**Ковкель Никита Викторович, ПОИТ - 4/2**

**Лабораторная работа №1**

**Критерий согласия Пирсона**

**Вариант 12**

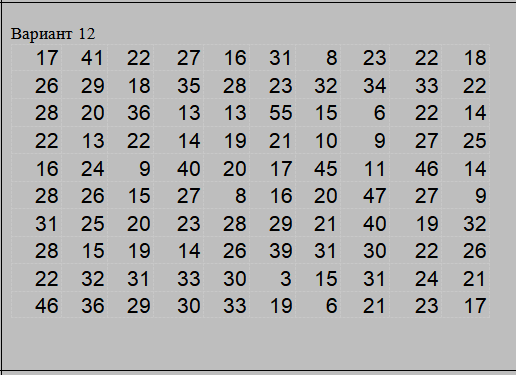
1. Составить интервальный статистический ряд. Величину интервалов округлить с точностью до 0,1 в большую сторону.

2. Найти эмпирическую функцию распределения и построить ее график.   
3. Построить гистограмму относительных частот. Можно ли предположить, что данная выборка взята из нормального распределения?

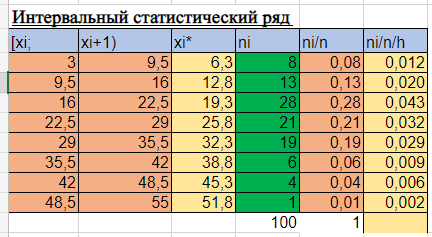
4. Определить выборочное среднее и несмещенную оценку дисперсии по сгруппированному статистическому ряду.

5. Записать предполагаемую плотность закона распределения.

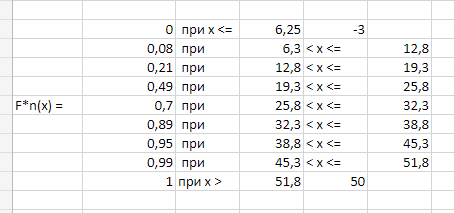
6. Проверить по критерию  Пирсона гипотезу о законе распределения. Уровень значимости принять равным α = 0,05.



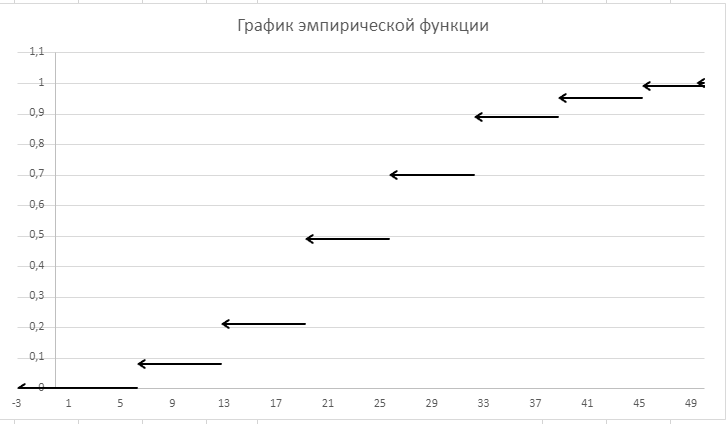
1. Объем выборки n = 100. Построим интервальный статистический ряд. Количество интервалов определим по формуле Стерджесса k ≈ 1 + log2(n) = 1+ log2(100) = 7,6439. Принимаем k = 8. Размах выборки W = xmax – xmin = 55 – 3 = 52. Длина каждого интервала будет h = W/k ≈ 52/8 = 6.5. Находим количество элементов выборки в каждом интервале.



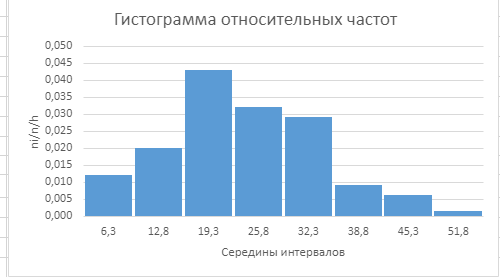
2. Для построения эмпирической функции распределения и гистограммы относительных частот дополним интервальный статистический ряд столбцами ni/n (относительные частоты нужны для построения эмпирической функции распределения) и ni/nh (высоты прямоугольников гистограммы). Запишем эмпирическую функцию распределения, накапливая относительные частоты ni/n (отметим, что при построении эмпирической функции распределения по интервальному статистическому ряду изменения ее значений (скачки) происходят в точках, соответствующих серединам интервалов группировки):



Построим график Fn\* (x) .

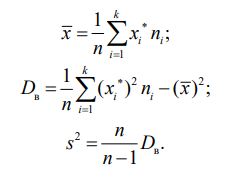


3. Построим гистограмму относительных частот, состоящую из прямоугольников шириной h = 6,5 и высотой ni/nh , По виду гистограммы можно выдвинуть гипотезу о том, что выборка взята из нормального распределения. Для проверки этой гипотезы по критерию согласия X2 Пирсона нужно рассчитать оценки параметров распределения по сгруппированному статистическому ряду.



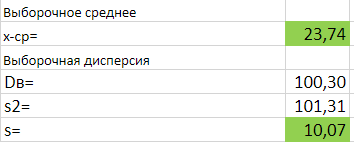
4. Рассчитаем оценки параметров предполагаемого нормального закона распределения по сгруппированному статистическому ряду. Данный закон содержит два параметра a и σ, которые имеют смысл математического ожидания и среднего квадратического отклонения СВ ξ: Mξ= a, D ξ = σ2.

В качестве оценок для математического ожидания a и дисперсии σ2 наблюдаемой случайной величины рассчитаем соответственно выборочное среднее и несмещенную оценку дисперсии s2, для вычисления s2 предварительно найдем выборочную дисперсию в Dв:

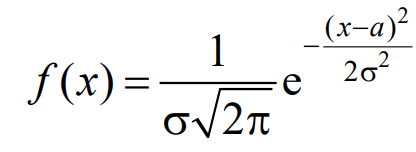


Используя интервальный статистический ряд, получим:

Тогда оценкой для среднего квадратического отклонения σ будет



5. Функция плотности нормального закона распределения имеет вид



Следовательно, выдвигаем гипотезу о том, что выборка взята из нормального распределения с плотностью

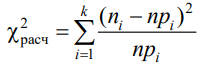
6. Проверим с помощью критерия согласия χ2 Пирсона гипотезу

H0: наблюдаемая СВ имеет нормальное распределение с параметрами

a =, σ = при альтернативе

: *наблюдаемая СВ имеет другое распределение.*

Для расчета статистики критерия Пирсона

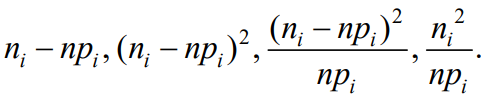


составим новую таблицу, содержащую следующие столбцы:

интервалы [xi-1; xi) − (при этом крайние интервалы должны быть расширены до −∞ и +∞ соответственно; а интервалы с количеством наблюдений меньше 5 объединяются с соседними);

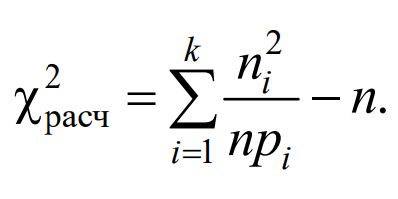
ni – эмпирическая частота наблюдения значений из интервала [xi-1; xi);

pi = P(ξ∈[xi-1; xi)) – теоретическая вероятность попадания СВ в интервал [xi-1; xi), в случае нормального распределения с параметрами a =, σ = эта вероятность рассчитывается как разность значений функции Лапласа:

npi – теоретическое значение соответствующей частоты,

а также столбцы со значениями

Последний столбец используется для контроля вычислений по формуле

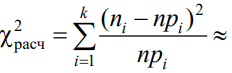


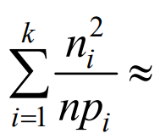
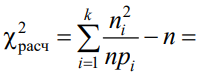
Все вычисления заносим в таблицу.



Суммирования значения в предпоследнем столбце, вычисляем

выборочное значение статистики критерия χ2 Пирсона:

3.326065. Сумма элементов последнего столбца равна

 103.3261. Это позволяет провести контроль вычислений: x`

103.3261 – 100 = 3.3261.

Определим критическое значение χ2крит = χ2α; k – r –1, где α = 0,05 – заданный уровень значимости; k = 8 – число интервалов после объединения малочисленных групп с соседними; r = 2, поскольку при расчете теоретических вероятностей pi использовались две полученные по выборке оценки x и s параметров нормального распределения. По таблице квантилей распределения χ2 получаем χ2крит = χ20.05; 2 = 9.487729.

Таким образом, χ2расч = 3.326065 < χ2крит = 9.487729, поэтому на уровне значимости α = 0,05 нет оснований отвергнуть гипотезу H0, согласно которой выборка взята из нормального распределения с параметрами a =, σ = .

Ответы на вопросы

1. Множество значений результатов наблюдений над одной и той же СВ ξ при одних и тех же условиях называется выборкой.

Количество проведенных наблюдений называется объемом выборки.

1. Эмпирической функцией распределения называется функция, определяющая для каждого значения х относительную частоту наблюдения значений, меньших х.
2. Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длины h , а высоты равны ni/nh .
3. 1
4. Под статистической гипотезой понимают всякое высказывание о виде или параметрах неизвестного распределения.
5. Статистическая гипотеза называется простой, если она полностью определяет функцию распределения. В противном случае гипотеза называется сложной
6. Одну из гипотез выделяют в качестве основной (или нулевой) H0 , а другую, являющуюся логическим отрицанием H0 , – в качестве конкурирующей (или альтернативной) гипотезы H
7. Правило, по которому принимается решение принять или отклонить проверяемую гипотезу, называется критерием проверки статистической гипотезы (статистическим критерием). При этом заранее выбирают допустимое значение ошибки вывода, которое называется уровнем значимости статистического критерия и обозначается α (это вероятность отвергнуть нулевую гипотезу, когда она верна).
8. Статистические критерии, с помощью которых проверяются гипотезы о виде распределения, называются критериями согласия или непараметрическими критериями.