Министерство образования и науки Российской Федерации
Федеральное агентство по образованию

Государственное образовательное учреждение
высшего профессионального образования
«Оренбургский государственный университет»

Кафедра прикладной математики

В.П. Матвейкина

СБОРНИК ЗАДАЧ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

Методические указания

Рекомендовано к изданию Редакционно-издательским советом Государственного образовательного учреждения высшего профессионального образования «Оренбургский государственный университет»

Оренбург
ИПК ГОУ ОГУ
2010

УДК 519.2(07)
ББК 22.17я7
М 33

Рецензент –доцент, кандидат технических наук, И.В. Влацкая

Матвейкина, В. П.
М 33 Сборник задач по теории вероятностей и математической статистике: методические указания /В.П. Матвейкина; Оренбургский гос. ун-т. – Оренбург: ОГУ, 2010. – 37с.

Методические указания представляют собой сборник задач, для проведения практических занятий по разделам теории вероятностей и математической статистике со студентами специальностей «География», «Геологическая съемка, поиск и разведка месторождений полезных ископаемых», «Геология нефти и газа» геолого-географического факультета.

В сборник включены задачи профессиональной направленности, цель которых формирование математической компетентности студентов геолого-географических специальностей, являющейся важной составляющей профессиональной компетентности будущих специалистов.

Методические указания предназначены как для преподавателей, так и для студентов естественно-научных специальностей. Рекомендуются для самостоятельной работы студентов всех форм обучения: очной, очно-заочной и заочной.

УДК 519.2(07)
ББК 22.17я7

Содержание

Введение....4
1 Понятие вероятности случайного события, формулы сложения и умножения вероятностей, формула полной вероятности, формула Байеса....5
2 Последовательность независимых испытаний (схема Бернулли), локальная и интегральная теоремы Муавра – Лапласа, теорема Пуассона....10
3 Случайные величины и функции распределения вероятностей. Числовые характеристики распределений....13
4 Закон больших чисел....18
5 Математическая статистика....21
Список использованных источников....31
Приложение А....32

Введение

Исследования природных процессов и изучение их закономерностей приводят к построению математических моделей, в основе которых лежат вероятностные и статистические методы. Теоретико-вероятностная интуиция будущего специалиста, умение строить математические модели реальных процессов вырабатывается при решении профессионально-направленных задач.

Методы теории вероятностей были впервые применены в геологии сибирским золоторазведчиком Н. Псаревым, который в 1899 г. сформулировал и решил две задачи, являющиеся важнейшими в разведочном деле: об ошибке оценки среднего содержания золота в россыпи и о числе шурфов, необходимых для того, чтобы эта ошибка не вышла (при заданной вероятности) за допустимые пределы.

Дальнейшую статистическую обработку разведочного материала осуществил томский профессор С. Ю. Доборжинский, его труды были опубликованы в 1908-1912 гг. Математические методы с разведочном деле применяли также В.И. Бауман, Г. Гувер, Э. Бейтли, А.К. Болдырев и др.

В дальнейшем математические методы нашли применение в геоморфологии, структурной геологии, нефтяной геологии, геохимии, минералогии, инженерной геологии, сейсмологии, географии, регионоведении и т. д.

Данные методические указания представляют собой сборник задач по теории вероятностей и математической статистике, для специальностей «Геология нефти и газа», «Геологическая съемка, поиск и разведка месторождений полезных ископаемых», «География».

Сборник состоит из пяти разделов. В каждом разделе приведены краткие теоретические сведения и необходимые формулы для решения задач, которые имеют профессиональную направленность, что особенно важно для формирования профессиональной компетентности будущих специалистов геолого-географического профиля.

Все задачи снабжены ответами, а некоторые из них указаниями и решениями.

1 Понятие вероятности случайного события, формулы сложения и умножения вероятностей, формула полной вероятности, формула Байеса

Каждый «эксперимент» завершается некоторым исходом или событием. *Случайным событием (или просто событием)* называется событие, которое может произойти, а может и не произойти в результате осуществления данного эксперимента. События обозначают латинскими буквами *A, B, C,* .... Если каждое осуществление данного эксперимента *a priori* вызывает появление события *A*, это событие называется *достоверным.* Если событие *A* заведомо не произойдет при осуществлении данного эксперимента, то оно называется *невозможным*. Достоверное событие будем обозначать буквой *U*; невозможное – буквой *V*.

Суммой (объединением) двух событий *A* и *B* называется событие $A+B$ (или $A \cup B$), состоящее в том, что происходит хотя бы одно из них (т. е. либо *A*, либо *B*, либо *A* и *B*). *Произведением* (пересечением) *AB* (или $A \cap B$) двух событий называется событие, состоящее в совместном появлении *A* и *B*. Если каждое появление события *A* сопровождается появлением *B*, то пишут $A \subset B$ и говорят, что *A влечет B*, или *A* есть частный случай *B*. События *A* и *B* называются *несовместимыми*, если $AB = V$. События *A* и $\overline{A}$ называются *противоположными*, если $A\overline{A} = V$ и $A + \overline{A} = U$. События $A_1, A_2, ..., A_n$ образуют *полную группу событий*, если $A_1 + A_2 + ...A_n = U$ и $A_i Aj = V$, $i = 1, 2, ..., n(i \neq j)$.

Аксиоматическое определение вероятности. Пусть имеется множество Ω некоторых объектов ω, которые назовем *элементарными событиями*. (Конечно, эти объекты всегда можно изображать точками в евклидовом пространстве нужного числа измерений). Образуем какую-нибудь совокупность **B** подмножеств множества Ω, обладающую следующими свойствами: 1) все множество Ω является элементом **B**; 2) если $A \in \mathbf{B}$, то и множество $\overline{A}$, состоящее из всех элементов Ω, не принадлежащих *A*, также принадлежит **B**; 3) если $A_1, A_2, ..., A_n, ...$ – любая конечная или счетная последовательность множеств из **B**, то их сумма и произведение (пересечение) также принадлежит **B**. Эта совокупность множеств **B** называется *борелевским телом множеств*, или σ - алгеброй множеств. Элементы A множества $\mathbf{B}^{g}$ называются случайными событиями. Тогда множество Ω будет достоверным событием U, а пустое множество – невозможным событием V.

Вероятностью случайного события A называется значение неотрицательной вполне аддитивной функции $P(A)$, определенной на множестве $\mathbf{B}$ и равной 1 для множества Ω. Это значит, что для любого $A \in \mathbf{B}$ $P(A) \geq 0$, и для любого конечного или счетного набора непересекающихся множеств A_i из $\mathbf{B}$ $P\left\{\sum_i A_i\right\} = \sum_i P(A_i)$.

Следствия: 1) $0 \leq P(A) \leq 1$ для любого A; 2) $P(V) = 0$; 3) $P(\overline{A}) = 1 - P(A)$; 4) если $A \subset B$, то $P(A) \leq P(B)$.

Классическое определение вероятности. Пусть в результате эксперимента возможны только n несовместимых и равновозможных исходов $A_1, A_2, \ldots, A_n$. Пусть событие B есть сумма определенных m из них, т. е. $B = A_{i_1} + A_{i_2} + \ldots + A_{i_m}$. Тогда $P(B) = \dfrac{m}{n}$.

Формула сложения вероятностей. Для любых двух событий A и B $P(A+B) = P(A) + P(B) - P(AB)$.

$P(A/B)$ обозначает условную вероятность события A при условии, что событие B произошло. Для независимых событий $P(A/B) = P(A)$; $P(B/A) = P(B)$.

Формула умножения вероятностей. Для любых двух событий

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

Формула полной вероятности. Вероятность события B, которое может наступить лишь при появлении одного из несовместных событий (гипотез) $A_1, A_2, \ldots A_n$, образующих полную группу, равна сумме произведений вероятностей каждой из гипотез на соответствующую условную вероятность события B:

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

Замечание: Поскольку заранее не известно, какое из событий $A_1, A_2, \dots A_n$ наступит, их называют гипотезами.

Формула Байеса. В тех же условиях

$$P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{\sum_{k=1}^{n} P(A_k) P(B/A_k)} .$$

Формула Байеса позволяет переоценить вероятность гипотезы A_i после того, как становиться известным результат испытания, в итоге которого появилось событие B.

Задачи:

1. С помощью специального прибора регистрируется направление φ и скорость ветра υ в данном месте Земли. Прибор устроен таким образом, что позволяет определять скорость ветра сколь угодно точно, а регистрация направления ветра возможна лишь с точностью до 2°. Установить, наблюдаемы ли в данном эксперименте события: $A = \left\{ (\upsilon, \varphi) \middle| \upsilon < 12 \text{км/ч}, \ \varphi = 343^\circ 35' \right\}$,

$B = \left\{ (\upsilon, \varphi) \middle| \upsilon = 15{,}5 \text{км/ч}, 340^\circ \le \varphi \le 350^\circ \right\}$, $C = \left\{ (\upsilon, \varphi) \middle| \upsilon \ge 3{,}25 \text{км/ч}, 48^\circ \le \varphi \le 51^\circ \right\}$.

2. Какова вероятность того, что случайно выбранная на глобусе точка лежит: а) за полярным кругом (66°33′ северной широты); б) между 60° и 30° северной широты; в) между 10° и 40° западной долготы?

3. В рюкзаке у геолога лежат пять одинаковых мешочков с образцами – 2 гранита и 3 селита. Наугад вынимают один мешочек. Какова вероятность, что в нем находится гранит?

4. На шарик нанесена сетка географических координат. Шарик брошен на плоскость. Найти вероятность того, что он прикоснется точкой:

а) которая находится в области между 0-м и 90-м градусами восточной долготы;

б) которая находится между 45-м и 90-м градусами северной широты. Предполагается, что выпадения областей, имеющих равные площади, равновероятны.

5. Сферическая частица радиуса r случайным образом вертикально падает на наклонное проволочное сито с квадратными ячейками. Угол наклона сита к горизонту α, диаметр проволоки d, расстояние между осевыми линиями проволок l. Найти вероятность того, что частица свободно пройдет через сито.

6. Необходимо двумя станками пробурить гидрологические скважины. На геологической карте этого района имеется всего 22 приблизительно равных по площади участка, в том числе 4 несмежных участка, показанные как площади юрских отложений. Точки для бурения скважин выбирают наугад поочередно, но так, чтобы на один и тот же участок не попали две скважины. Определить вероятность того, что точка для второй скважины попадем на участок распространения юрских отложений, если известно, что первая скважина уже задана на породы этого возраста.

7. На 6 полках находятся в одинаковых капсулах 85 измельченных проб вольфрамовой руды, в том числе на первой полке12 проб, на второй 12, на третьей 12, на четвертой 14, на пятой 14 и на шестой 21. При этом на первых трех полках лежат пробы, проанализированные в лаборатории Москвы, на следующих двух – в лаборатории Одессы и на последней – в лаборатории Иркутска. Все пробы были взяты с разных участков и обработаны одним способом. В числе участков есть западный участок, пробы в этого участка брались постепенно и поэтому попали на разные полки. Всего таких проб 50, в том числе на первой полке 10, на второй 10, на третьей 10, на четвертой 7, на пятой 7, на шестой 6. С какой-то полки произвольно взята одна проба, оказавшаяся с западного участка. Какова вероятность, что: 1) на полке, с которой взята проба, было 10 проб? 2) на полке, с которой взята проба, было 7 проб? 3) на полке с которой взята проба, было 6 проб?

8. В сфере радиуса R случайно и независимо друг от друга разбросано N точек.

а) Найти вероятность того, что расстояние от центра сферы до ближайшей точки будет не меньше a.

б) Найти предел этой вероятности, если $R \to \infty$ и $\frac{N}{R^3} \to \frac{4}{3}\pi\lambda$. (Эта задача взята из звездной астрономии, где R измеряется в парсеках, а λ в окрестности Солнца равна (приближенно) 0,0063).

9. Некоторое событие может произойти в любой из дней недели с одинаковыми вероятностями. Найти вероятность того, что 12 осуществлений этого события подряд выпадут только во вторники и четверги. Согласуется ли этот случай с предположением о равновероятности осуществления события в любой из дней недели?

10. Крупные метеориты падают на землю в среднем раз в месяц, причем попадания их в любые области, имеющие равные площади, равновероятны.

Найти вероятность того, что в течение 10 лет упадет не менее двух метеоритов в область, ограниченную: а) меридианами $30°$ в.д. и $60°$ в.д. и параллелями $40°$ с. ш. и $60°$ с. ш.; б) на территории России.

2 Последовательность независимых испытаний (схема Бернулли), локальная и интегральная теоремы Муавра – Лапласа, теорема Пуассона

Пусть осуществляется n независимых повторений некоторого эксперимента, в результате каждого из которых может произойти событие A. μ – число фактических появлений события A в этой серии. Тогда $\frac{\mu}{n}$ называется частотой события A в данной серии из n независимых испытаний. Пусть вероятность события A в каждом испытании равна p; вероятность не наступления события A равна $q = 1 - p$; m – число возможных появлений событий A в этой серии (до ее осуществления). $P_n(m)$

– вероятность того, что событие A произойдет ровно m раз в данной серии из n испытаний.

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

Последовательность чисел $P_n(m)$, $m = 0,1,2,...,n$, называется *биномиальным законом распределения вероятностей*. Если $np - q$ есть число целое, то наибольшие значения $P_n(m)$ будут при $m = m_0 = np - q$ и $m = m_0 + 1 = p(n+1)$. Если $np - q$ не целое, то $P_n(m)$ будет наибольшим при

$$m = m_0 = \left[(n+1)\, p \right] \quad \sum_{m=0}^{n} P_n(m) = 1.$$

Число $m = m_0$ называется наивероятнейшим числом.

Пусть в результате одного испытания все исходы $A_1,...,A_k$ образуют полную группу событий. $P(A_i) = p_i$, $\sum_{i=1}^{k} p_i = 1$; $p_n(m_1, m_2, ..., m_k)$ – вероятность появления m_1 раз события $A_1,...,m_k$ раз события A_k в серии из n независимых повторений данного испытания. Тогда

$$P_n(m_1, m_2, ..., m_k) = \frac{n!}{m_1!m_2!...m_k!} p_1^{m_1} p_2^{m_2} ... p_k^{m_k} \left(m_1 + ... + m_k = n \right).$$

Числа $P_n(m_1, m_2, ..., m_k)$ для всех возможных неотрицательных значений m_i, таких, что $m_1 + m_2 + ... + m_k = n$, образуют полиномиальное (или мультиномиальное) распределение вероятностей.

Локальная теорема Муавра-Лапласа. Пусть дана серия из n независимых испытаний. $0 < p < 1$, $q = p - 1$, $x = \dfrac{m - np}{\sqrt{npq}}$. Если $n \to \infty$ и $m \to \infty$ так, что величина x остается ограниченной, то

$$P_n(m) \sim \frac{1}{\sqrt{2\pi npq}} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{npq}} \varphi(x).$$

Знак асимптотического равенства ~ означает, что отношение левой части этого «равенства» к правой стремится к 1 при $n \to \infty$.

Интегральная теорема Муавра-Лапласа.

$$P\left\{ a \le \frac{m - np}{\sqrt{npq}} < b \right\} \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \text{ при } n \to \infty$$

равномерно относительно a и b $-\infty \le a, b \le \infty$.

Теорема Пуассона. Пусть $0 < a < b < \infty$, $a_n = np$. Если при $n \to \infty$ $p \to 0$ и $a < a_n < b$, то

$$P_n(m) \sim \frac{a_n^m}{m!} e^{-a_n}.$$

В частности, если при $n \to \infty$ $a_n \to a$, $0 < a < \infty$, то

$$P_n(m) \sim \frac{a^m}{m!} e^{-a}.$$

Положим

$$P(m) = \frac{a^m}{m!} e^{-a}, \quad \sum_{m=0}^{\infty} P(m) = 1.$$

Последовательность чисел $P(m)$ для $m = 0,1,2,...$ образует *закон распределения вероятностей Пуассона*.

Задачи:

1. На основании исследования многих тысяч кристаллов кварца из 8 месторождений, расположенных в разных странах, оценена вероятность встречи правых кристаллов p=0,49. Пусть из этой совокупности наудачу взято 10 кристаллов кварца. Требуется найти вероятность того, что среди этих 10 кристаллов окажется не более трех правых.

2. Три четверти территории района представляют собой солонец. Если взять 10 образцов почвы (в 10 наудачу выбранных точках), то из них несколько образцов могут оказаться солонцовыми. Найти наиболее вероятное число появления солонцовых образцов.

3. Наблюдениями установлено, что в некоторой местности в сентябре в среднем бывает 12 дождливых дней. Какова вероятность, что из случайно взятых в этом месяце 8 дней 3 дня окажутся дождливыми?

4. Вероятность появления события в каждом из независимых испытаний равна 0,8. сколько нужно произвести испытаний, чтобы с вероятностью 0,9 можно было ожидать, что событие появится не менее 75 раз?

5. Для космического корабля вероятность столкновения в течение часа полета с метеоритом, масса которого не меньше m_0, равна 0,001. Найти практически достоверные границы числа столкновений с таким метеоритом в течение 3 месяцев полета – с 1 июня по 31 августа, если вероятность практической достоверности принимается в данном случае равной 0,9995.

3 Случайные величины и функции распределения вероятностей. Числовые характеристики распределений

Пусть дано пространство элементарных событий $\Omega = \{\omega\}$. $\mathbf{B} = \{A\}$ – борелевское тело множеств, построенное на Ω. Пусть на $\mathbf{B}$ задано распределение вероятностей, т.е. определена такая неотрицательная вполне аддитивная функция $P(A)$, что $P(\Omega) = 1$. Тогда все элементы из $\mathbf{B}$ будем называть *P-измеримыми* подмножествами Ω.

Случайной величиной называется вещественная *P*-измеримая функция, определенная на $\mathbf{B}: \xi = f(A)$. Это значит, что для любого вещественного x множество $A_x = \{\omega\}_x$, для которого $f(A_x) < x$ имеет определенную вероятность $P\{\xi < x\} = P(A_x) = F(x)$. $F(x)$ называется функцией распределения вероятностей случайной величины ξ.

Пусть ξ величина, в зависимости от случая принимающая одно из *n возможных значений* $a_{١}, a_{٢}, ..., a_n$ с вероятностями соответственно $p_{١}, p_{٢}, ..., p_n$; $\sum_{i=١}^{n} p_i = ١$. Тогда ξ называется *дискретной случайной величиной с конечным множеством возможных значений.*

ξ называется дискретной случайной величиной *со счетным множеством возможных значений*, если ξ может принимать значения $a_{١}, a_{٢}, ..., a_n, ...$; $P\{\xi = a_i\} = p_i$; $\sum_{i=١}^{\infty} p_i = ١$.

Существуют случайные величины, имеющие несчетное множество возможных значений, например все числа данного интервала или все вообще вещественные числа. Эти (а также и дискретные) случайные величины задаются с помощью функции распределения $F(x)$:

$$F(x) = P\{\xi < x\}.$$

Функция распределения всякой случайной величины обладает следующими свойствами:

1. $0 \le F(x) \le 1 \ (-\infty \le x \le \infty)$;
2. $P\{a \le \xi < b\} = F(b) - F(a)$ для любых a и b.
3. $F(x)$ монотонно возрастает на всей числовой оси;
4. $F(-\infty) = 0, F(+\infty) = 1$;
5. непрерывна слева в каждой точке и имеет не более счетного множества разрывов первого рода.

Всякая функция обладающая свойствами 1, 3, 4, 5 является функцией распределения некоторой случайной величины.

Если существует (почти везде) неотрицательная функция

$$f(x) = F'(x), f(x) \ge 0, F(x) = \int_{-\infty}^{x} f(z)dz,$$

то, $f(x)$ называется плотностью распределения вероятностей величины ξ, а величина ξ и ее закон распределения называются непрерывными.

Тогда

$$P\{a \le \xi < b\} = \int_a^b f(x)dx; \ \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Математическим ожиданием (или *средним значением*) случайной величины ξ называется число $M\xi$. Для непрерывной случайной величины $M\xi = \int_{-\infty}^{\infty} xf(x)\,dx$. Для дискретной случайной величины $M\xi = \sum_i a_i p_i$ (в случае абсолютной сходимости интеграла и ряда).

Свойства математического ожидания:

1. $MC = C$, если C – постоянная.
2. $M(\xi + \eta) = M\xi + M\eta$.

3. Если ξ и η независимые случайные величины, то $M(\xi\eta) = M\xi \cdot M\eta$.

4. $M(C\xi) = CM\xi$, если C – постоянная.

Дисперсия $D\xi$ случайной величины ξ определяется формулой:

$$D\xi = M(\xi - M\xi)^2$$

Дисперсия характеризует рассеивание возможных значений случайной величины вокруг ее среднего значения.

Иногда дисперсию удобно вычислять по формуле:

$$D\xi = M(\xi^2) - (M\xi)^2$$

Свойства дисперсии:

1. Для любой случайной величины ξ справедливо: $D\xi \geq 0$.
2. $DC = 0$,если C – постоянная.
3. Если ξ и η независимы, то $D(\xi + \eta) = D\xi + D\eta$. В противном случае $D(\xi + \eta) = D\xi + D\eta + 2M\left[(\xi - M\xi)(\eta - M\eta)\right]$.
4. $D(C\xi) = C^2 D\xi$ если C постоянная.

Для непрерывных случайных величин

$$D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 f(x)\,dx.$$

Для дискретных величин

$$D\xi = \sum_i (a_i - M\xi)^2 p_i.$$

$M\left[(\xi - M\xi)(\eta - M\eta)\right] = K_{\xi\eta}$ называется *вторым смешанным центральным моментом* случайных величин ξ и η, или их *корреляционным моментом*.

Коэффициентом корреляции случайных величин ξ и η называется число:

$$r = \frac{K_{\xi\eta}}{\sqrt{D\xi D\eta}}.$$

Коэффициент корреляции служит для оценки тесноты линейной связи между случайными величинами ξ и η: чем ближе абсолютная величина коэффициента корреляции к единице, тем связь сильнее; чем ближе абсолютная величина коэффициента корреляции к нулю, тем связь слабее.

Свойства коэффициента корреляции:

1. для любых случайных величин ξ и η справедливо неравенство: $-1 \le r \le 1$, причем $|r| = 1$ в том и только в том случае, если ξ и η связаны линейной зависимостью: $\eta = a\xi + b$ ($r = +1$, если $a > 0$, и $r = -1$, если $a < 0$), и обратно.
2. если ξ и η независимы, то $r = 0$, обратное утверждение неверно.

Случайные величины называются некоррелированными, если $r = 0$. Случайные величины называются коррелированными, если $r \neq 0$.

Задачи:

1. Радиостанция ведет передачу информации в течение 10 мксек. Работа ее происходит при наличии хаотической импульсной помехи, среднее число импульсов которой в одну секунду составляет 10^4 . Для срыва передачи достаточно попадания одного импульса помехи в период работы станции. Считая, что число импульсов помехи, попадающих в данный интервал времени, распределено по закону Пуассона, найти вероятность срыва передачи информации.

2. На плоскую сетку равноотстоящих параллельных прямых брошен жесткий контур (многоугольник правильный или неправильный, замкнутый или нет и т.п.). Найти среднее число пересечений с параллелями.

3. Город состоит из n кварталов, причем в n_j них по x_j жителей $\left(n_1 + n_2 + \ldots + n_k = n\right)$. С помощью случайного выбора без возвращения отобраны r кварталов, и в каждом из них подсчитано число жителей. Пусть $X_1, \ldots, X_r$ – соответствующие числа. Вычислить $M\left(X_1 + \ldots + X_r\right)$ и $D\left(X_1 + \ldots + X_r\right)$.

4. Точка брошена наудачу внутрь круга радиуса *R*. Вероятность попадания точки в любую область, расположенную внутри круга, пропорциональна площади области. Найти функцию распределения, математическое ожидание и дисперсию расстояния точки до центра круга.

5. Случайная величина ξ - ошибка измерительного прибора – распределена по нормальному закону с дисперсией 16 мк. Систематическая ошибка прибора отсутствует. Найти вероятность того, что в пяти независимых измерениях ошибка ξ а) превзойдет по модулю 6 мк не более трех раз; б) хотя бы один раз окажется в интервале 0,5 мк – 3,5 мк.

6. Закон ошибок при наблюдении температуры выражен по шкале Фаренгейта формулой $\varphi(t)=\sqrt{\frac{1,2}{\pi}}\cdot e^{-1,2(t-23)^2}$. Написать этот закон, приспособив его к шкале Цельсия.

7. Некоторая величина отклоняется от своего среднего значения под воздействием двух случайных факторов *A* и *B*. Среднее квадратическое отклонение, вызванное фактором *A*, равно1,2, а фактором *B* – 1,1. Коэффициент корреляции между этими уклонениями *r* равен $\frac{1}{3}$. Найти среднее квадратическое уклонение этой величины, вызываемое совместным действием обоих факторов.

4 Закон больших чисел

Неравенства Чебышева. Первая форма: если случайная величина ξ неотрицательна и имеет математическое ожидание, то

$$P\{\xi>\alpha\}<\frac{M\xi}{\alpha}.$$

Вторая форма: если $D\xi<+\infty$, то

$$P\{|\xi - M\xi| > \varepsilon\} < \frac{D\xi}{\varepsilon^2}.$$

Пусть имеется последовательность случайных величин

$$\xi_1, \xi_2, ... \xi_n, ... \tag{1}$$

Говорят, что последовательность $\xi_1, \xi_2, ... \xi_n, ...$ подчиняется закону больших чисел, если

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k \right| < \varepsilon \right\} = 1. \tag{2}$$

Теорема Чебышева. Пусть случайные величины последовательности $\xi_1, \xi_2, ... \xi_n, ...$ независимы и имеют конечные дисперсии $D\xi_n$. Тогда для $\lim\limits_{n \to \infty} P\left\{ \left| \frac{1}{n}\sum\limits_{k=1}^{n} \xi_k - \frac{1}{n}\sum\limits_{k=1}^{n} M\xi_k \right| < \varepsilon \right\} = 1$ достаточно

$$D\xi_n < C, \ n = 1, 2, ...$$

Теорема Хинчина. Если все случайные величины $\xi_1, \xi_2, ... \xi_n, ...$ независимы, одинаково распределены и имеют конечное математическое ожидание $M\xi_n = a$, то

$$\lim_{n \to \infty} P\left\{ \left| \frac{١}{n}\sum_{k=١}^{n} \xi_k - a \right| < \varepsilon \right\} = ١.$$

Теорема Маркова. Пусть случайные величины (1) как угодно зависимы. Для выполнения (2) достаточно, чтобы

$$\frac{١}{n^{٢}} D\sum_{k=١}^{n} \xi_k \to ٠, \text{ при } n \to \infty.$$

Теорема Бернулли. Пусть имеем серию из n независимых испытаний. В каждом из них $P(A) = p > 0$. $\dfrac{\mu}{n}$ - частота события А в данной серии. Тогда

$$\lim_{n\to\infty} P\left\{\left|p - \frac{\mu}{n}\right| < \varepsilon\right\} = 1.$$

Говорят, что последовательность случайных величин (1) сходится по вероятности к величине $\xi : \xi_n \xrightarrow{p} \xi$ при $n \to \infty$, если $P\{|\xi_n - \xi| \geq \varepsilon\} \to 0$ при $n \to \infty$ для любого $\varepsilon > 0$.

Говорят, что последовательность (1) сходится к величине ξ с вероятностью единица (или «почти наверное»): $P\{\xi_n \to \xi\} = 1$ при $n \to \infty$, если $\lim\limits_{n\to\infty} P\{|\xi_m - \xi| > \varepsilon$ хотя бы при одном $m > n$ и любом $\varepsilon > 0\} = 0$.

Последовательность (1) подчиняется усиленному закону больших чисел, если

$$P\left\{\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k \to 0\right\} = 1.$$

Теорема Колмогорова. Если случайные величины (1) независимы и удовлетворяют условию $\sum\limits_{n=1}^{\infty}\dfrac{D\xi_n}{n^2} < \infty$, то эта последовательность подчиняется усиленному закону больших чисел.

Задачи:

1. Среднее значение скорости ветра у земли в данном пункте равно 16 км/ч. Оценить с помощью первой формы неравенства Чебышева вероятность того, что в этом пункте скорость ветра (при одном наблюдении) не превысит 80 км/ч.

2. Средний расход воды в населенном пункте составляет 50 000 л в день. Оценить вероятность того, что в этом населенном пункте в данный день расход воды не превысит 150 000 л.

3. Математическое ожидание количества выпадающих в течение года в данной местности осадков составляет 55 см. Оценить вероятность того, что в этой местности осадков выпадает 175 см.

4. Число солнечных дней в году для данной местности является случайной величиной с математическим ожиданием, равным 75 дням. Оценить вероятность того, что в течение года в этой местности будет не более 200 солнечных дней.

5. Математическое ожидание скорости ветра на данной высоте равно 25 км/ч. Какие скорости ветра можно ожидать на этой высоте с вероятностью, не меньшей 0,9?

6. Среднее квадратическое отклонение ошибки измерения азимута равно 30′ (а математическое ожидание равно нулю). Оценить вероятность того, что ошибка среднего арифметического трех независимых измерений не превзойдет 1°.

7. Измеряется скорость ветра в данном пункте Земли. Случайная величина X – проекция вектора скорости ветра на фиксированное направление. Оценить вероятность события $A = \left\{ X \geq 80\text{км/ч} \right\}$, если путем многолетних измерений установлено, что $m_X = 16\text{км/ч}$.

5 Математическая статистика

В математической статистике исследуются выводы, которые могут быть сделаны на основе эмпирических данных.

Выборочной совокупностью или просто *выборкой* называют совокупность случайно отобранный объектов.

Генеральной совокупностью называют совокупность объектов, из которых производится выборка.

Объемом совокупности (выборочной или генеральной) называют число этой совокупности объектов.

Пусть имеется случайная величина ξ с функцией распределения $F(x)$ и некоторый эксперимент E, осуществляя который мы наблюдаем значение *x*, принятое случайной величиной ξ. Осуществив *n* независимых повторений эксперимента E, мы получим последовательность $(x_1, x_2, ..., x_n)$, называемую *выборкой объема n* из распределения $F(x)$ или из *генеральной совокупности* с функцией распределения $F(x)$.

Выборка, расположенная в порядке возрастания, называется вариационным рядом.

Если выборка объемом n содержит k различных элементов, причем элемент x_i встречается n_i раз, то число n_i называется частотой элемента x_i, а отношение $\omega_i = \frac{n_i}{n}$ называется относительной частотой элемента x_i. Очевидно, что $\sum_{i=1}^{k} n_i = n$.

Приписав каждому члену вариационного ряда вероятность $\frac{n_i}{n}$, получим дискретное распределение вероятностей, называемое *статистическим распределением выборки*. Функция распределения выборки $F_n^*(x) = \frac{\nu}{n}$, где ν - число членов выборки, меньших *x*, т.е. $F_n^*(x)$ есть частота события $\xi < x$ в серии из *n* независимых повторений эксперимента E. Из теоремы Бернулли (закон больших чисел) следует: $F_n^*(x) \xrightarrow{P} F(x)$, $-\infty < x < \infty$, $n \to \infty$, т.е. функция распределения выборки сходится по вероятности к функции распределения генеральной совокупности, когда объем выборки *n* неограниченно возрастает.

Пусть величина ξ имеет плотность $f(x)$. Разобьем всю ось *ox* на интервалы длиной *h* и на каждом из них построим прямоугольник с высотой n_i, где *n* – объем выборки, а n_i – число выборочных значений x_i, попавших в данный интервал. Тогда ступенчатая ломанная, ограничивающая сверху построенную фигуру, называется *гистограммой* выборки и является статистической аппроксимацией плотности $f\left(x\right)$ генеральной совокупности. На рисунке 1 изображена гистограмма относительных частот распределения выборки объема *n=50*, приведенной в таблице 1.

Таблица 1

i	$x_i < x \le x_{i+1}$	n_i	$\frac{n_i}{nh}$
1	8-10	5	00,5
2	10-12	16	0,16
3	12-14	11	0,11
4	14-16	8	0,08
5	16-18	10	0,1

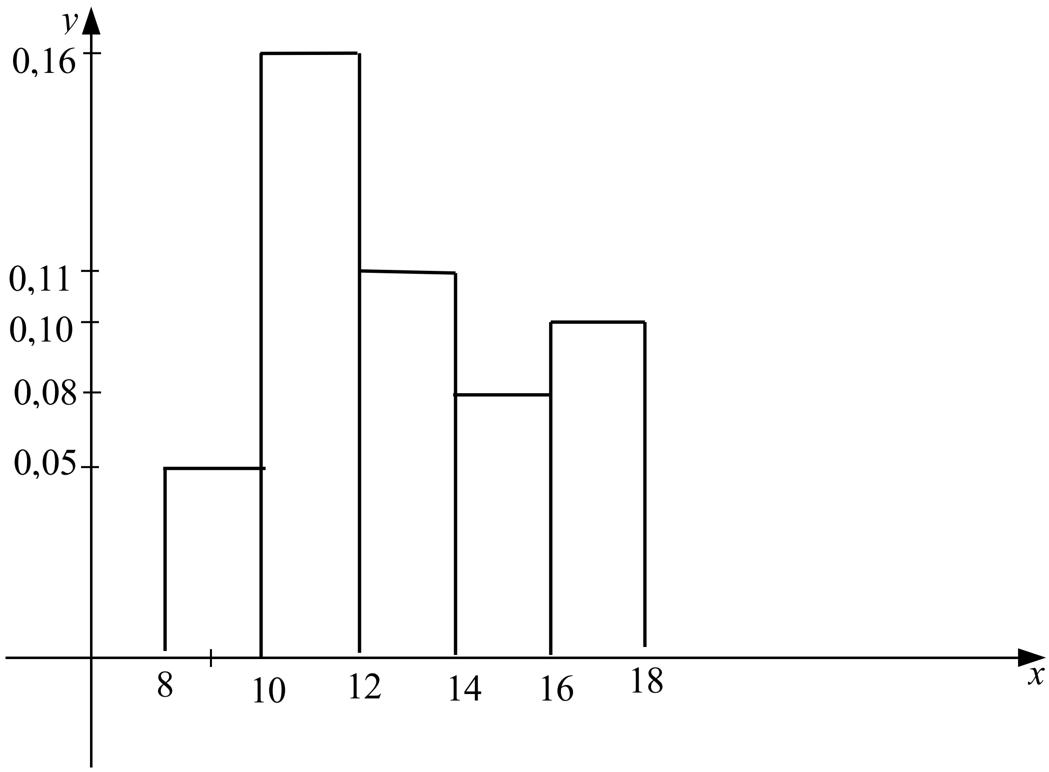

Рисунок 1

Пусть

$$M\xi = m,\ D\xi = M\left(\xi - m\right)^2 = \sigma^2,\ M\xi^\nu = \alpha_\nu.$$

– моменты генеральной совокупности. Соответствующие моменты распределения выборки – **выборочное среднее** $\overline{x}$**, выборочная дисперсия** s^2**, выборочные начальные моменты** a_ν определяются равенствами

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i;\ s^2 = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \overline{x}\right)^2 \cdot n_i;\ a_\nu = \frac{1}{n}\sum_{i=1}^{k} x_i^\nu \cdot n_i,$$

где x_i - варианта выборки, n_i - частота варианта x_i, $n = \sum_{i=1}^{k} n_i$ - объем выборки

Методом моментов называется прием получения оценок параметров распределения генеральной совокупности, который в случае известного функционального выражения плотности генеральной совокупности, содержащей несколько неизвестных параметров, сводится к следующему. За оценки $\alpha_1^*,...,\alpha_k^*$ параметров $\alpha_1,...,\alpha_k$ плотности $f(x) = f(x,\alpha_1,...,\alpha_k)$ берутся решения системы уравнений, получающиеся приравниванием попарно выборочных значений и выражений, полученных с помощью плотности $f(x,\alpha_1,...,\alpha_k)$ каких-нибудь k моментов (например, среднего, дисперсии и т.д.).

Пусть относительно распределения генеральной совокупности имеется какая-либо гипотеза. Проверка того, согласуется ли эта гипотеза с опытными данными, содержащимися в выборке, осуществляется с помощью *критериев значимости* и *критериев согласия*.

Нулевой (основной) называют выдвинутую гипотезу H_0.

Конкурирующей (альтернативной) называют гипотезу H_1, которая противоречит нулевой.

Для проверки нулевой гипотезы используют специально подобранную величину, точное или приближенное распределение известно. Эту величину обозначают через *U* или *Z*, если она распределена нормально, *F* или v^2 - по закону Фишера-Снедекора,

T - по закону Стьюдента, χ^2 - по закону «хи квадрат» и т.д. Обозначим эту величину в целях общности через K.

Статистическим критерием (или просто *критерием*) называют случайную величину K, которая служит для проверки нулевой гипотезы.

Наблюдаемым значением $K_{\text{набл}}$ называют значение критерия, вычисленное по выборкам.

Критической областью называют совокупность значений критерия, при которых нулевую гипотезу отвергают.

Областью принятия гипотезы (областью допустимых значений) называют совокупность значений критерия, при которых гипотезу принимают.

Основной принцип проверки статистических гипотез: если наблюдаемое значение критерия принадлежит критической области, то гипотезу отвергают; если наблюдаемое значение критерия принадлежит области принятия гипотезы, то гипотезу принимают.

Критическими точками (границами) $k_{\text{кр}}$ называют точки, отделяющие критическую область от области принятия гипотезы. Для каждого критерия имеются таблицы, по которым находят критические точки. Различают одностороннюю (правостороннюю или левостороннюю) и двустороннюю критические области.

Пусть требуется статистически проверить гипотезу H_0 о том, что данная выборка $x_1, x_2, ..., x_n$ извлечена из генеральной совокупности ξ с функцией распределения $F(x)$ (которая точно известна).

Критерий χ^2, применяющийся для этой цели, состоит в следующем. Все множество возможных значений ξ разбивается на r непересекающихся частей $S_1, S_2, ..., S_r$. Пусть $p_1, ..., p_r$ - вероятности попадания ξ в эти части соответственно $(p_i > 0)$ (эти вероятности вычисляются через $F(x)$), $\sum_{i=1}^{r} p_i = 1$. Пусть $\nu_1, ..., \nu_r$ – число выборочных значений, попавших фактически в $S_1, ..., S_r$ соответственно, $\sum_{i=1}^{r} \nu_i = n$. За меру уклонения распределения выборки $F_n^*(x)$ от гипотетического $F(x)$ принимается величина

$$\chi^2 = \sum_{i=1}^{r} \frac{\left(\nu_i - np_i\right)^2}{np_i} = \sum_{i=1}^{r} \frac{\nu_i^2}{np_i} - n .$$

Если ν_i рассматривать как случайные величины, то χ^2 будет случайной величиной, которая при $n \to \infty$ асимптотически распределена по закону χ^2 с $r-1$ степенями свободы. Зададимся теперь уровнем значимости ε, т.е. вероятностью, столь малой, что событие с такой вероятностью будем считать практически невозможным при одном данном испытании. Обычно ε выражают в процентах:

$$\varepsilon = p\% = \frac{p}{100}.$$

Найдем по таблицам распределения χ^2 с $r-1$ степенями свободы такое значение χ_ε^2, что $P\left\{\chi^2 > \chi_\varepsilon^2\right\} = p\%$. Далее вычислим по имеющейся выборке значение

$$\chi^2 = \sum_{i=1}^{r} \frac{\nu_i^2}{np_i} - n .$$

Если при этом окажется $\chi^2 > \chi_\varepsilon^2$, то такое уклонение значимо, и мы с *p-процентным уровнем значимости* отвергаем гипотезу H_0 как не согласующуюся с опытными данными. Если же вычисления дадут $\chi^2 \le \chi_\varepsilon^2$, то это значит, что данная выборка согласуется с гипотезой H_0. На практике, разбивая область значений ξ на части $S_1,\ldots,S_r$, желательно, чтобы выполнялось условие $np_i \ge 10 \left(i = 1, 2, \ldots, r\right)$.

Для проверки той же гипотезы H_0, в особенности, когда распределение ξ не полностью неизвестно или даже полностью неизвестно, применяется критерий А.Н.

Колмогорова. Пусть $D_n = \sup\limits_{-\infty<x<\infty} \left| F_n^*(x) - F(x) \right|$. Теорема Колмогорова утверждает, что при $n \to \infty$

$$P\left\{ \sqrt{n} D_n < x \right\} \to K(x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ \sum\limits_{\lambda=-\infty}^{\infty} (-1)^{\lambda} e^{-2\lambda^2 x^2} & \text{при} \quad x > 0. \end{cases}$$

Задаваясь уровнем значимости $\varepsilon = p\%$, мы по таблицам функции $K(x)$ находим значение x_ε так, чтобы

$$P\left\{ \sqrt{n} D_n \ge x_\varepsilon \right\} \le \varepsilon.$$

Вычислив теперь по имеющейся выборке $\sqrt{n} D_n$, сравниваем его с x_ε и в случае $\sqrt{n} D_n \ge x_\xi$ отвергаем гипотезу H_0. В противном случае гипотеза H_0 принимается.

Статистическое исследование наличия или отсутствия зависимости между случайными величинами и выделение линейной части этой зависимости производится с помощью выборочных коэффициентов корреляции и регрессии, выборочного уравнения линейной регрессии и оценки среднего квадратического уклонения нелинейной составляющей изучаемой зависимости.

Пусть в результате осуществления эксперимента E наблюдаются две величины ξ и η. Тогда *n* независимых повторений эксперимента E дадут нам *n* пар наблюденных значений $(x_1, y_1), \ldots, (x_n, y_n)$. При этом y_i являются выборочными значениями случайной величины η, а x_i может быть как случайной, так и неслучайной переменной.

Выборочный корреляционный момент K_{xy} величин ξ и η определяется формулой

$$K_{xy} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}).$$

Выборочный коэффициент корреляции $r_{xy} = \dfrac{K_{xy}}{s_x s_y}$, где s_x и s_y – статистические оценки стандартных уклонений величин ξ и η.

Выборочный коэффициент регрессии η по ξ: $l_0 = r_{xy}\dfrac{s_y}{s_x} = \dfrac{K_{xy}}{s_x^2}$.

Выборочное уравнение регрессии: $y - \overline{y} = l_0\left(x - \overline{x}\right)$.

Уклонение выборочного коэффициента корреляции от истинного $\rho_{\xi\eta}$ можно, в случае выборки из нормальной генеральной совокупности $\left(\xi, \eta\right)$, оценить подсчетом среднего квадратического уклонения r_{xy} по формуле

$$s_r \cong \frac{1 - r_{xy}^2}{\sqrt{n}}.$$

Положим $\zeta = \eta - \lambda. - \xi$, где $\lambda.$ – истинный коэффициент регрессии η по ξ. Статистической оценкой ζ будет $z = y - l_0 x$. О степени уклонения зависимости, связывающей η и ξ, от линейной судят по среднему квадратическому уклонению величины ζ, т.е. по σ_ζ. За статистическую оценку σ_ζ принимают $s_z^2 = \left(1 - r_{xy}^2\right)s_y^2$.

Если с некоторым уровнем значимости окажется, что уклонение выборочного коэффициента корреляции от нуля не значимо, то мы можем принять гипотезу о некоррелированности величин ξ и η и лишь для нормальных величин ξ и η – гипотезу об их независимости.

Задачи:

1. Уровни воды в реке по отношению к номиналу измерялись в течение 44 весенних паводков, и данные измерений приведены в следующей таблице 2:

Таблица 2

Уровень (в см)	0-24	25-49	50-74	75-99	100-124	125-149
Число случаев	0	1	3	6	7	6

Уровень (в см)	150-174	175-199	200-300	300-400	>400
Число случаев	5	4	8	4	0

Считая, что высота уровня ξ распределена по закону χ^2 с плотностью $f(x)=\dfrac{k^{a+1}x^a e^{-kx}}{\Gamma(a+1)}$ $(x\geq 0)$, с помощью метода моментов найти оценки параметров a и k этого распределения. (Использовать среднее и дисперсию).

2. При 296 пробах руды Каданского рудника были получены следующие данные о процентном содержании в руде свинца и серебра.

Таблица 3

Содержание серебра в руде (в %)	Содержание свинца в руде (в %)								
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
0-4	119	9	–	–	–	–	–	–	–
4-8	9	59	7	–	–	–	–	–	–
8-12	1	4	28	3	–	–	–	–	–
12-16	–	–	8	12	4	–	–	–	–
16-20	–	–	1	6	7	1	1	–	–
20-24	–	–	–	1	1	8	3	–	–
24-28	–	–	–	–	–	2	1	–	–
28-32	–	–	–	–	–	–	3	2	1
32-36	–	–	–	–	–	–	–	–	–
36-40	–	–	–	–	–	–	–	–	1

а) Найти выборочный коэффициент корреляции процентного содержания серебра (η) и свинца (ξ) в руде и выборочный коэффициент регрессии η по ξ; написать выборочное уравнение линии регрессии η по ξ.

б) оценить среднее квадратическое уклонение коэффициента корреляции η и ξ; охарактеризовать связь между η и ξ.

3. При снятии показаний измерительного прибора десятые доли деления шкалы прибора оцениваются на глаз наблюдателем. В нижеследующей таблице приведены количества цифр, записанных наблюдателем в качестве десятых долей при двухстах независимых измерениях.

Таблица 4

Цифры	0	1	2	3	4	5	6	7	8	9
v_i	35	16	15	17	17	19	11	16	30	24

Проверить, согласуются ли эти данные с гипотезой о том, что каждая цифра при этом могла появиться с равной вероятностью.

4. Средняя температура в г. Саратове (x) и в г. Алатыре (Чувашск. респ.) (y) измерялась в течение 13 лет и данные приведены в следующей таблице 5:

Таблица 5

Год	1891	1892	1893	1894	1895	1896	1897
x	-19,2	-14,8	-19,6	-11,1	-9,4	-16,9	-13,7
y	-21,8	-15,4	-20,8	-11,3	-11,6	-19,2	-13,0

Год	1899	1911	1912	1913	1914	1915
x	-4,9	-13,9	-9,4	-8,3	-7,9	-5,3
y	-7,4	-15,1	-14,4	-11,1	-10,5	-7,2

Найти выборочный коэффициент корреляции средних январских температур в Саратове и Алатыре, написать выборочное уравнение линейной регрессии y по x и оценить характер связи y с x.

5. Средняя температура июня в г. Москве (x) и Ярославле (y) измерялась в течение 40 лет и данные приведены в следующей таблице 6:

Таблица 6

x	y	x	y	x	y	x	y	x	y
12,0	10,8	13,9	10,1	15,0	13,8	17,2	13,9	18,1	16,0
12,0	11,3	14,2	10,0	15,0	16,0	16,9	14,8	18,4	17,8
12,0	12,0	14,0	10,0	15,5	13,9	16,9	15,0	19,2	15,0
12,0	13,0	14,0	12,0	15,9	14,7	17,0	16,0	19,3	16,1
12,8	10,9	13,9	12,4	16,0	13,0	16,8	17,0	20,0	17,0
13,8	10,0	15,0	11,0	15,9	15,0	17,5	16,0	20,1	17,7
13,1	11,5	14,9	13,0	16,0	16,0	18,0	14,0	14,0	14,8
13,0	13,0	14,9	14,2	16,9	12,9	18,0	14,8	14,0	15,2

Найти выборочные средние июньские температуры в Москве и Ярославле и их средние квадратические уклонения. С помощью критерия Колмогорова проверить согласие опытных данных с гипотезой о нормальности распределения средних июньских температур. Найти выборочный коэффициент корреляции x и y, написать выборочное уравнение линейной регрессии y по x. Охарактеризовать зависимость y и x.

6. Через равные промежутки времени в тонком слое раствора золота регистрировалось число частиц золота, попадавших в поле зрения микроскопа. В результате наблюдений было получено следующее эмпирическое распределение:

ξ_i	0	1	2	3	4	5	6	7
v_i	112	168	130	68	32	5	1	1

В первой строке приведено число ξ_i частиц золота, а во второй строке – частота v_i т.е. число интервалов времени, в течение которых в поле зрения попало ровно ξ_i частиц; объем выборки $n = \sum v_i = 517$.

Проверить, используя критерий χ^2, согласие с законом распределения Пуассона, приняв за уровень значимости $\alpha = 0{,}05$.

Список использованных источников

1 Агапов, Г.И. Задачник по теории вероятностей: учебное пособие для студентов втузов / Г.И. Агапов. – М.: Высш. шк., 1986. – 80 с.

2 Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для студентов вузов / В.Е. Гмурман. – М.: Высш. шк., 2002. – 405 с.

3 Гмурман, В.Е. Теория вероятностей и математическая статистика: учебное пособие для вузов/ В.Е. Гмурман. – М.: Высш. шк., 2002. – 476 с.

4 Емельянов, Г.В. Задачник по теории вероятностей и математической статистике: учебное пособие / Г.В. Емельянов, В.П. Скитович. – СПб.: Издательство «Лань», 2007. – 336с.

5 Писменный, Д.Т. Конспект лекций по теории вероятностей и математической статистике: учебное пособие/ Д.Т. Письменный. – М.: Айрис-пресс, 2005. – 256 с.

6 Руцкова, И.Г. Основные понятия теории вероятностей: терминологический словарь/ И.Г. Руцкова. – Оренбург: ОГУ, 2001. – 51 с.

7 Сборник задач по математике для втузов: в 4 ч. Ч. 3. Теория вероятностей и математическая статистика: учебное пособие для втузов/ под ред. А.В. Ефимова. – М.: Издательство «Наука», 1990. – 428 с.

8 Шарапов, И.П. Применение математической статистики в геологии: учебное пособие для студентов вузов / И.П. Шарапов. – М.: Издательство «Недра», 1971. – 345 с.

Приложение А

(*справочное*)

Ответы к задачам

А. 1 Понятие вероятности случайного события, формулы сложения и умножения вероятностей, формула полной вероятности, формула Байеса

1. Событие *B* наблюдаемо, *A* и *C* не наблюдаемы. (Множество элементарных исходов данного эксперимента можно записать в виде $\Omega=\left\{\left(\upsilon,\varphi\right)\middle|\upsilon\geq 0,\left(2n\right)^{\circ}\leq\varphi<\left(2n+2\right)^{\circ},n=0,1,...,179\right\}$).

2. а) $\approx 0{,}08$; б) $\frac{\sqrt{3}}{4}\approx 0{,}433$; в) $\frac{\pi}{12}$.

3. 0,4.

4. а) $p=\frac{1}{4}$**;** б) $p=\frac{\sqrt{2}-1}{2\sqrt{2}}$.

5. Если $\alpha>\arccos\frac{2r+\alpha}{l}$, то $p=0$, если $٠\leq\alpha\leq\arccos\frac{٢r+\alpha}{l}$, то

$$p=\frac{١}{l^{٢}\cos\alpha}\cdot\left(l\cos\alpha-٢r-\alpha\right)\left(l-٢r-\alpha\right).$$

6. $\frac{١}{٧}$.

7. 0,660; 0,264; 0,076.

8. а) Пусть *V* – любой объем внутри сферы, тогда $P\left(V\right)$ – вероятность одной точке попасть в объем *V*, поэтому вероятность одной точке попасть в сферу радиуса *a*, концентрическую основной, равна $\left(\frac{a}{R}\right)^{٣}$, т. е. вероятность не попасть ни одной из точек в эту сферу $P_N=\left[١-\left(\frac{a}{R}\right)^{٣}\right]^N$; б) $p=\lim\limits_{N\to\infty}P_N=\lim\limits_{N\to\infty}\left(١-\frac{a^{٣}}{N^{٣}}\right)^N=e^{-\frac{٤}{٣}\pi a\lambda}$

9. $p=\left(\dfrac{2}{7}\right)^{12}=0{,}0000003$**;** так как эта вероятность весьма мала, то это не согласуется с гипотезой равновероятности.

10. а) $p\cong 0{,}1621$; б) $p\cong 1$.

А. 2 Последовательность независимых испытаний (схема Бернулли), локальная и интегральная теоремы Муавра – Лапласа, теорема Пуассона

1. 0,1887.

2. 8.

3. 0,2787.

4. 100.

5. $P\left(np-a\sqrt{npq}<m<np+a\sqrt{npq}\right)\cong 2\Phi(a)\geq 0{,}9995$**,** $a\geq 3{,}6$**,** $n=2206$**,** $0<m<6$**.**

А. 3 Случайные величины и функции распределения вероятностей. Числовые характеристики распределений

1. Среднее число импульсов помехи, попадающих в интервал времени длиной 10 *мксек*, $\lambda=0{,}1$. Пусть ξ – число импульсов помехи попавших в интервал работы радиостанции, тогда $P\{\xi=m\}=\dfrac{(٠,١)^m}{m!}e^{-١}$, $m=٠,١,٢,\ldots$ Вероятность срыва передачи $P(\xi>٠)=١-P(\xi=٠)=١-e^{-٠,١}=٠,٠٩٥١٦$.

2. Разобьем контур длины L на части $l_١+l_٢+\ldots+l_n$ с длинами, не превышающими расстояние $٢a$ между параллелями; тогда $MV=MV_1+MV_2+\ldots+MV_n$, где V_i – число пересечений i-го участка контура с параллелями, $MV_i=١\cdot\dfrac{l_i}{\pi a}+٠\left(١-\dfrac{l_i}{\pi a}\right)=\dfrac{l_i}{\pi a}$; $MV=\sum\limits_{i=١}^{n}\dfrac{l_i}{\pi a}=\dfrac{L}{\pi a}$.

3. X_j можно рассматривать как случайную величину, принимающую значения $x_1, x_2, \ldots, x_k$ с вероятностями соответственно $\frac{n_1}{n}, \frac{n_2}{n}, \ldots, \frac{n_k}{n}$. Тогда $MX_j = m$ – среднее число жителей на квартал: $M\left(X_1 + X_2 + \ldots + X_r\right) = mr$;

$$DX_j = MX_j^2 - m^2 = \sum_{i=1}^{k} x_i^2 \frac{n_i}{n} - m^2 = \sum_{i=1}^{k} x_i^2 \frac{n_i}{n} - \left(\sum_{i=1}^{n} x_i \frac{n_i}{n}\right)^2 = \sum_{i=1}^{k} \frac{x_i^2 n_i \left(n - n_i\right)}{n^2} - 2\sum_{\substack{i,l=1 \\ i>l}}^{k} \frac{x_i x_l n_i n_l}{n^2};$$

так как все X_j одинаково распределены, то

$$D\left(X_1 + X_2 + \ldots + X_r\right) = rDX_1 + r\left(r-1\right)M\left\{\left(X_1 - m\right)\left(X_2 - m\right)\right\} = r\left\{MX_1^2 + \left(r-1\right)M\left(X_1 X_2\right) - \right.$$

$$\left. -rm^2\right\} = \left(\sum_{i=1}^{r} \frac{x_i^2 n_i}{n} - m\right)\frac{r\left(n-r\right)}{n-1} = \frac{DX_i r\left(n-r\right)}{n-1}$$, где i – любое.

4. $$F\left(r\right) = \begin{cases} ٠ \text{ при } & r \leq ٠, \\ \left(\frac{r}{R}\right)^{٢} & ,, \quad ٠ < r \leq R, \\ ١ & ,, r > R \end{cases}$$

$M\xi = \frac{٢}{٣}R$; $D\xi = \frac{R^2}{18}$.

5. $f_\xi\left(x\right) = \frac{١}{٤\sqrt{٢\pi}} e^{-\frac{x^{٢}}{٣٢}}$; $P\left\{\left|\xi\right| < ٦\right\} = ٠,٨٦٦٣٨$; $P\left\{\left|\xi\right| > ٦\right\} = ٠,١٣٣٦٢$. а) $p_1 = 0,9972$;

б) $P\left\{٠,٥ \leq \xi < ٣,٥\right\} \cong ٠,٢٥٩٥$; $p_{٢} = ٠,٢٢٢٧$.

6. $T_1 = \frac{5}{9}\left(T - 32\right)$. Обозначим через $\psi\left(t_{١}\right)$ плотность распределения ошибок, приспособленную к шкале Цельсия; тогда $\psi\left(t_{١}\right) = \frac{٩}{٥}\sqrt{\frac{١,٢}{\pi}} e^{-٣,٨\left(t_{١}+٥\right)^{٢}}$, т. к.

$$\psi\left(t_{١}\right) = \frac{٩}{٥}\varphi\left(\frac{٢}{٥}t_{١} + ٣٢\right) = \frac{٩}{٥}\sqrt{\frac{١,٢}{\pi}} e^{-١,٢\left(\frac{٢}{٥}t_{١}+٩\right)^{٢}}.$$

7. $\sigma_A = 1,2$; $\sigma_B = 1,1$; $K_{AB} = r\sigma_A\sigma_B = 0,44$; $\sigma_{A+B} = \sqrt{\sigma_A^2 + \sigma_B^2 + 2K_{AB}} \cong 1,88$.

А. 4 Закон больших чисел

1. По первой форме неравенства Чебышева находим $P\{\xi > 80\} \le \frac{M\xi}{80} = \frac{1}{5}$, т. е. $P\{\xi \le 80\} \ge \frac{4}{5}$.

2. $\frac{2}{3}$.

3. 0,3.

4. $\frac{5}{8} \cong 0,625$.

5. Пусть ξ – скорость ветра $P\{|\xi - 25| \le \varepsilon\} \ge 1 - \frac{4,5^2}{\varepsilon^2} > 0,9$, т. е. $\varepsilon > 14,2$ *км/час*, следовательно, с вероятностью, большей 0,9, имеем $١٠,٨ \le \xi \le ٣٩,٢$ *км/час*.

6. Пусть $\xi_١, \xi_٢, \xi_٣$ – ошибки трех наблюдений; тогда ошибка среднего арифметического этих наблюдений $\eta = \frac{\xi_1 + \xi_2 + \xi_3}{3}$; $D\eta = \frac{D\xi_1 + D\xi_2 + D\xi_3}{9} = 300$;

$P\{|\eta| \le ٦٠\} \ge ١ - \frac{D\eta}{٦٠^٢} \cong ٠,٩١٧$.

7. $P(A) \le ٠,٢$.

А. 5 Математическая статистика

1. $M\xi = \frac{\alpha + ١}{k}$; $D\xi = \frac{a+1}{k^2}$; выборочные значения: $\bar{x} = ١٦٥$, $s^٢ = ٦٩٤٨$; приравнивая попарно $M\xi = \bar{x}$ и $D\xi = s^2$, находим $a \cong ٣$, $k = ٠,٠٢٤$.

2. Пусть выборочное распределение ξ и η будет соответственно x и y. Тогда $\bar{x} = ٩,٨٨$; $\bar{y} = ٧,٦٣$; $s_x = ٨,٥٢٦$; $s_y = ٧,٧٤٢$; $r_{xy} = ٠,٨٦٥$; $l_{y/x} = 0,78$; уравнение регрессии $y-7,63=0,78(x-9,38)$. Обозначим $\eta=\lambda\xi+\zeta$, где λ– истинный коэффициент регрессии η по ξ; его статистической оценкой является $l=0,78$. Выборочным распределением ζ будет $z=y-lx$. Тогда $s_z=\sqrt{1-r^2}\cdot\sigma_y=3,88$, следовательно, с точностью до случайной величины, имеющей среднее квадратическое уклонение 3,88, связь между η и ξ можно считать линейной и ее статистической оценкой является выборочное уравнение регрессии. Так как число наблюдений $n=296$ достаточно велико, то приближенно $\sigma_r=\dfrac{1-r^2}{\sqrt{n}}=0,015$.

3. $p=0,1$; $n=200$; $np=20$; $\chi^٢=٢٤,٩٠$. По таблицам распределения $\chi^٢$ с девятью степенями свободы находим для 5%-ного уровня $\chi^٢_{٠,٠٥}=١٦,٩$. Для 1%-ного уровня $\chi^2_{0,01}=21,7$, т. е. уклонение $\chi^٢$ превосходит значения 5- и 1%-ного уровней, следовательно, гипотезу о равномерном распределении определяемых на глаз делений шкалы прибора следует отбросить. Это указывает на то, что наблюдатель отдает предпочтение (подсознательно) некоторым определенным цифрам, в нашем случае – цифрам 0 и 8.

4. $\bar{x}\cong-١٢$; $\bar{y}=-١٣,٨$; $s_x=4,59$; $s_y=٤,٤٦$; $r=٠,٩$; $y+13,8=0,87(x+12)$; $z=y-lx$; $s_z=١,٩٥$. Так как коэффициент корреляции близок к единице, а среднее квадратическое уклонение нелинейной составляющей зависимости, связывающей y и x, т. е. s_z, мало, то эту зависимость можно считать линейной.

5. $\bar{x}=15,6$; $\bar{y}=١٣,٨$; $s_x=2,20$; $s_y=١,٨٧$. Согласие с гипотезой нормальности эмпирических данных очень хорошее; так, для температуры в Москве имеем $\sqrt{n}D_n=0,74$. По таблицам $K(\lambda)$ находим, что в 65% случаев следует ожидать не меньшего отклонения. $r_{xy}=0,997$; $y-13,8=0,213(x-15,6)$. Полагая $z=y-lx=y-٠,٢١٣x$, находим $s_z=٠,٠٢١$. Зависимость y от x явно линейная.

6. $\bar{x}=١,٥٤$; $\lambda^*=١,٥٤$; теоретические вероятности: 0,2144; 0,3301; 0,2542; 0,1305; 0,0502; 0,0155; 0,0040; 0,0009. $\chi^٢_{\text{набл}}=٢,٨$. $m=r-1-1=8-1-1=6$.

$\chi^2_{\text{кр}} = 12,6$. Так как $\chi^2_{набл} < \chi^2_{\text{кр}}$, то нет основания отвергать гипотезу о распределении случайной величины ξ по закону Пуассона.

Замечание. Поскольку по выборке определен один параметр, то вместо $r-1$ следует брать $r-1-1$.