

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ
ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А. И. ГЕРЦЕНА**

Институт информационных технологий и технологического
образования Кафедра компьютерные технологии и электронного
обучения

Основная профессиональная образовательная программа
Направление подготовки 09.03.01 Информатика и вычислительная
техника Направленность (профиль) «Технологии разработки программного обеспечения»
форма обучения - очная

ЛАБОРАТОРНАЯ РАБОТА №4
по дисциплине: «Анализ данных и основы Data science»
НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Руководитель: кандидат педагогических
наук, доцент, Светлана
Викторовна Гончарова

Авторы работы:
Воложанин В.О.
Максимова А. В.

2 курс, 2 группа 1 подгруппа

Цель: проверить данные на соответствие нормальному закону распределения.

Оборудование: ПК, Excel.

Задание 1

Постановка задачи:

Рассчитать теоретические частоты для нормального распределения.

Математические модели:

$$\varphi = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \quad \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}} \quad \bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

$$t_i = \frac{x_i - \bar{x}}{\sigma} \quad f_m = \varphi(t) \frac{Nd}{\sigma}$$

Таблица:

сумма затрат предприятий на производство, тыс. руб	количество предприятий, fi	средний интервал, xi	$x_i f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 f_i$	σ	ti	$\varphi(t)$	N	d
A	1	2				29,63773037			108	10
30 - 40	2	35	70	3871,604938	7743,209877		-2,10	0,0440		
40 - 50	4	45	180	2727,160494	10908,64198		-1,76	0,0845		
50 - 60	6	55	330	1782,716049	10696,2963		-1,42	0,1446		
60 - 70	8	65	520	1038,271605	8306,17284		-1,09	0,2209		
70 - 80	11	75	825	493,8271605	5432,098765		-0,75	0,3012		
80 - 90	14	85	1190	149,382716	2091,358025		-0,41	0,3664		
90 - 100	15	95	1425	4,938271605	74,07407407		-0,07	0,3978		
100 - 110	13	105	1365	60,49382716	786,4197531		0,26	0,3854		
110 - 120	11	115	1265	316,0493827	3476,54321		0,60	0,3333		
120 - 130	8	125	1000	771,6049383	6172,839506		0,94	0,2571		
130 - 140	6	135	810	1427,160494	8562,962963		1,27	0,1771		
140 - 150	5	145	725	2282,716049	11413,58025		1,61	0,1088		
150 - 160	3	155	465	3338,271605	10014,81481		1,95	0,0597		
160 - 170	2	165	330	4593,82716	9187,654321		2,29	0,0292		
Итого	108		\bar{x}	-	-	-	-	-	-	-
			97,222							

Задание 2.

Постановка задачи:

Определить является ли распределение (из Задания 1) нормальным. Использовать критерий Колмогорова.

Математическая модель:

$$\lambda = \frac{D_{max}}{\sqrt{N}}$$

Таблица:

f_m	F_i	F_m	D_i	D_{max}	λ
				3	0.3052892906
2	2	2	0		
3	6	5	1		
5	12	10	2		
8	20	18	2		
11	31	29	2		
13	45	42	3		
14	60	57	3		
14	73	71	2		
12	84	83	1		
9	92	92	0		
6	98	99	-1		
4	103	103	0		
2	106	105	1		
1	108	106	2		
106	-	-	-	-	-

Вывод: в ходе работы мы выяснили, что вычисленные теоретические частоты достаточно близки к эмпирическим. Так же распределение исследуемых данных достаточно близко к нормальному закону.