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

Отчёт по лабораторной работе №2
По дисциплине «Математическая статистика»
Тема: «Интервальные оценки»
Вариант: 1

Студент: ______________________________ Барсуков Н.М.

Группа: ______________________________ ИУ7-62

Преподаватель: ______________________________ Саркисян П.С.

Москва, 2020

Содержание

Введение

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

Содержание работы:

1) для выборки объёма n из нормальной 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 $\hat{\mu}$ и S^2 для выборки из индивидуального варианта;
3) для заданного пользователем уровня доверия γ и N - объёма выборки из индивидуального варианта:
   а) на координатной плоскости O_{yn} построить прямую $y = \hat{\mu}(\overrightarrow{x_N})$, также графики функций $y = \hat{\mu}(\overrightarrow{x_n})$, $y = \underline{\mu}(\overrightarrow{x_n})$ и $y = \overline{\mu}(\overrightarrow{x_n})$ как функций объёма n выборки, где n изменяется от 1 до N;
   б) на другой координатной плоскости O_{zn} 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 изменяетя от 1 до N.

Теоретическая часть

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

Предположим, что для параметра θ построен интервал $(\underline{\theta}(\overrightarrow{X_n}); \overline{\theta}(\overrightarrow{X_n}))$, где $\underline{\theta}(\overrightarrow{X_n})$ и $\overline{\theta}(\overrightarrow{X_n})$ являются функциями случайной выборки $\overrightarrow{X_n}$, такой, что выполняется равенство $P\{\underline{\theta}(\overrightarrow{X_n}) < \theta < \overline{\theta}(\overrightarrow{X_n})\} = \gamma$.

Интервал $(\underline{\theta}(\overrightarrow{X_n}); \overline{\theta}(\overrightarrow{X_n}))$ называют интервальной оценкой для параметра θ с коэффициентом доверия γ, а $\underline{\theta}(\overrightarrow{X_n})$ и $\overline{\theta}(\overrightarrow{X_n})$ называют, соответственно, нижней и верхней границами интервальной оценки. Интервальная оценка $(\underline{\theta}(\overrightarrow{X_n}); \overline{\theta}(\overrightarrow{X_n}))$ представляет собой интервал со случайными границами, который с заданной вероятностью γ накрывает неизвестное истинное значение параметра θ.

Определение: γ-доверительным интервалом (доверительным интервалом с коэффициентом вероятности γ) называют пару статистик $\underline{\theta}(\overrightarrow{X_n})$, $\overline{\theta}(\overrightarrow{X_n})$ таких, что для оцениваемого параметра θ $P\{\underline{\theta}(\overrightarrow{X_n}) < \theta < \overline{\theta}(\overrightarrow{X_n})\} \geq \gamma$.

Формулы для вычисления границ γ-доверительного интервала для математического ожидания и дисперсии нормальной случайной величины представлены в таблице:

Таблица 1. Формулы для вычисления границ γ-доверительного интервала для математического ожидания и дисперсии нормальной случайной величины

Параметры	Центральная статистика	Границы
μ - неизв., σ - изв.; оценить μ	$\frac{\mu-\overline{X}}{\sigma} * \sqrt{n} \sim N(0,1)$	$\underline{\mu}(\overrightarrow{X_n}) = \overline{X} + \frac{h_\alpha * \sigma}{\sqrt{n}}$ $\overline{\mu}(\overrightarrow{X_n}) = \overline{X} + \frac{h_{1-\alpha} * \sigma}{\sqrt{n}}$
μ, σ - неизвестны; оценить μ	$\frac{\mu-\overline{X}}{S^2(\overrightarrow{x_N})} * \sqrt{n} \sim St(n-1)$	$\underline{\mu}(\overrightarrow{X_n}) = \overline{X} - \frac{{h^t}_{1-\alpha} * \sigma}{\sqrt{n}}$ $\overline{\mu}(\overrightarrow{X_n}) = \overline{X} + \frac{{h^t}_{1-\alpha} * \sigma}{\sqrt{n}}$
μ, σ - неизвестны; оценить σ^2	$\frac{S^2(\overrightarrow{x_N})}{\sigma^2} * (n-1) \sim \chi^2(n-1)$	$\underline{\sigma^2}(\overrightarrow{X_n}) = \frac{S^2(\overrightarrow{X_n}) * (n-1)}{h_{1-\alpha}^{\chi^2(n-1)}}$ $\overline{\sigma^2}(\overrightarrow{X_n}) = \frac{S^2(\overrightarrow{X_n}) * (n-1)}{h_{\alpha}^{\chi^2(n-1)}}$

$\gamma = 1 - 2 * \alpha$, h_α - квантиль уровня α нормального распределения, ${h^t}_\alpha$ - квантиль уровня α распределения Стьюдента, $h_\alpha^{\chi^2(n-1)}$ - квантиль уровня α распределения χ^2.

Определение: статистика $g(\overrightarrow{X}, \theta)$ называется центральной, если закон её распределения не зависит от закона распределения θ, то есть функция распределения $F_g(X)$ не зависит от θ.

Пусть $\overrightarrow{X_n}$ - случайная выборка объёма n из генеральной совокупности X, распределённой по нормальному закону с параметрами μ и σ^2. Тогда оценить математическое ожидание при неизвестной дисперсии можно следующим образом:

$$\underline{\mu}(\overrightarrow{X_n}) = \overline{X} - \frac{S(\overrightarrow{X_n}) * t_{1-\alpha}(n-1)}{\sqrt{n}}; \quad (1)$$

$$\overline{\mu}(\overrightarrow{X_n}) = \overline{X} + \frac{S(\overrightarrow{X_n}) * t_{1-\alpha}(n-1)}{\sqrt{n}}, \quad (2)$$

где $\overline{X}$ - оценка математического ожидания, n - число опытов, $S(\overrightarrow{X_n})$ - точечная оценка дисперсии, $t_{1-\alpha}(n-1)$ - квантиль уровня (1 - α) для распределения Стьюдента с (n - 1) степенями свободы, $\alpha = \frac{1-\gamma}{2}$.

Дисперсию в данном случае можно оценить следующим образом:

$$\underline{\sigma^2}(\overrightarrow{X_n}) = \frac{S^2(\overrightarrow{X_n}) * (n-1)}{\chi^2_{1-\alpha}(n-1)}; \quad (3)$$

$$\overline{\sigma^2}(\overrightarrow{X_n}) = \frac{S^2(\overrightarrow{X_n}) * (n-1)}{\chi^2_{\alpha}(n-1)}, \quad (4)$$

где n - объём выборки, $\chi^2_\alpha(n-1)$ - квантиль уровня α для распределения χ^2 с (n - 1) степенями свободы, $\alpha = \frac{1-\gamma}{2}$.

Листинг программы

Текст программы (*labwork2.m*):

```
function lab2()
X = [-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]


N = 1:length(X);

gamma = 0.9;
alpha = (1 - gamma)/2;

mu = expectation(X);
sSqr = variance(X);

fprintf('mu = %.4f\n', mu);
fprintf('S^2 = %.4f\n\n', sSqr);

muArray = expectationArray(X, N);
varArray = varianceArray(X, N);

figure
plot([N(1), N(end)], [mu, mu], 'm');
hold on;
plot(N, muArray, 'g');

Ml = muArray - sqrt(varArray./N).*tinv(1 - alpha, N - 1);
plot(N, Ml, 'b');

fprintf('mu_low = %.4f\n', Ml(end));
```

```
Mh = muArray + sqrt(varArray./N).*tinv(1 - alpha, N - 1);
plot(N, Mh, 'r'), legend('y=mu', 'y=mu_n', 'y=mu-low_n', 'y=mu-high_n');
grid on;
hold off;

fprintf('mu_high = %.4f\n', Mh(end));

figure
plot([N(1), N(end)], [sSqr, sSqr], 'm');
hold on;
plot(N, varArray, 'g');

Sl = varArray.*(N - 1)./chi2inv(1 - alpha, N - 1);
plot(N, Sl, 'b');

Sh = varArray.*(N - 1)./chi2inv(alpha, N - 1);
plot(N, Sh, 'r'), legend('z=S^2', 'z=S^2_n', 'z=S^2-low_n', 'z=S^2-high_n');
grid on;
hold off;


fprintf('sigma^2_low = %.4f\n', Sl(end));
fprintf('sigma^2_high = %.4f\n', Sh(end));
end

function mu = expectation(X)
mu = mean(X);
end

function sSqr = variance(X)
sSqr = var(X);
end

function muArray = expectationArray(X, N)
muArray = zeros(1, length(N));
for i = 1:length(N)
muArray(i) = expectation(X(1:N(i)));
end
end

function varArray = varianceArray(X, N)
varArray = zeros(1, length(N));
for i = 1:length(N)
```

```
varArray(i) = variance(X(1:N(i)));
end
end
```

Результаты расчётов для выборки из индивидуального варианта

Для выборки объёма 120 и $\gamma = 0.9$ согласно варианту были получены:

1) точечные оценки математического ожидания $\hat{\mu}(\overrightarrow{x_n})$ и дисперсии $S^2(\overrightarrow{x_n})$;

2) нижняя и верхняя границы γ-доверительного интервала для математического ожидания $(\underline{\mu}(\overrightarrow{x_n}); \overline{\mu}(\overrightarrow{x_n})) = (-1.7585; -1.4507)$;

3) нижняя и верхняя границы γ-доверительного интервала для дисперсии $(\underline{\sigma}^2(\overrightarrow{x_n}); \overline{\sigma}^2(\overrightarrow{x_n})) = (0.8460; 1.2979)$;

4) выборочное среднее $\hat{\mu} = -1.6046$;

5) несмещённая выборочная дисперсия $S^2 = 1.0341$.

В результате проведённых вычислений для выборки объёма 120 и $\gamma = 0.9$ согласно варианту были построены соответствующие графики. На графике 1 представлены прямая $y = \hat{\mu}(\overrightarrow{x_N})$ (голубая линия), полученное значение математического ожидания $y = \hat{\mu}(\overrightarrow{x_n})$ (красная линия), нижняя граница доверительного интервала $y = \underline{\mu}(\overrightarrow{x_n})$ (жёлтая линия) и верхняя граница доверительного интервала $y = \overline{\mu}(\overrightarrow{x_n})$ (фиолетовая линия), на графике 2 - прямая $z = S^2(\overrightarrow{x_N})$ (голубая линия), полученное значение несмещённой дисперсии $z = S^2(\overrightarrow{x_n})$ (красная линия), нижняя интервальная граница $z = \underline{\sigma}^2(\overrightarrow{x_n})$ (жёлтая линия) и верхняя интервальная граница $z = \overline{\sigma}^2(\overrightarrow{x_n})$ (фиолетовая линия).

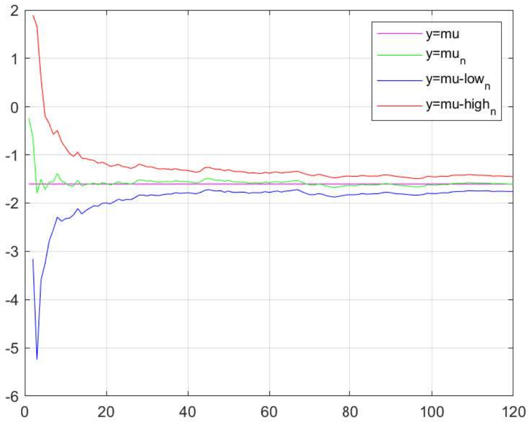

Рис. 1. График доверительного интервала для математического ожидания

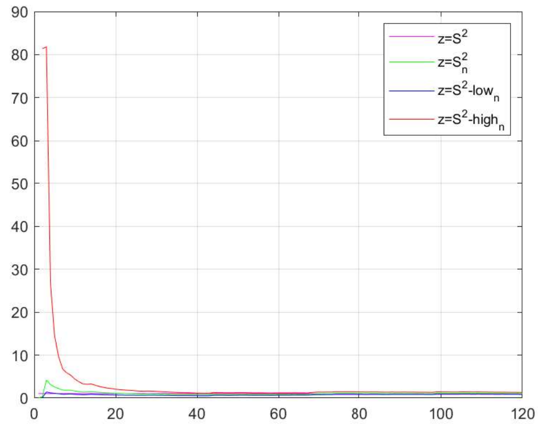

Рис. 2. График доверительного интервала для дисперсии

Заключение

В результате выполнения лабораторной работы для заданной согласно варианту выборки были получены точечные оценки математического ожидания и дисперсии, границы γ-доверительных интервалов математического ожидания и дисперсии, выборочное среднее, несмещённая выборочная дисперсия. Для выполнения вычислений был написан код MatLab, позволяющий вычислить искомые величины и построить графики γ-доверительных интервалов.