есть данное уравнение линейной регрессии на 84,2% описывает зависимость двух величин.

7. Нажать клавишу "Enter", выветится малое меню, войдя в первую позицию (Plot fitted Line), нажав клавишу "Enter", на экране получим график линии регрессии и 95% доверительный интервал для нее, точками на экране представлены экспериментальные значения (связи величин  $x$  и  $y$ ).



8. Под графиком (после скобки) набрать значение  $x$ , нажав клавишу "Enter", получим значения  $y$ ; это отобразится и на графике.

Вторая позиция позволяет построить график остатков, т.е. разностей значений  $y$ , вычисленных по уравнению регрессии, и экспериментальных  $y_t$ .

Для оценки достоверности  $P_r$  коэффициента корреляции  $r$  применяются различные критерии [1]. При этом можно применить следующий критерий

- если  $|r| \cdot \sqrt{N-1} > 3$ , то  $P_r > 0,99$ ,
- если же  $|r| \cdot \sqrt{N-1} > 1,96$ , то  $P_r > 0,95$ .

Здесь  $N$  – объем выборки. Полезно помнить о том, что строить уравнение линейной регрессии, связывающее две случайные величины, имеет смысл, если коэффициент корреляции значимо отличается от нуля ( $P_r \geq 0,95$ ).

## Тема 4

# ПРОВЕРКА МЕТЕОРОЛОГИЧЕСКИХ РЯДОВ НА ИХ ОДНОРОДНОСТЬ

Если в каком-либо ряду есть ярко выраженные экстремумы (*max* и *min*), то с определенной вероятностью можно считать ряд неоднородным и говорить о влиянии каких-то факторов, определяющих появление этих экстремумов. Это могут быть внешние причины (например, влияние солнечной активности на температуру верхних слоев атмосферы) и внутренние причины.

Отсутствие эффекта влияния какого-либо фактора на случайную величину (или его наличие) можно оценить с помощью критерия Краскела-Уоллиса. С помощью этого критерия одновременно можно проверить однородность рядов наблюдений. Для приобретения практических навыков по проверке однородности метеорологических рядов необходимо выполнить задание 7.

**ЗАДАНИЕ 7. Проверить гипотезу об отсутствии эффекта влияния данного фактора с помощью критерия Краскела-Уоллиса (гипотеза об однородности данных наблюдений).**

### Рекомендации к выполнению задания.

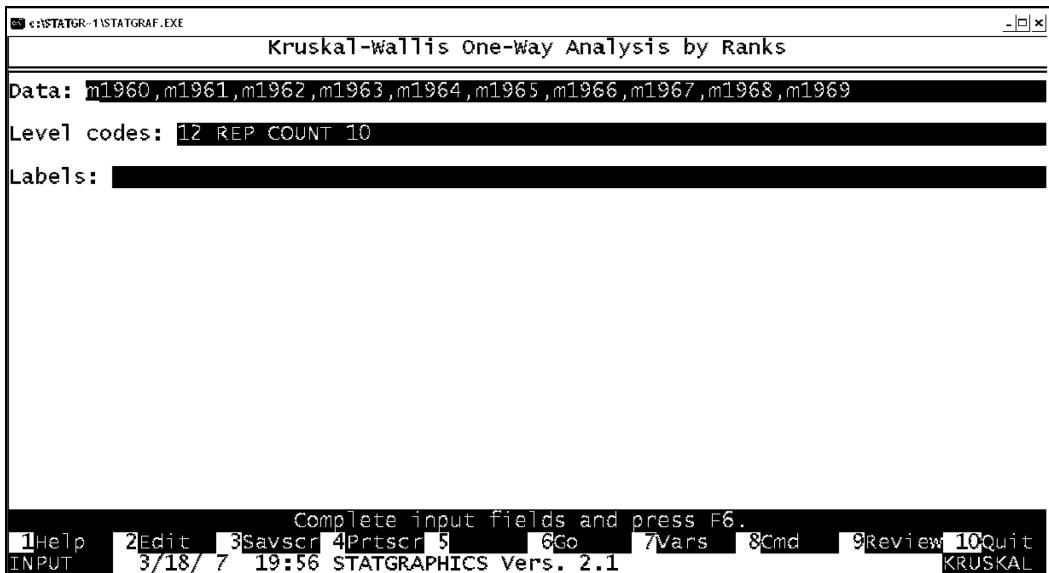
1. Выбрать процедуру "J" (Analysis of Variance – Анализ Вариаций, т.е. дисперсионный анализ), нажать клавишу "Enter".
2. На экране появится меню из пяти процедур.
3. Выбрать процедуру 4 (Kruskal-Wallis One-Way Analysis by Ranks), нажать клавишу "Enter".
4. Появится экран с полями ввода данных.
5. В поле Data (Данные) ввести имена всех переменных, отделяя их друг от друга запятой (ввод вектора переменных).

Если взять в качестве примера ежемесячные числа Вольфа (индекс солнечной активности с 1960 по 1969 гг.,  $N = 12$  мес.,  $n = 10$  лет, тогда вектор переменных надо ввести как  $m1960,m1961,m1962,\dots,m1969$ .

6. Ввести в поле Level codes (Коды уровня) для данного примера  
**12 REP COUNT 10**

Таким образом, в компьютере вектор кодов уровня будет иметь вид 1111111111122222222233333333...10101010101010101010,

т.е. первый год наблюдений 1960 помечен символом "1", второй – 1961 символом "2" и т.д.



Поле Labels (Метка) по умолчанию можно не заполнять.

7. Нажать клавишу "F6", на экране появятся результаты обработки.

Kruskal-Wallis analysis of m1960,m1961,m1962,m1963,m1964,m1965,m1966,m1967,m1968		
Level	Sample Size	Average Rank
1	12	81.1250
2	12	51.7500
3	12	44.2500
4	12	34.0833
5	12	10.9167
6	12	18.0833
7	12	64.3333
8	12	97.7500
9	12	101.250
10	12	101.458

Test statistic = 103.756 Significance level = 0

Use Esc to Quit, Cursor keys or Page Number: Page 1.1 of 1.1  
 1Help 2Edit 3Savscr 4Prtscr 5Prtptxt 6Go 7Vars 8Cmd 9Review 10Quit  
 INPUT 3/18/ 7 20:00 STATGRAPHICS Vers. 2.1 KRUSKAL

Здесь в столбце Level стоят символы, соответствующие номеру переменной (1 – m1960 , 2 – m1961, ..., 10 – m1969); в столбце Sample Size (размер выборки) – число наблюдений для каждой переменной, Average Rank (средний ранг) – указывается для каждой переменной.

Под таблицей приводится Test statistic – статистика Краскела-Уоллиса и Significance level – вероятность  $\alpha$  однородности данных; в данном примере  $\alpha = 0$ , т.е. вероятность того, что используемые данные чисел Вольфа неоднородны – эта вероятность равна  $P = 1 - \alpha = 100\%$ , вероятность ошибки равна нулю. Так, из результатов обработки видно, что уровню 5 (m1964 – 1964 год) соответствует минимальный средний ранг, уровню 10 (m1969 – 1969 год) соответствует максимальный средний ранг. По результатам наблюдений установлено, что в данном ряду есть год минимальной солнечной активности – 1964 – и год с максимальной солнечной активностью

– 1969, т.е. в целом ряд (вектор) данных чисел Вольфа неоднородный, что и доказано с помощью критерия Краскела-Уоллиса.

## Тема 5

### АНАЛИЗ ВНУТРЕННЕЙ СТРУКТУРЫ ВРЕМЕННЫХ РЯДОВ

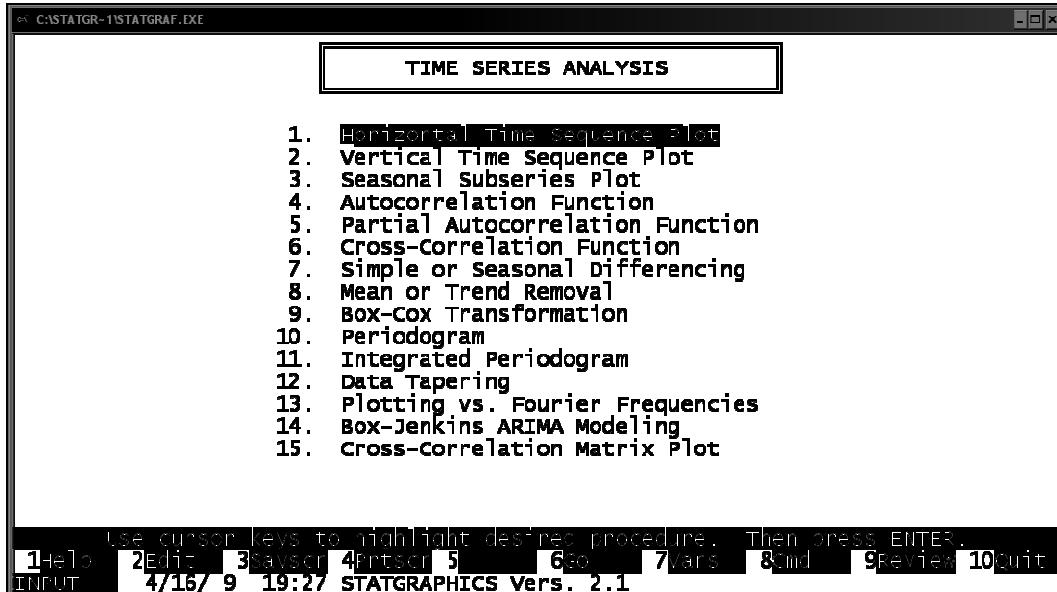
При анализе временных рядов, кроме выявления тренда, часто необходимо определить время (интервал) корреляции, периоды колебаний (флуктуаций) величины. Если в ряду имеется цикличность, то ее можно выявить с помощью различных методов. Определить периоды колебаний можно с помощью методов гармонического анализа (если ряд имеет четко выраженную периодичность, например, суточный и годовой ход), либо с помощью спектрально-корреляционного анализа (если ряд не имеет четко выраженной периодичности, длина цикла меняется во времени); последний применим для выявления периодов колебаний и в строго периодических процессах.

Для определения времени корреляции и периодов наиболее мощных флуктуаций необходимо вычислить автокорреляционную функцию и функцию спектра флуктуаций. Для этого необходимо совершить ряд операций, описанных в задании 8. В более сложных процедурах вычисления спектра флуктуаций оценивается и вероятность достоверности пиков на кривой спектра флуктуаций и соответствующих им периодов флуктуаций. Кроме того, нередко требуется выполнить фильтрацию временного ряда, позволяющую устраниТЬ или уменьшить влияние тренда, выявить пики в спектре флуктуаций в различных частотных полосах (в высокочастотной, среднечастотной или низкочастотной) [4].

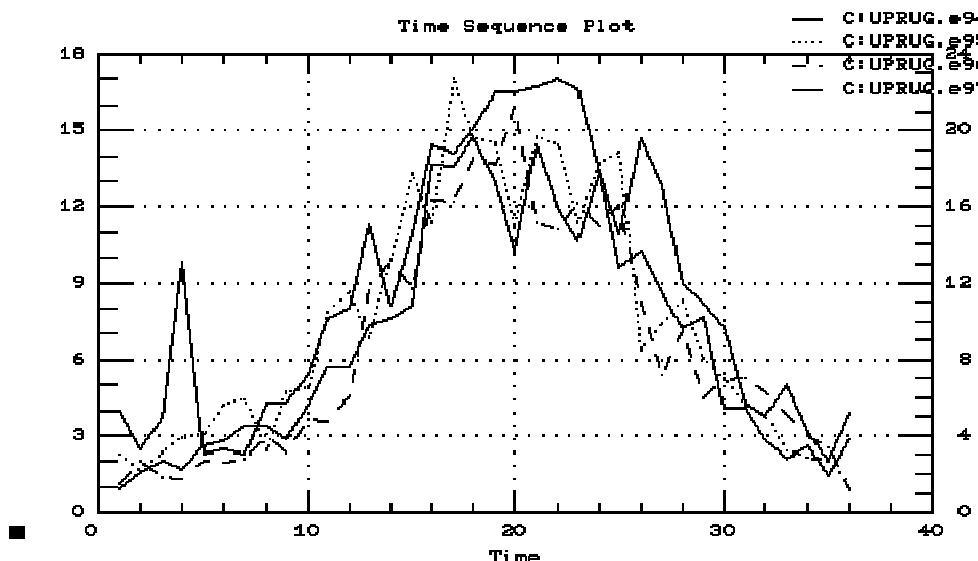
**ЗАДАНИЕ 8.** Вычислить автокорреляционную функцию, периодограмму (спектр флуктуаций), определить время корреляции и периоды наиболее мощных флуктуаций.

#### **Рекомендации к выполнению задания.**

1. Войти в процедуру "O" (Time Series Analysis – Анализ Временных Рядов), нажать клавишу "Enter".
2. Появится таблица из 15 позиций.



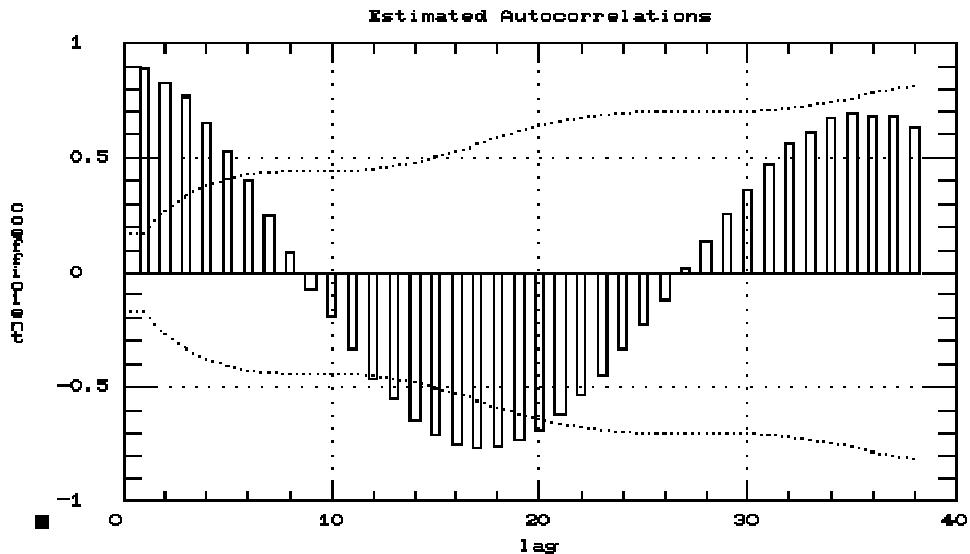
3. Выбрать позицию 1 для просмотра графика изменения случайной величины во времени, нажать клавишу "Enter". Ввести имя переменной в первой строке (можно ввести 12 переменных построчно). Нажать клавишу "F6", на запрос ввести 1 (старт с первого значения), нажать клавишу "Enter". На экране появится график.



(В примере – график упругости водяного пара на ст. Казань – университет по ежедекадным данным с 1994 по 1997 гг.).

4. Выйти из графика, нажав клавишу "Esc" 2 раза.
5. Войти в позицию 4 (Автокорреляционная функция), нажав клавишу "Enter".
6. Ввести в строку Data vector имя переменной (в примере ekgu).
7. В строке Number of lags ввести максимальный лаг (максимальную задержку, сдвиг  $L = \frac{1}{4}N$ ;  $L = \frac{1}{5}N$ ;  $L = \frac{1}{10}N$ , в примере  $L = \frac{1}{4}N = 38$ ), нажать клавишу "F6". На экране выскажется малое окно.

8. Выбрать в малом окне позицию Plot ACF – график автокорреляционной функции или позицию Display table – таблица значений ACF на дисплее. Нажать клавишу "Enter". На экране появится график



или соответствующая таблица ACF.

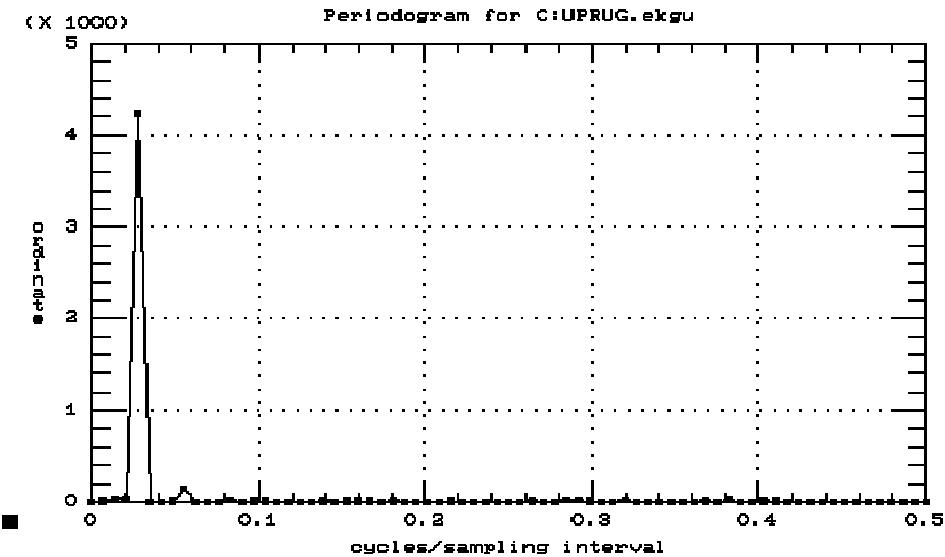
Estimated autocorrelations for C:UPRUG.ekgu					
Lag	Estimate	Stnd. Error	Lag	Estimate	Stnd. Error
1	.89372	.08333	2	.82781	.13431
3	.76630	.16600	4	.65133	.18897
5	.53315	.20397	6	.39866	.21343
7	.25101	.21854	8	.09349	.22053
9	-.07247	.22081	10	-.19307	.22097
11	-.33644	.22214	12	-.46176	.22565
13	-.55253	.23212	14	-.64746	.24108
15	-.71204	.25287	16	-.75368	.26643
17	-.77137	.28084	18	-.75872	.29519
19	-.73295	.30844	20	-.68742	.32030
21	-.61340	.33039	22	-.53846	.33820
23	-.45201	.34411	24	-.33629	.34821
25	-.22905	.35045	26	-.11674	.35149
27	.02011	.35176	28	.13623	.35177
29	.25615	.35214	30	.36053	.35343
31	.47027	.35597	32	.56003	.36026

Use Esc to exit, cursor keys or Page Number:  
**1** Help **2** Edit **3** Options **4** Print **5** Stop **6** Go **7** View **8** End **9** Review **10** Exit  
 INPUT **4/20/ 9 10:56 STATGRAPHICS Vers. 2.1** Page 1,1 of 21 ACF

Пунктирные линии на графике – границы доверительного интервала, соответствующие вероятности  $P = 0,95$ . Значения ACF, выходящие за эти границы, значимо ( $P \geq 0,95$ ) отличаются от нулевого значения. Время пересечения ACF нулевого значения можно считать интервалом корреляции, временем корреляции на уровне  $ACF = 0$ ,  $\tau_{00} = 9$  декад – в приведенном примере; время корреляции на уровне  $ACF = 0,5$ ,  $\tau_{0,5} = 5$  декад. Из графика следует, что в течение 5 декад значения случайной величины достаточно тесно связаны между собой ( $P \geq 0,95$ ). Кроме того, видно, что колебания случайной величины близки к периодическим – между двумя  $max$  ACF примерно 36 декад (т.е. имеется четкий годовой ход).

9. Выйти из графика и позиции 4 и войти в позицию 10 – Periodogram, нажать клавишу "Enter".

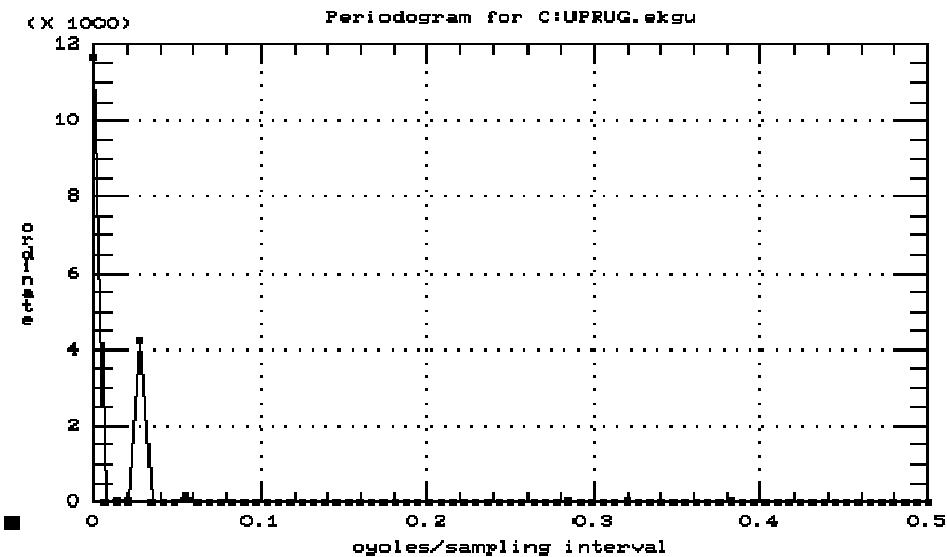
10. В строке Data vector ввести имя переменной, в строке Remove mean (Исключить среднее) вставить "Yes", остальные строки оставить без изменения, нажать клавишу "F6", появится малое окно, выбрать в малом окне строку Plot ordinates (График ординат, т.е. значений периодограммы), или строку Display results (Таблица результатов вычисления периодограммы).



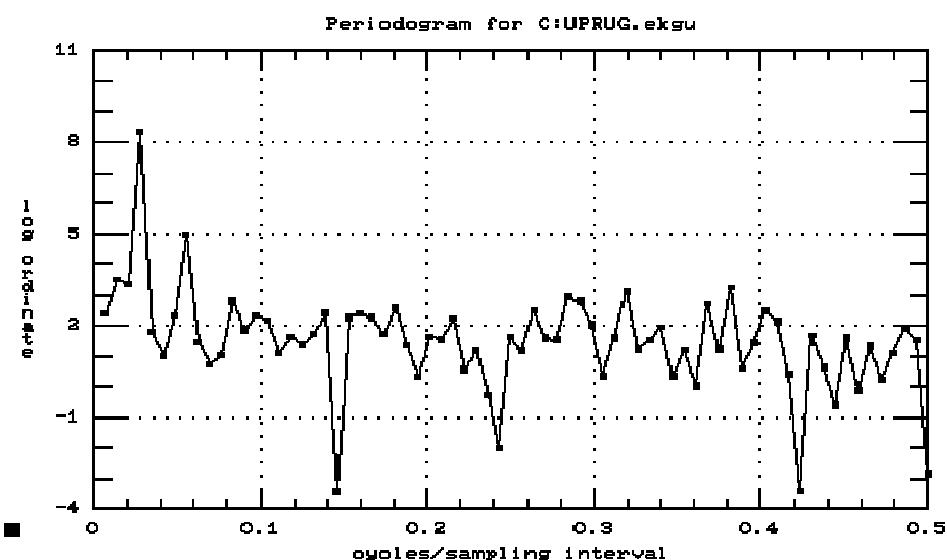
По оси абсцисс здесь частота в циклах на интервал дискретности. Из графика следует, что выделяется один мощный пик на частоте  $f_1 \approx 0,027$  цикл/декада (точное значение можно уточнить по Display results), кроме того, выделяется слабый пик на частоте  $f_2 \approx 0,06$  цикл/декада. Эти частоты соответствуют периодам  $T_1 = \frac{1}{f_1} \approx 37$  декад и  $T_2 \approx 18$  декад, т.е. выделен

годовой ход упругости водяного пара  $T_1 \approx 37$  декад  $\approx 12$  мес. и более слабый полугодовой. Можно исключить годовой ход и тогда проявятся более высокочастотные колебания.

Если в строке Remote mean ввести "No" (тренд не исключается), то в примере максимальное значение спектра будет на нулевой частоте, т.е. преобладают по мощности низкочастотные колебания, период которых сравним или больше длины всего ряда ( $T_1 \geq 144$  дек.); в тоже время остался и пик на частоте, соответствующей годовому периоду ( $T_2$ ) и маленький пик с  $T_3 \approx 18$  дек. (полугодовой цикл).



Если в строке Log for y-axis (логарифм ординаты) поставить "Yes", то ось ординат (значения спектра, периодограммы) будет в логарифмическом масштабе. Это позволяет выровнять спектр и более четко выделить менее мощные пики.



Так, в примере, более резко проявился второй по мощности пик (в данном примере  $T_2 = 18$  декад = 6 мес.).

Автокорреляционная функция и периодограмма также могут быть рассчитаны в позиции 14 (Box-Jenkins ARIMA Modeling). Кроме того, здесь приводятся оценка модели, график остатков (разница фактических значений случайной величины и значений, вычисленных по модели), автокорреляционная функция и периодограмма остатков, график прогностических значений. Расчеты в позиции 14 – более мощный анализ временных рядов.

## Литература

1. Верещагин М.А. Статистические методы в метеорологии / М.А. Верещагин, Э.П. Наумов, К.М. Шангалинский. – Казань: Изд-во Казанск. ун-та, 1990. – 107с.
2. Тюрин Ю.Н. Статистический анализ данных на компьютере / Под ред. В.Э. Фигурнова / Ю.Н. Тюрин, А.А. Макаров. – М.: Инфра-М, 1998. – 528с.
3. Тудрий В.Д. Методы статистической обработки гидрометеорологической информации (Краткий курс лекций) / В.Д. Тудрий. – Казань: Изд-во Казанск. ун-та, 2007. – 162с.
4. Тудрий В.Д. Флуктуации циклонических процессов в Северном полушарии Земли / В.Д. Тудрий, Н.В. Колобов. – Казань: Изд-во Казанск. Ун-та, 1984. – 164с.

## **ПРИЛОЖЕНИЕ**

### **Основные операции в пакете Statgraphics**

**DATA MANAGEMENT AND**

**SYSTEM UTILITIES**

A. Data Management

B. System Environment

C. Report Writer and Graphics

Replay

D. Plotter Interface

**PLOTTING AND DESCRIPTIVE**

**STATISTICS**

E. Plotting Functions

F. Descriptive Methods

G. Estimation and Testing

H. Distribution Functions

I. Exploratory Data Analysis

**ANOVA AND REGRESSION**

**ANALYSIS**

J. Analysis of Variance

K. Regression Analysis

**TIME SERIES PROCEDURES**

L. Forecasting

M. Quality Control

N. Smoothing

O. Time Series Analysis

**ADVANCED PROCEDURES**

P. Categorical Data Analysis

Q. Multivariate Methods

R. Nonparametric Methods

S. Sampling

T. Experimental Design

**MATHEMATICAL AND USER**

**PROCEDURES**

U. Mathematical Functions

**УПРАВЛЕНИЕ ДАННЫМИ И**

**СИСТЕМНЫЕ УТИЛИТЫ**

А. Управление Данными

В. Системное Окружение

С. Печать Отчетов и Графических  
Рисунков

D. Графические Атрибуты

**РИСОВКА И ОПИСАТЕЛЬНАЯ**  
**СТАТИСТИКА**

E. Функции Рисовки

F. Описательные Методы

G. Оценка и Тестирование

H. Функции Распределения

I. Анализ Экспериментальных Данных  
**ДИСПЕРСИОННЫЙ И**  
**РЕГРЕССИОННЫЙ АНАЛИЗ**

J. Дисперсионный Анализ

K. Регрессионный Анализ

**ВРЕМЕННЫЕ ПОСЛЕДОВАТ-ТИ**

L. Прогнозирование

M. Контроль Качества

N. Сглаживание

O. Временной Анализ

**ДОПОЛНИТ. ПРОЦЕДУРЫ**

P. Анализ Категориальных Данных

Q. Многомерные Методы

R. Непараметрические Методы

S. Моделирование

T. Планирование Эксперимента

**МАТЕМАТИЧЕСКИЕ ПРОЦЕДУРЫ**

U. Математические Функции

V. Supplementary Operations	
DATA MANAGEMENT	
1. Display Data Directory	
2. File Operations	
3. Import Data Files	
4. Export Data Files	
PLOTTING FUNCTIONS	
1. X-Y Line and Scatterplots	
2. Multiple X-Y Plots	
3. X-Y-Z Line and Scatterplots	
4. Multiple X-Y-Z Plots	
5. Barcharts	
6. Piecharts	
7. Component Line Charts	
DESCRIPTIVE METHODS	
1. Summary Statistics	
2. Frequency Tabulation	
3. Frequency Histogram	
4. Weighted Averages	
5. Percentiles	
6. Codebook Procedure	
7. Three-Dimensional Histogram	
ESTIMATION AND TESTING	
1. One-Sample Analysis	
2. Two-Sample Analysis	
3. Normal Probability Plot	
4. Hanging Histobars	
5. Comparison of Poisson Rates	
DISTRIBUTION FUNCTIONS	
1. Distribution Fitting	
2. Distribution Plotting	
3. Tail Area Probabilities	

V. Дополнительные Операции	
УПРАВЛЕНИЕ ДАННЫМИ	
1. Экран Каталога Данных	
2. Операции над Файлами	
3. Импорт Файлов Данных	
4. Экспорт Файлов Данных	
ФУНКЦИИ РИСОВКИ	
1. Графики Функций Одной Переменной	
2. Несколько Графиков Функций Одной Переменной	
3. Трехмерные графики	
4. Несколько Трехмерных Графиков	
5. Столбчатые Диаграммы	
6. Круговые Диаграммы	
7. Линейные Диаграммы	
ОПИСАТЕЛЬНЫЕ МЕТОДЫ	
1. Элементарная Статистика	
2. Табуляция Частот	
3. Гистограмма Частот	
4. Взвешенные Средние	
5. Процентили	
6. Статистика в Классах	
7. Трехмерная Гистограмма	
ОЦЕНКА И ТЕСТИРОВАНИЕ	
1. Одно-Выборочный Анализ	
2. Двух-Выборочный Анализ	
3. Рис. Нормального Распределения	
4. Висячая Гистограмма	
5. Тест Пуассона	
ФУНКЦИИ РАСПРЕДЕЛЕНИЯ	
1. Подбор Распределения	
2. Графики Распределения	
3. «Хвосты» Распределения	

- 4. Critical Values
  - 5. Random Number Generation
- EXPLORATORY DATA ANALYSIS**
- 1. Box-and-Whisker Plot
  - 2. Multiple Box-and-Whisker Plot
  - 3. Notched Box-and-Whisker Plot
  - 4. Median Polish of Two-Way Table
  - 5. Resistant Nonlinear Smoothing
- 
- 6. Suspended Rootogram
  - 7. Stem-and-Leaf Display
- ANALYSIS OF VARIANCE**
- 1. One-Way Analysis of Variance
  - 2. Multifactor Analysis of Variance
- 
- 3. Analysis of Nested Designs
  - 4. Kruskal-Wallis One-Way Analysis by Ranks
  - 5. Friedman Two-Way Analysis by Ranks
- REGRESSION ANALYSIS**
- 1. Simple Regression
  - 2. Interactive Outlier Rejection
  - 3. Multiple Regression
  - 4. Stepwise Variable Selection
  - 5. Ridge Regression
  - 6. Nonlinear Regression
- FORECASTING**
- 1. Brown's Exponential Smoothing

- 4. Критические Значения
  - 5. Генерация Случайных Чисел
- АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ**
- 1. Обзорный Рисунок
  - 2. Составной Обзорный Рисунок
  - 3. Усеченный Обзорный Рисунок
  - 4. Медианный Анализ Двумерн. Табл.
- 
- 5. Устойчивое Нелинейное Сглаживание
  - 6. Подвесная Корнеграмма
  - 7. Анализ Симметрии
- ДИСПЕРСИОННЫЙ АНАЛИЗ**
- 1. Однофакторный Дисперсионный Анализ
  - 2. Многофакторный Дисперсионный Анализ
  - 3. Анализ Грунтовых Планов
  - 4. Краскела-Уоллиса Однофакторный Ранговый Анализ
  - 5. Двухфакторный Ранговый Анализ Фридмана
- РЕГРЕССИОННЫЙ АНАЛИЗ**
- 1. Простая Регрессия
  - 2. Интерактивное Отбрасывание
  - 3. Множественная Регрессия
  - 4. Шаговая Регрессия
  - 5. Гребневая Регрессия
  - 6. Нелинейная Регрессия
- ПРОГНОЗ**
- 1. Экспоненциальное Сглаживание по Брауну

- |   |  |
|---|--|
| <p>2. Holt's Linear Exponential Smoothing</p> <p>3. Winter's Seasonal Smoothing</p> <p>4. Trend Analysis</p> <p>5. Seasonal Decomposition</p> <p><b>QUALITY CONTROL</b></p> <p>1. X-Bar Charts for Sample Averages</p> <p>2. R Charts for Sample Ranges</p> <p>3. S Charts for Standard Deviations</p> <p>4. C Charts for Poisson Counts</p> <p>5. U Charts for Counts per Unit</p> <p>6. NP Charts for Binomial Counts</p> <p>7. P Charts for Binomial Proportions</p> <p>8. Cusum Charts for Sample Averages</p> <p>9. Pareto Charts</p> <p>10. Multivariate Control Charts</p> <p><b>SMOOTHING</b></p> <p>1. Simple Moving Average</p> <p>2. Weighted Moving Average</p> <p>3. Polynomial Smoothing</p> <p>4. Open and Closed Q-Splines</p> <p>5. Poisson Rate Function Estimation</p> <p><b>TIME SERIES ANALYSIS</b></p> <p>1. Horizontal Time Sequence Plot</p> <p>2. Vertical Time Sequence Plot</p> <p>3. Seasonal Subseries Plot</p> <p>4. Autocorrelation Function</p> <p>5. Partial Autocorrelation Function</p> <p>6. Cross-Correlation Function</p> | <p>2. Линейное Экспоненциальное Сглаживание по Хольту</p> <p>3. Сезонное Сглаживание по Винтеру</p> <p>4. Тренд Анализ</p> <p>5. Сезонное Разложение</p> <p><b>КОНТРОЛЬ КАЧЕСТВА</b></p> <p>1. X-диаг. для Выборочных Средних</p> <p>2. R-диаг. для Размахов Выборок</p> <p>3. S-диаг. для Стандартного Отклонения</p> <p>4. С-диаг. для Счета Пуассона</p> <p>5. U-диаг. для Счета на Единицу</p> <p>6. NP-диаг. для Биномиального Счета</p> <p>7. P-диаг. для Биномиальных Пропорций</p> <p>8. Кусума-Диаграмма для Выборочных Средних</p> <p>9. Парето-Диаграмма</p> <p>10. Многовариантный Контроль СГЛАЖИВАНИЕ</p> <p>1. Скользящее Среднее</p> <p>2. Взвешенное Скользящее Среднее</p> <p>3. Полиномиальное Сглаживание</p> <p>4. Q-Сплайны</p> <p>5. Оценка Функции Частот Пуассона</p> <p><b>ВРЕМЕННОЙ АНАЛИЗ</b></p> <p>1. Горизонтальный График</p> <p>2. Вертикальный График</p> <p>3. График Сезонных посл-ей</p> <p>4. Автокорреляционная Функция</p> <p>5. Частная Автокорреляционная Функция</p> <p>6. Кросс-корреляционная Функция</p> |
|---|--|

- 7. Simple or Seasonal Differencing
- 8. Mean or Trend Removal
- 9. Box-Cox Transformation
- 10. Periodogram
- 11. Integrated Periodogram
- 12. Data Tapering
- 13. Plotting vs. Fourier Frequencies
- 14. Box-Jenkins ARIMA Modeling
- 15. Cross-Correlation Matrix Plot
- 7. Простое Сезон. Выч-е Послед. Раз.
- 8. Удаление Среднего/Тренда
- 9. Стабилизирующая Трансформация
- 10. Периодограмма
- 11. Интегрирование Периодограмма
- 12. Сужение Данных
- 13. Рисовка Частот Фурье
- 14. Моделирование Бокса-Дженкинса ARIMA
- 15. Рис. Матрицы Перекрест. Коррел.

