45,4	45,3	44,1	46,6	44,8	45,6	43,7	46,8	45,2	46,1
44,5	45,4	45,1	46,2	44,2	46,4	45,7	43,9	47,2	45,0
43,9	45,6	44,9	44,5	46,2	46,7	44,3	46,1	47,7	45,8
45,6	45,2	44,2	46,0	44,7	46,5	43,5	45,4	47,1	44,0
46,2	44,2	45,5	46,0	45,7	46,4	44,6	47,0	45,2	46,9

Требуется:

а) записать значения результатов эксперимента в виде вариационного ряда;

б) найти размах варьирования и разбить его на 9 интервалов;

в) построить полигон частот, гистограмму относительных частот и график эмпирической функции распределения;

г) найти числовые характеристики выборки $\bar{x}$, $D_{\text{в}}$;

д) приняв в качестве нулевой гипотезу H_0: генеральная совокупность, из которой извлечена выборка, имеет нормальное распределение, проверить ее, пользуясь критерием Пирсона при уровне значимости $\alpha = 0{,}01$;

е) найти доверительные интервалы для математического ожидания и среднего квадратичного отклонения при надежности $\gamma = 0{,}95$.

▸ а) Располагаем значения результатов эксперимента в порядке возрастания, т.е. записываем вариационный ряд:

43,4	43,5	43,7	43,8	43,9	43,9	43,9	44,0	44,0	44,1
44,2	44,2	44,2	44,3	44,3	44,3	44,4	44,5	44,5	44,5
44,6	44,6	44,7	44,7	44,8	44,8	44,8	44,9	44,9	44,9
45,0	45,0	45,1	45,2	45,2	45,2	45,2	45,2	45,3	45,3
45,3	45,4	45,4	45,4	45,4	45,4	45,4	45,5	45,5	45,6
45,6	45,6	45,6	45,6	45,7	45,7	45,7	45,7	45,7	45,7
45,8	45,8	45,9	45,9	46,0	46,0	46,0	46,0	46,0	46,0
46,1	46,1	46,1	46,1	46,2	46,2	46,2	46,2	46,2	46,4
46,4	46,4	46,4	46,4	46,5	46,5	46,5	46,6	46,7	46,7
46,7	46,7	46,7	46,8	46,9	47,0	47,1	47,1	47,2	47,7

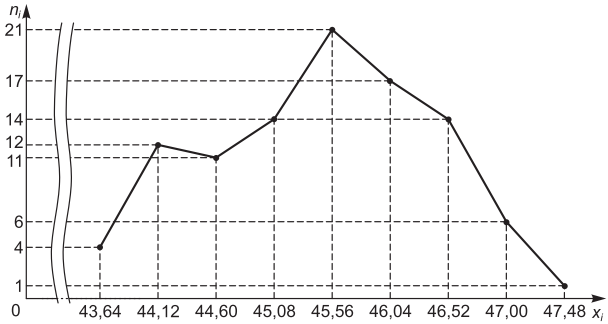
б) Находим размах варьирования: $\omega = x_{\max} - x_{\min} = 47{,}7 - 43{,}4 = 4{,}3$. По формуле $h = \omega / l$, где l – число интервалов, вычисляем длину частичного интервала $h = 4{,}3/9 = 0{,}4(7) = 0{,}48$. В качестве границы первого интервала можно выбрать значение $x_{\min}$. Тогда границы следующих частичных интервалов вычисляем по формуле $x_{\min} + dh$, $d = \overline{1, l}$. Находим середины интервалов: $x'_i = (x_i + x_{i+1})/2$. Подсчитываем число значений результатов эксперимента, попавших в каждый интервал, т.е. находим частоты интервалов n_i. Далее вычисляем относительные частоты $W_i = n_i/n$ $(n = 100)$ и их плотности W_i/h. Все полученные результаты помещаем в таблицу (табл. 19.24).

Таблица 19.24

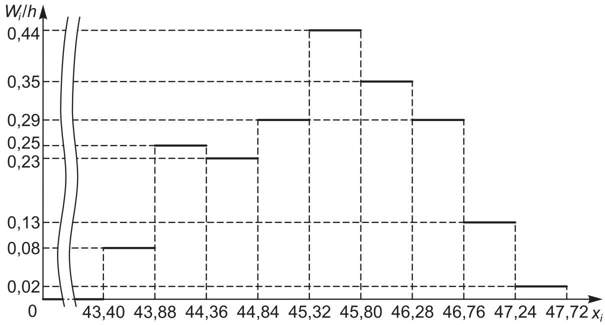
Номер частичного интервала l_i	Границы интервала $x_i - x_{i+1}$	Середина интервала $x'_i = (x_i + x_{i+1})/2$	Частота интервала n_i	Относительная частота $W_i = n_i/n$	Плотность относительной частоты W_i/h
1	43,40–43,88	43,64	4	0,04	0,083
2	43,88–44,36	44,12	12	0,12	0,25
3	44,36–44,84	44,60	11	0,11	0,23
4	44,84–45,32	45,08	14	0,14	0,29
5	45,32–45,80	45,56	21	0,21	0,44
6	45,80–46,28	46,04	17	0,17	0,35
7	46,28–46,76	46,52	14	0,14	0,29
8	46,76–47,24	47,00	6	0,06	0,13
9	47,24–47,72	47,48	1	0,01	0,02
$\sum_i$	–	–	100	–	–

в) Строим полигон частот и гистограмму относительных частот (рис. 19.3, 19.4 соответственно; масштабы на осях берем разные).

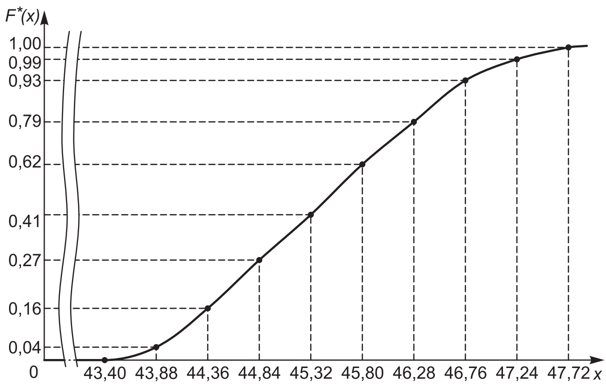
Находим значения эмпирической функции распределения $F^*(x) = n_x/n$: $F^*(43{,}40) = 0$, $F^*(43{,}88) = 0{,}04$, $F^*(44{,}36) = = 0{,}16$, $F^*(44{,}84) = 0{,}27$, $F^*(45{,}32) = 0{,}41$, $F^*(45{,}80) = = 0{,}62$, $F^*(46{,}28) = 0{,}79$, $F^*(46{,}76) = 0{,}93$, $F^*(47{,}24) = = 0{,}99$, $F^*(47{,}72) = 1$.



Р и с. 19.3



Р и с. 19.4



Р и с. 19.5

Строим график эмпирической функции распределения (рис. 19.5).

г) Находим выборочное среднее:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x'_i n_i$$

и выборочную дисперсию:

$$D_{\text{в}} = \frac{1}{n}\sum_{i=1}^{k} (x'_i - \bar{x})^2 n_i = \frac{1}{n}\sum_{i=1}^{k} (x'_i)^2 n_i - \bar{x}^2.$$

Для этого составляем расчетную таблицу (табл. 19.25). Из нее получаем:

$$\bar{x} = 4545{,}92/100 = 45{,}46,$$

$$D_{\text{в}} = 206\ 738{,}7/100 - 45{,}46^2 = 0{,}85, \ \sigma_{\text{в}} = \sqrt{D_{\text{в}}} = 0{,}92.$$

Таблица 19.25

m_i	Границы интервала x_i; x_{i+1}	Середина интервала x'_i	Частота интервала n_i	$n_i x'_i$	$(x'_i)^2$	$n_i(x'_i)^2$
1	43,40–43,88	43,64	4	174,56	1904,45	7617,80
2	43,88–44,36	44,12	12	529,44	1946,57	23 358,84
3	44,36–44,84	44,60	11	490,60	1989,16	21 880,76
4	44,84–45,32	45,08	14	631,12	2032,21	28 450,94
5	45,32–45,80	45,56	21	956,76	2075,71	43 589,91
6	45,80–46,28	46,04	17	782,68	2119,68	36 034,56
7	46,28–46,76	46,52	14	651,28	2164,11	30 297,54
8	46,76–47,24	47,00	6	282,00	2209,00	13 254,00
9	47,24–47,48	47,48	1	47,48	2254,35	2 254,35
$\sum_i$	–	–	100	4545,92	–	206 738,7

Выборочная дисперсия является *смещенной оценкой* генеральной дисперсии, а исправленная дисперсия – *несмещенной оценкой*:

$$\tilde{D}_{\text{в}} = \frac{n}{n-1} D_{\text{в}} = \frac{100}{99} \cdot 0{,}85 = 0{,}867 \,, \; \tilde{\sigma}_{\text{в}} = \sqrt{D_{\text{в}}} = 0{,}93 \,.$$

д) Согласно критерию Пирсона необходимо сравнить эмпирические и теоретические частоты. Эмпирические частоты даны. Найдем теоретические частоты. Для этого пронумеруем X, т.е. перейдем к СВ $z = (x - \bar{x})/\sigma_{\text{в}}$ и вычислим концы интервалов: $z_i = (x_i - \bar{x})/\sigma_{\text{в}}$, $z_{i+1} = (x_{i+1} - \bar{x})/\sigma_{\text{в}}$, причем наименьшее значение z, т.е. z_1, положим стремящимся к $-\infty$, а наибольшее, т.е. z_{m+1}, – к $+\infty$. Результаты занесем в таблицу (табл. 19.26). Так как $n_1 = 4 < 5$, то первый интервал объединяем со вторым и получаем интервал (43,40; 44,36) с частотой $n_1 = 16$. Далее объединим восьмой и девятый интервалы и получим интервал (46,76; 47,72) с частотой $n_7 = 7$.

Таблица 19.26

i	Границы интервала x_i; x_{i+1}		$x_i - \bar{x}$	$x_{i+1} + \bar{x}$	Границы интервала (z_i; z_{i+1})	
	x_i	x_{i+1}			$z_i = (x_i - \bar{x})/\sigma_{\text{в}}$	$z_{i+1} = (x_{i+1} - \bar{x})/\sigma_{\text{в}}$
1	43,40	44,36	–	–1,10	–	–1,19
2	44,36	44,84	–1,10	–0,62	–1,19	–0,67
3	44,84	45,32	–0,62	–0,14	–0,67	–0,15
4	45,32	45,80	–0,14	0,34	–0,15	0,37
5	45,80	46,28	0,34	0,82	0,37	0,89
6	46,28	46,76	0,82	1,30	0,89	1,40
7	46,76	47,72	1,30	–	1,40	–

Находим теоретические вероятности P_i и теоретические частоты: $n'_i = nP_i = 100P_i$. Составляем расчетную таблицу (табл. 19.27).

Таблица 19.27

i	Границы интервала $(z_i; z_{i+1})$		$\Phi(z_i)$	$\Phi(z_{i+1})$	$P_i = \Phi(z_{i+1}) - \Phi(z_i)$	$n'_i = 100P_i$
	z_i	z_{i+1}				
1	–	–1,19	–0,5000	–0,3830	0,1170	11,70
2	–1,19	–0,67	–0,3830	–0,2486	0,1344	13,34
3	–0,67	–0,15	–0,2486	–0,0596	0,1890	18,90
4	–0,15	0,37	–0,0596	0,1443	0,2039	20,39
5	0,37	0,89	0,1443	0,3133	0,1690	16,90
6	0,89	1,40	0,3133	0,4192	0,1059	10,59
7	1,40	–	0,4192	0,5000	0,0808	8,08
$\sum_i$	–	–	–	–	1	100

Вычислим наблюдаемое значение критерия Пирсона. Для этого составим расчетную таблицу (табл. 19.28). Последние два столбца служат для контроля вычислений по формуле

$$\chi^2_{\text{набл}} = \frac{1}{n}\sum_{i=1}^{k} n_i^2 - n.$$

Таблица 19.28

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$	n_i^2	$\frac{n_i^2}{n'_i}$
1	16	11,70	4,30	18,49	1,5803	256	21,8803
2	11	13,44	–2,44	5,9536	0,4430	121	9,0030
3	14	18,90	–4,90	24,01	1,2704	196	10,3704
4	21	20,39	0,61	0,3721	0,0182	441	21,6282
5	17	16,90	0,10	0,010	0,0006	289	17,1006
6	14	10,59	3,41	11,6281	1,0980	196	18,5080
7	7	8,08	–1,08	1,1664	0,1444	49	6,0644
$\sum_i$	100	100	–	–	$\chi^2_{\text{набл}} = 4,5549$	–	104,5549

Контроль: $\dfrac{\sum n_i^2}{n'_i} - n = \dfrac{\sum (n_i - n'_i)^2}{n} = 104{,}5549 - 100 =$ $= 4{,}5549$. По таблице критических точек распределения χ^2 (см. прил. 10), уровню значимости $\alpha = 0{,}01$ и числу степеней свободы $k = l - 3 = 7 - 3 = 4$ (l – число интервалов) находим: $\chi^2_{\text{кр}} = 13{,}3$.

Так как $\chi^2_{\text{набл}} < \chi^2_{\text{кр}}$, то гипотеза H_0 о нормальном распределении генеральной совокупности принимается.

е) Если СВ X генеральной совокупности распределена нормально, то с надежностью γ можно утверждать, что математическое ожидание a СВ X покрывается доверительным интервалом $\left(\bar{x} - \dfrac{\tilde{\sigma}_{\text{в}}}{\sqrt{n}} t_\gamma;\ \bar{x} + \dfrac{\tilde{\sigma}_{\text{в}}}{\sqrt{n}} t_\gamma\right)$, где $\delta = \dfrac{\tilde{\sigma}_{\text{в}}}{\sqrt{n}} t_\gamma$ – *точность оценки.* В нашем случае $\bar{x} = 45{,}46$, $\tilde{\sigma}_{\text{в}} = 0{,}93$, $n = 100$. Из прил. 4 для $\gamma = 0{,}95$ находим: $t_\gamma = 1{,}984$, $\delta = 0{,}1843$. Доверительным интервалом для a будет (45,2757; 45,6443). Доверительный интервал, покрывающий среднее квадратичное отклонение σ с заданной надежностью γ, $(\tilde{\sigma}_{\text{в}}(1 - q)$; $\tilde{\sigma}_{\text{в}}(1 + q))$, где q находится по данным γ и n из прил. 9. При $\gamma = 0{,}95$ и $n = 100$ имеем: $q = 0{,}143$. Доверительным интервалом для σ будет (0,7970; 1,0630). ◂

ИДЗ-19.2

Дана таблица распределения 100 заводов по производственным средствам X (тыс. ден. ед.) и по суточной выработке Y (т). Известно, что между X и Y существует линейная корреляционная зависимость. Требуется:

а) найти уравнение прямой регрессии y на x;

б) построить уравнение эмпирической линии регрессии и случайные точки выборки (X, Y).