

Санкт-Петербургский Политехнический Университет имени Петра Великого
Институт Прикладной Математики и Механики
Кафедра "Прикладная Математика"

Отчет по лабораторной работе №2
по дисциплине
"Математическая Статистика"

Выполнил студент:
Тырыкин Я. А.
группа 3630102/80401
Проверил:
к.ф.-м.н., доцент
Баженов А. Н.

Санкт-Петербург
2021

Содержание

1	Постановка задачи	4
2	Теория	4
2.1	Распределения	4
2.2	Выборочные числовые характеристики	5
2.2.1	Характеристики положения	5
2.2.2	Характеристики рассеяния	5
3	Модульная структура программы	6
4	Результаты	6
4.1	Характеристики положения и рассеяния	6
5	Обсуждение	11
5.1	Характеристики положения и рассеяния	11
6	Ресурсы	11

Список таблиц

1	Характеристики нормального распределения (3)	7
2	Характеристики распределения Коши (4)	8
3	Характеристики распределения Лапласа (5)	9
4	Характеристики распределения Пуассона (6)	10
5	Характеристики равномерного распределения (7)	11

1 Постановка задачи

Для 5 распределений:

1. Нормальное распределение $N(x, 0, 1)$
2. Распределение Коши $C(x, 0, 1)$
3. Распределение Лапласа $L(x, 0, \frac{1}{\sqrt{2}})$
4. Распределение Пуассона $P(k, 10)$
5. Равномерное распределение $U(x, -\sqrt{3}, \sqrt{3})$

Сгенерировать выборки размером 10, 100 и 1000 элементов. Для каждой выборки вычислить следующие статистические характеристики положения данных: \bar{x} , med , x , z_R , z_Q , z_{tr} . Повторить такие вычисления 1000 раз для каждой выборки и найти среднее характеристик положения и их квадратов:

$$E(z) = \bar{z} \quad (1)$$

Вычислить оценку дисперсии по формуле:

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

Представить полученные данные в виде таблиц.

2 Теория

2.1 Распределения

- Нормальное распределение:

$$N(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (3)$$

- Распределение Коши:

$$C(x, 0, 1) = \frac{1}{\pi} \frac{1}{x^2 + 1} \quad (4)$$

- Распределение Лапласа:

$$L(x, 0, \frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|} \quad (5)$$

- Распределение Пуассона:

$$P(k, 10) = \frac{10^k}{k!} e^{-10} \quad (6)$$

- Равномерное распределение:

$$U(x, -\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{при } |x| \leq \sqrt{3} \\ 0 & \text{при } |x| > \sqrt{3} \end{cases} \quad (7)$$

2.2 Выборочные числовые характеристики

2.2.1 Характеристики положения

- Выборочное среднее:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (8)$$

- Выборочная медиана:

$$\text{med } x = \begin{cases} x_{(l+1)} & \text{при } n = 2l + 1 \\ \frac{x_{(l)} + x_{(l+1)}}{2} & \text{при } n = 2l \end{cases} \quad (9)$$

- Полусумма экстремальных выборочных элементов:

$$z_R = \frac{x_{(1)} + x_{(n)}}{2} \quad (10)$$

- Полусумма квартилей:

Выборочная квартиль z_p порядка p определяется формулой

$$z_p = \begin{cases} x_{([np]+1)} & \text{при } np \text{ дробном,} \\ x_{([np])} & \text{при } np \text{ целом.} \end{cases} \quad (11)$$

Полусумма квартилей

$$z_Q = \frac{x_{1/4} + x_{3/4}}{2} \quad (12)$$

- Усечённое среднее:

$$z_{tr} = \frac{1}{n-2r} \sum_{i=r+1}^{n-r} x_{(i)}, r \approx \frac{n}{4} \quad (13)$$

2.2.2 Характеристики рассеяния

Выборочная дисперсия

$$D(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (14)$$

3 Модульная структура программы

Лабораторная работа выполнена с применением средств языка Python версии 3.7 в среде разработки PyCharm IDE (в частности, с применением встроенных методов библиотеки SciPy и Numpy). Исходный код лабораторной работы находится по ссылке в приложении к отчёту.

4 Результаты

4.1 Характеристики положения и рассеяния

Как было проведено округление:

В оценке $x = E \pm D$ вариации подлежит первая цифра после точки.

В данном случае $x = 0.0 \pm 0.1k$,

k зависит от доверительной вероятности и вида распределения (рассматривается в дальнейшем цикле лабораторных работ)

Округление сделано для $k = 1$

normal n = 10					
	$\bar{x}(8)$	$med\ x\ (9)$	$z_R(10)$	$z_Q(12)$	$z_{tr}(13)$
$E(z)\ (1)$	0.0004	0.2475	-0.0034	0.3104	0.422
$D(z)\ (2)$	0.1005	0.1319	0.1828	0.1266	0.1929
$E - \sqrt{D}$	-0.3189	-0.1405	-0.4405	-0.0414	-0.0238
$E + \sqrt{D}$	0.2976	0.61	0.4235	0.6499	0.84
Estimation	0.0	0.0	0.0	0^{+1}	0^{+1}
normal n = 100					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	0.0011	0.0224	0.0181	0.0145	0.6287
$D(z)\ (2)$	0.0093	0.0147	0.093	0.0119	0.023
$E - \sqrt{D}$	-0.1066	-0.1037	-0.3262	-0.1038	0.4613
$E + \sqrt{D}$	0.0956	0.1454	0.3184	0.1228	0.7845
Estimation	0.0	0.0	0.0	0.0	0^{+1}
normal n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	0.0007	0.0031	0.0063	0.0022	0.6367
$D(z)\ (2)$	0.0011	0.0016	0.0657	0.0014	0.0026
$E - \sqrt{D}$	-0.0328	-0.0391	-0.2587	-0.0348	0.5826
$E + \sqrt{D}$	0.0321	0.0423	0.2433	0.0372	0.6849
Estimation	0.0	0.0	0.0	0.0	0^1

Таблица 1: Характеристики нормального распределения (3)

cauchy n = 10					
	$\bar{x}(8)$	$med\ x\ (9)$	$z_R(10)$	$z_Q(12)$	$z_{tr}(13)$
$E(z)\ (1)$	2.0874	0.4028	10.6049	1.0676	7.0882
$D(z)\ (2)$	1734.692	0.4305	43204.7097	4.703	4606.8807
$E - \sqrt{D}$	-39.5622	-0.2534	-197.2525	-1.101	-60.7858
$E + \sqrt{D}$	43.737	1.0589	218.4624	3.2363	74.9622
Estimation	-	0_{-1}^{+1}	-	1_{-1}^{+2}	-
cauchy n = 100					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	-1.9491	0.027	-96.1502	0.0249	7.6371
$D(z)\ (2)$	3963.3208	0.0247	9781880.7879	0.0556	641.2068
$E - \sqrt{D}$	-64.904	-0.1302	-3223.7501	-0.2109	-17.6849
$E + \sqrt{D}$	61.0058	0.1841	3031.4496	0.2607	32.9592
Estimation	-	0.0	-	0.0	-
cauchy n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	-2.557	0.0023	-1304.1555	0.0033	7.3359
$D(z)\ (2)$	1585.6463	0.0024	390457865.8927	0.0048	280.3181
$E - \sqrt{D}$	-42.3772	-0.0467	-21064.1622	-0.0663	-9.4068
$E + \sqrt{D}$	37.2632	0.0513	18455.8513	0.0729	24.0786
Estimation	-	0.0	-	0.0	-

Таблица 2: Характеристики распределения Коши (4)

laplace n = 10					
	$\bar{x}(8)$	$med\ x\ (9)$	$z_R(10)$	$z_Q(12)$	$z_{tr}(13)$
$E(z)\ (1)$	0.0112	0.1825	0.0383	0.3002	0.4213
$D(z)\ (2)$	0.1024	0.0843	0.4016	0.118	0.1708
$E - \sqrt{D}$	-0.3089	-0.108	-0.5955	-0.0432	0.008
$E + \sqrt{D}$	0.3313	0.4729	0.672	0.6437	0.8346
Estimation	0	0	0	0	0^{+1}
laplace n = 100					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	-0.0038	0.0104	0.016	0.0087	0.586
$D(z)\ (2)$	0.0103	0.0057	0.4165	0.0099	0.0216
$E - \sqrt{D}$	-0.1054	-0.0648	-0.6294	-0.091	0.4389
$E + \sqrt{D}$	0.0978	0.0856	0.6614	0.1085	0.7331
Estimation	0.0	0.0	0_{-1}^{+1}	0.0	0^{+1}
laplace n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	-0.0013	0.0007	-0.0087	0.0002	0.5971
$D(z)\ (2)$	0.001	0.0005	0.3856	0.0009	0.002
$E - \sqrt{D}$	-0.0322	-0.0218	-0.6297	-0.0306	0.5526
$E + \sqrt{D}$	0.0296	0.0232	0.6123	0.0311	0.6416
Estimation	0.0	0.0	0_{-1}^{+1}	0.0	0_{-1}^{+1}

Таблица 3: Характеристики распределения Лапласа (5)

pois n = 10					
	$\bar{x}(8)$	$med\ x\ (9)$	$z_R(10)$	$z_Q(12)$	$z_{tr}(13)$
$E(z)\ (1)$	10.0283	10.6825	10.3165	10.9835	14.6477
$D(z)\ (2)$	0.965	1.4219	2.0436	1.3155	1.9667
$E - \sqrt{D}$	9.0459	9.49	8.887	9.8366	13.2453
$E + \sqrt{D}$	11.0107	11.875	11.746	12.1304	16.0501
Estimation	10_{-1}^{+1}	10_{-1}^{+1}	10_{-1}^{+1}	10_{-1}^{+2}	14_{-1}^{+2}
pois n = 100					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	9.9989	9.909	10.8785	9.967	16.9031
$D(z)\ (2)$	0.0997	0.2022	0.9735	0.1529	0.2782
$E - \sqrt{D}$	9.6831	9.4593	9.8918	9.576	16.3756
$E + \sqrt{D}$	10.3146	10.3587	11.8652	10.358	17.4305
Estimation	9_{-1}^{+1}	9_{-1}^{+1}	9_{-1}^{+1}	9_{-1}^{+1}	16_{-1}^{+1}
pois n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	9.9969	9.997	11.6315	9.9925	16.9141
$D(z)\ (2)$	0.0097	0.0025	0.6285	0.0037	0.0253
$E - \sqrt{D}$	9.8983	9.9471	10.8387	9.9317	16.7551
$E + \sqrt{D}$	10.0955	10.0469	12.4243	10.0533	17.0732
Estimation	9_{-1}^{+1}	9_{-1}^{+1}	11_{-1}^{+1}	9_{-1}^{+1}	16_{-1}^{+1}

Таблица 4: Характеристики распределения Пуассона (6)

uniform n = 10					
	$\bar{x}(8)$	$med\ x\ (9)$	$z_R(10)$	$z_Q(12)$	$z_{tr}(13)$
$E(z)\ (1)$	-0.0016	0.3037	0.0024	0.3178	0.4173
$D(z)\ (2)$	0.1001	0.2241	0.0488	0.1233	0.2134
$E - \sqrt{D}$	-0.318	-0.1697	-0.2186	-0.0333	-0.0447
$E + \sqrt{D}$	0.3148	0.7771	0.2234	0.6689	0.8792
Estimation	0.0	0 ⁺	0.0	0.0	0 ⁺
uniform n = 100					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	-0.0011	0.0328	0.0009	0.0143	0.6392
$D(z)\ (2)$	0.0102	0.0302	0.0006	0.0148	0.0301
$E - \sqrt{D}$	-0.1022	-0.1409	-0.0238	-0.1076	0.4657
$E + \sqrt{D}$	0.0999	0.2065	0.0257	0.1361	0.8127
Estimation	0.0	0.0	0.0	0.0	0 ⁺
uniform n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)\ (1)$	-0.0007	0.0022	-0.0001	0.0017	0.6479
$D(z)\ (2)$	0.0009	0.003	0.0	0.0014	0.0028
$E - \sqrt{D}$	-0.0315	-0.0522	-0.0026	-0.0356	0.5945
$E + \sqrt{D}$	0.0301	0.0567	0.0024	0.039	0.7012
Estimation	0.0	0.0	0.0	0.0	0 ⁺

Таблица 5: Характеристики равномерного распределения (7)

5 Обсуждение

5.1 Характеристики положения и рассеяния

Исходя из данных, приведенных в таблицах, можно судить о том, что дисперсия характеристик рассеяния для распределения Коши является некой аномалией: значения слишком большие даже при увеличении размера выборки - понятно, что это результат выбросов, которые мы могли наблюдать в результатах предыдущего задания.

6 Ресурсы

Код программы, реализующей отрисовку обозначенных распределений:

<https://github.com/YaroslavAggressive/Mathematical-statistics-lab-works>