

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра Алгоритмической математики

ОТЧЕТ
по лабораторной работе 4
по дисциплине «Статистический анализ»
Тема: Вычисление мощности критериев согласия.

Студент гр. 8374

Пихтовников К.С.

Преподаватель

Чирина А.В.

Санкт-Петербург

2020

Цель работы:

Научиться вычислять мощности критериев согласия.

Задание:

За основную гипотезу принимается распределение Лапласа $L(a,u)$;

Рассматриваются следующие альтернативы:

Альтернатива сдвига:

$$F(x) = F_0(x)$$

Рассмотреть варианты:

$$\theta = \beta, \theta = \beta/3, \theta = \beta/10, \text{ где } \beta -$$

является дисперсией при нулевой гипотезе

Вариант 15:

1. Равномерное распределение

$a=4, d=5$

2. Распределение Лапласа

$a=3, u=4$

Уровень значимости = 0.95

Выполнение работы:

1. Пример генерации выборок.

10x1000 double													
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	9.6684	7.5776	2.0570	13.1484	1.8422	5.3584	3.7472	4.0408	6.6325	-6.6017	-3.5025	2.6165	-0.2934
2	2.7600	10.6175	-8.8816	-5.3811	0.1756	2.2057	1.9912	-2.1487	2.7500	-4.1347	8.1675	-6.8699	7.1280
3	-4.7263	4.0872	6.7378	1.1485	5.6722	2.5169	5.4044	2.6306	16.5532	1.6561	7.5556	-1.6810	-16.0725
4	5.5022	3.0256	4.9042	-4.2705	4.0915	4.9902	4.9034	-2.6388	4.8904	3.6368	2.3647	7.3129	12.7582
5	4.7618	6.4469	6.4313	-4.7768	1.8772	-4.7239	6.7338	13.1821	8.3911	7.6897	4.8690	-2.4850	5.7550
6	1.9577	0.9518	7.2180	2.4300	5.0934	6.1436	3.3553	6.9732	-1.9471	-0.8175	1.8672	3.2197	8.1871
7	-1.8034	2.7331	7.2023	2.5686	15.5504	11.7775	20.9506	-3.4320	2.5536	1.1265	-1.9344	9.2967	-4.0253
8	21.3394	9.3877	0.8162	0.2152	5.4843	0.3346	-3.4569	7.0932	2.7892	13.7115	5.9543	5.1680	0.1950
9	0.5460	6.5732	3.0275	3.0274	0.4300	1.0432	-0.7693	2.5688	-0.4454	-13.2766	6.5804	-6.6300	1.5028

1000x1000 double													
	169	170	171	172	173	174	175	176	177	178	179	180	181
1 32	1.9309	1.4471	6.4119	8.0367	-3.9493	5.7314	-2.4725	-1.1154	1.3282	9.7224	3.7896	5.8470	1.7
2 85	12.2166	-5.5658	9.3061	-11.3347	1.8465	10.8203	0.3469	5.6729	8.5015	2.4663	1.4737	3.5063	5.9
3 41	3.6038	6.6672	-4.6640	4.5585	5.4699	-3.4571	11.2509	6.0282	8.1614	-0.0731	1.5235	0.3517	-2.0
4 10	3.8615	3.3633	7.9055	20.9448	-2.4854	13.5929	0.6149	3.0148	7.9682	1.6480	-1.5395	-3.8429	9.4
5 45	3.4610	17.0888	3.8639	-2.3272	-1.4047	-3.6791	2.4301	6.8292	3.2068	6.3104	1.3165	6.9994	5.5
6 14	4.4370	-12.8326	0.9508	-4.8057	-1.6483	5.1595	4.7956	-1.8555	-3.7455	2.1489	1.0389	14.0916	2.5
7 31	-10.0502	3.0633	-7.8913	1.1519	-4.2903	1.8925	6.6560	3.0566	-0.1779	4.7399	-2.4892	8.3390	8.3
8 06	8.5562	-4.3151	3.2380	0.8463	-8.6571	10.3281	2.7902	9.8650	15.3110	11.7550	0.2581	16.7719	3.2
9 07	-7.7880	3.1048	10.4760	-0.1862	8.5847	13.4466	4.4460	-10.2763	3.0534	0.8132	7.3252	11.8211	12.0

- 1) $\Theta=2/u^2$
- 2) $\Theta=2/3u^2$
- 3) $\Theta=2/10u^2$
- 4) $\Theta=0$

2. Вычисление статистики критерия.

Алгоритм вычисления критерия Колмогорова-Смирнова:

1. Сортируем выборку.
2. Вычисляем теоретические значения распределения Лапласа.
3. Вычисляем статистику критерия с помощью следующей формулы:

$$\sqrt{n} \max_{i=1, \dots, n} \max \left(\left| \frac{i}{n} - F_0(X_{(i)}) \right|, \left| \frac{i-1}{n} - F_0(X_{(i)}) \right| \right)$$

4. Сравниваем полученное значение с квантилью распределения Колмогорова уровня 0.95 ($K=1.36$).

Алгоритм вычисления критерия Хи-квадрат:

1. Сортируем выборку.
2. По формуле Стерджесса разбиваем выборку на n групп и k элементов:

$$n = 1 + 3,322 \cdot \lg N$$

$$K = \frac{X_{max} - X_{min}}{1 + 3,2 \cdot \lg n},$$

3. Рассчитываем вероятности попадания в интервалы
4. Вычисляем статистику критерия по формуле (здесь $n=r$):

$$\sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i}$$

5. Сравниваем полученное значение с квантилью распределения хи-квадрат уровня 0.95 и числом степеней свободы $r-1$ ($n-1$).

3. Процент выборок, для которых нулевая гипотеза была отвергнута.

Критерий Колмогорова-Смирнова:

```
k = 0;
for i=1:m
    criterion_statistics_kolmogorov = sqrt(n) * maximum(i);
    if criterion_statistics_kolmogorov > 1.36
        k = k + 1;
    end
end
```

Критерий Хи-квадрат:

```
Pirs = zeros(1,m);
for i=1:m
    for j=1:r
        Pirs(1,i) = Pirs(1,i) + ((nh(j,i) - n * P(j,i))^2 / (n * P(j,i)));
    end
end
Pirs;
k = 0;
```

```

for i=1:m
    if Pirs(1,i) > chi2inv(0.95, r - 1)
        k = k + 1;
    end
end
end

```

4. Мощность критерия в зависимости от альтернативы.

Мощность критерия Колмогорова ($\alpha=1-\gamma=0.05$)

	$\Theta=2/16$	$\Theta=2/48$	$\Theta=2/160$	0
n=10	0.078000	0.062000	0.057000	0.054000
n=100	0.395000	0.304000	0.299000	0.237
n=1000	1.000000	1.000000	1.000000	1.000000

Мощность критерия Хи-квадрат ($\alpha=1-\gamma=0.05$)

	$\Theta=2/16$	$\Theta=2/48$	$\Theta=2/160$	0
n=10	0.377000	0.372000	0.361000	0.234
n=100	0.709000	0.705000	0.677000	0.627000
n=1000	1.000000	1.000000	1.000000	1.000000

5. Вывод.

При большем увеличении сдвига и объеме выборки, вероятность принятия гипотез критериев Колмогорова-Смирнова и Хи-квадрат уменьшаются, соответственно увеличивается вероятность отвержения гипотез. Также критерий Хи-квадрат является более мощным по сравнению с критерием Колмогорова.

Значение мощности данных критериев совпадает с уровнем значимости при нулевом сдвиге.