Задание

Для данной случайной выборки объемом *п* = 100:

1. составить вариационный ряд;

2. составить статистический ряд;

3. построить гистограмму частот;

4. вычислить статистическое среднее и статистическую дисперсию;

5. составить статистическую функцию распределения *F(x);*

6. построить группированную выборку и с ее помощью:

а) вычислить статистическое среднее и статистическую дисперсию;

б) составить статистическую функцию распределения *F* (*x*) и построить ее график;

7. найти доверительный интервал для математического ожидания и дисперсии;

8. с помощью критерия Пирсона $x^{2}$проверить гипотезу о нормальном законе распределения данной выборки;

9. составить уравнение линии регрессии (кривой, «сглаживающей» гистограмму частот) и построить ее график.

Дана случайная выборка n=100

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -1,752 | -0,206 | -0,266 | -0,578 | 0,035 |
| -0,291 | 0,092 | 0,901 | 0,439 | 0,106 |
| -0,093 | -1.222 | -1,433 | -0,852 | 0,199 |
| -0.450 | 0,065 | 0,327 | 0,489 | -1,990 |
| 0,512 | 0,183 | 0,248 | 0,675 | 0,710 |
| -0,702 | -0.811 | -0,401 | -1,210 | 0,340 |
| 0,284 | -1.019 | 0,344 | 0,131 | -0,594 |
| -0,509 | 1,453 | 0,441 | -1,202 | -1,527 |
| -1.776 | 0,759 | 0,824 | 0,894 | 0,362 |
| -0.044 | 0,287 | 1,385 | -0,780 | -0,570 |
| 0,263 | -0,669 | -0,329 | -0,195 | -1,309 |
| 0,986 | 0,392 | 0,085 | -0,927 | 1,531 |
| -0,441 | -о,зз7 | 0,130 | -1,582 | -1,008 |
| -0.866 | 0,369 | -0,244 | 0,075 | 0,763 |
| -1.215 | 1,694 | -0,882 | 1,600 | 0,788 |
| -0,475 | -0,985 | 0,472 | 2,904 | -0,679 |
| 1,2 | -1,063 | 0,039 | 0,149 | -0,824 |
| -0,498 | 0,033 | 1,42 | 1,210 | -0,372 |
| -0,743 | 0,597 | -1,033 | -0,838 | 0,049 |
| 0,779 | -1.601 | 1,807 | 0,278 | 1,320 |

Таблица 1

Решение:

Из данной выборки определяем максимальную варианту  и минимальную варианту 

=2,904, =-1,99

Расположив варианты в порядке возрастания, начиная с  получим вариаионный ряд:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -1,99 | -0,852 | -0,329 | 0,183 | 0,71 |
| -1,776 | -0,838 | -0,291 | 0,199 | 0,759 |
| -1,752 | -0,824 | -0,266 | 0,248 | 0,763 |
| -1,601 | -0,811 | -0,244 | 0,263 | 0,779 |
| -1,582 | -0,78 | -0,206 | 0,278 | 0,788 |
| -1,527 | -0,743 | -0,195 | 0,284 | 0,824 |
| -1,433 | -0,702 | -0,093 | 0,287 | 0,894 |
| -1,309 | -0,679 | -0,044 | 0,327 | 0,901 |
| -1,222 | -0,669 | 0,033 | 0,34 | 0,986 |
| -1,215 | -0,594 | 0,035 | 0,344 | 1,2 |
| -1,21 | -0,578 | 0,039 | 0,362 | 1,21 |
| -1,202 | -0,57 | 0,049 | 0,369 | 1,32 |
| -1,063 | -0,509 | 0,065 | 0,392 | 1,385 |
| -1,033 | -0,498 | 0,075 | 0,439 | 1,42 |
| -1,019 | -0,475 | 0,085 | 0,441 | 1,453 |
| -1,008 | -0,45 | 0,092 | 0,472 | 1,531 |
| -0,985 | -0,441 | 0,106 | 0,489 | 1,6 |
| -0,927 | -0,401 | 0,130 | 0,512 | 1,694 |
| -0,882 | -0,372 | 0,131 | 0,597 | 1,807 |
| -0,866 | -0,337 | 0,149 | 0,675 | 2,904 |

Таблица 2

Для построения статистического ряда разобьем вариационный ряд на конечное число интервалов (зарядов). Длину интервала определим по формуле Стэрджеса:

∆=(-)/(1+3,32lgn)

где *п* - объем данной случайной выборки, ∆ - длина интервала;

∆ = (2,904 +1,99)/(7,64) = 0,640

Примем ∆ = 0,7. От  отступим влево на 0,01. Величину 0,01 выбрали так, чтобы округлить значение левого конца интервала. Откладываем вправо интервалы длиной ∆ = 0,7 до тех пор, пока не по­кроется вся выборка, и считаем число вариант, попавших на каждый интер­вал. В вариационном ряду удобно отделить варианты одного разряда от вари­ант другого разряда чертой. По результатам разбиения составим таблицу 1.

|  |  |  |
| --- | --- | --- |
| Интервалы I, (,) | Число вариант  | Частоты  |
| (-2;-1,31) | 8 | 0,08 |
| (-1,31;-0,588) | 23 | 0,23 |
| (-0,588;0,116) | 27 | 0,27 |
| (0,116;0,79) | 27 | 0,27 |
| (0,79;1,46) | 10 | 0,1 |
| (1,46;1,81) | 4 | 0,04 |
| (1,81;2,91) | 1 | 0,01 |

Таблица 3

Эта таблица называется статистическим рядом. В таблице 3 *тi —* это число вариант, попавших в i-й интервал; *ai,ai+x —* соответственно начало и конец i -го интервала.

3. Для построения графического изображения статистического ряда (гистограммы) отложим на оси *Ох* интервалы из таблицы 1, и на каждом *i* -м

интервале построим прямоугольник с высотой *уi: уi* =*mi/∆n,* тогда

y1 = 0,114; у2=0,328; у3 = 0,385; у4 = 0,385; у5 = 0,142; у6= 0,057;

у7 = 0,014;

На рисунке 1 представлено графическое изображение статистического ряда (гистограмма). Для удобства обозначения гистограммы ее основание должно быть в 1,5 - 2 раза больше высоты.

Рис.1

4. Найдем статистическое среднее  и статистичекую дисперсию :

-0,06993



— исправленная статистическая дисперсия;

— выборочная дисперсия.

5. Составим статистическую функцию распределения. Ее значения, составленные по данной случаной выборке, занесены в таблицу 4.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x1 | F(x1) | x1 | F(x1) | x1 | F(x1) | x1 | F(x1) | x1 | F(x1) |
| -1,99 | 0 | -0,852 | 0,2 | -0,329 | 0,4 | 0,183 | 0,6 | 0,71 | 0,8 |
| -1,77 | 0,01 | -0,838 | 0,21 | -0,291 | 0,41 | 0,199 | 0,61 | 0,759 | 0,81 |
| -1,75 | 0,02 | -0,824 | 0,22 | -0,266 | 0,42 | 0,248 | 0,62 | 0,763 | 0,82 |
| -1,60 | 0,03 | -0,811 | 0,23 | -0,244 | 0,43 | 0,263 | 0,63 | 0,779 | 0,83 |
| -1,58 | 0,04 | -0,78 | 0,24 | -0,206 | 0,44 | 0,278 | 0,64 | 0,788 | 0,84 |
| -1,52 | 0,05 | -0,743 | 0,25 | -0,195 | 0,45 | 0,284 | 0,65 | 0,824 | 0,85 |
| -1,43 | 0,06 | -0,702 | 0,26 | -0,093 | 0,46 | 0,287 | 0,66 | 0,894 | 0,86 |
| -1,30 | 0,07 | -0,679 | 0,27 | -0,044 | 0,47 | 0,327 | 0,67 | 0,901 | 0,87 |
| -1,22 | 0,08 | -0,669 | 0,28 | 0,033 | 0,48 | 0,34 | 0,68 | 0,986 | 0,88 |
| -1,215 | 0,09 | -0,594 | 0,29 | 0,035 | 0,49 | 0,344 | 0,69 | 1,2 | 0,89 |
| -1,21 | 0,1 | -0,578 | 0,3 | 0,039 | 0,5 | 0,362 | 0,7 | 1,21 | 0,9 |
| -1,202 | 0,11 | -0,57 | 0,31 | 0,049 | 0,51 | 0,369 | 0,71 | 1,32 | 0,91 |
| -1,063 | 0,12 | -0,509 | 0,32 | 0,065 | 0,52 | 0,392 | 0,72 | 1,385 | 0,92 |
| -1,033 | 0,13 | -0,498 | 0,33 | 0,075 | 0,53 | 0,439 | 0,73 | 1,42 | 0,93 |
| -1,019 | 0,14 | -0,475 | 0,34 | 0,085 | 0,54 | 0,441 | 0,74 | 1,453 | 0,94 |
| -1,008 | 0,15 | -0,45 | 0,35 | 0,092 | 0,55 | 0,472 | 0,75 | 1,531 | 0,95 |
| -0,985 | 0,16 | -0,441 | 0,36 | 0,106 | 0,56 | 0,489 | 0,76 | 1,6 | 0,96 |
| -0,927 | 0,17 | -0,401 | 0,37 | 0,130 | 0,57 | 0,512 | 0,77 | 1,694 | 0,97 |
| -0,882 | 0,18 | -0,372 | 0,38 | 0,131 | 0,58 | 0,597 | 0,78 | 1,807 | 0,98 |
| -0,866 | 0,19 | -0,337 | 0,39 | 0,149 | 0,59 | 0,675 | 0,79 | 2,904 | 0,99 |

Таблица 4

6. Группированная выборка.

Так как объем данной случайной выборки велик (n = 100), то удобнее пользоваться группированной выборкой, для построения которой в статистическом ряде (табл. 1) заменим каждый интервал его представителем. В качестве представителя i-го интервала возьмем его середину xi\*(табл.5).

|  |  |  |
| --- | --- | --- |
| xi\* | Число вариант  | Частоты  |
| -1,62 | 8 | 0,08 |
| -0,825 | 23 | 0,23 |
| -0,119 | 27 | 0,27 |
| 0,508 | 27 | 0,27 |
| 1,28 | 10 | 0,1 |
| 1,658 | 4 | 0,04 |
| 2,904 | 1 | 0,01 |

Таблица 5

1. Вычислим статистическое среднее и статистическую дисперсию по группированной выборке:



=0,079

*Здесь к-* число интервалов, на которые разбита выборка. С учетом поправок

Шеппарда 0,0448, статистическое среднее =-0,09, статистическая дисперсия =0,0448 .

1. Составим статистическую функцию распределения *F\*(x):*

Построим график статистической функции распределения , составленной по сгруппированной выборке

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  | 0,97 | 1 |  |
|  |  |  |  |  |  | 0,895 |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | 0,705 |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | 0,43 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  | 0,185 |  |  |  |  |  |  |
|  | 0,035 |  |  |  |  |  |  |  |  |
| 0 |  |  |  |  |  |  |  |  |  |
|  | -1,62 | -0,825 | -0,119 |  | 0,508 | 1,28 | 1,658 | 2,904 | X |

7. Найдем доверительный интервал для математического ожидания . Воспользуемся формулой

,

где  - статистика, распределенная асимптотически по нормальному закону. При можно считать, что . Значение статистики  определяем из условия

.

Пользуясь таблицей значений функции Лапласа для доверительной вероятности  находим

.

Значения  и  вычислены ранее по случайной выборке.

Подставляя указанные значения в формулу, получаем искомый доверительный интервал для математического ожидания

.

Найдем доверительный интервал для дисперсии . Воспользуемся формулой

,

Ранее было найдено , .

Подставляя указанные значения в формулу, получаем искомый доверительный интервал для дисперсии

.

8. Предположим, что данная случайная выборка распределена по нормальному закону с параметрами  и . Проверим гипотезу о нормальном законе распределения данной случайной выборки при помощи критерия Пирсона:

.

Для расчета попадания случайной величины  в интервал  используем функцию Лапласа в соответствии со свойствами нормального распределения:

.











Составим расчетную таблицу

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Интервалы | Частоты | Вероятности | Теоретические частоты  |  |  |
| (-3.8;-2,95) | 1 | 0,0025 | 0,25 2,42 |  |  |
| (-2,95;-2,1) | 0 | 0,0217 | 2,17 | 2,0164 | 0,83 |
| (-2,1;-1,25) | 12 | 0,1027 | 10,27 | 2,9929 | 0,29 |
| (-1,25;-0,4) | 26 | 0,2474 | 24,74 | 1,5876 | 0,06 |
| (-0,4;0,45) | 32 | 0,317 | 31,7 | 0,09 | 0,003 |
| (0,45;1,3) | 18 | 0,2167 | 21,67 | 13,4689 | 0,62 |
| (1,3;2,15) | 10 | 0,0758 | 7,58 9,02 | 3,9204 | 0,43 |
| (2,15;3) | 1 | 0,0144 | 1,44 |  |  |
|  |  |  |  |  |  |
| сумма |  |  |  |  | ***2,233*** |

Таким образом, получаем опытное значение критерия 

Найдем пороговое значение критерия Пирсона . Для этого вычислим число степеней свободы .

По найденному числу степеней свободы и заданному уровню значимости  находим пороговое значение критерия Пирсона .Так как , то гипотеза о нормальном законе распределения данной случайной выборки принимается.

9. Составим уравнение линии регрессии, т.е. уравнение кривой, «сглаживающей» гистограмму частот. Это уравнение будем искать в виде

.

Прологарифмируем последнее выражение по основанию :



Обозначим

, , , .

В этих обозначениях уравнение регрессии примет вид

. Так как объем выборки велик , то для нахождения коэффициентов ,  и  воспользуемся группированной выборкой, дополнив ее значениями  и 

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   |  |  |  |  |
|   | -3,375 | 1 | 0,01 | -4,61 |
|   | -2,525 | 1 | 0,01 | -4,61 |
|   | -1,675 | 11 | 0,13 | -2,04 |
|   | -0,825 | 26 | 0,31 | -1,17 |
|   | 0,025 | 32 | 0,38 | -0,97 |
|   | 0,875 | 18 | 0,21 | -1,56 |
|   | 1,725 | 10 | 0,12 | -2,12 |
|   | 2,575 | 1 | 0,01 | -4,61 |

Коэффициенты ,  и  уравнения находим, решая следующую систему линейных алгебраических уравнений



Составим расчетную таблицу

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|   |  |  |  |  |  |  |  |
|   | -3,375 | -4,61 | 15,55875 | 11,39063 | -52,51078125 | -38,44335938 | 129,7463379 |
|   | -2,525 | -4,61 | 11,64025 | 6,375625 | -29,39163125 | -16,09845313 | 40,64859414 |
|   | -1,675 | -2,04 | 3,417 | 2,805625 | -5,723475 | -4,699421875 | 7,871531641 |
|   | -0,825 | -1,17 | 0,96525 | 0,680625 | -0,79633125 | -0,561515625 | 0,463250391 |
|   | 0,025 | -0,97 | -0,02425 | 0,000625 | -0,00060625 | 1,5625E-05 | 0,0000004 |
|   | 0,875 | -1,56 | -1,365 | 0,765625 | -1,194375 | 0,669921875 | 0,586181641 |
|   | 1,725 | -2,12 | -3,657 | 2,975625 | -6,308325 | 5,132953125 | 8,854344141 |
|   | 2,575 | -4,61 | -11,8708 | 6,630625 | -30,56718125 | 17,07385938 | 43,96518789 |
| сумма | -3,2 | -21,69 | 14,66425 | 31,625 | -126,4927063 | -36,926 | 232,1354281 |

Таким образом, получено, что

, , , , , , .

В результате получаем систему уравнений



Решая систему методом Гаусса, имеем

|  |  |  |  |
| --- | --- | --- | --- |
| 232,14 | -36,926 | 31,625 | -126,49 |
| -36,926 | 31,625 | -3,2 | 14,66425 |
| 31,625 | -3,2 | 8 | -21,69 |
|  |  |  |  |
| 232,14 | -36,926 | 31,625 | -126,49 |
| 0 | 25,75126227 | 1,830519 | -5,45624 |
| 0 | 1,830519299 | 3,691649 | -4,45796 |
|  |  |  |  |
|  |  |  |  |
| 232,14 | -36,926 | 31,625 | -126,49 |
| 0 | 25,75126227 | 1,830519 | -5,45624 |
| 0 | 0 | 3,561527 | -4,0701 |
|  |  |  |  |
| 232,14 | -36,926 | 31,625 | -126,49 |
| 0 | 25,75126227 | 1,830519 | -5,45624 |
| 0 | 0 | 1 | -1,1428 |
|  |  |  |  |
|  |  |  |  |
| 232,14 | -36,926 | 0 | -90,349 |
| 0 | 25,75126227 | 0 | -3,36432 |
| 0 | 0 | 1 | -1,1428 |
|  |  |  |  |
|  |  |  |  |
| 232,14 | 0 | 0 | -95,1733 |
| 0 | 1 | 0 | -0,13065 |
| 0 | 0 | 1 | -1,1428 |
|  |  |  |  |
|  |  |  |  |
| 1 | 0 | 0 | -0,40998 |
| 0 | 1 | 0 | -0,13065 |
| 0 | 0 | 1 | -1,1428 |

Т.е. ,  и .

В результате находим уравнение параболической регрессии

.

Чтобы вернуться к исходному уравнению регрессии, вычислим коэффициенты ,  и , решив систему уравнений



Таким образом, получаем уравнение регрессии

.

Изобразим график полученной функции