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»

Отчёт

по лабораторной работе №2

Название «Интервальные оценки»

Дисциплина «Математическая статистика»

Студент ИУ7-65Б _______________ Бугаенко А.П.
(подпись, дата) (Фамилия И.О.)

Преподователь _______________ Андреева Т.В.
(подпись, дата) (Фамилия И.О.)

Москва, 2022

1 Цели и задачи работы

Цель работы — построение доверительных интервалов для математического ожидания и дисперсии нормальной случайной величины. Содержание работы:

1) Для выборки объёма n из генеральной совокупности X реализовать в виде программы на ЭВМ:

а) вычисление точечных оценок $\hat{\mu}(\overrightarrow{x_n})$ и $S^2(\overrightarrow{x_n})$ математического ожидания MX и дисперсии DX соответственно;

б) вычисление нижней и верхней границ $\underline{\mu}(\overrightarrow{x_n})$, $\overline{\mu}(\overrightarrow{x_n})$ для γ-доверительного интервала для математического ожидания MX;

в) вычисление нижней и верхней границ $\underline{\sigma}^2(\overrightarrow{x_n})$, $\overline{\sigma}^2(\overrightarrow{x_n})$ для γ-доверительного интервала для дисперсии DX;

2) вычислить $\hat{\mu}$, S^2 из индивидуального варианта;

3) для заданного пользователем уровня доверия γ и N - объёма выборки из индивидуального варианта:

а) на координатной плоскости Oyn построить прямую $y = \hat{\mu}(\overrightarrow{x}_N)$, также графики функций $y = \hat{\mu}(\overrightarrow{x_n}), y = \underline{\mu}(\overrightarrow{x_n}), y = \overline{\mu}(\overrightarrow{x_n})$ как функций объёма n выборки, где n изменяется от 1 до N.

б) на другой координатной плоскости Ozn построить прямую $z = S^2(\overrightarrow{x_N})$, также графики функций $z = S^2(\overrightarrow{x_n}))$, $z = \underline{\sigma}^2(\overrightarrow{x_n})$ и $z = \overline{\sigma^2}(\overrightarrow{x}_n)$ как функций объёма n выборки, где n изменяется от 1 до N.

2 Теоретическая часть

2.1 Формулы для вычисления точечных оценок математического ожидания и дисперсии

Пусть $(x_1, ..., x_n)$ - некая случайная выборка.

Оценка $\hat{\mu}$ (u_x) для математического ожидания MX вычисляется как $\hat{\mu}(\overrightarrow{X_n}) = \overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Оценка S^2 (S_2_x) для дисперсии (несмещённая) вычисляется как $S^2(\overrightarrow{X_n}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2$.

2.2 Определение γ-доверительного интервала для значения параметра распределения случайной величины

Пусть $\overrightarrow{X_n}$ - случайная выборка n из генеральной совокупности X с функцией распределения $F(x;\theta)$, зависящей от параметра θ, значение которого неизвестно.

Тогда интервальной оценкой с коэффициентов доверия γ (γ-доверительной интервальной оценкой) параметра θ называют пару статистик $\underline{\theta}\overrightarrow{X_n}, \overline{\theta}\overrightarrow{X_n}$ таких, что:
$P\{\underline{\theta}\overrightarrow{X_n} < \theta < \overline{\theta}\overrightarrow{X_n}\} = \gamma$

Интервал $(\underline{\theta}(\overrightarrow{X_n}), \overline{\theta}(\overrightarrow{X_n}))$ называют интервальной оценкой для параметра θ с коэффициентов доверия γ, а $\underline{\theta}(\overrightarrow{X_n}), \overline{\theta}(\overrightarrow{X_n})$ называют соответственно нижней и верхней графницами интервальной оценки. Смысл интервальной оценки состоит в том, что она представляет интервал со случайными границами, который с заданной вероятностью γ накрывает неизвестное истинное значение параметра γ.

Интервал $(\underline{\theta}(\overrightarrow{X_n}), \overline{\theta}(\overrightarrow{X_n}))$ называют доверительным интервалом для параметра с коэффициентов доверия γ, где $\overrightarrow{x_n}$ - любая реализация случайной выборки $\overrightarrow{X_n}$.

2.3 Формулы для вычисления границ γ-доверительного интервала для математического ожидания и дисперсии

Для вычисления верхней и нижней границы γ-доверительного интервала для математического ожидания используются следующие формулы:
$\underline{\mu}(\overrightarrow{X_n}) = \overline{X} - \frac{S(\overrightarrow{X})}{\sqrt{n}} t_{1-\alpha}(n-1)$ (u_low)
$\overline{\mu}(\overrightarrow{X_n}) = \overline{X} - \frac{S(\overrightarrow{X})}{\sqrt{n}} t_{1-\alpha}(n-1)$ (u_high)
$\overline{X}$ - точечная оценка математического ожидания
$S^2(\overrightarrow{X})$ - точечная оценка дисперсии
n - объём выборки
γ - уровень доверия
$t_{1-\alpha}$ - квантиль уровня $1-\alpha$ для распределения Стьюдента с n - 1 степенями свободы, $\alpha = \frac{1-\gamma}{2}$

Для вычисления верхний и нижней границы γ-доверительного интервала для дисперсии:

$\underline{\sigma}^2(\overrightarrow{x_n}) = \frac{(n-1)S^2(\overrightarrow{X})}{\chi^2_{1-\alpha}(n-1)}$ (sigma_2_low)

$\overline{\sigma}^2(\overrightarrow{x_n}) = \frac{(n-1)S^2(\overrightarrow{X})}{\chi^2_{\alpha}(n-1)}$ (sigma_2_high)

$S^2(\overrightarrow{X})$ - точечная оценка дисперсии

n - объём выборки

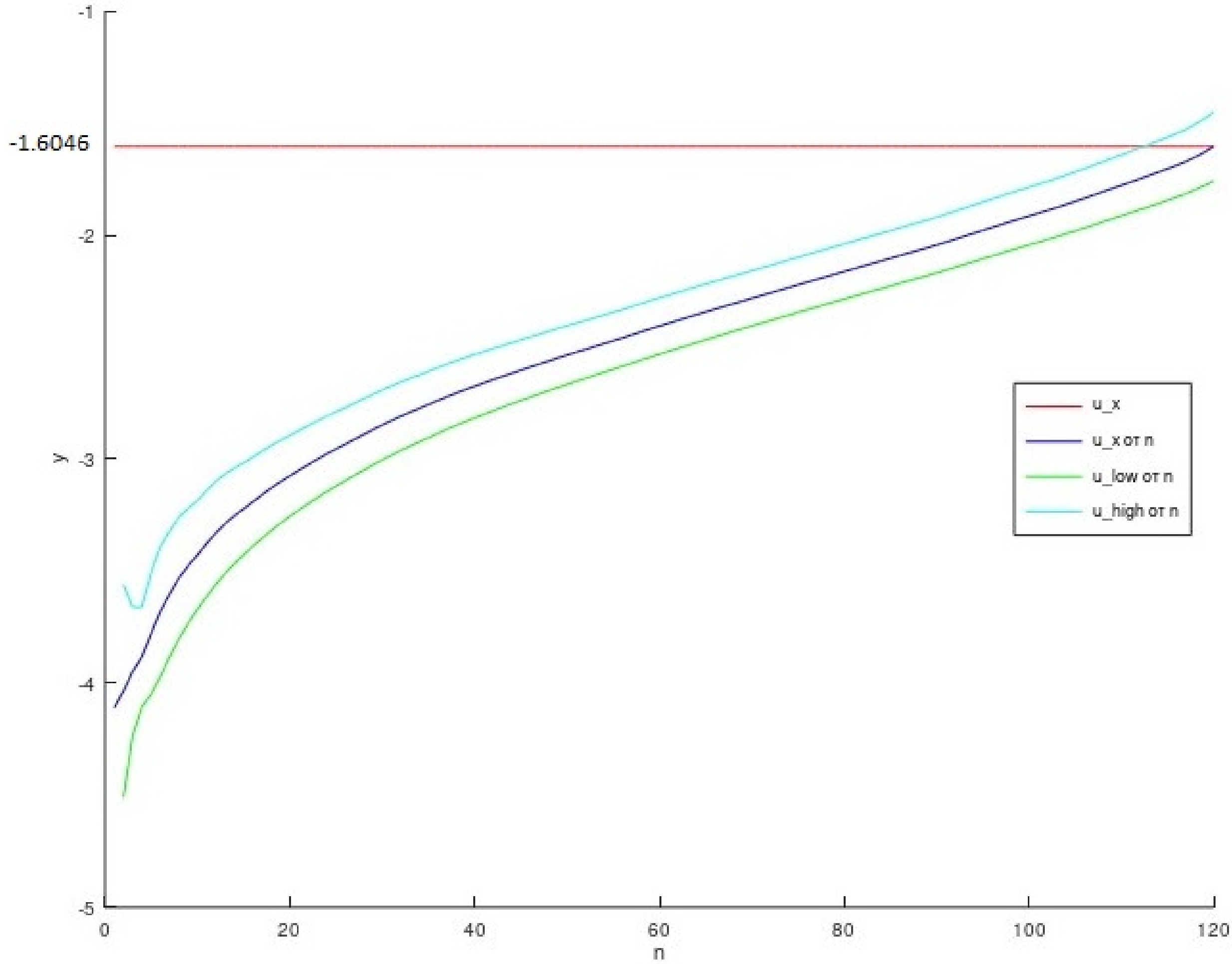
γ - уровень доверия

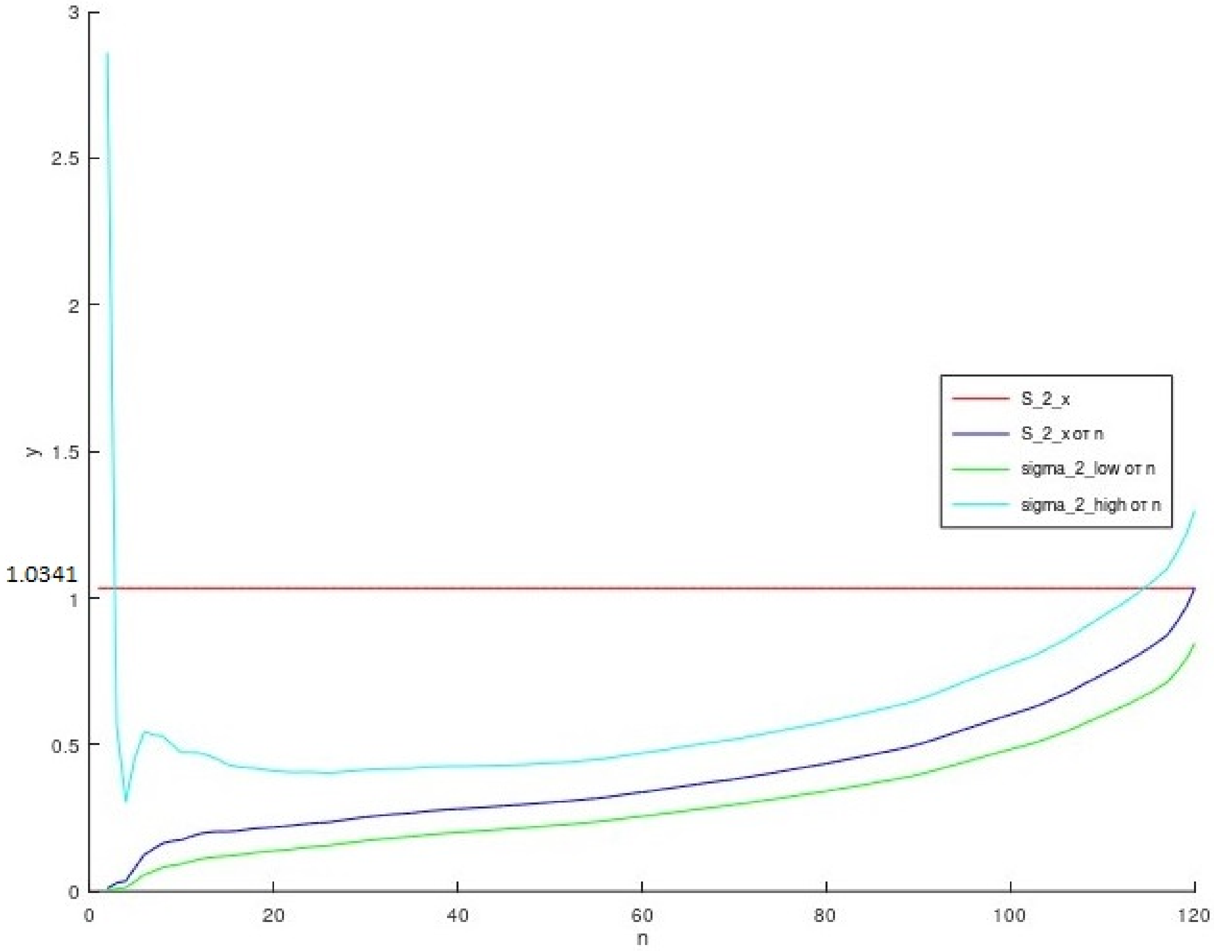
$t_{1-\alpha}(n-1)$ - квантиль уровня $1-\alpha$ для распределения Стьюдента с n - 1 степенями свободы, $\alpha = \frac{1-\gamma}{2}$

3 Практическая часть

3.1 Результаты работы для выборки по варианту

```
Lab 2
n = 120
u_x = −1.6046
S_2_x = 1.0341
u_low = −1.7585
u_high = −1.4507
sigma_2_low = 0.8460
sigma_2_high = 1.2979
```





3.2 Листинг программы

```
disp("Lab 2")
pkg load statistics

# Выборка
x_unsorted = [-0.23,-1.03,-4.11,-0.65,-2.58,-0.79,-1.53,-0.18,
-2.79,-1.97,-2.21,-1.59,-0.22,-3.18,-1.18,
-1.42,-1.29,-2.22,-0.82,-1.87,-2.30,-0.94,-0.74,-2.45,-1.40,
-2.09,-0.68,0.02,-1.80,-2.25,-1.19,-2.17,-1.89,
-1.14,-1.50,-1.76,-0.69,-2.21,-1.65,-1.51,-2.11,-2.24,-0.72,
0.94,-0.67,-2.44,-2.27,-1.33,-3.03,-0.42,-2.86,
-2.00,-1.37,-1.90,-2.80,-0.89,-2.04,-1.66,-0.14,-2.79,-0.21,
-1.29,-2.81,-0.29,-1.55,-0.45,-1.16,-3.96,-3.77,
-3.36,-1.81,0.13,-2.61,-3.69,-3.00,-2.61,-0.74,-0.41,-0.78,
-1.49,-1.89,-1.24,-0.00,-2.72,-1.69,-1.25,-1.59, 0.20,-1.08,
-2.42,-3.14,-2.54,-2.09,-2.51,-2.65,-2.42,-1.30,-0.65,1.40,
-2.33,-1.97,-0.54,-1.13,-2.04,0.77,-1.03,
-1.55,-1.47,-0.09,-2.11,-2.08,-1.79,-1.36,-1.92,-3.04,
-1.08,-1.67,-2.11,-1.99,-1.64];

# Сортировка массива данных
```

```
x = sort(x_unsorted);

# Нахождение длины массива
n = length(x)

# Оценка матожидания
function res = u_x_f(x)
  res = (1/numel(x))*sum(x);
endfunction
u_x = u_x_f(x)

# Несмещённая оценка дисперсии
function res = S_2_x_f(x, u_x)
  val = 0;
  for i = 1:numel(x)
    val +=(x(i) - u_x)*(x(i) - u_x);
  endfor
  res = val*(1/(numel(x) - 1));
endfunction
S_2_x = S_2_x_f(x, u_x)

gamma = 0.9;

# Нижняя граница доверительного интервала для матожидания
# tinv - функция распределения инверсии t студента
function res = u_low_f(u_x, S_2_x, n, gamma)
  alpha = (1 - gamma)/2;
  res = u_x - sqrt(S_2_x) / sqrt(n) * tinv(1 - alpha, n - 1);
endfunction

# Верхняя граница доверительного интервала для матожидания
function res = u_high_f(u_x, S_2_x, n, gamma)
  alpha = (1 - gamma)/2;
  res = u_x + sqrt(S_2_x) / sqrt(n) * tinv(1 - alpha, n - 1);
endfunction

u_low = u_low_f(u_x, S_2_x, n, gamma)
u_high = u_high_f(u_x, S_2_x, n, gamma)

# Нижняя граница доверительного интервала для дисперсии
function res = sigma_2_low_f(S_2_x, gamma, n)
  alpha = (1 - gamma)/2;
  res = ((n - 1) * S_2_x) / chi2inv(1 - alpha, n - 1);
endfunction

# Верхняя граница доверительного интервала для дисперсии
function res = sigma_2_high_f(S_2_x, gamma, n)
```

```
  alpha = (1 - gamma)/2;
  res = ((n - 1) * S_2_x)/chi2inv(alpha, n - 1);
endfunction

sigma_2_low = sigma_2_low_f(S_2_x, gamma, n)
sigma_2_high = sigma_2_high_f(S_2_x, gamma, n)

# Построение графиков
figure();
hold on;
x_n = 1:1:n;
for i = 1:numel(x)
  y(i) = u_x_f(x);
endfor
plot(x_n, y, 'r');
xlabel('n')
ylabel('y')

x_n = 1:1:n;
for i = 1:numel(x)
  y(i) = u_x_f(x(1:i));
endfor
plot(x_n, y, 'b');
xlabel('n')
ylabel('y')

for i = 1:numel(x)
  u_x = u_x_f(x(1:i));
  S_2_x = S_2_x_f(x(1:i), u_x);
  y(i) = u_low_f(u_x, S_2_x, i, gamma);
endfor

plot(x_n, y, 'g');
xlabel('n')
ylabel('y')

for i = 1:numel(x)
  u_x = u_x_f(x(1:i));
  S_2_x = S_2_x_f(x(1:i), u_x);
  y(i) = u_high_f(u_x, S_2_x, i, gamma);
endfor

plot(x_n, y, 'c');
xlabel('n')
ylabel('y')
legend({"u\\_x", "u\\_x от n", "u\\_low от n", "u\\_high от n"}, "location",
    "east");
```

```
hold off;


# Графики для дисперсии
figure();
hold on;
x_n = 1:1:n;
for i = 1:numel(x)
  y(i) = S_2_x_f(x, u_x);
endfor
plot(x_n, y, 'r');
xlabel('n')
ylabel('y')

x_n = 1:1:n;
for i = 1:numel(x)
  u_x = u_x_f(x(1:i));
  y(i) = S_2_x_f(x(1:i), u_x);
endfor
plot(x_n, y, 'b');
xlabel('n')
ylabel('y')

for i = 1:numel(x)
  u_x = u_x_f(x(1:i));
  S_2_x = S_2_x_f(x(1:i), u_x);
  y(i) = sigma_2_low_f(S_2_x, gamma, i);
endfor

plot(x_n, y, 'g');
xlabel('n')
ylabel('y')

for i = 1:numel(x)
  u_x = u_x_f(x(1:i));
  S_2_x = S_2_x_f(x(1:i), u_x);
  y(i) = sigma_2_high_f(S_2_x, gamma, i);
endfor

plot(x_n, y, 'c');
xlabel('n')
ylabel('y')
legend ({"S\\_2\\_x", "S\\_2\\_x от n", "sigma\\_2\\_low от n", "sigma\\_2\\_high от
    n"}, "location", "east");
hold off;
```