Министерство образования Республики Беларусь

Учреждение образования
«Гомельский государственный технический
университет имени П. О. Сухого»

Кафедра «Электроснабжение»

МАТЕМАТИЧЕСКИЕ ЗАДАЧИ ЭЛЕКТРОЭНЕРГЕТИКИ

ПОСОБИЕ
для студентов специальности
1-43 01 03 «Электроснабжение»
дневной и заочной форм обучения

Гомель 2016

УДК 621.31.51(075.8)
ББК 65.305.142я73
М34

Рекомендовано научно-методическим советом
энергетического факультета ГГТУ им. П. О. Сухого
(*протокол № 6 от 24.02.2015 г.*)

Составители: О. М. Попова, Т. В. Алферова, А. А. Алферов

Рецензент: канд. физ.-мат. наук, доц. каф. «Высшая математика» ГГТУ им. П. О. Сухого
С. Л. Авакян

Математические задачи электроэнергетики : пособие для студентов
М34 специальности 1-43 01 03 «Электроснабжение» днев. и заоч. форм обучения / сост.: О. М. Попова, Т. В. Алферова, А. А. Алферов. – Гомель : ГГТУ им. П. О. Сухого, 2016. – 65 с. – Систем. требования: PC не ниже Intel Celeron 300 МГц ; 32 Mb RAM ; свободное место на HDD 16 Mb ; Windows 98 и выше ; Adobe Acrobat Reader. – Режим доступа: https://elib.gstu.by. – Загл. с титул. экрана.

Содержит краткие теоретические сведения, задания и порядок их выполнения по восьми лабораторным работам, позволяющим студентам закрепить теоретические знания и получить практические навыки по проведению прикладных математических исследований с использованием специализированных программ в рамках основных разделов рабочей программы дисциплины «Математические задачи электроэнергетики».

Для студентов специальности 1-43 01 03 «Электроснабжение» дневной и заочной форм обучения.

УДК 621.31.51(075.8)
ББК 65.305.142я73

ПРЕДИСЛОВИЕ

Лабораторные работы по дисциплине «Математические задачи электроэнергетики» имеют целью приобретение практических навыков, необходимых для самостоятельного проведения прикладных математических исследований в области получаемой в ВУЗе специальности, и выполняются на персональных ЭВМ.

Лабораторная работа №1 посвящена знакомству с теоремами теории вероятности и составлению алгебраических выражений, определяющих вероятность безотказной работы или отказа для заданной схемы.

В лабораторной работе №2 рассматривается последовательность независимых испытаний как модель повреждаемости однотипных элементов системы и определяется вероятность сложных событий с использованием формулы Бернулли.

В лабораторной работе №3 изучаются случайные величины, представлены их законы распределения и определяются числовые характеристики по статистическим данным.

В лабораторной работе №4 рассматривается обработка исходных данных эргодической стационарной случайной функции и определяются её характеристики.

Лабораторная работа №5 посвящена статистическим критериям и их применению. В работе рассматриваются: распределения Стьюдента, Фишера, Кохрена,Пирсона,τ-распределение.

В лабораторных работах №6-7 рассмотрен корреляционный и регрессионный анализ случайных величин.

В лабораторной работе №8 исследуются логические схемы и уравнения на безотказность и отказ в системах электроснабжения.

Отчёт по каждой лабораторной работе содержит:

1. Тему;
2. Цель работы;
3. Краткие теоретические сведения;
4. Исходные данные и расчетные параметры;
5. Результаты расчетов (графики, рисунки) ;
6. Выводы по работе.

ЛАБОРАТОРНАЯ РАБОТА № 1
Теоремы теории вероятности. Определение вероятностей отказа и безотказной работы схемы электроснабжения

1.1 Цель работы

На основе теорем теории вероятностей для заданной схемы составить алгебраические выражения, определяющие вероятность безотказной работы или отказа.

1.2 Краткие теоретические сведения

Основные понятия

Случайным событием называется всякий факт, который в результате опыта может произойти или не произойти.

Вероятностью события называется численная мера степени объективной возможности этого события. Вероятность события A обозначается $P(A)$.

Достоверным называется событие U, которое в результате опыта непременно должно произойти.

$$P(U) = 1.$$

Невозможным называется событие, которое в результате опыта не может произойти.

$$P(V) = 0.$$

Вероятность любого события A заключена между нулем и единицей:

$$0 \le P(A) \le 1.$$

Полной группой событий называется несколько событий, таких, что в результате опыта непременно должно произойти хотя бы одно из них.

Несколько событий в данном опыте называются ***несовместными,*** если никакие два из них не могут появиться вместе.

Несколько событий в данном опыте называются ***равновозможными,*** если по условиям симметрии опыта нет оснований считать какое-либо из них более возможным, чем любое другое.

Если несколько событий: 1) образуют полную группу; 2) несовместны; 3) равновозможные, то они называются ***случаями.***

Теоремы сложения и умножения вероятностей

Суммой двух событий A и B называется событие C, состоящее в появлении хотя бы одного из событий A или B.

Суммой нескольких событий называется событие, состоящее в появлении хотя бы одного из этих событий.

Произведением двух событий A и B называется событие C, состоящее в совместном появлении события A и события B.

Произведением нескольких событий называется событие, состоящее в совместном появлении всех этих событий.

Теорема сложения вероятностей

Вероятность суммы двух несовместных событий равна сумме вероятностей этих событий:

$$P(A+B) = P(A) + P(B).$$

В случае, когда события A и B совместны, вероятность их суммы выражается формулой

$$P(A+B) = P(A) + P(B) - P(AB),$$

где AB - произведение событий A и B.

Вероятность суммы нескольких (более двух) несовместных событий равна сумме их вероятностей:

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

В случае, когда события A_i совместны, вероятность их суммы выражается формулой

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i} P(A_i) - \sum_{i,j} P(A_i A_j) + \sum_{i,j,k} P(A_i A_j A_k) - \ldots$$
$$+ (-1)^{n-1} \cdot P\left(A_1 A_2 A_3 \ldots A_n\right),$$

где суммы распространяются на все возможные комбинации различных индексов $i, j, k, \ldots$, взятых по одному, по два, по три и т.д.

Если события $A_1, A_2, \ldots, A_n$ несовместны и образуют полную группу, то сумма их вероятностей равна единице:

$$\sum_{i=1}^{n} P(A_i) = 1.$$

Событие $\overline{A}$ называется ***противоположным*** событию A, если оно состоит в не появлении события A.

Сумма вероятностей противоположных событий равна единице:

$$P(A) + P(\overline{A}) = 1.$$

Условной вероятностью события А при наличии В называется вероятность события, вычисленная при условии, что событие B произошло. Эта вероятность обозначается $P(A|_B)$.

События A и B называются ***независимыми,*** если появление одного из них не меняет вероятности появления другого. Для независимых событий

$$P\left(A\big|_B\right) = P(A); \qquad P\left(B\big|_A\right) = P(B).$$

Теорема умножения вероятностей

Вероятность произведения двух событий равна вероятности одного из них, умноженной на условную вероятность другого при наличии первого:

$$P(A\,B) = P(A)\,P\left(B\big|_A\right),$$

или

$$P(A\,B) = P(B)\,P\left(A\big|_B\right).$$

Вероятность произведения более двух событий определяется формулой

$$P(A_1\,A_2\,A_3 \cdots A_n) = P(A_1)P\left(A_2\big|_{A_1}\right)P\left(A_3\big|_{A_1\,A_2}\right)\cdots P\left(A_n\big|_{A_1\,A_2\,A_3 \cdots A_{n-1}}\right).$$

Для независимых событий A и B, т.е. появление события A не меняет вероятности появления события B, вероятность их произведения определяется

$$P(AB) = P(A) \cdot P(B).$$

В случае определения вероятности произведения n независимых событий, т.е. появление любого из них не меняет вероятностей появления остальных, определяется формулой

$$P\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i).$$

1.3 Задание

Записать алгебраическое выражение вероятности соответствующего состояния электрической цепи АВ.

Состояния: *а*) отказ цепи АВ,

б) безотказная работа цепи АВ.

Схемы состояния электрической цепи АВ приведены на рис.1.1

Вероятность включения i-го контакта выбирается по таблице 1.1

Таблица 1.1

Вероятность включения контакта

№ варианта	Вероятность включения контакта			
	1	2	3	4
1	p_1	p_2	p_3	p_4
2	p_1	p_1	p_2	p_2
3	p	p	p	p
4	p_3	p_3	p_4	p_4
5	p_2	p_2	p_3	p_3
6	p	p	p_3	p_4
7	p_1	p_2	p	p
8	p_1	p	p	p_4
9	p	p_2	p_3	p
0	p_1	p_2	p_4	p_4

Схема состояния и вариант вероятностей включения контактов определяется преподавателем.

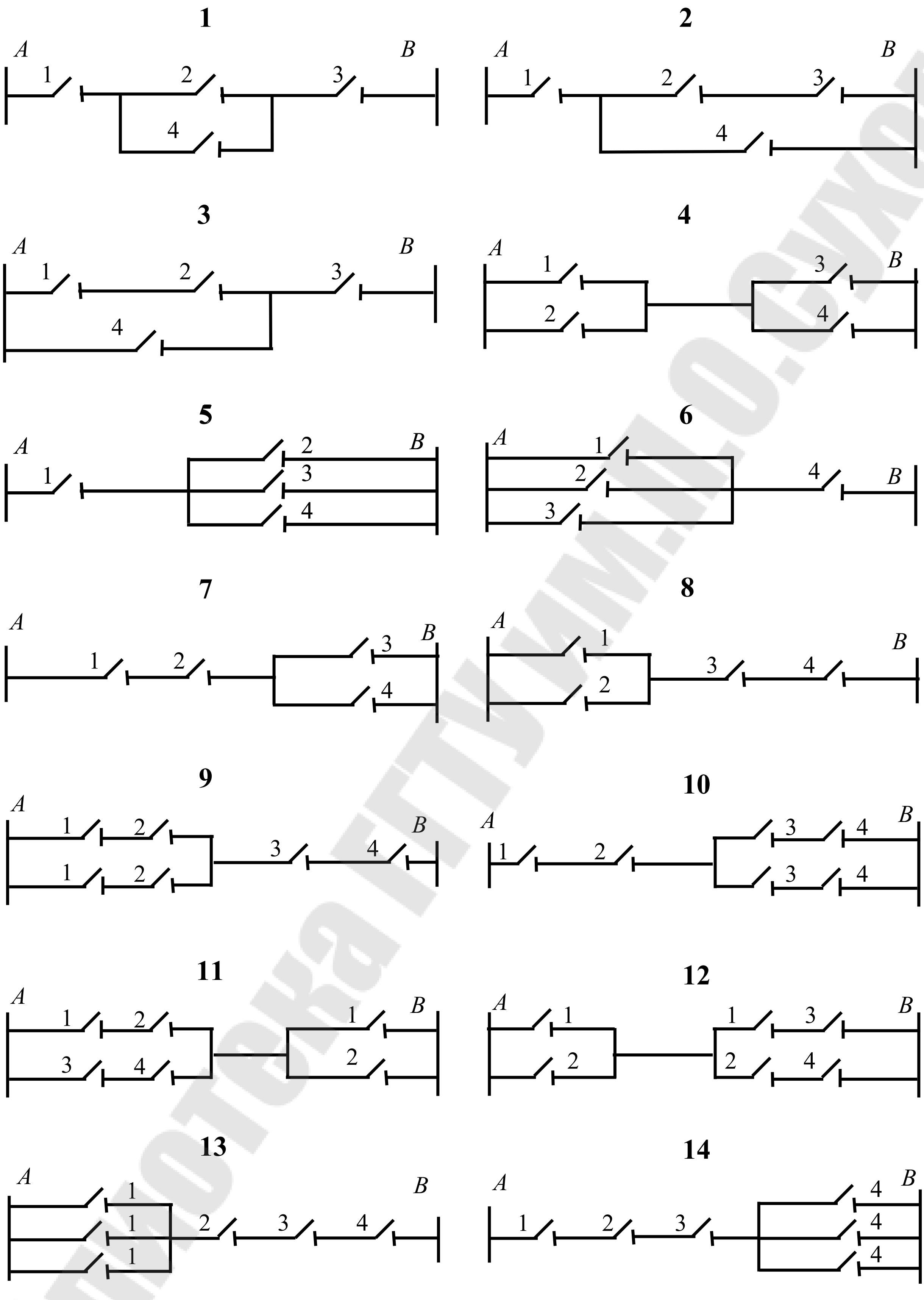
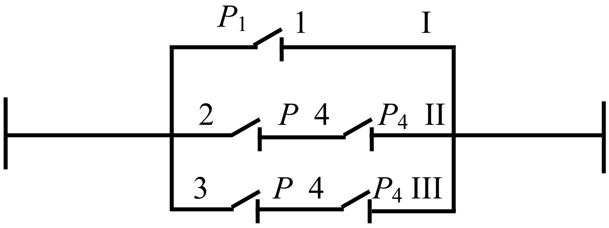


Рис. 1.1. Схемы состояния электрической цепи АВ

1.4 Методические указания к выполнению задания

Записать алгебраическое выражение вероятности безотказного состояния для приведенной схемы



Решение

Безотказное состояние цепи может быть представлено некоторым количеством независимых несовместных событий, в которых будет обеспечиваться одновременно совместное включенное и отключенное состояние различных контактов, например, включенный контакт 1 и отключенные контакты 2, 3 и 4.

Для записи соответствующих противоположных состояний примем обозначения для контактов: включен контакт, например, 1 и отключен контакт - $\bar{1}$. Контакт 4 имеется во втором и третьем участках, поэтому обозначим их 4 и 4′ соответственно включены, а $\bar{4}$ и $\overline{4'}$ - отключены.

Вероятность i-го ранее указанного состояния будет иметь вид

$$P_i\left(1\cdot\bar{2}\cdot\bar{4}\cdot\bar{3}\cdot\overline{4'}\right).$$

На основании теоремы умножения независимых событий P_i будет определяться выражением

$$P_i = p(1)\cdot p(\bar{2})\cdot p(\bar{4})\cdot p(\bar{3})\cdot p(\overline{4'}).$$

В последнее выражение можно подставить исходные данные вероятности соответствующих контактов

$$P_i = p_1(1\text{-}p)(1\text{-}p_4)(1\text{-}p)(1\text{-}p_4) = p_1(1\text{-}p)^2(1\text{-}p_4)^2.$$

Представим все возможные состояния, обеспечивающие безотказность цепи:

1: $(1\cdot 2\cdot 4\cdot 3\cdot 4')$ $\quad P_1 = p_1\cdot p\cdot p_4\cdot p\cdot p_4$;

2: $(\bar{1}\cdot 2\cdot 4\cdot 3\cdot 4')$ $\quad P_2 = (1-p_1)\cdot p\cdot p_4\cdot p\cdot p_4$;

3: $(1\cdot\bar{2}\cdot 4\cdot 3\cdot 4')$ $\quad P_3 = p_1\cdot(1-p)\cdot p_4\cdot p\cdot p_4$;

4: $(1\cdot 2\cdot\bar{4}\cdot 3\cdot 4')$ $\quad P_4 = p_1\cdot p\cdot(1-p_4)\cdot p\cdot p_4$;

5: $(1\cdot 2\cdot 4\cdot\bar{3}\cdot 4')$ $\quad P_5 = p_1\cdot p\cdot p_4\cdot(1-p)\cdot p_4$;

$6{:}\ (1\cdot 2\cdot 4\cdot 3\cdot \overline{4'})\quad P_6 = p_1\cdot p\cdot p_4\cdot p\cdot (1-p_4)$;

$7{:}\ (\bar{1}\cdot \bar{2}\cdot 4\cdot 3\cdot 4')\quad P_7 = (1-p_1)\cdot (1-p)\cdot p_4\cdot p\cdot p_4$;

$8{:}\ (\bar{1}\cdot 2\cdot \bar{4}\cdot 3\cdot 4')\quad P_8 = (1-p_1)\cdot p\cdot (1-p_4)\cdot p\cdot p_4$;

$9{:}\ (\bar{1}\cdot 2\cdot 4\cdot \bar{3}\cdot 4')\quad P_9 = (1-p_1)\cdot p\cdot p_4\cdot (1-p)\cdot p_4$;

$10{:}\ (\bar{1}\cdot 2\cdot 4\cdot 3\cdot \overline{4'})\quad P_{10} = (1-p_1)\cdot p\cdot p_4\cdot p\cdot (1-p_4)$;

$11{:}\ (1\cdot \bar{2}\cdot \bar{4}\cdot 3\cdot 4')\quad P_{11} = p_1\cdot (1-p)\cdot (1-p_4)\cdot p\cdot p_4$;

$12{:}\ (1\cdot \bar{2}\cdot 4\cdot \bar{3}\cdot 4')\quad P_{12} = p_1\cdot (1-p)\cdot p_4\cdot (1-p)\cdot p_4$;

$13{:}\ (1\cdot \bar{2}\cdot 4\cdot 3\cdot \overline{4'})\quad P_{13} = p_1\cdot (1-p)\cdot p_4\cdot p\cdot (1-p_4)$;

$14{:}\ (1\cdot 2\cdot \bar{4}\cdot \bar{3}\cdot 4')\quad P_{14} = p_1\cdot p\cdot (1-p_4)\cdot (1-p)\cdot p_4$;

$15{:}\ (1\cdot 2\cdot \bar{4}\cdot 3\cdot \overline{4'})\quad P_{15} = p_1\cdot p\cdot (1-p_4)\cdot p\cdot (1-p_4)$;

$16{:}\ (1\cdot 2\cdot 4\cdot \bar{3}\cdot \overline{4'})\quad P_{16} = p_1\cdot p\cdot p_4\cdot (1-p)\cdot (1-p_4)$;

$17{:}\ (\bar{1}\cdot \bar{2}\cdot \bar{4}\cdot 3\cdot 4')\quad P_{17} = (1-p_1)\cdot (1-p)\cdot (1-p_4)\cdot p\cdot p_4$;

$18{:}\ (\bar{1}\cdot 2\cdot 4\cdot \bar{3}\cdot \overline{4'})\quad P_{18} = (1-p_1)\cdot p\cdot p_4\cdot (1-p)\cdot (1-p_4)$;

$19{:}\ (1\cdot \bar{2}\cdot \bar{4}\cdot \bar{3}\cdot 4')\quad P_{19} = p_1\cdot (1-p)\cdot (1-p_4)\cdot (1-p)\cdot p_4$;

$20{:}\ (1\cdot 2\cdot \bar{4}\cdot \bar{3}\cdot \overline{4'})\quad P_{20} = p_1\cdot p\cdot (1-p_4)\cdot (1-p)\cdot (1-p_4)$;

$21{:}\ (1\cdot \bar{2}\cdot \bar{4}\cdot \bar{3}\cdot \overline{4'})\quad P_{21} = p_1\cdot (1-p)\cdot (1-p_4)\cdot (1-p)\cdot (1-p_4)$.

Вероятность безотказности цепи будет равна сумме вероятности перечисленных состояний.

$$P_\Sigma = \sum_{i=1}^{21} P_i .$$

После преобразований приведенных выражений получим

$$P_\Sigma = p_1 + 2pp_4 - 2pp_1p_4 - p^2p_4^2 + p^2p_1p_4^2 .$$

1.5 Контрольные вопросы

1. Дайте определения: а) случайного события; б) случайной величины.

2. Определение «вероятности события».

3. Приведите пояснения «достоверного» и «невозможного» событий.

4. Вычисление вероятностей.

5. Приведите определения а) дискретной случайной величины, б) непрерывной случайной величины.

6. Приведите пояснения понятий «сумма» и «произведение» событий и их представление на диаграмме Вьенна.

7. Определение вероятностей суммы несовместных и совместных событий.

8. Определение вероятностей противоположных событий.

9. Понятие условной вероятности событий *A* и *B*.

10. Определение вероятностей произведения событий *A, B* и *C* при условии зависимых и независимых событий.

ЛАБОРАТОРНАЯ РАБОТА № 2
Последовательность независимых испытаний как модель повреждаемости однотипных элементов системы

2.1 Цель работы

Определить вероятности сложных событий, используя схему независимых испытаний (формулу Бернулли)

2.2 Краткие теоретические сведения

Последовательность независимых испытаний представляет собой производство многократных независимых испытаний в одних и тех же условиях, при этом в каждом из испытаний возможно одно и то же число несовместных событий. Например, станок выпускает детали с различными допусками. Оценка работы станка производится следующим образом: из партии деталей берётся 100 штук и оценивается количество деталей с различными отклонениями. Опыты повторяются не однократно. Из 10 тысяч деталей (изделий) брак будет постоянным числом.

Вероятности повреждения различного количества агрегатов в системе электроснабжения при большом числе однотипных агрегатов могут быть определены по биноминальной формуле для схемы независимых испытаний (схема Бернулли).

Во многих практических случаях при многократных независимых испытаниях могут быть только два исхода: случайное событие А произойдет или не произойдет. Пусть вероятность того, что в каждом из этих независимых испытаний событие A произойдет, равна p, где p – статистическая вероятность. Тогда вероятность противоположного события (событие A не произойдёт)

$$q = 1 - p.$$

Зная p или q, можно определить вероятность того, что в n независимых испытаниях событие A, например, повреждение агрегата, случится m раз. Обозначим эту вероятность P_n^m, которая равна произведению числа комбинаций из n по m на вероятность события в степени m и на противоположную вероятность в степени $(n-m)$:

$$P_n^m = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m}. \tag{2.1}$$

Выражение (1.1) называют формулой Бернулли, определяет вероятность биномиального распределения. Очевидно, что

$$\sum_{m=0}^{n} P_n^m = 1, \tag{2.2}$$

т. к. эта сумма охватывает все возможные события (m варьируется от 0 до n).

Если число однотипных агрегатов n, вероятность не рабочего состояния p, то можно найти:

1) Вероятность повреждения ровно m элементов из n

$$P_n^m = C_n^m p^m q^{n-m};$$

2) Вероятность повреждения элементов от m_1 до m_2 из n

$$P_n^{m_1 \le m \le m_2} = \sum_{m_1}^{m_2} C_n^m p^m q^{n-m};$$

3) Вероятность повреждения не более k элементов из n

$$P_n^{0<m\le k} = \sum_{0}^{k} C_n^m p^m q^{n-m};$$

4) Вероятность повреждения более k элементов из n

$$P_n^{m>k} = \sum_{k+1}^{n} C_n^m p^m q^{n-m};$$

5) Вероятность того, что m элементов находятся в рабочем состоянии из n

$$P_n^m = C_n^m p^{n-m} q^m;$$

6) Вероятность того, что все элементы находятся в рабочем состоянии ($m = 0$)

$$P_n^0 = C_n^0 p^0 q^{n-0} = q^n.$$

Формула Бернулли применяется при определении вероятностей сложных событий.

2.3 Задание

Определить вероятности нагрузок головного участка, по заданному варианту исходных данных, используя программу «WEROT». Вариант исходных данных определяется преподавателем по таблицам 2.1, 2.2.

Таблица 2.1

Вариант исходных данных

№ вар.	СП$_1$			СП$_2$			СП$_3$			СП$_4$		
	n_1	S_1	$P_{в1}$	n_2	S_2	$P_{в2}$	n_3	S_3	$P_{в3}$	n_4	S_4	$P_{в4}$
1	2	3	0,5	3	4	0,6	4	5	0,7	2	7	0,4
2	4	3	0,4	5	5	0,3	3	2	0,6	2	5	0,6
3	2	10	0,3	3	5	0,4	3	15	0,5	4	10	0,6
4	3	5	0,4	2	10	0,5	4	5	0,3	3	15	0,7
5	2	3	0,8	3	5	0,4	4	2	0,6	5	5	0,4
6	5	10	0,7	3	5	0,6	2	15	0,5	4	5	0,4
7	4	15	0,6	6	10	0,5	3	15	0,4	2	10	0,7
8	2	7	0,5	4	5	0,4	6	3	0,7	3	4	0,6
9	3	15	0,4	2	10	0,7	4	15	0,6	6	10	0,5
0	5	5	0,6	3	5	0,6	2	3	0,2	4	2	0,4

Таблица 2.2

P(H≥S)=?; P(H≤S)=?; P(H=S)=?;

№ вар.	А О Э	Б П Ю	В Р Я	Г С	Д Т	Е У	Ж З Ф	И Х	К Ц	Л Ч	М Ш	Н Щ
1	>10	=22	<10	=21	=12	>54	>56	=54	=56	<54	<56	=25
2	>11	=11	<11	=12	>48	=43	<48	=48	=18	>10	<10	=10
3	<15	=15	>15	=55	=45	>25	=25	<25	≥40	≤85	=80	>55
4	=45	>25	=25	<25	>90	<90	=40	>15	<15	=60	=15	=105
5	>5	=5	<15	=15	>15	=8	=11	=25	=39	>39	<39	>50
6	>20	<20	=30	=50	=80	=40	=90	>95	=95	<95	>100	<100
7	>25	<25	>35	<35	=100	=120	=150	=140	=90	>150	<150	>140
8	=25	<56	<54	=56	=54	>56	>54	=12	=21	<10	=22	>10
9	>140	<150	>150	=90	=140	=150	=120	=100	=35	>35	<25	>25
0	>50	<39	>39	=39	=25	=11	=8	>15	=15	<15	=5	>5

2.4 Методические указания к выполнению задания

На рис. 2.1 приведена схема питания приёмников от силовых шкафов по магистральной схеме.

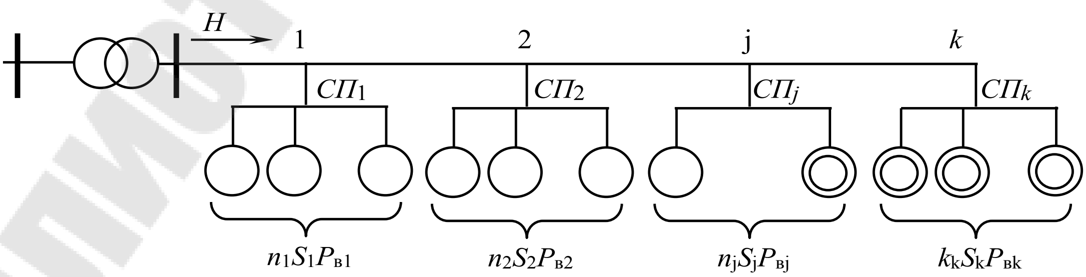


Рис. 2.1. Схема питания приемников от силовых шкафов по магистральной схеме

где $СП_j$ – j-ый силовой пункт, к которому присоединено n_j приёмников мощностью S_j каждый и вероятностью включения p_j, $j = 1, ..., k$.

Определить вероятность нагрузки (H) головного участка магистральной схемы.

Решение

Математически задача может быть записана следующим выражением: $P(H \geq S1)$, где $S1$ – численное значение. Нагрузка головного участка зависит от количества включенных приёмников в каждом СП. Численное значение $S1$ определяется $\sum\limits_{1}^{k} n_j \cdot S_j$, $j = 1, k$. Предполагается, что мощности приемников S_j однотипны. Вычисление вероятности выполним по выражению

$$P(H \geq S1) = \sum_{1}^{M} \prod_{1}^{K} C_{\mathrm{n}_j}^{\mathrm{m}_j} p_j^{\mathrm{m}_j} q_j^{\mathrm{n}_j - \mathrm{m}_j},$$

где m_j – количество включенных приемников в j-ом СП; М – количество случаев удовлетворяющих условию $H \geq S1$; K – количество групп электродвигателей (СП). Приведем численный пример.

Пусть $K = 3$; $n_1 = 3$; $S_1 = 10$ кВ·А; $p_1 = 0,7$;

$n_2 = 2$; $S_2 = 15$ кВ·А; $p_2 = 0,6$;

$n_3 = 4$; $S_3 = 5$ кВ·А; $p_3 = 0,5$.

Определить $P(H > 15\ \text{кВ} \cdot \text{А}) = ?$

Число случаев с нагрузкой более 15 кВ·А очень велико, поэтому выявляем все случаи, когда нагрузка головного участка ≤ 15 кВ·А.

$$\begin{aligned} P(H \leq 15) = {} & P(\mathrm{m}_1 = 0, \mathrm{m}_2 = 0, \mathrm{m}_3 = 0) + P(\mathrm{m}_1 = 1, \mathrm{m}_2 = 0, \mathrm{m}_3 = 0) + \\ & + P(\mathrm{m}_1 = 0, \mathrm{m}_2 = 0, \mathrm{m}_3 = 1) + P(\mathrm{m}_1 = 0, \mathrm{m}_2 = 1, \mathrm{m}_3 = 0) + \\ & + P(\mathrm{m}_1 = 1, \mathrm{m}_2 = 0, \mathrm{m}_3 = 1) + P(\mathrm{m}_1 = 0, \mathrm{m}_2 = 0, \mathrm{m}_3 = 2) + \\ & + P(\mathrm{m}_1 = 0, \mathrm{m}_2 = 0, \mathrm{m}_3 = 3) = \end{aligned}$$

$$\begin{aligned} & = C_3^0 p_1^0 q_1^3 \cdot C_2^0 p_2^0 q_2^2 \cdot C_4^0 p_3^0 q_3^4 + C_3^1 p_1^1 q_1^2 \cdot C_2^0 p_2^0 q_2^2 \cdot C_4^0 p_3^0 q_3^4 + \\ & + C_3^0 p_1^0 q_1^3 \cdot C_2^0 p_2^0 q_2^2 \cdot C_4^1 p_3^1 q_3^3 + C_3^0 p_1^0 q_1^3 \cdot C_2^1 p_2^1 q_2^1 \cdot C_4^0 p_3^0 q_3^4 + \\ & + C_3^1 p_1^1 q_1^2 \cdot C_2^0 p_2^0 q_2^2 \cdot C_4^1 p_3^1 q_3^3 + C_3^0 p_1^0 q_1^3 \cdot C_2^0 p_2^0 q_2^2 \cdot C_4^2 p_3^2 q_3^2 + \\ & + C_3^0 p_1^0 q_1^3 \cdot C_2^0 p_2^0 q_2^2 \cdot C_4^3 p_3^3 q_3^1 = \end{aligned}$$

$$\begin{aligned} & = 0,3^3 \cdot 0,4^2 \cdot 0,5^4 + 3 \cdot 0,7 \cdot 0,3^2 \cdot 0,4^2 \cdot 0,5^4 + 0,3^3 \cdot 0,4^2 \cdot 4 \cdot 0,5 \cdot 0,5^3 + \\ & + 0,3^3 \cdot 2 \cdot 0,6^1 \cdot 0,4^1 \cdot 0,5^4 + 3 \cdot 0,7 \cdot 0,3^2 \cdot 0,4^2 \cdot 4 \cdot 0,5 \cdot 0,5^3 + \\ & + 0,3^2 \cdot 0,4^2 \cdot 6 \cdot 0,5^2 \cdot 0,5^2 + 0,3^3 \cdot 0,4^2 \cdot 4 \cdot 0,5^3 \cdot 0,5 = 1,431 \cdot 10^{-2}. \end{aligned}$$

$$P(H > 15) = 1 - P(H \leq 15) = 1 - 0,01431 = 0,98569.$$

1.5 Контрольные вопросы

1. Что представляет собой последовательность независимых испытаний?

2. Поясните схему Бернулли.

3. При каких условиях формула Бернулли применима для определения вероятностей:
1) повреждения ровно m элементов из n;
2) повреждения элементов от m_1 до m_2 из n;
3) повреждения не более k элементов из n;
4) повреждения более k элементов из n;
5) того, что m элементов находятся в рабочем состоянии из n;
6) того, что все элементы находятся в рабочем состоянии.

Приведите соответствующие аналитические выражения.

4. На основании каких теорем теории вероятностей определяются вероятности сложных событий.

5. Приведите на диаграмме Венна одно из состояний (определяется преподавателем).

ЛАБОРАТОРНАЯ РАБОТА № 3
Случайные величины в энергетике. Числовые характеристики случайных величин. Показатели вариации

3.1 Цель работы

Изучить случайные величины – параметры в энергетике; представление их законов распределения. Определить числовые характеристики по статистическим данным.

3.2 Основные понятия

К случайным величинам в энергетике относятся такие важные параметры режима, как спрос электрической мощности и энергии, отклонение частоты и напряжения в электрических сетях от номинальных значений, располагаемая мощность электростанций, мощности агрегатов в аварийном ремонте, длительности безаварийной работы и аварийного ремонта отдельных агрегатов, напор на гидростанциях и т.д. Знание закономерностей изменения этих случайных величин необходимо как при проектировании, так и при эксплуатации энергетических систем.

Случайные величины можно разделить на два класса: дискретные и непрерывные. Дискретная случайная величина может принимать только дискретные (разрозненные) значения, например число агрегатов, вышедших аварийно из работы. Это число в ограниченном интервале является конечным. Значение непрерывных случайных величин могут изменяться непрерывно, т.е. даже в ограниченных интервалах такие величины могут иметь бесконечно большое число значений, например ошибка прогнозирования суммарного спроса мощности.

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями. Про случайную величину мы будем говорить, что она подчинена данному закону распределения.

Закон распределения для дискретных случайных величин может быть представлен рядом распределения (таблицей) или графически (многоугольником распределения). Для описания закона распределения непрерывных случайных величин чаще используется функция распределения случайной величины и гистограмма.

К числовым характеристикам случайных величин относятся: математическое ожидание, дисперсия, среднее квадратическое отклонение.

Математическое ожидание $M(x)$ для дискретной случайной величины определим по выражению:

$$M(x)=\sum_{k=1}^{n} x_i p_i ,$$

причем суммирование происходит по всем значениям дискретной величины x_i, имеющим вероятности p_i.

Аналогично для непрерывной случайной величины

$$M(x) = \int_{-\infty}^{+\infty} x\varphi(x)dx$$

где $\varphi(x)$ - плотность вероятности.

В качестве меры отклонения случайной величины от ее математического ожидания принимают величину, равную математическому ожиданию (м. о.) квадрата отклонения случайной величины от ее м. о., которую называют дисперсией случайной величины x и обозначают $D(x)$:

$$D(x) = M[x - M(x)]^2,$$

Определение дисперсии для дискретных случайных величин

$$D(x) = \sum_k \left[x_k - M(x) \right]^2 p_k,$$

где суммирование распространяется на все значения случайной величины $x_к$, имеющие соответствующие вероятности $p_к$.

Для непрерывных случайных величин дисперсия определяется

$$D(x) = \int_{-\infty}^{+\infty} \left[x_k - M(x) \right]^2 \varphi(x)dx,$$

Квадратный корень из величины дисперсии называется средним квадратическим отклонением или стандартным отклонением случайной величины.

$$\sigma(x) = \sqrt{D(x)}$$

Пример Пусть в энергосистеме возможны дефициты мощности 50, 100 и 150 МВт, причем вероятности этих дефицитов соответственно равны 0,001, 0,0004, 0,0002. Требуется определить м.о. недоотпуска энергии за год.

Решение

При постоянном дефиците 50 МВт недоотпуск электроэнергии за год составил бы $50 \cdot 8760$ МВт·ч, при дефиците 100 МВт - $100 \cdot 8760$ МВт·ч и т.д. Поэтому м.о. недоотпуска:

М (х)=50·8760·0,001+100·8760·0,0004+150·8760·0,0002=1051 МВт·ч.

Показатели вариации

Средние величины дают обобщающую характеристику совокупности по варьирующим признакам. Но наряду со средними величинами большое практическое и теоретическое значение имеет изучение отклонений от средних. Поэтому средние характеристики дополняют показателями измерениями отклонения от средних, т.е. показателями вариации признака.

Наиболее простым из этих показателей является показатель размаха вариации R. Его исчисляют как разность между наибольшим и наименьшим значениями варьирующего признака

$$R = x_{\max} - x_{\min}.$$

Относительной мерой вариации, позволяющей сравнивать степень варьирования признаков в вариационных рядах с разным уровнем средних, является коэффициент вариации, определяемый по выражению

$$\upsilon = \frac{\sigma}{x} \cdot 100,$$

где σ - среднее квадратическое отклонение;

$\bar{x}$ - среднее значение (математическое ожидание).

3.3 Задание

Имеются следующие данные 10% выборочного обследования рабочих электротехнического цеха (см. таблицу 3.1)

Графа А – тарифный разряд рабочего;

Графа Б – производственный стаж работы, полных лет;

Графа В – заработная плата, тыс. ден. ед.

Таблица 3.1

Исходные данные по вариантам

№ варианта	1			2			3			4		
№ рабочего	А	Б	В	А	Б	В	А	Б	В	А	Б	В
1	3	5	209	2	5	174	3	3	180	2	1	140
2	1	1	131	1	0	126	4	18	246	1	19	299
3	4	7	282	4	8	223	4	11	238	6	1	120
4	2	7	161	4	12	196	5	9	260	2	3	148
5	1	1	195	2	4	181	2	3	130	4	8	225
6	2	5	178	3	7	194	1	1	143	5	15	254
7	3	8	204	3	6	180	2	4	152	2	1	152
8	5	10	256	1	1	112	3	8	222	4	5	212
9	2	0	146	3	8	187	4	7	220	2	6	182
10	3	7	190	1	0	116	1	2	152	2	4	164
11	2	2	172	4	3	229	3	8	240	4	7	220
12	2	3	186	2	4	177	1	0	101	3	9	190
13	4	5	230	1	0	128	2	4	160	4	10	230
14	1	1	121	2	3	162	3	5	208	2	3	159
15	2	4	164	3	3	160	3	9	200	3	5	201
16	3	8	223	5	17	267	4	16	230	3	6	184
17	1	3	169	2	1	159	3	3	217	4	10	220

18	2	1	145	6	21	301	2	1	128	5	21	302
19	1	0	135	3	4	192	3	7	175	2	1	149
20	4	6	246	3	11	207	3	4	183	3	6	180
21	3	5	202	3	3	191	1	1	141	2	4	176
22	2	1	149	4	12	250	2	1	102	3	5	203
23	2	3	175	3	2	188	4	3	170	3	2	152
24	6	19	288	5	6	204	3	9	221	4	15	229
25	2	1	148	5	8	216	6	12	281	5	9	260
26	4	9	243	3	3	184	5	12	272	2	2	129
27	1	0	120	6	29	257	2	2	136	4	2	215
28	3	5	208	2	3	147	3	6	184	3	4	183
29	3	7	203	4	13	224	1	0	99	3	7	201
30	1	1	151	3	8	195	4	14	199	4	10	237

Продолжение таблицы

№ варианта	5			6			7			8		
№ рабочего	А	Б	В	А	Б	В	А	Б	В	А	Б	В
1	2	7	158	1	1	115	3	4	183	3	8	215
2	4	12	200	1	0	102	6	18	270	5	12	261
3	2	5	163	6	11	298	2	3	142	1	3	142
4	3	8	202	3	4	196	4	9	200	3	6	199
5	4	6	247	5	10	229	5	8	231	2	6	164
6	2	3	175	3	4	194	3	4	190	4	3	181
7	2	4	180	3	10	162	3	6	184	5	8	227
8	5	9	271	2	2	160	2	4	149	2	7	170
9	1	1	112	4	5	238	1	2	136	3	8	200
10	4	9	186	3	7	171	1	0	104	4	6	205
11	3	6	173	4	9	207	2	3	168	1	2	112
12	2	5	144	5	9	223	4	5	186	3	4	172
13	1	0	103	5	15	243	5	6	208	2	2	149
14	1	2	114	3	8	185	4	4	165	4	4	184
15	6	18	306	2	4	147	6	23	300	6	12	251
16	3	6	179	1	1	119	3	5	179	3	5	180
17	2	4	157	2	3	139	2	4	163	1	0	110
18	4	16	210	3	4	177	1	0	112	2	2	131
19	5	15	275	3	8	182	4	10	220	3	9	204
20	4	3	182	4	11	209	4	5	204	4	10	233
21	2	4	159	4	11	219	5	17	267	1	3	138
22	2	4	168	2	1	157	3	5	192	2	4	153
23	3	11	228	3	9	200	3	7	191	3	3	187
24	3	5	181	4	10	228	1	1	144	4	6	207

25	4	7	211	3	3	199	1	2	150	2	4	157
26	5	15	292	5	21	295	2	3	171	4	3	198
27	3	4	171	3	8	181	3	8	205	3	5	131
28	3	8	233	2	0	131	4	10	211	2	6	173
29	3	9	186	5	13	252	3	9	215	6	11	285
30	2	4	158	2	2	163	3	7	206	5	6	240

Продолжение таблицы

№ варианта	9			10			11			12		
№ рабочего	А	Б	В	А	Б	В	А	Б	В	А	Б	В
1	6	21	319	5	7	501	6	16	612	1	2	120
2	5	12	297	4	3	457	3	4	315	4	7	485
3	3	9	211	4	2	450	3	5	325	3	7	297
4	2	1	207	3	4	375	2	2	295	6	19	698
5	4	6	251	2	3	302	1	2	195	5	10	575
6	6	19	372	2	3	305	4	8	435	4	8	490
7	1	1	176	5	16	505	2	5	297	1	3	125
8	1	1	178	6	12	508	1	4	180	2	4	205
9	2	2	203	1	3	207	5	12	512	3	5	285
10	3	7	207	3	5	315	2	4	270	5	11	578
11	6	17	384	4	6	401	3	9	350	6	18	698
12	5	10	305	4	7	405	4	8	438	2	5	218
13	4	5	312	3	2	310	4	7	430	3	5	290
14	4	3	302	2	1	212	5	12	510	1	4	130
15	1	2	180	2	1	217	3	5	330	1	3	135
16	1	2	185	5	8	495	1	4	185	3	4	287
17	3	7	176	1	2	200	6	9	605	3	3	280
18	2	3	161	1	2	205	1	3	175	4	9	497
19	5	11	309	6	20	512	1	2	170	6	17	670
20	1	2	150	4	7	380	3	9	355	5	12	580
21	5	11	311	4	8	385	4	10	470	1	2	137
22	4	5	295	3	6	308	4	11	475	2	2	215
23	2	3	151	1	3	201	5	13	515	2	4	270
24	2	4	157	6	18	505	2	2	215	6	15	612
25	3	6	185	2	5	206	1	2	190	1	4	140
26	1	2	132	3	5	302	3	4	302	4	12	502
27	1	3	135	1	4	203	2	3	195	3	5	315
28	3	6	187	2	3	201	1	2	173	2	6	290
29	3	5	190	2	2	197	4	12	472	1	7	150
30	4	4	195	3	8	307	5	15	525	5	15	595

Продолжение таблицы

№ варианта	13			14			15			16		
№ рабочего	А	Б	В	А	Б	В	А	Б	В	А	Б	В
1	1	4	163	2	4	187	4	5	401	3	5	304
2	4	7	397	3	4	300	3	3	305	2	4	187
3	6	17	705	5	11	500	3	4	308	1	2	120
4	5	10	518	6	12	812	5	6	501	2	3	180
5	3	7	301	4	10	405	1	1	120	1	3	122
6	2	4	276	4	9	400	2	2	201	1	4	128
7	2	5	278	3	4	302	2	3	205	2	5	190
8	1	3	160	2	4	189	5	8	505	2	6	195
9	1	2	158	2	3	180	4	5	405	5	10	515
10	4	8	399	5	10	502	3	4	309	5	11	518
11	6	16	700	4	8	398	4	4	400	4	5	436
12	5	9	512	4	9	402	3	2	300	3	4	301
13	3	7	308	1	2	201	2	1	200	2	4	185
14	3	6	300	3	5	304	1	0	115	1	2	118
15	2	4	285	3	6	312	6	10	705	4	3	420
16	1	2	150	2	4	182	6	12	750	3	5	308
17	6	15	701	4	10	408	4	6	408	2	4	186
18	4	3	351	5	9	497	5	14	550	2	3	181
19	5	7	521	6	12	815	6	15	740	1	3	115
20	5	6	517	5	7	490	1	2	125	6	1	805
21	3	5	302	4	9	410	1	2	127	5	12	520
22	2	5	270	3	7	318	2	3	206	5	13	527
23	4	8	390	3	6	314	1	0	115	6	18	809
24	1	3	158	2	4	181	2	4	207	1	4	116
25	3	4	301	2	5	185	3	5	307	2	5	188
26	2	4	250	3	4	300	4	6	406	2	4	181
27	4	6	370	2	5	184	5	13	546	3	2	302
28	5	6	515	1	3	174	6	18	781	1	6	120
29	1	5	170	5	6	480	1	2	125	4	5	401
30	2	6	272	6	19	850	2	2	197	4	6	470

Продолжение таблицы

№ варианта	17			18			19			20		
№ рабочего	А	Б	В	А	Б	В	А	Б	В	А	Б	В
1	5	6	550	4	4	410	1	4	135	2	3	210
2	6	6	612	4	5	420	2	4	241	4	4	470
3	1	4	140	6	10	600	2	3	240	6	11	780

4	3	2	300	5	11	520	3	4	315	4	5	480
5	2	3	220	3	4	370	5	6	515	4	4	472
6	4	7	451	2	2	237	6	7	701	2	3	215
7	2	4	235	2	3	241	1	2	130	3	4	350
8	2	3	227	1	0	122	6	8	690	5	10	588
9	3	5	312	3	4	365	5	7	520	5	6	555
10	2	6	240	6	10	602	4	4	407	1	2	110
11	6	14	641	4	5	426	3	5	318	3	4	352
12	4	6	445	5	10	510	4	5	410	2	5	218
13	3	4	307	5	12	530	4	6	416	2	6	220
14	2	4	220	3	7	380	1	2	132	6	8	470
15	1	3	137	2	5	242	1	3	133	1	0	100
16	1	6	156	4	5	424	3	2	310	1	1	108
17	3	5	310	6	18	620	2	2	230	4	5	484
18	3	4	308	4	6	430	1	0	120	3	3	350
19	2	4	218	3	8	385	1	4	140	2	2	206
20	1	2	124	3	6	377	5	5	510	4	6	487
21	6	15	618	2	3	240	3	4	312	5	7	561
22	4	3	437	1	2	130	4	7	418	6	9	722
23	6	18	660	5	17	550	2	3	232	3	4	360
24	1	1	118	4	5	422	5	8	530	3	5	370
25	2	2	215	6	9	615	1	0	115	1	2	108
26	3	4	306	4	12	450	2	4	240	1	0	101
27	5	16	502	5	14	540	4	6	410	2	2	204
28	5	17	508	6	21	640	3	5	318	2	3	205
29	6	19	665	4	10	444	2	5	236	5	8	575
30	1	0	105	1	0	120	1	2	120	5	9	580

Требуется:

1. Построить ряды распределения рабочих цеха:

1)по квалификации (разрядам);

2) по размеру месячной заработной платы, выделив семь групп с равными интервалами;

3) по общему стажу работы, выделив группы со следующими специализированными интервалами: а) менее 1 года; б) от 1 до 2 лет; в) от 3 до 5 лет; г) от 6 до 10 лет; д) от 11 лет и выше.

Результаты представить в виде таблиц.

2. По результатам группировок вычислить: а) математическое ожидание тарифного разряда и заработной платы рабочих; б) общую дисперсию заработной платы; в) среднее квадратическое отклонение и коэффициент вариации заработной платы рабочих.

Выбор варианта осуществляется по последней цифре зачетной книжки студента или определяется преподавателем.

3.4 Методические указания к выполнению задания

Пример Для заданного вариационного ряда вычислить математическое ожидание дисперсию, среднее квадратическое отклонение и коэффициент вариации.

Решение

Решение целесообразно представить в виде таблицы.

Таблица 3.2

Вычисление математического ожидания, дисперсии, среднего квадратического отклонения и коэффициента вариации

Группа рабочих по размеру месячной платы, тыс.ден.ед.	Варианты (х)	Число рабочих (f)	$(x-\bar{x})$	$(x-\bar{x})^2$	$(x-\bar{x})^2 \cdot f$
130-140	135	10	-30,8	948,64	9486,4
140-150	145	50	-20,8	432,64	21632
150-160	155	100	-10,8	116,64	11664
160-170	165	115	-0,8	0,64	73,6
170-180	175	180	9,2	84,64	15235,2
180-190	185	45	19,2	386,64	16588,8
Сумма	-	500	-	-	74680

Математическое ожидание вычислялось по формуле:

$$\bar{x} = \frac{\sum xf}{\sum f} = 165,8\,руб.$$

Дисперсия:

$$\sigma^2 = \frac{\sum (x-\bar{x})^2 f}{\sum f} = \frac{74680}{500} = 149,36\,руб.$$

Среднее квадратическое отклонение:

$$\sigma = \sqrt{149,36} = 12,22\,руб.$$

Коэффициент вариации:

$$\vartheta = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{12,22}{165,8} \cdot 100 = 7,4\%$$

Среднее арифметическое вычисляем по выражению:

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}.$$

3.5 Контрольные вопросы

1. Какие параметры режима относятся к случайным величинам?

2. Какие случайные величины (параметры) относятся к дискретным, непрерывным.

3. Законом распределения случайных величин называется …. ?

4. Представление законов распределения дискретных и непрерывных случайных величин.

5. Числовые характеристики дискретных и непрерывных случайных величин и их математическое определение.

6. Укажите условие, при котором математическое ожидание и среднее арифметическое равны между собой; причины их расхождения

7. Укажите числовые характеристики случайных величин, характеризующие положение на числовой оси.

8. Укажите характеристики случайных величин, описывающих то или иное свойство распределения.

9. Дайте определение начальных и центральных моментов *s*-го порядка.

ЛАБОРАТОРНАЯ РАБОТА № 4
Определение характеристик эргодической стационарной случайной функции по одной реализации

4.1 Цель работы

Обработать исходные данные стационарной случайной функции и определить ее характеристики.

4.2 Краткие теоретические сведения

Для эргодической стационарной случайной функции одна реализация достаточно большой продолжительности практически эквивалентна (в смысле объема сведений о случайной функции) множеству реализаций той же, общей продолжительности; характеристики случайной функции могут быть приближенно определены не как средние по множеству наблюдений, а *как средние по времени t.* В частности, при достаточно большом T математическое ожидание m_x может быть приближенно вычислено по формуле

$$m_x \approx \frac{1}{T}\int_0^T x(t)dt.$$

Аналогично может быть приближенно найдена корреляционная функция $k_x(\tau)$ при любом τ. Действительно, корреляционная функция, по определению, представляет собой не что иное, как *математическое ожидание* случайной функции $\overset{\circ}{x}(t)\cdot \overset{\circ}{x}(t+\tau)$:

$$k_x(\tau) = M\left[\overset{\circ}{x}(t)\cdot \overset{\circ}{x}(t+\tau)\right].$$

Это математическое ожидание также, очевидно, может быть приближенно вычислено как среднее по времени.

Фиксируем некоторое значение τ и вычислим указанным способом корреляционную функцию $k_x(\tau)$. Для этого удобно предварительно «центрировать» данную реализацию $x(t)$, т. е. вычесть из нее математическое ожидание (4.1):

$$\overset{\circ}{x}(t) = x(t) - m_x. \qquad (4.1)$$

Вычислим при заданном τ математическое ожидание случайной функции $\overset{\circ}{x}(t)\cdot \overset{\circ}{x}(t+\tau)$ как среднее по времени. При этом, очевидно, нам придется учитывать не весь участок времени от 0 до T, а несколько меньший, т. к. второй сомножитель $\overset{\circ}{x}(t+\tau)$ известен нам не для всех

t, а только для тех, для которых $t + \tau \le T$. Вычисляя среднее по времени указанным выше способом, получим:

$$k_x(\tau) \approx \frac{1}{T-\tau} \int\limits_0^{T-\tau} \overset{\circ}{x}(t)\overset{\circ}{x}(t+\tau)dt. \qquad (4.2)$$

Вычислив интеграл (4.2) для ряда значений τ, можно приближенно воспроизвести по точкам весь ход корреляционной функции.

На практике обычно интегралы (4.1) и (4.2) заменяют конечными суммами. Покажем как это делается. Разобьем интервал записи случайной функции на n равных частей длиной Δt и обозначим середины полученных участков $t_1,\ t_2, ..., t_n$ (рис. 4.1).

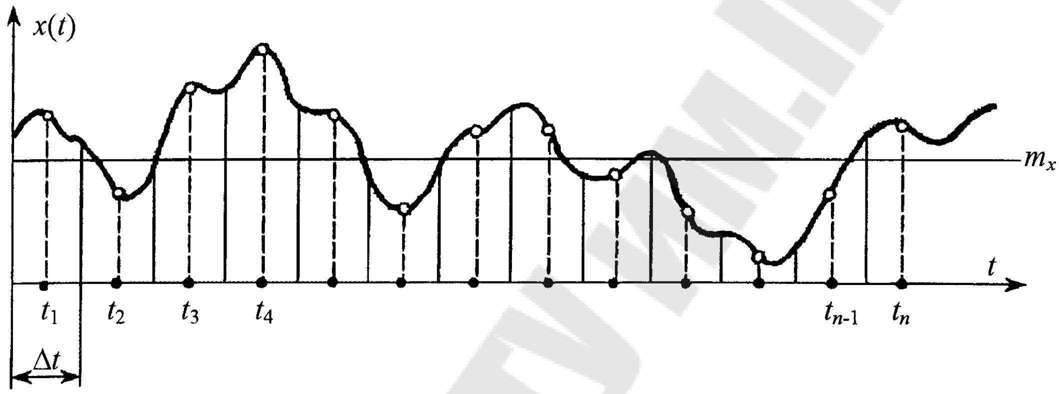


Рис. 4.1

Представим интеграл (6.7) как сумму интегралов по элементарным участкам Δt и на каждом из них вынесем функцию $x(t)$ из-под знака интеграла средним значением, соответствующим центру интервала $x(t_i)$. Получим приближенно:

$$m_x \approx \frac{1}{T} \cdot \frac{T}{n} \sum_{i=1}^{n} x(t_i),$$

или

$$m_x \approx \frac{1}{n} \cdot \sum_{i=1}^{n} x(t_i). \qquad (4.3)$$

Аналогично можно вычислить корреляционную функцию для значений τ, равных $0, \Delta t, 2\Delta t, ...$. Придадим, например, величине τ значение

$$\tau = m\Delta t = \frac{mT}{n}$$

и вычислим интеграл (4.2), деля интервал интегрирования

$$T - \tau = T - \frac{mT}{n} = \frac{n-m}{n} T$$

на $n-m$ равных участков длиной Δt и вынося на каждом из них функцию $\overset{\circ}{x}(t)\overset{\circ}{x}(t+\tau)$ за знак интеграла средним значением. Получим:

$$k_x\left(\frac{mT}{n}\right) \approx \frac{n}{(n-m)T} \cdot \frac{T}{n} \cdot \sum_{i=1}^{n-m} \overset{\circ}{x}(t_i) \cdot \overset{\circ}{x}(t_{i+m}),$$

или окончательно

$$k_x\left(\frac{mT}{n}\right) \approx \frac{1}{(n-m)} \sum_{i=1}^{n-m} \overset{\circ}{x}(t_i) \cdot \overset{\circ}{x}(t_{i+m}). \tag{4.4}$$

Вычисление корреляционной функции по формуле (4.4) производят для $m = 0, 1, 2, \ldots$ последовательно вплоть до таких значений *m,* при которых корреляционная функция становится практически равной нулю или начинает совершать небольшие нерегулярные колебания около нуля. Общий ход функции $k_x(t)$ воспроизводится по отдельным точкам (рис. 4.2).

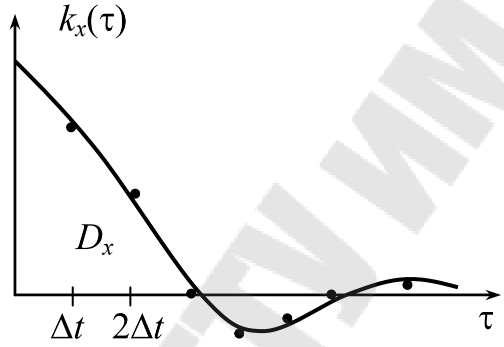


Рис. 4.2

Для того чтобы математическое ожидание m_x и корреляционная функция $k_x(\tau)$ были определены с удовлетворительной точностью нужно, чтобы число точек *n* было достаточно велико (порядка сотни, а в некоторых случаях даже нескольких сотен). Выбор длины элементарного участка Δt определяется характером изменения случайной функции. Если случайная функция изменяется сравнительно плавно, участок Δt можно выбирать большим, чем когда она совершает резкие и частые колебания. Чем более высокочастотный состав имеют колебания, образующие случайную функцию, тем чаще нужно располагать опорные точки при обработке. Ориентировочно можно рекомендовать выбирать элементарный участок Δt так, чтобы на полный период самой высокочастотной гармоники в составе случайной функции приходилось порядка 5–10 опорных точек.

Часто выбор опорных точек вообще не зависит от обрабатывающего, а диктуется темпом работы записывающей аппаратуры. В этом случае следует вести обработку непосредственно полученного из опыта материала, не пытаясь вставить между наблюденными значениями промежуточные, т. к. это не может повысить точности результата, а излишне осложнит обработку.

4.3 Задание

Для заданного исходного статистического материала, представляющего одну реализацию эргодической стационарной функции:

1) Представить график исходной реализации.
2) Определить характеристики случайной функции: m_x; D_x ; $k_x(\tau)$ и $\rho_x(\tau)$.
3) Представить график нормированной корреляционной функции $\rho_x(\tau)$.

Исходные данные задаются преподавателем из таблицы 4.1:

Таблица 4.1

Исходные данные

t, сек	X(t)	t, сек	X(t)	t, сек	X(t)	t, сек	X(t)
0	1	50	1	100	1,2	150	0,8
2	1,3	52	1,1	102	1,4	152	0,6
4	1,1	54	1,5	104	0,8	154	0,9
6	0,7	56	1	106	0,9	156	1,2
8	0,7	58	0,8	108	1	158	1,3
10	1,1	60	1,1	110	0,8	160	0,9
12	1,3	62	1,1	112	0,8	162	1,3
14	0,8	64	1,2	114	1,4	164	1,5
16	0,8	66	1	116	1,6	166	1,2
18	0,4	68	0,8	118	1,7	168	1,4
20	0,3	70	0,8	120	1,3	170	1,4
22	0,3	72	1,2	122	1,6	172	0,8
24	0,6	74	0,7	124	0,8	174	0,8
26	0,3	76	0,7	126	1,2	176	1,3
28	0,5	78	1,1	128	0,6	178	1
30	0,5	80	1,5	130	1	180	0,7
32	0,7	82	1	132	0,6	182	1,1
34	0,8	84	0,6	134	0,8	184	0,9
36	0,6	86	0,9	136	0,7	186	0,9
38	1	88	0,8	138	0,9	188	1,1
40	0,5	90	0,8	140	1,3	190	1,2
42	1	92	0,9	142	1,5	192	1,3
44	0,9	94	0,9	144	1,1	194	1,3
46	1,4	96	0,6	146	0,7	196	1,6
48	1,4	98	0,4	148	1	198	1,5

4.4 Методические указания к выполнению задания

Определить характеристики случайной функции: m_x; D_x и $k_x(\tau)$ по следующим выражениям: $m_x = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$; $D_x = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\overset{\circ}{x}(t_i) \right)^2$, где $\overset{\circ}{x}(t_i) = x(t) - m_x$ и $k_x(\tau) = \frac{1}{n-m} \cdot \sum_{i=1}^{n-m} \overset{\circ}{x}(t_i) \cdot \overset{\circ}{x}(t_{i+m})$.

Определить нормированную корреляционную функцию $\rho_x = \frac{k_x(\tau)}{D_x}$.

Все вычисления выполнить, используя программный пакет Excel
Выполнить графические построения.

4.5 Контрольные вопросы:

1. Дайте определение стационарной случайной функции.

2. В чем состоит эргодическое свойство стационарной случайной функции?

3. Какие вероятностные характеристики стационарной случайной функции необходимо определить?

4. Как определяется элемент в расчетной таблице если $\kappa=0\div10$?

5. К чему стремится нормированная корреляционная функция ρ_N ·если $\kappa\to\infty$?

ЛАБОРАТОРНАЯ РАБОТА № 5
Статистические критерии и их применение

5.1 Цель работы

Изучить различные статистические критерии, выполнить расчеты по определению критериев и сделать правильные выводы.

5.2 Краткие теоретические сведения

Критерии согласия применяют с целью проверки правдоподобия гипотез о согласованности теоретического и статистического распределения. Чтобы принять или опровергнуть предполагаемую гипотезу, определяют некоторые параметры и сравнивают их с критериями, соответствующими значениями (квантилями) определенного закона распределения.

Распределение Стьюдента (t - критерий)

Распределение Стьюдента дает возможность находить генеральное среднее или проверять статистические гипотезы при очень малых выборках. Распределение Стьюдента определяется по формуле:

$$T = \frac{\bar{x} - a}{S}\sqrt{N},$$

где $\bar{x}$ - случайная величина, определенная на выборке объемом N;

a - математическое ожидание случайной величины;

S - среднее квадратическое отклонение, определенное на основе выборочной дисперсии S^2;

N – объем выборки.

Очевидно, что распределение T должно существенно зависеть от объема выборки или числа степеней свободы f=N-1, с которым определена S^2. При больших N S^2 приближается к σ^2. Распределение T симметрично относительно начала координат, т.е. $t_{q/2} - t_{1-q/2}$. Квантили распределения Стьюдента для q = 0,05 приведены в таблице х; при $q \neq 0{,}05$ следует пользоваться более полными таблицами.

Таблица 5.1

Квантили распределения Стьюдента при q = 0,05

f=N-1	1	2	3	4	5	6	7	8
$t_{1-q/2}$	12,71	4,30	3,18	2,78	2,57	2,45	2,37	2,31

Продолжение табл. 5.1

f=N-1	9	10	20	30	∞	-
$t_{1-q/2}$	2,26	2,23	2,09	2,04	1,96	-

Распределение Пирсона (χ^2-критерий)

Применяется для оценки генеральной дисперсии σ^2 по выборочной S^2. Пирсон в 1900 г. ввел случайную величину

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{x_i - \bar{x}}{\sigma} \right)^2$$

и нашел ее распределение, зависящее лишь от f=N-1. Оно несимметрично, следовательно, $\chi^2_{q/2} \neq -\chi^2_{1-q/2}$; некоторые квантили приведены в таблице 5.2; более полные таблицы содержатся в [2].

Таблица 5.2

Квантили распределения Пирсона χ^2_{1-q}

f	1	2	3	4	5	6	7	8
$\chi^2_{1-0,95}$	0,0039	0,103	0,352	0,71	1,14	1,63	2,17	2,73
$\chi^2_{1-0,05}$	3,8	6,0	7,8	9,5	11,1	12,6	14,1	15,5

Продолжение табл. 5.2

f	9	10	15	20	-
$\chi^2_{1-0,95}$	3,32	3,94	7,3	10,9	-
$\chi^2_{1-0,05}$	16,9	18,3	25,0	31,4	-

Связь между S^2 и χ^2 определяется выражением

$$S^2 = \frac{\sigma^2 \chi^2}{N-1} = \frac{\sigma^2 \chi^2}{f},$$

которое позволяет решать статистические задачи о дисперсиях.

Распределение Фишера (F- критерий)

Используется для сравнения двух выборочных дисперсий S_1^2 и S_2^2 (обычно принимают $S_1^2 > S_2^2$ и используют односторонние оценки), найденных соответственно с f_1 и f_2, с целью установить иди отвергнуть их принадлежность одной генеральной совокупности. Фишером введена случайная величина

$$F = S_1^2 \big/ S_2^2$$

и построено ее распределение, зависящее от f_1 и f_2. Некоторые квантили распределения Фишера для q = 0,05 приведены в таблице 5.3.

Таблица 5.3

Квантили распределения Фишера $F_{1-0,05}$

f_2	f_1							
	1	2	3	4	5	6	12	24
1	164,4	199,5	215,7	224,6	230,2	234,0	244,9	249,0
2	18,5	19,2	19,2	19,3	19,3	19,3	19,4	19,5
3	10,1	9,6	9,3	9,1	9,0	8,9	8,7	8,6
4	7,7	6,9	6,6	6,4	6,3	6,2	5,9	5,8
5	6,6	5,8	5,4	5,2	5,1	5,0	4,7	4,5
6	6,0	5,1	4,8	4,5	4,4	4,3	4,0	3,8
7	5,6	4,7	4,4	4,1	4,0	3,9	3,6	3,4
8	5,3	4,5	4,1	3,8	3,7	3,6	3,3	3,1
9	5,1	4,3	3,9	3,6	3,5	3,4	3,1	2,9
10	5,0	4,1	3,7	3,5	3,3	3,2	2,9	2,7
15	4,5	3,7	3,3	3,1	2,9	2,8	2,5	2,3
20	4,4	3,5	3,1	2,9	2,7	2,6	2,3	2,1
60	4,0	3,2	2,8	2,5	2,4	2,3	1,9	1,7

Распределение Кохрена (G - критерий)

Используется для проверки однородности k выборочных дисперсий, найденных с одинаковыми числами степеней свободы $f_j = N$-1. Кохреном введена случайная величина

$$G = \frac{S^2_{j,\max}}{\sum\limits_{j=1}^{k} S_j^2},$$

где $S^2_{j,\max}$ наибольшая из сравниваемых дисперсий.

Распределение G зависит от f_j, и k. В таблице X_4 приведены некоторые квантили G_{1-q} для q = 0,05.

Таблица 5.4

Квантили распределения Кохрена $G_{1-0,05} \times 10^{-2}$

k	f_j										
	1	2	3	4	5	6	7	8	9	10	∞
2	100	98	94	91	88	85	83	82	80	79	50
3	97	87	80	75	71	68	65	63	62	60	33
4	91	77	68	63	59	56	52	50	49	49	25
5	84	68	60	54	51	48	46	44	42	41	20
6	78	62	53	48	44	42	40	38	37	36	17
7	73	56	48	43	40	37	35	34	33	32	14
8	68	52	44	39	36	34	32	30	29	28	13

9	64	48	40	36	33	31	29	28	27	26	11
10	60	45	37	33	30	28	27	25	24	24	10
20	39	27	22	19	17	16	15	14	14	13	5
60	17	11	9	8	7	6	6	6	5	5	1,7
∞	0	0	0	0	0	0	0	0	0	0	0

Отметим, что наилучшей оценкой k однородных дисперсий служит дисперсия, определяемая как

$$S^2 = \frac{\sum_{j=1}^{k} S_j^2}{k}$$

с числом степеней свободы $f = kf_j = k(N\text{-}1)$. Эта оценка может использоваться для определения доверительного интервала для σ^2.

τ - распределение (τ - критерий)

Используется для проверки однородности наблюдений, исключения грубых ошибок или выбросов.

Квантили распределения случайной величины

$$\tau = \frac{\left| x_{\text{кр}} - \bar{x} \right|}{S},$$

зависящие лишь от объема выборки N, по которой определяются $\bar{x}$ и S ($x_{\text{кр}}$ - крайний элемент выборки, наименьший по значению), приведены для q = 0,05 в таблице 5.5.

Таблица 5.5

Квантили распределения при q = 0,05

N	3	4	5	6	7	8	9	10	15	20
$\tau_{1\text{-}q}$	1,41	1,69	1,87	2,0	2,09	2,17	2,24	2,29	2,29	2,62

Все рассмотренные распределения и построенные на их основе критерии предполагают нормальность закона распределения случайной величины. Проверка нормальности при $N > 20$ осуществляется по критерию А.Н. Колмогорова. При малых выборках могут применяться критерии согласия, основанные на сравнении выборочных ассиметрии и эксцесса с их дисперсиями.

5.3 Задание

(Номер варианта задания задается преподавателем из таблиц, приведенных ниже).

№ 1

На основании критерия Стьюдента проверить гипотезу, состоящую в том, что нормально распределенная случайная величина X имеет $M(X) = 10$ на основании результатов двух испытаний x_1 и x_2.

№ вар.	0	1	2	3	4	5	6	7	8	9
x_1	4,5	4,7	5,3	6,2	6,9	7,3	8,1	8,3	8,7	9
x_2	4,9	5,1	5,7	6,7	7,7	8	8,4	8,9	9,3	9,9

№ вар.	10	11	12	13	14	15	16	17	18	19
x_1	4,6	4,8	6,6	5,5	8,8	10,1	12,1	7,2	8,3	9,1
x_2	5,1	5,3	7,1	5,9	9,2	10,5	12,5	7,5	8,9	9,7

№ 2

Пользуясь критерием Пирсона найти с вероятностью $p = 0{,}95$ доверительный интервал для σ^2 нормально распределенной случайной величины, если при f, S^2.

№ вар.	0	1	2	3	4	5	6	7	8	9
f	1	2	3	4	5	6	7	8	9	15
S^2	2	3	4	5	6	7	8	9	10	1
№ вар.	10	11	12	13	14	15	16	17	18	19
f	13	14	16	17	18	19	21	22	23	24
S^2	8	8	9	9	10	10	11	11	12	12

№ 3

Пользуясь критерием Фишера проверить гипотезу об однородности двух выборочных дисперсий нормально распределенной случайной величины. Результаты опытов: S_1^2 при f_1 и S_2^2 при f_2.

№ вар.	0	1	2	3	4	5	6	7	8	9
S_1^2	1,4	0,4	0,5	0,6	0,7	0,8	0,9	1	1,1	1,2
f_1	24	1	2	3	4	5	6	1	1	12
S_2^2	0,2	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,8	0,2
f_2	60	5	6	7	8	9	9	10	15	20

№ вар.	10	11	12	13	14	15	16	17	18	19
S_1^2	1,1	1,1	1,2	1,2	4,3	1,3	1,3	2,8	1,4	1,4
f_1	24	6	4	2	1	5	12	4	5	3
S_2^2	0,7	0,8	0,8	0,9	0,9	1,0	1,0	1,1	1,1	1,1
f_2	10	8	7	4	3	60	7	3	9	5

№ 4

На основании критерия Кохрена проверить гипотезу об однородности выборочных дисперсий S_1^2; S_2^2; S_3^2; S_4^2; S_5^2, каждая из которых определена с вероятностью f, на уровне $q = 0{,}05$.

№ вар.	0	1	2	3	4	5	6	7	8	9
S_1^2	3	4	5	6	7	7	8	8	3	3
S_2^2	5	6	6	8	9	7	4	4	5	6
S_3^2	15	15	12	10	8	8	10	12	16	15
S_4^2	2	2	3	4	6	6	6	11	13	14
S_5^2	4	4	5	6	7	8	9	13	11	12
f	1	2	3	4	5	6	7	8	9	10
№ вар.	10	11	12	13	14	15	16	17	18	19
S_1^2	6	6	6	6	6	6	6	6	7	7
S_2^2	5	5	5	5	5	5	4	4	4	4
S_3^2	12	12	12	12	13	13	13	13	13	13
S_4^2	14	16	17	19	20	21	23	24	26	27
S_5^2	14	15	16	17	18	19	20	21	22	23
f	1	2	3	4	5	6	7	8	9	10

№ 5

На одном из пяти одинаковых агрегатов, выполняющих однотипные технологические операции, были внедрены мероприятия по экономии электроэнергии. На основе τ-критерия определить их эффективность, если зарегистрированное месячное потребление энергии каждым агрегатом составляет x_1; x_2; x_3; x_4; x_5.

№ вар.	0	1	2	3	4	5	6	7	8	9
x_1	10	11	12	13	14	15	16	18	19	20
x_2	12	13	14	14	15	17	18	19	20	23
x_3	13	15	11	12	15	12	14	15	13	17
x_4	14	12	13	15	15	11	16	11	12	13
x_5	11	13	15	17	16	19	13	19	15	19
№ вар.	10	11	12	13	14	15	16	17	18	19
x_1	21	22	23	24	26	27	28	29	30	31
x_2	23	24	25	26	27	28	30	31	32	33
x_3	16	16	17	17	18	18	19	19	20	20
x_4	14	14	14	14	14	14	14	15	15	15
x_5	19	20	20	21	21	22	22	23	24	24

5.4 Методические указания к выполнению заданий

Пример решения задания № 1: Известны три значения нормально распределенной случайной величины X: $x_1 = 10{,}0$; $x_2 = 9{,}5$;. $x_3 = 10{,}2$. Требуется определить генеральное среднее с $p = 0{,}95$.

<u>*Решение*</u>

Определим $\bar{x}$ по формуле

$$\bar{x} = \sum_{i=1}^{3} x_i / 3 = 9{,}9 .$$

Найдем S^2 по формуле с $f = N$-1 = 2

$$S^2 = \left[\frac{\sum_{i=1}^{3} x_i^2}{N} - \left(\frac{\sum_{i=1}^{3} x_i}{N} \right)^2 \right] = 0{,}08 .$$

Запишем выражение для T в соответствии с критерием Стьюдента

$$T = \frac{\bar{x} - a}{S}\sqrt{N} = \frac{9{,}9 - a}{0{,}28}\sqrt{3} .$$

По таблице 5.1 определим квантильные границы
$t_{1-q/2} = 4{,}30$, $t_{q/2} = t_{1-q/2} = -4{,}30$ и запишем неравенство

$$-4{,}30 < \frac{9{,}9 - a}{0{,}28}\sqrt{3} < 4{,}30.$$

Решив неравенство относительно a, получим с доверительной вероятностью $p = 0{,}95$

$$9,2 < a < 10,6.$$

Пример решения задания № 5: По величине месячного потребления электроэнергии определяется агрегат, на котором внедрены энергосберегающие мероприятия $x_{\text{кр}} = x_{\text{i}}$.

Определим S по следующей формуле:

$$S^2 = \left[\frac{\sum_{i=1}^{N} x_i^2}{N} - \left(\frac{\sum_{i=1}^{N} x_i}{N} \right)^2 \right].$$

Определяется среднее месячное потребление электроэнергии $\bar{x}$:

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}.$$

Квантиль распределения случайной величины по τ-критерию:

$$\tau = \frac{\left| x_{\text{кр}} - \bar{x} \right|}{S}.$$

По таблице 5.5 выбирается $\tau_{\text{табл}}$ и сравнивается с расчетным значением τ на основании чего и делается вывод о эффективности энергосберегающих мероприятий.

Примечание: При определении S^2 и $\bar{x}$ необходимо $x_{кр}$ исключить!

5.5 Контрольные вопросы

1. Цель применения критериев согласия?
2. Условия принятия (отвержения) гипотез:
   - а) по критерию Стьюдента;
   - в) по критерию Пирсона;
   - с) по критерию Фишера;
   - d) по критерию Кохрена;
   - е) по τ-критерию.

ЛАБОРАТОРНАЯ РАБОТА № 6

Корреляционный анализ

6.1 Цель работы

Исследовать связи между параметрами статистической совокупности на основе корреляционного анализа. Определить тесноту связи между параметрами.

6.2 Краткие теоретические сведения

Связи между различными явлениями в природе и трудовой деятельности человека сложны и многообразны, но их можно определенным образом классифицировать. Часто речь может идти о функциональной связи между переменными x и y (когда возможному значению x однозначно соответствует значение y), или о статистической (когда одна случайная переменная реагирует на изменение другой изменением своего закона распределения). При исследовании связей и зависимостей между случайными переменными X и Y обычно ограничиваются изучением зависимости между одной из них и условным математическим ожиданием другой, т.е. $M(Y|_{X=x}) = \varphi(x)$ или $M(X|_{Y=y}) = \varphi(y)$, где $M(Y|_{X=x})$ – математическое ожидание случайной величины Y при условии, что случайная величина X приняла значение x. Аналогично для $M(X|_{Y=y})$.

Знание статистической зависимости между случайными переменными имеет большое практическое значение: с ее помощью можно прогнозировать значение зависимой случайной переменной в предположении, что независимая переменная примет определенное значение.

Статистические связи между переменными можно изучать методом корреляционного и регрессионного анализа. С помощью этих методов решают разные задачи; требования, предъявляемые к исследуемым переменным, в каждом методе различны.

Основная задача корреляционного анализа – выявление связи между случайными переменными путем точечной и интервальной оценки парных коэффициентов корреляции, вычисления и проверки значимости множественных коэффициентов корреляции и детерминации, оценки частных коэффициентов корреляции. Корреляционный анализ позволяет также оценить функцию регрессии одной случайной переменной на другую. Предпосылки корреляционного анализа следующие: 1) переменные величины должны быть случайными; 2) случайные величины должны иметь совместное нормальное распределение.

Основные понятия и принципы проведения корреляционного анализа. Если случайные переменные X и Y имеют совместное нормальное распределение, то связь между X и Y можно описать коэффициентом корреляции ρ. Этот коэффициент определяется как ковариация между X и Y, отнесенная к их средним квадратическим отклонениям:

$$\rho = \frac{K_{XY}}{\sigma_x \sigma_y}, \text{ или } \rho = M\left\{\frac{X - M(x)}{\sigma_x} \quad \frac{Y - M(y)}{\sigma_y}\right\}$$

Оценкой коэффициента корреляции является выборочный коэффициент корреляции r. Для его нахождения необходимо знать оценки следующих параметров: M(x), M(y), σ_x, σ_y. Наилучшей оценкой математического ожидания является среднее арифметическое, а оценкой дисперсии служит выборочная дисперсия, т. е.

$$\text{M(x)} = \bar{\text{x}} = \sum_{i=1}^{n} x_i / n; \qquad \sigma_x^2 = s_x^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 / n.$$

Тогда выборочный коэффициент корреляции

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n\, s_x s_y}.$$

Коэффициент ρ называют также *парным коэффициентом корреляции*, а *r – выборочным парным коэффициентом корреляции*.

Используя рассмотренные выше параметры распределения и коэффициент корреляции, можно получить выражение для условного математического ожидания, т. е. записать выражение для функции регрессии одной случайной величины на другую. Так, функция регрессии Y на X имеет вид

$$M(Y|_{X=x}) = M(y) + \rho \frac{\sigma_y}{\sigma_x}[X - M(x)];$$

функция регрессии X на Y – следующий вид:

$$M(X|_{Y=y}) = M(x) + \rho \frac{\sigma_x}{\sigma_y}[Y - M(y)].$$

Выражения $\rho \frac{\sigma_y}{\sigma_x}$ и $\rho \frac{\sigma_x}{\sigma_y}$ называют *коэффициентами регрессии*. Подставив в функцию регрессии Y на X соответствующие оценки параметров, получим уравнения регрессии, график которой – прямая линия, проходящая через точку $(\bar{x}; \quad \bar{y})$.

Запишем уравнение регрессии y на x и x на y:

$$\bar{y}(x)=\bar{y}+r\frac{s_y}{s_x}(x-\bar{x}), \quad \bar{x}(y)=\bar{x}+r\frac{s_x}{s_y}(y-\bar{y}).$$

В корреляционном анализе на основе оценок параметров двумерной нормальной совокупности получаем оценки тесноты связи между случайными переменными. Особенностью корреляционного анализа является строго линейная зависимость между переменными. Это обусловливается исходными предпосылками.

Свойства коэффициента корреляции.

Коэффициент корреляции принимает значения на интервале (-1, $+1$).

Если коэффициент корреляции положителен, то связь между переменными также положительна и значения переменных увеличиваются или уменьшаются одновременно. Если коэффициент корреляции имеет отрицательное значение, то при увеличении одной переменной уменьшается другая.

Коэффициент корреляции не зависит от выбора начала отсчета и единицы измерения.

Величина коэффициента корреляции не изменится, если переменные X и Y увеличить или уменьшить в a раз, а также вычитать или прибавлять к значениям X и Y одно и то же число b.

Если коэффициент корреляции $\rho_{xy}=0$, то случайные переменные не коррелированы. Понятие не коррелированности не следует смешивать с понятием независимости, независимые величины всегда не коррелированы. Однако обратное утверждение невероятно: некоррелированные величины могут быть зависимы и даже функционально, однако эта связь не линейная.

Величина выборочного коэффициента корреляции не зависит от порядка следования переменных, т.е. $r_{xy} = r_{yx}$, поэтому обычно пишут просто r.

Поле корреляции. Для вычисления оценок параметров двумерной модели удобно использовать *корреляционную таблицу и поле корреляции*. При изучении зависимости между Y и X имеется, как правило, выборка из генеральной совокупности n пар переменных (x_i, y_i). Считаем, что предпосылки корреляционного анализа выполнены.

Пару случайных чисел (x_i, y_i) можно изобразить графически в виде точки с координатами (x_i, y_i). Аналогично можно изобразить весь набор пар случайных чисел (всю выборку). Однако при большом объеме выборки это затруднительно. Задача упрощается, если выборку упорядочить, т.е. переменные сгруппировать. Сгруппированные ряды могут быть как дискретными, так и интервальными.

По осям координат откладывают или дискретные значения переменных, или интервалы их изменения. Для интервального ряда наносят координатную сетку. Каждую пару переменных из данной выборки изображают в виде точки с соответствующими координатами для дискретного ряда или в виде точки в соответствующей клетке для интервального ряда.

Такое изображение корреляционной зависимости называют *полем корреляции.* На рисунке изображено поле корреляции для выборки, состоящей из n пар переменных (ряд интервальный). Если вычислить средние значения y в каждом интервале изменения x [обозначим их $y_i(x)$], нанести эти точки на рис.6.1 и соединить между собой, то получим ломаную линию, по виду которой можно судить, как в среднем меняются y в зависимости от изменения x. По виду этой линии можно также сделать предположение о форме связи между переменными. В данном случае ломаную линию можно аппроксимировать прямой линией, так как она достаточно хорошо приближается к ней.

По выборочным данным можно построить также корреляционную таблицу (табл. 6.1). Корреляционную таблицу, как и поле корреляции, строят по сгруппированному ряду (дискретному или интервальному).

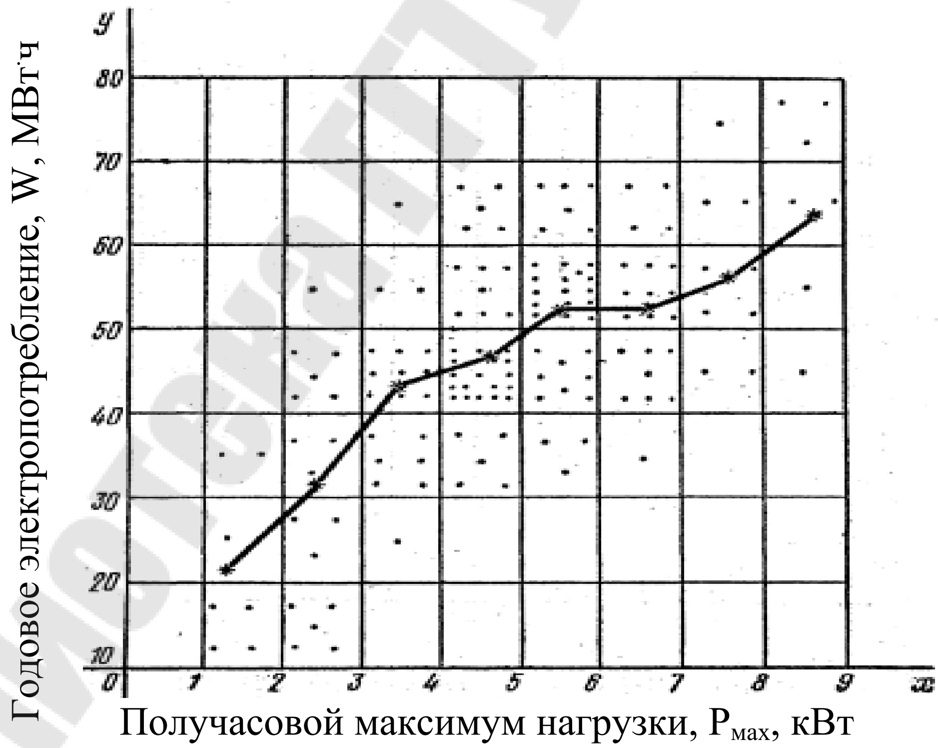


Рис. 6.1 Поле корреляции

Таблица 6.1

Корреляционная таблица

Годовое электропотребление, y, МВт·ч	Получасовой максимум нагрузки, x, кВт								
	1-2 1,5	2-3 2,5	3-4 3,5	4-5 4,5	5-6 5,5	6-7 6,5	7-8 7,5	8-9 8,5	m_y
10-20 15	4	5							9
20-30 25	1	3	1						5
30-40 35	2	3	6	5	3	1			20
40-50 45		5	9	19	8	7	2	1	51
50-60 55		1	2	7	16	9	4	2	41
60-70 65			1	5	6	4	2	2	20
70-80 75							1	3	4
m_x	7	17	19	36	33	21	9	8	150

Примечание: В таблице 6.1 приведено $n=150$.

Корреляционной таблицей удобно пользоваться при вычислении коэффициентов корреляции и параметров уравнений регрессии.

Корреляционная таблица построена на основе интервального ряда, поэтому для оценок параметров воспользуемся формулами для вычисления средней арифметической и дисперсии. Имеем:

$$\overline{\mathrm{x}} = \frac{\sum\limits_x x m_x}{\sum\limits_x m_x}; \overline{\mathrm{y}} = \frac{\sum\limits_y y m_y}{\sum\limits_y m_y};$$

$$s_x^2 = \frac{\sum\limits_x x^2 m_x}{\sum\limits_x m_x} - \left(\frac{\sum\limits_x x m_x}{\sum\limits_x m_x}\right)^2; s_y^2 = \frac{\sum\limits_y y^2 m_y}{\sum\limits_y m_y} - \left(\frac{\sum\limits_y y m_y}{\sum\limits_y m_y}\right)^2.$$

Корреляционное отношение. Часто предпосылки корреляционного анализа нарушаются: один из признаков оказывается величиной не случайной, или признаки не имеют совместного нормального распределения. Но статистическая зависимость между ними существует. Для изучения связи между признаками в этом случае существует общий показатель зависимости признаков, основанный на показателе изменчивости – общей (или полной) дисперсии.

Полной называется дисперсия признака относительно его математического ожидания. Так, для признака Y это $\sigma_Y^2 = M[Y - M(Y)]^2$.

Общая дисперсия σ_Y^2 может быть представлена суммой следующих дисперсий: $\delta_{Y|_x}^2$ - *групповая* дисперсия, измеряющая влияние признака X на Y, и $\sigma_{Y|_X}^2$ - *остаточная* дисперсия, измеряющая влияние на Y прочих факторов.

Тесноту связи удобно оценивать в единицах общей дисперсии $\sigma_{y|}^2$, т.е. рассматривать отношение $M\left[\overline{y}(x) - M(Y)\right]^2 / \sigma_y^2$.Эту величину обозначают $\eta_{T\,Y|_x}^2$ и называют *теоретическим корреляционным* отношением и определяю:

$$\eta_{T\,Y|_x}^2 = \frac{M\left[\overline{y}(x) - M(Y)\right]^2}{\sigma_Y^2} = \frac{\delta_{Y|X}^2}{\sigma_Y^2}.$$

Значения $\eta_{T\,Y|_x}^2$ всегда заключены между нулем и единицей. Обозначим $\eta_{T\,Y|_x}^2$ - η^2. При вычислении η^2 по выборочным данным получаем *выборочное корреляционное отношение*. Обозначим его $\overline{\eta^2}$. Вместо дисперсий в этом случае используются их оценки. Тогда, $\overline{\eta^2} = \frac{s_{Y|_x}^2}{s_Y^2}$.

Эмпирическое корреляционное отношение является одним из показателей, характеризующих тесноту корреляционной зависимости, то есть степень ее приближения к функциональной связи, определяется по формуле:

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}},$$

где δ^2 – групповая дисперсия; а σ^2 – общая дисперсия.

Корреляционное отношение изменяется от 0 до 1.

Для качественной оценки тесноты связи на основе показателя эмпирического корреляционного отношения можно пользоваться таблицей 6.2

Таблица 6.2

Оценка тесноты связи

Величина 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

6.3 Задание

Имеются следующие данные по 25 предприятиям о получасовом максимуме нагрузке P_{max} (МВт) – *Х* и расходе электроэнергии W (ГВт*ч) – *У*. Вариант исходных данных определяется преподавателем по таблице Х.Х

Требуется: 1) Представить поле корреляции. 2) Определить эмпирическое корреляционное отношение (разделив предприятия на 5 групп по P_{max}).

Таблица исходных данных по вариантам представлена в табл. 6.3

6.4 Методические указания к выполнению задания

По варианту исходных данных 20 предприятий: $P_{мах}$ – Х; W – У:

Х	455	267	203	168	140	99	63
У	2319	2095	1284	1385	928	653	357
Х	51	47	40	31	28	25	18
У	356	264	310	209	172	131	121
Х	17	15	13	11	9	8	-
У	87	87	77	111	21	56	-

Исходные данные по *Х* разобьем на пять групп с интервалом: $(Х_{мах}$-$Х_{мин})/5$= (455-8)/5 = 89,4; поэтому первый интервал будет: 8÷97,4.

Для выполнения п.2 воспользуемся программным пакетом Excel. Результаты представим в таблице:

Группы предприятий по уровню получасового максимума, $P_{мах}$, МВт	Число предприятий, (*n*)	Расход электроэнергии предприятиями, У(n_i) W, ГВт·ч	Расход на одно предприятие, $у_{гр}$, W, ГВт·ч	$(у_{гр}$-$у_{ср})$	$(у_{гр}$-$у_{ср})^2$	$(у_{гр}$-$у_{ср})^2 \cdot n$
8-97,4	14	2359	168,5	-385,65	148725,92	2082162,92
97,4-186,8	3	2966	988,67	437,52	191423,75	574271,25
186,8-276,2	2	3379	1689,5	1138,35	1295840,72	2591681,45
276,2-365,6	0	0	0	-551,15	303766,32	0
365,6-455	1	2319	2319	1767,85	3125293,62	3125293,62
Итого	20	11023	551,15			8373409,24

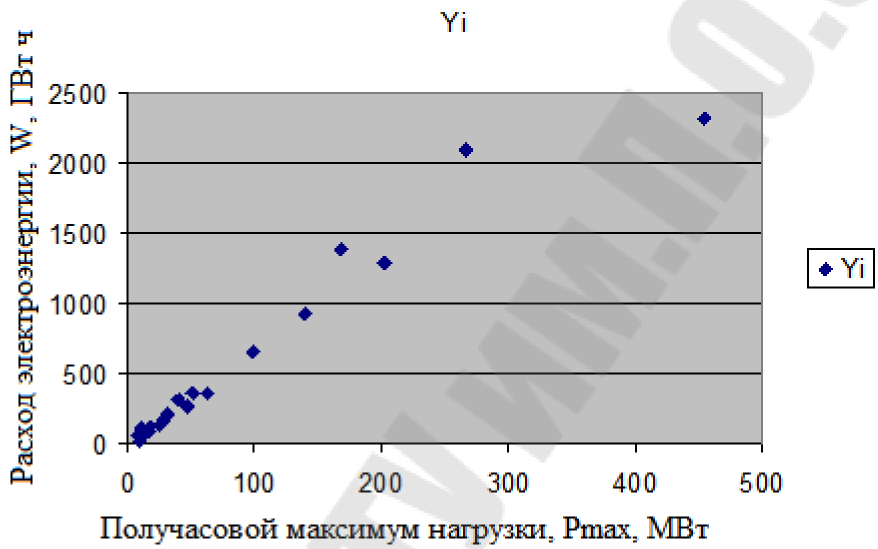
Групповая дисперсия расхода электроэнергии: δ^2= 8373409,24/ 20 = 418670,462.

Общая дисперсия: $\sigma^2 = [\sum(У_i)^2]/N - (у_{ср})^2 = 15183273/20 - 551{,}15^2$ =455397,33

Коэффициент $\acute{\eta}^2 = \delta^2/\sigma^2$=0,9194; $\acute{\eta}$ = 0,9589.

Сравнивая с данными табл.6.2, видим, что связь между У и X весьма тесная.

Представляем поле корреляции:



6.5 Контрольные вопросы

1. Область применения статистических зависимостей между случайными переменными.

2. Какими методами изучаются статистические связи между переменными?

3. Основная задача корреляционного анализа?

4. Предпосылки корреляционного анализа.

5. Как описывается и определяется связь между случайными переменными X и Y, имеющих совместное нормальное распределение?

6. Что является оценкой коэффициента корреляции?

7. Какие параметры и их оценки необходимо знать для определения выборочного коэффициента корреляции?

8. Как еще называют коэффициент корреляции ρ и выборочный коэффициент корреляции r ?

9. В чем особенность корреляционного анализа, чем это обусловлено?

10. Основные свойства коэффициента корреляции.

11 Назначение и определение поля корреляции, корреляционной таблицы.

12. Как определяются групповая, общая дисперсии и корреляционное отношение?

13. Групповая и общая дисперсии зависят от одной и той же случайной переменной (т.е. $\delta^2=f(y)$ и $\sigma^2=f(y)$). Объясните, почему их отношение (т.е. δ^2/σ^2) определяет связь между X и Y ?

Таблица 6.3 Исходные данные для расчета

Nвар	1		2		3		4		5		6		7		8		9		10	
Nпред	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y
1	51	370	75	417	95	665	81	595	45	315	62	475	37	290	49	350	68	455	59	380
2	50	365	74	410	93	660	80	590	41	312	61	472	35	287	47	348	67	450	58	365
3	49	361	74	408	92	654	76	585	40	310	60	470	34	284	46	347	66	445	56	361
4	47	327	72	405	90	651	74	581	40	307	58	467	31	282	44	340	65	442	55	360
5	42	321	71	403	87	650	71	580	39	300	54	465	29	274	42	330	63	440	51	358
6	42	319	70	400	86	645	70	575	37	295	53	463	28	271	41	328	61	437	50	356
7	41	305	69	395	84	643	65	570	36	292	51	462	26	270	40	321	58	435	48	350
8	38	301	68	390	81	640	60	565	34	290	50	458	24	268	38	317	57	430	47	345
9	37	300	65	387	80	635	58	562	32	287	48	450	22	265	37	315	56	427	45	340
10	35	295	62	386	74	631	57	561	31	285	47	447	21	261	35	310	53	425	43	335
11	32	290	61	384	71	628	55	560	30	284	46	441	21	260	34	307	51	420	42	330
12	31	285	60	379	70	625	53	558	28	280	45	440	20	258	32	305	50	415	41	327
13	30	283	58	354	68	622	51	554	27	279	44	435	19	254	31	301	49	410	40	326
14	28	280	57	350	67	620	31	503	24	275	42	431	18	251	29	300	46	407	39	325
15	27	274	54	348	64	618	30	501	23	270	41	430	16	250	27	295	43	403	36	320
16	25	271	53	345	63	615	30	497	21	265	40	428	15	243	25	290	42	401	31	310
17	21	265	51	340	60	610	28	480	20	260	39	421	13	240	24	287	41	400	29	305
18	20	260	50	338	58	605	27	478	18	255	37	420	12	234	23	285	38	395	27	301
19	20	258	45	335	56	601	27	475	17	251	34	419	11	232	21	281	37	390	26	300
20	19	252	42	331	55	600	26	460	15	250	31	415	9	215	20	280	35	385	23	295
21	18	250	41	328	43	595	26	458	13	249	29	401	8	210	18	274	34	380	21	291
22	16	249	40	326	40	593	25	450	12	247	27	400	7	208	17	273	31	375	20	290
23	15	247	39	321	38	590	24	445	11	245	25	397	6	204	15	270	29	370	18	287
24	14	241	38	320	35	585	23	440	10	240	23	395	5	201	13	265	28	364	16	285
25	13	235	35	315	31	582	22	435	7	235	21	391	3	197	11	261	26	361	13	284

Продолжение таблицы 6.3

Nвар	11		12		13		14		15		16		17		18		19		20	
Nпред	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y
1	74	515	85	695	105	715	121	755	117	735	27	324	39	337	41	498	56	510	64	540
2	72	512	84	691	103	712	119	751	110	734	27	321	37	333	40	497	55	508	63	537
3	71	510	83	690	101	711	118	750	105	730	26	320	36	332	40	496	51	506	61	535
4	68	505	81	687	100	706	115	749	103	728	25	319	34	330	39	495	50	504	60	533
5	65	503	80	685	98	704	110	695	102	727	21	307	31	328	37	490	45	502	58	531
6	61	501	76	684	95	702	105	691	101	725	20	305	30	326	35	488	43	500	56	529
7	59	500	72	681	91	697	102	690	100	721	19	303	28	325	33	487	42	498	55	521
8	57	497	70	675	90	693	101	667	100	719	18	301	27	320	32	485	40	496	50	520
9	53	493	68	671	85	690	97	664	97	717	17	300	25	310	31	484	39	494	48	519
10	51	491	64	669	80	684	95	661	95	714	16	297	24	305	30	481	36	492	44	517
11	50	485	63	664	75	681	92	658	94	710	15	295	23	300	28	480	35	490	43	516
12	48	481	61	662	74	680	91	651	93	700	13	294	20	295	27	478	31	485	42	515
13	47	474	60	657	72	677	90	650	91	695	11	293	19	290	25	476	30	480	40	513
14	45	471	57	651	70	683	87	647	90	690	10	290	18	287	21	474	29	475	39	511
15	43	470	55	647	66	671	85	641	87	687	9	289	15	285	20	472	28	470	38	508
16	41	467	54	643	63	665	84	640	84	681	8	287	13	282	18	470	27	465	37	506
17	39	465	52	635	61	664	83	638	83	665	7	285	13	280	16	468	26	460	36	504
18	37	460	51	631	60	662	81	636	81	661	6	283	12	278	15	466	25	455	35	502
19	36	455	50	625	59	658	80	628	80	655	6	281	11	275	12	464	20	453	34	500
20	35	450	49	621	57	654	75	621	79	650	5	280	9	273	10	462	18	451	31	454
21	34	445	48	620	55	651	70	620	76	645	5	279	8	271	9	460	16	450	30	450
22	30	440	47	615	51	650	69	610	75	641	4	277	7	270	7	458	15	449	29	445
23	28	435	45	612	50	648	68	607	70	640	3	276	6	268	6	456	14	448	28	440
24	27	430	40	610	48	641	65	602	68	635	2	274	5	265	5	454	13	447	27	435
25	25	427	35	608	46	630	63	601	65	630	2	273	3	264	2	452	11	446	25	430

Продолжение таблицы 6.3

Nвар	21		22		23		24		25		26		27		28		29		30	
Nпред	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y
1	71	699	83	718	94	738	103	761	111	794	137	815	145	837	154	849	163	811	125	929
2	70	698	81	708	94	734	102	758	108	792	135	813	143	836	151	841	162	810	124	927
3	68	697	80	707	91	731	101	756	107	791	130	811	141	835	150	840	160	808	123	925
4	65	696	79	706	90	730	99	754	106	790	125	801	137	833	149	838	157	806	120	924
5	64	695	76	703	87	724	98	751	101	771	120	800	134	831	147	837	155	804	110	921
6	63	693	71	700	86	721	96	749	100	770	118	799	133	829	146	836	154	802	105	920
7	61	691	70	698	85	720	95	741	97	768	117	789	132	827	140	817	153	801	100	917
8	60	690	69	691	81	711	94	739	95	764	115	779	131	825	138	807	152	798	97	916
9	59	685	68	690	80	704	93	731	94	762	113	771	130	824	137	801	150	788	95	914
10	57	681	66	688	75	701	91	721	93	761	111	769	129	823	135	800	148	778	93	912
11	54	680	64	687	71	691	90	720	91	758	109	768	121	821	134	797	147	768	92	909
12	53	675	62	681	70	681	89	719	90	756	108	767	120	819	130	791	140	758	91	907
13	52	670	60	680	65	680	87	718	90	755	106	761	118	818	128	790	135	751	90	906
14	51	665	58	671	60	677	86	717	87	754	105	760	116	816	125	787	133	750	86	904
15	49	660	54	661	59	671	84	716	85	751	103	757	114	814	122	781	131	747	85	902
16	47	658	53	651	58	670	81	715	83	749	101	751	112	812	120	780	130	746	84	901
17	46	657	51	641	57	666	80	714	82	747	98	749	110	810	118	767	128	745	81	891
18	45	656	50	634	55	661	79	704	81	741	94	747	107	808	116	761	126	744	80	890
19	40	655	49	633	50	660	77	701	80	740	90	746	106	806	115	760	125	742	79	888
20	35	651	47	632	49	656	76	700	68	739	87	745	105	804	114	757	121	741	77	884
21	30	647	45	630	48	654	75	681	66	737	84	744	103	802	111	751	120	738	75	882
22	25	644	44	621	45	651	74	671	63	736	81	742	102	800	110	747	119	736	74	880
23	20	642	42	620	40	650	71	665	61	734	79	741	101	780	108	741	118	734	71	878
24	19	640	41	611	35	640	69	663	60	733	77	738	100	760	106	737	117	730	70	876
25	18	636	40	602	31	634	67	661	59	731	74	731	89	740	104	731	115	727	68	874

ЛАБОРАТОРНАЯ РАБОТА № 7
Регрессионный анализ

7.1 Цель работы

Применить методы регрессионного анализа для определения параметров уравнения регрессии между получасовым максимумом нагрузки P_{max} (МВт) и расходе электроэнергии W (ГВт·ч); оценить точность подбора вида функции регрессии

7.2 Краткие теоретические сведения

Основные понятия регрессии

При изучении статистических зависимостей форму связей форму связи можно характеризовать функцией регрессии (линейной, квадратичной, показательной и т. д.)

Функция регрессии имеет значение при статистическом анализе зависимостей между переменными и может быть использована для прогнозирования одной из случайных переменных, если известно значение другой случайной переменной. Точность такого прогноза определяется дисперсией условного распределения.

Для оценки линейной функции регрессии необходимо знать аналитический вид двухмерного распределения *(X,Y)*, а далее оценить параметры двухмерного распределения. На практике, для оценки чаще всего располагают выборкой ограниченного объема, по которой нужно найти вид функции регрессии. Для характеристики формы связи при изучении зависимости используют понятие кривой регрессии.

Кривой регрессии *Y* по *X* (или *X* по *Y*) называют условное среднее значение случайной переменной *Y*, рассматриваемое как функция определенного класса, параметры которой находят методом наименьших квадратов по наблюденным значениям двумерной случайной величины (*x*, *y*), т.е.

$$\overline{\mathrm{y}}(x) = \varphi(x, b_1, b_2, ..., b_m). \quad \text{или} \quad \overline{\mathrm{x}}(y) = \varphi(y, c_1, c_2, ..., c_m).$$

Кривую регрессии называют также эмпирическим уравнением регрессии или просто уравнением регрессии. Кривая регрессии, выражаемая как $\overline{\mathrm{y}}(x)$, минимизирует среднюю квадратическую погрешность прогноза *Y* по *X*.

Под регрессией понимается функция, предназначенная для описания зависимости изменения результативных признаков под влиянием колебаний признаков-факторов. Наличие корреляционной связи между параметрическими признаками позволяет приближенно предоставить значения результативного признака в виде некоторой функции от величины одного или нескольких факторных признаков.

Функцию, показывающую корреляционную зависимость между признаками, принято называть *уравнением регрессии.* Если оно связывает лишь два признака, то представляет собой *уравнение парной регрессии;* если отражает зависимость результативного признака от двух, трех и более факторных признаков – это *уравнение множественной регрессии.*

При выявлении корреляционной формы, связывающей результативный признак с одним факторным, помогает графическое изображение связи в виде поля корреляции. Обычно считают, что увеличение результативного и факторного признаков в арифметической прогрессии при прямой связи требует применения прямолинейной, а при обратной – гиперболической регрессии.

Составление уравнения регрессии означает прежде всего определение его параметров, используя для этого, где возможно, способ наименьших квадратов, согласно которому сумма квадратов отклонений фактических значений результативного признака от теоретических значений, рассчитанных по уравнению регрессии, должна быть наименьшей, т. е.

$$\sum (y - y_x)^2 \to \min,$$

где y – фактические варианты признака-результата; y_x – теоретические значения признака-результата, рассчитанных по уравнению регрессии.

Это условие приводит к системе нормальных уравнений, решение которых позволяет определить параметры уравнения регрессии. При этом число нормальных уравнений на одно больше числа входящих в уравнение регрессии факторов. Если известны параметры уравнения, то, подставляя в него принятые значения факторных признаков, можно рассчитать теоретическое значение результативного признака, что делает удобным применение корреляционных уравнений при прогнозировании результативных признаков.

Уравнение регрессии может показать связь между признаками более точно, если оно построено на основании достаточно большой статистической совокупности. Уравнение регрессии нередко называют *моделью* связи между признаками.

Основная задача регрессионного анализа – изучение зависимости между результативным признаком Y и наблюдавшимся признаком X, оценка функции регрессии.

Предпосылки регрессионного анализа:

1) $\boldsymbol{Y}$ – независимые случайные величины, имеющие постоянную дисперсию;

2) $\boldsymbol{X}$ – величины наблюдаемого признака – величины не случайные;

3) условное математическое ожидание $\boldsymbol{M(Y|_{X=x})}$ можно представить в виде

$$\boldsymbol{M(Y|_{X=x}) = \varphi(x) = \beta_0 + \beta_1 x}. \qquad (7.1)$$

Выражение (7.1) называется функцией регрессии (или модельным уравнением регрессии) $\boldsymbol{Y}$ на $\boldsymbol{X}$. Оценке в этом выражении подлежат параметры $\boldsymbol{\beta_0}$ и $\boldsymbol{\beta_1}$, называемые *коэффициентами регрессии*, а также $\boldsymbol{\sigma^2_{ост}}$ – *остаточная дисперсия.*

Остаточной дисперсией называется та часть рассеивания результативного признака, которую нельзя объяснить действием наблюдаемого признака. Остаточная дисперсия может служить для оценки точности подбора вида функции регрессии (модельного уравнения регрессии). Оценки параметров функции регрессии находят, используя метод наименьших квадратов.

Рассмотрим линейный регрессионный анализ. Линейным он называется потому, что изучает лишь те виды зависимостей $\varphi(x)$, которые линейны по оцениваемым параметрам, хотя могут быть не линейны по переменным X. Например, зависимости $M(Y|_{X=x})=\varphi_1(x)=\beta_0+\beta_1\ x$; $M(Y|_{X=x}) = \varphi_2(x) = \beta_0+\beta_1/x$, $M(Y|_{X=x}) = \varphi_3(x) = \beta_0+\beta_1\ x+\beta_2\ x^2$ линейны относительно параметров $\beta_0, \beta_1, \beta_2$, хотя вторая и третья зависимости не линейны относительно переменных x. Вид зависимости $\varphi(x)$ выбирают, исходя из визуальной оценки характера расположения точек на поле корреляции.

В работе будем рассматривать так называемую «нормальную регрессию», т.е. будем считать, что случайная величина подчиняется нормальному закону распределения. Предпосылки «нормальной регрессии»:

1) Y – независимые случайные величины, имеющие постоянную дисперсию и распределенные по нормальному закону;

2) X – величины наблюдаемого признака – величины не случайные;

3) условное математическое ожидание $M(Y|_{X=x})$ можно представить в виде (7.1).

В этом случае оценки коэффициентов регрессии – несмещенные с минимальной дисперсией и нормальным законом распределения. Из этого положения следует, что при «нормальной регрессии» имеется возможность оценить значимость оценок коэффициентов регрессии, а также построить доверительный интервал для коэффициентов регрессии и условного математического ожидания M(Y|X=x).

Линейная регрессия. Модель уравнения регрессии вида (7.1), когда зависимость $\varphi(x)$ линейна и по оцениваемым параметрам и по переменным. Оценки параметров модели (7.1) β_0 и β_1 обозначим b_0 и b_1, а оценку остаточной дисперсии $\sigma^2_{ост}$ – $s^2_{ост}$. Подставив формулу (7.1) вместо параметров их оценки, получим уравнение регрессии $\overline{y}(x) = b_0 + b_1 \cdot x_1$, коэффициенты которого находят из условия минимума суммы квадратов отклонений измеренных значений результативного признака y_i от вычисленных по уравнению регрессии $y(x_i)$:

$$Q = \sum_{i=1}^{n} \left[y_i - y(x_i) \right]^2 = \min, \quad \text{или}$$

$$Q = \sum_{i=1}^{n} \left[y_i - b_0 - b_1 \cdot x_i \right]^2 = \min.$$

Для определения b_0 и b_1 находим $\dfrac{\partial Q}{\partial b_0} = 0$, $\dfrac{\partial Q}{\partial b_1} = 0$; это позволяет получить систему двух нормальных уравнений:

$$\begin{cases} nb_0 + b_1 \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} y_i \\ b_0 \sum\limits_{i=1}^{n} x_i + b_1 \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i y_i \end{cases}$$

Оценки, полученные по способу наименьших квадратов, обладают минимальной дисперсией в классе линейных оценок. Решая систему относительно b_0 и b_1, найдем оценки параметров β_0 и β_1:

$$b_1 = \frac{\sum\limits_{i=1}^{n} x_i y_i - \frac{1}{n} \sum\limits_{i=1}^{n} x_i \sum\limits_{i=1}^{n} y_i}{\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2 / n},$$

$$b_0 = \left(\sum_{i=1}^{n} y_i \right) / n - b_1 \left(\sum_{i=1}^{n} x_i \right) / n = \bar{y} - b_1 \bar{x}.$$

Оценка остаточной дисперсии $\sigma^2_{ост}$ определяется по формуле

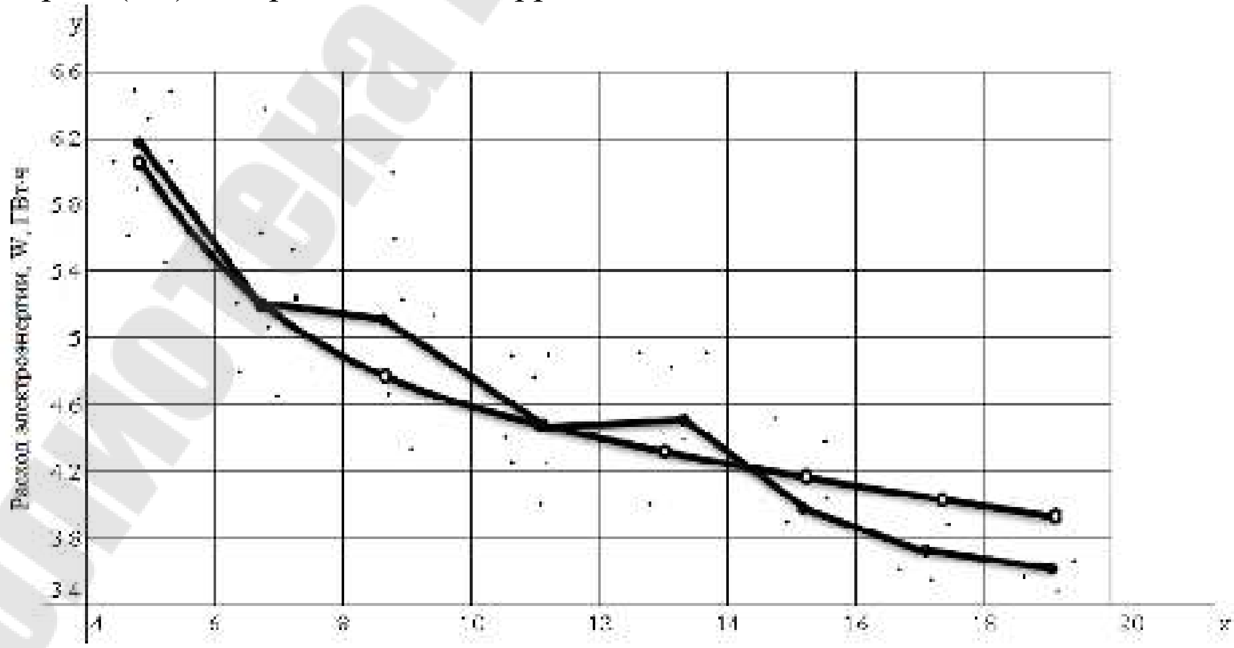
$$s^2_{ост} = \frac{1}{n\text{-}2} \sum_{i=1}^{n} \left[y_i - y(x_i) \right]^2,$$

где n – количество наблюдений.

Нелинейная регрессия. Рассмотрим модель уравнения вида, когда зависимость нелинейна по переменным x, например

$$M(Y|_{X=x}) = \varphi(x) = \beta_0 + \beta_1 / x. \qquad (7.2)$$

На рис. (7.1) изображено поле корреляции



Получасовой максимум нагрузки, Pmax, МВт

Рис. 7.1 Поле корреляции

Очевидно, что зависимость между y и x нелинейна и ее графическим изображением является не прямая, а кривая. Оценкой последнего выражения является уравнение регрессии

$$\overline{y}(x) = b_0 + b_1 / x ,$$

где b_0 и b_1 – оценки коэффициентов регрессии β_0 и β_1.

Принцип нахождения коэффициентов тот же – метод наименьших квадратов, т.е.

$$Q = \sum_{i=1}^{n} \left[y_i - b_0 - b_1 / x_i) \right]^2 = Q \min$$

Дифференцируя последнее равенство по b_0 и b_1 и приравнивая правые части нулю, получаем так называемую систему нормальных уравнений:

$$\begin{cases} \sum\limits_{i=1}^{\text{n}} b_0 + \sum\limits_{i=1}^{n} b_1 / x_1 = \sum\limits_{i=1}^{n} y_i \\ \sum\limits_{i=1}^{n} b_0 / x_i + \sum\limits_{i=1}^{n} b_1 / x_i^2 = \sum\limits_{i=1}^{n} y_i / x_i \end{cases} .$$

В общем случае нелинейной зависимости между переменными Y и X связь может выражаться многочленом k-й степени от x:

$$M(Y | X = x) = \beta_0 + \beta_1 x + \beta_2 x^2 + ... + \beta_k x^k \qquad (7.3)$$

Коэффициенты регрессии определяют по принципу наименьших квадратов. Система нормальных уравнений имеет вид

$$\left\{ \begin{array}{l} b_0 n + b_1 \sum\limits_{i=1}^{n} x_i + b_2 \sum\limits_{i=1}^{n} x_i^2 + ... + b_k \sum\limits_{i=1}^{n} x_i^k = \sum\limits_{i=1}^{n} y_i \\ b_0 \sum\limits_{i=1}^{n} x_i + b_1 \sum\limits_{i=1}^{n} x_i^{\ 2} + b_2 \sum\limits_{i=1}^{n} x_i^{\ 3} + ... + b_k \sum\limits_{i=1}^{n} x_i^{k+1} = \sum\limits_{i=1}^{n} y_i x_i \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ b_0 \sum\limits_{i=1}^{n} x_i^k + b_1 \sum\limits_{i=1}^{n} x_i^{k+1} + b_2 \sum\limits_{i=1}^{n} x_i^{\ k+2} + ... + b_k \sum\limits_{i=1}^{n} x_i^{2k} = \sum\limits_{i=1}^{n} y_i x_i^{\ k} \end{array} \right\}$$

Вычислив коэффициенты системы, ее можно решить любым известным способом и получить аналитическое уравнение регрессии. Оценка остаточной дисперсии в общем виде:

$$s^2 = \frac{\sum\limits_{i=1}^{n} \left[y_i - y(x_i) \right]^2}{n - m - 1} ,$$

где m – количество параметров уравнения регрессии.

7.3 Задание

На основе корреляционного поля (по исходным данным предыдущей лабораторной работы) и вида уравнений (заданных преподавателем) определить: 1) коэффициенты уравнений регрессии; 2) остаточные дисперсии; 3) теоретическое корреляционное отношение или индекс корреляции.

Виды уравнений:

1) $\square(x)=a_0+a_1\cdot x$;
2) $\square(x)=a_0+a_1/x$;
3) $\square(x)=a_0+a_1\cdot x+ a_2\cdot x^2$;
4) $\square(x)=a_0+a_1/x+ a_2/x^2$

7.4 Методические указания к выполнению задания

При выполнении лабораторной работы рекомендуется воспользоваться программным пакетом Excel.

Исходные данные приняты по предыдущей работе и представлены в виде таблицы 7.1

Таблица 7.1

Исходные данные

Исходые данные		Расчетные данные				
Xi	*Yi*	*Xi^2*	*Xi*Yi*	*Yx*	*Yi-Yx*	*(Yi-Yx)^2*
455	2319	207025	1055145	2734,17	-415,171586	172367,4
267	2095	71289	559365	1623,76	471,2398266	222067
203	1284	41209	260652	1245,75	38,25222252	1463,233
168	1385	28224	232680	1039,02	345,9777515	119700,6
140	928	19600	129920	873,642	54,35817474	2954,811
99	653	9801	64647	631,478	21,52236587	463,2122
63	357	3969	22491	418,846	-61,8456614	3824,886
51	356	2601	18156	347,968	8,031662799	64,50761
47	264	2209	12408	324,343	-60,3425625	3641,225
40	310	1600	12400	282,997	27,00254335	729,1373
31	209	961	6479	229,839	-20,8394635	434,2832
28	172	784	4816	212,12	-40,1201324	1609,625
25	131	625	3275	194,401	-63,4008014	4019,662
18	121	324	2178	153,056	-32,0556956	1027,568
17	87	289	1479	147,149	-60,1492519	3617,933
15	87	225	1305	135,336	-48,3363645	2336,404
13	77	169	1001	123,523	-46,5234771	2164,434
11	111	121	1221	111,711	-0,71058976	0,504938
9	21	81	189	99,8977	-78,8977024	6224,847
8	56	64	448	93,9913	-37,9912587	1443,336
1708	11023	391170	2390255	11023	1,27898E-13	550154,6
b1	5,906					

b0	46,74					
Общая дисперсия			455397,33			
Регрессионная дисперсия			27507,7316			
Теор.коррел отношен			0,969328			

Решение системы алгебраических уравнений

$$\begin{cases} 20 \cdot b_0 + 1708 \cdot b_1 = 11023 \\ 1708 \cdot b_0 + 391170 \cdot b_1 = 2390255 \end{cases}$$

позволяет получить оценки коэффициентов уравнений регрессии:

□(x)= b_0+ $b_1 \cdot x_i$=46,74+5,906·x_i

Графическая интерпретация исходных данных и регрессионной зависимости расхода электроэнергии от получасового максимума нагрузки приведена на рис.7.2

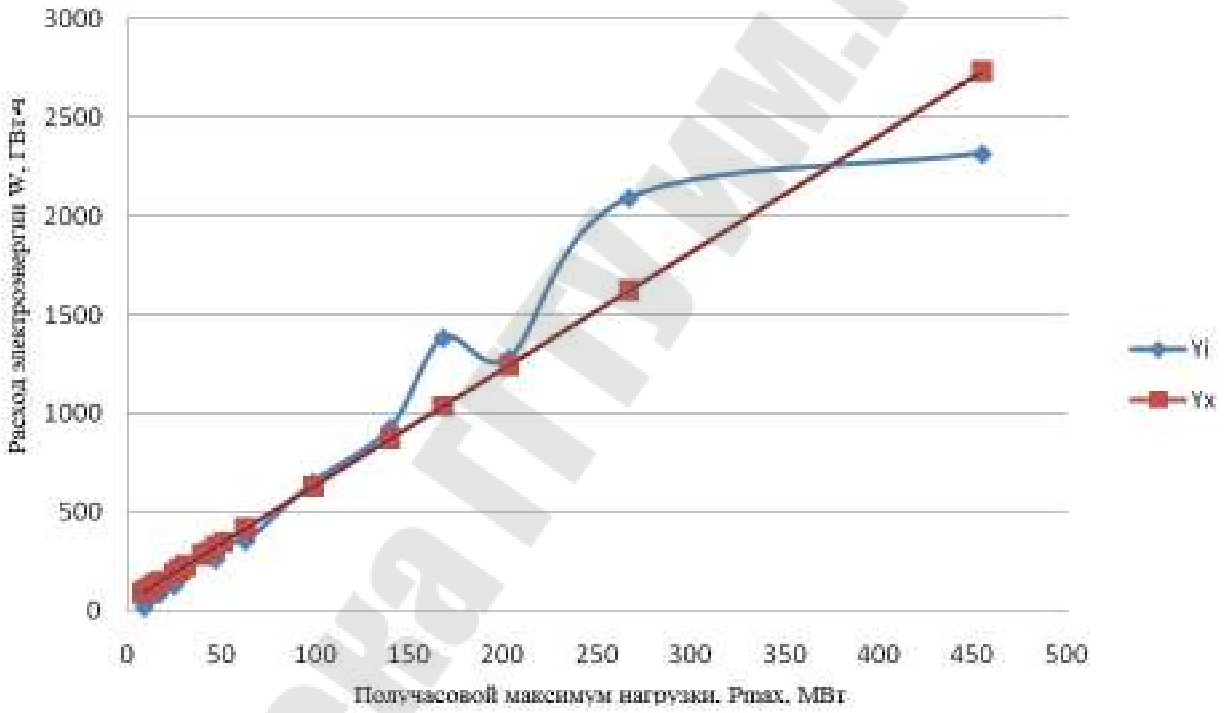


Рис. 7.2 Графическая интерпретация регрессионной зависимости расхода электроэнергии от получасового максимума нагрузки

Общая дисперсия определена в работе № 6. Для определения теоретического корреляционного отношения или индекса корреляции по формуле (7.4) воспользуемся результатами расчета, представленными в таблице (7.1):

$$R = \sqrt{\frac{\sigma_y^2 - \sigma_{y-y_x}^2}{\sigma_y^2}}. \tag{7.4}$$

$$R = \sqrt{\frac{455397{,}33 - 27507{,}73}{455397{,}33}} = 0{,}969328.$$

Вывод: теоретическое корреляционное отношение R хорошо согласуется с результатом, полученным в работе №6, □=0,9589.

7.5 Контрольные вопросы

1. Приведите выражение для функции регрессии одной случайной величины на другую.
2. Поясните элементы, входящие в функцию регрессии.
3. Чем располагают при оценке линейной (квадратичной, показательной и т.д.) функции регрессии?
4. Что определяет точность прогноза по функции регрессии $M(Y|_{X=x})$?
5. Дайте определение *кривой регрессии Y по X.*
6. Предпосылки регрессионного анализа.
7. Укажите виды зависимостей $\varphi(x)$, которые линейны по оцениваемым параметрам?
8. Укажите виды зависимостей $\varphi(x)$, которые не линейны по переменным x?
9. Какие параметры моделей уравнений регрессии подлежат оценке?
10. Из какого условия определяют оценки параметров уравнения регрессии?
11. В чем заключается метод наименьших квадратов?

12. Как определяется система нормальных уравнения

ЛАБОРАТОРНАЯ РАБОТА №8
Логические схемы и уравнения на безотказность и отказ в системах электроснабжения

8.1 Цель работы

Усвоить логические рассуждения на безотказность и отказ. По технологическим схемам составить логические уравнения на безотказность и отказ.

8.2 Краткие теоретические сведения

Надежность объекта систем электроснабжения характеризуется количественными показателями, имеющими вероятностный смысл.

Перечислим некоторые показатели:

$R(t)$ – интегральная функция распределения вероятностей безотказной работы:

$$R(t) = P(t_{\text{отк}} > t^*) = \frac{\mathrm{n}(t_{\text{отк}} > t^*)}{N(t=0)} = \begin{cases} 0, & t^* = 0 \\ 1, & t^* = \infty. \end{cases}$$

$R(t)$ численно равна доле начального количества объектов $N(t = 0)$, не отказавших до произвольного, но фиксированного момента времени t^* объектов $\mathrm{n}(t_{\text{отк}} > t^*)$;

$F(t)$ – интегральная функция распределения вероятностей отказа:

$$F(t) = P(t_{\text{отк}} \le t^*) = \frac{\mathrm{n}(t_{\text{отк}} \le t^*)}{N(t=0)} = \begin{cases} 0, & t^* = 0 \\ 1, & t^* = \infty. \end{cases}$$

$F(t)$ численно равна доле начального количества объектов $N(t = 0)$, отказавших до произвольного, но фиксированного момента времени t^* объектов $\mathrm{n}(t_{\text{отк}} \le t^*)$.

Составление уравнений для определения безотказной работы потребителя предполагает понимание логических рассуждений по безотказности и по отказу.

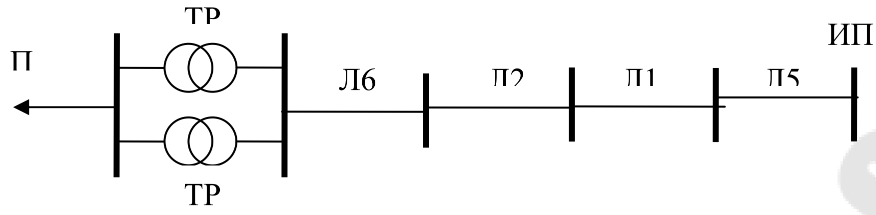
Логические схемы расчета надежности получаются из технологических схем на основе логических рассуждений, ведущих к цели расчета надежности. Логические рассуждения в расчетах надежности бывают двух типов – "по безотказности" и "по отказу". Строго говоря, расчет надежности возможен при использовании любого одного типа рассуждения. Два их типа необходимы только для того, чтобы выбрать простейшую логическую схему и, соответственно, простейшие расчетные соотношения.

Практически используются два трафарета логических рассуждений – "по теореме умножения вероятностей (союз И)" и "по теореме сложения вероятностей (союз ИЛИ)", причем, оба трафарета применимы к обоим типам логических рассуждений – "по безотказности" и по "отказу".

С учетом изложенного, логические рассуждения при составлении логических схем расчета надежности по технологическим схемам систем электроснабжения имеют следующий обобщенный вид: "система не откажет

(откажет), если не откажет (откажет) этот ее элемент И (ИЛИ) тот, И (ИЛИ) вот этот, И (ИЛИ) …" и так далее.

Приведем конфигурацию системы электроснабжения, представляющую технологическую схему относительно узла $П_i$ с одним источником питания, для которой необходимо составить логические схемы и логические уравнения при определении безотказной работы потребителя.



Уравнения безотказной работы потребителя составляются на основании логических схем, полученных на основе логических рассуждений, ведущих к цели расчета надежности. Логические рассуждения в расчетах надежности бывают двух типов: - «по безотказности» и «по отказу».

Элементы системы электроснабжения в логических схемах расчета надежности будем обозначать символами R_i при логических рассуждениях «по безотказности» и символами F_i – «по отказу».

Для расчета безотказной работы потребителя, исходную схему (технологическую) заменим логической по условию «безотказной» работы, рис. 8.1, а по условию "отказа" – рис. 8.2.

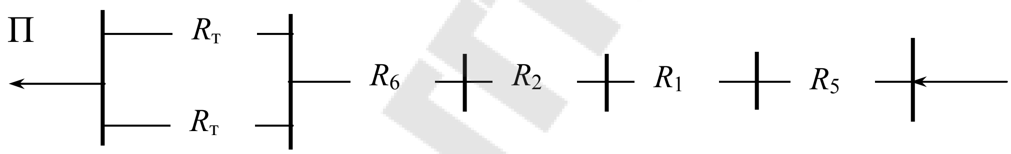


Рис. 8.1 Логическая схема электроснабжения по условию безотказной работы

Уравнение безотказной работы потребителя по условию "безотказности" в соответствии с логической схемой рис. 8.1 примет вид

$$R_\Sigma = R_5 \cdot R_1 \cdot R_2 \cdot R_6 \cdot (R_т + R_т - R_т \cdot R_т).$$

Если предположить, что трансформаторы однотипные, то R_Σ преобразуется:

$$R_\Sigma = R_5 \cdot R_1 \cdot R_2 \cdot R_6 \cdot (2R_т - R_т^2).$$

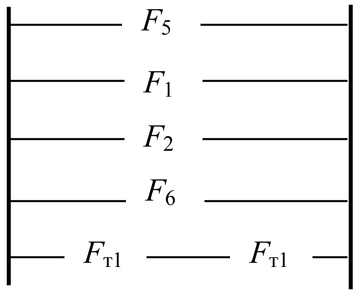


Рис. 8.2 Логическая схема электроснабжения по условию отказа

Уравнение по условию отказа в соответствии с логической схемой рис. 8.2 примет вид:

$$F_\Sigma = \left[(F_5 + F_1 - F_1F_5) + (F_2 + F_6 - F_2F_6) - (F_5 + F_1 - F_1F_5)(F_2 + F_6 - F_2F_6)\right] + F_{\text{т}}^2 -$$
$$- F_{\text{т}}^2\left[(F_5 + F_1 - F_1F_5) + (F_2 + F_6 - F_2F_6) - (F_5 + F_1 - F_1F_5)(F_2 + F_6 - F_2F_6)\right].$$

После алгебраических преобразований получим:

$$F_\Sigma = F_5 + F_1 + F_2 + F_6 + F_{\text{т}}^2 - F_5 \cdot F_1 - F_5 \cdot F_2 - F_5 \cdot F_6 - F_5 \cdot F_{\text{т}}^2 - F_1 \cdot F_2 - F_1 \cdot F_6 -$$
$$- F_1 \cdot F_{\text{т}}^2 - F_2 \cdot F_6 - F_2 \cdot F_{\text{т}}^2 - F_6 \cdot F_{\text{т}}^2 + F_1 \cdot F_2 \cdot F_5 + F_1 \cdot F_5 \cdot F_6 + F_1 \cdot F_5 \cdot F_{\text{т}}^2 +$$
$$+ F_1 \cdot F_6 \cdot F_{\text{т}}^2 + F_2 \cdot F_5 \cdot F_6 + F_2 \cdot F_5 \cdot F_{\text{т}}^2 + F_2 \cdot F_6 \cdot F_{\text{т}}^2 + F_5 \cdot F_6 \cdot F_{\text{т}}^2 - F_1 \cdot F_5 \cdot F_{\text{т}}^2 -$$
$$- F_1 \cdot F_2 \cdot F_5 \cdot F_6 - F_1 \cdot F_2 \cdot F_5 \cdot F_{\text{т}}^2 - F_1 \cdot F_5 \cdot F_6 \cdot F_{\text{т}}^2 - F_2 \cdot F_5 \cdot F_6 \cdot F_{\text{т}}^2 -$$
$$- F_1 \cdot F_2 \cdot F_6 \cdot F_{\text{т}}^2 + F_1 \cdot F_2 \cdot F_5 \cdot F_6 \cdot F_{\text{т}}^2.$$

Поскольку состояния "безотказность" и "отказ" противоположны друг другу, то $R_\Sigma + F_\Sigma = 1$.

Проверим правильность составленных уравнений.

Выразим $R_i = 1\text{-}F_i$ и поставим в R_Σ:

$$1 - F_\Sigma = (1 - F_5)(1 - F_1)(1 - F_2)(1 - F_6)\left[2(1 - F_{\text{т}}) - (1 - F_{\text{т}})^2\right] =$$
$$= (1 - F_5 - F_1 + F_1F_5)(1 - F_2 - F_6 + F_2F_6)\left[2 - 2F_{\text{т}} - 1 + 2F_{\text{т}} - F_{\text{т}}^2\right] =$$
$$= 1 - \big(F_5 + F_1 + F_2 + F_6 - F_5 \cdot F_1 - F_5 \cdot F_2 - F_5 \cdot F_6 - F_5 \cdot F_{\text{т}}^2 - F_1 \cdot F_2 -$$
$$- F_1 \cdot F_6 - F_1 \cdot F_{\text{т}}^2 - F_2 \cdot F_6 - F_2 \cdot F_{\text{т}}^2 - F_6 \cdot F_{\text{т}}^2 + F_1 \cdot F_2 \cdot F_5 + F_1 \cdot F_5 \cdot F_6 +$$
$$+ F_1 \cdot F_5 \cdot F_{\text{т}}^2 + F_1 \cdot F_2 \cdot F_6 + F_1 \cdot F_2 \cdot F_{\text{т}}^2 + F_1 \cdot F_6 \cdot F_{\text{т}}^2 + F_2 \cdot F_5 \cdot F_6 + F_2 \cdot F_5 \cdot F_{\text{т}}^2 +$$
$$+ F_2 \cdot F_6 \cdot F_{\text{т}}^2 + F_5 \cdot F_6 \cdot F_{\text{т}}^2 - F_1 \cdot F_2 \cdot F_5 \cdot F_6 - F_1 \cdot F_2 \cdot F_5 \cdot F_{\text{т}}^2 - F_1 \cdot F_5 \cdot F_6 \cdot F_{\text{т}}^2 -$$
$$- F_2 \cdot F_5 \cdot F_6 \cdot F_{\text{т}}^2 - F_1 \cdot F_2 \cdot F_6 \cdot F_{\text{т}}^2 + F_1 \cdot F_2 \cdot F_5 \cdot F_6 \cdot F_{\text{т}}^2\big).$$

Задание студенту по технологической конфигурации электрической сети определяет преподаватель.

8.3 Контрольные вопросы

1. Приведите качественные определения основных показателей надежности.

2. Какими количественными показателями характеризуется надежность элемента, система электроснабжения?

3. Дайте определение интегральной функции распределения вероятностей безотказной работы.

4. Дайте определение интегральной функции распределения вероятностей отказа.

5. Какие типы логических рассуждений бывают в расчетах надежности?

6. Что представляет собой технологическая схема в расчетах надежности?

7. Какие теоремы теории вероятностей используются при логических рассуждениях «по безотказности»?

8. Какие теоремы теории вероятностей используются при логических рассуждениях «по отказу»?

9. Приведите обобщенный вид логических рассуждений при составлении логических схем расчета надежности (применительно индивидуального задания).

10. На основании чего составляются уравнения безотказной работы потребителя?

11. Приведите символы, используемые в логических схемах и уравнениях, при расчетах надежности.

12. Как проверяется правильность составленных уравнений «безотказность» и «отказ»?

ЛИТЕРАТУРА

1. Электрические системы. Математические задачи электроэнергетики / В. А. Веников [и др.]. – Москва : Высш. шк., 1981. – 288 с.
2. Математическая статистика/В.М.Иванова, В.Н. Калинина [и др.] – Москва : Высш. шк., 1981. – 371 с.
3. Вентцель, Е. С. Теория вероятностей / Е. С. Вентцель. – Москва: Наука, 1969. – 576 с.
4. Колесникова И.Н., Круглякова Д.В. Статистика – Минск: Вышэйшая школа, 2011. – 285с.
5. Шундалов Б.М. Статистика. Общая теория. – Минск: ИВЦ Минфина, 2006. – 288с.

Содержание

МАТЕМАТИЧЕСКИЕ ЗАДАЧИ ЭЛЕКТРОЭНЕРГЕТИКИ

ПОСОБИЕ
для студентов специальности
1-43 01 03 «Электроснабжение»
дневной и заочной форм обучения

Составители: **Попова** Ольга Михайловна
Алферова Тамара Викторовна
Алферов Александр Александрович

Подписано к размещению в электронную библиотеку
ГГТУ им. П. О. Сухого в качестве электронного
учебно-методического документа 12.02.16.

Рег. № 26Е.
http://www.gstu.by