Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н. Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ «Информатика и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

ОТЧЕТ

по лабораторной работе №2.

по курсу «Математическая статистика.»

на тему: «Интервальные оценки.»

Студент ИУ7-62Б (Группа) ____________ (Подпись, дата) Трунов А. Р. (И. О. Фамилия)

Преподаватель ____________ (Подпись, дата) Власов П. А. (И. О. Фамилия)

2022 г.

Теоретическая часть

Определение γ-доверительного интервала для значения параметра распределения случайной величины

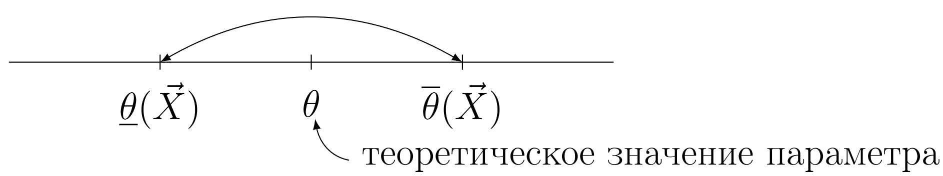
Пусть X – случайная величина, закон распределения которой зависит от вектора $\vec{\theta} = (\theta_1, \, .., \, \theta_r)$ неизвестных параметров (**вторая задача математической статистики**).

Интервальной оценкой параметра θ **уровня** γ называется пара статистик $\underline{\theta}(\vec{X})$ и $\overline{\theta}(\vec{X})$ таких, что

$$P\{\, \theta \in (\underline{\theta}(\vec{X}), \, \overline{\theta}(\vec{X}))\} = \gamma.$$

<u>Замечания</u>:

1. Интервальная оценка уровня γ – интервал со случайными границами $\underline{\theta}(\vec{X})$ и $\overline{\theta}(\vec{X})$, который «накрывает» неизвестное теоретическое значение параметра с вероятностью γ.



2. Вероятность совершить ошибку при построении интервальной оценки уровня γ:

$$1 - \gamma = P\{\, \gamma \notin (\underline{\theta}(\vec{X}), \, \overline{\theta}(\vec{X}))\}.$$

3. Вероятностной характеристикой интервальной оценки уровня γ является случайная величина $l(\vec{X}) = \overline{\theta}(\vec{X}) - \underline{\theta}(\vec{X})$, которая называется размахом интервальной оценки.

4. Иногда удобно строить односторонние интервальные оценки.

   Односторонней нижней γ-доверительной границей для параметра θ называют статистику $\underline{\theta}(\vec{X})$ такую, что

$$P\{\ \theta\ >\ \underline{\theta}(\vec{X})\}\ =\ \gamma.$$

   Односторонней верхней γ-доверительной границей для параметра θ называют статистику $\overline{\theta}(\vec{X})$ такую, что

$$P\{\ \theta\ <\ \overline{\theta}(\vec{X})\}\ =\ \gamma.$$

γ-**доверительным интервалом** (доверительный интервал уровня γ) для параметра θ называют реализацию (выборочное значение) интервальной оценки уровня γ для этого параметра, то есть интервал с детерминированными границами $(\underline{\theta}(\vec{x}),\ \overline{\theta}(\vec{x}))$.

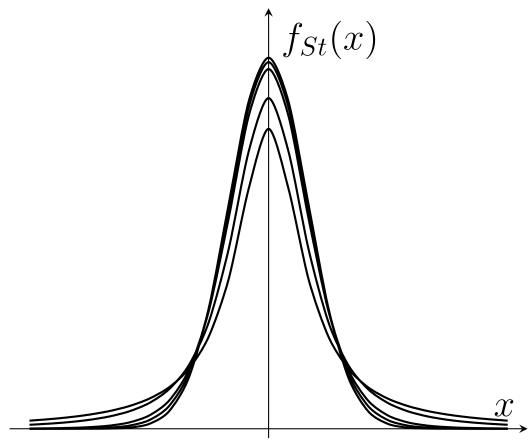
Формулы для вычисления границ γ-доверительного интервала для математического ожидания и дисперсии нормальной случайной величины

Формулы для вычисления границ γ-доверительного интервала для математического ожидания:

$$\underline{\mu}(\vec{X})\ =\ \bar{X}\ +\ \frac{S(\vec{X})t^{(n-1)}_{\frac{1-\gamma}{2}}}{\sqrt{n}}.$$

$$\overline{\mu}(\vec{X})\ =\ \bar{X}\ +\ \frac{S(\vec{X})t^{(n-1)}_{\frac{1+\gamma}{2}}}{\sqrt{n}}.$$

Однако функция плотности распределения Стьюдента – чётная функция (видно из графика ниже).



Таким образом

$$t^{n-1}_{\frac{1-\gamma}{2}} = -t^{n-1}_{\frac{1+\gamma}{2}}.$$

Тогда формулы для вычисления границ γ-доверительного интервала для математического ожидания можно представить в виде:

$$\underline{\mu}(\vec{X}) = \bar{X} - \frac{S(\vec{X})t^{(n-1)}_{\frac{1+\gamma}{2}}}{\sqrt{n}}.$$

$$\overline{\mu}(\vec{X}) = \bar{X} + \frac{S(\vec{X})t^{(n-1)}_{\frac{1+\gamma}{2}}}{\sqrt{n}}.$$

n – объём выборки;

γ – уровень доверия;

$t^{n-1}_{\frac{1+\gamma}{2}}$ – квантиль соответствующих уровней распределения Стьюдента (табличное значение $St(n-1)$);

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ – выборочное среднее (точечная оценка математического ожидания);

$S^2(\vec{X}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ – исправленная выборочная дисперсия (точечная оценка дисперсии).

Формулы для вычисления границ γ-доверительного интервала для дисперсии:

$$\underline{\sigma}(\vec{X}) = \sqrt{\frac{(n-1)S^2(\vec{X})}{h^{(n-1)}_{\frac{1-\gamma}{2}}}}.$$

$$\overline{\sigma}(\vec{X}) = \sqrt{\frac{(n-1)S^2(\vec{X})}{h^{(n-1)}_{\frac{1+\gamma}{2}}}}.$$

n – объём выборки;

γ – уровень доверия;

$h^{n-1}_{\frac{1-\gamma}{2}}$, $h^{n-1}_{\frac{1+\gamma}{2}}$ – квантили соответствующих уровней распределения Хи-квадрат (табличное значение $\chi^2(n-1)$);

$S^2(\vec{X}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ – исправленная выборочная дисперсия (точечная оценка дисперсии).

Практическая часть

Текст программы

```
%Лабораторная работа №2. Вариант 22.

function main()
    pkg load statistics;
     %Генеральной совокупностью называется множество возможных значений
     %случайного величины X с учётом её закона распределения.

     %Создание выборки объёма n из генеральной совокупности X.
     X = [7.76 5.96 4.58 6.13 5.05 6.40 7.46 5.55 5.01 ...
            3.79 7.65 8.87 5.94 7.25 6.76 6.92 6.68 4.89 ...
            7.47 6.53 6.76 6.96 6.58 7.92 8.47 6.27 8.05 ...
            5.24 5.60 6.69 7.55 6.02 7.34 6.81 7.22 6.39 ...
            6.40 8.28 5.39 5.68 6.71 7.89 5.69 5.18 7.84 ...
            7.18 7.54 6.04 4.58 6.82 4.45 6.75 5.28 7.42 ...
            6.88 7.10 5.24 9.12 7.37 5.50 5.52 6.34 5.31 ...
            7.71 6.88 6.45 7.51 6.21 7.44 6.15 6.25 5.59 ...
```

```
        6.68 6.52 4.03 5.35 6.53 3.68 5.91 6.68 6.18 ...
        7.80 7.17 7.31 4.48 5.69 7.11 6.87 6.14 4.73 ...
        6.60 5.61 7.32 6.75 6.28 6.41 7.31 6.68 7.26 ...
        7.94 7.67 4.72 6.01 5.79 7.38 5.98 5.36 6.43 ...
        7.25 5.54 6.66 6.47 6.84 6.13 6.21 5.52 6.33 ...
        7.55 6.24 7.84];

    %Уровень доверия
    gamma = 0.9;

    %Объём выборки
    len_X = length(X);

    %Точечная оценка математического ожидания
    mu = mean(X);

    %Точечная оценка дисперсии
    S_2 = var(X);

    %Нахождение границ доверительного интервала для математического ожидания
    mu_low = Find_Mu_Low(len_X, mu, S_2, gamma);
    mu_high = Find_Mu_High(len_X, mu, S_2, gamma);

    %Нахождение границ доверительного интервала для дисперсии
    sigma_low = Find_Sigma_Low(len_X, S_2, gamma);
    sigma_high = Find_Sigma_High(len_X, S_2, gamma);

    fprintf('Нахождение точечной оценки математического ожидания.\nmu = %f\n\n', mu);
    fprintf('Нахождение точечной оценки дисперсии.\nS_2 = %f\n\n', S_2);
    fprintf('Нахождение доверительного интервала для математического ожидания.\n( %f , %f )\n\n', mu_low, mu_high);
    fprintf('Нахождение доверительного интервала для дисперсии.\n( %f , %f )\n\n', sigma_high, sigma_low);

    mu_array = zeros(1, len_X);
    S_2_array = zeros(1, len_X);
```

```
    mu_low_array = zeros(1, len_X);
    mu_high_array = zeros(1, len_X);
    S_2_low_array = zeros(1, len_X);
    S_2_high_array = zeros(1, len_X);

    for i = 1 : len_X
        mu = mean(X(1:i));
        S_2 = var(X(1:i));

        mu_array(i) = mu;
        S_2_array(i) = S_2;

        mu_low_array(i) = Find_Mu_Low(i, mu, S_2, gamma);
        mu_high_array(i) = Find_Mu_High(i, mu, S_2, gamma);

        S_2_low_array(i) = Find_Sigma_Low(i, S_2, gamma);
        S_2_high_array(i) = Find_Sigma_High(i, S_2, gamma);
    end

    figure("numbertitle", "off", ...
           "name", "Графики доверительных интервалов для математического ожидания.", ...
           "position", [10, 100, 1000, 850]);
    plot(1 : len_X, [(zeros(1, len_X) + mu)', mu_array', mu_low_array', mu_high_array']);
    xlabel('n', "fontsize", 15);
    ylabel('y', "fontsize", 15);
    title("Графики доверительных интервалов для математического ожидания.", ...
          "fontsize", 15);

    graph_1 = 'y = hat \mu(vec x_N)';
    graph_2 = 'y = hat \mu(vec x_n)';
    graph_3 = 'y = underline \mu(vec x_n)';
    graph_4 = 'y = overline \mu(vec x_n)';
    legend(graph_1, graph_2, graph_3, graph_4, "fontsize", 15);
    grid;
    hold off;
```

```

    figure("numbertitle", "off", ...
            "name", "Графики доверительных интервалов для дисперсии.", ...
            "position", [10, 100, 1000, 850]);
    plot(1 : len_X, [(zeros(1, len_X) + S_2)', S_2_array', S_2_low_array', S_2_high_array']);
    xlabel('n', "fontsize", 15);
    ylabel('z', "fontsize", 15);
    title("Графики доверительных интервалов для дисперсии.", ...
          "fontsize", 15);

    graph_1 = 'y = hat S^2(vec x_N)';
    graph_2 = 'y = hat S^2(vec x_n)';
    graph_3 = 'y = underline \sigma^2(vec x_n)';
    graph_4 = 'y = overline \sigma^2(vec x_n)';
    legend(graph_1, graph_2, graph_3, graph_4, "fontsize", 15);
    grid;
    hold off;
end

function mu_low = Find_Mu_Low(n, mu, S_2, gamma)
    mu_low = mu + ((sqrt(S_2) * tinv((1 - gamma) / 2, n - 1)) / sqrt(n));
end

function mu_high = Find_Mu_High(n, mu, S_2, gamma)
    mu_high = mu + ((sqrt(S_2) * tinv((1 + gamma) / 2, n - 1)) / sqrt(n));
end

function sigma_low = Find_Sigma_Low(n, S_2, gamma)
    sigma_low = ((n - 1) * S_2) / chi2inv((1 + gamma) / 2, n - 1);
end

function sigma_high = Find_Sigma_High(n, S_2, gamma)
    sigma_high = ((n - 1) * S_2) / chi2inv((1 - gamma) / 2, n - 1);
end
```

Результаты расчётов и графики для выборки из индивидуального варианта – 22 вариант (для $\gamma = 0.9$)

Ниже представлен результат работы программы, описанной выше.

```
>> main

Нахождение точечной оценки математического ожидания.
mu = 6.459583

Нахождение точечной оценки дисперсии.
S_2 = 1.101315

Нахождение доверительного интервала для математического ожидания.
( 6.300770 , 6.618397 )

Нахождение доверительного интервала для дисперсии.
( 0.900975 , 1.382288 )
```

Рисунок 1 – Вывод информации в консоль при запуске программы.

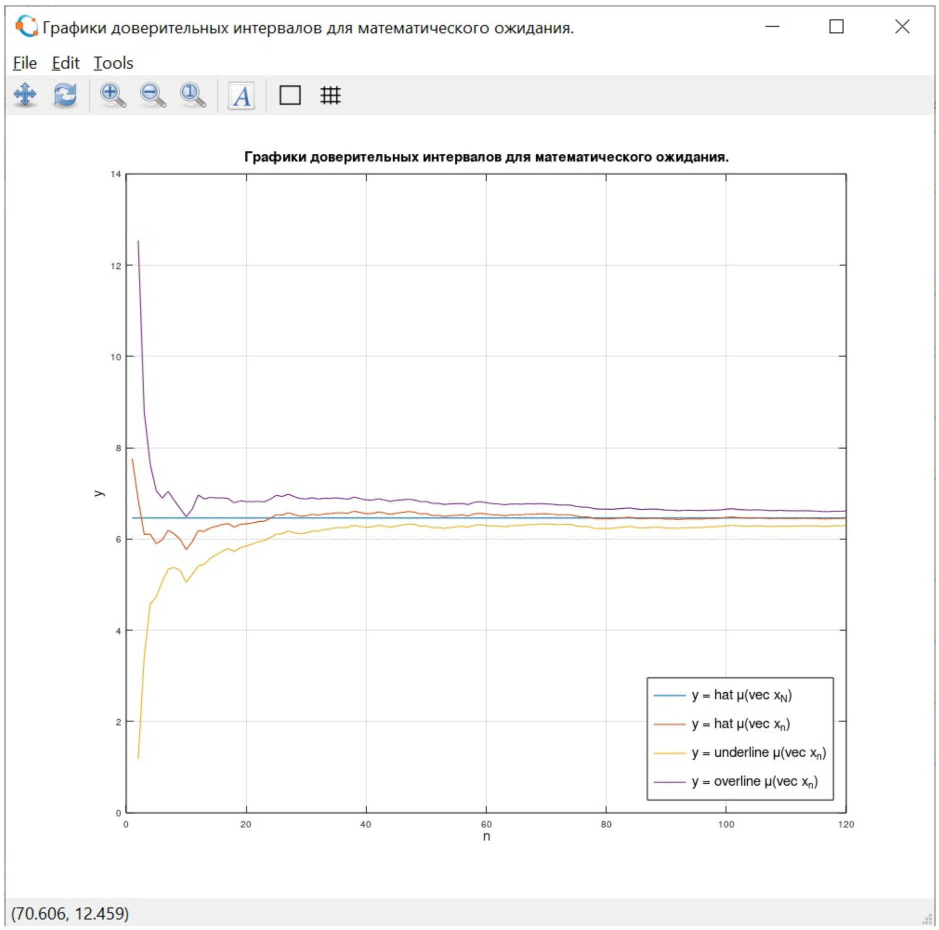


Рисунок 2 – График прямой $y = \hat{\mu}(\vec{x}_N)$, а также графики функций $y = \hat{\mu}(\vec{x}_n)$, $y = \underline{\mu}(\vec{x}_n)$ и $y = \overline{\mu}(\vec{x}_n)$, как функций объёма n выборки, где $n = \overline{1,\ N}$.

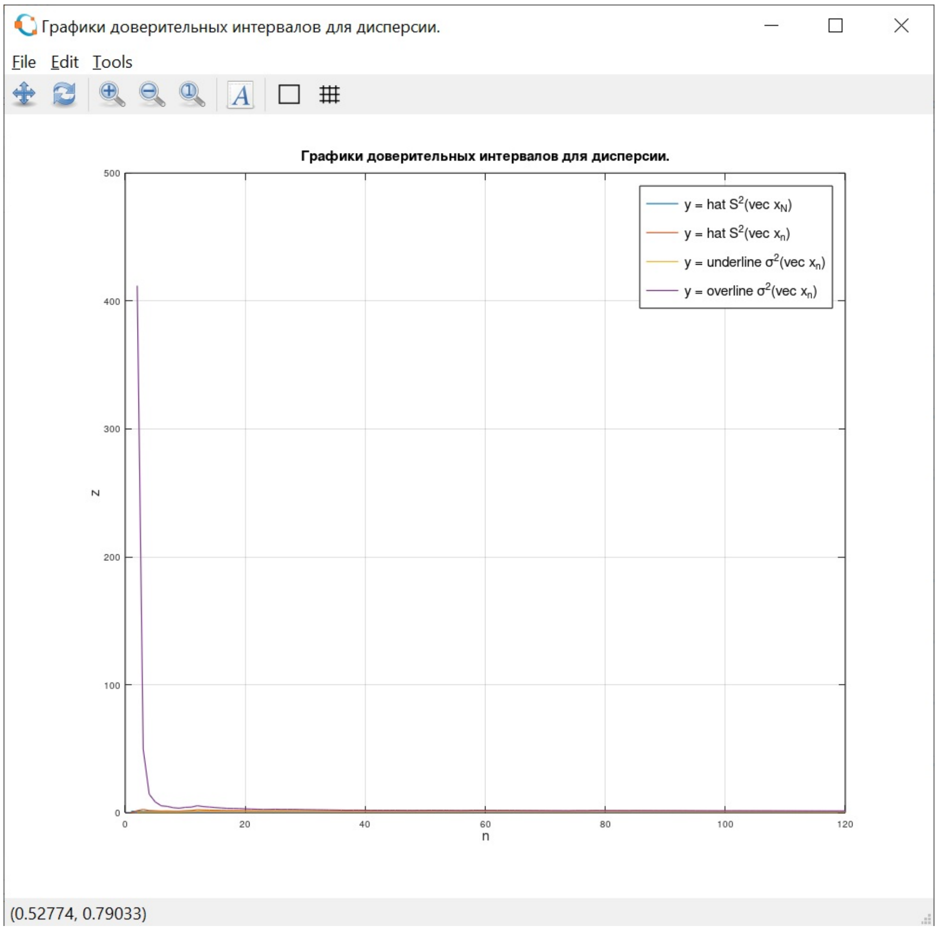


Рисунок 3 – График прямой $z = S^2(\vec{x}_N)$, а также графики функций $z = S^2(\vec{x}_n)$, $z = \underline{\sigma}^2(\vec{x}_n)$ и $z = \overline{\sigma}^2(\vec{x}_n)$, как функций объёма n выборки, где $n = \overline{1,\ N}$.