Министерство науки и высшего образования Российской Федерации  
Южно-Уральский государственный университет  
(национальный исследовательский университет)  
Высшая школа электроники и компьютерных наук  
Кафедра «Информационно-измерительная техника»

ОТЧЕТ  
по практической работе №1  
на тему «Эмпирическая функция распределения случайной величины»  
по дисциплине «Теория вероятностей и математическая статистика»

Проверил, доцент  
 / А.П. Лапин /  
«\_\_\_\_» 2023 г.  
  
Автор работы  
студент группы КЭ-215  
 / Д.А. Суринский /  
«\_\_\_\_» 2023 г.

ОГЛАВЛЕНИЕ

[ВВЕДЕНИЕ 3](#_Toc127119485)

[1. ИСХОДНЫЕ ДАННЫЕ 5](#_Toc127119486)

[2. АЛГОРИТМ ОБРАБОТКИ ДАННЫХ 6](#_Toc127119487)

[3. ПРИМЕНЕНИЕ АЛГОРИТМА ДЛЯ ОБРАБОТКИ ДАННЫХ 8](#_Toc127119488)

[заключение 15](#_Toc127119489)

[библиографический список 16](#_Toc127119490)

ВВЕДЕНИЕ

Математическая статистика – раздел математики, опирающийся на теорию вероятностей. В ней рассматриваются приближённые методы нахождения закономерностей поведения и числовых характеристик случайных величин по результатам экспериментов и наблюдений. [1, стр. 142]

Эмпирической функцией распределения выборки *Xn* называется функция *Fn(x)*, определяющая для любого *x ∈ R* относительную частоту события *(Xn< x)*. [1, стр. 147]

Случайная величина – числовая величина, которая в результате эксперимента принимает некоторое, заранее неизвестное значение. Случайные величины обозначаются заглавными буквами конца латинского алфавита: X, Y, Z, … [1, стр. 71]

Генеральная совокупность – вся совокупность реализации случайной величины, все возможные наблюдения некоторого показателя, все возможные исходы некоторого испытания. [1, стр.  143]

Выборка – часть генеральной совокупности, а именно конечное подмножество значений случайной величины из множества элементов генеральной совокупности. [1, стр.  143]

Объём выборки – количество содержащихся в ней случайных величин. [1, стр.  143]

Размах выборки – разность между наибольшим и наименьшим значениями случайной величины. [1, стр.  143]

Относительная частота события – отношения числа испытаний, в которых событие уже появилось, к общему числу проведённых испытаний. [1, стр.  8]

Полигон частот ‒ один из способов графического представления плотности вероятности случайной величины. Представляет собой ломаную, соединяющую точки, соответствующие срединным значениям интервалов группировки и частотам этих интервалов.

Гистограмма распределения ‒ наглядное представление функции плотности вероятности некоторой случайной величины, построенное по выборке.

Эмпирическая функция распределения ‒ функция, которая определяет для каждого значения *x* частоту событий *(X < x)* и предназначена для оценки теоретической функции распределения генеральной совокупности.

В данной практической работе необходимо построить эмпирическую функцию распределения для случайной выборки с помощью специально разработанного алгоритма построения и сделать выводы об особенностях распределения случайной величины.

Целью работы является выполнение построения эмпирических распределений случайной величины, для выборки из 100 элементов.

Данная работа выполнена в соответствии с СТО ЮУрГУ 04-2008. [2].

# ИСХОДНЫЕ ДАННЫЕ

Таблица 1 – Исходная выборка

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 127.52  135.94  153.89  148.23  146.00  111.43  147.00  137.17  156.32  136.86  143.39  136.20  135.33  139.15  138.42  147.89  139.40  124.90  152.45  157.44 | 148.64  126.48  136.83  161.33  139.75  147.79  143.66  130.72  131.66  162.61  155.01  158.43  144.19  156.85  151.98  153.19  123.19  161.01  142.75  159.31 | 149.92  147.26  149.57  147.05  140.21  132.59  130.70  145.60  141.37  120.72  111.55  130.81  132.34  143.38  145.05  155.10  132.40  137.83  151.94  134.06 | 155.22  141.81  121.39  150.38  169.52  137.93  146.24  146.63  139.44  139.35  123.25  142.35  148.99  140.66  139.68  124.46  136.49  148.43  141.50  148.34 | 135.02  145.72  145.08  115.97  124.47  160.25  130.23  143.04  152.99  133.97  142.36  126.57  156.54  142.23  135.26  145.57  148.40  131.38  170.18  143.29 |

# АЛГОРИТМ ОБРАБОТКИ ДАННЫХ

При построении эмпирической функции распределения случайной величины используется следующий порядок (алгоритм) обработки данных:

1. Выборка (*x1*, *x2*, … , *xN*) преобразуется в вариационный ряд, т. е. ряд данных, расположенных в порядке возрастания;
2. Определяется размах выборки *R* по формуле:

, (1)

где *xmax* – максимальное значение выборки;  
*xmin* – минимальное значение выборки.

1. Весь размах выборки делится на *k* равных интервалов, определяемые с помощью формулы Стерджесса:

; (2)

 (3)

где *N* – число элементов выборки.

1. Находится длина каждого из *k* интервала:

. (4)

1. Находятся границы интервалов и их середины:

 (5)

, (6)

где граница интервала (*i* + 1), *i* = .

, (7)

где середина интервала.

1. Определяется количество элементов выборки, попавших в каждый интервал (частота попадания в интервал) – *m*i, где *i* – номер интервала, *i* = .
2. Находятся относительные частоты попадания:

причем  (8)

где ‒ относительная частота для -го интервала, .

Также находятся относительные накопленные частоты:

 (9)

где ‒ относительная накопленная частота -го интервала, .

1. Графическое изображение эмпирических распределений в виде полигона частот, гистограммы распределения, эмпирической функции распределения.

# ПРИМЕНЕНИЕ АЛГОРИТМА ДЛЯ ОБРАБОТКИ ДАННЫХ

1. Расположим исходные данные из таблицы 1 в вариационный ряд ( в порядке возрастания):

Таблица – Вариационный ряд

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 111,43  111,45  115,97  120,72  121,39  123,19  123,25  124,46  124,47  124,90  126,48  126,57  127,52  130,23  130,70  130,72  130,81  131,38  131,66  132,34 | 132,40  132,59  133,97  134,06  135,02  135,26  135,33  135,94  136,20  136,49  136,83  136,86  137,17  137,83  137,93  138,42  139,15  139,35  139,40  139,44 | 139,68  139,75  140,21  140,66  141,37  141,50  141,81  142,23  142,35  142,36  142,75  143,04  143,29  143,38  143,39  143,66  144,19  145,05  145,08  145,57 | 145,60  145,72  146,00  146,24  146,63  147,00  147,05  147,26  147,79  147,89  148,23  148,34  148,40  148,43  148,64  148,99  149,57  149,92  150,38  151,94 | 151,98  152,45  152,99  153,19  153,89  155,01  155,10  155,22  156,32  156,54  156,85  157,44  158,43  159,31  160,25  161,01  161,33  162,61  169,52  170,18 |

1. Определим размах выборки *R* по формуле (1):

Исходя из вариационного ряда из п.1 следует, что максимальная величина выборки xmax равна 170,18. Минимальная величина выборки xmin равна 111,43. Следовательно, размах выборки R из формулы (1):

*R* = 170,18 – 111,43;

*R* = 58,75.

1. Найдем количество равных интервалов *k* по формуле (2)

*N* = 100, *k* = 3 + [3,2 × 2];

*k* = 9.

1. Определим длину интервала *h* по формуле (4):

*h* = 

*h =* 6,53.

1. Найдем границы интервалов и их середины по формулам (5) и (7):

Таблица – Границы и середины интервалов

|  |  |  |
| --- | --- | --- |
| № интервала | Границы интервалов | Середина интервалов |
| 1 | 111,43  117,96 | 114,70 |
| 2 | 117,96  124,49 | 121,23 |
| 3 | 124,49  131,02 | 127,76 |
| 4 | 131,02  137,55 | 134,29 |
| 5 | 137,55  144,08 | 140,82 |
| 6 | 144,08  150,61 | 147,35 |
| 7 | 150,61  157,14 | 153,88 |

Продолжение таблицы 3

|  |  |  |
| --- | --- | --- |
| 8 | 157,14  163,67 | 160,41 |
| 9 | 163,67  170,20 | 166,94 |

1. Определим количество элементов выборки, попадающих в каждый интервал – *mi*, где *i* – номер интервала, .

Таблица – Частота попадания элементов выборки в интервал

|  |  |  |  |
| --- | --- | --- | --- |
| № интервала | Границы интервалов | Середина интервалов | Частота в интервале |
| 1 | 111,43 117,96 | 114,70 | 3 |
| 2 | 117,96 124,49 | 121,23 | 6 |
| 3 | 124,49 131,02 | 127,76 | 8 |
| 4 | 131,02 137,55 | 134,29 | 16 |
| 5 | 137,55 144,08 | 140,82 | 23 |
| 6 | 144,08 150,61 | 147,35 | 23 |
| 7 | 150,61 157,14 | 153,88 | 12 |
| 8 | 157,14 163,67 | 160,41 | 7 |
| 9 | 163,67 170,20 | 166,94 | 2 |

1. Найдем относительные частоты и относительные накопленные частоты по формулам (8) и (9):

Таблица – Относительные частоты и относительные накопленные частоты

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № интервала | Границы интервалов | Середина интервалов | Частота в интервале | Отн. частота | Накоп. частота | Отн. накоп. частота |
| 1 | 111,43 117,96 | 114,70 | 3 | 0,03 | 3 | 0,03 |
| 2 | 117,96 124,49 | 121,23 | 6 | 0,06 | 9 | 0,09 |
| 3 | 124,49 131,02 | 127,76 | 8 | 0,08 | 17 | 0,17 |
| 4 | 131,02 137,55 | 134,29 | 16 | 0,16 | 33 | 0,33 |
| 5 | 137,55 144,08 | 140,82 | 23 | 0,23 | 56 | 0,56 |
| 6 | 144,08 150,61 | 147,35 | 23 | 0,23 | 79 | 0,79 |
| 7 | 150,61 157,14 | 153,88 | 12 | 0,12 | 91 | 0,91 |
| 8 | 157,14 163,67 | 160,41 | 7 | 0,07 | 98 | 0,98 |
| 9 | 163,67 170,20 | 166,94 | 2 | 0,02 | 100 | 1 |

1. Изобразим эмпирические распределения в виде полигона частот, гистограммы распределения и эмпирической функции распределения:

Для построения полигона частот следует по оси абсцисс отложить рассчитанные границы интервалов. В серединах интервалов строятся ординаты, пропорциональные частотам или относительным частотам (рисунок 1).

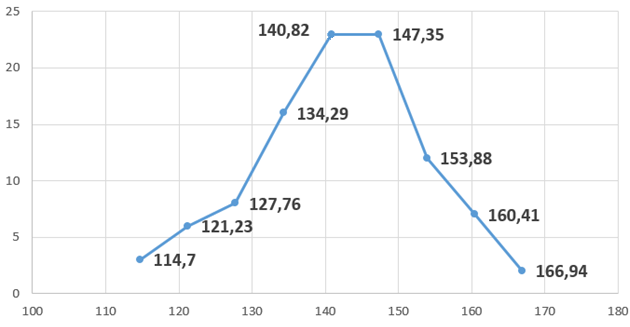


Рисунок – Полигон частот СВ

Чтобы построить гистограмму распределения, нужно по оси абсцисс отложить границы интервалов, далее на каждом из интервалов, как на основании, строится прямоугольник, площадь которого равна относительной частоте попадания СВ в данный интервал (рисунок 2).

Таблица 6 – Относительная частота попадания

|  |  |  |
| --- | --- | --- |
| Границы интервалов | Относительная частота попадания (*pi*) | *pi* /*h (h =* 6,53*)* |
| 111,43 117,96 | 0,03 | 0,004594 |
| 117,96 124,49 | 0,06 | 0,009188 |
| 124,49 131,02 | 0,08 | 0,012251 |
| 131,02 137,55 | 0,16 | 0,024502 |
| 137,55 144,08 | 0,23 | 0,035222 |
| 144,08 150,61 | 0,23 | 0,035222 |

Продолжение таблицы 6

|  |  |  |
| --- | --- | --- |
| 150,61 157,14 | 0,12 | 0,018377 |
| 157,14 163,67 | 0,07 | 0,01072 |
| 163,67 170,20 | 0,02 | 0,003063 |

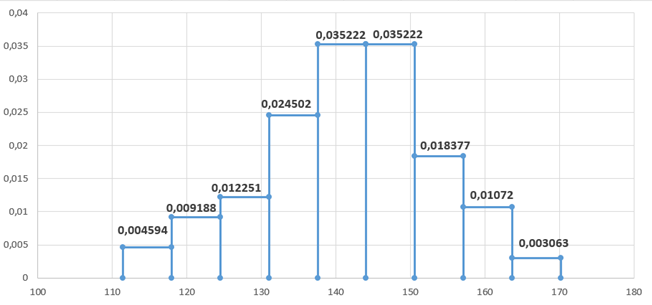


Рисунок – Гистограмма распределения СВ

Для построения эмпирической функции распределения нужно по оси абсцисс отложить границы интервалов, а по оси ординат отложить относительные накопленные частоты (рисунок 3).

Таблица - Относительные накопленные частоты

|  |  |
| --- | --- |
| Границы интервалов | Относительная накопленная частота |
| 111,43 117,96 | 0,03 |
| 117,96 124,49 | 0,09 |
| 124,49 131,02 | 0,17 |
| 131,02 137,55 | 0,33 |
| 137,55 144,08 | 0,56 |
| 144,08 150,61 | 0,79 |
| 150,61 157,14 | 0,91 |
| 157,14 163,67 | 0,98 |
| 163,67 170,20 | 1 |

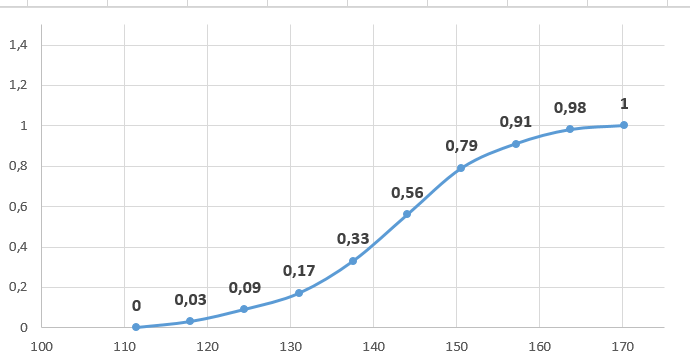


Рисунок - Эмпирическая функция распределения СВ

заключение

Результаты обработки данных непрерывной случайной величины дают полную её характеристику и приводят к выводам об особенностях распределения этой величины.

По результатам проделанной работы справедливы следующие утверждения:

1. Наибольшая плотность вероятности сосредоточена на пятом и шестом участках. [137,55; 144,08] и [144,08; 150,61] соответственно;
2. Основная часть выборки сконцентрирована в середине размаха из-за чего можно сделать вывод что она не симметрична;
3. Размах выборки составляет 58,75 на границах [111,43; 170,20];
4. Наименьшее количество элементов выборки находится на девятом интервале, границами которого является отрезок [163,67; 170,20].

Эмпирическая функция распределения, полученная в итоге работы, показывает характер распределения величины на всем размахе выборки.

библиографический список

1. Горлач, Б. А. Теория вероятностей и математическая статистика: учебно-методическое пособие / Б. А. Горлач. – Санкт-Петербург: Лань, 2013. – 320 с.
2. СТО ЮУрГУ 04-2008. Стандарт организации. Курсовое и дипломное проектирование. Общие требования к содержанию и оформлению / сост. Т.И. Парубочая, Н.В. Сырейщикова, В.И. Гузеев, Л.В. Винокурова. – Челябинск: Изд-во ЮУрГУ, 2008. – 56 с.