◆◆◆◆◆ 第2章 練習問題 ◆◆◆◆◆

2.1 ＜代表値の間隔＞　所得だけでなく，貯蓄などの多くの経済量の分布が右裾の長い非対称な形をもっている．このとき，Mo（モード），Me（メディアン），平均 $\bar{x}$ の間隔を

$$Me-Mo \fallingdotseq k(\bar{x}-Me)$$

で表してみよう．経済統計の教科書，参考書で k の値を調べてみよ．

（参考書）中村，新家，美添，豊田『統計入門』，同『経済統計入門』，津村，淵脇，築林『社会統計入門』（以上，東京大学出版会），養谷『統計のはなし』東京図書，Encyclopedia of Statistics.

2.2 ＜ジニ係数＞　1次元データ $x_1, x_2, \cdots, x_n$ の散らばりの尺度として，データの各対ごとの隔り $|x_i-x_j|$ の平均

$$\sum_i \sum_j |x_i-x_j|/n^2$$

を考え，これを平均差という．また，平均差と平均値 $\bar{x}\times 2$ の比を

$$GI=\sum_i \sum_j |x_i-x_j|/2n^2\bar{x}$$

のように定義し，ジニ係数という．不平等度の指標として用いられる．

35ページのA, B, Cのデータに対して，平均差，ジニ係数を計算せよ．

---

*) S_x のかわりに，(9.5)の不偏分散 s^2 による s を用いることが多い．

2.3 〈エントロピー〉　カテゴリー別に分類されたデータにおいて，各カテゴリーの相対頻度を $p_i=f_i/n$ としよう．いま

$$H(p_1, p_2, \cdots, p_k)=-\sum_{i=1}^{k} p_i \log p_i$$

と定義すると，

i）$H(p_1, p_2, \cdots, p_k) \geqq 0$

ii）どれかの $p_i=1$ のとき，$H(p_1, p_2, \cdots, p_k)=0$

iii）$H(p_1, p_2, \cdots, p_k) \leqq H(1/k, 1/k, \cdots, 1/k)$　（最大値）

が成り立つ．この量を**エントロピー**という．H の大きいほど分布は一様となり，小さいほど集中性が見られる．（なお，$0\log 0=0$ と約束する．）

100人の学生に出身地をたずねた．10年前と本年にたいして下の結果を得た．集中性という見地から，この出身地の分布を比較せよ．（架空例）

地　域	A	B	C	D	E	計
本　年	32	19	10	24	15	100
10年前	28	13	18	29	12	100

2.4 〈諸得点〉　2.3節のデータBについて，標準得点，偏差値得点を計算せよ．