ИДЗ-18.2

Найти закон распределения указанной случайной величины X и ее функцию распределения $F(X)$. Вычислить математическое ожидание $M(X)$, дисперсию $D(X)$ и среднее квадратическое отклонение $\sigma(X)$. Построить график функции распределения $F(X)$.

1) Вероятность безотказной работы в течение гарантийного срока для телевизоров первого типа равна 0,9, второго типа – 0,7, третьего типа – 0,8. Случайная величина X – число телевизоров, проработавших гарантийный срок, среди трех телевизоров разных типов.

Решение: По условию $p_1 = 0{,}9$, $p_2 = 0{,}7$, $p_3 = 0{,}8$ – вероятности безотказной работы в течение гарантийного срока телевизоров соответствующих типов. Тогда вероятности их отказа:

$$q_1 = 1 - p_1 = 1 - 0{,}9 = 0{,}1$$
$$q_2 = 1 - p_2 = 1 - 0{,}7 = 0{,}3$$
$$q_3 = 1 - p_3 = 1 - 0{,}8 = 0{,}2$$

Используя теоремы умножения независимых и сложения несовместных событий, составим закон распределения случайной величины X – количества телевизоров, проработавших гарантийный срок, среди трех телевизоров разных типов.

0) $X = 0$ (все телевизоры вышли из строя)

$$p(0) = q_1 q_2 q_3 = 0{,}1 \cdot 0{,}3 \cdot 0{,}2 = 0{,}006$$

1) $X = 1$

$$p(1) = p_1 q_2 q_3 + q_1 p_2 q_3 + q_1 q_2 p_3 = 0{,}9 \cdot 0{,}3 \cdot 0{,}2 + 0{,}1 \cdot 0{,}7 \cdot 0{,}2 + 0{,}1 \cdot 0{,}3 \cdot 0{,}8 =$$
$$= 0{,}054 + 0{,}014 + 0{,}024 = 0{,}092$$

2) $X = 2$

$$p(2) = p_1 p_2 q_3 + p_1 q_2 p_3 + q_1 p_2 p_3 = 0{,}9 \cdot 0{,}7 \cdot 0{,}2 + 0{,}9 \cdot 0{,}3 \cdot 0{,}8 + 0{,}1 \cdot 0{,}7 \cdot 0{,}8 =$$
$$= 0{,}126 + 0{,}216 + 0{,}056 = 0{,}398$$

3) $X = 3$ (все телевизоры проработали гарантийный срок)

$$p(3) = p_1 p_2 p_3 = 0{,}9 \cdot 0{,}7 \cdot 0{,}8 = 0{,}504$$

Таким образом, искомый закон распределения:

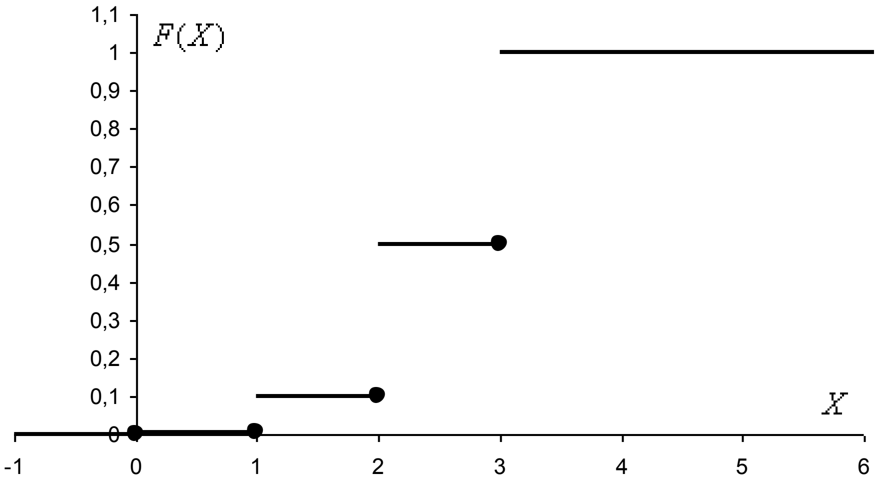
x_i	0	1	2	3
$p(i)$	0,006	0,092	0,398	0,504

Проверка: 0,006 + 0,092 + 0,398 + 0,504 = 1

Составим функцию распределения:

$$F(X)=\begin{cases}0 \ \textit{при}\ x\le 0;\\ 0{,}006\ \textit{при}\ 0<x\le 1;\\ 0{,}098\ \textit{при}\ 1<x\le 2;\\ 0{,}4968\ \textit{при}\ 2<x\le 3;\\ 1\ \textit{при}\ x>3.\end{cases}$$

Выполним чертеж:



Вычислим математическое ожидание $M(X)$, дисперсию $D(X)$ и среднее квадратическое отклонение $\sigma(X)$. Заполним расчетную таблицу:

x_i	0	1	2	3	Суммы:
p_i	0,006	0,092	0,398	0,504	1
$x_i p_i$	0	0,092	0,796	1,512	2,4
$x_i^2 p_i$	0	0,092	1,592	4,536	6,22

Математическое ожидание: $M(X)=2{,}4$

Дисперсию вычислим по формуле:
$D(X)=M(X^2)-(M(X))^2=6{,}22-(2{,}4)^2=6{,}22-5{,}76=0{,}46$.

Среднее квадратическое отклонение: $\sigma(X)=\sqrt{D(X)}=\sqrt{0{,}46}\approx 0{,}68$

2) Задана функция распределения случайной величины X:

$$F(x)=\begin{cases}0, x<-1\\ \frac{1}{5}(x+1), -1\le x\le 4\\ 1, x>4\end{cases}$$

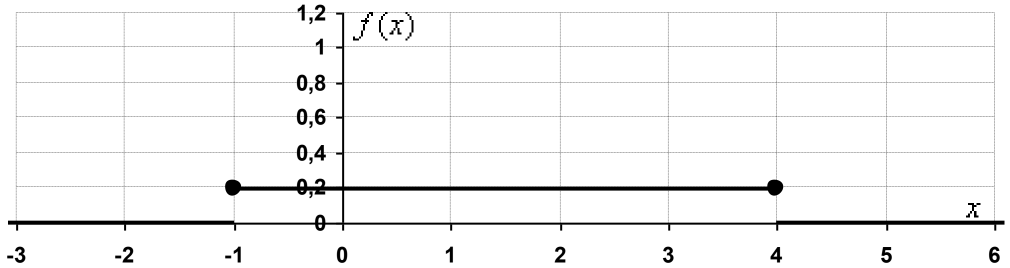
Найти плотность распределения вероятностей $f(x)$, математическое ожидание $M(X)$, дисперсию $D(X)$ и вероятность попадания случайной величины X на отрезок $[0;3]$. Построить графики функций $F(x)$ и $f(x)$.

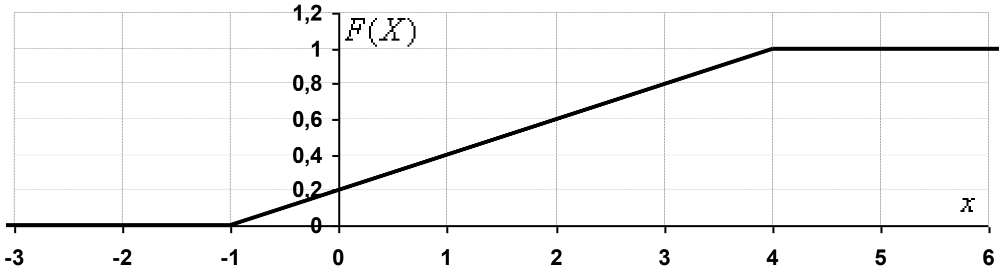
Решение:

Найдем функцию плотности распределения вероятностей:

$$f(x)=F'(x)=\begin{cases}0, x<-1\\ \frac{1}{5}, -1\le x\le 4\\ 0, x>4\end{cases}$$

Построим графики функций $f(x)$ и $F(x)$.





Вычислим математическое ожидание и дисперсию.

Математическое ожидание:

$$M(x)=\int_{-\infty}^{+\infty}xf(x)dx=\frac{1}{5}\int_{-1}^{4}xdx=\frac{1}{5}\cdot\frac{1}{2}(x^2)\Big|_{-1}^{4}=\frac{1}{10}\cdot(16-1)=\frac{15}{10}=\frac{3}{2}=1{,}5$$

Дисперсия:

$$D(x)=\int_{-\infty}^{+\infty}x^2 f(x)dx-(M(x))^2=\frac{1}{5}\int_{-1}^{4}x^2dx-\left(\frac{3}{2}\right)^2=$$

$$=\frac{1}{5}\cdot\frac{1}{3}(x^3)\Big|_{-1}^{4}-\frac{9}{4}=\frac{1}{15}(64+1)-\frac{9}{4}=\frac{13}{3}-\frac{9}{4}=\frac{25}{12}\approx 2{,}083$$

Найдем вероятность того, что X примет значение из отрезка $[0;3]$.

$P(0\le x\le 3)=F(3)-F(0)=\frac{4}{5}-\frac{1}{5}=\frac{3}{5}=0{,}6$ – искомая вероятность.

3) Считается, что изделие – высшего качества, если отклонение его размеров от номинальных не превосходит по абсолютной величине 3,6 мм. Случайные отклонения размера изделия от номинального подчиняется нормальному закону со средним квадратическим отклонением, равным 3 мм. Систематические отклонения отсутствуют. Определить среднее число изделий высшего качества среди 100 изготовленных.

Решение: Используем формулу:

$P(|X-M(x)|<\delta)\approx 2\Phi\left(\frac{\delta}{\sigma}\right)$, где $\Phi(x)$ – интегральная функция нормально распределенной случайной величины X; значения данной функции находим по соответствующей таблице.

В данном случае:

$P(|X-M(x)|<3{,}6)\approx 2\Phi\left(\frac{3{,}6}{3}\right)=2\Phi(1{,}2)=2\cdot 0{,}3849\ =0{,}7699\ \approx 0{,}77$ –вероятность того, что изделие – высшего качества.

Таким образом, среднее число изделий высшего качества среди 100 изготовленных: $0{,}77\cdot 100=77$

Ответ: 77

4) Производится выборочный контроль партии электролампочек для определения средней продолжительности горения. Каким должен быть объем выборки, чтобы с вероятностью не меньшей 0,9876, можно было утверждать, что средняя продолжительность эксплуатации лампочки по всей партии отклонилась от средней, полученной в выборке, не более, чем на 10 ч, если среднее квадратическое отклонение продолжительности эксплуатации лампочки равно 80 ч?

Решение: Используем теорему Чебышева:

$$P\left(|\bar{x}-x_{\Gamma}|\le\varepsilon\right)\ge 1-\frac{D}{n\varepsilon^2}$$

В данном случае:

$$1-\frac{80^2}{n\cdot 10^2}=0{,}9876$$

$$\frac{64}{n}=0{,}0124$$

$$n=\frac{64}{0{,}0124}\approx 5161{,}3$$

Ответ: не менее чем 5162

ИДЗ 19.1 Задача по статистике

В результате эксперимента получены данные, записанные в виде статистического ряда.

Решение:

а-б) Построим интервальный вариационный ряд распределения.

По статистическим данным находим: $\min x_i = 17$, $\max x_i = 143$

Размах вариации: $\delta = x_{\max} - x_{\min} = 143 - 17 = 126$

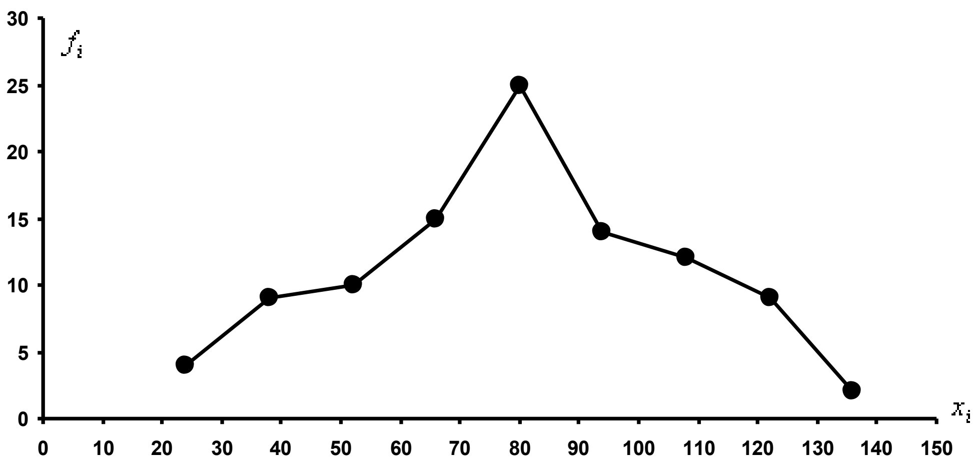
Выборку разобьем на 9 равных интервалов. Величина отдельного интервала: $h = \frac{\delta}{9} = \frac{126}{9} = 14$. Подсчитаем частоту f_i по каждому интервалу.

Перейдем к дискретному ряду распределения, выбрав в качестве вариант x_i середины интервалов.

Заполним расчетную таблицу:

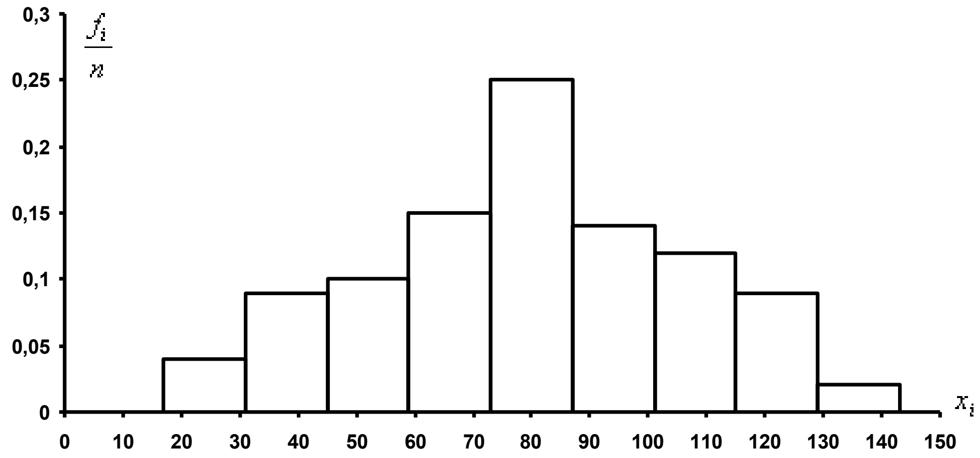
Интервал		x_i	f_i	$\omega_i = \frac{f_i}{n}$	$\omega_i^{НАК}$	$x_i f_i$	$x_i^2 f_i$
17	31	24	4	0,04	0,04	96	2304
31	45	38	9	0,09	0,13	342	12996
45	59	52	10	0,1	0,23	520	27040
59	73	66	15	0,15	0,38	990	65340
73	87	80	25	0,25	0,63	2000	160000
87	101	94	14	0,14	0,77	1316	123704
101	115	108	12	0,12	0,89	1296	139968
115	129	122	9	0,09	0,98	1098	133956
129	143	136	2	0,02	1	272	36992
		$\sum$	$n = 100$	1		7930	702300

в) Построим полигон частот:



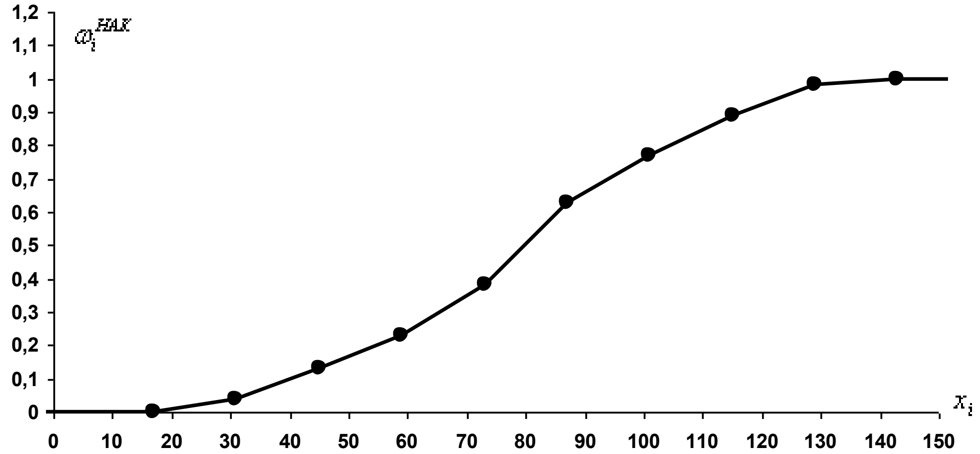
Полигон частот

Постоим гистограмму относительных частот:



Гистограмма относительных частот

Построим эмпирическую функцию распределения:



Эмпирическая функция распределения

г) Вычислим среднюю арифметическую:

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{7930}{100} = 79{,}3$$

Вычислим выборочную дисперсию:

$$D_в = \frac{\sum x_i^2 f_i}{n} - (\bar{x})^2 = \frac{702300}{100} - (79{,}3)^2 = 7023 - 6288{,}49 = 734{,}51 .$$

д) Проведем проверку гипотезы H_0: генеральная совокупность распределена по нормальному закону; конкурирующая гипотеза: H_1: генеральная совокупность не распределена по нормальному закону.

Используем критерий Пирсона; уровень значимости $\alpha = 0{,}025$

Найдем значения теоретических частот f_i^T. Используем формулу: $f_i^T = \frac{h\sum f_i}{\sigma} f(t)$,

где $f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$, $t = \frac{x_i - \bar{x}}{\sigma}$.

В данном случае среднее квадратическое отклонение:

$\sigma = \sqrt{D(x)} = \sqrt{734{,}51} \approx 27{,}10$

Заполним расчетную таблицу:

x_i	f_i	t_i	$f(t_i)$	f_i^T
24	4	-2,04045	0,049754	3
38	9	-1,52388	0,124924	6
52	10	-1,00731	0,240202	12
66	15	-0,49074	0,353684	18
80	25	0,025828	0,398809	21
94	14	0,542398	0,344371	18
108	12	1,058968	0,227718	12
122	9	1,575538	0,115314	6
136	2	2,092108	0,044717	2

Расчетное значение критерия вычислим по формуле: $\chi^2_{расч} = \sum_{i=1}^{l} \frac{(f_i - f_i^T)^2}{f_i^T}$, где l - количество интервалов эмпирического распределения.

Заполним расчетную таблицу:

f_i	f_i^T	$\frac{(f_i - f_i^T)^2}{f_i^T}$
4	3	0,3333
9	6	1,5000
10	12	0,3333
15	18	0,5000
25	21	0,7619
14	18	0,8889
12	12	0,0000
9	6	1,5000
2	2	0,0000

Таким образом: $\chi^2_{расч} = 5{,}8175$

По соответствующей таблице найдем теоретическое значение критерия: $\chi^2_{теор} = \chi^2_{\alpha,k}$.

В данном случае $\alpha = 0{,}025$, количество степеней свободы: $k = l - s - 1$, где s - количество параметров теоретического распределения; в данном случае $k = 9 - 2 - 1 = 6$

По соответствующей таблице: $\chi^2_{теор} = \chi^2_{0{,}025;6} = 14{,}4$.

Таким образом $\chi^2_{расч} < \chi^2_{теор}$, то есть на уровне значимости $\alpha = 0{,}025$ принимаем нулевую гипотезу о том, что генеральная совокупность, из которой извлечена выборка, имеет нормальное распределение.

е) Доверительный интервал для оценки истинного значения генеральной средней $\tilde{x}$ измеряемой величины вычислим по формуле:

$$\bar{x} - \frac{t_\gamma s}{\sqrt{n}} < \tilde{x} < \bar{x} + \frac{t_\gamma s}{\sqrt{n}}$$

Выберем уровень доверительной вероятности $\gamma = 0{,}05$. По количеству степеней свободы $f = n - 1 = 9 - 1 = 8$ и уровню значимости $\gamma = 0{,}05$ по таблице распределения Стьюдента находим: $t_\gamma = 2{,}31$.

Вычислим точность оценки:

$$\frac{t_\gamma s}{\sqrt{n}} = \frac{2{,}31 \cdot 27{,}10}{\sqrt{100}} \approx 6{,}261 \ .$$

Таким образом:

$$79{,}3 - 6{,}261 < \bar{x} < 79{,}3 + 6{,}261$$

Искомый доверительный интервал: $73{,}039 < \bar{x} < 85{,}561$, то есть с вероятностью 95% данный интервал накроет истинное значение $\tilde{x}$ генеральной средней.

При размере выборки $n > 30$ доверительный интервал для генерального среднего квадратического отклонения $\tilde{\sigma}$ определяется по формуле:

$$\frac{\sqrt{2n}}{\sqrt{2n-3} + t_\gamma} \cdot \sigma < \tilde{\sigma} < \frac{\sqrt{2n}}{\sqrt{2n-3} - t_\gamma} \cdot \sigma$$

Ввиду достаточно большого количества наблюдений смещенностью найденного значения σ пренебрегаем.

Таким образом, искомый доверительный интервал:

$$\frac{\sqrt{2 \cdot 100}}{\sqrt{2 \cdot 100 - 3} + 2{,}31} \cdot 27{,}10 < \tilde{\sigma} < \frac{\sqrt{2 \cdot 100}}{\sqrt{2 \cdot 100 - 3} - 2{,}31} \cdot 27{,}10$$

$23{,}45 < \tilde{\sigma} < 32{,}69$, то есть с вероятностью 95% данный интервал накроет истинное значение $\tilde{\sigma}$ генерального среднего квадратического отклонения.

ИДЗ 19.2 (скриншоты из Экселя)

K13

ИДЗ 19.2 Решение: а, б)

X	Y										
	2400	2440	2480	2520	2560	2600	2640	2680	0	0	Ni
300	5	4	2								11
305		1	3	3							7
310			7	10	14						31
315				9	6	4					19
320						8	5	7			20
325							6	6			12
Mj	5	5	12	22	20	12	11	13	0	0	100

Вычислим коэфициент корреляции

Заполним расчетную таблицу для компоненты X:

Xi	300	305	310	315	320	325	Суммы:
Ni	11	7	31	19	20	12	100
Xi*Ni	3300	2135	9610	5985	6400	3900	31330
Xi*Xi	90000	93025	96100	99225	102400	105625	
Xi*Xi*N	990000	651175	2979100	1885275	2048000	1267500	9821050

n = 100

Вычислим Xср., Dx, Sx:

Средняя: Xср. = ∑(Xi*Ni) / n = 313,3

Дисперсия: Dx = ∑(Xi*Xi*Ni) / n - Xср.* Xср. = 53,61

Среднее квадратическое отклонение: $S_x = \sqrt{D_x} =$ 7,32

Заполним расчетную таблицу для компоненты Y:

Yj	2400	2440	2480	2520	2560	2600	2640	2680	0	0	Суммы:
Mj	5	5	12	22	20	12	11	13	0	0	100
Yj*Mj	12000	12200	29760	55440	51200	31200	29040	34840	0	0	255680
Yj*Yj	5760000	5953600	6150400	6350400	6553600	6760000	6969600	7182400	0	0	
Yj*Yj*M	28800000	29768000	73804800	139708800	131072000	81120000	76665600	93371200	0	0	654310400

n = 100

Вычислим Yср., Dy, Sy:

Средняя: Yср. = ∑(Yj*Mj) / n = 2556,8

Дисперсия: Dy = ∑(Yj*Yj*Mj) / n - Yср.* Yср. = 5877,76

Среднее квадратическое отклонение: $S_y = \sqrt{D_y} =$ 76,67

Заполним таблицу произведениями всех возможных значений X и Y

на соответствующие частоты:

на соответствующие частоты:

X	Y									
	2400	2440	2480	2520	2560	2600	2640	2680	0	0
300	3600000	2928000	1488000	0	0	0	0	0	0	0
305	0	744200	2269200	2305800	0	0	0	0	0	0
310	0	0	5381600	7812000	11110400	0	0	0	0	0
315	0	0	0	7144200	4838400	3276000	0	0	0	0
320	0	0	0	0	0	6656000	4224000	6003200	0	0
325	0	0	0	0	0	0	5148000	5226000	0	0

XYср. =∑∑ (Xi*Yj*NMij) / n= 801550

Коэффициент корреляции:

Rxy = (XYср. - Xср.* Yср.) / (Sx * Sy)= **0,90**

Составим уравнения линий регрессии Y на X:

Y - Yср. = Rxy * Sy / Sx * (X - Xср.)

Y - 2557 = 9,41 * (X - 313,3)

Y = 9,41 * X -391,88

Точки для графика:

X	300	325
Y	2431,6	2666,9

Изобразим поле рассеяния и построим линию регрессии: