

第14回

Excel分析ツールの活用(2) —実際に使ってみよう—

(一財)宮城県成人病予防協会学術・研究開発室長
小島 光洋

Excelにデータを入力するには、最初の行に、個人のID、性別、年齢、測定項目などの名称(ラベル)を記載します。すなわち、1人の個人のデータは横方向に入力することになります。縦横を逆にしても構いませんが、その後の操作がやりにくく感じます。

基本統計量では、データが入力されているセル範囲を指定すれば、平均と標準偏差、分散、最大値、最小値、中央値、最頻値などを得ることができます。

ヒストグラムでは、データを階級に分類し、それぞれの階級のデータ数を表とグラフで表示できます。階級は区切りの値をデータ区間として別に設定します。データ区間として、5と10を隣り合ったセルに入力すると、5(5以下)、10($5 < \leq 10$)、次の階級($10 <$)の3つの階級に分類されます。Excelでは、「以下」で区切るため、「以上」「未満」の階級を作成するにはデータの有効数字を考えた工夫

が必要です。グラフの作成は、作成された度数分布表を使って、データ区間を書き直してからグラフ機能を使うと、分析ツールで作るよりも見やすいグラフができます。t検定(一对の標本による平均の検定)を使うのは、教室指導の効果測定のように、同じ人に2回の測定を行い教室参加者全体の変化の有無を把握しようとする場合です。欠席者のデータは欠損値となるのでチェックする必要があります。

同じ測定を2つ以上の集団に対して行い、集団間の差の有無を調べるのが、t検定(分散が等しくないと仮定した2標本の平均の検定)と分散分析(一元配置)です。2つの集団ならばt検定、3つ以上ならば分散分析を使います。t検定には等分散を仮定する方法の設定もありますが、等分散の場合でも分散が等しくないと仮定した方法が使えます。