

А.М. Сазонова

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ  
И МАТЕМАТИЧЕСКАЯ  
СТАТИСТИКА**

Могилев 2011

Электронный архив библиотеки МГУ имени А.А. Кулешова

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ**

**УЧРЕЖДЕНИЕ ОБРАЗОВАНИЯ  
«МОГИЛЕВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
им. А.А. КУЛЕШОВА»**

**А.М. Сазонова**

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ  
И МАТЕМАТИЧЕСКАЯ  
СТАТИСТИКА**

**Контрольные задания  
и методические рекомендации**



**Могилев 2011**

УДК 519.21(075)  
ББК 22.3  
С12

*Печатается по решению редакционно-издательского совета  
УО «МГУ им. А.А. Кулешова»*

### **Рецензент**

кандидат физико-математических наук доцент  
УО «Могилевский государственный университет им. А.А. Кулешова»  
*В.Н. Борбат*

**Сазонова, А.М.**

**С12** Теория вероятностей и математическая статистика: контрольные задания и методические рекомендации / А.М. Сазонова. – Могилев: УО «МГУ им. А.А. Кулешова», 2011. – 28 с.

В издании отражена практика преподавания предмета в Могилевском государственном университете им. А.А. Кулешова.

Предназначается для студентов вузов, обучающихся по экономическим специальностям, по социологии, преподавателей и слушателей курсов повышения квалификации, применяющих в своей практике вероятностные и статистические методы.

**УДК 519.21(075)  
ББК 22.3**

© Сазонова А.М., 2011  
© УО «МГУ им. А.А. Кулешова», 2011

## ВЫБОР КОНТРОЛЬНОЙ РАБОТЫ

Номер варианта контрольной работы совпадает с двумя последними цифрами номера зачетной книжки. В контрольную работу входит 6 задач, содержание данных в которых студент формирует самостоятельно, согласно следующим пояснениям.

### Пояснения к обозначениям:

$N$  – номер варианта равен последним двум цифрам зачетной книжки.

Выражение  $N \bmod K$  равно остатку от деления целого числа  $N$  на число  $K$ . Если число  $N$  на число  $K$  делится без остатка, тогда  $N \bmod K = 0$ . Если  $N < K$ , тогда  $N \bmod K = N$ .

### ПРИМЕРЫ:

1. Найти значение выражения  $32 \bmod 5$ . Для этого найдем остаток при делении числа 32 на число 5.

$32:5 = 5 \cdot 6 + 2 = 30 + 2$ . Остаток от деления равен 2. Следовательно,  $32 \bmod 5 = 2$ .

2. Найти значение выражения  $28 \bmod 7$ . Так как 28 делится на 7 без остатка, следовательно,  $28 \bmod 7 = 0$ .

3. Найти значение выражения  $6 \bmod 9$ .

Так как  $6 < 9$ , следовательно,  $6 \bmod 9 = 6$ .

# ОБРАЗЕЦ ОФОРМЛЕНИЯ ТИТУЛЬНОГО ЛИСТА КОНТРОЛЬНОЙ РАБОТЫ

УО «Могилевский государственный университет им. А.А. Кулешова»  
Кафедра алгебры и геометрии

Контрольная работа  
по теории вероятностей и математической статистике

студента 2 курса ФЭиП, гр.ЭУП-111  
номер зачетной книжки № 001158  
Иванова Ивана Ивановича

Место проживания:  
212030, г. Могилев,  
ул. Первомайская, д. 31, кв.8,  
дом. тел. 23 23 23,  
служ. тел. 26 16 16 ,  
моб. тел. 8 0296 777 888

# ПРОГРАММА ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

Правила суммы и произведения в комбинаторике.

Размещение с повторениями.

Размещение без повторения.

Перестановки.

Сочетание.

Предмет теории вероятностей. События, виды событий.

Действия над событиями.

Аксиоматическое определение вероятности.

Классическое определение вероятности.

Статистическое определение вероятности. Аналог вероятности в статистике. Теорема Бернулли.

Геометрическое определение вероятности.

Теорема сложения вероятностей.

Условные вероятности. Независимые события.

Теорема умножения вероятностей.

Формула полной вероятности.

Формула Байеса.

Формула Бернулли.

Вероятность наступления события в схеме Бернулли менее, более, не менее, не более  $m$  раз в  $n$  испытаниях, хотя бы один раз, после  $k$  неудач.

Наивероятнейшее число наступивших событий в схеме Бернулли.

Локальная теорема Муавра-Лапласа.

Формула Пуассона.

Интегральная теорема Муавра-Лапласа.

Случайная величина. Виды случайных величин. Функция распределения случайной величины.

Дискретная случайная величина, ее распределение.

Независимые случайные величины. Сумма случайных величин.

Произведение случайной величины на константу. Произведение случайных величин.

Математическое ожидание случайной величины, его свойства, его аналог в статистике.

Дисперсия случайной величины, ее свойства.

Среднее квадратическое отклонение, его аналог в статистике.

Функция распределения дискретной случайной величины, ее свойства, график. Аналог в статистике.

Непрерывная случайная величина, ее распределение, плотность распределения, их свойства.

Биномиальное распределение. Закон Пуассона.

Геометрическое, гипергеометрическое распределения.

Равномерное распределение.

Показательное (экспоненциальное) распределение.

Нормальное распределение. Стандартное нормальное распределение. Функция Лапласа. Попадание случайной величины в заданный интервал.

Дискретная двумерная случайная величина, ее таблица распределения. Ряды распределения составляющих.

Функция распределения двумерной случайной величины, свойства, геометрическая интерпретация.

Независимые случайные величины, коэффициент корреляции, его свойства.

Условные математические ожидания. Линии регрессии.

Предельные теоремы. Закон больших чисел. Центральная предельная теорема.

Предмет математической статистики. Генеральная совокупность.

Выборочная совокупность. Выборочный метод.

Вариационный ряд (дискретный, непрерывный (интервальный)).

Изображение вариационных рядов (полигон, гистограмма).

Кумулята, эмпирическая функция распределения, эмпирическая плотность распределения.

Числовые характеристики вариационных рядов.

Мода. Медиана. Асимметрия, эксцесс.

Точечные оценки параметров. Определение, требования к оценке.

Метод моментов. Метод наибольшего правдоподобия.

Основные статистические распределения. Критические точки.

Интервальные оценки параметров. Предельная ошибка выборки.

Доверительные интервалы для генеральной средней и генеральной доли признака.

Доверительный интервал для дисперсии.

Объем выборки. Формула необходимого объема выборки.

Статистическая гипотеза. Ошибки первого и второго рода, мощность критерия. Критическая область.

Гипотеза о численной величине средней генеральной совокупности (дисперсия известна).

Гипотеза о численной величине средней генеральной совокупности (дисперсия неизвестна).

Гипотеза о числовом значении дисперсии.

Гипотеза о числовом значении доли признака в генеральной совокупности.

Гипотеза о равенстве двух средних значений.

Однофакторный дисперсионный анализ.

Критерии согласия.

Модели и основные понятия корреляционного и регрессионного анализа.

Коэффициент корреляции и его свойства.

Линейная корреляционная зависимость и прямые регрессии. Метод наименьших квадратов.

Корреляционное отношение и его свойства.

Анализ соответствия регрессионной модели статистическим данным.

## КОНТРОЛЬНЫЕ ЗАДАНИЯ

### ЗАДАЧА № 1.

Три стрелка одновременно производят залп по мишени. Вероятность того, что 1-й стрелок попадет в цель, равна  $p_1 = 1 - \frac{N \bmod 7 + 1}{100}$ , вероятность того, что 2-й стрелок попадет в цель, равна  $p_2 = 1 - \frac{N \bmod 5 + 1}{10}$ , вероятность того, что 3-й стрелок попадет в цель, равна  $p_3 = 1 - \frac{N \bmod 3 + 1}{10}$ .

Найти вероятность того, что:

1. Все попадут в цель.
2. Менее двух промахнутся.
3. Только один попадет.
4. Хотя бы один попадет.

### ЗАДАЧА № 2.

Монета бросается 17 раз. Найти вероятность того, что «герб» появится:

1. Ровно  $k$  раз.
2. Не более  $k$  раз.
3. Не менее  $k$  раз, где  $k = N \bmod 4 + 2$ .

### ЗАДАЧА № 3.

В партии находятся 15 деталей. Из них  $M = N \bmod 3 + 10$  стандартных. Наудачу извлекаются  $n = N \bmod 6 + 4$  деталей. Составить закон распределения случайной величины  $X$  – числа стандартных деталей среди отобранных.

### ЗАДАЧА № 4.

По данному статистическому материалу опыта требуется:

- 1) составить статистический ряд распределения;
- 2) составить интервальный вариационный ряд относительных частот, разбив размах варьирования на  $k$  интервалов; найти эмпирическую функцию распределения и построить ее график;
- 3) построить гистограмму и полигон относительных частот;

4) проверить гипотезу о нормальном распределении генеральной совокупности;

В случае нормального распределения по заданному уровню значимости  $\alpha$ :

5) найти интервальные оценки параметров распределения;

6) проверить нулевую гипотезу  $H_0: a = a_0$  о математическом ожидании при альтернативной гипотезе  $H_1: a \neq a_0$  ( $a > a_0, a < a_0$ );

7) проверить нулевую гипотезу  $H_0: \sigma^2 = \sigma_0^2$  о дисперсии против альтернативной  $H_1: \sigma^2 \neq \sigma_0^2$  ( $\sigma^2 > \sigma_0^2, \sigma^2 < \sigma_0^2$ ).

Статистические данные исследуемой генеральной совокупности представлены в таблице:

|    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|
| 53 | 60 | 72 | 58 | 46 | 44 | 54 | 60 | 80 | 66 | 74 | 46 |
| 41 | 58 | 80 | 66 | 58 | 80 | 41 | 58 | 79 | 41 | 58 | 54 |
| 48 | 80 | 70 | 62 | 62 | 62 | 54 | 68 | 78 | 52 | 81 | 62 |
| 41 | 72 | 60 | 72 | 75 | 54 | 75 | 54 | 75 | 41 | 76 | 82 |
| 75 | 52 | 40 | 41 | 54 | 66 | 40 | 44 | 78 | 80 | 74 | 72 |

|                                      |   |   |                                      |   |                                      |   |   |   |  |
|--------------------------------------|---|---|--------------------------------------|---|--------------------------------------|---|---|---|--|
| вариант<br>данных $i$<br>$N \bmod 9$ | 1                                       | 2                                       | 3                                    | 4                                       | 5                                    | 6                                       | 7                                       | 8                                       | 9  |
|                                      | $i \equiv N \bmod 9 + 1$                |   |                                      |   |                                      |   |   |   |  |
| $k$                                  | 10                                      | 8                                       | 9                                    | 8                                       | 10                                   | 9                                       | 10                                      | 9                                       | 10   |
| $\alpha$ -уровень<br>значимости      | 0,01                                    | 0,01                                    | 0,05                                 | 0,05                                    | 0,01                                 | 0,01                                    | 0,05                                    | 0,05                                    | 0,05                                       |
| $a_0$                                | $a_1$                                   | $a_2$                                   | $a_2$                                | $a_1$                                   | $a_1$                                | $a_2$                                   | $a_1$                                   | $a_2$                                   | $a_1$                                      |
| $\sigma_0^2$                         | $S_2^2$                                 | $S_1^2$                                 | $S_2^2$                              | $S_2^2$                                 | $S_1^2$                              | $S_2^2$                                 | $S_2^2$                                 | $S_2^2$                                 | $S_2^2$                                    |
| гипотезы<br>$H_1$                    | $a \neq a_0$<br>$\sigma^2 < \sigma_0^2$ | $a > a_0$<br>$\sigma^2 \neq \sigma_0^2$ | $a < a_0$<br>$\sigma^2 > \sigma_0^2$ | $a \neq a_0$<br>$\sigma^2 < \sigma_0^2$ | $a < a_0$<br>$\sigma^2 < \sigma_0^2$ | $a > a_0$<br>$\sigma^2 \neq \sigma_0^2$ | $a \neq a_0$<br>$\sigma^2 > \sigma_0^2$ | $a \neq a_0$<br>$\sigma^2 < \sigma_0^2$ | $a \neq a_0$<br>$\sigma^2 \neq \sigma_0^2$ |

$a_1, S_1$  – соответственно значения доверительного интервала  $a$  и  $\sigma$  в левом конце,  $a_2, S_2$  – в правом конце. Объем выборки  $n = 50$ ;  $i$  – порядковый номер  $x_i$ , от которого ведется отсчет значений случайной величины  $X$ , считая по строке,  $i = N \bmod 9 + 1$ .

### ЗАДАЧА № 5.

Имеется следующая частотная таблица, построенная на основе результатов опроса студентов трех факультетов относительно того, сколько часов в неделю они проводят в библиотеке.

| Временные интервалы | Число студентов, которые проводят в библиотеке указанное количество времени |                |                |
|---------------------|---|----------------|----------------|
|                     | Факультет А   | Факультет В    | Факультет С    |
| 0 – 4               | 11  | 8              | 10             |
| 4 – 8               | 18  | $N \bmod 3+20$ | 19             |
| 8 – 12              | 13  | $N \bmod 4+17$ | $N \bmod 3+21$ |
| 12 – 16             | $N \bmod 5+4$   | $N \bmod 3+10$ | 15             |

Можно ли утверждать, что количество времени, проведенного студентом в библиотеке, не зависит от того, на каком факультете учится студент?

### ЗАДАЧА № 6.

В корреляционной таблице представлены данные  $n$  предприятий по стоимости основных промышленно-производственных фондов ( $X$  млн \$) и объема выпуска продукции ( $Y$  т):

| Объем выпуска продукции ( $Y$ ) \\<br>Стоимость основных фондов ( $X$ ) | Объем выпуска продукции ( $Y$ ) |     |          |          |      | $m_{.j}$                               |
|---|---------------------------------|-----|----------|----------|------|--|
|   | 0-2                             | 2-4 | 4-6      | 6-8      | 8-10 |  |
| 0-0,2   | 2                               | 2   |          |          |      | 4                                      |
| 0,2-0,4   | 2                               | 8   | $m_{23}$ |          |      | $m_{2.}$                               |
| 0,4-0,6   |                                 | 2   | $m_{33}$ | 8        |      | $m_{3.}$                               |
| 0,6-0,8   |                                 |     | 4        | 3        |      | 7                                      |
| 0,8-1,0   |                                 |     |          | $m_{54}$ | 2    | $m_{5.}$                               |
| 1,0-1,2   |                                 |     |          |          | 3    | 3                                      |
| $m_{.j}$  | 4                               | 12  | $m_{.3}$ | $m_{.4}$ | 5    | $\sum_{i=1}^5 \sum_{j=1}^6 m_{ij} = n$ |

где  $m_{23} = N \bmod 3 + 12$ ,  $m_{33} = N \bmod 4 + 18$ ,  $m_{54} = N \bmod 3 + 2$ .

По данным корреляционной таблицы найти:

1. Числовые характеристики случайных величин  $X$  и  $Y$ .
2. Выборочный коэффициент корреляции  $\hat{r}$ . Проверить значимость коэффициента корреляции при уровне значимости  $\alpha = \frac{N \bmod 2 + 1}{20}$ .
3. Вычислить коэффициент детерминации и объяснить его смысл.
4. Уравнение линейной регрессии  $Y$  на  $X$ . Построить корреляционное поле и график линейного уравнения регрессии  $Y$  на  $X$ .

5. Определить степень рассеивания экспериментальных точек вокруг линии регрессии.

## ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ ВАРИАНТА

### ЗАДАЧА № 1.

Три стрелка одновременно производят залп по мишени. Вероятность того, что первый стрелок попадет в цель, равна  $p_1 = 0,8$ , вероятность того, что второй стрелок попадет в цель, равна  $p_2 = 0,9$ , вероятность того, что третий стрелок попадет в цель, равна  $p_3 = 0,7$ . Найти вероятность того, что:

1. Все попадут в цель.
2. Менее двух промахнутся.
3. Только один попадет.
4. Хотя бы один попадет

**Решение.** Вероятности промаха первым, вторым, третьим стрелком соответственно равны:  $q_1 = 1 - p_1 = 1 - 0,8 = 0,2$ ;  $q_2 = 1 - p_2 = 1 - 0,9 = 0,1$ ;  $q_3 = 1 - p_3 = 1 - 0,7 = 0,3$ .

1. Найти вероятность того, что все попадут в цель, – событие  $A$ .

Обозначим через  $A_1, A_2, A_3$  элементарные (независимые) события: попали в цель соответственно первый, второй и третий стрелки. Тогда  $A = A_1 A_2 A_3$ .  $P(A) = P(A_1 A_2 A_3) = P(A_1) P(A_2) P(A_3) = p_1 p_2 p_3 = 0,8 \cdot 0,9 \cdot 0,7 = 0,504$ .

2. Найти вероятность того, что менее двух промахнутся – событие  $B$  (то есть промахнется один или ноль стрелков).

Тогда  $B = \bar{A}_1 A_2 A_3 + A_1 \bar{A}_2 A_3 + A_1 A_2 \bar{A}_3 + A_1 A_2 A_3$ .

$$\begin{aligned} P(B) &= P(\bar{A}_1 A_2 A_3 + A_1 \bar{A}_2 A_3 + A_1 A_2 \bar{A}_3 + A_1 A_2 A_3) = \\ &= q_1 p_2 p_3 + p_1 q_2 p_3 + p_1 p_2 q_3 + p_1 p_2 p_3 = 0,2 \cdot 0,9 \cdot 0,7 + 0,8 \cdot 0,1 \cdot 0,7 + \\ &+ 0,8 \cdot 0,9 \cdot 0,3 + 0,8 \cdot 0,9 \cdot 0,7 = 0,902. \end{aligned}$$

3. Найти вероятность того, что попадет в цель только один стрелок, – событие  $C$ .  $C = A_1 \bar{A}_2 \bar{A}_3 + \bar{A}_1 A_2 \bar{A}_3 + \bar{A}_1 \bar{A}_2 A_3$ .  $P(C) = P(A_1 \bar{A}_2 \bar{A}_3 + \bar{A}_1 A_2 \bar{A}_3 + \bar{A}_1 \bar{A}_2 A_3) = p_1 q_2 q_3 + q_1 p_2 q_3 + q_1 q_2 p_3 = 0,8 \cdot 0,1 \cdot 0,3 + 0,2 \cdot 0,9 \cdot 0,3 + 0,2 \cdot 0,1 \cdot 0,7 = 0,092$ .

4. Найти вероятность того, что хотя бы один попадет, – событие  $D$ . Событие представляет собой всевозможные события, кроме

события, ему противоположного,  $\bar{D}$  – «никто не попадет в цель». Тогда  $P(D) = 1 - P(\bar{D}) = 1 - P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = 1 - 0,2 \cdot 0,1 \cdot 0,3 = 0,006$ .

### ЗАДАЧА № 2.

Монета бросается 8 раз. Найти вероятность того, что «герб» появится:

1. Ровно 6 раз.
2. Не более 2 раз.
3. Не менее 6 раз.

**Решение.** Испытание проводится  $n = 8$  раз, в каждом из которых событие  $A$  («появление герба») появляется с одной и той же вероятностью  $p = 0,5$ . Таким образом, имеем схему Бернулли.

1. Найти вероятность того, что событие  $A$  появится точно 6 раз. По формуле Бернулли

$$P_8(6) = C_8^6 p^6 (1-p)^{8-6} = \frac{8!}{6!2!} 0,5^6 0,5^2 = 28 \cdot \left(\frac{1}{2}\right)^8 = \frac{28}{216} = \frac{7}{54}.$$

2. Найти вероятность того, что событие  $A$  появится не более двух раз (то есть ноль раз, или один раз, или два раза):  $P_8(0) + P_8(1) + P_8(2) =$   
 $= C_8^0 p^0 (1-p)^8 + C_8^1 p^1 (1-p)^7 + C_8^2 p^2 (1-p)^6 = 1 \cdot \left(\frac{1}{2}\right)^8 + 8 \cdot \left(\frac{1}{2}\right)^8 + 28 \cdot \left(\frac{1}{2}\right)^8 =$   
 $= \frac{37}{216}.$

3. Найти вероятность того, что событие  $A$  появится не менее 6 раз (то есть шесть раз, или семь раз, или восемь раз):  $P_8(6) + P_8(7) +$   
 $+ P_8(8) = C_8^6 p^6 (1-p)^2 + C_8^7 p^7 (1-p) + C_8^8 p^8 (1-p)^0 = 28 \cdot \left(\frac{1}{2}\right)^8 + 8 \cdot \left(\frac{1}{2}\right)^8 +$   
 $+ 1 \cdot \left(\frac{1}{2}\right)^8 = \frac{37}{216}.$

### ЗАДАЧА № 3.

В партии находится 8 деталей. Из них 5 стандартных. Наудачу извлекаются 3 детали. Составить закон распределения случайной величины  $X$  – числа стандартных деталей среди отобранных.

**Решение.**

Случайная величина  $X$  может принимать значения  $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$ . Найдем соответствующие вероятности:

$$P(X=0) = \frac{C_5^0 C_3^3}{C_8^3} = \frac{1}{56}.$$

$$P(X=1) = \frac{C_5^1 C_3^2}{C_8^3} = \frac{15}{56}.$$

$$P(X=2) = \frac{C_5^2 C_3^1}{C_8^3} = \frac{30}{56}.$$

$$P(X=3) = \frac{C_5^3 C_3^0}{C_8^3} = \frac{10}{56}.$$

Контроль:  $\frac{1}{56} + \frac{15}{56} + \frac{30}{56} + \frac{10}{56} = 1.$

Ряд распределения представляет собой следующую таблицу:

|       |      |       |       |       |
|-------|------|-------|-------|-------|
| $x_i$ | 0    | 1     | 2     | 3     |
| $p_i$ | 1/56 | 15/56 | 30/56 | 10/56 |

или

|       |      |       |       |      |
|-------|------|-------|-------|------|
| $x_i$ | 0    | 1     | 2     | 3    |
| $p_i$ | 1/56 | 15/56 | 15/28 | 5/28 |

#### ЗАДАЧА № 4.

По данному статистическому материалу опыта требуется:

- 1) составить статистический ряд распределения;
- 2) составить интервальный вариационный ряд относительных частот, разбив размах варьирования на  $k$  интервалов; найти эмпирическую функцию распределения и построить ее график;
- 3) построить гистограмму и полигон относительных частот;
- 4) проверить гипотезу о нормальном распределении генеральной совокупности;

В случае нормального распределения по заданному уровню значимости  $\alpha$ :

- 5) найти интервальные оценки параметров распределения;
- 6) проверить нулевую гипотезу  $H_0: a = a_0$  о математическом ожидании при альтернативной гипотезе  $H_1: a \neq a_0$  ( $a > a_0$ ,  $a < a_0$ );
- 7) проверить нулевую гипотезу  $H_0: \sigma^2 = \sigma_0^2$  о дисперсии против альтернативной  $H_1: \sigma^2 \neq \sigma_0^2$  ( $\sigma^2 > \sigma_0^2$ ,  $\sigma^2 < \sigma_0^2$ ).

### Решение.

Пусть имеем следующие статистические данные исследуемой генеральной совокупности:

|     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 201 | 202 | 214 | 208 | 210 | 209 | 212 | 213 | 206 | 209 |
| 209 | 206 | 202 | 211 | 209 | 211 | 208 | 213 | 209 | 211 |
| 211 | 209 | 207 | 208 | 210 | 214 | 203 | 209 | 207 | 210 |
| 205 | 211 | 206 | 210 | 208 | 211 | 208 | 205 | 212 | 207 |
| 209 | 207 | 213 | 207 | 210 | 215 | 204 | 209 | 207 | 210 |
| 212 | 210 | 207 | 204 | 206 | 208 | 207 | 210 | 211 | 208 |
| 206 | 205 | 209 | 207 | 211 | 205 | 209 | 208 | 207 | 204 |
| 209 | 212 | 208 | 210 | 208 | 206 | 207 | 211 | 212 | 207 |
| 206 | 208 | 206 | 213 | 212 | 209 | 215 | 206 | 213 | 205 |
| 208 | 209 | 206 | 204 | 209 | 203 | 217 | 208 | 210 | 215 |

Объем выборки  $n = 100$ .

1. Составить статистический ряд распределения:

|                             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Наблюдаемые значения СВ $X$ | 201 | 202 | 203 | 204 | 205 | 206 | 207 | 208 | 209 | 210 | 211 | 212 | 213 | 214 | 215 | 217 |
| Подсчет частот              | •   | •   | •   | ••  | ••• | ⊗   | ⊗   | ⊗   | ⊗   | ⊗   | ⊗   | ••• | ••• | •   | ••  | •   |
| Частоты ( $m_i$ )           | 1   | 2   | 2   | 4   | 5   | 10  | 12  | 13  | 15  | 10  | 9   | 6   | 5   | 2   | 3   | 1   |

Проверка  $\sum m_i = n = 100$  – несгруппированных значений.

2. Составить интервальный вариационный ряд относительных частот, разбив размах варьирования на  $k$  интервалов; найти эмпирическую функцию распределения и построить ее график.

Пусть  $k = 9$ . Размах варьирования  $R = x_{\max} - x_{\min} = 217 - 201 = 17$ . Длина интервала  $d = \frac{R}{k} = \frac{17}{9} = 1,9 \approx 2$ . Составим интервальный вариационный ряд относительных частот, учитывая условие, чтобы наименьшее значение варианта  $x_{\min}$  принадлежало первому промежутку, а наибольшее значение варианта  $x_{\max}$  принадлежало девятому промежутку:

$$x_{\min} \in [a_1; a_2), x_{\max} \in [a_9; a_{10})$$

| Интервалы                     | [200-202) | [202-204) | [204-206) | [206-208) | [208-210) | [210-212) | [212-214) | [214-216) | [216-218) |
|-------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Частота $m_i$                 | 1         | 4         | 9         | 22        | 28        | 19        | 11        | 5         | 1         |
| Относительная частота $m_i/n$ | 0,01      | 0,04      | 0,09      | 0,22      | 0,28      | 0,19      | 0,11      | 0,05      | 0,01      |

Проверяем:

$$\sum_{i=1}^{10} \frac{m_i}{n} = 0,01 + 0,04 + 0,09 + 0,22 + 0,28 + 0,19 + 0,11 + 0,05 + 0,01 = 1.$$

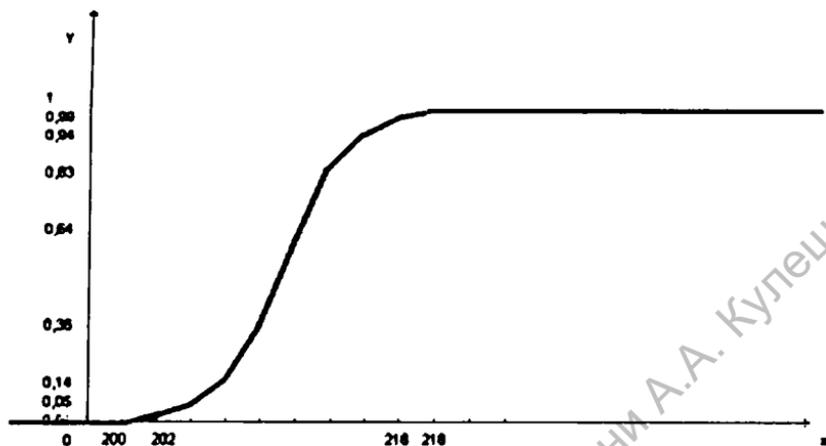
Для построения графика эмпирической функции распределения составляем таблицу:

| Интервалы   | $(-\infty; 200]$ | $(200; 202]$     | $(202; 204]$     | $(204; 206]$     |
|---|------------------|------------------|------------------|------------------|
| Накопленная относительная частота $\frac{m_x}{n}$ | 0                | 0,01             | $0,05=0,01+0,04$ | $0,14=0,05+0,09$ |
| Интервалы   | $(206; 208]$     | $(208; 210]$     | $(210; 212]$     | $(212; 214]$     |
| Накопленная относительная частота $\frac{m_x}{n}$ | $0,36=0,14+0,22$ | $0,64=0,36+0,28$ | $0,83=0,64+0,19$ | $0,94=0,83+0,11$ |
| Интервалы   | $(214; 216]$     | $(216; 218]$     | $(218; +\infty)$ |                  |
| Накопленная относительная частота $\frac{m_x}{n}$ | $0,99=0,94+0,05$ | $1=0,99+0,01$    | $1=1+0$          |                  |

Эмпирическая функция распределения:

$$F(x) = \begin{cases} 0 & x \leq 200; \\ 0,01 & \text{при } 200 < x \leq 202; \\ 0,05 & \text{при } 202 < x \leq 204; \\ 0,14 & \text{при } 204 < x \leq 206; \\ 0,36 & \text{при } 206 < x \leq 208; \\ 0,64 & \text{при } 208 < x \leq 210; \\ 0,83 & \text{при } 210 < x \leq 212; \\ 0,94 & \text{при } 212 < x \leq 214; \\ 0,99 & \text{при } 214 < x \leq 216; \\ 1 & \text{при } x > 216. \end{cases}$$

Для интервального ряда графиком эмпирической функции является ломаная, проходящая через точки  $(a_i; F(a_i))$ ,  $i=1, 2, \dots, k$ .



3. Построить гистограмму и полигон относительных частот.

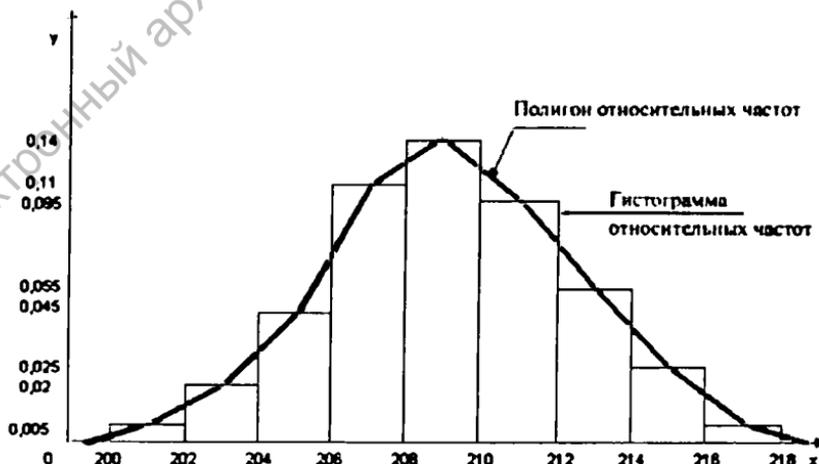
Для гистограммы на интервалах строим прямоугольники

высотой  $h_i = \frac{m_i}{nd}$ .

$$h_1 = \frac{0,01}{2} = 0,005; \quad h_2 = \frac{0,04}{2} = 0,02; \quad h_3 = \frac{0,09}{2} = 0,045; \quad h_4 = \frac{0,22}{2} = 0,07;$$

$$h_5 = \frac{0,28}{2} = 0,11; \quad h_6 = \frac{0,19}{2} = 0,095; \quad h_7 = \frac{0,11}{2} = 0,055; \quad h_8 = \frac{0,05}{2} = 0,025;$$

$$h_9 = \frac{0,01}{2} = 0,005;$$



Для полигона находим середины интервалов:

$$x_1 = \frac{200 + 202}{2} = 201, \quad x_2 = 203, \quad x_3 = 205, \quad x_4 = 207, \quad x_5 = 209, \quad x_6 = 211,$$

$x_7 = 213, \quad x_8 = 215, \quad x_9 = 217$  и соединяем ломаной точки с координа-

тами  $\left(x_i, \frac{m_i}{nd}\right), \quad i = \overline{1, 9}$ .

4. Проверить гипотезу о нормальном распределении генеральной совокупности.

Алгоритм проверки гипотезы:

- весь диапазон выборочных значений (вариант  $x_p$ , объем выборки  $n \geq 50$ ) разбивают на  $k (k \geq 5)$  полуинтервалов  $\Delta_i = [a_i, a_{i+1}), i = 1, 2, \dots, k$ ,

длиной  $h$ , где число  $m_i$  вариант в  $i$ -ом интервале и  $\sum_{i=1}^k m_i = n$ .

*Рекомендация:*  $a_1 = x_{\min} - 0,5 h, a_k \leq x_{\max} < a_{k+1}$ .

- Для каждого интервала вычисляют теоретические вероятности  $p_i$  попадания случайной величины  $X$ , растянув крайние интервалы вариационного ряда до бесконечности.  $p_i = P(X \in \Delta_i) = F(a_{i+1}) - F(a_i), \sum_{i=1}^k p_i = 1$ .

- Подсчитывают соответствующие теоретические частоты  $np_i$ , причем если для некоторых интервалов они меньше 5, то их объединяют с соседними, чтобы  $np_i \geq 5$ . Новое число интервалов обозначим  $k'$ .

- При  $n \rightarrow \infty$  статистика  $\sum_{i=1}^{k'} \frac{(m_i - np_i)^2}{np_i}$  имеет распределение  $\chi^2$  с  $\nu = k' - l - 1$  степенями свободы,  $l$  – число неизвестных параметров предполагаемой функции распределения  $F(x)$ , оцениваемых по результатам наблюдений (если все параметры предполагаемого закона известны точно, то  $l = 0$ ).

Величина  $\chi^2 = \sum_{i=1}^{k'} \frac{(m_i - np_i)^2}{np_i}$  называется критерием согласия  $\chi^2$ , или

критерием согласия Пирсона. Чем ближе к нулю наблюдаемое значение критерия, тем вероятнее, что нулевая гипотеза справедлива. Поэтому для проверки нулевой гипотезы используется **правосторонняя**

**критическая область.** Для заданного уровня значимости  $\alpha$  определяем ее  $[\chi_{\alpha, \nu}^2, +\infty)$ .

- Если расчетное значение критерия согласия  $\chi_{расчет}^2 < \chi_{\alpha, \nu}^2$ , то принимают нулевую гипотезу. В противном случае нулевая гипотеза отклоняется в пользу альтернативной.

Нормальное распределение имеет два параметра – математическое ожидание  $MX = a$  и дисперсия  $DX = \sigma^2$ . В задаче они не известны, поэтому найдем их оценки:

$$\hat{a} = \bar{x} = \frac{1}{100} (201 \cdot 1 + 203 \cdot 4 + 205 \cdot 9 + 207 \cdot 22 + 209 \cdot 28 + 211 \cdot 19 + 213 \cdot 11 + 215 \cdot 5 + 217 \cdot 1) = 209,08.$$

$$\hat{\sigma}^2 = s_{несмещ}^2 = \frac{1}{99} ((201 - 209,08)^2 \cdot 1 + (203 - 209,08)^2 \cdot 4 + (205 - 209,08)^2 \cdot 9 + (207 - 209,08)^2 \cdot 22 + (209 - 209,08)^2 \cdot 28 + (211 - 209,08)^2 \cdot 19 + (213 - 209,08)^2 \cdot 11 + (215 - 209,08)^2 \cdot 5 + (217 - 209,08)^2 \cdot 1) = 9,448080, \quad \hat{\sigma} = 3,0738.$$

Для определения расчетного значения критерия  $\chi_{расчет}^2$  составим таблицу в соответствии с алгоритмом:

| № i-го интервала | $a_i - a_{i-1}$ | $m_i$ | $F(a_i) = -1/2(1 - \alpha(\frac{a_i - \hat{a}}{\hat{\sigma}}))$ |  | $p_i = -F(a_{i+1}) - F(a_i)$ | $np_i$ | $\frac{\chi^2}{np_i}$        |
|------------------|-----------------|-------|---|--|------------------------------|--------|------------------------------|
|                  |                 |       | $-1/2(1 - \alpha(\frac{a_i - \hat{a}}{\hat{\sigma}}))$          | $-1/2(1 - \alpha(\frac{a_{i+1} - \hat{a}}{\hat{\sigma}}))$ |                              |        |                              |
| 1                | 200-202         | 1     | 0,0000  | 0,0081   | 0,0081                       | 0,81   | 0,20399                      |
| 2                | 202-204         | 4     |   |  |                              |        |                              |
| 3                | 204-206         | 9     |   |  |                              |        |                              |
| 4                | 206-208         | 22    | 0,15795   | 0,3624   | 0,20445                      | 20,445 | 0,11827                      |
| 5                | 208-210         | 28    | 0,3624  | 0,6175   | 0,2551                       | 25,51  | 0,01574                      |
| 6                | 210-212         | 19    | 0,6175  | 0,8292   | 0,2117                       | 21,17  | 0,22243                      |
| 7                | 212-214         | 11    | 0,8292  | 0,9454   | 0,1162                       | 11,62  | 0,000375                     |
| 8                | 214-216         | 5     |   |  |                              |        |                              |
| 9                | 216-218         | 1     |   |  |                              |        |                              |
| Всего            |                 | 100   |   |  | 1,0000                       | 100    | $\chi_{расчет}^2 = 0,560805$ |

По таблице распределения Пирсона для уровня значимости  $\alpha = 0,05$  и  $\nu = k' - l - 1 = 5 - 2 - 1 = 2$  степенями свободы найдем  $\chi_{0,005; 2}^2 = 5,99$ . Так как  $\chi_{расчет}^2 = 0,560805 < 5,99$ , то гипотезу о нормальном распределении средней заработной платы принимаем.

5. Найти интервальные оценки параметров  $\alpha$  и  $\sigma$  нормального распределения.

Точечные оценки  $\hat{a} = \bar{x} = 209,08$ ,  $\hat{\sigma} = s_{\text{несмещ}} = 3,0738$ .

Вычислим доверительный интервал, покрывающий неизвестное математическое ожидание  $a$  по заданному уровню значимости  $\alpha = 0,05$  и числу степеней свободы  $\nu = n - 1 = 100 - 1 = 99$ . По таблицам распределения Стьюдента находим  $t_{\frac{\alpha}{2}, \nu} = t_{0,025, 99} = 1,96$ .

$$\text{Предельная погрешность } \Delta = t_{\frac{\alpha}{2}, \nu} \cdot \frac{\hat{\sigma}}{\sqrt{n-1}} = 1,96 \cdot \frac{3,0738}{\sqrt{100-1}} = 0,605.$$

$$\bar{x} - \Delta < a < \bar{x} + \Delta,$$

$$209,08 - 0,605 < a < 209,08 + 0,605,$$

$$a_1 = 208,475 < a < 209,685 = a_2.$$

Находим доверительный интервал, накрывающий неизвестное среднеквадратическое отклонение  $\sigma$  с заданной точностью надежности  $\gamma = 1 - \alpha = 1 - 0,05 = 0,95$  и степенями свободы  $\nu = n - 1 = 100 - 1 = 99$ . По статистическим таблицам определяем  $q = q(\gamma; \nu) = 0,144$ .

6. Проверить нулевую гипотезу  $H_0: a = a_0 = a_1 = 208,475$  о математическом ожидании при альтернативной гипотезе  $H_1: a < a_1$ .

$$\text{Вычислим } t_{\text{расчет}} = \frac{\bar{x} - a_0}{s} \sqrt{n-1} = \frac{209,08 - 208,475}{3,0738} \sqrt{100-1} = 1,958.$$

По виду альтернативной гипотезы применяем левосторонний  $t$ -критерий, т.е. критическая точка  $t_{\alpha, n-1}$  находится из условия  $P(t < t_{\text{крит}}) = \alpha$ .  $t_{\text{левая крит}} = -t_{\alpha, n-1} = -2,62$ .

Расчетное значение критерия  $1,958 > -2,62$  – значение критической точки, поэтому нет оснований для отклонения нулевой гипотезы с вероятностью доверия 95 %.

7. Проверить нулевую гипотезу  $H_0: \sigma^2 = \sigma_0^2 = s_2^2 = 3,516$  о дисперсии против альтернативной  $H_1: \sigma^2 \neq \sigma_0^2$ .

По виду альтернативной гипотезы критическая область двусторонняя. Надо определить правую и левую критические точки для критерия  $\chi^2$  по уровню значимости  $\alpha = 0,05$  и числу степеней свободы  $\nu = n - 1 = 100 - 1 = 99 > 30$ . Поэтому воспользуемся равенством Уилсона-Гилфerti

$$\chi_{\alpha, \nu}^2 = \nu \left( 1 - \frac{2}{9\nu} + u_{\alpha} \sqrt{\frac{2}{9\nu}} \right)^3, \text{ где значение функции Лапласа } \Phi(u_{\alpha}) =$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{u_\alpha} e^{-\frac{t^2}{2}} dt = 1 - 2\alpha$$
 для  $\alpha$ , принимающего значения соответственно 0,025 (правой) и  $1 - 0,025 = 0,975$  (левой критических точек).

$\Phi(u_{0,025}) = 0,95, \quad u_{0,025} = 1,96. \quad \Phi(u_{0,975}) = -0,95, \quad u_{0,975} = -1,96.$

$$\chi_{\text{правая кр. т.}}^2 = \chi_{0,025,99}^2 = 99 \left( 1 - \frac{2}{9,99} + 1,96 \cdot \sqrt{\frac{2}{9,99}} \right)^3 = 128,43,$$

$$\chi_{\text{левая кр. т.}}^2 = \chi_{0,975,99}^2 = 99 \left( 1 - \frac{2}{9,99} - 1,96 \cdot \sqrt{\frac{2}{9,99}} \right)^3 = 73,35.$$

Вычислим расчетное значение критерия  $\chi_{\text{расчет}}^2 = \frac{n\hat{\sigma}^2}{\sigma_0^2} =$

$$= \frac{100 \cdot 9,448080}{3,516^2} = 76,43. \quad 73,35 < 76,43 < 128,43, \text{ поэтому нулевая гипотеза}$$

принимается.

### ЗАДАЧА № 5.

Пусть частотная таблица, построенная на основе результатов опроса студентов трех факультетов относительно того, сколько часов в неделю они проводят в библиотеке, имеет следующий вид:

| Временные интервалы                           | Середина интервала | Число студентов, которые проводят в библиотеке указанное количество времени                                      |             |             |
|---|--------------------|--|-------------|-------------|
|   |                    | Факультет А  | Факультет В | Факультет С |
| 0 – 4   | 2                  | 11   | 8           | 10          |
| 4 – 8   | 6                  | 18   | 20          | 21          |
| 8 – 12  | 10                 | 14   | 15          | 19          |
| 12 – 16                                       | 14                 | 7  | 12          | 15          |
| $m_i$   |                    | 11 + 18 + 14 + 7 = 50  | 55          | 65          |
| $\bar{x}_i$                                   |                    | $\frac{1}{50}(2 \cdot 11 + 6 \cdot 18 + 10 \cdot 14 + 14 \cdot 7) = 7,36$  | 8,25        | 8,4         |
| $S_{i(\text{несмещен})}^2 = \hat{\sigma}_i^2$ |                    | $\frac{1}{50-1}(11 \cdot (2-7,36)^2 + 18 \cdot (6-7,36)^2 + 14 \cdot (10-7,36)^2 + 7 \cdot (14-7,36)^2) = 15,42$ | 15,69       | 16,4        |

Можно ли утверждать, что количество времени, проведенного студентом в библиотеке, не зависит от того, на каком факультете учится студент?

### Решение.

Используем метод однофакторного дисперсионного анализа. Проверим гипотезу о равенстве среднего количества времени, проведенного студентами разных факультетов в библиотеке:  $H_0: a_1 = a_2 = a_3$ . В альтернативной гипотезе равенство нарушено.

Фактором является факультет. Даны три уровня фактора ( $k=3$ ).

При этом должны быть выполнены условия. Допустим, что независимость наблюдений гарантируется организацией эксперимента, а время проведения студентами в библиотеке для каждого факультета имеет нормальный закон распределения со своими математическими ожиданиями и одинаковыми дисперсиями. Используя критерий Бартлетта, убедимся предварительно в справедливости гипотезы  $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ . В альтернативной гипотезе равенство нарушено. Вычислим значение выборочного критерия:

$$\chi^2_{\text{расчет}} = \frac{\sum_{i=1}^3 (m_i - 1) \ln(s^2 / s_i^2)}{1 + \frac{1}{3(k-1)} \cdot c} = \frac{49 \cdot \ln \frac{15,88}{15,42} + 54 \cdot \ln \frac{15,88}{15,69} + 64 \cdot \ln \frac{15,88}{16,4}}{1,09} = \frac{49 \cdot 0,0862 + 54 \cdot 0,0118 - 64 \cdot 0,0304}{1,09} = 2,67,$$

$$\text{где } c = \sum_{i=1}^k \frac{1}{m_i - 1} - \frac{1}{\sum_{i=1}^k (m_i - 1)} = \left( \frac{1}{50-1} + \frac{1}{55-1} + \frac{1}{65-1} \right) - \frac{1}{49+54+64} = 0,54;$$

$$s^2 = \frac{\sum_{i=1}^3 (m_i - 1) s_i^2}{\sum_{i=1}^3 (m_i - 1)} = \frac{49 \cdot 15,42 + 54 \cdot 15,69 + 64 \cdot 16,4}{49 + 54 + 64} = 15,88.$$

По таблице определяем критическое значение критерия  $\chi_{0,05,2} = 5,991$ , которое больше расчетного, поэтому принимаем предварительную нулевую гипотезу о равенстве дисперсий.

Проверим теперь основную гипотезу однофакторного дисперсионного анализа:

$H_0: a_1 = a_2 = a_3$ . В альтернативной гипотезе равенство нарушено.

Вычислим выборочную среднюю:

$$\bar{x} = \frac{1}{50+55+65} (2 \cdot (11+8+10) + 6 \cdot (18+20+21) + 10 \cdot (14+15+19) +$$

$$+ 14 \cdot (7 + 12 + 15)) = 8,047.$$

Определим суммы квадратов отклонений от средних (факторную и остаточную):

$$S_{\text{факт}}^2 = (7,36 - 8,047)^2 \cdot 50 + (8,25 - 8,047)^2 \cdot 55 + (8,4 - 8,047)^2 \cdot 65 = 33,96;$$

$$S_{\text{остат}}^2 = \sum_{i=1}^3 s_i^2 (m_i - 1) = 15,42 \cdot 49 + 15,69 \cdot 54 + 16,4 \cdot 64 = 2652,44.$$

Вычисляем расчетное значение критерия

$$F_{\text{расчет}} = \frac{\frac{1}{k-1} S_{\text{факт}}^2}{\frac{1}{n-k} S_{\text{остат}}^2} = \frac{\frac{1}{3-1} \cdot 33,96}{\frac{1}{(50+55+65)-3} \cdot 2652,44} = \frac{16,98}{15,88} = 1,069.$$

По таблицам распределения Фишера-Снедекора с  $\nu_1 = k - 1 = 3 - 1 = 2$  и  $\nu_2 = n - k = 170 - 3 = 167$  степенями свободы для уровня значимости  $\alpha = 0,05$  определим критическое значение выборочной статистики  $F_{0,05,2,167} = 3,00$ . Расчетное значение меньше критического, поэтому нулевую гипотезу принимаем. Это значит, что количество времени, проведенного студентом в библиотеке, не зависит от того, на каком факультете учится студент. Уменьшая уровень значимости, тем более результат подтвердится.

### ЗАДАЧА № 7.

Пусть исходная корреляционная таблица имеет вид

| Объем выпуска продукции (Y)   | 0-2 | 2-4 | 4-6 | 6-8 | 8-10 | $m_{i \cdot}$                               |
|-------------------------------|-----|-----|-----|-----|------|---|
| Стоимость основных фондов (X) |     |     |     |     |      |   |
| 0-0,2                         | 2   | 2   |     |     |      | 4   |
| 0,2-0,4                       | 2   | 8   | 12  |     |      | 22  |
| 0,4-0,6                       |     | 2   | 18  | 8   |      | 28  |
| 0,6-0,8                       |     |     | 4   | 3   |      | 7   |
| 0,8-1,0                       |     |     |     | 3   | 2    | 5   |
| 1,0-1,2                       |     |     |     |     | 3    | 3   |
| $m_{\cdot j}$                 | 4   | 12  | 34  | 14  | 5    | $\sum_{i=1}^5 \sum_{j=1}^6 m_{ij} = n = 69$ |

По данным корреляционной таблицы найдем середины интервалов:

| Объем выпуска продукции (Y)<br>Средина интервала y   |         | 0-2 | 2-4 | 4-6 | 6-8 | 8-10 | $m_{.j}$                                    |
|--|---------|-----|-----|-----|-----|------|---|
|  |         | 1   | 3   | 5   | 7   | 9    |   |
| Стоимость основных фондов (X)<br>Средина интервала x | 0-0,2   | 2   | 2   |     |     |      | 4   |
|  | 0,2-0,4 | 2   | 8   | 12  |     |      | 22  |
|  | 0,4-0,6 |     | 2   | 18  | 8   |      | 28  |
|  | 0,6-0,8 |     |     | 4   | 3   |      | 7   |
|  | 0,8-1,0 |     |     |     | 3   | 2    | 5   |
|  | 1,0-1,2 |     |     |     |     | 3    | 3   |
| $m_{.j}$   |         | 4   | 12  | 34  | 14  | 5    | $\sum_{i=1}^2 \sum_{j=1}^5 m_{ij} = n = 69$ |

По данным корреляционной таблицы найти:

1. Числовые характеристики случайных величин  $X$  и  $Y$ .
2. Выборочный коэффициент корреляции  $\hat{r}$ . Проверить значимость коэффициента корреляции при уровне значимости 5 %.
3. Вычислить коэффициент детерминации и объяснить его смысл.
4. Уравнение линейной регрессии  $Y$  на  $X$ . Построить корреляционное поле и график линейного уравнения регрессии  $Y$  на  $X$ .
5. Определить степень рассеивания экспериментальных точек вокруг линии регрессии.

1. Числовыми характеристиками случайных величин  $X$  и  $Y$  являются соответственно их математические ожидания  $MX$ ,  $MY$ , дисперсии  $\sigma_x^2$ ,  $\sigma_y^2$ . По выборочным данным можно определить оценки числовых характеристик. Это соответственно выборочные средние  $\bar{x}$ ,  $\bar{y}$ , выборочные дисперсии  $s_x^2$ ,  $s_y^2$  (обычно берут несмещенные оценки дисперсии при малом объеме выборки):

$$\bar{x} = \frac{1}{69}(0,1 \cdot 4 + 0,3 \cdot 22 + 0,5 \cdot 28 + 0,7 \cdot 7 + 0,9 \cdot 5 + 1,1 \cdot 3) = 0,49;$$

$$\bar{y} = \frac{1}{69}(1 \cdot 4 + 3 \cdot 12 + 5 \cdot 34 + 7 \cdot 14 + 9 \cdot 5) = 5,12;$$

$$\hat{\sigma}_x^2 = s_x^2 = \overline{x^2} - (\bar{x})^2 = \frac{1}{69}(0,1^2 \cdot 4 + 0,3^2 \cdot 22 + 0,5^2 \cdot 28 + 0,7^2 \cdot 7 + 0,9^2 \cdot 5 + 1,1^2 \cdot 3) - 0,49^2 = 0,0516; s_x = 0,23;$$

$$\hat{\sigma}_y^2 = s_y^2 = \overline{y^2} - (\bar{y})^2 = \frac{1}{69}(1 \cdot 4 + 9 \cdot 12 + 25 \cdot 34 + 49 \cdot 14 + 81 \cdot 5) - 5,12^2 = 3,44; s_y = 1,88.$$

2. Вычислим выборочный коэффициент корреляции, составив для удобства таблицу:

| $i, j$   | $x_i$ | $m_i$ | $y_i$ | $m_i$ | $y_i m_i$ | $x_i m_i$ | $y_i^2$ | $y_i^2 m_i$ | $x_i^2$ | $x_i^2 m_i$ |
|----------|-------|-------|-------|-------|-----------|-----------|---------|-------------|---------|-------------|
| 1        | 0,1   | 4     | 1     | 4     | 4         | 0,4       | 1       | 4           | 0,01    | 0,04        |
| 2        | 0,3   | 22    | 3     | 12    | 36        | 6,6       | 9       | 108         | 0,09    | 1,98        |
| 3        | 0,5   | 28    | 5     | 34    | 170       | 14        | 25      | 850         | 0,25    | 7           |
| 4        | 0,7   | 7     | 7     | 14    | 98        | 4,9       | 49      | 686         | 0,49    | 3,43        |
| 5        | 0,9   | 5     | 9     | 5     | 45        | 4,5       | 81      | 405         | 0,81    | 4,05        |
| 6        | 1,1   | 3     |       |       |           | 3,3       |         |             | 1,21    | 3,63        |
| $\Sigma$ |       | 69    |       | 69    | 353       | 33,7      |         | 2053        |         | 20,13       |

$$\sum_{i=1}^6 \sum_{j=1}^6 x_i y_j m_{ij} = 0,1 \cdot 1 \cdot 2 + 0,1 \cdot 3 \cdot 2 + 0,3 \cdot 1 \cdot 2 + 0,3 \cdot 3 \cdot 8 + 0,3 \cdot 5 \cdot 12 + 0,5 \cdot 3 \cdot 2 +$$

$$+ 0,5 \cdot 5 \cdot 18 + 0,5 \cdot 7 \cdot 8 + 0,7 \cdot 5 \cdot 4 + 0,7 \cdot 7 \cdot 3 + 0,9 \cdot 7 \cdot 3 + 0,9 \cdot 9 \cdot 2 + 1,1 \cdot 9 \cdot 3 = 196,1.$$

Выборочный коэффициент корреляции

$$\hat{r} = \frac{\overline{xy} - \bar{x}\bar{y}}{\hat{\sigma}_x \hat{\sigma}_y} = \frac{\frac{1}{69} \cdot 196,1 - 0,49 \cdot 5,12}{0,23 \cdot 1,88} = 0,771.$$

Проверим значимость  $\hat{r}$  на уровне  $\alpha = 0,05$ .

Формулируем основную и альтернативную гипотезы:

$$H_0: \hat{r} = 0.$$

$$H_1: \hat{r} \neq 0.$$

Критическая область двусторонняя.

$$\text{Вычислим выборочную статистику (критерий)} \quad t_{\text{расчет}} = \frac{\hat{r} \sqrt{n-2}}{\sqrt{1-\hat{r}^2}} =$$

$$= \frac{0,771 \cdot \sqrt{67}}{\sqrt{1-0,771^2}} = 9,903. \text{ По таблицам распределения Стьюдента при числе}$$

степеней свободы  $v = n - 2 = 67$  находим критическое значение  $t_{\text{крит}} = t_{0,025; 67} = 2,00$ . Расчетное значение критерия  $t_{\text{расчет}} = 9,903 > t_{\text{крит}} = 2,00$ , поэтому принимаем альтернативную гипотезу и считаем выборочное значение коэффициента корреляции статистически значимым.

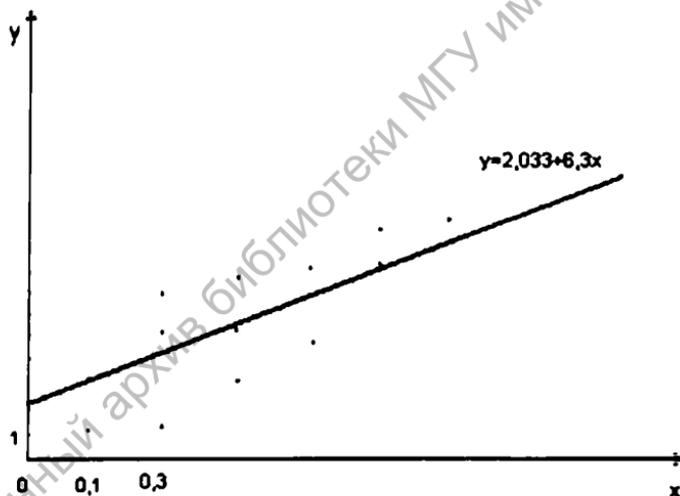
3. Для исследования зависимости объема выпуска продукции ( $Y$ ) от стоимости основных промышленно-производственных фондов ( $X$ ) вычислим коэффициент детерминации  $\hat{r}^2 = 0,594$ . Коэффициент детерминации показывает, что 59,4 % различий в объеме выпуска продукции объясняется зависимостью от стоимости основных фондов.

4. Определим уравнение регрессии  $y - \bar{y} = \hat{r} \frac{s_y}{s_x} (x - \bar{x})$ :  $y - 5,12 = 0,771 \cdot \frac{1,88}{0,23} (x - 0,49)$ , или  $y = 2,033 + 6,3x$  (общий вид уравнения  $y = \hat{a} + \hat{b}x$ ),  $\hat{a} = 2,033$ ,  $\hat{b} = 6,3$ .

Построим корреляционное поле и график уравнения регрессии  $Y$  на  $X$ .

|       |       |      |
|-------|-------|------|
| $x_i$ | 0     | 0,49 |
| $y_i$ | 2,033 | 5,12 |

По двум точкам строим линию регрессии.



5. Найдем среднюю ошибку  $s_{y|x} = s_y \sqrt{1 - \hat{r}^2}$ , характеризующую рассеивание экспериментальных точек вокруг линии регрессии  $y = 2,033 + 6,3x$  при изменении аргумента  $x$ :  $s_{y|x} = 1,88 \cdot 0,637 = 1,198$ .

## ЛИТЕРАТУРА

1. Белько, И.В., Кузьмич, К.К. Высшая математика для экономистов. III семестр: Экспресс-курс. – М.: Новое знание, 2002. – 144 с.
2. Большев, Л.Н., Смирнов, Н.В. Таблицы математической статистики. – М.: Наука, 1983.
3. Гмурман, В.Е. Теория вероятностей и математическая статистика. – М.: Высш. школа, 1997.
4. Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учеб. пособие для студентов вузов. Изд. 4-е стер. – М.: Высш. шк., 1998. – 400 с.: ил.
5. Гнеденко, Б.В. Курс теории вероятностей. – М.: Физматгиз, 1961.
6. Горелова, Г.В., Кацко, И.А. Теория вероятностей и математическая статистика в примерах и задачах с применением Excel: учебное пособие для вузов. – 2-е изд., исправленное и дополненное. – Ростов н/Д: Феникс, 2002. – 400 с.: ил.
7. Калинина, В.Н., Панкин, В.Ф. Математическая статистика: учеб. для студ. сред. спец. учеб. заведений. – 3-е изд., испр. – М.: Высш. шк., 2001. – 336 с.: ил.
8. Кремер, Н.Ш. Теория вероятностей и математическая статистика. – М.: ЮНИТИ, 2001.
9. Сазонова, А.М. Теория вероятностей и математическая статистика: методические рекомендации. Модуль 1 (Модуль 2, Модуль 3). – Могилев: МГУ им. А.А. Кулешова, 2004 (2007, 2010).

## СОДЕРЖАНИЕ

|   |    |
|---|----|
| ВЫБОР КОНТРОЛЬНОЙ РАБОТЫ .....  | 3  |
| ОБРАЗЕЦ ОФОРМЛЕНИЯ ТИТУЛЬНОГО ЛИСТА<br>КОНТРОЛЬНОЙ РАБОТЫ .....       | 4  |
| ПРОГРАММА ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ<br>И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ ..... | 5  |
| КОНТРОЛЬНЫЕ ЗАДАНИЯ .....   | 8  |
| ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ ВАРИАНТА .....                                  | 11 |
| ЛИТЕРАТУРА .....  | 26 |

Учебное издание

Сазонова Алла Михайловна

ТЕОРИЯ ВЕРОЯТНОСТЕЙ  
И МАТЕМАТИЧЕСКАЯ  
СТАТИСТИКА

Контрольные задания  
и методические рекомендации

Технический редактор *А.Л. Позняков*  
Компьютерная верстка *Н.А. Наумович*  
Корректор *Г.В. Карпенкова*

Подписано в печать *15.12*. 2011. Формат 60x84/16.  
Гарнитура Times New Roman. Усл.-печ. л. 1,6.  
Уч.-изд. л. 1,2. Тираж *57* экз. Заказ № *513*.

Учреждение образования «Могилевский государственный университет  
им. А.А. Кулешова», 212022, Могилев, Космонавтов, 1.

ЛИ № 02330/278 от 30.04.2004

Отпечатано в отделе оперативной полиграфии  
УО «МГУ им. А.А. Кулешова». 212022, Могилев, Космонавтов, 1.