

Университет ИТМО

Факультет программной инженерии и компьютерной техники

Индивидуальное домашнее задание №1

по «Математической статистике»

Вариант 9

Выполнили:

Студенты группы Р3233

Хасаншин Марат

Шикунов Максим

Номер команд: 9

Санкт-Петербург

2024

Цель работы

На основании анализа выборки:

- 1) Построить интервальный ряд исследуемого признака;
- 2) Полигон частот, выборочную функцию распределения и гистограмму;
- 3) Найти точечные оценки математического ожидания и дисперсии;
- 4) Построить доверительные интервалы для математического ожидания и дисперсии с доверительной вероятностью 0.95;

Исходные данные

0.462	-0.603	0.264	-0.373	2.173	-1.875	0.261	0.064	-0.814	-0.456
1.288	1.833	0.292	-0.294	0.572	0.917	0.743	-1.727	0.990	-1.903
-0.956	-0.965	0.781	-1.717	0.815	-0.546	-0.162	0.716	-1.781	-0.392
1.195	-0.397	0.404	-0.053	-1.078	-0.605	0.435	0.036	-0.044	-1.107
-0.405	0.089	-0.325	0.217	-0.579	0.025	0.861	-0.184	0.890	1.757
-0.719	1.202	-1.083	0.606	1.244	-1.547	-0.108	0.856	1.034	-0.127
-0.219	-0.112	0.157	0.074	0.029	-1.071	-0.300	3.343	-0.618	1.019
-0.030	0.673	-0.662	-0.685	-1.675	0.737	1.279	0.894	0.987	0.170
-0.495	-1.322	0.362	0.475	-0.043	-1.698	-0.404	-0.741	-0.237	-0.420
-0.333	-0.216	1.170	0.757	-0.691	-0.591	1.444	1.695	0.307	2.096

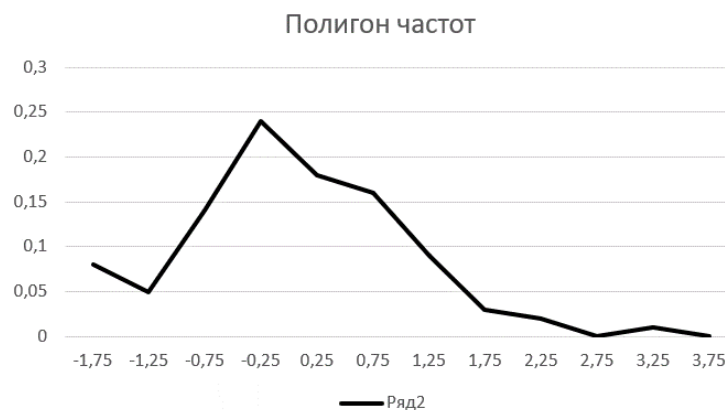
Ход выполнения

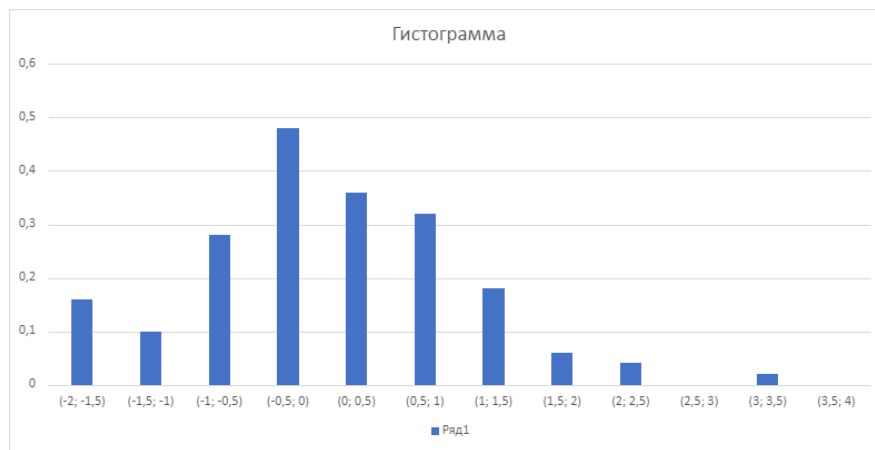
Наименьшее значение: -1.903

Наибольшее значение: 3.343

Интервальный ряд:

Номера интервалов	1	2	3	4	5	6	7	8	9	10	11	12
Границы интервалов	(-2; -1,5)	(-1,5; -1)	(-1; -0,5)	(-0,5; 0)	(0; 0,5)	(0,5; 1)	(1; 1,5)	(1,5; 2)	(2; 2,5)	(2,5; 3)	(3; 3,5)	(3,5; 4)
x^*	-1,75	-1,25	-0,75	-0,25	0,25	0,75	1,25	1,75	2,25	2,75	3,25	3,75
n	8	5	14	24	18	16	9	3	2	0	1	0
p^*	0,08	0,05	0,14	0,24	0,18	0,16	0,09	0,03	0,02	0	0,01	0
p^*/h	0,16	0,1	0,28	0,48	0,36	0,32	0,18	0,06	0,04	0	0,02	0





Точечная оценка математического ожидания и дисперсии

$$\begin{aligned}
 \hat{m} &= \sum_{i=1}^n x_i^* \cdot p_i^* = \\
 &= -0,14 - 0,0625 - 0,105 - 0,06 + 0,045 + 0,12 + 0,1125 \\
 &\quad + 0,0525 + 0,045 + 0 + 0,0325 + 0 = 0,04 \\
 \hat{d} &= \sum_{i=1}^n (x_i^* - \hat{m})^2 \cdot p_i^* = \sum_{i=1}^n (x_i^*)^2 \cdot p_i^* - \hat{m}^2 \\
 &= (0,24 + 0,07813 + 0,07875 + 0,015 + 0,01125 + 0,09 \\
 &\quad + 0,14063 + 0,09188 + 0,10125 + 0 + 0,10563 + 0) - 0,0016 \\
 &= 0,9559
 \end{aligned}$$

Доверительные интервалы

Математическое ожидание

Центрируем и нормируем \hat{m} :

$$\frac{\sqrt{n}(\hat{m} - m)}{\hat{\sigma}} = \frac{\sqrt{n}(\hat{m} - m)}{\sqrt{\hat{d}}} = \tau_{n-1}$$

Случайная величину $t = 1.95$, $\gamma = 0.95$, $\Phi(t) = 0.975$

И подставим ее в формулу:

$$P\left(\hat{m} - t \cdot \sqrt{\frac{\hat{d}}{n}} < m < \hat{m} + t \cdot \sqrt{\frac{\hat{d}}{n}}\right) = 0.95$$
$$0.04 - 1.95 \cdot \sqrt{\frac{0.9559}{100}} < m < 0.04 + 1.95 \cdot \sqrt{\frac{0.9559}{100}}$$
$$-0.150652 < m < 0.23065$$

Итак, доверительный интервал с доверительной вероятностью 0.95 для математического ожидания – $(-0.150652; 0.23065)$

Дисперсия

Рассмотрим следующую случайную величину:

$$\frac{(n-1) \cdot \hat{\sigma}^2}{\sigma^2} = \chi_{n-1}^2$$

Она имеет распределение «хи-квадрат» с 11 степенями свободы. По таблице распределения «хи-квадрат» находим такие h_1 и h_2 , что:

$$P(\chi_{11}^2 < h_1) = P(\chi_{11}^2 > h_2) = \frac{1 - \beta}{2} = \frac{0.05}{2} = 0.025$$

$$h_1 = \chi_{11, \frac{1+\beta}{2}}^2 = \chi_{11, 0.975}^2 = 3.81575$$

$$h_2 = \chi_{11, \frac{1-\beta}{2}}^2 = \chi_{11, 0.025}^2 = 21.92$$

Тогда:

$$P(h_1 < \chi_{11}^2 < h_2) = \beta$$

$$h_1 < \frac{(n-1) \cdot \hat{\sigma}^2}{\sigma^2} < h_2$$

$$\frac{(n-1) \cdot \hat{\sigma}^2}{h_2} < \sigma^2 < \frac{(n-1) \cdot \hat{\sigma}^2}{h_1}$$

$$\frac{11 \cdot 0.9559}{21.92} < \sigma^2 < \frac{11 \cdot 0.9559}{3.81575}$$

$$0.4797 < \sigma^2 < 2.7557$$

Лабораторная №5

Проверка статистической гипотезы

Формулировка основной и альтернативной гипотез:

H_0 : генеральная совокупность имеет нормальное распределение

H_1 : генеральная совокупность имеет распределение, отличное от нормального

Уровень значимости:

$$\alpha = 0.05$$

Критерий Пирсона:

$$U = \sum_{i=1}^k \frac{n_i^2}{np_i} - 100, \text{ где } p_i = F(x_{i+1}) - F(x_i) = \Phi\left(\frac{x_{i+1} - \hat{m}}{\hat{\sigma}}\right) - \Phi\left(\frac{x_i - \hat{m}}{\hat{\sigma}}\right)$$

F – функция распределения нормального закона, Φ – функция Лапласа

Если H_0 верна, то U имеет распределение «хи-квадрат» распределение с числом степеней свободы $k - 3$ ($k = 6$).

Вычислим значение критерия:

	x =		S^2 =											
Интервалы	(-2; -1,5)	(-1,5; -1)	(-1; -0,5)	(-0,5; 0)	(0; 0,5)	(0,5; 1)	(1; 1,5)	(1,5; 2)	(2; 2,5)	(2,5; 3)	(3; 3,5)	(3,5; 4)		
Zi	-1,53	-1,09	-0,48	0,04	0,56	1,09	1,61	2,13	2,66	3,18	3,7	4,23		
Ф(Zi)	0,063	0,1379	0,3156	0,526	0,7123	0,8621	0,9463	0,9834	0,9961	0,9993	0,9999	1	Сумма:	
pi	0,063	0,0749	0,1777	0,2004	0,1963	0,1498	0,0842	0,0371	0,0127	0,0032	0,0006	0,0001	1	
	0,1379						0,1379							
ni	8	5	14	24	18	16	9	3	2	0	1	0	Сумма:	
	13						15							100
ni^2	169		196	576	324	256	225							Сумма:
npi	13,79		17,77	20,04	19,63	14,98	13,79							100
ni^2/npi	12,25525743		11,0298	28,7425	16,5053	17,0895	16,31617114							101,939

$$U = 1.93857$$

Находим значение критической точки по таблице:

$$t_{кр} = 7.81473$$

Видим, что значение критерия попадает в допустимую область ($U < t_{кр}$). Гипотеза принимается на уровне значимости 0.05.

Вывод

Мы нашли частоту наших значений, доверительные интервалы для математического ожидания. В пятой лабораторной мы проверяем и доказываем гипотезу о нормальном распределении наших данных. Данная гипотеза принимается по нашим данным.