Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

Отчет по лабораторной работе № 2
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

Список таблиц

Постановка задачи

Для каждого из 5 распределений:

1. Нормального $\mathcal{N}(x, 0, 1)$
2. Коши $\mathcal{C}(x, 0, 1)$
3. Лапласа $\mathcal{L}(x, 0, \frac{1}{\sqrt{2}})$
4. Пуассона $\mathcal{P}\ (k,\ 10)$
5. Равномерного $\mathcal{U}\left(x, -\sqrt{3}, \sqrt{3}\right)$

Сгенерировать выборки размеров: 10, 100, 1000. Вычислить следующие статистические характеристики для каждой выборки:

$$\bar{x}, med, x, z_R, z_Q, z_{tr}.$$

Для каждой выборки провести подобные вычисления по 1000 раз и найти среднее значение их характеристик положения и их квадратов:

$$E(z) = \bar{z}; \tag{1}$$

Вычислить оценку дисперсии:

$$D(z) = \overline{z^2} - \bar{z}^2; \tag{2}$$

Представить полученные данные в виде таблицы.

Теория

1. Вариационный ряд

Вариационный ряд (или упорядоченная выборка) – последовательность

$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n-1)} \leq X_{(n)}$,

состоящая из одинаково распределенных случайных величин: $X_1, X_2, X_3, \cdots, X_n$, расположенных в неубывающем порядке. (Может быть получена из исходной выборки, в результате расположения элементов в неубывающем порядке)

Значение k-го элемента вариационного ряда $x_{(k)}$ называется k-ой порядковой статистикой.

2. Выборочные числовые характеристики

Математическая статистика рассматривает приближенные методы отыскания законов распределения и числовых характеристик по результатам экспериментов.

Выборочные характеристики – случайные величины как борелевские функции от случайных величин.

Рассматриваются следующе характеристики:

- Выборочное среднее:

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{n} x_i \tag{3}$$

- Выборочная медиана:

$$med\ x = \begin{cases} x_{(l+1)} & \text{при} \quad n = 2l + 1 \\ \dfrac{x_{(l)} + x_{(l+1)}}{2} & \text{при} \quad n = 2l \end{cases} \tag{4}$$

- Полусумма экстремальных выборочных элементов:

$$z_R = \frac{x_{(1)} + x_{(n)}}{2} \tag{5}$$

- Выборочный квантиль порядка p:

$$z_p = \begin{cases} x_{([np]+1)} & \text{при} \quad np \ \text{дробном} \\ x_{(np)} & \text{при} \quad np \ \text{целом} \end{cases} \tag{6}$$

- Полусумма квантилей:

$$z_Q = \frac{z_{0.25} + z_{0.75}}{2} \tag{7}$$

- Усечённое среднее:

$$z_{tr} = \frac{1}{n-2r}\sum_{i=r+1}^{n-r} x_{(i)}\text{, где } r = \left\lceil\frac{n}{4}\right\rceil \tag{8}$$

Реализация

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

Были использованы библиотеки:

- **Numpy** – для работы с данными и построения выборок.

Результаты

	$\bar{x}(3)$	$med\ x(4)$	$z_R(5)$	$z_Q(7)$	$z_{tr}(8)$
$n = 10$					
$E(z)$	-0.01	0.0	0.0	0.3	0.0
$D(z)$	0.090461	0.145400	0.190432	0.122063	0.1860126
$n = 100$					
$E(z)$	0.001	0.00	-0.00	0.02	-0.00
$D(z)$	0.0092806	0.0162908	0.09482238	0.01335049	0.0214439
$n = 1000$					
$E(z)$	-1.147	0.000	-0.00	0.001	0.003
$D(z)$	0.00101	0.00156	0.061354	0.0012	0.0019

Таблица 1: Нормальное распределение

	$\bar{x}(3)$	$med\ x(4)$	$z_R(5)$	$z_Q(7)$	$z_{tr}(8)$
$n = 10$					
$E(z)$	-4.8	-0.0	-14.4	1.2	0.2
$D(z)$	22978.4851	0.35161	134103.5860	5.9101	304.4308
$n = 100$					
$E(z)$	-3.0	-0.00	51.4	0.04	0.3
$D(z)$	22963.8426	0.0224	3810135.8719	0.05342	1000.4721
$n = 1000$					
$E(z)$	1.8	0.001	819.1	0.005	5.5
$D(z)$	10823.6848	0.0025	1589854120.7	0.0055	25780.6205

Таблица 2: Распределение Коши

	$\bar{x}$(3)	$med\ x$(4)	z_R(5)	z_Q(7)	z_{tr}(8)
$n = 10$					
$E(z)$	0.0	0.00	0.0	0.3	-0.0
$D(z)$	0.10458	0.07528	0.39273	0.123057	0.192249
$n = 100$					
$E(z)$	0.002	-0.001	0.0	0.02	0.00
$D(z)$	0.009718	0.005793	0.398961	0.0110906	0.02108
$n = 1000$					
$E(z)$	-0.002	-0.00	0.02	0.001	0.00
$D(z)$	0.001094	0.000493	0.439785	0.000954	0.002094

Таблица 3: Распределение Лапласа

	$\bar{x}$(3)	$med\ x$(4)	z_R(5)	z_Q(7)	z_{tr}(8)
$n = 10$					
$E(z)$	10	9.8	10.3	11	10
$D(z)$	0.96096	1.5011	1.9474	1.4409	2.0389
$n = 100$					
$E(z)$	10.0	9.9	10	10	9.9
$D(z)$	0.10364	0.19059	0.96729	0.16532	0.20728
$n = 1000$					
$E(z)$	9.996	9.99	11.7	9.994	10.00
$D(z)$	0.0099	0.0091	0.61238	0.00320	0.0190

Таблица 4: Распределение Пуассона

	$\bar{x}(9)$	$med\ x$	z_R	z_Q	z_{tr}
$n = 10$					
$E(z)$	0.01	-0.0	-0.00	0.3	-0.0
$D(z)$	0.098	0.2237	0.05196	0.13023	0.20549
$n = 100$					
$E(z)$	-0.003	-0.00	0.0014	0.01	-0.00
$D(z)$	0.0093	0.0295	0.000543	0.015398	0.0205049
$n = 1000$					
$E(z)$	0.0008	0.001	-8e-06	0.001	0.000
$D(z)$	0.000983	0.00318	6e-06	0.001490	0.00193

Таблица 5: Равномерное распределение

Заключение

Исследование показало:

- действительно, математическое ожидание для распределения Коши не определено, а дисперсия – бесконечна;
- выборочное среднее при увеличении n стремится к математическому ожиданию;
- медиана у всех распределений определена;
- z_Q и z_R оценивают центр симметрии распределения

Упорядочение характеристик

- Нормального $\mathcal{N}(x, 0, 1)$: $\bar{x} < z_R < med\ x < z_Q < z_{tr}$
- Коши $\mathcal{C}(x, 0, 1)$: $med\ x < z_Q < \bar{x} < z_{tr} < z_R$
- Лапласа $\mathcal{L}(x, 0, \frac{1}{\sqrt{2}})$: $\bar{x} < med\ x < z_{tr} < z_Q < z_R$
- Пуассона $\mathcal{P}\ (k,\ 10)$: $med\ x < z_Q < \bar{x} < z_{tr} < z_R$
- Равномерного $\mathcal{U}\left(x, -\sqrt{3}, \sqrt{3}\right)$: $z_R < z_{tr} < \bar{x} < med\ x \leq z_Q$

Список литературы

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

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