

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ (ФИЛИАЛ)**

**ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

**В Г.ТАГАНРОГЕ РОСТОВСКОЙ ОБЛАСТИ**

**ПИ (филиал) ДГТУ в г. Таганроге**

 УТВЕРЖДАЮ

Директор

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.К. Исаев

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г

Рег. № \_\_\_\_\_\_\_\_\_\_\_\_

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ

к практическим занятиям

по учебной дисциплине ЕН.03 Теория вероятностей и математическая статистика

по специальности 09.02.07 «Информационные системы и программирование»

Таганрог

2020

**Лист согласования**

Учебно-методическое пособие по учебные Теория вероятностей и математическая статистикаразработана на основе Федерального государственного образовательного стандарта (далее – ФГОС) для специальности среднего профессионального образования (далее – СПО) 09.02.07 «Информационные системы и программирование».

**Разработчик(и):**

Преподаватель \_\_\_\_\_\_\_\_\_\_\_\_\_ Т.М. Марданова

Методические указания рассмотрены и одобрены на заседании цикловой методической комиссии специальности 09.02.05Прикладная информатика (по отраслям)

Протокол № 7 от «04» февраля 2020г

Председатель цикловой методической комиссии О.В. Андриян

**Рецензенты:**

ЧОУ ВО «ТИУиЭ» начальник информационно-аналитического управления, к.т.н., доцент О.И. Овчаренко

АО «Красный гидропресс»зам. начальника отдела ИТ С.С. Пирожков

**Согласовано:**

Заведующий УМО

Т. В. Воловская

**Введение**

В учебно-методическом пособии к практикуму по курсу «Теория вероятностей и математическая статистика» изложены сведения, необходимые для успешного выполнения практических занятий по данному курсу. Описан процесс работы с инструментарием, применяемым на практических занятиях, представлен ряд типичных задач и подходы к их решению. Практические занятия посвящены углубленному знакомству обучающихся с формирование у обучающихся научного представления о случайных событиях, величинах и случайных процессах, а также о методах их исследования; способствовать усвоению методов количественной оценки характеристик случайных событий и величин; способствовать приобретению практических навыков и знаний для решения задач по теории вероятностей, случайным процессам и математической статистике.

Цель настоящего пособия – помочь обучающимся при выполнении практических работ, выполняемых для закрепления знаний по теоретическим основам и получения практических навыков

**Обучающийся должен** **уметь:**

В результате освоения учебной дисциплины обучающийся должен уметь:

-вычислять вероятность событий с использованием элементов комбинаторики;

-использовать методы математической статистики

знать:

-основы теории вероятностей и математической статистики; основные понятия теории графов.

Данное учебно-методическое пособие предназначено для обучающихся 2 курса.

**Правила выполнения практических занятий**

Практические занятия выполняются каждым обучающимся самостоятельно в полном объеме и согласно содержанию методических указаний.

Перед выполнением обучающийся должен отчитаться перед преподавателем за выполнение предыдущего занятия (сдать отчет).

Обучающийся должен на уровне понимания и воспроизведения предварительно усвоить необходимую для выполнения практических занятий теоретическую и информацию.

Обучающийся, получивший положительную оценку и сдавший отчет по предыдущему практическому занятию, допускается к выполнению следующему занятию.

Обучающийся, пропустивший практическое занятие по уважительной либо неуважительной причине, закрывает задолженность в процессе выполнения последующих практических занятий.

**Тема 1.1**

**Основы комбинаторики.**

**Решение задач на расчет числа перестановок, размещений, сочетаний без повторений и с повторениями.**

***Задача 1.*** В корзине лежат 5 кубиков разного цвета. Сколько цветовых комбинаций можно из них составить, если кубики выкладывать в одну линию?

***Задача 2.*** Сколько существует перестановок из букв слова «фонарь», в которых буква «р» на первом месте, а буква «о» - в конце слова?

***Задача 3.*** Сколько 3- буквенных «слов» можно составить из букв слова «ВОЛАН»? Словом считается любая последовательность букв.

***Задача 4***. В ящике 2 шара белого цвета, 2 шара синего цвета и 1 шар желтого цвета. Сколькими способами можно выбрать 3 шара?

**Методика вычисления вероятностей событий.**

**Решение задач на вычисление вероятностей событий**

**по классической формуле вероятности.**

**Классификация событий**

**1 вариант.** Основным понятием теории вероятностей является понятие *события*.

*Случайным событием* (возможным событием или просто событием) называется любой факт, который в результате испытания (или опыта) может произойти или не произойти.

Под *испытанием* (опытом, экспериментом) понимается выполнение определенного комплекса условий, в которых наблюдается то или иное явление или фиксируется тот или иной результат. Испытание может проводиться человеком, но может происходить независимо от человека, который при этом выступает в роли наблюдателя. Событие не является происшествием, а является возможным исходом, результатом опыта.

 Если при каждом испытании, при котором происходит событие А, происходит и событие В, то говорят , *что А влечет за собой В*(входит в В), В включает А, и обозначается .Например, А – изделие 1сорта,В – изделие 2 сорта, С – стандартное изделие: . Если одновременно  то события А и В называются *равносильными*. События называются *несовместными*, если наступление одного из них исключает наступление другого. В противном случае события называются *совместными*.

Событие называется *достоверным*, если оно происходит в результате каждого испытания.

Событие называется *невозможным*, если оно не происходит ни в одном испытании.

События называются *равновозможными*, если в результате испытания ни одно из них не является объективно более возможным, чем другое. Равновозможные события не могут появляться иначе, чем в испытаниях, обладающих *симметрией исходов*, и наше незнание того, какое событие объективно более возможно при отсутствии симметрии исходов, не может служить основанием, чтобы считать события равновозможными.

Несколько событий называются *единственно возможными*, если в результате испытания обязательно произойдет одно из них Например, в семье , имеющей двух детей, единственно возможные события имеют вид: А – «два мальчика», В – «две девочки», С – «мальчик и девочка».

Несколько событий образуют *полную группу событий*, если они являются единственно возможными и образуют полную систему событий. Это означает , что в результате опыта *должно произойти одно и только одно из этих событий.*

Частным случаем событий, образующих полную группу, являются *противоположные* события. Два несовместных события, образующих полную группу, из которых одно обязательно должно произойти, называются *противоположными*.

**СЛУЧАЙНОЕ СОБЫТИЕ, ЧАСТОТА И ВЕРОЯТНОСТЬ.**

**2 вариант.** Основными понятиями в теории вероятностей являются понятия *события* и *вероятности* *события*.

*Под событием понимают результат эксперимента или опыта, который при реализации данного комплекса условий может произойти или не произойти, повторяясь многократно в одних и тех же условиях.*

События обозначаются буквами А,В,С,.. Если событие неизбежно произойдет при реализации опыта, то оно называется *достоверным*; если же оно не может произойти, то оно называется *невозможным*. Событие, которое может произойти, а может не произойти, называется *случайным*.

Событие, которое состоит в наступлении *хотя бы одного из событий А и В*, называется *суммой* (объединением) событий А+В.

Событие, состоящее в наступлении *обоих событий А и В*, называется *произведением* (совмещением) событий А и В, АВ.

События называются *несовместными*, если появление одного из них исключает появление другого события в одном и том же испытании.

Для каждого связанного с опытом события А можно выделить совокупность тех элементарных исходов, наступление которых влечет за собор появление события А, совокупность всех элементарных событий образует *пространство элементарных исходов.*

Элементарные события взаимно исключают друг друга и в результате данного опыта обязательно произойдет одно из них. Пространство элементарных исходов образует *полную группу попарно несовместных событий*, т.к. появление хотя бы одного из событий полной группы есть достоверное событие.

Два несовместных события, образующих полную группу, называются *противоположными*

Для противоположных событий выполняются одновременно два условия:

*.*

*Для количественной оценки возможности появления случайного события А вводится понятие вероятности.*

Вероятностью события А называется отношение числа исходов, благоприятствующих появлению события А, к числу всех равновозможных несовместных элементарных исходов, образующих полную группу: (*классическое* *определение вероятности*).

К числу основных понятий теории вероятности относится *частота события*, под которой понимают отношение числа испытаний, в которых произошло событие, к общему числу испытаний. Частота события называется *статистической вероятностью*. Для нахождения частоты события опыты должны быть реально произведены.

*Массовые события обладают свойством устойчивой частоты: при больших количествах испытаний частота колеблется в тесных пределах и стремится к постоянному числу. При этом частота события считается ее вероятностью.*

В случаях, когда не возможно вычисление вероятности по формуле классического определения, используют *геометрическое определение вероятности, т.е. вероятность попадания точки в некоторую область, вне зависимости от ее формы.*

***Задача 1.*** В ящике лежит 10 шаров. Из них 3 белых шара, 5 желтых шаров и 2 красных шара. Какова вероятность вынуть из урны красный шар?

***Задача 2.*** В коробке лежит 10 конфет. Из них 3 карамели, 5 конфет «Мишка на севере» и 2 конфеты «Трюфель». Какова вероятность наугад вынуть из коробки шоколадную конфету?

***Задача 3.*** В коробке лежит 10 конфет. Из них 3 карамели, 5 конфет «Мишка на севере» и 2 конфеты «Трюфель». Какова вероятность наугад вынуть из коробки две шоколадные конфеты?

***Задача 4.*** В партии из N деталей имеется n стандартных. Наудачу отобраны m деталей. Найти вероятность того, что среди отобранных деталей ровно k стандартных.

***Задача 5.*** В группе 15 студентов, среди которых 6 отличников. По списку наудачу отобраны 10 студентов. Найти вероятность того, что среди отобранных студентов 4 отличника.

***Задача 6.*** Подбрасывается два игральных кубика, отмечается число очков на верхней грани каждого кубика. Найти вероятность того, что на обоих кубиках выпало число очков, большее двух.

***Задача 7***. Игральный кубик бросают два раза. Какова вероятность того, что на верхней грани два раза выпадет четное число очков, большее 2?

***Задача 8.*** Стрелок стреляет по мишени дважды. Вероятность попадания в мишень 0,7. Какова вероятность того, что стрелок хотя бы один раз попал в мишень?

**Вероятности сложных событий.**

**Решение задач на нахождение вероятностей сложных событий.**

***Задача 1***. В ящике 10 красных и 5 синих пуговиц. Вынимаются наудачу две пуговицы. Какова вероятность, что пуговицы будут одноцветными*?*

***Задача 2***. Среди сотрудников фирмы 28% знают английский язык, 30% – немецкий, 42% – французский; английский и немецкий – 8%, английский и французский – 10%, немецкий и французский – 5%, все три языка – 3%. Найти вероятность того, что случайно выбранный сотрудник фирмы: а) знает английский или немецкий; б) знает английский, немецкий или французский; в) не знает ни один из перечисленных языков.

***Задача 3****.*В семье – двое детей. Какова вероятность, что старший ребенок – мальчик, если известно, что в семье есть дети обоего пола?

***Задача 4****.* Мастер, имея 10 деталей, из которых 3 – нестандартных, проверяет детали одну за другой, пока ему не попадется стандартная. Какова вероятность, что он проверит ровно две детали?

**Формула полной вероятности. Формула Байеса.**

Следствием двух основных теорем теории вероятностей - теоремы сложения и теоремы умножения - являются формула полной вероятности и формула Байеса.

**Теорема.** *Если событие F может произойти только при появлении одного из событий (гипотез) А1,А2,…,Аn, образующих полную группу, то вероятность события F равна сумме произведений вероятностей каждого из этих событий ( гипотез) на соответствующие условные вероятности события F:*

 .

Следствием теоремы умножения и формулы полной вероятности является **формула** **Байеса.**

Она применяется тогда, когда событие F, которое может появиться только с одной из гипотез А1,А2,…Аn, образующих полную группу событий, произошло и необходимо произвести *количественную переоценку априорных вероятностей* этих гипотез P(A1),P(A2),...,P(An), известных до испытания, т.е. надо найти апостериорные (получаемые после проведения испытания) условные вероятности гипотез PF(A1), PF(A2),...,PF(An).

Формула Байеса имеет вид:

 PF(Ai)=

Значение формулы Байеса в том, что при наступлении события F, т.е. по мере получения новой информации, можно проверять и корректировать выдвинутые до испытания гипотезы. Такой подход, называемый *байесовским*, дает возможность корректировать управленческие решения в экономике, оценки неизвестных параметров распределения изучаемых признаков в статистическом анализе и др.

**Пример 1.**

В торговую фирму поступили телевизоры от трех поставщиков в отношении 1:4:5. Практика показала, что телевизоры, поступающие от 1-го, 2-го и 3-го поставщиков, не требуют ремонта в течение гарантийного срока соответственно в 98%, 88% и 92% случаев.

1) Найти вероятность того, что поступивший в торговую фирму телевизор не потребует ремонта в течение гарантийного срока. 2) Проданный телевизор потребовал ремонта в течение гарантийного срока. От какого поставщика вероятнее всего поступил этот телевизор?

Решение.

Обозначим события Аi-телевизор поступил от i-го поставщика (i=1,2,3).

F – телевизор не потребует ремонта в течение гарантийного срока.

По условию

 

Результат получен по формуле полной вероятности.

2) Обозначим событие  - телевизор потребует ремонта в течение гарантийного срока.



По условию 

По формуле Байеса

 

 Таким образом, вероятность гипотезы А2 увеличилась с 0,4 до0,б533, а гипотезы А3 уменьшилась с 0,5 до 0,444; если до наступления события F наиболее вероятной была гипотеза А3, то после наступления события F , наиболее вероятна гипотеза А2 – поступление телевизора от 2-го поставщика.

**Пример 2.**

Известно, что 95% продукции, выпускаемой заводом, является стандартным. Упрощенная схема контроля признает продукцию пригодной с вероятностью 0,98, если она стандартная, и с вероятностью 0,06, если она нестандартная. Найти вероятность того, что:

1. взятое наудачу изделие пройдет упрощенный контроль;
2. изделие стандартное, если оно: а) прошло упрощенный контроль; б) дважды прошло упрощенный контроль.

Решение.

1. Обозначим события:

А1- наудачу взятое изделие стандартное; А2-изделие нестандартное.

F- изделие прошло упрощенный контроль.

По условию Р(А1)=0,95; Р(А2)=0,05; 

Вероятность того, что взятое наудачу изделие пройдет упрощенный контроль, по формуле полной вероятности P(F)=0,95⋅0,98+0,05⋅0,06=0,934.

1. а) Вероятность того, что изделие, прошедшее упрощенный контроль, стандартное,

найдем по формуле Байеса:

 

 б) Пусть событие F\*- изделие дважды прошло упрощенный контроль. Тогда по теореме умножения вероятностей



По формуле Байеса 

Так как  очень мала, то вероятность того, что изделие, пошедшее упрощенный контроль дважды, окажется бракованной, следует отбросить как практически невозможное.

**Пример 3.**

Два стрелка стреляют по мишени независимо друг от друга, делают по одному выстрелу. Вероятность попадания в мишень для первого стрелка равна 0,8; для второго 0,4. После стрельбы в мишени обнаружена одна пробоина. Какова вероятность того, что она принадлежит а) 1-му стрелку; б) 2-му стрелку?

Решение.

Обозначим события:

А1-оба стрелка не попали в мишень;

А2-оба стрелка попали в мишень;

А3- 1-ый стрелок попал в мишень, второй не попал;

А4- второй стрелок попал в мишень, первый не попал;

F- в мишени одно попадание.

Определим вероятности гипотез и условные вероятности события F для этих гипотез:

Р(А1)=0,2•0,6=0,12; ;

Р(А2)=0,8•0,4=0,32; ;

Р(А3)=0,8•0,6=0,48; ;

Р(А4)=0,2•0.4=0,08; .

По формуле Байеса получаем: ;

.

Получаем, что вероятность попадания в цель первого стрелка при наличии одной пробоины выше, чем для второго стрелка (  ) почти в 6 раз.

***Задача 5****.*В одном ящике 3 белых и 5 черных шаров, в другом ящике – 6 белых и 4 черных шара. Найти вероятность того, что хотя бы из одного ящика будет вынут белый шар, если из каждого ящика вынуто по одному шару.

**Схема Бернулли.**

**Решение задач на вычисление вероятностей событий**

**с помощью формулы Бернулли**

**Повторные независимые испытания.**

Практически часто встречаются задачи, в которых при многократно повторяющихся испытаниях при данном комплексе условий интерес представляет вероятность числа m наступления некоторого события А в n испытаниях.

 Если вероятность наступления события А в каждом испытании не меняется в зависимости от исходов других, то такие испытания называются *независимыми* *относительно события А*. Если независимые повторные испытания проводятся при одном и том же комплексе условий, то *вероятность наступления события А в каждом* *испытании одна и та же*. Описанная последовательность независимых испытаний *называется схемой Бернулли*.

**Формула Бернулли.**

**Теорема.**

Если вероятность р наступления события А в каждом испытании постоянна, то вероятность

Pm,n того, что событие А наступит m раз в n независимых испытаниях, равна

 .

***Задача 1*.** Игральная кость брошена 6 раз. Найти вероятность того, что ровно 3 раза выпадет «шестерка».

***Задача 2***. Монета бросается 6 раз. Найти вероятность того, что герб выпадет не более, чем 2 раза.

***Задача 3***. Аудитор обнаруживает финансовые нарушения у проверяемой фирмы с вероятностью 0,9. Найти вероятность того, что среди 4 фирм-нарушителей будет выявлено больше половины.

***Задача 4*.** Монета подбрасывается 3 раза. Найти наиболее вероятное число успехов (выпадений герба).

***Задача 5.*** В результате каждого визита страхового агента договор заключается с вероятностью 0,1. Найти наивероятнейшее число заключенных договоров после 25 визитов.

***Задача 6***. Известно, что процент брака для некоторой детали равен 0,5%. Контролер проверяет 1000 деталей. Какова вероятность обнаружить ровно три бракованные детали? Какова вероятность обнаружить не меньше трех бракованных деталей?

**Основы теории случайных величин.**

**Характеристики ДСВ и их свойства.**

**Решение задач на запись распределения ДСВ.**

**Понятие случайной величины. Закон распределения дискретной случайной величины.**

Понятие случайной величины является одним из важнейших в теории вероятностей.

Под случайной величиной понимают переменную, которая в результате испытаний в зависимости от случая принимает одно из возможного множества своих значений (какое именно – заранее не известно).

Примеры случайных величин:

1. число новорожденных в Таганроге;
2. количество бракованных изделий в партии;
3. число выстрелов до попадания;
4. дальность полета снаряда;
5. расход электроэнергии на предприятии за месяц.

**Случайная величина называется дискретной (прерывной), если множество ее значений конечно (или бесконечно, но счетно).**

Пример 1-3 – дискретные случайные величины; 1,2 – с конечным множеством значений; 3 – со счетным множеством; 4,5 – непрерывные случайные величины.

**Определение. Случайной величиной Х называется функция, заданная на множестве элементарных исходов (или в пространстве элементарных событий) X = f(), где  и является элементарным исходом.**

Случайные величины обозначают X,Y,Z...;х,y,z – их соответствующие значения.

**Определение. Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.**

 Для дискретной случайной величины закон распределения может быть задан в виде таблицы(матрицы), формулы (аналитически), графически. Простейший способ задания – матрица, в которой в порядке возрастания перечислены все возможные значения случайной величины и соответствующие им вероятности.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  х1 |  х2 |  ... |  хi |  ... |  xn |
|  p1 |  p2 |  … |  pi |  … |  pn |

Такая таблица называется **рядом распределения** случайной величины.

События х1,х2,…хn , состоящие в том, что в результате испытаний случайная величина Х примет все значения, являющиеся несовместными и единственно возможными, образуют полную группу. Следовательно, **сумма их вероятностей равна 1**. Для любой дискретной величины **** Эта 1 некоторым образом распределена между значениями случайной величины. Отсюда термин «распределение».

Ряд распределения может быть изображен графически, если по оси ОХ откладывать значения случайной величины, а по оси ОУ – их соответствующие вероятности.

 Соединение полученных точек образует ломаную, называемую **многоугольником** или **полигоном распределения** вероятностей.

 У

 p2

pi

 p1

 pn

 0 x1 x2 ... xi ... xn x

**Пример.** Вероятность того, что студент сдаст семестровый экзамен в сессию по дисциплинам А и Б, равны 0,7 и 0,9.Составить закон распределения числа семестровых экзаменов, которые сдает студент.

Решение. Возможные значения случайной величины Х – числа сданных экзаменов -0,1,2.

Пусть А1,А2,А3 – событие, состоящее в том, что студент сдаст 1,2,3 экзамен. А1 и А2 – независимы: Р(Х=0)=.





Ряд распределения случайной величины имеет вид

|  |  |  |  |
| --- | --- | --- | --- |
|  xi |  0 |  1 |  2 |
|  pi |  0,03 |  0,34 |  0,63 |

Построить по нему многоугольник (полигон) распределения вероятностей.

Две случайные величины называются **независимыми**, если закон распределения одной из них не меняется от того, какие возможные значения приняла другая величина.

**Математические операции над случайными величинами выполняются по правилам:**

**1.Произведение** кХ случайной величины Х на постоянную к называется случайная величина, которая принимает значения кхi с теми же вероятностями pi (i=1,2,...,n).

**2.Суммой (разностью или произведением)** случайных величин Х и У называется случайная величина, которая принимает все возможные значения вида , с вероятностями рij того, что случайная величина Х примет значение хi, а Y – значение yj.

**Замечание.** Операции нуждаются в уточнении в зависимости от задачи. Но в случае независимых случайных величин Х и У действия выполняются по теоремам вероятностей.

**Математическое ожидание дискретной случайной величины.**

Ряд распределения дискретной случайной величины дает полную информацию о ней, но бывает не всегда удобен для анализа. Для усреднения результатов исследования значений случайной величины используется математическое ожидание.

Происхождение термина «математическое ожидание» связано с периодом возникновения теории вероятностей, когда областью ее применения являлись азартные игры. Игроков интересовало среднее значение ожидаемого выигрыша, т.е. математическое ожидание.

**Определение.** **Математическим ожиданием**, или средним значением, М(Х) дискретной случайной величины Х называется произведение всех ее значений на соответствующие им вероятности:

 .

По условию примера вычислим математическое ожидание

М (Х)=0⋅0,03+1⋅0,34+2⋅0,63=1,6.

**Механическая интерпретация математического ожидания.** Если каждая материальная точка с абсциссой  имеет массу, равную , а вся единичная масса распределена между этими точками, то математическое ожидание представляет собой абсциссу центра масс системы материальных точек.

**Свойства математического ожидания.**

1. М (С)=С, с – постоянная.
2. М (к⋅Х)=к⋅М (Х), к – постоянный множитель.
3. М (Х±У)=М (Х)±М (У).
4. М (Х⋅У)=М (Х)⋅М(У).
5. Математическое ожидание отклонения случайной величины от ее математического ожидания равно 0: M[X-M(X)]=0.

**Дисперсия дискретной случайной величины.**

Слово «дисперсия» означает «рассеяние». Дисперсия является еще одной характеристикой случайной величины, указывающей на разброс значений относительно среднего значения.

**Определение. Дисперсией** D(Х) случайной величины Х называется математическое ожидание квадрата ее отклонения от математического ожидания: или D(X)=M(X-a)2, где а=М (Х).

Дисперсия имеет размерность квадрата случайной величины, что не всегда удобно. Поэтому в качестве показателя рассеяния также используют величину .

**Определение.** Арифметическое значение корня квадратного из дисперсии случайной величины называется **средним квадратическим отклонением** (**стандартным отклонением** или **стандартом**): .

**Свойства дисперсии случайной величины.**

1.Дисперися постоянной величины равна 0: D(C)=0, C=const.

2.Постоянный множитель можно выносить за знак дисперсии, возводя его при этом в квадрат: D(kX)=k2 D(X).

3.Дисперсия случайной величины равна разности между математическим ожиданием квадрата случайной величины и квадратом ее математического ожидания:



4.Дисперсия алгебраической суммы независимых случайных величин равна сумме их дисперсий: D(X±Y)=D(X)+D(Y).

**Замечание.** Математическое ожидание и дисперсия в финансовом анализе имеют следующие интерпретации. Пусть известно распределение доходности Х некоторого актива (например, акции), т.е. известны значения доходности хi и соответствующие им вероятности рi за рассматриваемый промежуток времени. Тогда математическое ожидание М (Х) выражает среднюю (прогнозную) доходность актива, а дисперсия D(X) или среднее квадратичное отклонение  - меру отклонения, колеблемость доходности от ожидаемого среднего значения, т.е. риск данного актива.

Математическое ожидание, дисперсия, среднее квадратическое отклонение, призванные в сжатой форме характеризовать наиболее существенные черты распределения, называются **числовыми характеристиками** случайной величины.

Следует отметить, что сама величина Х является **случайной**, а ее числовые характеристики являются величинами **неслучайными, постоянными**.

***Задача 1.*** В связке из 3 ключей только один ключ подходит к двери. Ключи перебирают до тех пор, пока не отыщется подходящий ключ. Построить закон распределения для случайной величины ξ – числа перепробованных ключей.

***Задача 2.*** В связке из 3 ключей только один ключ подходит к двери. Ключи перебирают до тех пор, пока не отыщется подходящий ключ. Построить закон распределения для случайной величины ξ – числа опробованных ключей.

***Задача 3.*** Построить функцию распределения Fξ(x) для случайной величины ξ из задачи 1.

***Задача 4.*** Совместный закон распределения случайных величин ξ и η задан c помощью таблицы

|  |  |  |
| --- | --- | --- |
| ξ η | 1 | 2 |
| –1 | 1/16 | 3/16 |
| 0 | 1/16 | 3/16 |
| 1 | 1/8 | 3/8 |

Вычислить частные законы распределения составляющих величин ξ и η. Определить, зависимы ли они. Вычислить вероятность .

***Задача 5*.** Пусть случайная величина ξ имеет следующий закон распределения:

|  |  |  |  |
| --- | --- | --- | --- |
| ξ | –1 | 0 | 2 |
| P | 1/4 | 1/4 | 1/2 |

Вычислить математическое ожидание Mξ, дисперсию Dξ и среднеквадратическое отклонение σ.

1. **Математическая статистика**

В теории вероятностей рассматривались вопросы, связанные с математическими моделями экспериментов со случайным исходом.

*Задачи математической статистики относятся к разработке методов обработки опытных данных*, относящихся к наблюдениям над случайными массовыми явлениями. В зависимости от характера измеряемой величины, цели измерений при обработке результатов измерений эти задачи могут принимать различные формы. Типичными задачами математической статистики являются следующие:

-оценка на основании результатов измерений неизвестной функции распределения;

-оценка неизвестных параметров распределения;

-статистическая проверка гипотез.

Начало интенсивного развития статистических методов исследования определяется работами К. Пирсона, выполненными в конце 19 столетия. Раздел математической статистики , изучающий задачи метода проверки статистических гипотез, создан Ю. Нейманом, Э. Пирсоном, Р. Фишером в конце 20 – 30 годов 20 века.

 Одним из центральных вопросов математической статистики является вопрос оценки параметров распределения. Первые результаты в этой области получены К. Гауссом (1809) и А. Марковым (1900). В дальнейшем методы статистического оценивания получили развитие в работах Р. Фишера.

1. **Генеральная совокупность и выборка. Статистический ряд. Статистическая функция распределения. Полигон и гистограмма.**

 В математической статистике изучение случайной величины связано с выполнением ряда независимых опытов, в которых она принимает определенные значения. Полученные значения случайной величины представляют собой *статистическую совокупность или статистический ряд,* подлежащий обработке и научному анализу.

 Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты. Например, качественным признаком может служить стандартность детали, а количественным – контролируемый размер детали.

 *Выборочной совокупностью, или выборкой, называют совокупность случайно отобранных объектов. Генеральной совокупностью называется совокупность объектов, из которой производится выборка.*

 *Объемом совокупности*  (генеральной или выборочной) называют число объектов этой совокупности. Например, если из 100000 деталей отобрано 100 деталей, то объем генеральной совокупности N=100000, объем выборки равен n = 100. Для того, чтобы выборка правильно представляла пропорции генеральной совокупности она должна быть *репрезентативной, т. е. представительной.* В силу закона больших чисел можно утверждать, что выборка будет репрезентативной, если ее осуществлять случайно, т. е. все объекты выборки имеют одинаковую вероятность попасть в выборку. Выборка может быть *бесповоротной*, если отобранные объекты не возвращаются в генеральную совокупность. Выборка называется *поворотной*, если выбранный объект возвращается в генеральную совокупность перед выбором следующего объекта.

 Пусть из генеральной совокупности извлечена выборка, причем х1 наблюдалось n1 раз, х2 – n2 раза, …, хк – nк раза и  - объем выборки.

 Наблюдаемые значения xi  называются *вариантами*. Последовательность вариант, записанная в порядке возрастания, называется *вариационным рядом*.

Числа наблюдений xi, обозначаемые ni, называются *частотами*, а величины  называются *относительными частотами*. Статистическим распределением выборки называется перечень вариант xi и соответствующих частот ni (или относительных частот Wi)

Рис.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X | x1 | x2 | … | xk |
| nx | n1 | n2 | … | nk |
| Wx | w1 | w2 | … | wk |

 В случае большого количества вариант и непрерывного распределения признака статистическое распределение выборки задают в виде последовательности интервалов и соответствующих им частот. Для этого интервал, в котором заключены все наблюдаемые значения признака, разбивают на определенное число частичных интервалов  длиной  и находят для каждого интервала ni  - сумму частот вариант, попавших в - й интервал. Таким образом получают  *интервальное статистическое распределение выборки:*

Рис.2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| интервалХ | (х0;х1) | (х1;х2) | … | (xk-1;xk) |
| nx | n1 | n2 | … | nk |
| Wx | w1 | w1 | … | wk |

 Статистическое распределение рис 1. или рис.2 называют *статистическим рядом*.

Для графического изображения статистического ряда используют *полигон* и *гистограмму*.

Для построения полигона на оси Ох откладывают значение вариант, на оси Оу - значения частот ni  (или относительных частот ). Построенную таким образом ломаную , отрезки которой соединяют точки (xi;ni) или (xi;wi) называют *полигоном частот* или *полигоном относительных частот*.

В случае непрерывного распределения признака на основании интервального распределения (рис.2) используют гистограмму. Устанавливающую зависимость частот от разряда интервалов, в которые попадают значения случайной величины. Предполагаем, что длины интервалов равны **- шаг распределения. На оси Ох отметим точки х1,х2,…хк с найденным шагом друг от друга. На каждом частичном интервале строим прямоугольник высотой  (плотность частоты). *Гистограммой частот* называют ступенчатую фигуру , состоящую из найденных прямоугольников. Поскольку площадь i –го частичного прямоугольника равна hni/h , то *площадь гистограммы частот* равна сумме всех частот, т. е. объему выборки n.

Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению  (плотность относительной частоты). Для построения гистограммы относительных частот на оси абсцисс откладывают частичные интервалы , а над ними строят прямоугольники высотой . Площадь i –го прямоугольника равна  - относительной частоте вариант, попавших в i –ый интервал. Следовательно, площадь гистограммы относительных частот равна сумме всех относительны частот, т. е. единице. Поэтому гистограмма относительных частот является статистическим аналогом плотности распределения случайной величины Х.

Определение 1. Статистической (эмпирической) функцией распределения (иначе функцией распределения выборки) называют функцию F\*(x), определяющую для каждого значения х относительную частоту события : , где nx - число наблюдений, при которых значение признака Х меньше х, n – объем выборки .

 В отличии от эмпирической функции распределения F\*(x) выборки функция распределения F(x) генеральной совокупности называется теоретической функцией распределения. Различие между эмпирической функцией F\*(x) и теоретической функцией F(x) функциями распределения состоит в том, что F(x)определяет вероятность события , а F\*(x), относительную частоту этого события. Поэтому F\*(x) можно использовать для приближенного представления теоретической функции распределения генеральной совокупности. Функция F\*(x) обладает свойствами F(x):

1. Значения эмпирической функции распределения принадлежат отрезку ;
2. F\*(x), является неубывающей функцией на промежутке ;
3. Если х1- наименьшая варианта, то F\*(x)=0 при ;
4. Если хк – наибольшая варианта, то F\*(x)=1 при х>xk.

**3.Точечные оценки параметров генеральной совокупности по выборочным совокупностям, их свойства. Точечные оценки для математического ожидания и дисперсии случайной величины.**

Пусть закон распределения случайной величины Х содержит неизвестный параметр **а**. Требуется на основании опытных данных найти подходящую оценку для параметра а.

Пусть х1, х2, …хn  - наблюдаемые значения случайной величины Х, получаемые в результате n независимых испытаний. При этом результат статистического распределения выборки , представленного на рис 2, можно представить как набор n независимых случайных величин Х1, Х2, …,Хn, представляющих собой n независимых копий случайной величины Х, именно  - случайная величина, представляющая собой результат i –го опыта, но имеющая тот же закон распределения, что исследуемая случайная величина Х. Случайная величина , построенная на основании статистических данных Х1, Х2, …,Хn,называется точечной оценкой параметра **a. ** является случайной величиной . закон распределения которой зависит от 1) закона распределения случайной величины Х; 2) от числа опытов n. Для того, чтобы оценка имела практическую ценность, она должна обладать свойствами:

1. *Несмещенность* . Оценка **** называется *несмещенной*, если математическое ожидание равно оцениваемому параметру а. т.е**. .**В противном случае оценка **** называется *смещенной*. Естественно в качестве оценки, т.е. приближенного значения неизвестного параметра, брать несмещенные оценки. В этом случае не происходит систематической ошибки в сторону завышения или занижения.
2. *Состоятельность.* Оценка **** называется *состоятельной*, если она сходится по вероятности к оцениваемому параметру а при неограниченном возрастании n: при . Состоятельность оценки означает, что при достаточно большом числе опытов n со сколь угодно большой достоверностью отклонения оценки от истинного значения параметра по модулю меньше любого заранее выбранного числа .
3. *Эффективность*. Оценки, обладающие свойством несмещенности и состоятельности, при ограниченном числе опытов могут отличаться дисперсиями. Чем меньше дисперсия оценки, тем меньше вероятность грубой ошибки при определении приближенного значения параметра. Поэтому необходимо, чтобы дисперсия оценки была минимальной, т.е. чтобы выполнялось условие . Оценка, обладающая этим свойством, называется *эффективной*.

Условия несмещенности, состоятельности и эффективности являются условиями доброкачественности оценки. Что является необходимым при обработке статистических данных.

 *Оценка для математического ожидания случайной величины.*

Пусть исследуется случайная величина Х с математическим ожиданием mx. Обозначим х1,х2, …хn значения случайной величины, полученной в результате n независимых равноточных опытов, т.е. измерений, проходивших в одинаковых условиях. В качестве оценки для mx  принимаем среднее арифметическое наблюдаемых значений , называемым *выборочной средней*

Оценка выборочной средней  является несмещенной, т.к.

.

Выборочная средняя является состоятельной, т.к. по теореме Чебышева  при .

Эффективность оценки выполняется только для узкого класса распределений, в частности для нормального распределения.

 *Оценка для дисперсии случайной величины.*

В качестве оценки для дисперсии рассмотрим следующую величину: .

Эту оценку называют *выборочной дисперсией* .

Эта оценка является состоятельной и несмещенной, что можно увидеть, проведя алгебраические преобразования выражения и получить =.

Первый член выражения есть среднее арифметическое n наблюдаемых значений случайной величины Х2, значит она сходится по вероятности к МХ2. Второй член  сходится по вероятности к mx.

Следовательно, правая часть сходится по вероятности к величине МХ2-mx2 = Dx , это означает состоятельность оценки.

Проверяя несмещенность оценки, в силу независимости случайных величин Х1,Х2,…Хn, 

получаем . Получаем, что выборочная дисперсия является смещенной оценкой. Умножим на  получим для дисперсии  оценку, обладающую свойством *несмещенности*

. Эта оценка называется *«исправленной» выборочной* *дисперсией* и определять формулой .

Величину s называют «исправленным» *средним квадратическим отклонением*.

Если имеем интервальное выборочное распределение, то формулы для выборочной средней, выборочной дисперсии, и «исправленной» выборочной дисперсии имеют вид:



Здесь - среднее значение случайной величины Х на интервале , т.е..

**Мода и медиана.**

*Модой* дискретной случайной величины Х называется ее наиболее вероятное значение.

*Модой* непрерывной случайной величины Х называется то ее значение, при котором плотность распределения максимальна

Мода обозначается символом .

*Медианой* непрерывной случайной величины Х называется такое ее значение µ, для которого одинаково вероятно, окажется ли случайная величина меньше или больше µ, т.е. .

Геометрически мода является абсциссой той точки распределения , ордината которой максимальна. Ордината, проведенная в точке х=µ, делит пополам площадь, ограниченную кривой распределения. Если прямая х =а является осью симметрии кривой распределения , то .

**Задача 1.**

Дана плотность вероятности непрерывной случайной величины .

Найти моду этой случайной величины.

 Найдем максимум функции . Для этого найдем производные первого и второго порядков: 

Из уравнения  получаем х=1.

Так как , то при х=1 функция имеет максимум, т.е..

Мы не определяли значение постоянной величины а, т.к. максимум данной функции  не зависит от числового значения а.

**Задача 2.**

Дана плотность вероятности случайной величины Х1

 0, при х<0,

f(x)= x-x3/4, при ,

 0,при x>2.

Найти медиану этой случайной величины.

Медиану µ найдем из условия . В данном случае .

Тогда получаем уравнение µ2/2-µ4/16=0,5. Следовательно µ4-8µ2+8=0, отсюда получим . Из четырех корней этого уравнения следует выбрать тот, который заключен между 0 и 2. Таким образом .

**Задача 3.**

Дан ряд распределения дискретной случайной величины:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  xi |  10 |  20 |  30 |  40 |  50 |  60 |
|  pi |  0,24 |  0,36 |  0,20 |  0,15 |  0,03 |  0,02 |

Найти моду.

**Задача 4.**

Дана плотность распределения непрерывной случайной величины:

 0,при x<2

 f(x)= , при 

 0,при x>4.

Определить значение а, моду и медиану.

**Решение задач на вычисление характеристик ДСВ, биномиального и геометрического распределения ДСВ**

***Задача 1.***В связке из 3 ключей только один ключ подходит к двери. Ключи перебирают до тех пор, пока не отыщется подходящий ключ. Построить закон распределения для случайной величины ξ – числа опробованных ключей*.*

***Задача 2.*** В партии 10% нестандартных деталей. Наугад отобраны 4 детали. Написать биноминальный закон распределения дискретной случайной величины – числа нестандартных деталей среди четырех отобранных и построить многоугольник полученного распределения.

***Задача* 3.** Две игральные кости одновременно бросают 2 раза. Написать биноминальный закон распределения дискретной случайной величины – числа выпадений четного числа очков на двух игральных костях.

***Задача 4.*** По цели производится 5 выстрелов. Вероятность попадания для каждого выстрела равна 0,4. Найти вероятности числа попаданий и построить многоугольник распределения.

**Функция плотности НСВ. Интегральная функция распределения НСВ. Характеристики НСВ.**

**Решение задач о вероятности попадания случайной величины в заданный интервал**.

***Задача 1.*** Найти вероятность попадания в заданный интервал [a,b] значения нормально распределенной случайной величины X, если известно её математическое ожидание M[X] и дисперсия D[X].

|  |  |  |  |
| --- | --- | --- | --- |
| M[X] | D[X] | https://arhivurokov.ru/multiurok/html/2017/05/11/s_591411e5775f0/625139_4.png | b |
| 4 | 25 | 2 | 7 |

***Задача 2.*** Случайна величина распределена по нормальному закону , а вероятность ее попадания в интервал равна 0,8. Найти вероятность попадания в интервал .

***Задача 3.*** Найти вероятность попадания случайной величины Х в заданный интервал (α,β), если она распределена по указанному закону:

1) равномерное распределение на интервале (a,b ) ;

2) показательное распределение с математическим ожиданием, равным b;

3) нормальное распределение с математическим ожиданием, равным a, и среднеквадратическим отклонением, равным α.

α=10,β=16

a=11,b=20

1. **Математическая статистика**

В теории вероятностей рассматривались вопросы, связанные с математическими моделями экспериментов со случайным исходом.

*Задачи математической статистики относятся к разработке методов обработки опытных данных*, относящихся к наблюдениям над случайными массовыми явлениями. В зависимости от характера измеряемой величины, цели измерений при обработке результатов измерений эти задачи могут принимать различные формы. Типичными задачами математической статистики являются следующие:

-оценка на основании результатов измерений неизвестной функции распределения;

-оценка неизвестных параметров распределения;

-статистическая проверка гипотез.

Начало интенсивного развития статистических методов исследования определяется работами К. Пирсона, выполненными в конце 19 столетия. Раздел математической статистики , изучающий задачи метода проверки статистических гипотез, создан Ю. Нейманом, Э. Пирсоном, Р. Фишером в конце 20 – 30 годов 20 века.

 Одним из центральных вопросов математической статистики является вопрос оценки параметров распределения. Первые результаты в этой области получены К. Гауссом (1809) и А. Марковым (1900). В дальнейшем методы статистического оценивания получили развитие в работах Р. Фишера.

1. **Генеральная совокупность и выборка. Статистический ряд. Статистическая функция распределения. Полигон и гистограмма.**

 В математической статистике изучение случайной величины связано с выполнением ряда независимых опытов, в которых она принимает определенные значения. Полученные значения случайной величины представляют собой *статистическую совокупность или статистический ряд,* подлежащий обработке и научному анализу.

 Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты. Например, качественным признаком может служить стандартность детали, а количественным – контролируемый размер детали.

 *Выборочной совокупностью, или выборкой, называют совокупность случайно отобранных объектов. Генеральной совокупностью называется совокупность объектов, из которой производится выборка.*

 *Объемом совокупности*  (генеральной или выборочной) называют число объектов этой совокупности. Например, если из 100000 деталей отобрано 100 деталей, то объем генеральной совокупности N=100000, объем выборки равен n = 100. Для того, чтобы выборка правильно представляла пропорции генеральной совокупности она должна быть *репрезентативной, т. е. представительной.* В силу закона больших чисел можно утверждать, что выборка будет репрезентативной, если ее осуществлять случайно, т. е. все объекты выборки имеют одинаковую вероятность попасть в выборку. Выборка может быть *бесповоротной*, если отобранные объекты не возвращаются в генеральную совокупность. Выборка называется *поворотной*, если выбранный объект возвращается в генеральную совокупность перед выбором следующего объекта.

 Пусть из генеральной совокупности извлечена выборка, причем х1 наблюдалось n1 раз, х2 – n2 раза, …, хк – nк раза и  - объем выборки.

 Наблюдаемые значения xi  называются *вариантами*. Последовательность вариант, записанная в порядке возрастания, называется *вариационным рядом*.

Числа наблюдений xi, обозначаемые ni, называются *частотами*, а величины  называются *относительными частотами*. Статистическим распределением выборки называется перечень вариант xi и соответствующих частот ni (или относительных частот Wi)

Рис.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X | x1 | x2 | … | xk |
| nx | n1 | n2 | … | nk |
| Wx | w1 | w2 | … | wk |

 В случае большого количества вариант и непрерывного распределения признака статистическое распределение выборки задают в виде последовательности интервалов и соответствующих им частот. Для этого интервал, в котором заключены все наблюдаемые значения признака, разбивают на определенное число частичных интервалов  длиной  и находят для каждого интервала ni  - сумму частот вариант, попавших в - й интервал. Таким образом получают  *интервальное статистическое распределение выборки:*

Рис.2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| интервалХ | (х0;х1) | (х1;х2) | … | (xk-1;xk) |
| nx | n1 | n2 | … | nk |
| Wx | w1 | w1 | … | wk |

 Статистическое распределение рис 1. или рис.2 называют *статистическим рядом*.

Для графического изображения статистического ряда используют *полигон* и *гистограмму*.

Для построения полигона на оси Ох откладывают значение вариант, на оси Оу - значения частот ni  (или относительных частот ). Построенную таким образом ломаную , отрезки которой соединяют точки (xi;ni) или (xi;wi) называют *полигоном частот* или *полигоном относительных частот*.

В случае непрерывного распределения признака на основании интервального распределения (рис.2) используют гистограмму. Устанавливающую зависимость частот от разряда интервалов, в которые попадают значения случайной величины. Предполагаем, что длины интервалов равны **- шаг распределения. На оси Ох отметим точки х1,х2,…хк с найденным шагом друг от друга. На каждом частичном интервале строим прямоугольник высотой  (плотность частоты). *Гистограммой частот* называют ступенчатую фигуру , состоящую из найденных прямоугольников. Поскольку площадь i –го частичного прямоугольника равна hni/h , то *площадь гистограммы частот* равна сумме всех частот, т. е. объему выборки n.

Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению  (плотность относительной частоты). Для построения гистограммы относительных частот на оси абсцисс откладывают частичные интервалы , а над ними строят прямоугольники высотой . Площадь i –го прямоугольника равна  - относительной частоте вариант, попавших в i –ый интервал. Следовательно, площадь гистограммы относительных частот равна сумме всех относительны частот, т. е. единице. Поэтому гистограмма относительных частот является статистическим аналогом плотности распределения случайной величины Х.

Определение 1. Статистической (эмпирической) функцией распределения (иначе функцией распределения выборки) называют функцию F\*(x), определяющую для каждого значения х относительную частоту события : , где nx - число наблюдений, при которых значение признака Х меньше х, n – объем выборки .

 В отличии от эмпирической функции распределения F\*(x) выборки функция распределения F(x) генеральной совокупности называется теоретической функцией распределения. Различие между эмпирической функцией F\*(x) и теоретической функцией F(x) функциями распределения состоит в том, что F(x)определяет вероятность события , а F\*(x), относительную частоту этого события. Поэтому F\*(x) можно использовать для приближенного представления теоретической функции распределения генеральной совокупности. Функция F\*(x) обладает свойствами F(x):

1. Значения эмпирической функции распределения принадлежат отрезку ;
2. F\*(x), является неубывающей функцией на промежутке ;
3. Если х1- наименьшая варианта, то F\*(x)=0 при ;
4. Если хк – наибольшая варианта, то F\*(x)=1 при х>xk.

**3.Точечные оценки параметров генеральной совокупности по выборочным совокупностям, их свойства. Точечные оценки для математического ожидания и дисперсии случайной величины.**

Пусть закон распределения случайной величины Х содержит неизвестный параметр **а**. Требуется на основании опытных данных найти подходящую оценку для параметра а.

Пусть х1, х2, …хn  - наблюдаемые значения случайной величины Х, получаемые в результате n независимых испытаний. При этом результат статистического распределения выборки , представленного на рис 2, можно представить как набор n независимых случайных величин Х1, Х2, …,Хn, представляющих собой n независимых копий случайной величины Х, именно  - случайная величина, представляющая собой результат i –го опыта, но имеющая тот же закон распределения, что исследуемая случайная величина Х. Случайная величина , построенная на основании статистических данных Х1, Х2, …,Хn,называется точечной оценкой параметра **a. ** является случайной величиной . закон распределения которой зависит от 1) закона распределения случайной величины Х; 2) от числа опытов n. Для того, чтобы оценка имела практическую ценность, она должна обладать свойствами:

1. *Несмещенность* . Оценка **** называется *несмещенной*, если математическое ожидание равно оцениваемому параметру а. т.е**. .**В противном случае оценка **** называется *смещенной*. Естественно в качестве оценки, т.е. приближенного значения неизвестного параметра, брать несмещенные оценки. В этом случае не происходит систематической ошибки в сторону завышения или занижения.
2. *Состоятельность.* Оценка **** называется *состоятельной*, если она сходится по вероятности к оцениваемому параметру а при неограниченном возрастании n: при . Состоятельность оценки означает, что при достаточно большом числе опытов n со сколь угодно большой достоверностью отклонения оценки от истинного значения параметра по модулю меньше любого заранее выбранного числа .
3. *Эффективность*. Оценки, обладающие свойством несмещенности и состоятельности, при ограниченном числе опытов могут отличаться дисперсиями. Чем меньше дисперсия оценки, тем меньше вероятность грубой ошибки при определении приближенного значения параметра. Поэтому необходимо, чтобы дисперсия оценки была минимальной, т.е. чтобы выполнялось условие . Оценка, обладающая этим свойством, называется *эффективной*.

Условия несмещенности, состоятельности и эффективности являются условиями доброкачественности оценки. Что является необходимым при обработке статистических данных.

 *Оценка для математического ожидания случайной величины.*

Пусть исследуется случайная величина Х с математическим ожиданием mx. Обозначим х1,х2, …хn значения случайной величины, полученной в результате n независимых равноточных опытов, т.е. измерений, проходивших в одинаковых условиях. В качестве оценки для mx  принимаем среднее арифметическое наблюдаемых значений , называемым *выборочной средней*

Оценка выборочной средней  является несмещенной, т.к.

.

Выборочная средняя является состоятельной, т.к. по теореме Чебышева  при .

Эффективность оценки выполняется только для узкого класса распределений, в частности для нормального распределения.

 *Оценка для дисперсии случайной величины.*

В качестве оценки для дисперсии рассмотрим следующую величину: .

Эту оценку называют *выборочной дисперсией* .

Эта оценка является состоятельной и несмещенной, что можно увидеть, проведя алгебраические преобразования выражения и получить =.

Первый член выражения есть среднее арифметическое n наблюдаемых значений случайной величины Х2, значит она сходится по вероятности к МХ2. Второй член  сходится по вероятности к mx.

Следовательно, правая часть сходится по вероятности к величине МХ2-mx2 = Dx , это означает состоятельность оценки.

Проверяя несмещенность оценки, в силу независимости случайных величин Х1,Х2,…Хn, 

получаем . Получаем, что выборочная дисперсия является смещенной оценкой. Умножим на  получим для дисперсии  оценку, обладающую свойством *несмещенности*

. Эта оценка называется *«исправленной» выборочной* *дисперсией* и определять формулой .

Величину s называют «исправленным» *средним квадратическим отклонением*.

Если имеем интервальное выборочное распределение, то формулы для выборочной средней, выборочной дисперсии, и «исправленной» выборочной дисперсии имеют вид:



Здесь - среднее значение случайной величины Х на интервале , т.е..

**Заключение. Использование математически-статистических методов. Развитие науки**

Экономико-статистические методы анализа и прогнозирования в процессе познания экономических закономерностей имеют два вида моделей:

- нормативный анализ, используемый для выработки управляющих решений в условиях определенных, или хотя бы понятных , для исследователя закономерностей;

- позитивного анализа, используемого для выявления неизвестных или недоступных в рамках эксперимента закономерностей и тенденций.

Экономико-статистические методы анализа и прогнозирования представляют математический аппарат построения моделей позитивного анализа на основе зафиксированных ранее (ретроспективных) данных о состоянии экономического объекта или процесса. Методы базируются на фундаментальных разделах математики – теории вероятностей и математической статистике. Эти методы изучаются в разделе эконометрики. Развитие эконометрики (новой математической экономики) началось примерно с 1950 года. На этой стадии получили признание работы Дж. Фон Неймана,

 К. Эрроу, М. Морришимы, В. Леонтьева, написанные им после эмиграции из России по вопросу метода «затраты-выпуск», а так же работы советских ученых Л. Кантаровича, Я. Понтрягина. Школа Канторовича положила начало развитию фундаментальных работ Агангебяна, Гранберга , Шаталина, Ясина. Сам Канторович был удостоен Нобелевской премии за исследования в теории оптимального распределения ресурсов, по анализу объективно обусловленных (двойственных) оценок в 1975 году. За последние зол е по этой науке были присуждены ; Нобелевские премии, наука является одной из самых молодых.

Используются понятия:

- модель- мера, мерило, образец, норма. В мат. под моделью понимается произвольное множество с заданными на нем свойствами отношениями.

- моделирование – исследование объектов не непосредственно, а косвенным образом, с помощью аналитических методов и моделей.

- объект моделирования – наблюдаемый процесс развития экономического объекта во времени

- математическая модель экономического объекта, экономико-математическая модель – совокупность математических уравнений, описывающих функционирование экономического объекта с заданной степенью точности

- основные переменные, с помощью которых описывается система – объемы различных товаров и услуг, которые производятся и потребляются, прибавляются и вычитаются из имеющихся запасов, продаются и покупаются.

Для построения уравнений нужны данные- имеющиеся ресурсы различного рода, уровень технических знаний, природа потребительских предпочтений, природа предпочтений регулирующих органов. Из этих данных и переменных формируются условия функционирования экономического объекта, т.е. система уравнений или неравенств.

Экономическая теория содержательно сформулировала основные цели, преследуемые людьми при оценке вариантов стратегий поведения:

- один человек – достичь наибольшего удовлетворения своих потребностей;

- группа организованных людей стремится к максимальной прибыли предприятия;

- национальное сообщество (государство) – обеспечить максимальное благосостояние населения.

В эконометрике выделяются три основных направления:

- разработка и исследование методов прикладной статистики с учетом специфики экономических данных;

- Разработка экономических моделей в соответствии с потребностями науки и практики;

- применение эконометрических методов для анализа и прогнозирования состояния конкретных экономических объектов.

При этом в эконометрике используются следующие методы:

- статистика случайных величин (выборочные методы, оценка законов распределения случайных величин); природа элементов исследования – числа;

- многомерный статистический анализ (корреляционно – регрессионный анализ); природа элементов исследования – векторы;

- анализ и прогнозирование временных рядов; природа элементов – функции;

- статистика объектов нечисловой природы; объекты исследования есть элементы пространств, в которых нет операций сложения и умножения, а используются качественные признаки, бинарные отношения (ранжировки, разбиения, толерантность), нечеткие множества, интервалы, тексты.

В эконометрике решаются следующие задачи:

- описание и усреднение данных; оценивание и проверка гипотез;

- установление зависимостей;

- классификация объектов и признаков; прогнозирование статистических решений.

Эконометрические данные отличаются от технических, следовательно, и методы работы специфические.

Информационной базой исследования являются эмпирические ряды числовых значений показателей, характеризующих состояние объекта в разные моменты времени. Они оформляются в таблицах или изображаются графически - в виде геометрических образов – точек, линий, фигур в различных сочетаниях.

Модель при этом представляет собой описание компонентов и функций, формирующих объект. В модели оперируют показателями, исчисленными для качественно однородных массовых явлений.

Выражение модели в виде функциональных уравнений используют для расчета средних значений моделируемого показателя по набору заданных величин и выявления степени влияния на них отдельных факторов.

Модели могут быть однофакторными и многофакторными.

Эконометрические методы показывают, как производится обработка первого статистического материала, чтобы в дальнейшем с достаточной достоверностью прогнозировать развитие событий.

Стат исследования устанавливают качественное и количественное влияние экономических показателей друг на друга. Но эти зависимости по своей природе стохастичны, т.е. позволяют установить лишь вероятностные, логические соотношения между явлениями.

Система описывается набором переменных:

-входные переменные, описывающие условия функционирования, часть из них поддается регулированию «х», называемые независимыми, экзогенными, объясняющими или факторными;

- выходные переменные, характеризующие поведение объекта или результат функционирования «у», называемые эндогенными, зависимыми, результирующими или откликами;

- ε - латентные случайные компоненты, отражающие влияние неучтенных факторов, а так же неточности измерений. Эконометрические методы используются как составная часть технико-экономического научного исследования: оценка точности и стабильности процессов, разработка методов контроля тех процессов, оптимизация выхода полезного продукта, изучение спроса, экспертных оценках принятия решений в стратегическом, инновационном, инвестиционном менеджменте. Перспективно их применение в изучении рисков инновационных проектов, задачах страхования, разработке стратегии антикризисного управления.

***Методы действий***

Моделирование механизма рыночных связей, взаимодействие спроса и предложения, влияние объема и структуры товарооборота на объем и состав продукции, формирование товарных запасов, издержек производства, прибыли и других качественных показателей имеет первостепенное значение для прогнозирования конъюнктуры рынка, рационализации производственных и торговых процессов, успешного ведения бизнеса.

В зависимости от цели модели подразделяются на структурные, динамические и модели связи.

В основе **структурной модели** лежит описание структуры связи между изучаемыми переменными в форме так называемых **одновременных уравнений**. Структурные (причинные) модели в эконометрике соединяют теорию объекта с эмпирическими данными на основе графа связи. Они формализуют гипотезы о причинных отношениях.

Модель является **динамической**, если в данный момент времени она учитывает значения входящих в нее переменных, относящихся как к текущему, так и к предыдущим моментам времени. Эта модель отражает динамику исследуемых процессов в каждый момент времени.

**Модели связи** представляют собой математические уравнения, связывающие эндогенные и экзогенные переменные и отражающие статистически значимые взаимосвязи.

Проблемы моделирования решаются методами корреляционного и регрессионного анализа.

Между различными явлениями и их признаками можно выделить связь двух типов: функциональную (жестко детерминированную) и стохастическую (вероятностную).

Связь называется **функциональной**, если каждому независимому признаку Х соответствует один или несколько определенных значений независимого признака У.

**Стохастическая связь** характеризуется тем, что зависимая величина У реагирует на изменение независимой величины Х изменением закона распределения.

**Корреляционная связь** существует там, где взаимосвязанные явления характеризуются случайными величинами. При такой связи среднее значение (математическое ожидание) случайной величины результативного признака У закономерно изменяется в зависимости от изменения случайной величины Х. Корреляционная связь проявляется не в каждом отдельном случае, а во всей совокупности в целом.

В общем виде задача изучения связей состоит не только в количественной оценке их наличия, но ив определении формы (аналитического выражения) влияния факторных признаков на результативные.

Односторонняя вероятностная зависимость между случайными величинами есть **регрессия**. Она устанавливает соответствие между величинами.

Односторонняя стохастическая зависимость с помощью функции, которая называется **регрессией**.

Перечислим виды регрессий:

*1.Регрессия относительно числа переменных:*

 -простая регрессия – регрессия между двумя переменными;

 -множественная регрессия – зависимость между зависимой переменной У несколькими объясняющими переменными Х1,Х2 ,..Хп , имеющая вид:

, где у – функция регрессии; х – независимые переменные; а1,2,.. – коэффициенты регрессии; а0 – свободный член уравнения; м – число факторов, включаемых в модель.

*2.Регрессия относительно форм зависимости:*

 -линейная регрессия, выражаемая линейной функцией;

 -нелинейная регрессия, выражаемая нелинейной функцией.

*3.Регрессия относительно ее характера:*

 -положительная регрессия, прямо пропорциональную зависимость между независимой и зависимой переменными.

 - отрицательная регрессия, имеющая обратную зависимость.

*4..Регрессия относительно типа соединения явлений:*

 -непосредственная зависимость

 -косвенная зависимость ( объясняющая переменная действует на зависимую через ряд других переменных)

 -ложная регрессия возникает при формальном подходе к исследуемым явлениям, без уяснения причин, обуславливающих данную связь.

Регрессия тесно связана с корреляцией. В широком смысле корреляция означает связь , соотношение между объективно существующими объектами. Связи между явлениями могут быть различны по силе. При измерении тесноты связи говорят о корреляции в узком смысле слова. Если случайные переменные причинно обусловлены и в вероятностном смысле можно говорить об их связи, то имеется корреляция.

*Корреляция имеет аналогичную регрессии классификацию по видам.*

Любое причинное влияние может выражаться либо функциональной зависимостью, либо корреляционной связью.

Но не каждая функция или корреляция соответствует причинной зависимости между явлениями, поэтому обязательно требуется выяснение причинно-следственных связей.

Исследование корреляционных связей называется *корреляционным анализом*, а исследование односторонних стохастических зависимостей называется *регрессионным* *анализом*.

**Задача корреляционного анализа** – измерение тесноты известной связи между варьирующими признаками и оценке факторов, оказывающих наибольшее влияние на результативный признак.

**Задача регрессионного анализа** – выбор типа модели (формы связи), установление степени влияния независимых переменных на зависимую и определение расчетных значений зависимой переменной (функции регрессии).

Основные термины

1.Если каждому значению одной переменной соответствует определенное(условное) распределение другой переменной, то эта зависимость называется стохастической(вероятностной) или статистическим исследованием.

2.Корреляционной зависимостью между двумя переменными величинами называется функциональная зависимость между значением одной величины и условным математическим ожиданием другой величины.

Mx (Y)=ϕ(x); My (X)=τ(Y): ϕ(X)≠const; τ(Y)≠const.

3.Ковариацией (или корреляционным моментом Cov(x,y) или Кху случайных величин х,у называется математическое ожидание произведения отклонений этих величин от своих математических ожиданий.

Ковариация двух случайных величин характеризует степень зависимости их друг от друга и их рассеяния вокруг точки (ах ,ау).

Ковариация двух независимых величин равна 0.

Корреляция есть уточнение ковариации.

4.Эконометрика есть раздел науки, находящийся на стыке математики и экономики.

Большинство экономических факторов представляют собой случайные величины. Их статистическое оценивание есть задача эконометрики.

5.Статистические связи между переменными изучаются методами корреляционного и регрессивного анализа.

6.Корреляционный анализ выявляет связи между случайными переменными и оценивает их тесноту.

  - коэффициент корреляции;

 -выборочный коэффициент корреляции.

9.Условное математическое ожидание случайной величины У при значении Х=х, т.е. Мх (У) есть функция от Х, называемая регрессией У по Х.

Аналогично определяется Му (Х).

8.Регрессионный анализ имеет задачу установления формы и изучения зависимости между переменными.

9.Запаздывание воздействия фактора на результат называется лагом.

10.Уравнения, сдвинутые на некоторый момент, называются лаговыми.

11.Функции двух переменных широко используются в экономической теории

Линия уровня производственной функции называется изоквантой.

Функция полезности называется кривой безразличия.

 **Свойства некоторых оценок выборочного метода**

|  |  |  |
| --- | --- | --- |
| Параметр (генеральная характеристика) | Оценка (выборочная характеристика) | Свойство оценки |
| Несмещенная | Состоятельная | Эффективная | Достаточная |
| Средняя |  |  | да | да | да | да |
| Дисперсия |  | ; | нетда | дада | нетнет | данет |
| Доля |  |  | да | да | да | да |
| Мода |  |  | да | да | нет | нет |
| Медиана |  |  | да | да | нет | нет |

**Формулы средних квадратических ошибок выборки**

|  |  |
| --- | --- |
| Оцениваемый параметр | Формулы средних квадратических ошибок выборки |
| Повторная выборка | Бесповторная выборка |
| Генеральная средняя |  |  |
| Генеральная доля |  |  |

**Методика расчета вероятностей и числовых характеристик НСВ.**

**Решение задач на вычисление числовых характеристик НСВ**

***Задача 1****.* Плотность распределения непрерывной случайной величины имеет вид:



Определить константу C, построить функцию распределения Fξ(x) и вычислить вероятность .

***Задача 2*.** Для случайной величины ξ из задачи 1 вычислить математическое ожидание и дисперсию*.*

***Задача 3****.* Пусть задана случайная величина . Вычислить вероятность .

**Статистические оценки параметров распределения по выборочным данным.**

**Тема 3.1. Генеральная совокупность и выборка.**

**Построение для заданной выборки ее графической диаграм­мы; расчёт по заданной выборке её числовых характеристик.**

***Задача 1.*** Из партии электроламп взята 20%-ная случайная бесповторная выборка для определения среднего веса спирали.

Результаты выборки следующие:



**Определите:** с вероятностью 0,95 доверительные пределы, в которых лежит средний вес спирали, для всей партии электроламп.

***Задача 2.*** На основе случайного повторного выборочного обследования в отделении связи города предполагается определить долю писем частных лиц в общем объеме отправляемой корреспонденции. Никаких предварительных данных об удельном весе этих писем в общей массе отправляемой корреспонденции не имеется.

**Определите:** численность выборки, если результаты выборки необходимо дать с точностью до 1% и гарантировать это с вероятностью 0,95.

***Задача 3.*** В городе 500 тыс. жителей. По материалам учета городского населения было обследовано 50 тыс. жителей методом случайного бесповторного отбора. В результате обследования установлено, что в городе 15% жителей старше 60 лет.

**Определите:** с вероятностью 0,683 пределы, в которых находится доля жителей в городе в возрасте старше 60 лет

**Интервальная оценка, интервальная оценка математического ожидания.**

**Интервальное оценивание математического ожидания нор­мального распределения.**

***Задача 1.*** В области, состоящей из 20 районов, проводилось выборочное обследование урожайности на основе бесповторного отбора серий (районов). Выборочные средние по районам составили соответственно 14,5 *ц/га*; 16,0; 15,5; 15,0 и 14,0 *ц/га*.

**Определите:** с вероятностью 0,954 пределы урожайности во всей области.

***Задача 2.*** При проверке веса импортируемого груза на таможне методом случайной повторной выборки отобрано 200 изделий. В результате был установлен средний вес изделия 30 *г* при среднем квадратическом отклонении 4 *г*.

**Определите:** c вероятностью 0,9973 пределы, в которых находится средний вес изделий в генеральной совокупности.

**Методика моделирования случайных величин.**

**Моделирование случайных величин.**

**Решение задач на моделирование сложных испытаний и их результатов.**

***Задача 1.*** Разыграть 6 возможных значений дискретной случайной величины Х, закон распределения которой задан в виде таблицы:

|  |  |  |  |
| --- | --- | --- | --- |
| Х | 2 | 10 | 18 |
| Р | 0,22 | 0,17 | 0,61 |

***Задача 2.*** Заданы вероятности 3 событий А1, А2, А3, образующих полную группу: р1= р(А1)= 0,22, р2= р(А2)= 0,31, р3= р(А3)= 0,47. Разыграть пять испытаний, в каждом из которых появляется одно из трех рассматриваемых событий.

***Задача 3.*** События А и В независимы и совместны. Разыграть 4 испытания, в каждом из которых вероятность появления события А равна 0,7, а событие В-0,4.

**Критерии оценки контрольной работы:**

Отметка «5» ставится, если: работа выполнена полностью;
• в логических рассуждениях и обосновании решения нет пробелов и ошибок;
• в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).
Отметка «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). Отметка «3» ставится, если: допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме. Отметка «2» ставится, если:
допущены существенные ошибки, показавшие, что учащийся не владеет
обязательными умениями по данной теме в полной мере.

**Правила построения доверительных интервалов для неизвестных параметров распределения** Таблица 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Неизвестный параметр | Условия оценки | Вид используемого распределения | Границы интервала | Доверительный интервал |
| Математическое ожидание  |  - известно |  Ф(t) – функция Лапласа для нормального распределения |   |    |
|  -неизвестно |  S(t) – распределение Стьюдента |  |
| Дисперсия  |    m - известно |   - распределение Пирсона |  |  |
|    m - неизвестно |   - распределение |  |  |
| Дисперсия  |  n > 30 | Ф(u) – функция Лапласа |    |
| Вероятность р |  | Ф(u) – функция Лапласа |  |  |

**Условия применения различных статистических гипотез** Таблица 5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № | Тип гипотезы Н0 |  Условия | Границы критической области на уровне значимости α | Статистика наблюдений |
| 1 | О числовом значении генерального среднего  или  | - известно для  | Функция Лапласа  |   |
| - неизвестно для  | Распределение Стьюдента с к = n – 1 степенями свободы |  |
| 2 | О числовом значении дисперсии  или  |  - гипотетическое значение для  | Распределение Пирсона с к = n – 1 степенями свободы |  |
| 3 | Сравнение средних двух совокупностей  |  -известны | Функция Лапласа ; Z – нормированная НСВ |  |
| - неизвестны, малые независимые выборки объема n, m | Распределение Стьюдента с к = n1+n2 – 2 степенями свободы |  |
| 4 | Сравнение относительной частоты с гипотетической вероятностью р = р0 |  - относительная частота | Функция Лапласа  |  |
| 5 | О законе распределения (критерий согласия гипотезы о теоретическом распределении с опытными данными) | r – число параметров теоретического распределения, вычисленных по выборке | Распределение  с k = s – r – 1степенями свободы, где s – число интервалов группировки |  |

 **Соответствие терминов (обозначений, формул) вариационного ряда**

 **и случайной величины**

|  |  |
| --- | --- |
| **Вариационный ряд** | **Случайная величина** |
| Обозначения, формулы | термин | Обозначения , формулы | термин |
| **-** | Дискретный ряд | - | Дискретная случайная величина |
| **-** | Интервальный ряд | **-** | Непрерывная случайная величина |
|  | Вариант | **,х** | Значение случайной величины |
|  | Частость | **,Р** | Вероятность |
| **-** | Полигон, гистограмма | **-** | Многоугольник распределения (полигон), кривая распределения |
|  | Эмпирическая функция |  | Функция распределения |
|  | Средняя арифметическая |  | Математическое ожидание |
|  | Дисперсия |  | Дисперсия |
|  | Среднее квадратическое отклонение |  | Среднее квадратическое отклонение |
|  | Мода |  | Мода |
|  | Медиана |  | Медиана |

**Применение метода Монте –Карло к вычислению определенных и кратных интегралов**

1. ***Вычисление определенных интегралов методом Монте-Карло***

Требуется вычислить интеграл . Пусть t – равномерно распределенная случайная величина, p(t)- плотность распределения вероятности этой случайной величины:

 0, при t<0

 P(t)= 1, при 0≤t≤1

 0, при t>1.

Тогда математическое ожидание случайной функции ϕ(t) определяется равенством .

Учитывая значение p(t), получим . (1)

Найдем приближенное значение математического ожидания. Пусть в результате N испытаний получено N значений случайной величины : t1,t2,…,t N. Эти значения можно взять из таблицы случайных чисел (см. таблицу). Тогда приближенное значение , по теореме Чебышева, определяется равенством . (2)

Из равенства (1) и(2) следует, что  . (3)

 Если рассмотреть общий случай, т. е. найти интеграл . Перейдем к новой переменной t с помощью равенства . Тогда , (4)

 где . Используя формулу (3) для приближенного вычисления интеграла в правой части равенства (4), получим , тогда

 (5)

Расчетная таблица для вычисления определенного интеграла по формуле (5) имеет вид:

|  |  |  |  |
| --- | --- | --- | --- |
|  i |  ti |  xi=a+(b-a)ti |  f(xi) |
|  1  2 3 … … N |  t1 t2 t3 … …  tN |  X1X2 X3 … … XN |  f(x1) f(x2) f(x3) … … F(xN) |
|  |   |

***Пример 1***

С помощью формулы (3) найти приближенное значение интеграла  , взяв из таблицы случайных чисел подряд 30 значений , ограничиваясь тремя цифрами.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  i | ti |  |  | i | ti |  |  | i | ti |  |
| 1 | 0,875 | 0,734 |  | 11 | 0,609 | 0,371 |  | 21 | 0,070 | 0,005 |
| 2 | 0,457 | 0,209 |  | 12 | 0,179 | 0,032 |  | 22 | 0,692 | 0,478 |
| 3 | 0,499 | 0,249 |  | 13 | 0,974 | 0,949 |  | 23 | 0,696 | 0484 |
| 4 | 0,762 | 0,581 |  | 14 | 0,011 | 0,0001 |  | 24 | 0,203 | 0,041 |
| 5 | 0,431 | 0,186 |  | 15 | 0,098 | 0,010 |  | 25 | 0,350 | 0,122 |
| 6 | 0,698 | 0,487 |  | 16 | 0,805 | 0,648 |  | 26 | 0,900 | 0,810 |
| 7 | 0,038 | 0,001 |  | 17 | 0,516 | 0,266 |  | 27 | 0,451 | 0,203 |
| 8 | 0,558 | 0,311 |  | 18 | 0,296 | 0,088 |  | 28 | 0,318 | 0,101 |
| 9 | 0,653 | 0,426 |  | 19 | 0,149 | 0,022 |  | 29 | 0,798 | 0,637 |
| 10 | 0,573 | 0,328 |  | 20 | 0,815 | 0,664 |  | 30 | 0,111 | 0,012 |

Таким образом, ,

 из этого по формуле (3) получаем 

Точное значение интеграла вычислим и определим погрешность приближенного вычисления :

.

Абсолютная погрешность составляет . Тогда относительная погрешность

.

***Пример 2***

Вычислить определенный интеграл, используя формулу (5).

Из таблицы случайных значений возьмем 20 значений, начиная с третьего. Расчетная таблица имеет вид:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| i | ti | Xi=2+ti | Xi2 | Xi3 | F(xi)=xi2+xi3 |
| 1 | 0,499 | 2,499 | 6,245 | 15,606 | 21,851 |
| 2 | 0,762 | 2,762 | 7,629 | 21,070 | 28,699 |
| 3 | 0,431 | 2,431 | 5,910 | 14,367 | 20,277 |
| 4 | 0,698 | 2,698 | 7,279 | 19,639 | 26,918 |
| 5 | 0,038 | 2,038 | 4,153 | 8,464 | 12,617 |
| 6 | 0,558 | 2,558 | 6,543 | 16,738 | 23,281 |
| 7 | 0,653 | 2,653 | 7,038 | 18,672 | 25,710 |
| 8 | 0,573 | 2,573 | 6,620 | 17,034 | 23,654 |
| 9 | 0,609 | 2,609 | 6,807 | 17,759 | 24,566 |
| 10 | 0,179 | 2,179 | 4,748 | 10,346 | 15,094 |
| 11 | 0,974 | 2,974 | 8,645 | 26,305 | 32,150 |
| 12 | 0,011 | 2,011 | 4,044 | 8,133 | 12,177 |
| 13 | 0,098 | 2,098 | 4,402 | 9,235 | 13,673 |
| 14 | 0,805 | 2,805 | 7,868 | 22,07 | 29,938 |
| 15 | 0,516 | 2,516 | 6,330 | 15,926 | 22,256 |
| 16 | 0,296 | 2,296 | 5,276 | 12,104 | 17,380 |
| 17 | 0,149 | 2,149 | 4,618 | 9,924 | 14,542 |
| 18 | 0,815 | 2,815 | 7,924 | 22,307 | 30,231 |
| 19 | 0,070 | 2,070 | 4,285 | 8,870 | 13,155 |
| 20 | 0,692 | 2,692 | 7,247 | 19,508 | 26,755 |

Используем формулу (5) при a=2,b=3, N=20, .

Получим .

Относительная погрешность составляет .

**Обобщенная линейная модель множественной регрессии.**

Если выполняются условия , приведенные в определении, то модель множественной линейной регрессии У=Хβ+ε называется обобщенной линейной моделью множественной регрессии.

При этом результаты анализа точности модели, оценки значимости и построение интервалов оказываются неверными, т. к. неверна оценка ковариационной матрицы вектора оценок.

Обобщенный метод наименьших квадратов.

**Обобщенный метод наименьших квадратов.**

Если выборочные данные представляют собой временные ряды, то часто оказывается, что результаты предыдущих наблюдений влияют на результаты последующих, т.е. ошибки регрессии εi не являются независимыми и условие ковариации Cov(εi ,εj )=0 не выполняется. В этих случаях ковариационная матрица  не является диагональной. Для ее оценки и дальнейшего применения метода следует принять предположения, уменьшающие количество оцениваемых параметров. Введем предположение о **стационарности** временного ряда.

 Временной ряд называется строго стационарным, если совместное распределение вероятностей наблюдений у1, у2, ...,ук одно и то же для любых наборов t1,t2,...,tk. В этом случае среднее значение  и дисперсия  не зависит от номера наблюдений и оценивается по формулам

, .

**Теорема Айткена.**

В классе линейных оценок параметра β для обобщенной регрессионной модели У=Хβ+ε эффективной является оценка .

При этом ковариационная матрица вектора оценок b имеет вид .

Приведенная оценка  называется оценкой обобщенного метода наименьших квадратов.

**Обобщенный метод наименьших квадратов для пространственной выборки. Гетеорскедастичность.**

 **Гетероскедастичность.**

Пусть имеется пространственная выборка (х;у).

Регрессионная модель У=Хβ+ε называется гетероскедастичной, если матрица

 является диагональной с различными элементами диагонали

D(εi )=σi 2 , т.е. дисперсии ошибок регрессии различны.

**Тесты на выявление гетероскедастичности.**

1). Тест ранговой корреляции Спирмена 

2) Тест Голдфельда-Квандта 

**Устранение гетероскедастичности-**

Оценка параметров и применение обобщенного метода наименьших квадратов.

**Пример.**

Рассматривается модель **** зависимости цены автомобиля У от его пробега Х (марка, год выпуска фиксированы).

Данные приведены в таблице:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| i |  У(цена, тыс.усл.ед)\Х(пробег,тыс.км) | 0-1 | 1-2 | 2-3 | 3-4 | Свыше 4 |
| 1 | 0-20 |  |  |  |  | 125 |
| 2 | 20-40 |  |  |  | 98 | 236 |
| 3 | 40-60 |  |  | 398 | 104 | 60 |
| 4 | 60-80 |  | 93 | 165 | 49 | 37 |
| 5 | Свыше 80 | 18 | 44 | 39 | 23 | 11 |

Вычислить групповые дисперсии и сделать предположение о гетероскедастичности модели.

Решение.

Имеем σ, где все σ есть выборочная дисперсия цены автомобиля с пробегом, принимающим значения из вариационного ряда Х. При этом групповые средние возрастают. Это позволяет предположить, что модель гетероскедастична.. При этом дисперсии ошибок ε возрастают с ростом значений регрессоров Х.

**Метод наименьших квадратов.**

1.*Постановка задачи*. Производится n наблюдений  переменных х;у .Предполагая, что между переменными х и у существует зависимость вида y=f(x), найти значения параметров a,b, наилучшим образом согласованные с экспериментальными данными.

Согласно методу наименьших квадратов параметры функции f(x) следует выбирать так, чтобы сумма квадратов невязок

S= была наименьшей.

2.*Если функция y=f(x) линейная,* т.е. у=ах+в, то s=, а неизвестные параметры а, в определяются из системы нормальных уравнений 

 .

3.*Если y=f(x) квадратичная функция,* т.е. у=ах2 +вх+с, то s=

и неизвестные параметры а,в,с определяются из системы нормальных уравнений : 



.

**Задача 1.** Имеются данные о величине пробега автомобиля х (тыс. км/ч) и у- расходе масла (л/тыс.км):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  xi |  50 |  70 |  90 |  110 |  130 |
|  yi |  0,2 |  0,5 |  0,8 |  1,1 |  1,3 |

Полагая, что между переменными х и у существует линейная зависимость, найти эмпирическую формулу у=ах+в методом наименьших квадратов.

Решение. Найдем необходимые для решения суммы . Промежуточные вычисления представим в виде таблицы:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  i |  xi |  yi |  xi yi |  Xi2 |
|  1 2 3 4 5 |  50 70 90 110 130 |  0,2 0,5 0,8 1,1 1,3 |  10 35 72 121 169 |  2500 4900 8100 12100 16900 |
|  Σ |  450 |  3,9 |  407 |  44500 |

Система нормальных уравнений имеет вид:

 44500а+450в=407

 450а+5в=3,9.

Ее решение имеет вид а=0,014, в=-0,48.

Тогда линейная зависимость примет вид: у = 0,014х-0,48.

**Задача 2.** Имеются четыре пары переменных (х;у), результаты которых приведены в таблице:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  х |  1 |  2 |  3 |  4 |
|  у |  0,2 |  0,3 |  1,0 |  1,2 |

Методом наименьших квадратов построить линейную зависимость у= ах + в.

Сравнить полученную зависимость с зависимостью .

Аналогично предыдущей задаче найдем уравнение линейной зависимости

У = 0,37х – 0,25.

Сравним величины  для найденной линейной зависимости и зависимости  . Промежуточные вычисления приведены в таблице:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| i | xi | yi | 1/8x2i | 0,37xi -0,25 | (1/8x2i –yi )2 | (0,37xi -0,25-yi )2 |
| 1234 | 1234 | 0,20,31,01,2 | 0,1250,51,1252 | 0,120,490,761,23 | 0,0056250,0400000,0156250,640000 | 0,00640,03610,05760,0009 |
| Σ | - | - | - | - | 0,701250 | 0,1010 |

Из вычисления видно, что Sлин < Sквадр , следовательно, линейная зависимость предпочтительнее.

Задача 3. Имеются следующие данные о цене на нефть *х*(усл.ед) и индексе акций нефтяных компаний *у*(усл.ед)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| xi | 17,3 | 17,0 | 18,3 | 18,8 | 19,2 | 18,5 |
| yi | 537 | 534 | 550 | 555 | 560 | 552 |

Считая, что между переменными существует линейная зависимость, найдем эмпирическую формулу у = ах + в методом наименьших квадратов

.

Вычислим необходимые суммы:







.

Составим систему уравнений:

1987,51а+109,1в=59831,1

 109,1а+6в=3288

Из нее найдем, что а = 11,946; в = 330, 78.

Тогда у = 11,946х + 330,78.

**Биномиальный закон распределения. Закон Пуассона.**

Если вероятность наступления случайного события в каждом испытании равна р ,то вероятность

того, что при n испытаниях событие осуществится m раз, определяется формулой Бернулли:

.

Закон распределения случайной величины Х, которая может принимать n+1 значение (0;1;2;…п), описываемый формулой Бернулли, называется биномиальным.

Закон распределения случайной величины Х, которая может принимать любые целые неотрицательные значения (1;2;3…п), описываемый формулой , называется законом Пуассона.

Закон распределения Пуассона является законом распределения вероятностей для следующих случайных величин.

a) Пусть на интервале (0,N) оси ОХ случайно размещаются п точек независимо друг от друга. Причем события, заключающиеся в попадании одной точки на любой наперед заданный отрезок постоянной (например, единичной) длины, равновероятны.

Если , то случайная величина Х, равная числу точек, попадающих на на заданный отрезок единичной длины (которая может принимать значения 0,1,2,…,m,…), распределяется по закону Пуассона.

b) Если п равно среднему числу вызовов абонентов, поступающих на за один час на данную телефонную станцию, то число вызовов , поступающих за одну минуту, приближенно распределяется по закону Пуассона и определяется по следующим формулам:

-для биномиального закона: ;

-для закона Пуассона: .

**Задача 1.**

Автоматическая телефонная станция получает в среднем за час 300 вызовов. Какова вероятность того, что за данную минуту она получит точно два вызова?

 За минуту АТС получает в среднем 300/60=5 вызовов, т. е. а=5. Требуется найти Р2. По формуле Пуассона найдем .

**Задача 2.**

 Книга в 1000 страниц имеет 100 опечаток. Какова вероятность того, что на случайно выбранной странице имеется не менее 4 опечаток?

 Среднее количество опечаток на страницу а=1000/100=0,1

По формуле Пуассона получаем: 

Если m=0, то , если m=1, то ; если m=2, то ; если m=3, то . Сумма  является вероятностью того, что на странице окажется не более трех опечаток. Тогда вероятность того, что на случайно выбранной странице не менее четырех опечаток , равна

  = 1-1,105167∙0,904837=1-0,99996=0,000004.

**Задача 3.**

Случайная величина Х подчинена биномиальному закону распределения

. Определить математическое ожидание этой случайной величины.

Имеем

 

Следовательно ,

т.е. M(X)=np.

**Равномерное распределение**

**Равномерным** называется распределение таких случайных величин, все значения которых лежат на некотором отрезке  и имеют постоянную плотность вероятности на этом отрезке.

Таким образом,

 0,при x<a,

 f(x)= h, при a≤x≤b

 0, при x>b.

Так как h(b-a)=1 ,то h=1/(b-a), следовательно

 o,при x<a

 f(x)= 1/(b-a),при a≤x≤b

 0,при x>b.

**Задача 1.**

Определить математическое ожидание случайной величины с равномерным распределением.

 ,

Т.е. М(Х)=(а+в)/2, как это и должно быть в силу симметрии распределения.

**Задача 2.**

Вычислить дисперсию и среднее квадратичное отклонение для случайной величины с равномерным распределением.

 Используем формулу  ,воспользовавшись найденным в задаче 1 значением М(Х)=(а+в)/2. Остается вычислить М(Х2); имеем.

Отсюда получим .

Следовательно, .

**Список литературы**

1. Бирюкова Л.Г., Бобрик Г.И., Матвеев В.И., Теория вероятностей и математическая статистика: Учебное пособие/ Бирюкова Л.Г., Бобрик Г.И. М.:НИЦ ИНФРА-М, 2017. — 312 с. —Текст : электронный // Электронно-библиотечная система «Знания» : [Электронный ресурс]. — URL: https://znanium.com/bookread2.php?book=370899.

2. Белько И.В., Морозова И.М., Криштапович Е.А.., Теория вероятностей, математическая статистика, математическое программирование: Учебное пособие / Белько И.В., Морозова И.М., Криштапович Е.А.. — М.:НИЦ ИНФРА-М, Нов. знани, 2016. — 351 с. — Текст : электронный // Электронно-библиотечная система «Знания» : [Электронный ресурс]. — URL: https://znanium.com/bookread2.php?book=542521

3. Сапожников П.Н., Макаров А.А., Радионова М.В., Теория вероятностей, математическая статистика в примерах, задачах и тестах: Учебное пособие / Сапожников П.Н., Макаров А.А., Радионова М.В. — М.:КУРС, НИЦ ИНФРА-М, 2016. — 140 с. —Текст : электронный // Электронно-библиотечная система «Знания» : [Электронный ресурс]. — URL: https://znanium.com/bookread2.php?book=548242.

4. Е.С. Кочетков, С.О. Смерчинская, В.В. Соколов. Теория вероятностей и математическая статистика : учебник / Е.С. Кочетков, С.О. Смерчинская, В.В. Соколов. – М. : ФОРУМ : ИНФРА-М, 2017. – 312 с. : ил. – [Электронный ресурс]. – URL: https://znanium.com/bookread2.php?book=760157.

5. Е.А. Коган, А.А. Юрченко, Теория вероятностей и математическая статистика : учебник / Е.А. Коган, А.А. Юрченко.— Москва : ИНФРА-М, 2019 .— 42 с. — 42 с. — [Электронный ресурс]. — Режим доступа: https://znanium.com/bookread2.php?book=971766.