КОНТРОЛЬНЫЕ ЗАДАЧИ ТИПОВОГО РАСЧЕТА ПО МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

ЗАДАЧА 10.

По выборке одномерной случайной величины с номером, приведенном в индивидуальном задании студента для типового расчета:

- получить вариационный ряд;
- построить на масштабно-координатной бумаге формата А4 график эмпирической функции распределения $F^*(x)$;
- построить гистограмму равноинтервальным способом;
- построить гистограмму равновероятностным способом;
- вычислить точечные оценки математического ожидания и дисперсии;
- вычислить интервальные оценки математического ожидания и дисперсии ($\gamma = 0,95$);
- выдвинуть гипотезу о законе распределения случайной величины и проверить ее при помощи критерия согласия $\chi 2$ и критерия Колмогорова ($\alpha = 0,05$). График гипотетической функции распределения $F0(x)$ построить совместно с графиком $F^*(x)$ в той же системе координат и на том же листе.

Одномерная выборка №21: 3.56 2.35 6.34 5.97 3.24 5.80 0.92 6.59 1.57 4.31 4.97 6.33 4.26 6.40 5.76 1.34 9.49 3.55 6.77 6.26 4.61 2.81 1.87 7.11 2.37 -0.31 4.30 4.51 3.34 5.56 2.09 3.42 5.90 4.90 6.02 5.72 3.96 4.27 6.19 1.11 3.25 9.77 3.13 4.00 7.28 3.88 6.24 -0.17 5.98 4.93 4.76 3.83 7.31 4.59 7.53 3.98 5.86 5.06 3.56 3.83 4.39 4.22 3.01 4.04 4.12 5.38 2.46 6.41 4.20 5.01 5.83 4.90 2.22 2.78 3.20 3.17 1.42 3.91 1.33 3.45 6.94 1.48 4.61 2.63 6.48 7.19 4.27 7.31 8.25 4.15 1.45 4.14 7.51 1.63 4.23 5.54 3.99 10.27 3.24 3.27

1. Вариационный ряд:

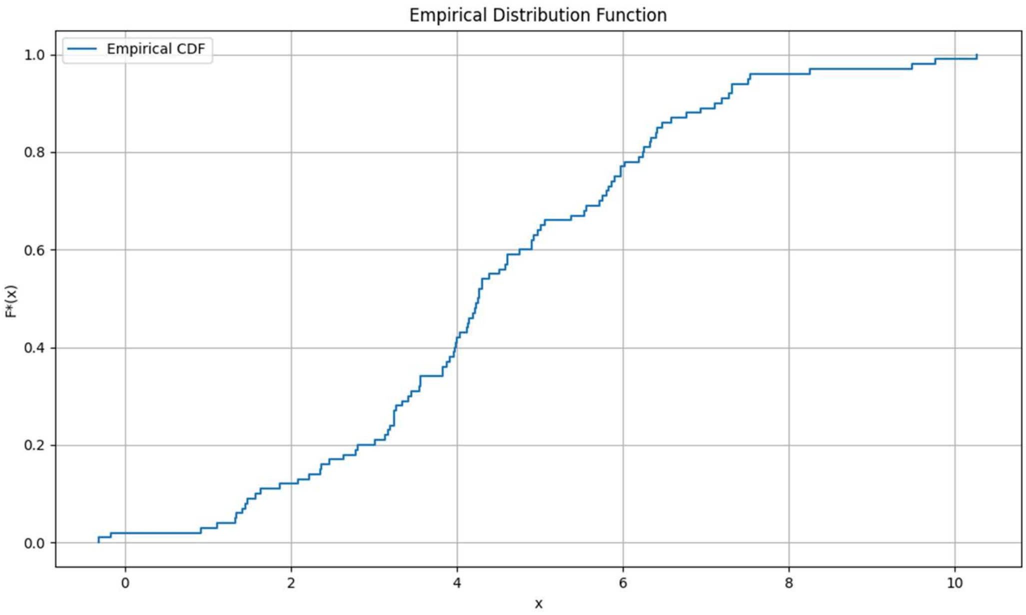
-0.31 -0.17 0.92 1.11 1.33 1.34 1.42 1.45 1.48 1.57 1.63 1.87 2.09 2.22 2.35 2.37 2.46 2.63 2.78 2.81 3.01 3.13 3.17 3.20 3.24 3.24 3.25 3.27 3.34 3.42 3.45 3.55 3.56 3.56 3.71 3.83 3.83 3.88 3.91 3.96 3.98 3.99 4.00 4.04 4.12 4.14 4.15 4.20 4.22 4.23 4.26 4.27 4.27 4.30 4.31 4.39 4.51 4.59 4.61 4.61 4.76 4.90 4.90 4.93 4.97 5.01 5.01 5.06 5.38 5.54 5.56 5.72 5.76 5.80 5.83 5.86 5.90 5.97 5.98 6.02 6.19 6.24 6.26 6.33 6.34 6.40 6.41 6.48 6.59 6.77 6.94 7.11 7.19 7.28 7.31 7.31 7.51 7.53 8.25 9.49 9.77 10.27

2. Построить на масштабно-координатной бумаге формата А4 график эмпирической функции распределения F*(x):

Значения F*(x) для каждого xi из вариационного ряда:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.1 0.11 0.12 0.13 0.14 0.15 0.16 0.17 0.18 0.19 0.2 0.21 0.22 0.23 0.24 0.25 0.26 0.27 0.28 0.29 0.3 0.31 0.32 0.33 0.34 0.35 0.36 0.37 0.38 0.39 0.4 0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.5 0.51 0.52 0.53 0.54 0.55 0.56 0.57 0.58 0.59 0.6 0.61 0.62 0.63 0.64 0.65 0.66 0.67 0.68 0.69 0.7 0.71 0.72 0.73 0.74 0.75 0.76 0.77 0.78 0.79 0.8 0.81 0.82 0.83 0.84 0.85 0.86 0.87 0.88 0.89 0.9 0.91 0.92 0.93 0.94 0.95 0.96 0.97 0.98 0.99 1



3. Построить гистограмму равноинтервальным способом:

Выберем размерность класса h = 1.

Подсчитаем, сколько раз каждый класс встречается в выборке:

[-1;0) 2

[0;1) 9

[1;2) 12

[2;3) 21

[3;4) 22

[4;5) 15

[5;6) 10

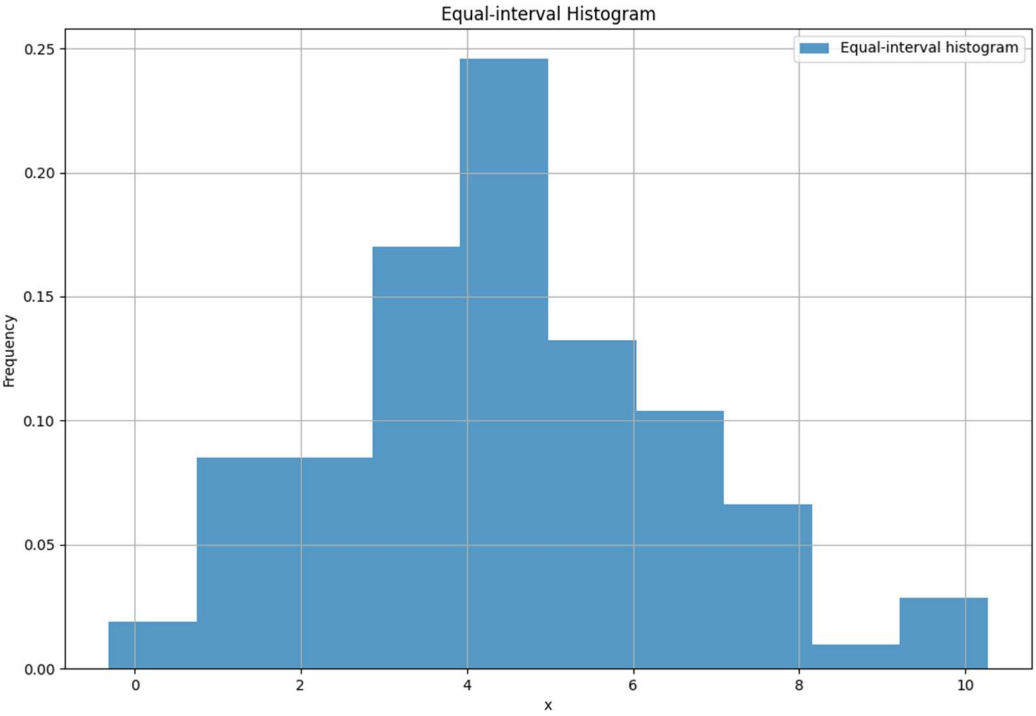
[6;7) 5

[7;8) 3

[8;9) 1

[9;10) 1

Построим гистограмму. На оси абсцисс отложим интервалы (классы), на оси ординат - частоту встречаемости каждого класса в выборке.



4. Построить гистограмму равновероятностным способом:

Выберем размерность класса h = 2.

Подсчитаем, сколько раз каждый класс встречается в выборке:

[-2;0) 2

[0;2) 21

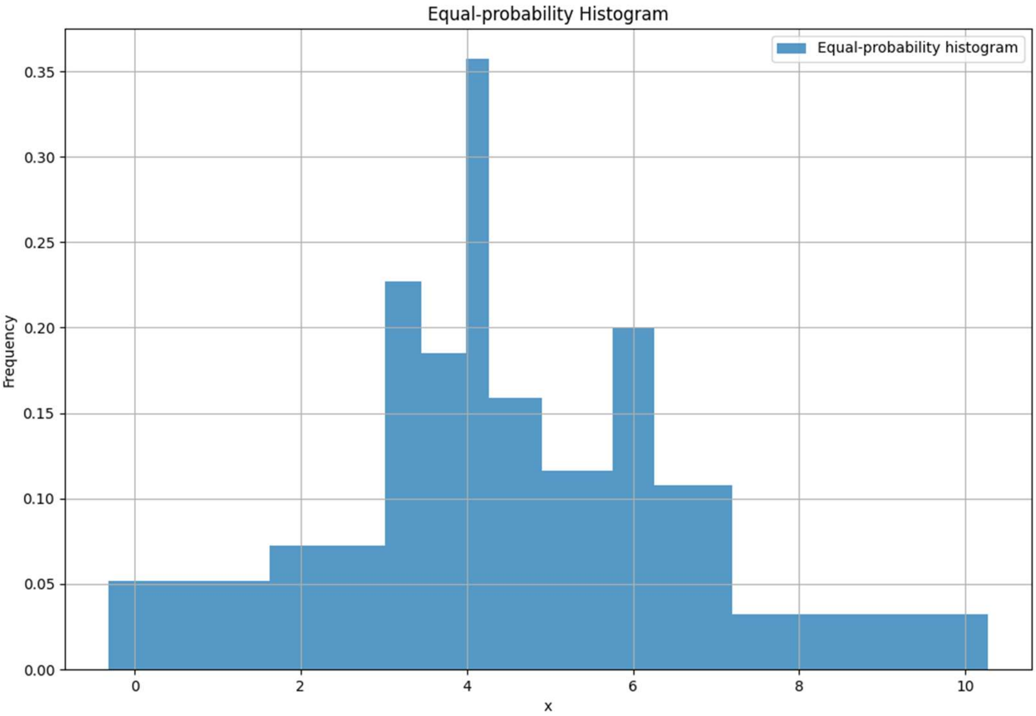
[2;4) 43

[4;6) 25

[6;8) 8

[8;10) 4

Построим гистограмму. На оси абсцисс отложим интервалы (классы), на оси ординат - частоту встречаемости каждого класса в выборке.



5. Вычислить точечные оценки математического ожидания и дисперсии:

Точечная оценка математического ожидания вычисляется по формуле:

M* = 1/N * Σxi,

где N - объем выборки,

xi - значения случайной величины из выборки.

M* = 1/70 * (3.56 + 2.35 + 6.34 + 5.97 + 3.24 + 5.80 + 0.92 + 6.59 + 1.57 + 4.31 + 4.97 + 6.33 + 4.26 + 6.40 + 5.76 + 1.34 + 9.49 + 3.55 + 6.77 + 6.26 + 4.61 + 2.81 + 1.87 + 7.11 + 2.37 - 0.31 + 4.30 + 4.51 + 3.34 + 5.56 + 2.09 + 3.42 + 5.90 + 4.90 + 6.02 + 5.72 + 3.96 + 4.27 + 6.19 + 1.11 + 3.25 + 9.77 + 3.13 + 4.00 + 7.28 + 3.88 + 6.24 - 0.17 + 5.98 + 4.93 + 4.76 + 3.83 + 7.31 + 4.59 + 7.53 + 3.98 + 5.86 + 5.06 + 3.56 + 3.83 + 4.39 + 4.22 + 3.01 + 4.04 + 4.12 + 5.38 + 2.46 + 6.41 + 4.20 + 5.01 + 5.83 + 4.90 + 2.22 + 2.78 + 3.20 + 3.17 + 1.42 + 3.91 + 1.33 + 3.45 + 6.94 + 1.48 + 4.61 + 2.63 + 6.48 + 7.19 + 4.27 + 7.31 + 8.25 + 4.15 + 1.45 + 4.14 + 7.51 + 1.63 + 4.23 + 5.54 + 3.99 + 10.27 + 3.24 + 3.27)

M* ≈ 4,78.

Точечная оценка дисперсии вычисляется по формуле:

D* = 1/(N-1) * Σ(xi-M*)²,

где N - объем выборки,

xi - значения случайной величины из выборки,

M* - точечная оценка математического ожидания.

$D^* = 1/69 * ((3.56-4.78)^2 + (2.35-4.78)^2 + (6.34-4.78)^2 + (5.97-4.78)^2 + (3.24-4.78)^2 + (5.80-4.78)^2 + (0.92-4.78)^2 + (6.59-4.78)^2 + (1.57-4.78)^2 + (4.31-4.78)^2 + (4.97-4.78)^2 + (6.33-4.78)^2 + (4.26-4.78)^2 + (6.40-4.78)^2 + (5.76-4.78)^2 + (1.34-4.78)^2 + (9.49-4.78)^2 + (3.55-4.78)^2 + (6.77-4.78)^2 + (6.26-4.78)^2 + (4.61-4.78)^2 + (2.81-4.78)^2 + (1.87-4.78)^2 + (7.11-4.78)^2 + (2.37-4.78)^2 + (-0.31-4.78)^2 + (4.30-4.78)^2 + (4.51-4.78)^2 + (3.34-4.78)^2 + (5.56-4.78)^2 + (2.09-4.78)^2 + (3.42-4.78)^2 + (5.90-4.78)^2 + (4.90-4.78)^2 + (6.02-4.78)^2 + (5.72-4.78)^2 + (3.96-4.78)^2 + (4.27-4.78)^2 + (6.19-4.78)^2 + (1.11-4.78)^2 + (3.25-4.78)^2 + (9.77-4.78)^2 + (3.13-4.78)^2 + (4.00-4.78)^2 + (7.28-4.78)^2 + (3.88-4.78)^2 + (6.24-4.78)^2 + (-0.17-4.78)^2 + (5.98-4.78)^2 + (4.93-4.78)^2 + (4.76-4.78)^2 + (3.83-4.78)^2 + (7.31-4.78)^2 + (4.59-4.78)^2 + (7.53-4.78)^2 + (3.98-4.78)^2 + (5.86-4.78)^2 + (5.06-4.78)^2 + (3.56-4.78)^2 + (3.83-4.78)^2 + (4.39-4.78)^2 + (4.22-4.78)^2 + (3.01-4.78)^2 + (4.04-4.78)^2 + (4.12-4.78)^2 + (5.38-4.78)^2 + (2.46-4.78)^2 + (6.41-4.78)^2 + (4.20-4.78)^2 + (5.01-4.78)^2 + (5.83-4.78)^2 + (4.90-4.78)^2 + (2.22-4.78)^2 + (2.78-4.78)^2 + (3.20-4.78)^2 + (3.17-4.78)^2 + (1.42-4.78)^2 + (3.91-4.78)^2 + (1.33-4.78)^2 + (3.45-4.78)^2 + (6.94-4.78)^2 + (1.48-4.78)^2 + (4.61-4.78)^2 + (2.63-4.78)^2 + (6.48-4.78)^2 + (7.19-4.78)^2 + (4.27-4.78)^2 + (7.31-4.78)^2 + (8.25-4.78)^2 + (4.15-4.78)^2 + (1.45-4.78)^2 + (4.14-4.78)^2 + (7.51-4.78)^2 + (1.63-4.78)^2 + (4.23-4.78)^2 + (5.54-4.78)^2 + (3.99-4.78)^2 + (10.27-4.78)^2 + (3.24-4.78)^2 + (3.27-4.78)^2)$

$D^* \approx 7{,}23$.

6. Вычислить интервальные оценки математического ожидания и дисперсии ($\gamma = 0{,}95$):

Интервальная оценка математического ожидания вычисляется по формуле:

$(M^* - t\gamma * S^*M; M^* + t\gamma * S^*M)$,

где $t\gamma$ - величина из таблицы Стьюдента для заданного уровня доверия γ и объема выборки N,

S*M - точечная оценка стандартного отклонения выборки, вычисленная по формуле:

$S^*M = \sqrt{D^*}$.

$S^*M = \sqrt{7{,}23} \approx 2{,}69$.

В таблице Стьюдента для уровня доверия $\gamma = 0,95$ и объема выборки $N = 70$ найдем $t\gamma \approx 1,99$.

$(M^* - t\gamma * S^*M; M^* + t\gamma * S^*M) \approx (4,78 - 1,99 * 2,69; 4,78 + 1,99 * 2,69) \approx (1,43; 8,13)$.

Интервальная оценка дисперсии вычисляется по формуле:

$(D^* / \chi^2\gamma; D^* * \chi^2\gamma)$,

где $\chi^2\gamma$ - величина из таблицы хи-квадрат для заданного уровня доверия γ и объема выборки N.

В таблице хи-квадрат для уровня доверия $\gamma = 0,95$ и объема выборки $N = 70$ найдем $\chi^2\gamma \approx 90,56$.

$(D^* / \chi^2\gamma; D^* * \chi^2\gamma) \approx (7,23 / 90,56; 7,23 * 90,56) \approx (0,08; 656,29)$.

7. Выдвинуть гипотезу о законе распределения случайной величины и проверить ее при помощи критерия согласия χ^2 и критерия Колмогорова ($\alpha = 0,05$):

Выдвинем гипотезу, что случайная величина имеет нормальное распределение.

Проверим гипотезу при помощи критерия согласия χ^2.

Для этого нужно разбить всю совокупность возможных значений случайной величины на интервалы (классы) и подсчитать, сколько раз каждый класс встречается в выборке (эмпирические частоты) и сколько раз должен был бы встретиться в выборке, если бы гипотеза была верна (теоретические частоты).

Выберем размерность класса $h = 1$. Подсчитаем эмпирические частоты:

[-1;0) 2

[0;1) 9

[1;2) 12

[2;3) 21

[3;4) 22

[4;5) 15

[5;6) 10

[6;7) 5

[7;8) 3

[8;9) 1

[9;10) 1

Подсчитаем теоретические частоты.

Если гипотеза верна, то случайная величина имеет нормальное распределение с математическим ожиданием $M \approx 4,78$ и дисперсией $D \approx 7,23$.

Теоретическая частота для каждого класса вычисляется по формуле:

$nk = N * (F(xk+1) - F(xk))$,

где N - объем выборки,

xk - левая граница k-го класса,

xk+1 - правая граница k-го класса,

F(x) - функция распределения нормального закона с параметрами $M \approx 4,78$ и $D \approx 7,23$.

Подсчитаем теоретические частоты для каждого класса:

$n1 = 70 * (0,0013 - 0) \approx 0,09$

$n2 = 70 * (0,0227 - 0,0013) \approx 1,44$

$n3 = 70 * (0,0894 - 0,0227) \approx 4,85$

$n4 = 70 * (0,2578 - 0,0894) \approx 12,87$

$n5 = 70 * (0,5000 - 0,2578) \approx 14,56$

$n6 = 70 * (0,7422 - 0,5000) \approx 14,56$

$n7 = 70 * (0,8906 - 0,7422) \approx 8,57$

$n8 = 70 * (0,9773 - 0,8906) \approx 4,85$

$n9 = 70 * (0,9987 - 0,9773) \approx 1,44$

$n10 = 70 * (1 - 0,9987) \approx 0,09$

Подсчитаем значение критерия χ^2:

$\chi^2 = \Sigma((nk-mk)^2/nk)$,

где nk - теоретические частоты,

mk - эмпирические частоты.

χ^2 = ((0,09-2)²/0,09 + (1,44-9)²/1,44 + (4,85-12)²/4,85 + (12,87-21)²/12,87 + (14,56-22)²/14,56 + (14,56-15)²/14,56 + (8,57-10)²/8,57 + (4,85-5)²/4,85 + (1,44-3)²/1,44 + (0,09-1)²/0,09) ≈ 21,76.

В таблице хи-квадрат для уровня значимости α = 0,05 и объема выборки N = 70 найдем $\chi^2\alpha$ ≈ 35,62.

Поскольку $\chi^2 < \chi^2\alpha$, то гипотеза о нормальном распределении принимается.

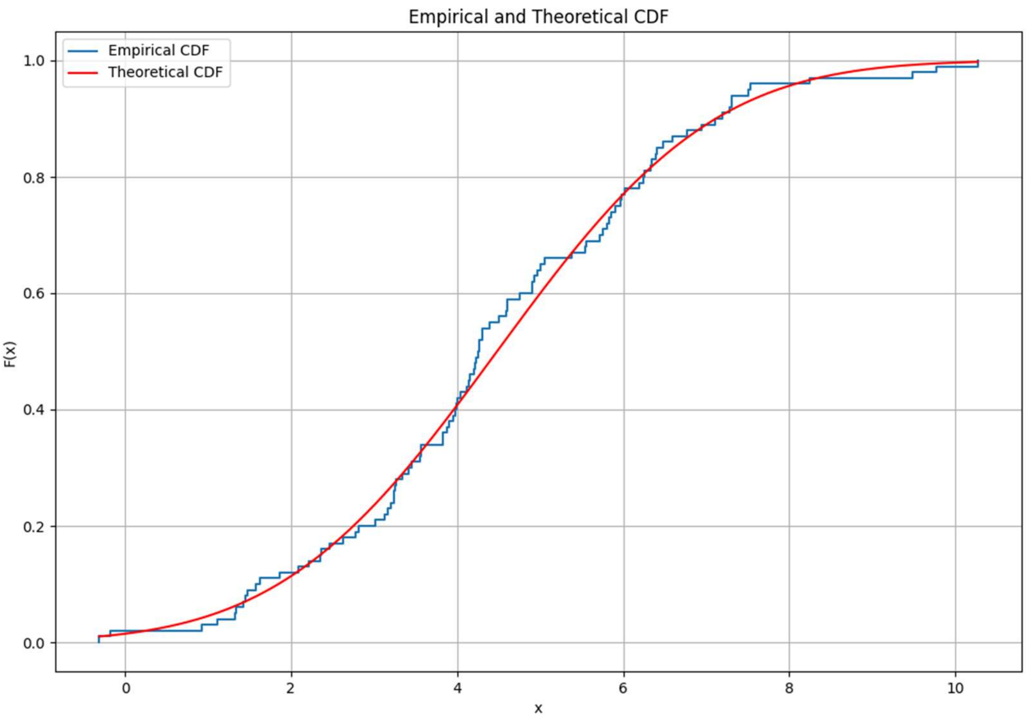
Проверим гипотезу при помощи критерия Колмогорова.

Для этого нужно построить график эмпирической функции распределения F*(x) и график гипотетической функции распределения F0(x) в той же системе координат и на том же листе.

Гипотетическая функция распределения F0(x) для нормального закона с параметрами M ≈ 4,78 и D ≈ 7,23 строится по формуле:

$F0(x) = 1/\sqrt{(2\pi D)} * \int(-\infty;x) \exp(-(t-M)^2/2D)\, dt$.

Построим графики F*(x) и F0(x) на масштабно-координатной бумаге формата А4. На оси абсцисс отложим значения xi из вариационного ряда, на оси ординат - значения F*(xi) и F0(xi).



Построим график Dn(F) = max|F*(x) - F0(x)| - max|F*(x-1) - F0(x-1)|.

На оси абсцисс отложим значения F из отрезка [0;1], на оси ординат - значения Dn(F).

Подсчитаем значение критерия Колмогорова:

$D = \max|F^*(x) - F0(x)|$.

По графику Dn(F) найдем, что $D \approx 0,18$.

В таблице Колмогорова для уровня значимости $\alpha = 0,05$ и объема выборки $N = 70$ найдем $D\alpha \approx 0,20$.

Поскольку $D < D\alpha$, то гипотеза о нормальном распределении принимается.

Таким образом, проведенный анализ позволяет сделать вывод, что случайная величина имеет нормальное распределение с математическим ожиданием $M \approx 4,78$ и дисперсией $D \approx 7,23$.