

地球物理学実験（気象学分野） 予習課題 1

プログラミングによるデータ解析

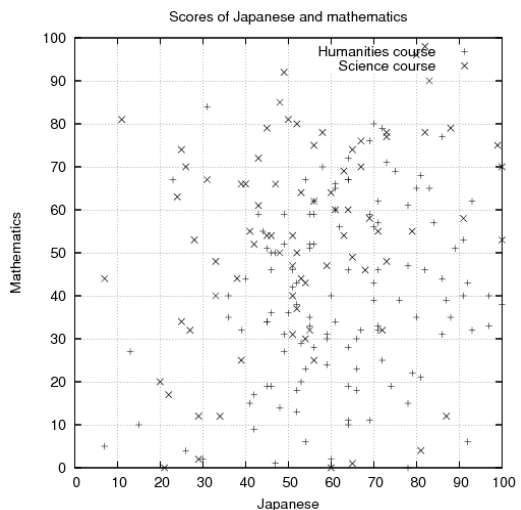
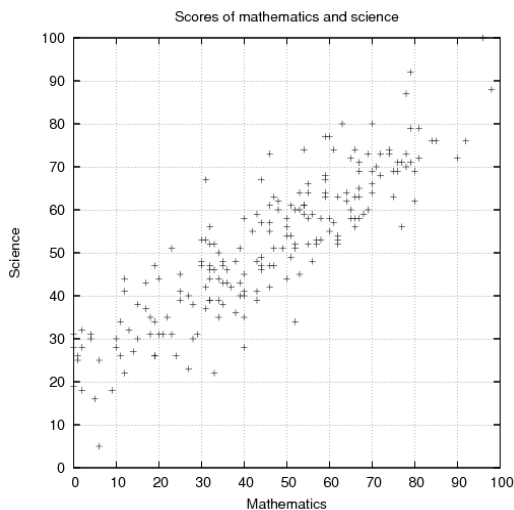
これは、地球物理学実験を行なうにあたって、プログラミングを復習しておくための課題です。

添付ファイルの中の「data.txt」は、200人の生徒の国語、社会、数学、理科、英語の試験の得点を記載した、データファイル（テキスト形式）である。本問においては、高等学校での成績処理を想定しているが、便宜上、地理歴史と公民を総称して「社会」と呼んでいる。このデータについて、以下の指示に従い、問1～4に答えよ。解答はメールの添付ファイルとして提出すること。

問1. (1) 数学の点数と理科の点数との間に相関があるか調べるために、全生徒のデータを対象にして散布図を `gnuplot` で描け。横軸を数学の点数、縦軸を理科の点数とする。凡例を含め、グラフ（散布図）の各種設定を適切に行なえ。ファイル入力は、`sample1.f`、`sample1.c` を参考にせよ。ファイル名は `fig11.ps` とする。

(2) 同様に、国語と数学の点数について、全生徒のデータを対象にして（1枚の）散布図を `gnuplot` で描け。ここでは、文系（1～120番）と理系（121～200番）の生徒を区別して作図せよ。横軸を国語の点数、縦軸を数学の点数とする。凡例を含め、グラフ（散布図）の各種設定を適切に行なえ。ファイル入力は、`sample1.f`、`sample1.c` を参考にせよ。ファイル名は `fig12.ps` とする。

ヒント：



問2. 各教科の平均点と標準偏差（不偏標準偏差ではない）を計算するプログラムを作成せよ。小数点第1位まで表示せよ。ファイル入力は、**sample1.f**、**sample1.c**を参考にせよ。実行結果ではなく、プログラムを提出せよ。ファイル名は **prog2.f** または **prog2.c** とする。

ヒント：国語の平均点は58.4点、標準偏差は19.8点である。

問3. 各生徒について、5教科の合計点、各教科の偏差値、5教科の偏差値の平均を計算するプログラムを提出せよ。いずれも1の位まで表示せよ。**hint3.txt** にならって、名簿番号、各教科の点数、5教科の合計点、各教科の偏差値、5教科の偏差値の平均の順に書き出すようにせよ（出力ファイル名は **output.txt** とする）。ファイル入力、ファイル出力は、**sample3.f**、**sample3.c**を参考にせよ。実行結果ではなく、プログラムを提出せよ。ファイル名は **prog3.f** または **prog3.c** とする。

偏差値の定義は、

$$\text{偏差値} = 50 + 10 \times (\text{点数} - \text{平均点}) / \text{標準偏差}$$

である。「5教科の偏差値の平均」は、5教科の合計点から直接偏差値を計算するのではなく、各教科の偏差値を単純に平均する点に注意せよ。この方式を科目偏差値方式といい、多くの予備校や模擬試験で使われている。

ヒント：**hint3.txt** は名簿番号1、2、200番の生徒の結果のみを例示しているが、これら3名に関しては正しい値であるから、全く同じ結果になればよい。

【以下は発展的な内容です。余裕のある人は挑戦してください。】

問4. (1) ある国立大学では、入学試験において国語、社会、数学、理科、英語の5教科を課している。この国立大学には、偏差値（平均）で見て、これら200人中の上位5人のみが合格可能であるとする。一方、ある私立大学では、入学試験において数学のみを課している。この私立大学には、偏差値（数学）で見て、これら200人中の上位5人のみが合格可能であるとする。ここで、合格可能な受験者の中の偏差値の最低値（便宜上、上位から5人目の生徒の偏差値とみなしてよい）を、その大学の入学試験の難易度と定義する。上記の国立大学と私立大学の入学試験の難易度は、それぞれ、どのような値になるか。プログラムを書いて求めよ。1の位まで表示せよ。実行結果ではなく、プログラムを提出せよ。ファイル名は **prog4.f** または **prog4.c** とする。

(2) 前問の結果において、入試難易度の値にはどのような違いがあるか。また、入学できる生徒の割合は同じであるにもかかわらず、このような違いが生じたのはなぜか、考察せよ。解答は テキストファイル に書いて提出せよ。