Excelで度数分布表を作成

表計算ソフトの「Microsoft Excel」を使って、度数分布表を作成する場合、関数を使わなくても、四則演算（＋－＊／）だけでも作成できます。

しかし、データ数が多い場合に度数を求めたり、度数などの合計を求めるときには、関数を使えばデータを処理しやすくなります。

度数分布表の作成で使用する関数

- 合計は、SUM関数を利用します。

SUM（合計を計算する）

- 書式：SUM(数値1, 数値2, ...)
- 引数：数値1, 数値2, ... ：平均を計算するセルの範囲
- 例：B1～B10セルまでのセルの数値の平均値を計算する

```
=SUM(B1:B10)
```

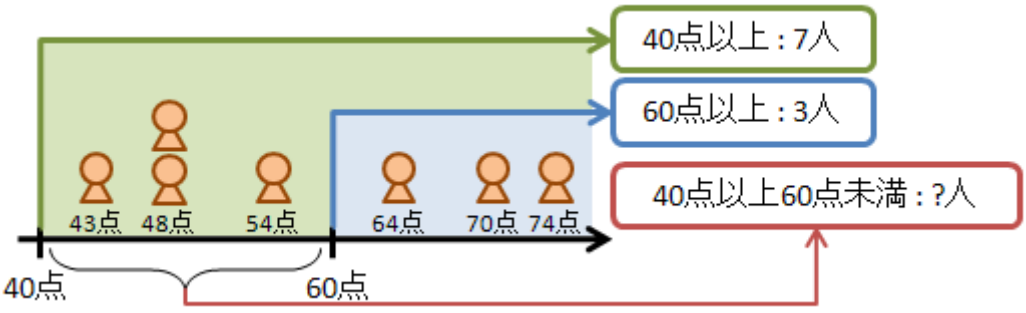
- 「値が60以上のセル」のように、特定の条件を満たすセルの個数を数えるには、COUNTIF関数を使います。

COUNTIF（指定された範囲のセルのうち、検索条件に一致するセルの個数を返す）

- 書式：COUNTIF(範囲, 検索条件)
- 引数：範囲：個数を求めるセルの範囲
- 引数：検索条件：個数を求めるセルの検索条件
- 例：W1～W10セルまでで値が「80以上」のセルの個数を数える

```
=COUNTIF(W1:W10,">=80")
```

COUNTIF関数を利用して数式をつくると、「40以上60未満の値があるセルの数」を求めることができます。ただし、COUNTIF 関数には条件は1つしか設定できないため、ひとつの数式でCOUNTIF 関数を2つ使います。



たとえば、C1～C7セルに上の図のように点数の値が入力されているとします。そのときに、「40以上60未満の値があるセルの数」を求める場合は、次のような数式になります

```
=COUNTIF(C1:C7,">=40")-COUNTIF(C1:C7,">=60")
```

つまり、「40以上の値があるセルの数」と「60以上の値があるセルの数」の差を求めればよいわけです。

Excelでヒストグラムを作成

はじめに

度数分布図（ヒストグラム）は、データの分布を視覚的（図的）にとらえることができ、ひと目でデータの状況を把握することができます。

しかし、表計算ソフトの「Microsoft Excel」には、直接ヒストグラムを作成する機能はありません。そのため、縦棒グラフを利用して、**すべての階級の階級幅が同じ**場合のヒストグラムを作成します。

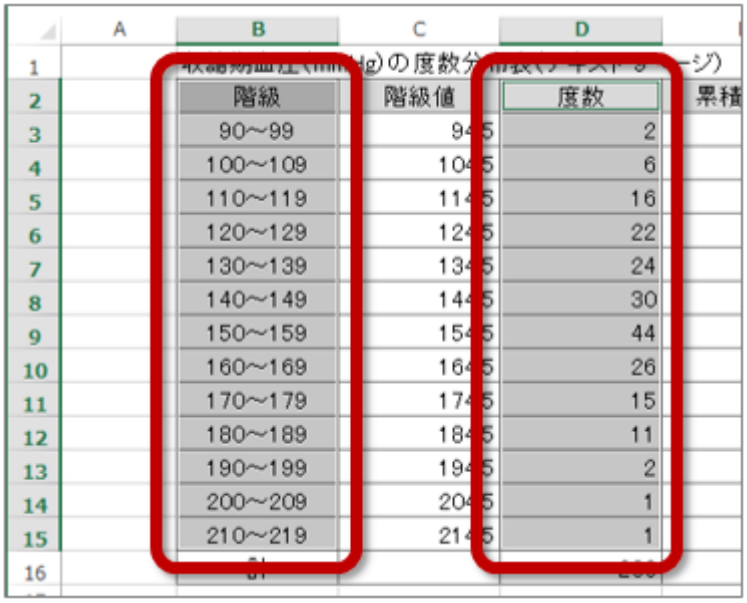
縦棒グラフの作成

Excel 2013での作成の仕方を説明します。

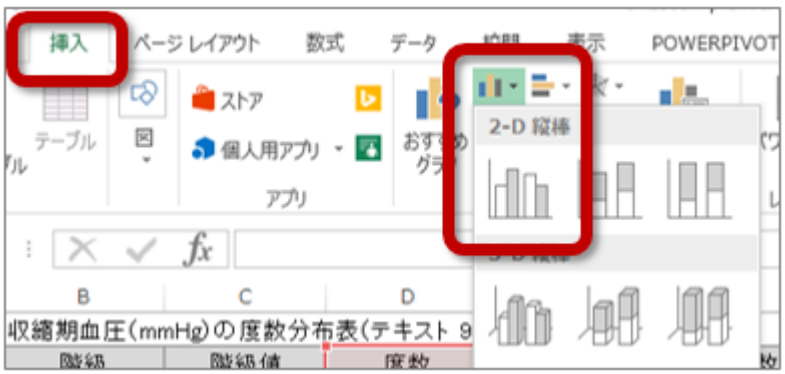
eラーニングの画面からダウンロードできるExcelのファイルを利用してみます。「度数分布図の練習(1)」というシートで作成します。

次のようにして、縦棒グラフを作成してみましょう。

1. **「Ctrl」キーを押しながらマウスをドラッグ**して、B2～B15セルとD2～D15セルを範囲選択する



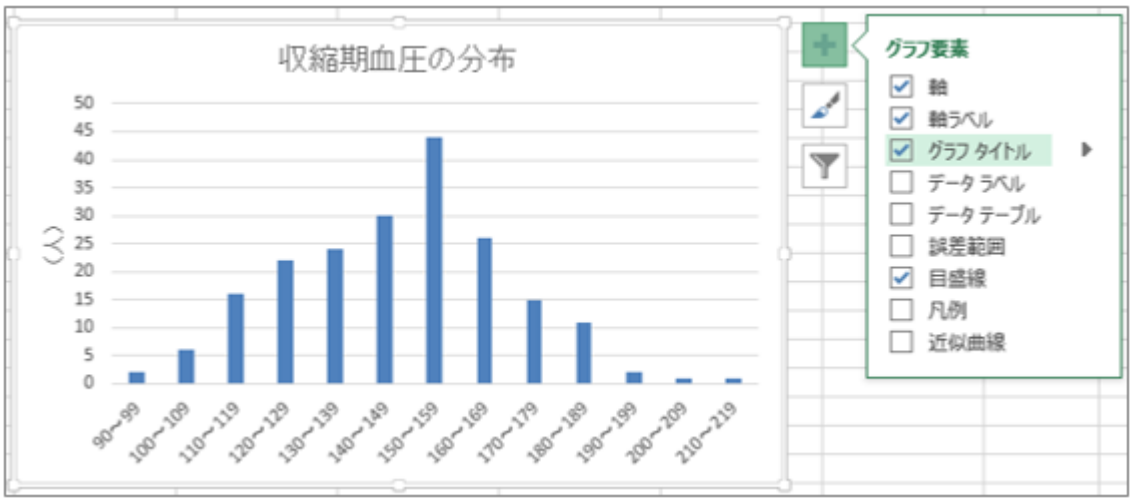
2. 「挿入」タブの「グラフ」グループにある「縦棒」ボタンをクリックする
3. メニューから「2-D縦棒」の「2-D集合縦棒」を選択する



4. グラフが作成される

作成できたら、グラフに次の設定をしてください。

- グラフのタイトルを設定する
- 縦軸の軸ラベルを「(人)」と設定する



棒グラフからヒストグラムへ

縦棒グラフのままでは、縦棒と縦棒の間に隙間があいてしまいます。

次のような設定をして、隙間をなくして、ヒストグラムの形にします。

1. グラフの縦棒部分をマウスで右クリックして、メニューから「データ系列の書式設定」を選択する

2. 「データ系列の書式設定」の書式設定が表示されたら、「系列オプション」の「要素の間隔」を「0%」に設定する

3. 設定すると、縦棒の隙間がなくなる

Excelで代表値と散布度を計算

数式の入力

Excelでは、セルに「**数式**」を入力することで、計算ができます。数式の基本的なルールは、次のとおりです。

- 最初は「=」ではじめる
- カッコ「()」を使って計算する順番を指定できる

演算	数学での記号	Excelでの記号	計算式の例	表示される結果
足し算	＋	+	=1+2	3
引き算	ー	-	=2-3	-1
掛け算	×	*	=4*5	20
割り算	÷	/	=1/2	0.5
べき乗	^	^	=2^3	8

数式の入力例

たとえば、身長と体重のデータから人の肥満度をはかる指標である、BMI（ボディマス指数）を計算する場合を考えてみましょう。

```
BMI＝体重（kg）÷身長（m）の2乗
```

身長のデータがB2～B11セルに、体重のデータがC2～C11セルに入力されており、それらから求めたBMIをD2～D11セルに表示させるには、次のように操作します。

1. D2セルに次の計算式を入力する

```
=C2/((B2/100)^2)　（「/100」としているのは、身長がcm単位のため）
```

2. 「Enter」キーを押すと、計算結果が表示される
3. D2セルの計算結果を、D3～D11セルへコピーする

平方根、n 乗根の計算

- 正の平方根（√）を計算するには、**SQRT関数**を利用します。

SQRT（平方根を計算する）

- 書式：SQRT(数値)
- 引数：平方根を求める数値
- 例：A12セルの数値の平方根を計算する

```
=SQRT(A12)
```

- n乗根を計算する関数はないため、べき乗（^）を利用（「n乗根の計算」は「1/n のべき乗の計算」と同じ意味）

- 例：A12セルの数値の4乗根 $\sqrt[4]{A12}$ を計算

```
=A12^(1/4)
```

代表値を計算

平均値

- 算術平均は、**AVERAGE**関数を利用します。

AVERAGE (平均値を計算する)

- 書式 : AVERAGE(数値1, 数値2, ...)
- 引数 : 数値1, 数値2, ... :平均を計算するセルの範囲
- 例:F1～F10セルまでのセルの数値の平均値を計算する

```
=AVERAGE(F1:F10)
```

中央値

- 中央値は、**MEDIAN**関数を利用します。

MEDIAN (引数に含まれる数値の中央値を求める)

- 書式 : MEDIAN(数値1, 数値2, ...)
- 引数 : 数値1, 数値2, ... :中央値を計算するセルの範囲
- 例:F1～F10セルまでのセルの中央値を求める

```
=MEDIAN(F1:F10)
```

四分位数

- 四分位数は、**QUARTILE**関数または**QUARTILE.INC**関数を利用します。

QUARTILE (配列に含まれるデータから四分位数を抽出する)

- 書式 : QUARTILE(配列, 戻り値)
- 引数 : 配列 :対象となるデータを含む配列 (セルの範囲)
- 引数 : 戻り値 :戻り値として返す四分位数の内容を指定
  - 戻り値: 0: 最小値
  - 戻り値: 1: 第1四分位数 (25%)
  - 戻り値: 2: 第2四分位数 (50%) =中央値
  - 戻り値: 3: 第3四分位数 (75%)
  - 戻り値: 4: 最大値

QUARTILE.INC (配列に含まれるデータから四分位数を抽出する)

- 書式 : QUARTILE.INC(配列, 戻り値)
- 引数 : 配列 :対象となるデータを含む配列 (セルの範囲)
- 引数 : 戻り値 :戻り値として返す四分位数の内容を指定
  - 戻り値: 0: 最小値
  - 戻り値: 1: 第1四分位数 (25%)
  - 戻り値: 2: 第2四分位数 (50%) =中央値
  - 戻り値: 3: 第3四分位数 (75%)
  - 戻り値: 4: 最大値

百分位数

- 百分位数は、**PERCENTILE**関数または**PERCENTILE.INC**関数を利用します。

PERCENTILE（配列に含まれるデータから百分位数（%）を抽出する）

- 書式：PERCENTILE(配列, 率)
- 引数：配列 ：対象となるデータを含む配列（セルの範囲）
- 引数：率 ：0～1の値で、目的の百分位の値（パーセンタイル値）を指定

PERCENTILE.INC（配列に含まれるデータから百分位数（%）を抽出する）

- 書式：PERCENTILE(配列, 率)
- 引数：配列 ：対象となるデータを含む配列（セルの範囲）
- 引数：率 ：0～1の値で、目的の百分位の値（パーセンタイル値）を指定

最頻値

- 最頻値は、**MODE**関数または**MODE.SNGL**関数を利用します。

MODE（引数に含まれるデータのなかで最も頻繁に出現する値を求める）

- 書式：MODE(数値1, 数値2, ...)
- 引数：数値1, 数値2, ... ：最頻値を計算するセルの範囲
- 例：F1～F10セルまでのセルの最頻値を求める

```
=MODE(F1:F10)
```

MODE.SNGL（引数に含まれるデータのなかで最も頻繁に出現する値を求める）

- 書式：MODE.SNGL(数値1, 数値2, ...)
- 引数：数値1, 数値2, ... ：最頻値を計算するセルの範囲
- 例：F1～F10セルまでのセルの最頻値を求める

```
=MODE(F1:F10)
```

散布度を計算

分散

- 分散は、**VAR.P**関数を利用します。

VAR.P（引数を母集団全体と見なし、母集団の分散（標本分散）を求める）

- 書式：VAR.P(数値1, 数値2, ...)
- 引数：数値1, 数値2, ... ：母集団に対応するセルの値、セルの範囲

標準偏差

- 標準偏差は、**STDEV.P**関数を利用します。

STDEV.P（引数を母集団全体であると見なして、母集団の標準偏差を求める）

- 書式：STDEV.P(数値1, 数値2, ...)
- 引数：数値1, 数値2, ... ：母集団に対応するセルの値、セルの範囲

不偏分散

- 不偏分散は、**VAR**関数または**VAR.S**関数を利用します。

VAR（引数を正規母集団の標本と見なし、標本に基づいて母集団の分散の推定値 (不偏分散) を求める）

- 書式 : VAR(数値1, 数値2, ...)
- 引数 : 数値1, 数値2, ... ：母集団の標本に対応するセルの値、セルの範囲

VAR.S （引数を正規母集団の標本と見なし、標本に基づいて母集団の分散の推定値 (不偏分散) を求める）

- 書式 : VAR.S(数値1, 数値2, ...)
- 引数 : 数値1, 数値2, ... ：母集団の標本に対応するセルの値、セルの範囲

不偏標準偏差

- 不偏標準偏差は、**STDEV**関数または**STDEV.S**関数を利用します。

STDEV（引数を標本と見なし、標本に基づいて母集団の標準偏差の推定値を求める）

- 書式 : STDEV(数値1, 数値2, ...)
- 引数 : 数値1, 数値2, ... ：母集団に対応するセルの値、セルの範囲

STDEV.S （引数を標本と見なし、標本に基づいて母集団の標準偏差の推定値を求める）

- 書式 : STDEV.S(数値1, 数値2, ...)
- 引数 : 数値1, 数値2, ... ：母集団に対応するセルの値、セルの範囲