Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

Отчет по лабораторной работе № 8
по дисциплине: Математическая статика.

Выполнила студентка:
Заболотских Екатерина Дмитриевна
группа: 3630102/70301

Проверил:
к.ф.-м.н., доцент
Баженов Александр Николаевич

Санкт-Петербург
2020 г.

Оглавление

Постановка задачи ... 3
Теория ... 4
Интервальное оценивание ... 4
Классическое оценивание ... 5
Для математического ожидания m ... 5
Для среднего квадратичного отклонения σ ... 5
Асимптотически нормальные оценки ... 5
Для математического ожидания m ... 5
Для среднего квадратичного отклонения σ ... 6
Реализация ... 7
Результаты ... 8
Классические оценки ... 8
Асимптотически нормальные оценки ... 8
Обсуждение ... 9
Список литературы ... 10

Список таблиц

Таблица 1: Классические оценки ... 8
Таблица 2: Асимптотически нормальные оценки ... 8

Постановка задачи

Для двух выборок размерами 20 и 100 элементов, сгенерированных согласно нормальному закону $N(x, 0, 1)$ для параметров положения масштаба построить асимптотически нормальные интервальные оценки на основе точечных оценок метода максимального правдоподобия и классические интервальные оценки на основе статистик χ^2 и Стьюдента. В качестве надёжности взять $\gamma = 0.95$.

Теория

Интервальное оценивание

Интервальной оценкой (доверительным интервалом) числовой характеристики или параметра распределения θ генеральной совокупности с доверительной вероятностью γ называется интервал (θ_1, θ_2), границы которого являются случайными функциями $\theta_1 = \theta_1(x_1, \dots x_n)$, который покрывает θ с вероятностью γ:

$$P(\theta_1 < \theta < \theta_2) = \gamma \tag{1}$$

Часто вместо доверительной вероятности γ рассматривается уровень значимости $\alpha = 1 - \gamma$. Важной характеристикой данной интервальной оценки является половина длины доверительного интервала, она называется точностью интервального оценивания.

$$\Delta = \frac{\theta_1 - \theta_2}{2} \tag{2}$$

<u>Общий вид интервальных оценок:</u>

Пусть известна статистика $Y(\hat{\theta}, \theta)$, содержащая оцениваемый параметр θ и его точечную оценку $\hat{\theta}$. Функция $Y(\hat{\theta}, \theta)$ непрерывна и строго монотонна (для определенности строго возрастает) по θ. Известна функция распределения $F_Y(x)$, и она зависит от θ.

Зададим уровень значимости α и будем строить доверительный интервал так, чтобы $(-\infty, \alpha_1), (\alpha_2, \infty)$ накрывали θ с вероятностью $\frac{\alpha}{2}$.

Пусть $y_{\alpha/2}, y_{1-\alpha/2}$ – квантили распределения Y соответствующих порядков, тогда:

$$\begin{aligned} P\left(y_{\alpha/2} < Y(\hat{\theta}, \theta) < y_{1-\alpha/2}\right) &= F_Y\left(y_{1-\alpha/2}\right) - F_Y\left(y_{\alpha/2}\right) \\ &= 1 - \alpha/2 - \alpha/2 = 1 - \alpha = \gamma \end{aligned} \tag{3}$$

Так как $Y(\hat{\theta}, \theta)$ строго возрастает по θ, то у нее есть обратная функция $Y^{-1}(y)$. Она в свою очередь строго возрастает и тоже зависит от θ, следовательно:

$$\begin{aligned} &y_{\alpha/2} < Y(\hat{\theta}, \theta) < y_{1-\alpha/2} \\ &Y^{-1}\left(y_{\alpha/2}\right) < \theta < Y^{-1}\left(y_{1-\alpha/2}\right) \end{aligned} \tag{4}$$

Получаем границы интервала: $\theta_1 = Y^{-1}\left(y_{\alpha/2}\right), \theta_2 = Y^{-1}\left(y_{1-\alpha/2}\right)$.

Классическое оценивание

Для математического ожидания m

Доказано, что случайная величина

$$T = \sqrt{n-1} * \frac{\overline{x} - m}{s} \tag{5}$$

называется статистикой Стьюдента, распределена по закону Стьюдента с $n-1$ степенями свободы. После некоторых выкладок имеем оценки границ интервала:

$$\begin{gathered} m_1 = \bar{x} - \frac{xt_{1-\alpha/2}(n-1)}{\sqrt{n-1}} \\ m_2 = \bar{x} + \frac{xt_{1-\alpha/2}(n-1)}{\sqrt{n-1}} \end{gathered} \tag{6}$$

, где $t_{1-\alpha/2}(n-1)$ – квантиль порядка $1-\alpha/2$ распределения Стьюдента с n – 1 степенями свободы.

Для среднего квадратичного отклонения σ

Доказано, что случайная величина ${ns^2}/{\sigma^2}$ распределена по закону χ^2 с n – 1 степенями свободы. Применяя общий метод построения интервальных оценок, получаем оценки границ интервала:

$$\begin{gathered} \sigma_1 = \frac{s\sqrt{n}}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}} \\ \sigma_2 = \frac{s\sqrt{n}}{\sqrt{\chi^2_{\alpha/2}(n-1)}} \end{gathered} \tag{7}$$

, где $\chi^2_{1-\alpha/2}(n-1), \chi^2_{\alpha/2}(n-1)$ – квантили соответствующих порядков χ^2-распределения с n – 1 степенями свободы.

Асимптотически нормальные оценки

Для математического ожидания m

В силу центральной предельной теоремы, центрированная и нормированная случайная величина ${\sqrt{n}(\bar{x}-m)}/{\sigma}$ распределена приблизительно нормально с параметрами 0 и 1. Исходя из этого получаем:

$$\begin{gathered} m_1 = \bar{x} - \frac{su_{1-\alpha/2}}{\sqrt{n}} \\ m_2 = \bar{x} + \frac{su_{1-\alpha/2}}{\sqrt{n}} \end{gathered} \tag{8}$$

, где $u_{1-\alpha/2}$ – квантиль нормального распределения $\mathcal{N}(0,1)$ порядка $1-\alpha/2$

<u>Для среднего квадратичного отклонения σ</u>

Аналогично, в силу ЦПТ, центрированная и нормированная случайная величина $(s^2 - M_{s^2})\big/\sqrt{D_{s^2}}$ при большом объеме выборки распределена приблизительно нормально с параметрами 0 и 1. Исходя из этого получаем оценку:

$$\begin{aligned} \sigma_1 &= s\left(1 + u_{1-\alpha/2}\sqrt{{(e+2)}/{n}}\right)^{-1/2} \\ \sigma_2 &= s\left(1 - u_{1-\alpha/2}\sqrt{{(e+2)}/{n}}\right)^{-1/2} \end{aligned} \tag{9}$$

, где e – выборочный эксцесс, определяемый по формуле:

$$e = \frac{m_4}{s^4} - 3 \tag{10}$$

, где m_4 – четвертый выборочный центральный момент, определяемый по формуле:

$$m_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{s})^4 \tag{11}$$

Реализация

Код программы, реализующий данную задачу, был написан на языке Python в интегрированной среде разработке PyCharm.

Были использованы библиотеки:

- **SciPy** – модуль “stats” для генерации данных, и встроенных вычислений.

Результаты

Классические оценки

	m (6)	σ (7)
n= 20	-0.57 < m < 0.25	0.67 < σ < 1.28
n = 100	-0.12 < m < 0.27	0.86 < σ < 1.14

Таблица 1: Классические оценки

Асимптотически нормальные оценки

	m (8)	σ (9)
n= 20	-0.54 < m < 0.21	0.69 < σ < 1.23
n = 100	-0.12 < m < 0.27	0.86 < σ < 1.14

Таблица 2: Асимптотически нормальные оценки

Обсуждение

Результаты показывают, что значения $m = 0, \sigma = 1$ лежат в соответствующих интервалах с вероятностью 0.95. Интервалы действительно накрывают значения параметров, причем при увеличении n асимптотические оценки почти совпадают с классическими.

Список литературы

1. Конспекты лекции
2. Википедия: https://ru.wikipedia.org/wiki

Ссылка на github: https://github.com/KateZabolotskih/MathStatLabs