Санкт-Петербургский политехнический университет Петра Великого
Институт прикладной математики и механики
Кафедра прикладной математики

ОТЧЕТ

Тема: ***Интервальные оценки характеристик распределения***

Направление: 01.03.02 Прикладная математика и информатика

Выполнил студент гр. 33631/4 Камалетдинова Ю.

Преподаватель Баженов А.

Санкт-Петербург
2019

Содержание

Постановка задачи

Рассматриваются методы интервальных оценок характеристик распределения. Требуется сгенерировать выборки объемами $n = 20,\ 100$ элементов для нормального распределения $N(x; 0, 1)$, затем для параметров положения и масштаба построить асимптотически нормальные интервальные оценки на основе точечных оценок метода максимального правдоподобия. Также необходимо оценить параметры распределения на основе статистик χ^2 и Стьюдента. В качестве параметра надежности взять $\gamma = 0.95$.

Описание алгоритма

Оценка на основе статистик Стьюдента и хи-квадрат

Пусть $x_1, \ldots, x_n$ — заданная выборка из нормального распределения $N(x; \mu, \sigma)$, по которой требуется оценить параметры μ, σ, генерального распределения. Построим на ее основе выборочные среднее $\overline{x}$ и среднее квадратическое отклонение s. Параметры распределения μ, σ не известны. В источнике [1] показано, что статистика Стьюдента

$$T = \sqrt{n-1}\,\frac{\overline{x}-\mu}{s} \tag{1}$$

распределена по закону Стьюдента с $n-1$ степенями свободы. Пусть $f_T(x)$ — плотность вероятности данного распределения. Тогда

$$\begin{aligned} P(-x < \sqrt{n-1}\,\frac{\overline{x}-\mu}{s} < x) = P(-x < \sqrt{n-1}\,\frac{\mu-\overline{x}}{s} < x) = \\ = \int_{-x}^{x} f_T(t)dt = 2F_T(x) - 1\ , \end{aligned} \tag{2}$$

где $F_T(t)$ — функция распределения Стьюдента с $n-1$ степенями свободы. Положим $2F_T(x) - 1 = 1 - \alpha$, где α — выбранный уровень значимости. Тогда $F_T(x) = 1 - \alpha/2$. Положим $t_{1-\alpha/2}(n-1)$ — квантиль распределения Стьюдента с $n-1$ степенями свободы и уровнем значимости $1-\alpha/2$. Из (1), (2) получаем

$$P\left(\overline{x}-\frac{st_{1-\alpha/2}(n-1)}{\sqrt{n-1}}<\mu<\overline{x}+\frac{st_{1-\alpha/2}(n-1)}{\sqrt{n-1}}\right)=1-\alpha\ , \tag{3}$$

что дает доверительный интервал для μ с вероятностью $\gamma=1-\alpha$

Для поиска оценки параметра σ воспользуемся источником [1], где показано, что случайная величина ns^2/σ^2 распределена по закону χ^2 с $n-1$ степенями свободы. Найдем квантили $\chi^2_{\alpha/2}(n-1)$, $\chi^2_{1-\alpha/2}(n-1)$ и приведем выражение для доверительного интервала для σ с доверительной вероятностью $\gamma=1-\alpha$

$$P\left(\frac{s\sqrt{n}}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}<\sigma<\frac{s\sqrt{n}}{\sqrt{\chi^2_{\alpha/2}(n-1)}}\right)=1-\alpha \tag{4}$$

Асимптотический подход при построении оценок

Данный метод оценивания параметров применяется в случае неизвестности закона распределения, или когда он не является нормальным. Асимптотический метод построения доверительных интервалов основан на центральный предельной теореме.

Пусть $\overline{x}$ — выборочное среднее из выборки большого объема n независимых одинаково распределенных случайных величин. Тогда в силу центральной предельной теоремы случайная величина $(\overline{x}-M\overline{x})/\sqrt{D\overline{x}}=\sqrt{n}(\overline{x}-\mu)/\sigma$ распределена приблизительно нормально с параметрами 0, 1. Из данных рассуждений получим выражение для доверительного интервала для μ с доверительной вероятностью $\gamma=1-\alpha$

$$P\left(\overline{x}-\frac{su_{1-\alpha/2}}{\sqrt{n}}<\mu<\overline{x}+\frac{su_{1-\alpha/2}}{\sqrt{n}}\right)\approx\gamma\ , \tag{5}$$

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

Приведем выражение для доверительного интервала для σ с доверительной вероятностью $\gamma=1-\alpha$

$$s\,(1-0.5U)<\sigma<s\,(1+0.5U)\ , \tag{6}$$

где $U = u_{1-\alpha/2}\sqrt{(e+2)/n}$; e — выборочный эксцесс; m_4 — четвертый выборочный центральный момент.

Формулы (1) — (6) и определения взяты из источника [1]

Реализация

Для выполнения поставленной задачи будем пользоваться библиотеками для языка Python: *numpy, scipy* – расчеты, законы распределения вероятностей; *matplotlib, seaborn* – визуализация результатов. Ход работы:

- Генерируем выборки из распределения $N(0,1)$ объемами $n = 20,\ 100$
- Вычисляем выборочные среднее, дисперсию, четвертый центральный момент, эксцесс по приведенным ниже формулам

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{7}$$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{8}$$

$$m_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4 \tag{9}$$

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

- Вычисляем границы доверительных интервалов по формулам (3), (4), (5), (6)

Результат

Представим табулированные значения квантилей распределений необходимых порядков, $\alpha = 0.05$

- $t_{0.95}(19) = 1.72,\ t_{0.95}(99) = 1.66$ — квантили распределения Стьюдента
- $\chi^2_{0.025}(19) = 8.91,\ \chi^2_{0.975}(19) = 32.85,\ \chi^2_{0.025}(99) = 73.12,\ \chi^2_{0.975}(99) = 128.4$ — квантили распределения хи-квадрат
- $u_{0.975} = 1.96$ — квантиль стандартного нормального распределения

n	Интервал для μ	Интервал для σ
20	$(-0.239; 0.686)$	$(0.891; 1.711)$
100	$(-0.067; 0.266)$	$(0.876; 1.160)$

Таблица 1: Таблица оценок на основе статистик Стьюдента и хи-квадрат

n	Интервал для μ	Интервал для σ
20	$(-0.290; 0.737)$	$(0.873; 1.471)$
100	$(-0.096; 0.295)$	$(0.868; 1.126)$

Таблица 2: Таблица оценок на основе на основе асимптотического подхода

Вывод

По полученным результатам можно судить о том, что асимптотический подход не имеет преимущества по обоим параметрам сразу в случае малой выборки ($n = 20$) при условиях, что закон распределения известен и является нормальным. При объеме выборки $n = 100$ можно заметить сокращение длин доверительных интервалов $(0.333 < 0.391$

для μ, $0.258 < 0.284$ для σ), что является преимуществом асимптотического подхода в оценке параметров распределения.

Недостатком интервальных оценок на основе статистик Стьюдента и хи-квадрат может выступать сложность получения точечных оценок параметра распределения, если оно не является нормальным, что в с свою очередь усложнит вычисления.

Список литературы

[1] *Амосова Н.Н., Куклин Б.А., Макарова С.Б., Максимов Ю.Д., Митрофанова Н.М., Полищук В.И., Шевляков Г.Л.* Вероятностные разделы математики. Учебник для бакалавров технических направлений. — СПб.: Иван Федоров, 2001. — 592 с.: илл. — ISBN 5-81940-050-X.