



# Excelによる相関係数分析

---



# 内容

---

1. 相関係数とは？
2. 相関係数の定義
3. 相関係数が何に役立つか？
4. 相関係数の求め方
5. 練習問題
6. 演習問題



# 相関係数とは？

---

- ある量とある量の間での線形な関係の度合いを表す指標のこと.
- $-1 \sim 1$ までの値をとる
  - 1に近ければ正の相関があるという.
  - $-1$ に近ければ負の相関があるという.
  - 0に近ければ, 相関がないという.



# 相関係数の定義

---

変数 $x, y$ のデータが $n$ 組あるとする。つまり,

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

このとき、両者の相関係数は次式で与えられる。

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

ここで、 $\bar{x}, \bar{y}$  はそれぞれの平均値を示す。



# 分散投資と相関係数

---

## ■ 分散投資

- 資産を複数の対象に投資することで、そのうちの一つが値下がりしたときに、他のものでカバーをしてリスクを回避する方法

## ■ 分散投資と相関係数

- 2つの対象の相関係数が1に近い場合、同時に騰落する可能性があるため、同時に保有しない方がよい。
- 分散投資のために、相関係数が0からマイナスになる対象に投資することがある。



# 練習問題

- 世帯数と新聞(朝刊)発行部数の関係

年	世帯数(千件)	発行部数
1990	41,156	51,908
1991	41,797	52,026
1992	42,458	51,938
1993	43,077	52,433
1994	43,666	52,601
1995	44,236	52,855
1996	44,631	53,556
1997	45,498	53,765



# 相関係数の求め方(1)

---

- 関数を用いる方法

1. 「ホーム」タブをクリック
2. 「オートSUM」→「その他の関数」を選択
3. 関数の分類「統計」→「CORREL」を選択
4. 配列1にデータ1(世帯数)を, 配列2にデータ2(発行部数)を選択
5. 「OK」をクリック



# 相関係数の求め方(2)

---

## ■ グラフから評価する方法

1. 相関係数を求めたい散布図を描く.
2. 最小自乗法で直線(単回帰式)を描く.
3. 上記の直線を右ボタンクリックし,「近似曲線の書式設定」を選択する.
4. 「オプション」タグで「グラフにR-2乗値を表示する」をチェックする.
5. 「OK」をクリックする.





# 相関係数の求め方(3-1)

---

## ■ 分析ツールを利用する設定

1. Officeボタンをクリックし、「Excelのオプション」を選択
2. 「アドイン」タブをクリック
3. 管理: の右横が「Excelアドイン」であることを確認して「設定」をクリック
4. 「分析ツール」をチェックし、「OK」をクリック  
→ これで、「分析ツールが利用可能となる。」



## 相関係数の求め方(3-2)

---

1. 「データ」タブをクリック
2. 右端の「データ分析」をクリック
3. 「相関」を選択し、「OK」をクリック
4. 入力範囲のところで相関係数を求めるデータの範囲を選択
5. データのグループが列毎か行毎かで「データ方向」を選択
6. 先頭行がラベルの場合は「先頭行をラベルとして使用」をチェック
7. 必要に応じて出力オプションを選択
8. 「OK」をクリック



# 演習問題1

- 次のデータについて相関係数を, 3つの方法で求めなさい.
- 散布図を描きなさい.
- 求めた相関係数の値とグラフの間に違和感はないか考えなさい.

No.	$x_i$	$y_i$
1	1.9	1.3
2	2.2	5.7
3	3.5	3.7
4	4.5	5.4
5	4.1	1.3
6	5.2	4.3
7	6.2	5.7
8	5.7	1.8
9	6.5	1.2
10	9.0	16.0



## 演習問題2

---

- 日本各地の月別平均降水量について、分析ツールを用いて全てのデータ間の相関係数を求めなさい。
- このグラフからわかることをレポートしなさい。