

Министерство образования, науки и молодежной политики Краснодарского края

**Государственное бюджетное профессиональное образовательное учреждение
Краснодарского края
«АРМАВИРСКИЙ МАШИНОСТРОИТЕЛЬНЫЙ ТЕХНИКУМ»**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ
ПО ВЫПОЛНЕНИЮ ПРАКТИЧЕСКИХ РАБОТ**

по учебной дисциплине

«Теория вероятностей и математическая статистика»

**для студентов 2 курса очной формы обучения
по специальности 09.02.07 «Информационные системы и программирование»**

Содержание

1.	Введение.....	4
2.	Порядок проведения практического занятия.....	4
3.	Требования к оформлению практической работы.....	4
4.	Критерии выставления оценок.....	4
5.	Практическая работа №1. «Решение комбинаторных задач и уравнений»	5
6.	Практическая работа №2. «Вычисление вероятностей событий» ..	7
7.	Практическая работа №3. «Нахождение закона распределения дискретной случайной величины. Вычисление ее числовых характеристик»	12
8.	Практическая работа №4. «Нахождение числовых характеристик непрерывных случайных величин»	15
9.	Практическая работа №5. «Построение вариационного ряда и его графической диаграммы. Расчет по заданной выборке ее числовых характеристик»	17
10.	Практическая работа №6. «Расчет сводных характеристик выборки различными методами»	20
11.	Практическая работа №7. «Нахождение выборочного уравнения прямой линии регрессии»	23
12.	Список рекомендуемой литературы.....	26

1. Введение

Цель методических указаний - обеспечить четкую организацию проведения практических занятий по дисциплине «Теория вероятностей и математическая статистика» и предоставить возможность обучающимся, отсутствовавшим на практическом занятии, самостоятельно выполнить работу. При этом студенты, отсутствовавшие на практическом занятии, имеют право на получение консультаций у преподавателя.

Неудовлетворительная оценка, полученная обучающимся при выполнении практической работы, должна быть исправлена и повторно проверена преподавателем. Студент, имеющий к концу семестра более 75% практических работ, написанных на неудовлетворительную оценку, не может иметь положительную оценку при зачете по предмету.

2. Порядок проведения практических занятий

1. Опрос обучающихся по теме практической работы в различных формах
2. Краткое сообщение преподавателя о целях практического занятия, порядке его проведения и оформления работы
3. Выдача вариантов заданий
4. Выполнение обучающимися практической работы
5. Подведение преподавателем итогов практического занятия

3. Требования к оформлению практических работ

1. Задания выполняются в специально отведенной тетради шариковой ручкой синего или черного цвета.
2. Обязательно указываются номера практической работы и варианта, а также тема работы.
3. Условия заданий переписываются полностью или делается краткая запись того, что дано и надо найти; выполняется решение, в конце которого записывается ответ. В некоторых заданиях ответ не записывается, о чем обучающиеся должны быть уведомлены в начале выполнения работы.
4. Все рисунки, таблицы и схемы выполняются аккуратно карандашом или ручкой, желательно, под линейку.
5. Задания могут выполняться в произвольном порядке.

4. Критерии выставления оценок

Оценка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обоснованиях решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Оценка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Оценка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но обучающийся владеет обязательными умениями по проверяемой теме.

Оценка «2» ставится, если:

- допущены существенные ошибки, показавшие, что обучающийся не владеет обязательными умениями по данной теме в полной мере.

5. Практическая работа № 1

Тема: «Решение комбинаторных задач и уравнений»

Цель: приобрести практические навыки по вычислению числа соединений – вариантов различных выборок для конечных множеств в зависимости от конкретного случая.

Методические указания

1. Определение факториала

Опр. Произведение всех натуральных чисел от 1 до n включительно называется **n -факториалом** и обозначается $n!$, т.е. $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$

(!!) Считается, что $0! = 1$.

2. Основные правила комбинаторики

Правило умножения: Если некоторый элемент a можно выбрать m числом способов, а затем элемент b – n числом способов, то элемент « a и b » можно выбрать $m \cdot n$ числом способов.

Правило сложения: Если некоторый элемент a можно выбрать m числом способов, а другой элемент b – n числом способов, то элемент «**либо a , либо b** » можно выбрать $m+n$ числом способов.

Задача. Сколько однозначных, двузначных и трехзначных четных чисел можно составить из цифр 0,1,2,3, если цифры могут повторяться?
/ Очевидно, что из данных цифр можно составить только одно четное однозначное число – 2.

При составлении двузначного числа ав из данных цифр вместо а можно взять любую цифру, кроме нуля (3 возможности), вместо в можно взять любую из цифр 0 и 2 (2 возможности). Т.о., согласно правилу умножения, имеем $3 \cdot 2 = 6$ способов составить нужное нам число.

При составлении трехзначного числа авс из данных цифр вместо а можно взять любую цифру, кроме нуля (3 возможности), вместо в можно взять любую из них (4 возможности), вместо с можно взять любую из цифр 0 и 2 (2 возможности). Т.о., согласно правилу умножения, имеем $3 \cdot 4 \cdot 2 = 24$ способа составить число, удовлетворяющее условию задачи.

Применяя правило сложения, получим: $1 + 6 + 24 = 31$ /

3. Алгоритм вычисления числа соединений

1) Установить количество элементов n всего множества и количество элементов m его подмножества.

2) Определить, влияет ли порядок расположения элементов в подмножестве на число вариантов различных подмножеств, состоящих из этих m элементов.

3) Выбрать, в зависимости от конкретного случая, комбинаторную операцию:

а) если число комбинаций всего множества зависит от порядка расположения элементов в нем, и нет повторяющихся элементов, то **перестановки без повторений**: $P_n = n!$;

б) если число комбинаций всего множества зависит от порядка расположения элементов в нем, и есть повторяющиеся элементы, то **перестановки с повторениями**: $\tilde{P}_{m_1, m_2, \dots, m_n} = \frac{(m_1 + m_2 + \dots + m_n)!}{m_1! \cdot m_2! \cdot \dots \cdot m_n!}$;

в) если число комбинаций в подмножестве (выборке) зависит от порядка расположения элементов в нем, и нет повторяющихся элементов, то **размещения без повторений**: $A_n^m = \frac{n!}{(n-m)!}$;

г) если число комбинаций в подмножестве (выборке) зависит от порядка расположения элементов в нем, и элементы повторяются, то **размещения с повторениями**: $\tilde{A}_n^m = n^m$;

д) если число комбинаций в подмножестве (выборке) не зависит от порядка расположения элементов в нем, и нет повторяющихся элементов, то **сочетания без повторений**: $C_n^m = \frac{n!}{(n-m)!m!}$;

е) если число комбинаций в подмножестве (выборке) не зависит от порядка расположения элементов в нем, и есть повторяющиеся элементы, то **сочетания с повторениями**: $\tilde{C}_n^m = \frac{(m+n-1)!}{m!(n-1)!}$.

Задача. Сколько различных натуральных чисел можно составить из цифр 1, 2, 3, 4, 5 при условии, что любая из цифр в написании числа встречается не более одного раза?

/ Однозначных - $A_5^1 = 5$

Двузначных - $A_5^2 = 5 \cdot 4 = 20$

Трёхзначных - $A_5^3 = 5 \cdot 4 \cdot 3 = 60$

Четырёхзначных - $A_5^4 = 5 \cdot 4 \cdot 3 \cdot 2 = 120$

Пятизначных - $A_5^5 = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

$5 + 20 + 60 + 120 + 120 = 325$ /

Задача. В соревнованиях по волейболу участвуют 8 команд. Насколько более продолжительным будет турнир, организованный по круговой системе, чем по олимпийской?

/При проведении турнира по круговой системе каждый участник встречался с каждым и порядок их вхождения в пару не важен. Следовательно, по круговой системе

потребуется провести 28 встреч (C_8^2), а по олимпийской только - 7 (четыре встречи в $\frac{1}{4}$ финала, две - в полуфинале и одна в финале)/

Задача. Сколько различных двузначных чисел можно составить из цифр 1, 2, 3?
/ $\widehat{A}_3^2 = 3^2 = 9$ /

Задача. Сколько различных пятизначных чисел можно составить из цифр 1, 2, 3, если 1 встречается 1 раз, 2 – 2 раза, 3 – 2 раза?

$$/ \widehat{P}_{1,2,2} = \frac{(1+2+2)!}{1! \cdot 2! \cdot 2!} = 30 /$$

Задача. Имеются конфеты трех сортов в коробках. Сколько можно составить различных наборов из пяти коробок?

$$/ \tilde{C}_3^5 = \frac{7!}{5! \cdot 2!} = \dots = 21 /$$

Содержание задания

Задание 1. В футбольном турнире участвуют семь команд.

- 1) Найти число вариантов возможного распределения мест между ними.
- 2) Найти число вариантов распределения призовых мест.
- 3) Сколько игр будет проведено, если каждая команда проводит с каждым из соперников по одной игре?

Задание 2. В корзине лежат 7 белых шаров и 8 черных.

- 1) Сколькими способами можно достать из этой корзины 3 белых шара и 3 черных?
- 2) Ученик достает 3 шара одинакового цвета. Сколькими способами он может это сделать?

Задание 3. Сколькими способами можно разбить группу из 15 студентов на три подгруппы А, В и С по 3, 5 и 7 человек соответственно?

Задание 4. Сколько всего четырехзначных чисел можно составить из цифр 1, 4, 5 и 7? Сколько среди них четных чисел?

Задание 5. Решите уравнение:

$$P_n = 2P_{n-2}$$

6. Практическая работа № 2

Тема: «Вычисление вероятностей событий»

Цель: приобрести практические навыки по вычислению:

- вероятностей случайных событий по определению;
- вероятностей случайных событий по известным вероятностям других событий, с ним связанных, используя теоремы сложения и умножения;

- вероятности события, которое может наступить только при условии появления одной из гипотез, образующих полную группу событий, а также по переоценке вероятностей гипотез после того, как событие произошло;
- вероятностей событий в независимых повторных испытаниях по формуле Бернулли, а также наивероятнейшего числа наступления события в схеме Бернулли.

Методические указания

1. Классическая и геометрическая вероятность

Опр. Вероятностью события A называется отношение числа M элементарных исходов, благоприятствующих наступлению данного события, к числу N всех элементарных исходов испытания.

Т.о.:

$$P(A) = \frac{M}{N}$$

Опр. Геометрической вероятностью события называется отношение меры области, благоприятствующей появлению события, к мере всей области.

2. Алгоритм вычисления вероятностей событий по определению

1. Ввести обозначения для заданных величин и вопроса задачи.
2. Выбрать формулу вероятности, соответствующую данному случаю:

а) **классическое определение**: если задано общее число N равновероятных исходов и число исходов M , благоприятствующих событию (которые можно сосчитать), то найти вероятность по формуле $p = \frac{M}{N}$;

б) **геометрическое определение**: если все возможные исходы можно изобразить с помощью геометрической фигуры, то надо:

- нарисовать эту фигуру;
- внутри нее нарисовать фигуру, соответствующую исходам, благоприятствующим событию;
- вычислить меры этих фигур (длины, площади, объемы);
- найти вероятность как отношение этих мер.

Задача. Брошена игральная кость. Найти вероятность следующих событий:

- 1) A – «выпало 3 очка»

/ По условию: $N = 6, M = 1$

$$p = \frac{M}{N} = \frac{1}{6}$$

- 2) B – «выпало четное число очков»

/ По условию: $N = 6, M = 3(2, 4, 6)$

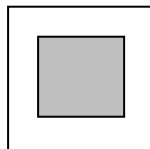
$$p = \frac{M}{N} = \frac{3}{6} = \frac{1}{2}$$

Задача. Из урны, в которой находятся 12 белых и 8 черных шаров, вынимают наудачу 2 шара. Какова вероятность того, что оба шара черные?

$$/ N = C_{20}^2 = 190, \quad M = C_8^2 = 28 \quad \Rightarrow \quad P(A) = \frac{14}{95} \quad /$$

Задача. На плоскости нанесена сетка квадратов со стороной 8 см. Найти вероятность, что брошенный на плоскость круг радиуса 1 см не пересечет ни одной стороны квадрата.

/



$$p = \frac{M}{N} = \frac{S_{\text{м.кв.}}}{S_{\text{б.кв.}}} = \frac{6^2}{8^2} = \frac{36}{64} = \frac{9}{16} \quad /$$

3. Алгоритм вычисления вероятностей событий по известным вероятностям других событий, с ними связанных

1. Обозначить все события, указанные в задаче. Известные вероятности представить в виде дробей.

2. Установить связи между событиями.

3. Вычислить требуемые вероятности, используя теоремы сложения и умножения вероятностей, а также формулу для вычисления вероятности противоположного события:

Т1. Вероятность появления одного из двух несовместных событий, неважно какого, равна сумме вероятностей этих событий, т.е.

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

Т2. Вероятность совместного появления двух независимых событий равна произведению их вероятностей, т.е.

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

Т3. Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного наступления, т.е.

$$P(A + B) = P(A) + P(B) - P(A \cdot B)$$

Т4. Вероятность совместного появления двух зависимых событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое событие уже наступило, т.е.

$$P(A \cdot B) = P(A) \cdot P_A(B)$$

Если A и \bar{A} - противоположные события, то $P(\bar{A}) = 1 - P(A)$

Задача. От коллектива бригады, которая состоит из 6 мужчин и 4 женщин, на профсоюзную конференцию выбираются 2 человека. Какова вероятность, что среди выбранных хотя бы одна женщина?

$$/ \text{ I способ: Пусть } A - \text{ «одна женщина»: } P(A) = \frac{C_6^1 \cdot C_4^1}{C_{10}^2} = \dots = \frac{8}{15}$$

$$B - \text{«две женщины»}: P(B) = \frac{C_4^2}{C_{10}^2} = \dots = \frac{2}{15},$$

$$\text{тогда } P(A+B) = \frac{8}{15} + \frac{2}{15} = \frac{10}{15} = \frac{2}{3}$$

$$\underline{2 \text{ способ:}} \text{ Пусть } A - \text{«два мужчины»}: P(A) = \frac{C_6^2}{C_{10}^2} = \frac{15}{45} = \frac{1}{3},$$

$$\text{тогда } P(\bar{A}) = 1 - P(A) = \frac{2}{3} /$$

Задача. Два стрелка одновременно производят выстрел по мишени. Первый из них поражает цель в 80%, а второй – в 70% выстрелов. Какова вероятность поражения цели?

$$\underline{1 \text{ способ:}} P(A) = 0,8 \quad P(\bar{A}) = 0,2$$

$$P(B) = 0,7 \quad P(\bar{B}) = 0,3$$

$$P(\bar{A} \cdot \bar{B}) = 0,2 \cdot 0,3 = 0,06 \Rightarrow P(A+B) = 1 - 0,06 = 0,94$$

$$\underline{2 \text{ способ:}} \text{ Пусть } P(A) = 0,7 \text{ и } P(B) = 0,8, \text{ тогда } P(A \cdot B) = 0,7 \cdot 0,8 = 0,56.$$

$$A \text{ значит, } P(A+B) = 0,7 + 0,8 - 0,56 = 0,94 /$$

Задача. В цехе 2 станка. Вероятность занятости каждого из них равна 0,7. Какова вероятность того, что один станок занят, а второй нет?

/ Пусть A – «занят только один станок»

A_1 – «занят 1-й станок»

$$A_2 - \text{«занят 2-й станок»}, \text{ тогда } A = A_1 \cdot \bar{A}_2 + \bar{A}_1 \cdot A_2. \text{ А сл-но, } P(A) = 0,7 \cdot 0,3 + 0,3 \cdot 0,7 = 0,42 /$$

Задача. В ящике находятся 12 деталей, из которых 8 стандартных. Рабочий берет 2 детали. Вычислить вероятность того, что обе детали окажутся стандартными.

$$A - \text{«1-я деталь стандартная»} \quad P(A) = \frac{8}{12} = \frac{2}{3}$$

$$B - \text{«2-я деталь стандартная»} \quad P_A(B) = \frac{7}{11}$$

$$P(A \cdot B) = \frac{2}{3} \cdot \frac{7}{11} = \frac{14}{33} /$$

4. Алгоритм вычисления вероятности события по формуле полной вероятности и вероятности одной из гипотез по формуле Байеса

1. Дать описание всех гипотез и вычислить их вероятности.
 2. Вычислить условные вероятности события по каждой гипотезе.
 3. Выбрать нужную формулу:
- А) для подсчета вероятности события - формулу полной вероятности

$$P(A) = \sum_{i=1}^n P(H_i) P_{H_i}(A)$$

Б) для подсчета вероятности гипотезы - формулу Байеса

$$P_A(H_i) = \frac{P(H_i) \cdot P_{H_i}(A)}{\sum_{i=1}^n P(H_i) P_{H_i}(A)}$$

Задача. Пусть имеются три одинаковые урны с таким составом шаров:

- 1) 2 белых и 1 черный;
- 2) 3 белых и 2 черных;

3) 1 белый и 3 черных.

Какова вероятность того, что извлеченный из произвольно взятой урны шар - белый?

/ Пусть A – «извлечен белый шар»,

H_i – «извлечен шар из i -ой урны», тогда

$$P(H_i) = \frac{1}{3}, P_{H_1}(A) = \frac{2}{3}, P_{H_2}(A) = \frac{3}{5} \text{ и } P_{H_3}(A) = \frac{1}{4}.$$

$$A \text{ значит, } P(A) = \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{3}{5} + \frac{1}{3} \cdot \frac{1}{4} = \dots = \frac{91}{180} /$$

Задача. Два автомата производят одинаковые детали, которые поступают на общий конвейер. Производительность первого автомата вдвое больше производительности второго. Первый производит, в среднем, 60% деталей отличного качества, а второй – 84%. Наудачу взятая с конвейера деталь оказалась отличного качества. Какова вероятность того, что эта деталь произведена первым автоматом?

/ Пусть A – «деталь отличного качества»,

H_1 – «деталь произведена 1-ым автоматом»

H_2 – «деталь произведена 2-ым автоматом»

$$P(H_1) = \frac{2}{3}, P(H_2) = \frac{1}{3}, P_{H_1}(A) = 0,6 \text{ и } P_{H_2}(A) = 0,84 =>$$

$$P_A(H_1) = \frac{P(H_1) \cdot P_{H_1}(A)}{P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A)} = \dots = \frac{10}{17} /$$

5. Алгоритм вычисления вероятности события и наимвероятнейшего числа наступления события при повторных испытаниях

1. Ввести обозначения для заданных величин

2. Выбрать нужную формулу:

А) для подсчёта вероятности наступления события m раз при n испытаниях формулу Бернулли:

$$p_n(m) = C_n^m p^m q^{n-m}$$

Б) для подсчета вероятности наступления события не менее m раз при n испытаниях: $p_n(m) + \dots + p_n(n)$

В) для подсчета вероятности наступления события менее m раз при n испытаниях: $p_n(0) + \dots + p_n(m-1)$

Г) для подсчета вероятности наступления события не более m раз при n испытаниях: $p_n(0) + \dots + p_n(m)$

Д) для подсчета вероятности наступления события более m раз при n испытаниях: $p_n(m+1) + \dots + p_n(n)$

Е) для подсчета наимвероятнейшего числа наступления события при n испытаниях:

$$np - q \leq \mu \leq np + p$$

Задача. В урне 20 шаров: 15 белых и 5 чёрных. Вынули подряд 5 шаров, причём каждый вынутый шар возвращается в урну и перед извлечением следующего шары в урне тщательно перемешиваются. Найти вероятность того, что из пяти вынутых шаров будет 2 белых.

$$/ n = 5, m = 2, p = \frac{15}{20} = \frac{3}{4}, q = \frac{1}{4}$$

$$p_5(2) = C_5^2 p^2 \cdot q^3 = \frac{5!}{2! \cdot 3!} \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^3 = \dots = \frac{45}{512} /$$

Задача. Вероятность изготовления на станке-автомате нестандартной детали равна 0,02. Определить вероятность того, что среди наудачу взятых шести деталей окажутся более 4-х стандартных.

$$/ n = 6, \quad p = 0,02, \quad q = 0,98$$

$$p_6(0 \leq m \leq 1) = p_6(0) + p_6(1) = C_6^0 p^0 q^6 + C_6^1 p^1 q^5 \approx 0,9943 /$$

Задача. Доля изделий высшего сорта на данном предприятии составляет 31%. Чему равно наивероятнейшее число изделий высшего сорта в случае отбора партии из 75 изделий?

$$/ p = 0,31, \quad q = 0,69, \quad n = 75$$

$$0,31 \cdot 75 - 0,69 \leq \mu \leq 0,31 \cdot 75 + 0,31$$

$$22,56 \leq \mu \leq 23,56$$

$$\mu = 23 /$$

Содержание задания

Задание 1. Студент знает ответы на двадцать вопросов зачета из двадцати пяти. Какова вероятность того, что:

- 1) ему достанется на зачете известный вопрос?
- 2) из двух вопросов на зачете оба окажутся неизвестными?
- 3) из четырёх случайно взятых вопросов он знает ровно два?

Задание 2. Точку бросают наугад в круг радиусом 12 см. Какова вероятность того, что точка окажется вне квадрата, вписанного в этот круг?

Задание 3. В двух залах кинотеатра идут два фильма. Вероятности того, что на определенный час в кассах первого и второго залов есть билеты, соответственно равны 0,3 и 0,1. Найти вероятность того, что на данный час:

- 1) нет билетов;
- 2) только в одной кассе есть билеты.

Задание 4. На склад поступает продукция от двух фабрик, причем продукция 1-ой фабрики составляет 65% от всей продукции. Известно, что средний процент бракованных изделий для 1-ой фабрики равен 2%, а для 2-ой – 3%. Найти вероятность того, что наудачу взятое со склада изделие окажется бракованным. Взятое изделие оказалось бракованным, какова вероятность, что оно произведено на 2-ой фабрике?

Задание 5. Прибор состоит из 10 узлов. Вероятность безотказной работы каждого узла за некоторое время равна 0,85. Узлы выходят из строя независимо друг от друга. Найти вероятность того, что за это время откажут:

- 1) ровно два узла;
- 2) не менее двух узлов.

7. Практическая работа № 3

Тема: «Нахождение закона распределения дискретной случайной величины. Вычисление ее числовых характеристик»

Цель: приобрести практические навыки по составлению ряда распределения, нахождению функции распределения и вычислению основных числовых характеристик ДСВ.

Методические указания

1. Понятие ДСВ

Опр. Соответствие между возможными значениями x_1, x_2, \dots, x_n случайной величины X и их вероятностями p_1, p_2, \dots, p_n называется **законом распределения** случайной величины X .

Закон распределения ДСВ может быть представлен таблицей, первая строка которой содержит возможные значения x_i , а вторая – вероятности p_i . Эта таблица называется **рядом распределения**. Причем, $\sum_{i=1}^n p_i = 1$.

Ряд распределения можно задать графически, если по оси абсцисс отложить возможные значения случайной величины, а по оси ординат – соответствующие вероятности этих значений. Соединив точки $(x_i; p_i)$ последовательно отрезками, получают ломанную, которую называют **многоугольником распределения вероятностей**.

2. Числовые характеристики ДСВ

Опр. **Модой ДСВ** называется такое значение ДСВ, вероятность которого наибольшая.

Обозначение: $M_o(X)$

Опр. **Медианой ДСВ** называется среднее по положению в пространстве событий значение ДСВ.

Обозначение: $M_c(X)$

Опр. **Математическим ожиданием ДСВ** называется сумма произведений всех возможных значений случайной величины на соответствующие вероятности появления этих значений, т.е.

Обозначение: $M(X)$

$$M(X) = \sum x_i p_i$$

Опр. **Дисперсией СВ** называется математическое ожидание квадрата отклонения СВ от ее математического ожидания.

Обозначение: $D(X)$

$$D(X) = M(X - M(X))^2$$

или

$$D(X) = M(X^2) - M^2(X)$$

Если X является ДСВ, то

$$D(X) = \sum (x_i - M(X))^2 p_i$$

Опр. **Среднеквадратичным отклонением СВ** называется арифметический квадратный корень из ее дисперсии.

Обозначение: $\sigma(X)$

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

3. Вычисление числовых характеристик случайных величин, распределение которых подчиняется определенному закону

Вид распределения ДСВ	Биномиальное распределение	Гипергеометрическое распределение
Вероятность	$C_n^k p^k q^{n-k}$	$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$
Математическое ожидание	np	$n \frac{M}{N}$
Дисперсия	npq	$n \frac{M}{N-1} (1 - \frac{M}{N}) (1 - \frac{n}{N})$

Задача. Найдите $M_0(X)$, $M_c(X)$, $M(X)$, $D(X)$ и $\sigma(X)$, если СВ задана следующим рядом распределения:

X	2	4	7	10	12
p	0,1	0,2	0,4	0,2	0,1

/ Очевидно: $M_0(X) = 7$, $M_c(X) = 7$

Для удобства вычислений остальных числовых характеристик все результаты можно свести в таблицу:

x_i	p_i	$x_i \cdot p_i$	$x_i - M(X)$	$(x_i - M(X))^2 \cdot p_i$	x_i^2	$x_i^2 \cdot p_i$
2	0,1	0,2	-5	$25 \cdot 0,1 = 2,5$	4	0,4
4	0,2	0,8	-3	$9 \cdot 0,2 = 1,8$	16	3,2
7	0,4	2,8	0	0	49	19,6
10	0,2	2	3	1,8	100	20
12	0,1	1,2	5	2,5	144	14,4
Σ	1	7		8,6		57,6
$M(X)$			$D(X)$		$M(X^2)$	

$$D(X) = 57,6 - 7^2 = 8,6$$

$$\sigma(X) = \sqrt{8,6} \approx 2,93$$

Задача. Проверка качества МК показала, что из каждых ста МК имеют дефекты в среднем 25 штук.

1) Составить ряд распределения СВХ– «число исправных МК» из взятых наудачу шести из них.

2) Найти числовые характеристики этой СВ.

/ По условию, $n = 6$, $p = 0,75$, $q = 0,25$

$$p_6(0) = 0,25^6 \approx 0,000$$

$$p_6(1) = 6 \cdot 0,75 \cdot 0,25^5 \approx 0,004$$

$$p_6(2) = C_6^2 \cdot 0,75^2 \cdot 0,25^4 \approx 0,033$$

$$p_6(3) = C_6^3 \cdot 0,75^3 \cdot 0,25^3 \approx 0,132$$

$$p_6(4) = C_6^4 \cdot 0,75^4 \cdot 0,25^2 \approx 0,297$$

$$p_6(5) = C_6^5 \cdot 0,75^5 \cdot 0,25 \approx 0,356$$

$$p_6(6) = 0,75^6 \approx 0,178$$

Т.о., закон распределения можно представить в виде ряда:

x	0	1	2	3	4	5	6
p	0,000	0,004	0,033	0,132	0,297	0,356	0,178

$$M(X) = 6 \cdot 0,75 = 4,5 \quad D(X) = 6 \cdot 0,75 \cdot 0,25 = 1,125 \quad \sigma(X) = \sqrt{1,125} \approx 1,06$$

Задача. В партии, состоящей из 10 МК, семь – стандартных. Контролер ОТК наудачу проверил два МК. Составить закон распределения числа обнаруженных стандартных МК и найти числовые характеристики этой СВ.

/ X – число стандартных МК из взятых 2-х

Возможные значения этой величины: 0, 1 и 2. Подсчитаем вероятности этих значений:

$$P(x=0) = \frac{C_7^0 \cdot C_3^2}{C_{10}^2} = \frac{1}{15} \quad P(x=1) = \frac{C_7^1 \cdot C_3^1}{C_{10}^2} = \frac{7}{15} \quad P(x=2) = \frac{C_7^2 \cdot C_3^0}{C_{10}^2} = \frac{7}{15}$$

Т.о., закон распределения этой ДСВ имеет вид:

X	0	1	2
p	$\frac{1}{15}$	$\frac{7}{15}$	$\frac{7}{15}$

1 способ: Найдем $M(X)$ и $D(X)$ по общей формуле:

$$M(X) = 0 \cdot \frac{1}{15} + 1 \cdot \frac{7}{15} + 2 \cdot \frac{7}{15} = 1,4, \quad M(X^2) = 0 \cdot \frac{1}{15} + 1 \cdot \frac{7}{15} + 4 \cdot \frac{7}{15} = 2\frac{1}{3} \Rightarrow D(X) = 2,33 - 1,96 = 0,37 \text{ и } \sigma(X) = \sqrt{0,37} \approx 0,6$$

$$\text{2 способ: } M(X) = 2 \cdot \frac{7}{10} = 1,4, \quad D(X) = 2 \cdot \frac{7}{9} \left(1 - \frac{7}{10}\right) \left(1 - \frac{2}{10}\right) = \frac{28}{75} = 0,37/$$

Содержание задания

Задание 1. Случайная величина X задана рядом распределения:

X	3	5	7	11	12
p	0,11	0,24	0,37	0,13	?

- Найдите недостающее значение вероятности.
- Постройте многоугольник распределения.
- Найдите функцию распределения $F(x)$ и постройте ее график.

Задание 2. Закон распределения случайной величины задан таблицей:

x_i	2	7	11
p_i	0,1	0,3	0,6

Найдите медиану, моду, математическое ожидание и дисперсию случайной величины.

Задание 3. Для следующих задач составьте закон распределения ДСВ и найдите ее математическое ожидание и среднеквадратичное отклонение.

1) В магазин вошли три покупателя. Вероятность сделать покупку для каждого из них равна 0,25. Случайная величина X – «число покупок».

2) В партии из двадцати деталей пятнадцать стандартных. Берут наугад три детали. Случайная величина X – «число стандартных деталей» среди отобранных.

8. Практическая работа № 4

Тема: «Нахождение числовых характеристик непрерывных случайных величин»

Цель: приобрести практические навыки по вычислению основных числовых характеристик НСВ, а также установлению связей между функцией распределения и плотностью распределения вероятностей для НСВ.

Методические указания

Опр. Функция $F(x)$, задающая вероятность того, что случайная величина X принимает значение, меньшее x , называется **функцией распределения случайной величины (интегральной функцией распределения)**.

Т.о.:

$$F(x) = P(X < x)$$

Опр. Первая производная интегральной функции распределения $F(x)$ называется **дифференциальной функцией распределения вероятностей $f(x)$ (плотностью вероятности)**.

Т.о.:

$$f(x) = F'(x)$$

$$F(x) = \int_{-\infty}^x f(x) dx$$

Вероятность нахождения величины X в интервале $a \leq X \leq b$

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$$

Опр. Математическим ожиданием НСВ с функцией плотности $f(x)$, возможные значения которой принадлежат отрезку $[a; b]$, называется определенный интеграл $\int_a^b xf(x)dx$, т.е.

$$M(X) = \int_a^b xf(x)dx$$

Опр. Дисперсией СВ называется математическое ожидание квадрата отклонения СВ от ее математического ожидания.

Для НСВ с плотностью распределения $f(x)$:

$$D(X) = \int_a^b (x - M(X))^2 f(x) dx$$

Задача. Дана интегральная функция распределения вероятности

$$F(x) = \begin{cases} 0, & \text{при } x \leq 0, \\ x^2, & \text{при } 0 < x \leq 1, \\ 1, & \text{при } x > 1. \end{cases}$$

Найти p ($0,25 < X < 0,5$) и плотность распределения вероятности.

$$p(0,25 < X < 0,5) = 0,25 - 0,0625 = 0,1875$$

$$f(x) = F'(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ 2x, & \text{если } 0 < x \leq 1, \\ 0, & \text{если } x > 1. \end{cases}$$

Задача. Дано: $f(x) = \begin{cases} 0, & \text{если } x \leq -\frac{\pi}{2} \\ \frac{1}{2} \cos x, & \text{если } -\frac{\pi}{2} < x \leq \frac{\pi}{2} \\ 0, & \text{если } x > \frac{\pi}{2} \end{cases}$. Найдите: $F(x)$

$$\text{При } x \leq -\frac{\pi}{2} \quad F(x) = \int_{-\infty}^x 0 \, dx = 0$$

$$\text{При } -\frac{\pi}{2} < x \leq \frac{\pi}{2} \quad F(x) = \int_{-\infty}^{-\frac{\pi}{2}} 0 \, dx + \int_{-\frac{\pi}{2}}^x \frac{1}{2} \cos x \, dx = 0 + \frac{1}{2} \sin x \Big|_{-\frac{\pi}{2}}^x = \frac{1}{2}(\sin x + 1)$$

$$\text{При } x > \frac{\pi}{2} \quad F(x) = 1$$

$$\text{Т.о.: } F(x) = \begin{cases} 0, & \text{если } x < -\frac{\pi}{2} \\ \frac{1}{2}(\sin x + 1), & \text{если } -\frac{\pi}{2} \leq x < \frac{\pi}{2} \\ 1, & \text{если } x \geq \frac{\pi}{2} \end{cases} \quad /$$

Задача. Найти математическое ожидание, дисперсию и среднее квадратичное отклонение НСВ, заданной дифференциальной функцией распределения

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0 \\ 3x^2 & \text{при } 0 < x \leq 1 \\ 0 & \text{при } x > 1. \end{cases}$$

$$M(X) = \int_0^1 x \cdot 3x^2 \, dx = \dots = \frac{3}{4} \quad D(X) = \int_0^1 \left(x - \frac{3}{4}\right)^2 3x^2 \, dx = \dots = 0,04 \quad \sigma(X) = 0,2 \quad /$$

Содержание задания

Задание 1. Случайная величина X задана функцией распределения:

$$F(x) = \begin{cases} 0, & \text{если } x \leq 1, \\ (x-1)^3, & \text{если } 1 < x \leq 2, \\ 1, & \text{если } x > 2. \end{cases}$$

а) Постройте график функции распределения $F(x)$.

б) Найдите плотность вероятности $f(x)$ и постройте ее график.

Задание 2. Непрерывная случайная величина X подчинена закону распределения с плотностью вероятности $f(x)$:

$$f(x) = \begin{cases} 0 & \text{при } x < 0, \\ \frac{2}{9}x & \text{при } 0 \leq x \leq 3, \\ 0 & \text{при } x > 3. \end{cases}$$

Найдите математическое ожидание $M(X)$, дисперсию $D(X)$, среднее квадратическое отклонение $\sigma(X)$, а также вероятность попадания СВ X в интервал $(1; 1,5)$.

9. Практическая работа № 5

Тема: «Построение вариационного ряда и его графической диаграммы.

Расчет по заданной выборке ее числовых характеристик»

Цель: приобрести практические навыки по построению дискретных вариационных рядов и их полигонов, а также вычислению числовых характеристик выборок различными методами.

Методические указания

Опр. Совокупность всех мысленно-возможных объектов данного вида, над которыми производится наблюдение с целью получения конкретных значений определенной случайной величины, называется **генеральной совокупностью**.

Опр. Часть отобранных объектов из генеральной совокупности называется **выборочной совокупностью (выборкой)**, а число объектов выборки – ее **объемом**.

Опр. **Дискретным вариационным рядом распределения** называется ранжированная совокупность вариантов x_i с соответствующими им частотами m_i (или частостями p_i^*).

Опр. Значение СВ, соответствующее определенной группе сгруппированного ряда наблюдаемых данных, называется **вариантой**, а численность отдельной группы сгруппированного ряда наблюдаемых данных – **частотой варианты (ее весом)**.

Опр. Разность между наибольшим и наименьшим значениями вариантов называют **размахом выборки**.

Опр. **Полигоном частот (относительных частот)** называется ломаная, отрезки которой соединяют точки $(x_1; m_1), (x_2; m_2), \dots, (x_k; p_k^*), \dots$.

Опр. **Интервальным вариационным рядом распределения** называется упорядоченная совокупность интервалов варьирования значений СВ с соответствующими частотами (или частостями) попадания значений величины в каждый из них.

Опр. **Гистограммой частот** называется ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы длины h , а высоты равны отношению $\frac{m_i}{h}$ (плотность частоты), где m_i – сумма частот вариантов, попавших в i -ый интервал.

Опр. **Средним арифметическим \bar{X}** наблюдаемых значений СВ X называется частное от деления суммы всех этих значений на их число, т.е.

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

Если данные наблюдений представлены в виде дискретного вариационного ряда, где x_1, x_2, \dots, x_k – наблюдаемые варианты, а m_1, m_2, \dots, m_k – соответствующие им частоты ($\sum_{i=1}^k m_i = n$), то по определению,

$$\bar{X} = \frac{\sum_{i=1}^k x_i m_i}{n}$$

Задача. Из генеральной совокупности извлечена выборка объема $n = 50$

x_i	2	5	7	10
m_i	16	12	8	14

Найти выборочную среднюю.

$$/ \bar{X} = \frac{2 \cdot 16 + 5 \cdot 12 + 7 \cdot 8 + 10 \cdot 14}{50} = 5,76 \quad /$$

(!!) Если первоначальные варианты x_i – большие числа, то для упрощения расчёта целесообразно вычесть из каждой варианты одно и то же число C , т.е. перейти к условным вариантам $u_i = x_i - C$, тогда $\bar{X} = C + \frac{\sum_{i=1}^k x_i \cdot m_i}{n}$

Задача. Найти выборочную среднюю по данному распределению выборки объема $n = 10$

x_i	1250	1270	1280
m_i	2	5	3

/ Перейдём к условным вариантам $u_i = x_i - 1270$, получим

u_i	-20	0	10
m_i	2	5	3

$$\bar{X} = 1270 + \frac{-20 \cdot 2 + 0 \cdot 5 + 10 \cdot 3}{10} = 1270 - 1 = 1269 \quad /$$

Опр. *Выборочной дисперсией $D^*(X)$ значений СВ X является среднее арифметическое квадратов отклонений наблюдаемых значений этой величины от их среднего арифметического.*

$$\text{Т.о.:} \quad D^*(X) = \frac{\sum_{i=1}^k (x_i - \bar{X})^2 m_i}{n} \quad \text{или} \quad D^*(X) = \sum_{i=1}^k (x_i - \bar{X})^2 p_i^*$$

Опр. *Выборочное среднеквадратическое отклонение:*

$$\sigma^*(X) = \sqrt{D^*(X)}$$

Наряду с выборочной дисперсией $D^*(X) = \frac{\sum_{i=1}^k (x_i - \bar{X})^2 m_i}{n}$ в качестве приближенного значения генеральной дисперсии $D(X)$ используют величину, которую называют *исправленной выборочной дисперсией*:

$$S^2 = \frac{\sum_{i=1}^k (x_i - \bar{X})^2 m_i}{n-1}$$

$$\text{Очевидно: } S^2 = \frac{n D^*(X)}{n-1}.$$

Задача. Найти выборочную и исправленную выборочную дисперсии по данному распределению выборки

x_i	-2	0	4
m_i	2	3	5

$$/ \bar{X} = \frac{-2 \cdot 2 + 0 \cdot 3 + 4 \cdot 5}{10} = 1,6$$

$$\bar{X}^2 = \frac{4 \cdot 2 + 0 \cdot 3 + 16 \cdot 5}{10} = 8,8$$

$$D^*(X) = 8,8 - 1,6^2 = 6,24$$

$$S^2 = \frac{10 \cdot 6,24}{9} = 6,93 \text{ /}$$

Содержание задания

Задание 1. Постройте вариационный ряд и полигон относительных частот по данным выборки: 5, 8, 7, 6, 7, 8, 9, 5, 10, 8, 7, 5, 6, 6, 9, 7, 8, 6, 7, 10.

Найдите размах выборки, ее моду и медиану.

Задание 2. В техникуме проводилось тестирование по философии, содержащее 60 вопросов. Данные о результатах тестирования группы из 25 студентов имеют вид: 44, 36, 56, 60, 50, 48, 55, 24, 52, 52, 54, 45, 43, 60, 40, 52, 54, 56, 49, 59, 58, 32, 50, 60, 60. Составьте интервальный вариационный ряд и постройте гистограмму частот.

Задание 3. Найдите выборочную среднюю, выборочную и исправленную выборочную дисперсии по данному распределению выборки:

x_i	1	3	4	7
m_i	15	10	20	5

10. Практическая работа № 6

Тема: «Расчет сводных характеристик выборки различными методами»

Цель: приобрести практические навыки по вычислению числовых характеристик выборок различными методами.

Методические указания

1. Метод произведений вычисления выборочных средней и дисперсии

Пусть выборка задана в виде распределения равноотстоящих вариантов и соответствующих им частот. В этом случае удобно находить выборочные среднюю и дисперсию методом произведений по формулам:

$$\bar{X} = M_1^* \cdot h + C \quad \text{и} \quad D^*(X) = (M_2^* - (M_1^*)^2) \cdot h^2, \quad \text{где}$$

h – шаг (разность между двумя соседними вариантами)

C – ложный нуль (варианта, которая расположена примерно в середине вариационного ряда)

$u_i = \frac{x_i - C}{h}$ – условная варианта

$M_1^* = \frac{\sum m_i u_i}{n}$ – условный момент 1-го порядка

$M_2^* = \frac{\sum m_i u_i^2}{n}$ – условный момент 2-го порядка

Задача. Найти методом произведений выборочную среднюю и выборочную дисперсии по заданному распределению выборки объема $n = 100$.

x_i	12	14	16	18	20	22
m_i	5	15	50	16	10	4

/ 1) Найдём шаг $h = 14 - 12 = 2$

2) В качестве ложного нуля возьмём варианту, имеющую наибольшую частоту, т.е. $C = 16$.

2) Составим расчётную таблицу

x_i	m_i	u_i	$m_i u_i$	$m_i u_i^2$	$m_i(u_i + 1)^2$
12	5	-2	-10	20	5
14	15	-1	-15	15	0
16	50	0	0	0	50
18	16	1	16	16	64
20	10	2	20	40	90
22	4	3	12	36	64
Итого	100		23	127	273

Контроль: $\sum m_i (u_i^2 + 2 u_i + 1) = \sum m_i u_i^2 + 2 \sum m_i u_i + n = 100 + 23 + 127 = 273$

3) Вычислим условные моменты $M_1^* = \frac{\sum m_i u_i}{n} = \frac{23}{100} = 0,23$

$$M_2^* = \frac{\sum m_i u_i^2}{n} = \frac{127}{100} = 1,27$$

4) Найдём искомые выборочные среднюю и дисперсию

$$\bar{X} = M_1^* \cdot h + C = 0,23 \cdot 2 + 16 = 16,46 \quad D^*(X) = (M_2^* - (M_1^*)^2) \cdot h^2 = (1,27 - 0,23^2) \cdot 4 = 4,87 /$$

(!!) Если первоначальные варианты не являются равноотстоящими, то интервал, в котором заключены все варианты выборки делят на несколько равных, длины h , частичных интервалов (каждый интервал должен содержать не менее 8 – 10 вариантов). Затем находят середины частичных интервалов, которые и образуют последовательность равноотстоящих вариантов. В качестве частоты каждой середины интервала принимают сумму частот вариантов, которые попали в соответствующий частичный интервал.

При вычислении выборочной дисперсии для уменьшения ошибки, вызванной группировкой, делают поправку Шеппарда, а именно: вычитают из найденной дисперсии $\frac{1}{12}$ квадрата длины частичного интервала.

Т.о., с учетом поправки: $D^* = D^* - \frac{h^2}{12}$

Задача. Найти методом произведений выборочную среднюю и выборочную дисперсии по заданному распределению выборки объёма $n = 100$.

x_i	2	3	7	9	11	12,5	16	18	23	25	26
m_i	3	5	10	6	10	4	12	13	8	20	9

/ 1) Разобьём интервал 2 - 26 на 4 частичных интервала длины $h = 6$: 2 - 8; 8 - 14; 14 - 20; 20 - 26. Приняв середины этих интервалов в качестве новых вариантов u_i , получим равноотстоящие варианты: 5; 11; 17 и 23.

2) Составим вариационный ряд:

y_i	5	11	17	23
m_i	18	20	25	37

3) Пользуясь методом произведений найдем \bar{X} (15,86) и D^* (45,14).

С учетом поправки Шеннарда $D^* = 45,14 - \frac{6^2}{12} = 42,14$ /

2. Метод сумм вычисления выборочных средних и дисперсии

При использовании метода сумм моменты 1-го и 2-го порядков находятся по формулам:

$$M_1^* = \frac{d_1}{n} \quad \text{и} \quad M_2^* = \frac{s_1 + 2s_2}{n}, \text{ где}$$

$$d_1 = a_1 - b_1, \quad s_1 = a_1 + b_1, \quad s_2 = a_2 + b_2$$

Задача. Найти методом сумм выборочную среднюю и выборочную дисперсии по заданному распределению выборки объёма $n = 100$.

x_i	48	52	56	60	64	68	72	76	80	84
m_i	2	4	6	8	12	30	18	8	7	5

/ 1) Составим расчётную таблицу:

x_i	m_i	$b_1 = 72$	$b_2 = 70$
48	2	2	2
52	4	6	8
56	6	12	20
60	8	20	40
64	12	32	0
68	30	0	0
72	18	38	0
76	8	20	37
80	7	12	17
84	5	5	5
	$n = 100$	$a_1 = 75$	$a_2 = 59$

В качестве ложного нуля C выберем варианту 68, которая имеет наибольшую частоту и находится примерно в середине таблицы. В клетках строки, содержащей ложный нуль, запишем нули. В 4-ом столбце над и под уже помещённым нулём запишем ещё по одному нулю. В оставшихся незаполненными над нулём клетках 3-го столбца (исключая самую верхнюю) запишем последовательно накопленные частоты: 2, $2 + 4 = 6$, $6 + 6 = 12$, ... Сложив все накопленные частоты, получим число $b_1 = 72$.

Аналогично поступим с клетками под нулём. При этом получим число a_1 .

Заполняя 4-ый столбец, суммируем частоты 3-го столбца. Т.о. находим b_2 и a_2 .

2) Найдём $d_1 = a_1 - b_1 = 3$

$$s_1 = a_1 + b_1 = 147$$

$$s_2 = a_2 + b_2 = 129$$

3) Вычислим условные моменты:

$$M_1^* = 0,03, \quad M_2^* = \frac{147 + 2 \cdot 129}{100} = 4,05$$

4) Учитывая, что $h = 4$ и ложный нуль $C = 68$, находим:

$$\bar{X} = M_1^* \cdot h + C = 0,03 \cdot 4 + 68 = 68,12$$

$$D^*(X) = (M_2^* - (M_1^*)^2) \cdot h^2 = (4,05 - 0,03^2) \cdot 14 = 64,87 \quad /$$

Содержание задания

Задание 1. Найдите методом произведений выборочную среднюю и выборочную дисперсии по заданному распределению выборки:

x_i	5	10	15	20	25	30
m_i	3	7	17	13	6	4

Задание 2. Найдите методом сумм выборочную среднюю и выборочную дисперсии по заданному распределению выборки:

x_i	32	34	36	38	40	42	44	46	48
m_i	5	9	10	15	21	22	11	6	1

11. Практическая работа № 7

Тема: «Нахождение выборочного уравнения прямой линии регрессии»

Цель: приобрести практические навыки по нахождению выборочного уравнения линии регрессии.

Методические указания

Опр. Корреляция – это статистическая взаимосвязь двух или нескольких случайных величин.

Опр. Зависимость одной случайной величины от значений, которые принимает другая случайная величина, называется **регрессией**, а аналитическая форма представления этой зависимости – **уравнением регрессии**.

Выборочное уравнение прямой линии регрессии имеет вид:

$$y - \bar{y} = r_b \frac{\sigma_y}{\sigma_x} (x - \bar{x}),$$

где \bar{x} и \bar{y} – выборочные средние признаков X и Y , σ_x и σ_y – выборочные средние квадратические отклонения, r_b – выборочный коэффициент корреляции, причем

$$r_b = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\sigma_x\sigma_y}$$

Если данные наблюдений над признаками X и Y заданы в виде корреляционной таблицы с равноотстоящими вариантами, то целесообразно перейти к условным вариантам:

$$u_i = \frac{x_i - C_1}{h_1} \quad \text{и} \quad v_i = \frac{y_i - C_2}{h_2},$$

где C_1 - ложный нуль, h_1 - шаг вариант X, а C_2 - ложный нуль, h_2 - шаг вариант Y.

В этом случае выборочный коэффициент корреляции

$$r_B = \frac{\sum n_{uv} uv - n \bar{u} \bar{v}}{n \sigma_u \sigma_v}$$

Слагаемое $\sum n_{uv} uv$ удобно вычислять, используя специальную расчетную таблицу.

Величины $\bar{u}, \bar{v}, \sigma_u, \sigma_v$ могут быть найдены либо методом произведения, либо непосредственно по формулам:

$$\bar{u} = \frac{\sum u_i m_i}{n}, \quad \bar{v} = \frac{\sum v_i m_i}{n}, \quad \sigma_u = \sqrt{\overline{u^2} - \bar{u}^2}, \quad \sigma_v = \sqrt{\overline{v^2} - \bar{v}^2}$$

Зная эти величины, можно определить входящие в уравнение регрессии величины по формулам:

$$\bar{x} = \bar{u} h_1 + C_1, \quad \bar{y} = \bar{v} h_2 + C_2, \quad \sigma_x = \sigma_u h_1, \quad \sigma_y = \sigma_v h_2$$

Пример. Найти выборочное уравнение прямой линии регрессии Y на X по данным, приведенным в корреляционной таблице:

Y	X					m _y
	20	25	30	35	40	
16	4	6	-	-	-	10
26	-	8	10	-	-	18
36	-	-	32	3	9	44
46	-	-	4	12	6	22
56	-	-	-	1	5	6
m _x	4	14	46	16	20	n = 100

/По условию: $h_1 = 5, h_2 = 10$.

Составим корреляционную таблицу в условных вариантах, выбрав в качестве ложных нулей $C_1 = 30$ и $C_2 = 36$.

V	U					m _v
	-2	-1	0	1	2	
-2	4	6	-	-	-	10
-1	-	8	10	-	-	18
0	-	-	32	3	9	44
1	-	-	4	12	6	22
2	-	-	-	1	5	6
m _U	4	14	46	16	20	n = 100

Найдем:

$$\bar{U} = \frac{\sum u_i m_i}{n} = \dots = 0,34$$

$$\bar{V} = \frac{\sum v_i m_i}{n} = \dots = -0,04$$

Найдем вспомогательные величины:

$$\overline{U^2} = \frac{\sum u_i^2 m_i}{n} = \dots = 1,26$$

$$\overline{V^2} = \frac{\sum v_i^2 m_i}{n} = \dots = 1,04$$

Найдем:

$$\sigma(U) = \sqrt{\overline{U^2} - \bar{U}^2} = \dots = 1,07$$

$$\sigma(V) = \sqrt{\overline{V^2} - \bar{V}^2} = \dots = 1,02$$

Найдем $\sum uv m_{uv}$, для чего составим расчетную таблицу:

V	U					$\sum um_{uv}$	$v \cdot \sum um_{uv}$
	-2	-1	0	1	2		
-2	-8 4	-6 6	-	-	-	-14	28
-1	-	-8 8	0 10	-	-	-8	8
0	-	-	0 32	3 3	18 9	21	0
1	-	-	0 4	12 12	12 6	24	24
2	-	-	-	1 1	10 5	11	22
$\sum vm_{uv}$	-8	-20	-6	14	16		82
$u \cdot \sum vm_{uv}$	16	20	0	14	32	82	

1) Произведение um_{uv} варианты u на частоту m_{uv} записывают в правом верхнем углу клетки, содержащей значение частоты

2) Складывают все числа, помещенные в правых верхних углах клеток одной строки.

3) Наконец, умножают варианту v на найденную сумму и складывают все числа последнего столбца, получая тем самым искомое число $\sum uvm_{uv}$.
Т.о.: $\sum uvm_{uv} = 82$

Для контроля аналогичные вычисления производят по столбцам: произведения vm_{uv} записывают в левый нижний угол соответствующей клетки; складывают все числа, помещенные в левых нижних углах клеток одного столбца; затем умножают полученные суммы на варианту u и складывают все числа последней строки. Полученное число также равно искомой сумме $\sum uvm_{uv}$.

Совпадение сумм свидетельствует о правильности вычислений.

Найдем выборочный коэффициент корреляции:

$$R = \frac{\sum uvm_{uv} - n\bar{U}\bar{V}}{n\sigma(U)\sigma(V)} = \frac{82 - 100 \cdot 0,34 \cdot (-0,04)}{100 \cdot 1,07 \cdot 1,02} = 0,76$$

$$\bar{X} = \bar{U}h_1 + C_1 = 0,34 \cdot 5 + 36 = 31,70$$

$$\bar{Y} = \bar{V}h_2 + C_2 = -0,04 \cdot 10 + 30 = 35,60$$

$$\sigma(X) = \sigma(U)h_1 = 1,07 \cdot 5 = 5,35$$

$$\sigma(Y) = \sigma(V)h_2 = 1,02 \cdot 10 = 10,2$$

$$\text{Подставим найденные величины в уравнение: } y - \bar{Y} = R \frac{\sigma(Y)}{\sigma(X)} (x - \bar{X})$$

$$\text{Получим: } y - 35,60 = 0,76 \cdot \frac{10,2}{5,35} (x - 31,70)$$

$$y = 1,45x - 10,36 /$$

Содержание задания

Задание Найдите выборочное уравнение прямой линии регрессии Y на X по данным, приведенным в корреляционной таблице:

Y	X						n_y
	2	7	12	17	22	27	
15	3	3					6
25		4	5				9
35			35	7	8		50
45			9	10	6		25
55				2	5	3	10
n_x	3	7	49	19	19	3	100

12. Список рекомендуемой литературы

1. М.С. Спирина, П.А. Спирин. Теория вероятностей и математическая статистика / Учебник/ - М: Издательский центр «Академия», 2021, 352 с.
2. Т.Ю. Беляева «Элементы комбинаторики» /Учебное пособие для студентов специальности 09.02.07/ - ГБПОУ КК «АМТ», 2023
3. Т.Ю. Беляева «Основы теории вероятностей. Раздел 1. Случайные события» /Учебное пособие для студентов специальности 09.02.07/ - ГБПОУ КК «АМТ», 2023
4. Т.Ю. Беляева «Основы теории вероятностей. Раздел 2. Случайные величины» /Учебное пособие для студентов специальности 09.02.07/ - ГБПОУ КК «АМТ», 2023
5. Т.Ю. Беляева «Элементы математической статистики» /Учебное пособие для студентов специальности 09.02.07/ - ГБПОУ КК «АМТ», 2023