

宿題解答 (第3章)

[問3]

$$\begin{aligned} \text{ラスパイレス物価指数} &= \frac{\sum_i p_{1i} q_{0i}}{\sum_i p_{0i} q_{0i}} \\ &= \frac{500 \times 1000 + 700 \times 2000 + 1200 \times 500}{300 \times 1000 + 500 \times 2000 + 1000 \times 500} \\ &= 139\% \end{aligned}$$

$$\begin{aligned} \text{パーシェ物価指数} &= \frac{\sum_i p_{1i} q_{1i}}{\sum_i p_{0i} q_{1i}} \\ &= \frac{500 \times 1500 + 700 \times 2800 + 1200 \times 600}{300 \times 1500 + 500 \times 2800 + 1000 \times 600} \\ &= 140\% \end{aligned}$$

[問4]

加重平均は次式のかたちで定義される.

$$\text{加重平均} = \sum_{i=1}^n w_i x_i,$$

ここで, w_i は非負値であり, 総和は1である.
まず, $(2m+1)$ 項移動平均を考える.

$$(2m+1) \text{ 項移動平均} = \frac{1}{2m+1} \sum_{i=-m}^m x_{t+i}$$

となるので,

$$w_i = \frac{1}{2m+1} > 0, \quad (i = 0, \pm 1, \pm 2, \dots, \pm m)$$

とすると,

$$\sum_{i=-m}^m w_i = (2m+1) \times \frac{1}{2m+1} = 1$$

となる. よって, $(2m+1)$ 項移動平均は加重平均の特殊な場合であることがわかる.
一方, 中心化 $2m$ 項移動平均を考えてみると,

$$\text{中心化 } 2m \text{ 項移動平均} = \frac{1}{4m} (x_{t-m} + x_{t+m}) + \frac{1}{2m} \sum_{i=-m+1}^{m-1} x_{t+i}$$

であるので,

$$w_{-m} = w_m = \frac{1}{4m},$$

$$w_i = \frac{1}{2m} > 0, \quad (i = 0, \pm 1, \pm 2, \dots, \pm(m-1))$$

とすると,

$$\sum_{i=-m}^m w_i = 2 \times \frac{1}{4m} + \frac{1}{2m} \times (2m-1) = 1$$

となることがわかる. したがって, 中心化 $2m$ 項移動平均も加重平均の特殊な場合である.

[問 6]

受験者数が n の場合, 次の結果が得られたとする.

投票番号	0	1	2	...	i	...	$n-1$	n
票数	V_0	V_1	V_2	...	V_i	...	V_{n-1}	V_n
累積票数	V_0	$V_0 + V_1$	$V_0 + V_1 + V_2$...	$\sum_{j=0}^i V_j$...	$\sum_{j=0}^{n-1} V_j$	$\sum_{j=0}^n V_j$

ここで, $\sum_{j=0}^n V_j = 67$. メディアンまでを合格とする方式で, 上位 i 位までの受験者を合格させることにすると,

$$\begin{cases} \sum_{j=0}^{i-1} V_j < \frac{1}{2} \sum_{j=0}^n V_j = 33.5 \\ \sum_{j=0}^i V_j > \frac{1}{2} \sum_{j=0}^n V_j = 33.5, \end{cases}$$

即ち,

$$\begin{cases} \sum_{j=0}^{i-1} V_j \leq 33 \\ \sum_{j=0}^i V_j \geq 34. \end{cases}$$

となる. 一方, 受験者一人ずつを合否の投票にかける場合, 上の結果から, 以下のように復元することができる.

順位成績	1	2	...	i	...	$n-1$	n
{合格}票	$\sum_{j=1}^n V_j$	$\sum_{j=2}^n V_j$...	$\sum_{j=i}^n V_j$...	$\sum_{j=n-1}^n V_j$	V_n
{不合格}票	V_0	$\sum_{j=0}^1 V_j$...	$\sum_{j=0}^{i-1} V_j$...	$\sum_{j=0}^{n-2} V_j$	$\sum_{j=0}^{n-1} V_j$

このとき, 上位 i 位の受験者の”合格”票数は $\sum_{j=i}^n V_j$ で, $\sum_{j=0}^{i-1} V_j \leq 33$ より, $\sum_{j=i}^n V_j \geq 34$ であることがわかる. そして, 上位 $i+1$ 位の受験者の”合格”票数は $\sum_{j=i+1}^n V_j$ で, $\sum_{j=0}^i V_j \geq 34$ より, $\sum_{j=i+1}^n V_j \leq 33$ となる. 多数決によって, 上位 i 位までの受験者が合格になる.

これは, 二つの方法が同じ結果をもたらすことを意味する.

[問 11]

$$A(i) = \{i : |x_i - \bar{x}| \geq ks_x\},$$

$$B(i) = \{i : |x_i - \bar{x}| < ks_x\}$$

とすると,

$$\begin{aligned} (n-1)s_x^2 &= \left(\sum_{A(i)} + \sum_{B(i)}\right)(x_i - \bar{x})^2 \\ &\geq \sum_{A(i)}(x_i - \bar{x})^2 \\ &\geq \sum_{A(i)}k^2s_x^2 \\ &= n_k k^2 s_x^2 \end{aligned}$$

が成り立つので,

$$\frac{1}{k^2} \geq \frac{n_k}{n-1} > \frac{n_k}{n}$$

を得る. よって, チェビシェフの不等式が証明された.

[問 13]

ヒストグラムを作成する場合の各階級の下限と上限を見てみよう.

	$y_i (= 50 + 10z_i)$ に対して	$z_i (= (x_i - \bar{x})/s_x)$ に対して	x_i に対して
図 2-3			[-15, -10) [-10, -5) [-5, 0) [0, 5) [5, 10) [10, 15)
図 3-4		[-3, -2) [-2, -1) [-1, 0) [0, 1) [1, 2) [2, 3)	[-16.00, -10.88) [-10.88, -5.76) [-5.76, -0.63) [-0.63, 4.50) [4.50, 9.62) [9.62, 14.74)
図 3-5	[20, 30) [30, 40) [40, 50) [50, 60) [60, 70) [70, 80)	[-3, -2) [-2, -1) [-1, 0) [0, 1) [1, 2) [2, 3)	[-16.00, -10.88) [-10.88, -5.76) [-5.76, -0.63) [-0.63, 4.50) [4.50, 9.62) [9.62, 14.74)

図 3-4 と 図 3-5 の x_i に対する各階級の下限と上限が一致するので, 形状も同一である. また, 図 2-3 の x_i に対する各階級の下限と上限は, 図 3-4 及び 図 3-5 と違うので, 形状も異なる.

[問 14]

	範囲	平均	標準偏差	変動係数
	$\max\{x_i\} - \min\{x_i\}$	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	$s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$	s_x / \bar{x}
A	21	138.4	5.91	0.0427
B	27	171.1	7.22	0.0422

変動係数の観点から、データ Aの方がばらつきが大きいと言える。