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

ФАКУЛЬТЕТ «Информатика и системы управления»
КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

Лабораторная работа № 2

Вариант 22

Дисциплина	Математическая статистика.
Тема	
Студент	Тимонин А. С.
Группа	ИУ7-62Б
Оценка (баллы)	
Преподаватель	Власов П.А.

Москва, 2020 г.

Теоретическая часть

Определение γ-доверительного интервала

Вторая задача математической статистики состоит в том, что зная общий вид закона распределения случайной величины X с точностью до вектора $\vec{\theta} = (\theta_1, ..., \theta_r)$ неизвестных параметров. Требуется оценить значение $\vec{\theta}$.

Если ранее для решения этой задачи использовались точечные оценки, недостаток которых состоит в том, что они не дают вероятностных характеристик точности оценивания параметров. Также для решения второй задачи математической статистики используюется другой подход, который заключается в построении доверительных интервалов.

Определение. γ-доверительным интервалом (доверительным интервалом уровня γ) для параметра θ называется пара статистик

$$\underline{\theta}(\vec{X}), \overline{\theta}(\vec{X})$$

таких, что

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

Замечание 1. Таким образом, γ-доверительный интервал — это интервал, который покрывает теоретическое значение параметра θ с вероятностью .

Замечание 2. Границы $\overline{\theta}(\vec{X})$, $\underline{\theta}(\vec{X})$ γ-доверительного интервала являются функциями случайной выборки, т.е. сами являются случайными величинами.

Замечание 3. Вероятностной характеристикой $gamma$-доверительного интервала является случайная величина $l(\vec{X}) = \overline{\theta}(\vec{X}) - \underline{\theta}(\vec{X})$, которая называется размахом этого интервала.

Замечание 4. Вероятность совершить ошибку при $gamma$-доверительном интервале:

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

Формулы для вычисления γ-доверительного интервала

Пусть $\vec{X}_n$ - случайная выборка объема n из генеральной совокупности X, $X \sim N(\mu, \sigma^2)$.

Оценка математического ожидания

$$\gamma = P\{-t_{1-\alpha} < g(\vec{X}_n, \mu) < t_{1-\alpha}\} = P\{-t_{1-\alpha} < \tfrac{\mu - \overline{X}}{S(\vec{X}_n)}\sqrt{n} < t_{1-\alpha}\} =$$
$$= P\{\underbrace{\overline{X} - \frac{S(\vec{X}_n)}{\sqrt{n}}t_{1-\alpha}}_{\underline{\mu}(\vec{X}_n)} < \mu < \underbrace{\overline{X} + \frac{S(\vec{X}_n)}{\sqrt{n}}t_{1-\alpha}}_{\overline{\mu}(\vec{X}_n)}\} =>$$
$$=> \underline{\mu}(\vec{X}_n) = \overline{X} - \tfrac{S(\vec{X}_n)}{\sqrt{n}}t_{1-\alpha}, \overline{\mu}(\vec{X}_n) = \overline{X} + \tfrac{S(\vec{X}_n)}{\sqrt{n}}t_{1-\alpha},$$

где:

$\overline{X}$ - оценка математического ожидания;

n - объем выборки;

$S^2(\vec{X}_n)$ - точечная оценка случайной выборки $\vec{X}_n$;

$t_{1-\alpha}$ - квантиль уровня $1-\alpha = \frac{1+\gamma}{2}$ для распределения $St(n-1)$;

Оценка дисперсии

$$\gamma = P\{h_{\frac{1-\gamma}{2}} < g(\vec{X}, \sigma^2) < h_{\frac{1+\gamma}{2}}\} = P\{h_{\frac{1-\gamma}{2}} < \frac{S^2(\vec{X})(n-1)}{\sigma^2} < h_{\frac{1+\gamma}{2}}\} =$$

$$P\{\frac{1}{h_{\frac{1+\gamma}{2}}} < \frac{\sigma^2}{S^2(\vec{X})(n-1)} < \frac{1}{h_{\frac{1-\gamma}{2}}}\} = P\{\underbrace{\frac{S^2(\vec{X})(n-1)}{h_{\frac{1+\gamma}{2}}}}_{\underline{\sigma}^2} < \sigma^2 < \underbrace{\frac{S^2(\vec{X})(n-1)}{h_{\frac{1-\gamma}{2}}}}_{\overline{\sigma}^2}\} =>$$

$$=> \underline{\sigma}^2 = \frac{S^2(\vec{X})(n-1)}{h_{\frac{1+\gamma}{2}}}, \overline{\sigma}^2 = \frac{S^2(\vec{X})(n-1)}{h_{\frac{1-\gamma}{2}}}$$

где:

$\overline{X}$ - оценка математического ожидания;

n - объем выборки;

$S^2(\vec{X})$ - точечная оценка случайной выборки $\vec{X}_n$;

$h_{\frac{1-\gamma}{2}}$ - квантиль уровня $\frac{1-\gamma}{2}$ для распределения $\chi^2(n-1)$

Практическая часть

Листинг 1: Код программы

```
function lab2()
  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];

  N = 6:120;
  gamma = 0.9;
  alpha = (1 - gamma) / 2;

  math_exp = mean(X);
  variance = var(X);

  math_exp_array = [];
  variance_array = [];

  for i = N
    math_exp_array = [math_exp_array, mean(X(1:i))];
    variance_array = [variance_array, var(X(1:i))];
  end

  figure
  plot([N(1), N(end)], [math_exp, math_exp], 'm');
  hold on;
```

```
    plot(N, math_exp_array, 'r');

    tmp_math_exp = sqrt(variance_array./N).*tinv(1 - alpha, N - 1);
    math_exp_low = math_exp_array - tmp_math_exp;
    math_exp_high = math_exp_array + tmp_math_exp;

    plot(N, math_exp_low, 'g');
    plot(N, math_exp_high, 'b');
    grid on; hold off;

    figure
    plot([N(1), N(end)], [variance, variance], 'm');
    hold on;
    plot(N, variance_array, 'r');

    tmp_variance = variance_array.*(N - 1);
    variance_low = tmp_variance./chi2inv(1 - alpha, N - 1);
    variance_high = tmp_variance./chi2inv(alpha, N - 1);

    plot(N, variance_low, 'g');
    plot(N, variance_high, 'b');
    grid on; hold off;

    fprintf('math_exp = %.2f\n', math_exp);
    fprintf('variance = %.2f\n\n', variance);

    fprintf('low math expectation = %.2f\n', math_exp_low(end));
    fprintf('high math exprectation = %.2f\n', math_exp_high(end));

    fprintf('low variance = %.2f\n', variance_low(end));
    fprintf('high variance = %.2f\n', variance_high(end));
end
```

Результат выполнения программы

$$\hat{\mu} = 6.46$$

$$S^2 = 1.10$$

$$(\underline{\mu}; \overline{\mu}) = (6.30; 6.62)$$

$$(\underline{\sigma}^2; \overline{\sigma}^2) = (0.90; 1.38)$$

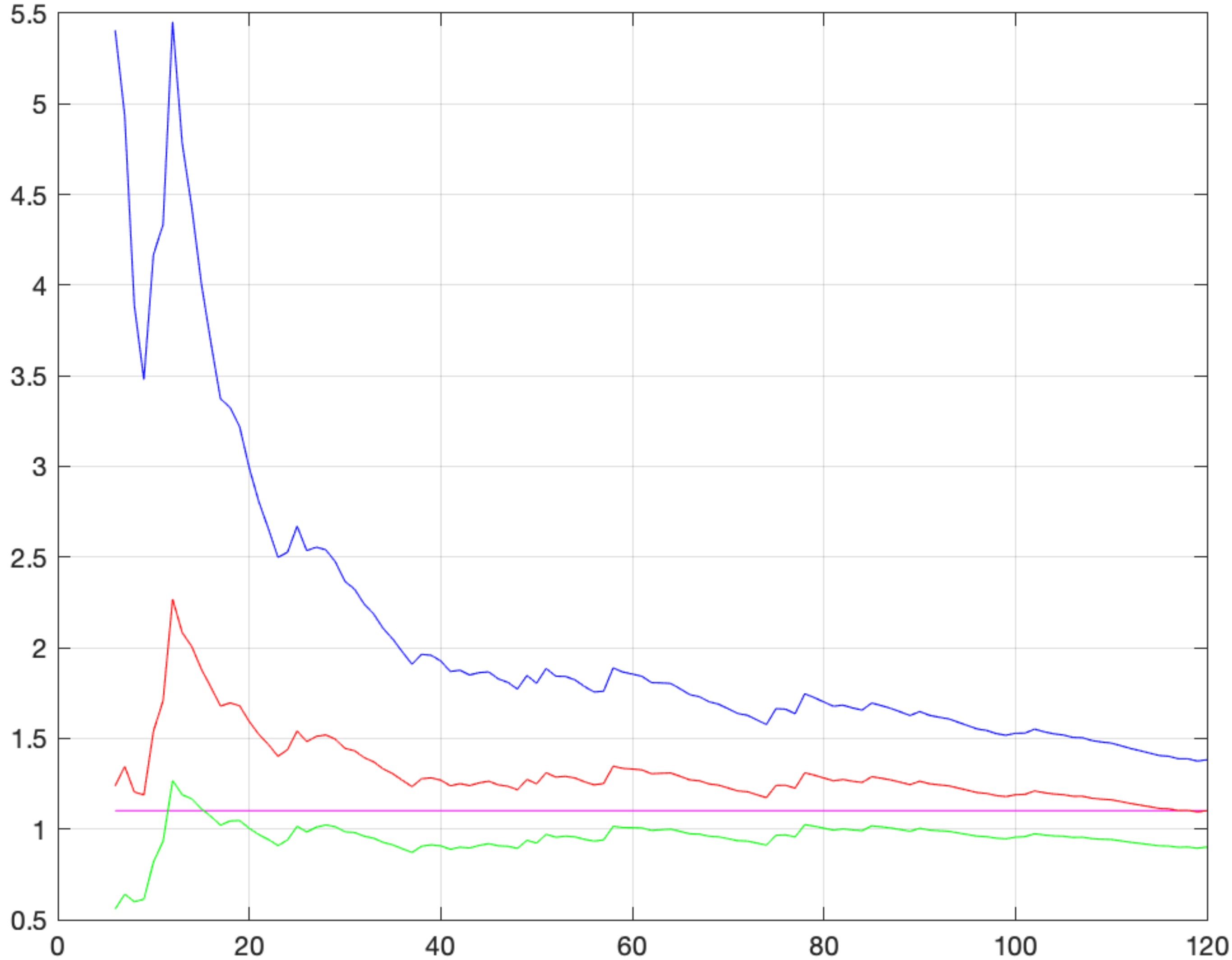

Рис. 1: Гистограмма

График функций: $y = \hat{\mu}(\vec{x}_n)$ (красный цвет), $y = \underline{\mu}(\vec{x}_N)$ (зеленый цвет), $y = \overline{\mu}(\vec{x}_N)$ (синий цвет) как функции объема n выборки, где $n = \overline{1; N}$;

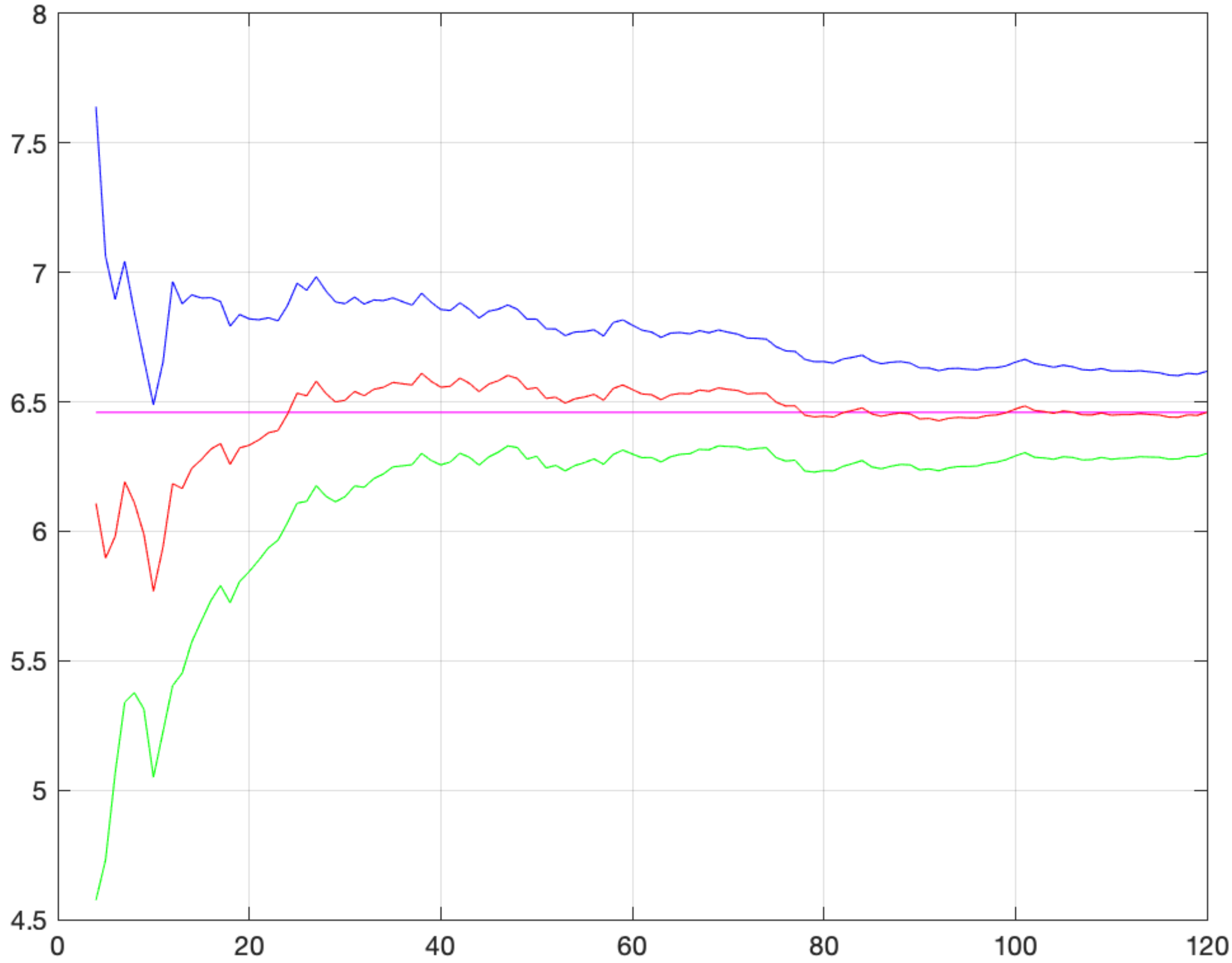

Рис. 2: Эмпирическая функция распределения

График функций: $z = S^2(\vec{x}_n)$ (красный цвет), $z = \underline{\sigma}^2(\vec{x}_n)$ (зеленый цвет), $z = \overline{\sigma}^2(\vec{x}_n)$ (синий цвет) как функции объема n выборки, где $n = \overline{1; N}$.