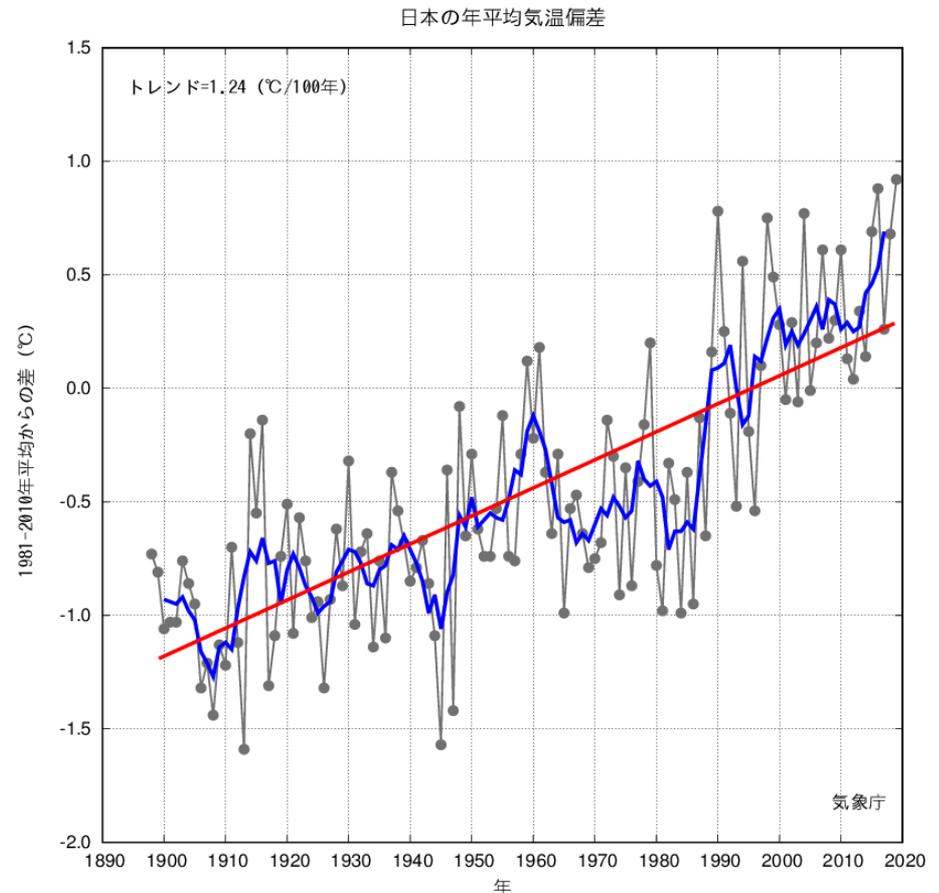


Fortran 入門

はじめに

- 今後の実験や実習では、取得したデータを様々な方法で処理することになるだろう。
- これまでに学んだ命令を組み合わせてデータ解析しよう。
 - 平均, 分散, 標準偏差, 共分散の計算
 - 最小二乗法による直線フィッティング(例えば右図の赤線)
 - 移動平均(右図の青線)



http://www.data.jma.go.jp/cpdinfo/temp/an_jpn.html
2020/01/07 にダウンロードしました。

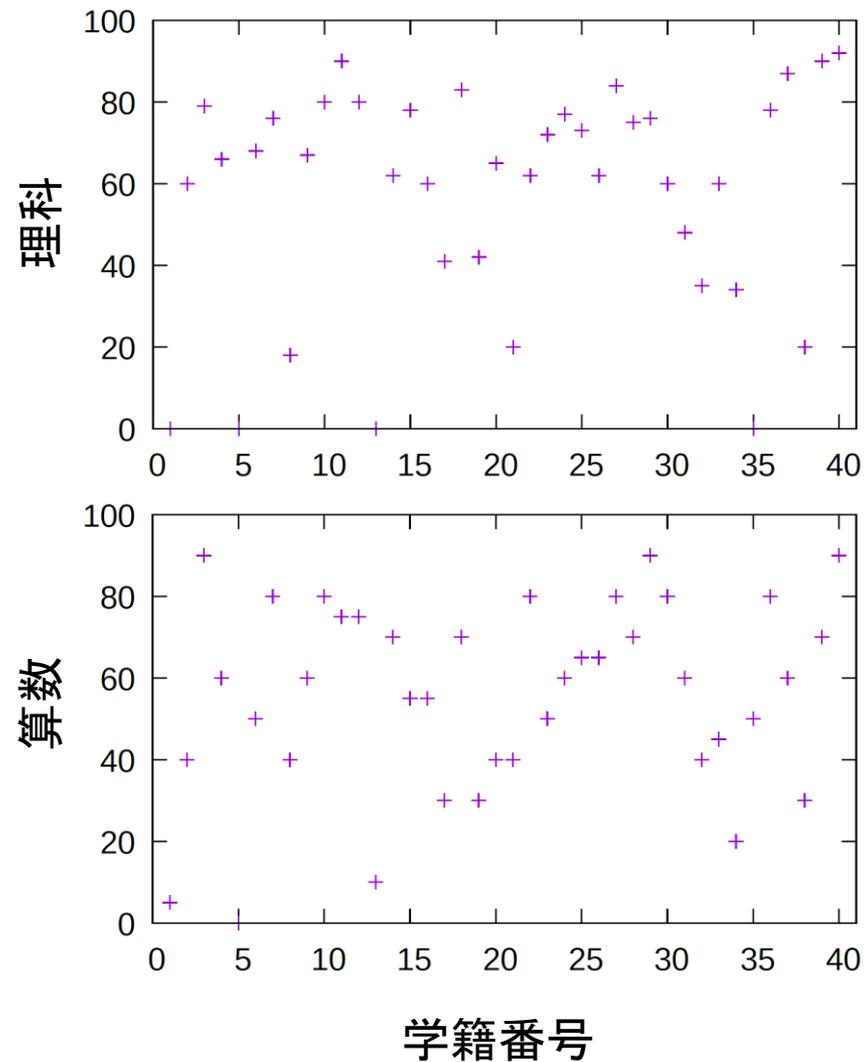
復習：統計処理

- 下のようなデータがあるとする。
 - 例：〇〇小学校〇年〇組のテストの点数

学籍番号	x : 理科	y : 算数
01	60	40
02	79	90
03	66	60
...
N	92	90

復習：統計処理

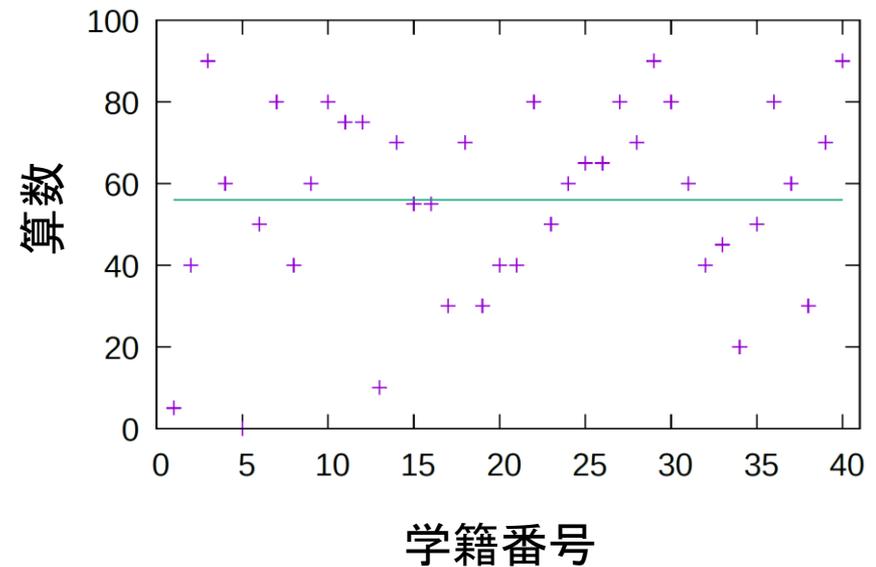
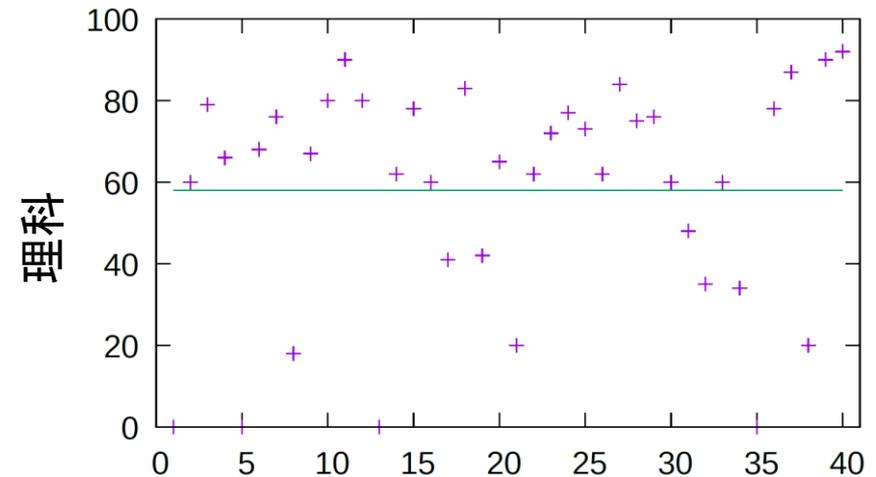
学籍番号	x : 理科	y : 算数
01	60	40
02	79	90
03	66	60
...
N	92	90



復習：平均

- 平均 (average, mean)

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$



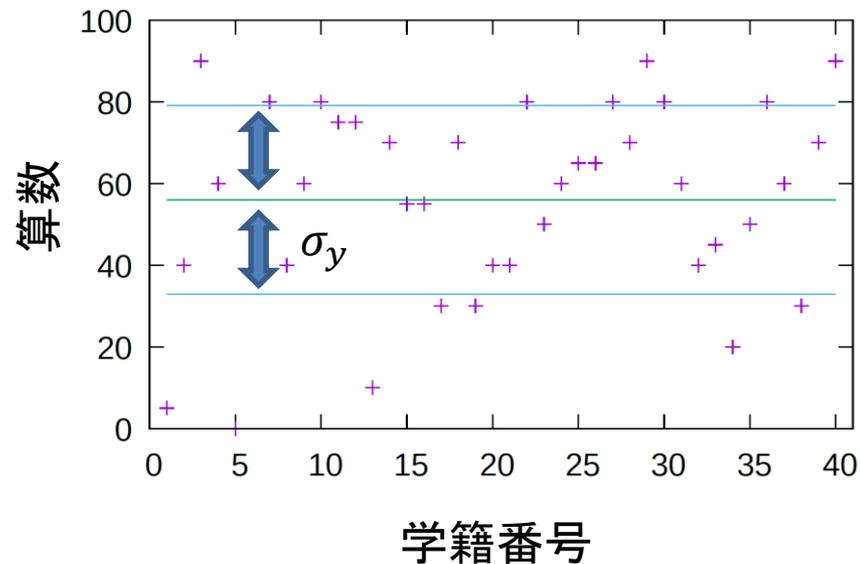
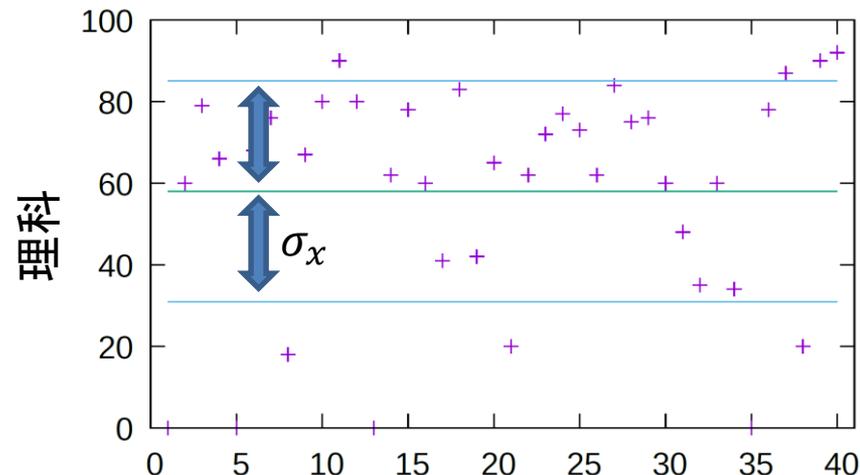
復習：分散, 標準偏差

- 分散 (variance)

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

- 標準偏差 (standard deviation)

$$\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$



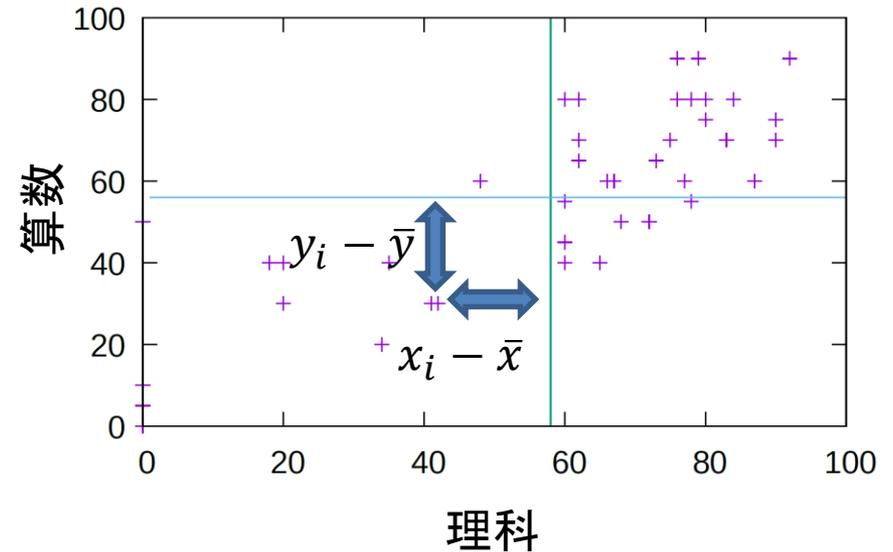
復習：共分散, 相関係数

- 共分散 (covariance)

$$C_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

- 相関係数 (correlation coefficient)

$$r = \frac{C_{xy}}{\sigma_x \sigma_y}$$



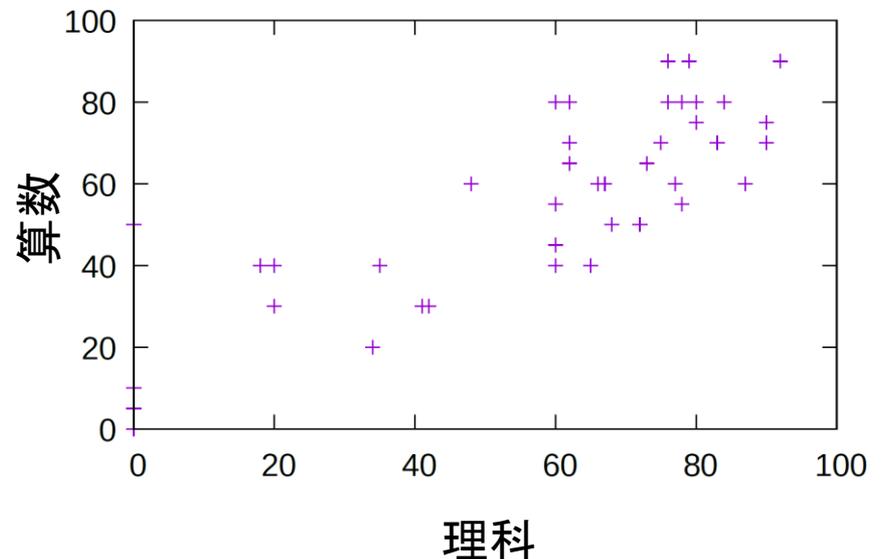
最小二乗法 概要 (1)

- ここまでに見てきた理科と算数の点数の組を考える.

$$(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$$

- このとき, 理科と算数の点数の関係を, 下の関数でフィッティングすることを考えよう.

$$y = f(x)$$



最小二乗法

概要 (2)

- 「フィッティングする」とは、言うなれば「よく合っているようにする」ことである。
- 「よく合っている」の基準の一つは、二乗誤差が小さいことである。
 - つまり、下の値が最小となれば「よく合っている」だろう。

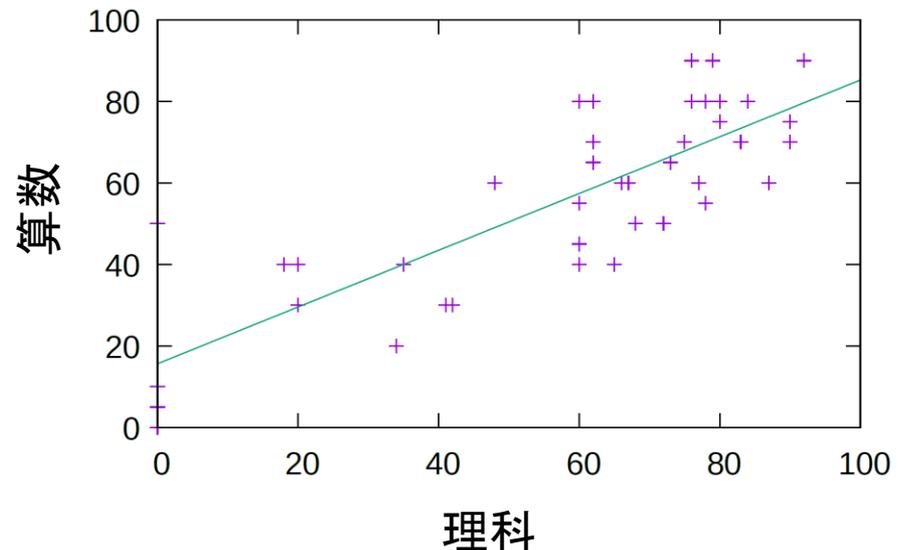
$$E = \sum_{i=1}^N (y_i - f(x_i))^2$$

- 二乗誤差が極小となるのは、その微分がゼロになるときである。この条件を基に、 $y = f(x)$ を求めればよい。

最小二乗法 一次関数 (1)

- ここから先は, 具体的な例を使って考える.
- フィッティングしたい関数が一次関数の時を考える.

$$y = ax + b$$



最小二乗法 一次関数 (2)

- 二乗誤差は下のようになる。

$$E = E(a, b) = \sum_{i=1}^N \{y_i - (ax_i + b)\}^2$$

- 一次関数の係数は, 下の条件を基に決めればよい。

$$\frac{\partial E(a, b)}{\partial a} = 0$$

$$\frac{\partial E(a, b)}{\partial b} = 0$$

最小二乗法 一次関数 (3)

- これは, 下のような, a, b の連立方程式となる.

$$\begin{aligned}\frac{\partial E(a, b)}{\partial a} &= - \sum_{i=1}^N 2x_i(y_i - ax_i - b) \\ &= \left(2 \sum_{i=1}^N x_i^2 \right) a + \left(2 \sum_{i=1}^N x_i \right) b - 2 \sum_{i=1}^N x_i y_i = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial E(a, b)}{\partial b} &= - \sum_{i=1}^N 2(y_i - ax_i - b) \\ &= \left(2 \sum_{i=1}^N x_i \right) a + 2Nb - 2 \sum_{i=1}^N y_i = 0\end{aligned}$$

最小二乗法 一次関数 (4)

- この連立一次方程式より, a, b は下のよう
に求められる.

$$a = \frac{(\sum_{i=1}^N x_i y_i)N - (\sum_{i=1}^N x_i)(\sum_{i=1}^N y_i)}{(\sum_{i=1}^N x_i^2)N - (\sum_{i=1}^N x_i)^2}$$

$$b = \frac{(\sum_{i=1}^N x_i^2)(\sum_{i=1}^N y_i) - (\sum_{i=1}^N x_i y_i)(\sum_{i=1}^N x_i)}{(\sum_{i=1}^N x_i^2)N - (\sum_{i=1}^N x_i)^2}$$

最小二乗法 一次関数 (5)

- 平均, 分散, 共分散を用いると, 下のよう整理できる.

平均

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

分散

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

$$\sigma_y^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2$$

共分散

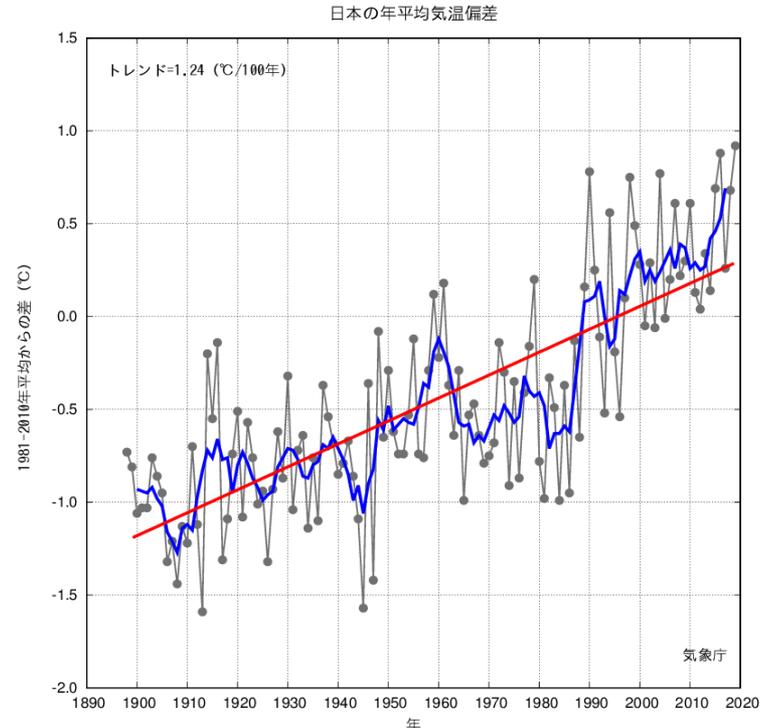
$$C_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

$$a = \frac{C_{xy}}{\sigma_x^2}$$

$$b = \bar{y} - a\bar{x}$$

移動平均

- 様々な時間・空間変動の中には, 生データのままだでは「ギザギザ」していて変動が見にくいこともあるだろう (右図黒線).
- 「ギザギザ」を除き, 長い時間・空間変動を見やすくするための方法が移動平均 (右図青線) である.



http://www.data.jma.go.jp/cpdinfo/temp/an_jpn.html
2020/01/07 にダウンロードしました.

青線は 5 年間の移動平均の結果

移動平均

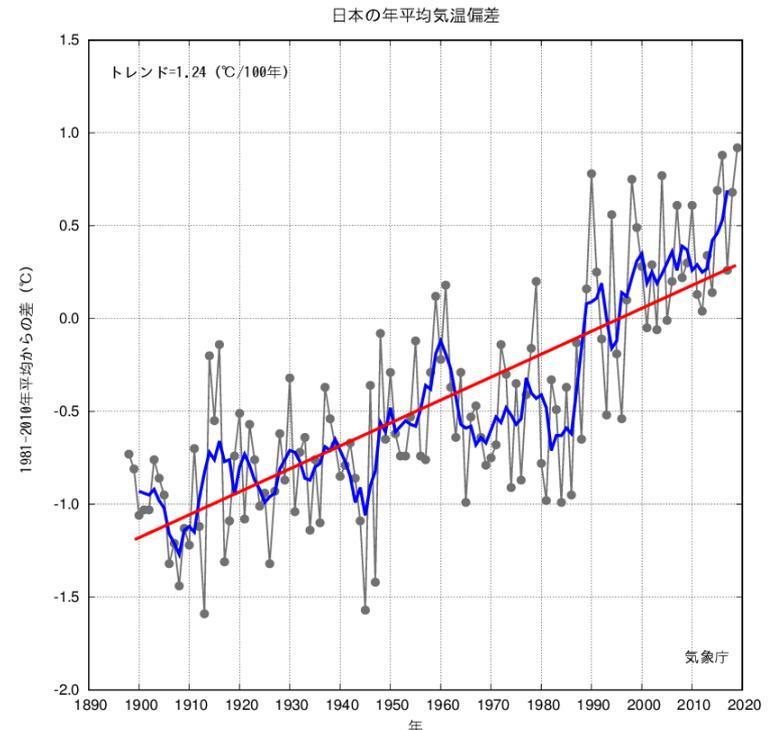
- 下のようなデータがあると
する。

$$(x_1, y_1), (x_2, y_1), \dots, (x_N, y_N)$$

- 区間ごとに平均値を求
める。

$$(\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2), \\ \dots, (\bar{x}_{N-n+1}, \bar{y}_{N-n+1})$$

$$\bar{x}_j = \frac{1}{n} \sum_{i=j}^{j+n-1} x_i \quad \bar{y}_j = \frac{1}{n} \sum_{i=j}^{j+n-1} y_i$$

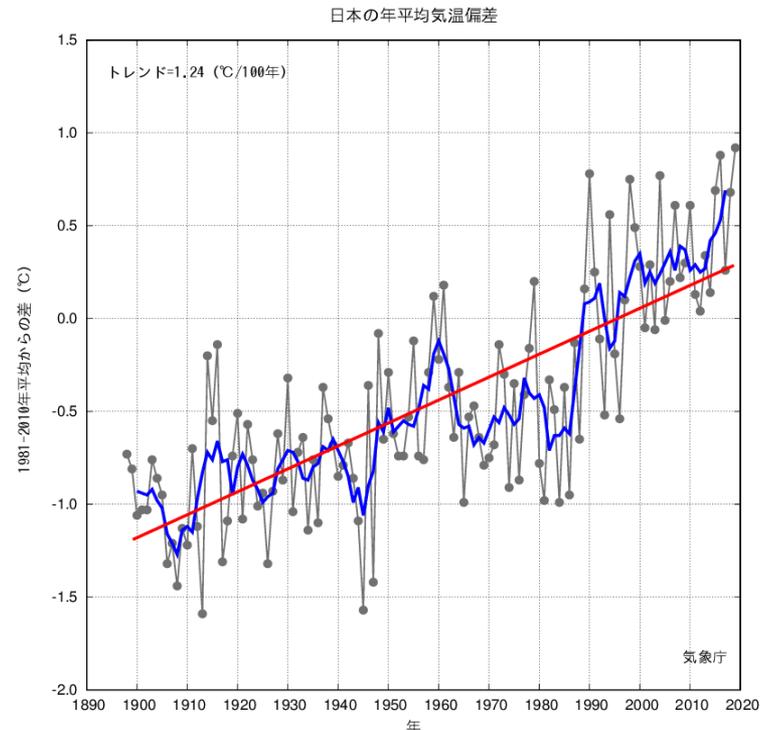


http://www.data.jma.go.jp/cpdinfo/temp/an_jpn.html
2020/01/07 にダウンロードしました。

青線は5年間の移動平均の結果

移動平均：補足

- 移動平均により, 短い時間・空間変動が取り除かれるが, どの程度短い時間・空間変動が取り除かれているのかはよくわからない.
- 取り除かれる周期を明確にするにはフーリエ変換が必要である.
- しかしながら, 移動平均は実用上便利なので, 良く使われている.



http://www.data.jma.go.jp/cpdinfo/temp/an_jpn.html
2020/01/07 にダウンロードしました.

青線は5年間の移動平均の結果

実習へ

- 実習で自分でプログラムを作って,
 - 平均, 分散, 標準偏差を計算してみよう.
 - 最小二乗法を使って一次関数を求めてみよう.
 - 移動平均を計算してみよう.