Лабораторные работы

по математической статистике

(составитель: Е.И.Латышева)Лабораторная работа №1.

Закон распределения количественного

случайного признака X.

1. Цель работы: Определение закона распределения случайного признака  по выборочным данным. Определение эмпирической функции распределения и функции

плотности относительных частот.

2. Основные теоретические сведения.

Закон распределения количественного случайного признака X представляет собой

соответствие наблюденных значений признака  и их относительных частот ,

которые равны отношению частоты  события к объему выборки N.

Эмпирической функцией распределения называется функция, определяющая

относительную частоту события  для каждого значения :



Величина  принадлежит интервалу значений признака в данном статистическом материале:



Для определения закона распределения при объеме статистических данных, большем 10-15, производится группировка данных с шагом , равным



Тогда  - частота события середина интервала группировки. Значит, относительная частота этого события . Таким образом определена статистическая вероятность события X=.

Относительные частоты  позволяют для каждого  определить относительную частоту события  по формуле:

 

где  - количество интервалов группировки.

Функция плотности относительных частот вычисляется по формуле:



Она аналогична функции плотности распределения вероятности непрерывных случайных величин. Эмпирические функция распределения и плотность распределения вероятностей полностью задают закон распределения случайного признака .

3. Порядок выполнения работы и оформления отчета.

3.1. Оформить заголовок рабочего листа как на расчетном листе 1.

3.2. Ввести исходный массив данных.

3.3 С помощью функций пакета EXEL вычислить  и :





3.4. Заполнить первые три столбца таблицы: 

3.5. Заполнить столбец  следующим образом:

* выделить столбец ;
* ввести функцию (*массив данных; массив карманов*), где массив данных – исходная выборка; массив карманов – значения столбца  с первой по предпоследнюю строки;
* произвести вычисление одновременным нажатием клавиш Ctrl – Shift – Enter.

3.6. Заполнить столбцы , , , произведя вычисления по приведенным формулам.

3.7. Построить графики функций , , .

3.8. Результаты вычислений перенести в отчет по лабораторной работе (распечатать),

сформулировать выводы.

4. Контрольные вопросы.

4.1. Определить вероятность попадания признака  данной выборки в интервал ,

где ; .

4.2. Какому интервалу принадлежат все значения функции ?

4.3. Чему равна сумма всех относительных частот для признака ?

4.4. В чем вероятностный смысл функции ?

4.5. Перечислите основные свойства функции .

Лабораторная работа №2

Статистические оценки параметров распределения признака X.

1. Цель работы: Вычисление статистических оценок распределения случайного признака *X* по выборке из работы №1.
2. Основные теоретические сведения:

Статистическими оценками параметров генеральной совокупности являются параметры распределения выборочной совокупности. Несмещенной оценкой генеральной средней является выборочное среднее , которое вычисляется по формуле:

;

где  - число интервалов группировки,  - вычислены в работе №1.

Несмещенной оценкой генеральной дисперсии является исправленная выборочная

дисперсия , формула которой

.

Генеральное среднеквадратичное отклонение оценивается по стандарту выборки, который равен:



Одной из важных оценок эмпирического распределения является асимметрия относительно среднего значения , которая вычисляется по формуле:

.

Если асимметрия положительна, правый склон распределения  более пологий, если отрицательная - левый склон более пологий.

Относительную остроконечность или сглаженность распределения признака *X* по сравнению с нормальным законом характеризует эксцесс, который вычисляется по формуле:



1. Порядок выполнения работы и оформления отчета.
   1. Оформить заголовок рабочего листа как показано на расчетном листе 2.

3.2.По приведенным формулам вычислить значения всех описанных параметров, составить таблицу.

3.3.Проверить правильность расчетов с использованием стандартных функций EXEL:

среднее  - функция =СРЗНАЧ (массив данных)

исправленная дисперсия D - функция =ДИСП (массив данных)

стандарт *S* - функция =СТАНДОТКЛ (массив данных)

асимметрия  - функция =СКОС (массив данных)

эксцесс  - функция =ЭКСЦЕСС (массив данных)

Результаты вычислений перенести в отчет по лабораторной работе.

1. Контрольные вопросы.
   1. Каково соответствие между параметрами выборки и генеральной совокупности?
   2. Какие особенности распределения характеризуют асимметрия и эксцесс?
   3. Запишите формулы для вычисления всех статистических оценок.
   4. Какие из статистических оценок могут быть отрицательными?

Лабораторная работа № 3. 

Вычисление интервальных оценок математического ожидания

и среднеквадратичного отклонения.

1. Цель работы: Определение доверительных интервалов, в которых с заданной надежностью содержится математическое ожидание и среднеквадратичное отклонение генеральной совокупности в предположении, что признак X распределен

по нормальному закону.

1. Основные теоретические сведения.

Статистические оценки параметров выборки, такие как выборочное среднее , исправленная дисперсия , стандарт *s* являются эффективными, то есть их

математическое ожидание равно оцениваемому параметру генеральной совокупности.

Для нормально распределенных признаков можно вычислить интервал, симметричный относительно  и *s*, в который с заданной вероятностью попадает среднее значение признака генеральной совокупности и его среднеквадратичное отклонение.

Этот интервал называется доверительным. Вероятность попадания генерального среднего *a* при распределении по нормальному закону в доверительный интервал обозначается . При этом основное уравнение для определения величины интервала  имеет вид:

,

где  - бесконечно малая величина.

Обычно величина  задается на уровне 0,9; 0,95; 0,99. Пусть =0,95. Тогда вероятность попадания в интервал, симметричный относительно , равна удвоенной функции Лапласа:



Величина  табулирована в зависимости от объема выборки и заданной

надежности

Доверительный интервал для среднеквадратичного отклонения нормально распределенного признака рассчитывается с использованием таблиц параметра q, где  - надежность, *N* – объем выборки. При этом среднеквадратичное отклонение лежит в интервале



Величина q табулирована в зависимости от объема выборки и заданной надежности.

3.Порядок выполнения работы и оформления отчета.

3.1.В отчетной тетради по данным предыдущих работ вычислить доверительный интервал для генерального среднего, положив =0,95.

.

3.2. Вычислить доверительный интервал для среднеквадратичного отклонения для той

же величины .

1. Контрольные вопросы.

4.1. Какой доверительный интервал для генерального среднего шире - при 

или при ?

4.2. Как определить доверительный интервал для генеральной дисперсии?

Лабораторная работа № 4.

Гипотеза о нормальном законе распределения признака X, построение теоретических кривых нормального закона распределения. Проверка гипотезы по критерию Пирсона.

1. Цель работы: Построение теоретических кривых функции распределения и плотности распределения вероятностей признака X согласно выдвинутой гипотезе о нормальном законе распределения; сопоставление полученных теоретических кривых и эмпирических кривых, полученных в работе №1. Оценка справедливости гипотезы о нормальном законе распределения по критерию Пирсона.

1. Основные теоретические сведения.

Форма кривых  и , полученные в работе №1, а также величины асимметрии  и эксцесса , полученные в работе №2, позволяют выдвинуть гипотезу о том, что признак X распределен по нормальному закону с математическим ожиданием а, равным выборочному среднему , и среднеквадратичным отклонением , равным стандарту *s*. Значит, для каждого конкретного задания можно записать функцию плотности распределения вероятностей в виде:



Для построения теоретических кривых распределения можно воспользоваться табулированной функцией Гаусса. Для этого рассмотрим нормально распределенную случайную величину:



Так как эта величина представляет собой разность, ее математическое ожидание равно нулю:



Среднеквадратичное отклонение величины, нормированной на стандарт, равно 1:



Значит, для *U* функция плотности распределения вероятности имеет вид:



Эта функция табулирована, ее значения для различных значений аргумента приведены в приложении к пособию. Очевидно, что она является четной:

* *

Тогда, рассчитав для каждого значения  величину  и найдя значение функции

плотности распределения вероятности , можно вычислить величину 



Теоретическая кривая относительных частот  - вероятность попадания в каждый интервал группировки, определяется по формуле:

,

где *h* – величина шага группировки.

Значения теоретической функции распределения  могут быть вычислены по формуле:

, где m – число интервалов группировки.

При совмещении эмпирических и теоретических зависимостей:  и ;  и ;  и  очевидно некоторое расхождение. Необходимо оценить значение этого расхождения по критерию Пирсона . Для этого определим теоретические частоты  и сопоставим их с эмпирическими частотами:

,

где *N* – объем выборки.

Критерий Пирсона  равен:

,

где *m* – число интервалов группировки.

Критическая величина критерия Пирсона находится из таблицы по заданной доверительной вероятности и количеству интервалов группировки.

1. Порядок выполнения работы и оформления отчета.
   1. Оформить заголовок расчетного листа как на рис. 3.
   2. Составить заголовок расчетной таблицы, взять нужные данные из работ №1 и №2.
   3. Вычислить значения случайной величины  для каждого интервала группировки.
   4. Найти интервальные значения функции  по таблице или с помощью функции EXEL: =НОРМРАСП(; ЛОЖЬ)
   5. По вышеприведенным формулам вычислить ,  и  для каждого интервала группировки, найти необходимые суммы.
   6. Вычислить величину критерия 
   7. В отчет по лабораторной работе перенести все полученные расчетные данные, графики теоретических кривых распределения нанести на графики эмпирических кривых из работы №1.
   8. Определить по таблице критическую величину критерия Пирсона , записать вывод о справедливости выдвинутой гипотезы о нормальном законе распределения признака X.
2. Контрольные вопросы.
   1. Почему выдвинута гипотеза о нормальном законе распределения признака Х?
   2. Запишите формулу нормального закона распределения
   3. Что такое уровень значимости при определении критической величины ?
   4. Какого вида критическая область строится при проверке гипотезы о нормальном законе распределения?
   5. Чему равно число степеней свободы при определении ?

Лабораторная работа №5.

Система двух случайных признаков X и Y.

Вычисление ковариации COV(x,y).

1. Цель работы: Вычисление простейшей меры корреляционной зависимости двух случайных признаков X и Y - ковариации COV(X;Y).
2. Основные теоретические сведения:

Корреляционной называется такая зависимость между признаками, при которой изменение одного из признаков влечет за собой изменение среднего значения другого признака.

Цель корреляционного анализа состоит в том, чтобы установить наличие корреляционной связи, определить ее характер (линейная или какая – либо нелинейная связь), и оценить силу (тесноту) этой связи.

Простейшей мерой корреляционной связи любого типа является ковариация COV(X;Y), которая вычисляется по формуле:

,

где *N* – объем выборки,

  - текущие значения признаков,

  - выборочные средние значения признаков в данном статистическом материале.

1. Порядок выполнения работы и оформления отчета.
   1. Оформить заголовок рабочего листа как на расчетном листе 4. Ввести значения признаков X и Y.
   2. Провести первичный статистический анализ признаков *X*и *Y*, вычислив их средние значения, найти исправленную дисперсию и стандарт с помощью функций EXEL
   3. Обозначив разности ; , составить расчетную таблицу как на расчетном листе 4, и вычислить величину ковариации.
   4. Проверить правильность вычислений с помощью функции EXEL: =КОВАP(массив Х; массив Y)
   5. Внести полученные расчетные данные в отчет по лабораторной работе.
2. Контрольные вопросы.
   1. Какая статистическая зависимость называется корреляционной?
   2. Какой знак может иметь ковариация двух случайных признаков?