Chapter IV

Measures of Dispersion (Solutions)

4. 1.

Working Table

b_i	B_i	f_i	$\sum_{j=1}^{i} f_j$	m_i	w_i	h_i
000	200	0.155	0.155	100	200	0.0008
200	300	0.356	0.511	250	100	0.0036
300	400	0.090	0.601	350	100	0.0009
400	500	0.066	0.667	450	100	0.0007
500	600	0.058	0.725	550	100	0.0006
600	800	0.116	0.841	700	200	0.0006
800	1200	0.159	1.000	1000	400	0.0004

1.

$$F(x) = \begin{cases} 0.000 & \text{forr} \quad -\infty < x \le 100 \\ 0.155 & \text{for} \quad 100 < x \le 250 \\ 0.511 & \text{for} \quad 250 < x \le 350 \\ 0.601 & \text{for} \quad 350 < x \le 450 \\ 0.667 & \text{for} \quad 450 < x \le 550 \\ 0.725 & \text{for} \quad 550 < x \le 700 \\ 0.841 & \text{for} \quad 700 < x \le 1000 \\ 1.000 & \text{for} \quad 1000 < x < +\infty \end{cases}$$

$F(700) - F(250) = 0.725 - 0.155 = 0.570$.

Nearly 57% of the rainfall is at least 250 mm but less than 700 mm.

2.

$$\tilde{x}_{0.25} \in C_2$$

$$x_{0.25} \approx 200 + \frac{0.25 - 0.155}{0.356} \cdot 100 \approx 226.69$$

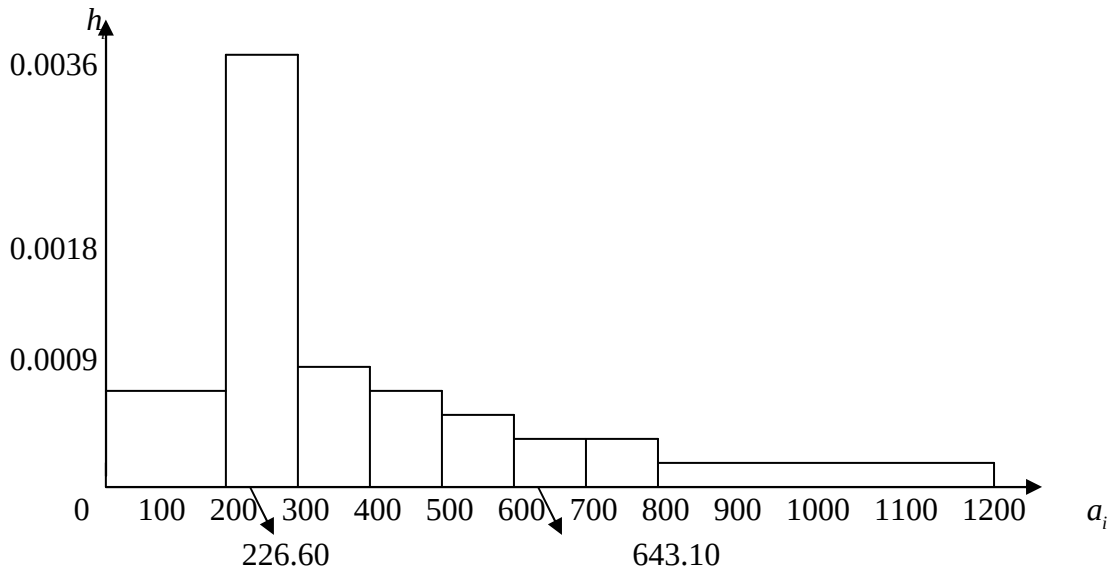
Nearly 25% of the rainfall is less than 226.69 mm.

$x_{0.75} \in C_6$.

$$x_{0.75} \approx 600 + \frac{0.75 - 0.725}{0.116} \cdot 200 \approx 643.10$$

Nearly 75% of the rainfall is less than 643.10 mm.

3.



The area between 0 and 266.60 is equal to 0.25 and the area between 0 and 643.10 is equal to 0.75.

4. 2.

Working Table

b_i	B_i	F_i	f_i	$\sum_{j=1}^{i} f_j$	w_i	$h_i := \frac{f_i}{w_i}$	m_i	$m_i \cdot F_i$	$b_i \cdot F_i$	$B_i \cdot F_i$	$(m_i - \mu)^2 \cdot F_i$
0	10	32	0.256	0.256	10	0.026	5.0	160.0	0	320	30356.48
10	20	40	0.320	0.576	10	0.032	15.0	600.0	400	800	17305.60
20	45	25	0.200	0.776	25	0.008	32.5	812.5	500	1125	272.25
45	90	13	0.104	0.880	45	0.002	67.5	877.5	585	1170	13063.57
90	180	15	0.120	1.000	90	0.001	135.0	2025.0	1350	2700	147609.60
		125	1.000					4475	2835	6115	208607.50

1.

$$F(x)=\begin{cases}0.000 & \text{when} \quad -\infty < x \le 5.0\\ 0.256 & \text{when} \quad 5.0 < x \le 15.0\\ 0.576 & \text{when} \quad 15.0 < x \le 32.5\\ 0.776 & \text{when} \quad 32.5 < x \le 67.5\\ 0.880 & \text{when} \quad 67.5 < x \le 135.0\\ 1.000 & \text{when} \quad 135.0 < x < +\infty\end{cases}$$

$F(23)=0.576$.

Nearly 57.6 % of the waiting time is shorter than 23 minutes.

2.

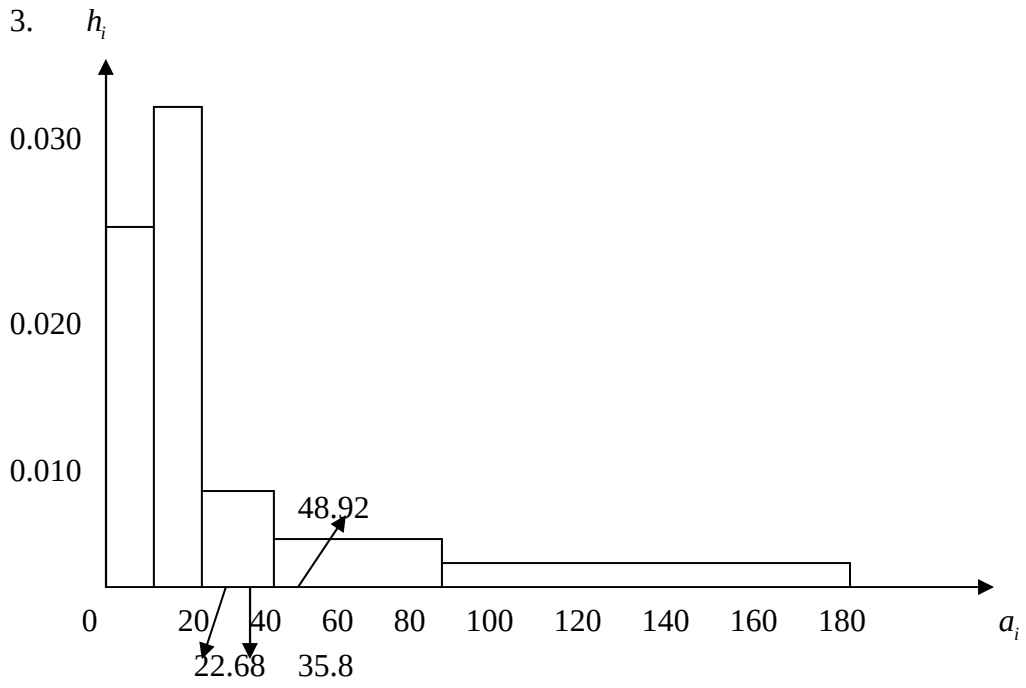
$$\mu \approx \frac{2836}{125} \approx 35.8, \qquad \frac{2836}{125} = 22.68 \le \mu \le 48.92 = \frac{6115}{125}$$

$$\sigma^2 \approx \frac{208607.5}{125} \approx 1668.86, \qquad \sigma \approx 40.85,$$

$$\mu \in [35.8-40.85, \quad 35.8+40.85] = [-5.05. \quad 76.65]$$

At least 50% of the waiting times lie in the above interval .

3.



4. 3.

a)

$n = 20$

b)

Number of close friends	Frequency	Relative frequency
0	2	0.10
1	3	0.15
2	5	0.25
3	5	0.25
4	4	0.20
5	1	0.05
Total	20	1.00

c)

50% of the students had 2 or fewer close friends.

d)

mean = 2.45 median = 2.5 standard deviation s = 1.39

4. 4.

1.

The variable is the height of 200 participants. It is a continuous variable.

2.

Working Table

b_i	B_i	F_i	f_i	w_i	h_i	$\sum_{j=1}^{i} f_j$	m_i
1.40	1.54	12	0.06	0.14	0.43	0.06	1.47
1.54	1.68	38	0.19	0.14	1.36	0.25	1.61
1.68	1.78	50	0.25	0.10	2.50	0.50	1.73
1.78	1.88	58	0.29	0.10	2.90	0.79	1.83
1.88	1.98	30	0.15	0.10	1.50	0.94	1.93
1.98	2.08	12	0.06	0.10	0.60	1.00	2.03
		200	1.00				

Exercise: Draw the histogram!

3.

$$F(x)=\begin{cases}0.00 & \text{when} \quad -\infty < x \le 1.47\\ 0.06 & \text{when} \quad 1.47 < x \le 1.61\\ 0.25 & \text{when} \quad 1.61 < x \le 1.73\\ 0.50 & \text{when} \quad 1.73 < x \le 1.83\\ 0.79 & \text{when} \quad 1.83 < x \le 1.93\\ 0.94 & \text{when} \quad 1.93 < x \le 2.03\\ 1.00 & \text{when} \quad 2.03 < x < +\infty\end{cases}$$

$F(1.75)=0.50,$ i.e. 50%.

4.

$F(1.90)-F(1.60)=0.79-0.06=0.73$, i.e. 73%.

5.

Working Table (contin.)

$b_i \cdot F_i$	$m_i \cdot F_i$	$B_i \cdot F_i$	$(m_i-1.77)^2 \cdot F_i$
16.80	17.64	18.48	1.069943520
58.52	61.18	63.84	0.955850480
84.00	86.50	89.00	0.074498000
103.24	106.14	109.04	0.218657680
56.40	57.90	59.40	0.781498800
23.76	24.36	24.96	0.819959520
342.72	353.72	364.72	3.920408000

$$\bar{x} \approx \frac{353.72}{200} \approx 1.77\,\text{m}, \qquad 1.7136=\frac{342.72}{200} \le \bar{x} \le \frac{364.72}{200}=1.8236\,.$$

$$s^2=\frac{3.920408000}{200}=0.0197005, \qquad s=0.14003584 \approx 0.14\,,$$

$$v=\frac{s}{\bar{x}}=\frac{0.14}{1.77}=0.0791014115 \approx 0.079 < 0.5\,.$$

The average 1.77m is, therefore, representative.

4. 5.

1.

The variable is the age of 100 guests of a hotel on a certain day. It is a (theoretically) a continuous variable.

2.

Working Table

b_i	B_i	F_i	f_i	w_i	h_i	$\sum_{j=1}^{i} f_j$	m_i	$m_i \cdot F_i$	$(m_i - \bar{x})^2 \cdot F_i$
1	20	20	0.20	19	0.105	0.20	10.5	210	13107.20
20	40	38	0.38	20	0.019	0.58	30.0	1140	1413.98
40	60	34	0.34	20	0.017	0.92	50.0	1700	6569.14
60	80	8	0.08	20	0.004	1.00	70.0	560	9193.68
		100	1.00					3610	30284.00

Exercise: Draw the histogram!

3.

$$F(x) = \begin{cases} 0.00 & \text{for} \quad -\infty < x \le 10.5 \\ 0.20 & \text{for} \quad 10.5 < x \le 30.0 \\ 0.58 & \text{for} \quad 30.0 < x \le 50.0 \\ 0.92 & \text{for} \quad 50.0 < x \le 70.0 \\ 1.00 & \text{for} \quad 70.0 < x < +\infty \end{cases}$$

$F(45) = 0.58,$ i.e. 58%.

4.

$F(42) - F(27) = 0.58 - 0.20 = 0.38,$ i.e. 38%.

5.

$$\bar{x} \approx \frac{3610}{100} = 36.10\,, \qquad s^2 \approx \frac{30284.00}{99} = 305.90\,, \qquad s \approx 17.49\,.$$

Because of $v \approx \frac{17.49}{36.10} \approx 0.48 < 0.50$ the mean is representative.

4. 6.

1.

The variable is the weekly wages of 100 entry-level workers in a firm. It is a (theoretically) a continuous variable.

2.

Working Table

b_i	B_i	F_i	f_i	w_i	h_i	$\sum_{j=1}^{i} f_j$	m_i	$m_i \cdot F_i$	$(m_i - \bar{x})^2 \cdot F_i$
240	260	7	0.07	20	0.0035	0.07	250	1750	14175
260	280	20	0.20	20	0.0100	0.27	270	5400	12500
280	300	33	0.33	20	0.0165	0.60	290	9570	825
300	320	25	0.25	20	0.0125	0.85	310	7750	5625
320	340	11	0.11	20	0.0055	0.96	330	3630	13475
340	360	4	0.04	20	0.0020	1.00	350	1400	12100
		100	1.00					29500	58700

Exercise: Draw the histogram!

3.

$$F(x) = \begin{cases} 0.00 & \text{for} \quad -\infty < x \le 250 \\ 0.07 & \text{for} \quad 250 < x \le 270 \\ 0.27 & \text{for} \quad 270 < x \le 290 \\ 0.60 & \text{for} \quad 290 < x \le 310 \\ 0.85 & \text{for} \quad 310 < x \le 330 \\ 0.96 & \text{for} \quad 330 < x \le 350 \\ 1.00 & \text{for} \quad 350 < x < +\infty \end{cases}$$

$F(300) = 0.60,$ i.e. 60%.

4.

$F(310.50) - F(250.45) = 0.85 - 0.07 = 0.78,$ i.e. 78%.

5.

$$\bar{x} \approx \frac{29500}{100} = 295.00\,€, \qquad s^2 \approx \frac{58700}{99} = 592.93\,,\ s \approx 24.35\,.$$

Because of $v \approx \frac{24.35}{295} \approx 0.083 < 0.500$ the mean is representative.

4. 7.

Working Table

x_i	$(x_i - \bar{x})^2$
8408	10481252.0800
1374	14413593.0400
1872	10880259.3300
8879	13752795.6500
2459	7352361.3720
11413	38968508.9800
608	20816623.5200
14138	80415629.2200
6452	1642181.2260
1850	11025878.3700
2818	5534368.2770
1356	14550591.9000
10498	28382002.5600
7478	5324446.3670
4019	1326007.0850
4341	688109.7514
739	19638403.2800
2127	9263037.1820
3653	2302878.5140
5794	388722.5595
8305	9824940.9860
108581	306972591.2000

1.

$$\bar{x} = \frac{108581}{21} = 5170.52381$$

2.

$$s^2 = \frac{306972591.2000}{20} = 15348629.56, \qquad s = 3917.732707$$

At least 50% of the annual sales lie in the interval

$$[5170.52381 - 3917.732707,\ 5170.52381 + 3917.732707] = [1252.79110,\ 9088.25652]$$

3.

$$v = \frac{3917.732707}{5170.52381} = 0.757705186 > 0.5.$$

Therefore, the mean is not representative.

4. 8.

1.

$$\bar{x} = \frac{139.96 + 141.06 + \ldots + 128.94}{14} = \frac{1946.66}{14} = 139.047143.$$

2. – 3.

$[b_i$	$B_i[$	F_i	m_i	$m_i \cdot F_i$
128	131	2	129.5	259.0
131	134	0	132.5	0.000
134	137	3	135.5	406.5
137	140	2	138.5	277.0
140	143	3A	141.5	424.5
143	146	4	144.5	578.0
				1945.0

$$\tilde{x} \approx \frac{1945}{14} \approx 138.9286$$

4.

The *approximate* mean of the grouped data set differs from the *exact* mean of the ungrouped data set. The reason is that the calculation of the mean of the grouped data set is based on the assumption that the observations in each class are equally dispersed. This is obviously not the case in our example.

(Last revised: 14.08.10)