

宇宙地球物理学実験（気象学分野） 相関性の解析

1. はじめに

エルニーニョ現象が発生すると、日本は冷夏や暖冬になりやすいと言われることがある。本実験では、気象庁による観測データを用いて、エルニーニョ現象が日本の天候に与える影響を調べる。エルニーニョ監視海域の月平均水温と、東京の月平均気温、月降水量、月日照時間のデータが、CSV形式の電子ファイルで与えられている。この実験では、OpenOffice.orgの表計算ソフトで解析することを想定しているが、各種プログラミング言語を用いて解析してもよい。ただし、OpenOffice.org以外の表計算ソフトを使ってはいけない。

☞ 高等学校の地学でエルニーニョ現象を取り上げる。日本の天候に与える影響にも言及する。

グラフや散布図は紙に印刷し、課題4と8はレポート用紙等を書いて提出しなさい（手書きでもワープロでもよい）。適切なタイトルと学籍番号、氏名を記載した表紙をつけ、ホッチキスでとじて提出すること。サイズはA4とする。

2. 用意するもの

筆記用具、レポート用紙、ノートパソコン

3. データファイル

使用するデータファイルは以下の4個である。

1. nino3.csv エルニーニョ監視海域（北緯5度～南緯5度、西経90～150度）の水温データ
2. t662.csv 東京の気温データ
3. r662.csv 東京の降水量データ
4. s662.csv 東京の日照時間データ

単位は℃、mm、時間である。99.9は無効なデータである。

テキスト形式のデータファイルも用意されている（*.txt）。FORTRAN やCでプログラムを書いて処理する場合は、テキスト形式のデータファイルを使用してもよい。

4. 課題

課題1：エルニーニョ監視海域の月平均海面水温の時系列（全期間）をグラフに示しなさい。

課題2：エルニーニョ監視海域の月平均海面水温の基準値との差を計算しなさい。ここでは海面水温の基準値を、その年の前年までの30年間の各月の平均値と定義する。たとえば、1979年1月における基準値は、1949年から1978年の1月の海面水温の平均値である。1979年から2008年までの結果をグラフで示しなさい。

課題3：計算した基準値との差の5か月移動平均値を求めなさい。5か月移動平均とは、その月の2か月前から2か月後までの5か月間の平均である。1979年から2008年までの結果を

グラフで示しなさい。

※課題2と課題3の結果は、別々のグラフに示すのではなく、1枚のグラフに重ねて示しなさい。どの線がどのデータを指しているか明確になるように、必要に応じて線の色や種類を変え、凡例をつけること。

課題4：課題3で計算した5か月移動平均が6か月以上連続して+0.5℃以上であったとき、+0.5℃以上であった期間をエルニーニョ現象の期間と定義する。この定義に従い、エルニーニョ現象の期間を○年○月～○年○月というような形ですべて指摘しなさい。

課題5：東京の8月の月平均気温の時系列（全期間にわたる年々変動）をグラフに示しなさい。

課題6：課題3で計算した5か月移動平均（8月における値）と、東京の8月の月平均気温との間の関係を散布図に示しなさい。また、相関係数を求めなさい。相関係数は図中に適切に書き入れること（手書きでもよい）。解析期間は1979～2008年とする。

相関係数はふたつの確率変数の関係性を定量的に表す統計学的な指標である。相関係数は-1から1までの値をとる。相関係数が正であるということは、一方の確率変数が大きくなるほど、もう一方の確率変数も大きくなる傾向があることを示している。負の場合はこの逆であり、ゼロの場合は相関がないということになる。2組の数値からなるデータ列 $(x_i, y_i)(i=1, 2, \dots, n)$ におけるxとyの相関係数rは、

$$r = \frac{\sum_{i=1}^n \{(x_i - \bar{x})(y_i - \bar{y})\}}{\sqrt{\sum_{i=1}^n \{(x_i - \bar{x})^2\}} \sqrt{\sum_{i=1}^n \{(y_i - \bar{y})^2\}}}$$

で定義される。ただし、

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

である。表計算ソフトで解析する場合は、相関係数は関数CORRELを使って計算できる。

課題7：課題6と同様の解析を、月平均気温の代わりに、月降水量と月日照時間についても行ないなさい。

課題8：以上の結果に基づいて、エルニーニョ期間中の夏季の天候の傾向について考察しなさい。

課題の解答は、学籍番号と氏名の記入を確認のうえ、次回の実験の開始時まで提出してください。

(参考)

12 か月分ごとに横一列に並んでいるデータを縦に並べ替える方法

表計算ソフトでは、いくつかの関数を用いて、12 か月分ごとに横一列に並んでいるデータを縦に並べなおすことができる。

1. INDEX 関数を使う

INDEX 関数を使って、指定したセル番号のデータを参照できる。たとえば、INDEX(Sheet1. \$A\$1:\$L\$60;2;3)は、シート Sheet1 の中の 1 行 A 列から 60 行 L 列までの範囲を指定し、その中の第 2 行、第 3 列（つまり C 2 セル）の値を参照する。INDEX(Sheet1. \$A\$1:\$L\$60;10;2)なら Sheet1 の B 10 セルの値を参照する。別のシート Sheet2 の適当な列において、

1 行目で、INDEX(Sheet1. \$A\$1:\$L\$60;1;1)、

2 行目で、INDEX(Sheet1. \$A\$1:\$L\$60;1;2)、

...

12 行目で、INDEX(Sheet1. \$A\$1:\$L\$60;1;12)、

13 行目で、INDEX(Sheet1. \$A\$1:\$L\$60;2;1)、

...

のように参照していけば、12 か月ごとに横一列に並んでいるデータを縦に並べ替えることができる。しかし、この方法では、セル番号を自分でひとつずつ書かなくていけない。

2. ROW 関数を使う

上の例で、1 行目から 12 行目までに限れば、INDEX 関数の第 3 引数には、Sheet2 上で自分がいるセルの行番号が入っている。実は、行番号は ROW() という関数で参照できる。したがって、

1 行目で、INDEX(Sheet1. \$A\$1:\$L\$60;1;ROW())、

2 行目で、INDEX(Sheet1. \$A\$1:\$L\$60;1;ROW())、

...

12 行目で、INDEX(Sheet1. \$A\$1:\$L\$60;1;ROW())

と書くこともできる。

3. INT 関数を使う

INT 関数は、切り捨てて整数にした値を返す。行番号が 1 から 12 だったら 1 を、13 から 24 だったら 2 を返すようにしたい場合、INT((ROW()+11)/12) とすればよい。また、行番号が 1 から 12 の区間では 1 から 12、13 から 24 の区間でも再び 1 から 12 というように周期的な値を得たい場合には、ROW()-12*(INT((ROW()+11)/12)-1) とすればよい。したがって、すべての行で、

INDEX(Sheet1. \$A\$1:\$L\$60;INT((ROW()+11)/12);ROW()-12*(INT((ROW()+11)/12)-1))

と書けば、縦一列に並べ替えることができるようになる。