

## 表計算 (Excel) とデータ処理 (10)

## まとめ

## 1. 尺度水準：計算方法からのデータの区別

名義尺度：等しい／等しくないだけわかるデータ（名前，車のナンバーなど）

順序尺度：上記に加えて，順序（大小前後など）がわかるデータ（名簿，競争の順位など）

間隔尺度：上記に加えて，加減しても意味のある（差が意味のある）データ（摂氏温度など）

比例尺度：上記に加えて，乗除しても意味のある（割合が意味のある）データ（身長，絶対温度など）

## 2. ヒストグラム（度数分布）

データの集まりを，カテゴリや部分範囲で分けしたとき，その範囲のデータ数を整理したもの。

階級：それぞれの分けのこと。

階級値：階級を示す値のこと。階級に範囲はある場合は，その範囲の真ん中を階級値とすることが多い。

〔課題〕 各尺度について，ヒストグラムの例を挙げなさい。

## 3. データの要約（要約統計量）

## 3.1 間隔尺度以上の要約統計量

## 平均 [ 代表値 ]

算術平均：データ値の合計をデータ数で割ったもの。

## 分散、標準偏差 [ 散布度 ]

分散：平均からの差（偏差）の二乗和（二乗した値の合計）をデータ数で割ったもの。

データの二乗和をデータ数で割ったものから平均の二乗を引いても求められる。

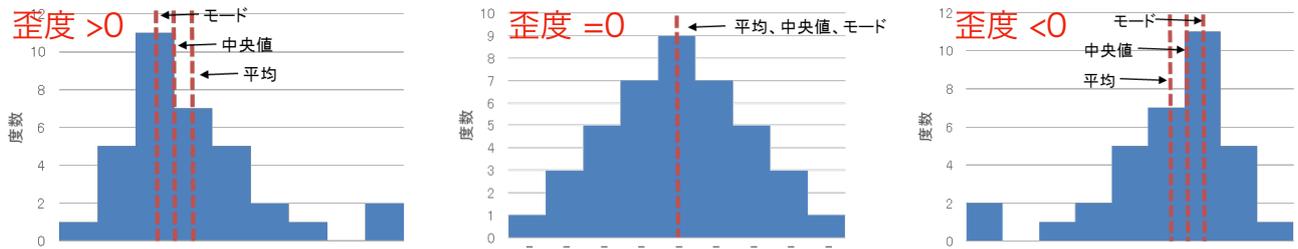
標準偏差：分散の正の平方根。

## 歪度 [ 散布度 ]（平均偏差の3乗の平均を標準偏差の3乗で割った値）

度数分布で左右比対称の違いを表す。

歪度が 0 の時は左右対称の場合が多く、正の場合は裾が正方向に広がり、負の場合は負方向に広がっていることを表す。「正規分布」の場合、歪度は 0。

課題：歪度が、負, 0, 正のグラフを描きなさい。



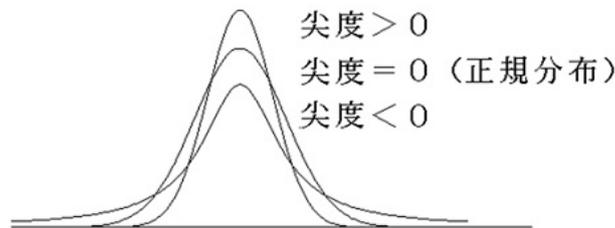
尖度 [ **散布度** ] (平均偏差の4乗の平均を標準偏差の4乗で割った値から3を引いた値)

度数分布で山の尖り(裾の広がり)の程度を表す

尖度が大きいほど尖りが急で裾野が広がっていることを表す

「正規分布」の場合、尖度は 0 になる。

課題：尖度が大きいグラフと小さいグラフを描きなさい。



### 3.2 順序尺度以上の要約統計量

中央値 (メジアン、メディアン (median)) [ **代表値** ]

データを並べたときに、真ん中の順序になるデータの値。第2四分位でもある。

第1四分位・第3四分位 [ **散布度** ]

データを小さい順に並べたときに、1/4, 3/4 の順位にあるデータの値。

最小値・最大値 [ **散布度** ]

データを並べたときの、最大、最小の値。

### 3.3 名義尺度以上の要約統計量

最頻値 (モード (mode)) [ **代表値** ]

度数分布で最も高い度数を示す階級の階級値。

[課題]

上記の統計量について、データ群の代表としての統計量 (代表値) と、データ群の広がりを示す統計量 (散布度) に区別し [ ] に書き入れなさい。