度数分布表の作成

(1) データ数、データの最大値と最小値から、階級の数を決める。
(2) （最大値－最小値）/ 階級数　から、きりのよい値の階級幅を決める
(3) 個々のデータについて、どの階級に含まれるかをカウントする。
(4) (3)の個数の計を度数に記入する。

度数分布表における階級数の決め方

(1) シャリエ(Charlier)
　大量観察から理想分布を考慮し、級間は標準偏差の 1/3 が望ましい。

(2) フィッシャ(Fisher)
　少数統計を考慮し、級間は標準偏差の 1/4 が望ましい。

(3) スタージェス(Sturges)
　Ｎこのデータからなる集団の階級個数ｎは、ｎ＝１＋logＮ / log２

(4) 度数検定を考慮し、平均 10 以上が望ましい。

(3),(4)の併用により、Ｎ＝２ｍとおき、ｎ＝１＋logＮ / log２　に代入することにより、ｎ＝１＋m、これにより、Sturges の階級数を得る。

Ｎ	３２	６４	１２８	２５６	５１２	１０２４	２０４８	４０９６	８１９２
ｎ	６	７	８	９	１０	１１	１２	１３	１４

	階級	度数(f)	階級値(x)	fx	u	fu	fu^2	x^2f
	40 ～ 45	10	42.5	425.0	−2	−20	40	18062.50
	45 ～ 50	17	47.5	807.5	−1	−17	17	38356.25
(仮平均)	50 ～ 55	36	52.5	1890.0	0	0	0	99225.00
	55 ～ 60	22	57.5	1265.0	1	22	22	72737.50
	60 ～ 65	15	62.5	937.5	2	30	60	58593.75
	計(Σ)	100		5325.0		15	139	286975.00

$\bar{x}$ = G8/E8
= 53.25

$\bar{x}$ = F5+I8/E8×(D3−B3)
= 53.25

S = (K8−I8^2/E8)×(D3−B3)^2
= 3418.75

σ^2 = 34.1875

σ^2 = N8/E8−E11^2
= 34.1875

度数分布表からの代表値の求め方

階級	中点値	度数	累積度数
40 ～ 45	42.5	10	10
45 ～ 50	47.5	17	27
50 ～ 55	52.5	36	63
55 ～ 60	57.5	22	85
60 ～ 65	62.5	15	100
		100	

階級	中点値	度数	累積度数
$c_1 \sim c_2$	x_1	f_1	F_1
$c_2 \sim c_3$	x_2	f_2	F_2
$c_3 \sim c_4$	x_3	f_3	F_3
$c_4 \sim c_5$	x_4	f_4	F_4
$c_5 \sim c_6$	x_5	f_5	F_5
		n	

$h = c_{j+1} - c_j$

①算術平均

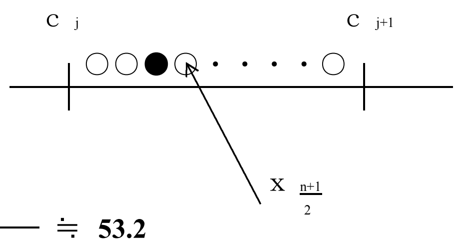
$$\bar{x} = \frac{\Sigma x_j \times_j}{\Sigma f_j}$$

$$= \frac{42.5 \times 10 + 47.5 \times 17 + 52.5 \times 36 + 57.5 \times 22 + 62.5 \times 15}{10 + 17 + 36 + 22 + 15}$$

$$= \mathbf{53.25}$$

②中央値

中央値を含む階級（$c_j \sim c_{j+1}$）を按分して求める。

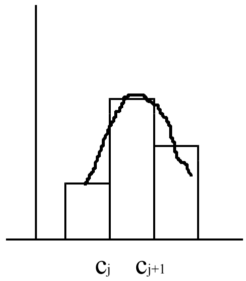


$F_j < \dfrac{n}{2} \leqq F_{j+1}$ となるので、

$$\tilde{x} = c_j + \frac{\left(\frac{n}{2} - F_{j-1}\right) h}{f_j} = 50 + \frac{(50 - 27) \times 5}{36} \fallingdotseq \mathbf{53.2}$$

③最頻値

最多度数 f_j の前後を考え、このあたりの度数分布が放物線と見なす。



$$\mathring{x} = c_j + \frac{(f_j - f_{j-1}) h}{2 f_j - f_{j-1} - f_{j+1}} = 50 + \frac{(36 - 17) \times 5}{2 \times 36 - 17 - 22} \fallingdotseq \mathbf{52.9}$$

④四分位偏差

資料 x_j がこの集団の q ％に位置するとき、この x_j を q ％ ile（または第 q 百分位数）という。

いま、百分比 q が与えられ、$F_j < \dfrac{nq}{100} \leqq F_{j+1}$ ならば、百分位数 P_q は以下のようになる。

$$P_q = P_j + \frac{\left(\frac{q}{100} \times n - F_j\right) h}{f_j}$$

q ＝ 50 のとき、中央値

q ＝ 25 のとき、第一四分位

q ＝ 75 のとき、第三四分位という。

$$\frac{q_1 n}{100} = \frac{25 \times 100}{100} = 25$$

$$\frac{q_3 n}{100} = \frac{75 \times 100}{100} = 75$$

$Q_1 = 45 + (25 - 10) \times 5 / 17 \fallingdotseq 49.4$

$Q_3 = 55 + (75 - 63) \times 5 / 22 \fallingdotseq 57.7$

$\therefore Q = (Q_3 - Q_1) / 2$

$= (57.7 - 49.4) / 2$

$= \mathbf{4.15}$