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 дисциплине «Математическая статистика»

Тема Интервальные оценки

Студент Леонов В.В.

Группа ИУ7-66Б

Оценка (баллы) ___________

Преподаватель Власов П.А., Андреева Т.В.

Москва — 2022 г.

Постановка задачи

Цель работы

Построение доверительных интервалов для математического ожидания и дисперсии нормальной случайной величины.

Содержание работы

1. Для выборки объема n из нормальной генеральной совокупности X реализовать в виде программы на ЭВМ:
   (a) вычисление точечных оценок $\hat{\mu}(\vec{X}_n)$ и $S^2(\vec{X}_n)$ математического ожидания MX и дисперсии DX соответственно;
   (b) вычисление нижней и верхней границ $\underline{\mu}(\vec{X}_n)$, $\overline{\mu}(\vec{X}_n)$ для γ-доверительного интервала для математического ожидания MX;
   (c) вычисление нижней и верхней границ $\underline{\sigma}^2(\vec{X}_n)$, $\overline{\sigma}^2(\vec{X}_n)$ для γ-доверительного интервала для дисперсии DX.
2. Вычислить $\hat{\mu}$ и S^2 для выборки из индивидуального варианта.
3. Для заданного пользователем уровня доверия γ (при построении графиков принять γ $= 0.9$) и N – объёма выборки из индивидуального варианта:
   (a) на координатной плоскости Oyn построить прямую $y = \hat{\mu}(\vec{x_N})$, также графики функций $y = \hat{\mu}(\vec{x}_n)$, $y = \underline{\mu}(\vec{x}_n)$ и $y = \overline{\mu}(\vec{x}_n)$ как функций объема n выборки, где n изменяется от 1 до N;
   (b) на другой координатной плоскости Ozn построить прямую $z = S^2(\vec{x_N})$, 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 объема n выборки, где n изменяется от 1 до N.

Вариант 9:

X = (-8.47, -7.45, -7.12, -8.30, -8.15, -6.01, -5.20, -7.38, -6.76, -9.18, -6.00, -8.08, -7.96, -8.34, -6.82, -8.46, -8.07, -7.04, -7.24, -8.16, -8.20, -8.27, -7.79, -7.37, -7.02, -7.13, -6.99, -7.65, -8.18, -6.71, -8.41, -6.71, -7.04, -9.15, -7.74, -10.11, -8.20, -7.07, -7.63, -8.99, -6.62, -6.23, -7.13, -6.41, -7.06, -7.72, -8.44, -8.85, -8.02, -6.98, -6.08, -7.20, -7.48, -7.82, -9.19, -8.31, -7.95, -7.97, -6.66, -6.59, -9.10, -7.87, -9.02, -8.77, -7.62, -9.44, -8.05, -7.60, -7.33, -6.94, -8.51, -7.39, -6.44, -8.88, -8.21, -7.66, -6.91, -8.39, -7.37, -7.26, -6.04, -7.58, -7.28, -7.02, -7.10, -7.33, -8.63, -8.21, -7.12, -8.11, -9.03, -8.11, -8.79, -9.22, -7.32, -5.97, -7.26, -6.39, -7.64, -8.38, -7.67, -7.70, -7.70, -8.95, -6.25, -8.09, -7.85, -8.10, -7.73, -6.78, -7.78, -8.20, -8.88, -8.51, -7.45, -7.14, -6.63, -7.38, -7.72, -6.25)

Общие теоретические сведения

Дана случайная величина X, закон распределения которой известен с точностью до неизвестного параметра θ.

Интервальной оценкой с уровнем доверия γ (γ-доверительной интервальной оценкой) параметра θ называют пару статистик $\underline{\theta}(\vec{X}), \overline{\theta}(\vec{X})$ таких, что

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

Поскольку границы интервала являются случайными величинами, то для различных реализаций случайной выборки $\vec{X}$ статистики $\underline{\theta}(\vec{X}), \overline{\theta}(\vec{X})$ могут принимать различные значения.

Доверительным интервалом с уровнем доверия γ (γ-доверительным интервалом) называют интервал $(\underline{\theta}(\vec{x}), \overline{\theta}(\vec{x}))$, отвечающий выборочным значениям статистик $\underline{\theta}(\vec{X}), \overline{\theta}(\vec{X})$.

Напишем формулы для вычисления границ γ-доверительного интервала для математического ожидания и дисперсии нормальной случайной величины при условии, что их теоретические значения не известны.

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

$$\underline{\mu}(\vec{X}_n) = \overline{X} + \frac{S(\vec{X}) t^{St(n-1)}_{\frac{1-\gamma}{2}}}{\sqrt{n}}; \overline{\mu}(\vec{X}_n) = \overline{X} + \frac{S(\vec{X}) t^{St(n-1)}_{\frac{1+\gamma}{2}}}{\sqrt{n}} \qquad (1)$$

$\overline{X}$ – точечная оценка математического ожидания;
$S(\vec{X}) = \sqrt{S^2(\vec{X})}$ – квадратный корень из точечной оценки дисперсии;
n – объем выборки;
γ – уровень доверия;
$t^{St(n-1)}_{\alpha}$ – квантиль уровня α распределения Стьюдента с $n-1$ степенями свободы.
Формулы для вычисления границ γ-доверительного интервала для дисперсии:

$$\underline{\sigma}(\vec{X}_n) = \frac{(n-1)S^2(\vec{X})}{t^{\chi^2(n-1)}_{\frac{1+\gamma}{2}}}; \overline{\sigma}(\vec{X}_n) = \frac{(n-1)S^2(\vec{X})}{t^{\chi^2(n-1)}_{\frac{1-\gamma}{2}}} \qquad (2)$$

$S^2(\vec{X})$ – точечная оценка дисперсии;
n – объем выборки;
γ – уровень доверия;
$t^{\chi^2(n-1)}_{\alpha}$ – квантиль уровня α распределения χ^2 с $n-1$ степенями свободы.

Ход работы

Для решения поставленной задачи была использованна программа *MATLAB 2022a*.

Листинг программы

```
X = [-8.47, -7.45, -7.12, -8.30, -8.15, -6.01, -5.20, -7.38, -6.76, -9.18,
    -6.00, -8.08, -7.96, -8.34, -6.82, -8.46, -8.07, -7.04, -7.24, -8.16,
    -8.20, -8.27, -7.79, -7.37, -7.02, -7.13, -6.99, -7.65, -8.18, -6.71,
    -8.41, -6.71, -7.04, -9.15, -7.74, -10.11, -8.20, -7.07, -7.63, -8.99,
    -6.62, -6.23, -7.13, -6.41, -7.06, -7.72, -8.44, -8.85, -8.02, -6.98,
    -6.08, -7.20, -7.48, -7.82, -9.19, -8.31, -7.95, -7.97, -6.66, -6.59,
    -9.10, -7.87, -9.02, -8.77, -7.62, -9.44, -8.05, -7.60, -7.33, -6.94,
    -8.51, -7.39, -6.44, -8.88, -8.21, -7.66, -6.91, -8.39, -7.37, -7.26,
    -6.04, -7.58, -7.28, -7.02, -7.10, -7.33, -8.63, -8.21, -7.12, -8.11,
    -9.03, -8.11, -8.79, -9.22, -7.32, -5.97, -7.26, -6.39, -7.64, -8.38,
    -7.67, -7.70, -7.70, -8.95, -6.25, -8.09, -7.85, -8.10, -7.73, -6.78,
    -7.78, -8.20, -8.88, -8.51, -7.45, -7.14, -6.63, -7.38, -7.72, -6.25];
gamma = 0.9;

% 1-2
[muhat, muci] = my_normfit_mu(X, 1 - gamma);
[s2hat, s2ci] = my_normfit_s2(X, 1 - gamma);
fprintf("mu = %f\nmu_bottom = %f\nmu_top = %f\n", muhat, muci(1), muci(2));
fprintf("s2 = %f\ns2_bottom = %f\ns2_top = %f\n", s2hat, s2ci(1), s2ci(2));
% 3
process_mu(X, gamma, muhat);
process_s2(X, gamma, s2hat);


function [muhat, muci] = normfit_mu(X, alpha)
  [muhat, ~, muci, ~] = normfit(X, alpha);
end

function [s2hat, s2ci] = normfit_s2(X, alpha)
  [~, sigmahat, ~, sigmaci] = normfit(X, alpha);
  s2hat = sigmahat ^ 2;
  s2ci = sigmaci .^ 2;
end


```

```
function [muhat, muci] = my_normfit_mu(X, alpha)
  muhat = mean(X);
  s = std(X);
  gamma = 1 - alpha;
  n = length(X);
  mu_bottom = muhat + s * tinv((1 - gamma) / 2, n - 1) / sqrt(n);
  mu_top = muhat + s * tinv((1 + gamma) / 2, n - 1) / sqrt(n);
  muci = [mu_bottom, mu_top];
end

function [s2hat, s2ci] = my_normfit_s2(X, alpha)
  s2hat = var(X);
  gamma = 1 - alpha;
  n = length(X);
  s2_top = (n - 1) * s2hat / chi2inv((1 - gamma) / 2, n - 1);
  s2_bottom = (n - 1) * s2hat / chi2inv((1 + gamma) / 2, n - 1);
  s2ci = [s2_bottom, s2_top];
end

function process_parameter(start, X, gamma, est, fit, line_legend,
    est_legend, top_legend, bottom_legend)
  N = length(X);
  figure;
  hold on;
  grid on;
  plot([1, N], [est, est]);
  ests = [];
  cis_bottom = [];
  cis_top = [];
  for n = start:N
    [est, cis] = fit(X(1:n), 1 - gamma);
    ests = [ests, est];
    cis_bottom = [cis_bottom, cis(1)];
    cis_top = [cis_top, cis(2)];
  end
  plot(start:N, ests);
  plot(start:N, cis_bottom);
  plot(start:N, cis_top);
  l = legend(line_legend, est_legend, top_legend, bottom_legend);
  set(l, 'Interpreter', 'latex', 'fontsize', 18);
  hold off;
end

function process_mu(X, gamma, muhat)
  process_parameter(1, X, gamma, muhat, @my_normfit_mu, '$\hat\mu(\vec x_N)$
      ', '$\hat\mu(\vec x_n)$', '$\underline\mu(\vec x_n)$', '$\overline\mu(\
      vec x_n)$');
end
```

```
function process_s2(X, gamma, S2)
  process_parameter(10, X, gamma, S2, @my_normfit_s2, '$\hat\sigma^2(\vec
    x_N)$', '$\hat\sigma^2(\vec x_n)$', '$\underline\sigma^2(\vec x_n)$', '
    $\overline\sigma^2(\vec x_n)$');
end
```

Результат работы программы

	$\hat\mu(\vec x_n)$	$S^2(\vec x_n)$	$\underline\mu(\vec x_n)$	$\overline\mu(\vec x_n)$	$\underline{S^2}(\vec x_n)$	$\overline{S^2}(\vec x_n)$
$\gamma = 0.95$	-7.660917	0.777892	-7.820342	-7.501492	0.612699	1.020613
$\gamma = 0.9$	-7.660917	0.777892	-7.794389	-7.527445	0.636386	0.976353
$\gamma = 0.8$	-7.660917	0.777892	-7.764675	-7.557158	0.665250	0.928415

Таким образом, при увелечении/уменьшении уровня доверия γ доверительный интервал увеличивается/уменьшается соответственно.

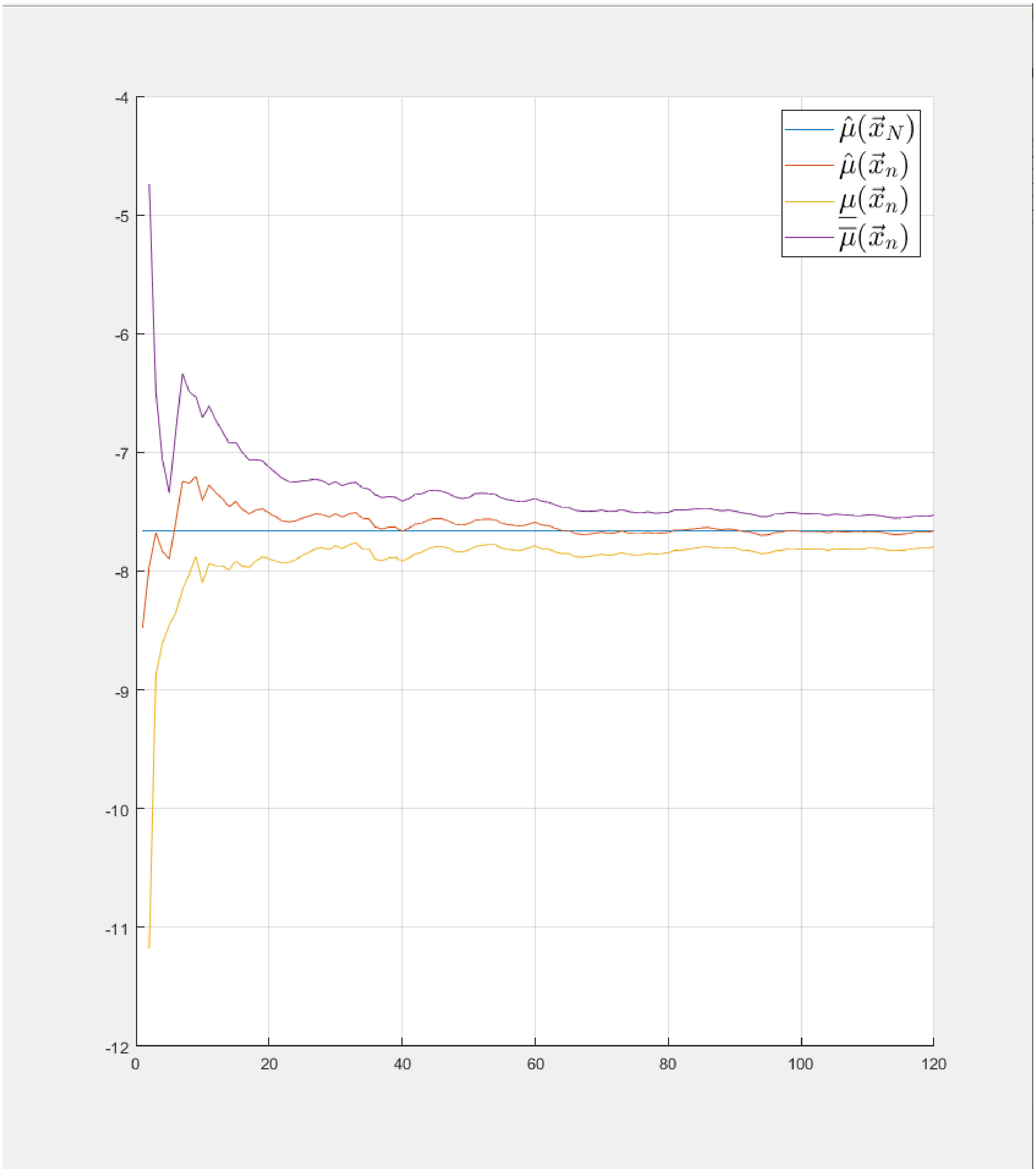

Рис. 1: Прямая $y(n) = \hat{\mu}(\vec{x}_N)$, а также графики функций $y(n) = \underline{\mu}(\vec{x}_n)$, $y(n) = \overline{\mu}(\vec{x}_n)$, $y(n) = \hat{\mu}(\vec{x}_n)$ как функций объема n выборки, где n изменяется от 1 до N

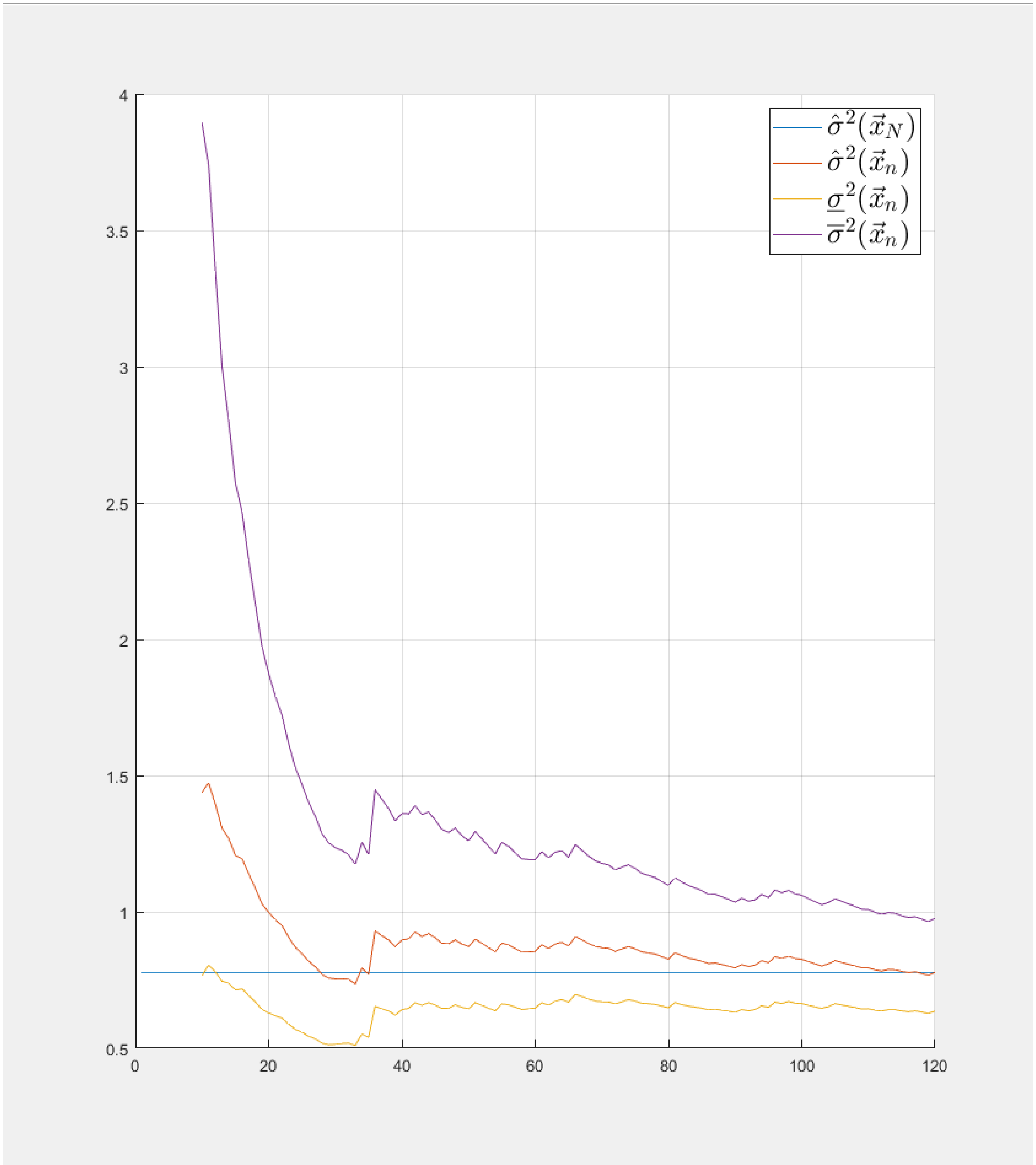

Рис. 2: Прямая $z(n) = \hat{S}^2(\vec{x}_N)$, а также графики функций $z(n) = \underline{S}^2(\vec{x}_n)$, $z(n) = \overline{S}^2(\vec{x}_n)$, $z(n) = \hat{S}^2(\vec{x}_n)$ как функций объема n выборки, где n изменяется от 10 до N