

統計・分析入門③

I コンピュータを用いた分析

(1) 表計算(スプレッドシート)ソフトウェア

数学の授業等しているようなデータを処理するための計算式を学習し、実際に計算を自分で行ってきたが、データの数が多くなったり、データの桁数が多くなったりすると自分で計算するのは大変である。

そのこともあり、実際にデータに関する処理を行う際には「表計算(スプレッドシート)ソフトウェア」を利用することがほとんどである。この表計算ソフトウェアは、次のようなことができる。

- ・自動的に計算する
- ・グラフを描く
- ・表によるデータ処理をする

有名なソフトウェアとしては、Microsoft 社が販売している「Excel」がある。この時間では「Excel」を使い、計算やグラフの作成、データの整理などを行う。



※ 実習ファイル「Ixxx_Excel トレーニング.xlsx」について

- ① 「PC」→「Read-Only」→「探究」→「1年」→「統計分析入門」にある実習ファイルを選択したあと、「右クリック」→「送る」→「s191xxx」をクリックする
- ② 「PC」→「s191xxx」を開く
- ③ 「Ixxx_Excel トレーニング.xlsx」を「右クリック」し、「名前の変更」から「xxx」を「クラス」「出席番号(2桁)」に変更する

(2) 計算式の入力 【⇒実習シート：計算基礎】

Excel で計算式を入力する場合は、必ず最初に(①)を入力する。四則演算は、次の記号を用いて行う。なお、入力は(②)で行う。

- たし算+ → (③) ひき算- → (④)
- かけ算× → (⑤) わり算÷ → (⑥)

問 Excel のセル F6 で「3+5」の答えを求める計算式を入力する方法を考えよう

(1) 直接数値を入力して計算させる

(2) 「3」と「5」が入力されたセルの番地を使って計算させる

Excel では、繰り返し利用できることなどを考え、問(2)のようなセル番地を使った計算式を使うのが一般的である。

(3) 関数とは? 【⇒実習シート：処理】

Excel では、必要な数値などを指定するだけでカンタンに計算結果を求められる「関数」という機能であり、次のように表現される。

$$=関数名(引数1, 引数2, 引数3, …… , 引数n)$$

※ (⑦) : 数値やセル番地、数値を入力した範囲、関数などのことである

- ・ 複数の引数がある場合は、「,」で区切って並べる
- ・ セル番地の範囲を指定する場合は、矩形領域(長方形上の領域)で、左上端と右下端のセル番地を「:」で区切って表す

※ 「=」や「(, , , ,)」はすべて半角文字

● 関数の挿入方法(一般的なもの)

①「ホーム」タブの「編集」グループ「オート SUM」→「その他の関数」

②出てきたウィンドウから、関数を選択
※関数の分類:「すべて表示」

③出てきたウィンドウで、引数などを設定

なお、①②については、「数式」タブの「関数ライブラリ」グループからある程度選択できる。



● データの分析で使う関数

I 変数のデータ

| | 関数名 | 特別な引数など |
|------|--------------|--|
| 平均値 | AVERAGE | |
| 中央値 | MEDIAN | |
| 最大値 | MAX | |
| 最小値 | MIN | |
| 四分位数 | QUARTILE.INC | 「戻り値」の設定は、以下のようになる 「1」…第1四分位数、「2」…第2四分位数 「3」…第3四分位数 (注意) Excel では、教科書と異なる定義で四分位数を計算している【→赤チャート P.297】 |
| 分散 | VAR.P | |
| 標準偏差 | STDEV.P | |

II 2変数のデータ

| | 関数名 | 特別な引数など |
|------|--------------|------------------------|
| 共分散 | COVARIANCE.P | 「配列1」には「1種類目のデータ」、 |
| 相関係数 | CORREL | 「配列2」には「2種類目のデータ」を指定する |

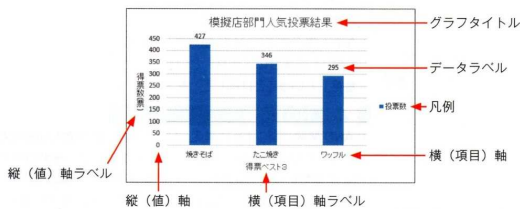
(4) グラフをかき 【⇒実習シート：処理】

● グラフのかき方(一般的なもの)

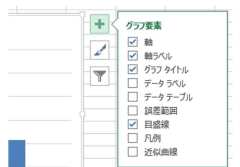
①グラフをかきたいデータの部分を選択

②「挿入」タブの「グラフ」グループからかきたいグラフを選択

● グラフの主な要素



Excel では、グラフでどのような要素を表示させるかを、グラフの右上にある「+」ボタンで選択することができる。



● ヒストグラムをかこう

①軸で右クリック

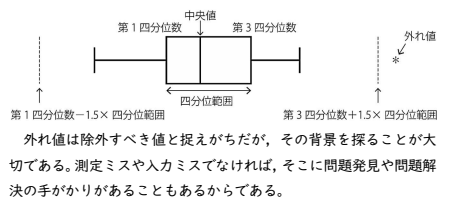
②軸の書式設定

③「BIN」の設定で「区間の幅」や「BINの数」を指定できる

● 箱ひげ図をかこう

【参考】外れ値 …他の値から極端にかけ離れたデータ

四分位範囲の1.5倍以上離れた値とされる



● 散布図をかこう