

古明地  
こいし

# 不思議な数字の世界

- 市民のための数理モデリングの基礎 -





# 古明地こいしと 不思議な数字の世界

市民のための数理モデリングの基礎

著：後藤和智（後藤和智事務所 OffLine）

表紙イラスト：杏飴（こいんとす）

発行：2015年12月30日（コミックマーケット89）

## 注意

1. 本書は、同人サークル「上海アリス幻楽団」の作品「東方 Project」の二次創作作品です。本書は東方 Project の二次創作ガイドラインに従って制作されているものであり、また著者と原作者及び作者のサークルとは一切関係がありません。そのほか、登場人物の口調などが原作と異なる場合があります。
2. 本書を著作権法の定める私的使用の範囲外で公開などを行うことを禁じます。また、本書の使用により生じた問題についての責任は負いかねます。

## 0.1 まえがき

——地霊殿にて。

**古明地さとり**（以下、さとり）：こいしの感情を育てる方法？

**霧雨魔理沙**（以下、魔理沙）：そうだ。悪い話じゃないだろう？

**さとり**：……あの子は心が読めるという能力によって傷つき、自らそれを封印して、心を閉ざしてしまっただけ。そして感情を持たない存在になったということはあなたも知っているはずよ。

**魔理沙**：そんなことを言ってる場合か？ お前だってわかってるはずだ。間欠泉の異変のあと、私や霊夢と会うことによって、またこころや董子の異変によって、あいつだって少しずつ変わりつつあることは、お前が最もよく知ってるはずだろう。

**さとり**：そうね。でも、どうやってあの子の感情を育てるというのよ。

**魔理沙**：最も大きな変化といえば、あいつにも久しぶりに友達ができたってことかな。それを利用すればいい。共通の話題や体験を持てば、感情を失ったあいつだっていろいろと学べるはずだ。

**さとり**：で、共通の話題というのは一体何を使うのかしら？

**魔理沙**：そうだな、以前に紅魔館で統計学の講座を行ったとき\*1、あいつも統計学や数学に興味を持ったよな。数理によって物事を記述するという世界があることは、あいつにとっても新鮮な話かもしれないなかっただろう。

**さとり**：数学の話か…。確かに私もあの子からも話題を振られることが多くなったわ。そして「数理によって物事を記述する」ということは、数理モデリングの話でもするのかしらね。あなたの心はそう語っているわ。

**魔理沙**：話が早いな。そう、数理モデリングだ。学術や実験からゲームに至るまで、数理モデリングの活躍する場は幅広い。数学に興味を持ち始めたあいつにも打ってつけの題材だろう。

**さとり**：そう。とりあえず、今後のことはあなたに任せるわ。

**魔理沙**：おう、任せておけ。

——数日後の地霊殿にて。

**魔理沙**：邪魔するぜ。

**古明地こいし**（以下、こいし）：魔理沙だー！ ねえねえ魔理沙あ、今日はどんな話を持ってきたの！？

**魔理沙**：今日はお前も最近お気に入りの数学の話だ。それも、あらゆる世界を記述するための数理モデリングっていう分野の話だぜ。

**こいし**：わあ、凄く面白そうだね！

**魔理沙**：しかも、今回はせっかくだからお前の友達も一緒に話を聞くことになったぜ。入っていいぞ。

**秦こころ**（以下、こころ）：誰が友達だ！ 古明地こいし、希望の面を奪った件、今度という今度は何がなんでも反省させてやるぞ！

**こいし**：あ、こころちゃん！ 久しぶりね。お元気してた？

**魔理沙**：まあまあ落ち着け。それにその件については一定の解決はしてるしお前だってそれほど深く

\*1 後藤和智『改訂増補版 紅魔館の統計学なティータイトム——市民のための統計学 Speical』後藤和智事務所 OffLine、2013年（コミックマーケット 85）

は考えなくなってるはずだろ。

**こころ**：まあ、それはそうなんだけど。

**魔理沙**：それに、いい機会なんだし二人で同じ話を聞いて、共通の話題を持ってみるっていうのもいい話じゃないか。

**こころ**：仕方ないな…。こいしも悪い奴じゃないようだし、まあ付き合ってもいいかな。

**こいし**：やったあ！ こころちゃん、今日ではよろしくね！ で、今日の話は魔理沙がしてくれるの？

**魔理沙**：いや、私じゃない。今日の話はここにある魔道書の中に詰まってる。

**こいし**：どうということ？

**魔理沙**：この魔道書を開くと、それぞれのテーマに沿った話をしてくれる奴の映像が浮かび上がるんだ。そして私はこの日のために、この魔道書の中で話をする奴を集めてきた。みんなとおきの話をしてくれると思うから、心して聞けよ。

**こいし**：ふーん。じゃあ、早速開いてみようよ！

**こころ**：相変わらずだな……少しは待つということを知らないのか。

**魔理沙**：おい、勝手に開くなよ！ ……ってまあいいか。どのみち開くことにはなるからな。

**アリス・マーガトロイド (以下、アリス)**：あー、案の定予定よりも早く開いちゃったみたいね。まあ、こいしちゃんの性格じゃ仕方ないか。こんにちは、こいしちゃん、こころちゃん。紅魔館での統計学の講義で会って以来かしら。

**パチュリー・ノーレッジ (以下、パチュリー)**：久しぶりだね、古明地こいし、秦こころ。最初の案内は私たちがするよ。ついでに言うと、この話で使う魔道書も私たちが作ったものなんだ。

**こいし**：魔法使いのお姉さんたち、久しぶりだね。

**パチュリー**：今回の趣旨は世界を記述する言葉としての数学の一つである、数理モデリングについて解説すること。数理モデリングが用いられる世界は、生物学や工学、経済学などと言った学術の世界から、マーケティングや災害避難計画などの実務の世界、あるいはゲームなどでも非常に多く用いられる。私たちが解説するのはそんな数理モデリングの世界のほんの一部に過ぎないけれど、あなたたちも数理に興味を持っているようだし、「実用的」な数学の世界の一端に触れて、新たな社会観を獲得してくれればと思う。

**こころ**：あなたたちが全部解説してくれるの？

**アリス**：いいえ。私たちが解説するのは、数理モデリングの中でも時系列解析の分野ね。これは本書の第1章で取り扱うわ。他にも、本書では最適化と、確率モデルについて取り扱うことになってるよね。結構意外性のある人選をしたから、楽しんでもらえると思うわ。

**パチュリー**：数理モデリングで取り扱うモデルには、一次関数のような極めて単純なものから、微分方程式をふんだんに使った複雑なものまでいろいろなものがある。それはモデルの抽象度によって決められるものなんだ。例えば真っ直ぐな通りを、時速4kmで歩いたとき、 $t$ 時間後の移動距離 $l$ は次のように示されるよね。

$$l = 4t \quad (1)$$

**こいし**：こういうのも立派なモデリングなんだね。

**アリス**：でも注意して欲しいのは、こういうモデリングによって捨象されるものもいっぱいあるの。常識的に考えて、ずっと一定の速度で歩けるわけがないし、また厳密に時速4kmで歩けるわけでもないわよね。でも、こういったモデリングには、ある運動をするものの将来の動きの予測などに非常に役立つのよ。

**魔理沙**：従って、数理モデリングと一言で言っても扱う範囲は非常に広い。そういったモデリングの分野を、1冊の本で網羅するというのは実質不可能だな。そのため本書で扱うモデリングはほんの一部に過ぎない。それでも、数学を用いてこんなこともできるということを知ってくれば幸いだな。

**パチュリー**：さて、おおよその説明が終わったところで、読者の皆様にも挨拶しないとね。

**魔理沙**：おう、そうだな。……というわけで、このたびは「後藤和智事務所 OffLine」通算 53 冊目、東方 Project の同人誌としては 12 冊目の同人誌を手にとってくれてありがとう。本書の基本的なスタンスはここまで解説したとおりだ。本書の講義はオムニバス形式で行われるから、基本的にどの章から読んでも問題はない。

**アリス**：また本書では、本書の版元サークルの数学・統計学同人誌ではおなじみの、フリーの統計解析ソフト「R」も使うわ。本書ではコマンドを用いた計算以外にも、プログラミングを行うこともあると思うから、今までの同人誌で説明しきれなかった R の新しい使い方を知ってくると嬉しいわね。

**こいし**：前に統計学の講義で魔理沙が教えてくれた R が出てくるんだね。しかも、プログラミングもするんだ。

**魔理沙**：R は統計解析パッケージではあるが、同時にプログラミング言語としても使えるからな。ただ一点だけ断っておくと、R のインストール方法やごく初歩の操作については、本書では解説しない。詳しくは、本書の版元サークルが出している『[普及版]R Maniax——フリーの統計ソフト「R」を使いこなす本』（後藤和智事務所 OffLine、2013 年/コミックマーケット 84）などの R の解説書か、あるいは日本の R の情報サイトの代表である「RjpWiki」\*2の解説を読んで欲しい。……それでは導入は以上だ。こいし、こころ、これからいろいろな講師が織りなす数理モデリングの話聞いてくれよ。

**こころ**：うん。楽しみにしようか。

---

\*2 <http://www.okada.jp.org/RWiki/>

# 目次

0.1	まえがき	2
<b>第 1 章</b>	<b>時系列解析</b>	<b>7</b>
1.1	はじめに——時系列データとは	7
1.2	時系列データの解析の基礎	8
1.3	時系列モデルの考え方	13
1.4	時系列解析のモデル	14
1.4.1	AR モデル	15
1.4.2	MA モデル、ARMA モデル	17
1.4.3	ARCH モデル、GARCH モデル	18
1.5	実際の時系列解析	19
<b>第 2 章</b>	<b>最適化</b>	<b>23</b>
2.1	はじめに——最適化とはなにか	23
2.2	ファイブ・ステップ法	24
2.3	感度分析	25
2.4	多変量の最適化とラグランジュ乗数	27
2.5	遺伝的アルゴリズム	29
2.5.1	遺伝的アルゴリズムの基礎	29
2.5.2	遺伝子によるデータの記述と第 1 世代の生成	30
2.5.3	適者生存	30
2.5.4	交配	31
2.5.5	突然変異	32
2.6	遺伝的アルゴリズムの利点と問題点	33
2.7	遺伝的アルゴリズムの R でのプログラミング	34
<b>第 3 章</b>	<b>確率モデル</b>	<b>41</b>
3.1	確率モデルのために——行列の入門	41
3.2	遷移確率行列	45
3.3	行列の対角化と収束	48
<b>第 4 章</b>	<b>付録</b>	<b>53</b>
4.1	第 1 回	54
4.2	第 2 回	59
4.3	第 3 回	65

---

第5章	エピローグ/あとがき	69
5.1	エピローグ . . . . .	69
5.2	参考文献 . . . . .	69
5.3	あとがき . . . . .	70



# 第1章

## 時系列解析

### 1.1 はじめに——時系列データとは

**魔理沙**：というわけで、最初の講師はアリスとパチュリーが務めることになる。二人とも、よろしくな。

**アリス**：わかったわ。パチェ、頑張りましょうね。

**パチュリー**：そうだね。頑張ろうか。というわけでここからは私たちの講義になるよ。私たちが説明するのは時系列解析。イメージで表すなら、時間を横軸に、そしてパラメータを縦軸に取ったグラフと書いていいかな。

**こいし**：そもそも時系列解析で扱うデータってどういうものを指すの？

**アリス**：まずは**時系列データ**について解説しておきましょうか。そもそも時系列データっていうのは、《対象のある側面を特定の時間間隔で観測した結果》<sup>\*1</sup>ね。例えばある時間ごとの東京の気温だったり、新宿駅南口改札の利用者数だったりするもので、身の回りに溢れているわ。

**こころ**：なるほど。そういうデータは、探してみるといろいろありそうだな。経済ニュースでよく見る為替や株式のデータも、時系列データと言えるかもしれない。

**パチュリー**：そういう時系列性を持ったデータというのは、最近流行りの「ビッグデータ」にも深く関わっていたりするものなんだ。そもそも「ビッグデータ」という言葉は、単純に膨大なデータという意味にとどまっていないの。それは、様々な種類や形式が含まれるという非定型性と、日々膨大に記録されるという時系列性という意味を持っている。この講義ではその時系列性に焦点を当てて取り扱うことになるだろうね。

**魔理沙**：総務省の『平成24年情報通信白書』では、ビッグデータの活用例として本田技研工業の例が挙げられているな<sup>\*2</sup>。これによると、ホンダは2003年から、車に付けられた装置から5分間隔で送信される情報を元に、渋滞情報などの情報がユーザーに提供されるというシステムを運用しているようだ。これと、主要幹線のデータを統合することにより、二酸化炭素の排出量の削減や、急ブレーキの減少などの効果が現れているらしい。

**こいし**：「ビッグデータ」ってよく聞くようになったけど、そのキーとなるのが時系列解析、ってことでのいの？

**アリス**：そうね。ただその前に、時系列データのリテラシーについて学びましょう。まず、時系列

<sup>\*1</sup> 横内大介、青木義充「現場ですぐ使える時系列データ分析——データサイエンティストのための基礎知識」（技術評論社、2014年）p.10

<sup>\*2</sup> <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h24/html/nc121430.html>

データというのは、一定の間隔ごとに観測されたデータのことを指すということよ。これと間違いやすいのが、株取引の世界などに見られるような**ティックデータ**ね。ティックデータというのは、取引が成立したときの取引件数や株価を指すデータなんだけど、なぜ時系列データと言えないのかというと、こういうデータはある時間に依存しているということね。

**こいし**：どうということ？

**パチュリー**：取引が成立する間隔は一定とは限らないよね。それに、特定の取引が成立するときの価格や件数というのは、その時々々の株価によって異なるじゃない。そういうデータは時系列データとは言えないんだ。

**こころ**：つまり、時系列データを考えるときは、一定の間隔で観測されているか、ということが重要なんだね。

## 1.2 時系列データの解析の基礎

**こころ**：ところで、ここは数理モデリングの解説なんでしょ？ 時系列解析のモデリングについて説明してくれるの？

**アリス**：時系列解析のモデリングを解説する前に、まずは時系列データの解析の基礎について学びましょう。ここでは、TOPIX(東証株価指数)の2014年7月から12月までの日ごとの終値を例として使ってみるわね。Rを使ってやっていきましょう。まずは、TOPIXのデータを、「TOPIX\_201507-12.csv」というcsvファイルに格納して、それを読み込みましょうか。

```
> dataset01_01 <- read.csv("TOPIX_201507-12.csv",header=T)
> dataset01_01
  day   rate
1   1 1265.32
2   2 1283.01
3   3 1284.70
4   4 1289.17
(略)
122 122 1419.89
123 123 1433.80
124 124 1423.98
```

**パチュリー**：でも、このままだと、このデータがどういう特徴を持っているかわかりづらいよね。二人に質問だけど、まずこのデータをわかりやすくするためにはどうすればいいと思う？

**こいし、こころ**：はい！

**パチュリー**：さすがにこれは簡単だったかな。ほぼ同時に手が上がったね。

**こいし**：こいしが先だよ！

**こころ**：いや、私だ！

**アリス**：こらこら。ここは仲良くやらないとせっかくさとりさんが用意してくれた機会も台無しでしょ。とりあえずこいしちゃん、どうすればいいかわかるわよね。

**こいし**：グラフを描けばいいんでしょ。折れ線グラフ。

**パチュリー**：そう。まず時系列データ解析の初歩として、折れ線グラフを描いてデータの特徴を調べ

るのが第一だね。R では、折れ線グラフは `plot` コマンドをちょっと工夫して描く。こんな感じに。

```
> plot(dataset01_01,type="l")
```

**パチュリー**：`plot` コマンドは基本的に  $x$  軸と  $y$  軸のデータがあればそれを元に散布図を描くコマンドなのだけど、中に `type="l"` というものをつけることにより折れ線グラフにすることができるんだ。

**こいし**：これで TOPIX の動きがわかりやすくなったね。最初は緩やかに上がってるけど、いったんがくっと下がって、その後はぐいっと上がってる感じかな。

**アリス**：このグラフを元に時系列解析の基礎の解説に入っていくわよ。まず、時刻  $t$  における時系列データが  $P_t$  で表されるとしましょう。時系列データで重要になるのは、以前の観測時間との「差」ね。もっともわかりやすいのは、直前の時間に観測されたデータとの差かしら。これは次のように表されるわね。

$$d_t = P_t - P_{t-1} \quad (1.1)$$

**パチュリー**：でも株価などの場合は、重要になってくるのは差ではなく変化率のほうかもしれないね。さて、これはちょっと難しいかもしれない質問だけど、差を使って変化率を表すにはどうすればいいと思う？ こいし、答えてみて。

**こいし**：うーん……わかんないなあ。なんだろう？

**こころ**：ははは！ 私にはわかったぞ！ 対数を使えばいいんでしょ？

**パチュリー**：こころも答えを先取りしすぎだよ。それはさておき、対数を用いるのが正しいね。というのも、対数の性質として、次のようなものがあるからね。

$$\log \alpha + \log \beta = \log \alpha \beta \quad (1.2)$$

$$\log \alpha - \log \beta = \log \frac{\alpha}{\beta} \quad (1.3)$$

**パチュリー**：この性質を用いれば、次のような値を導出できるはず。この値のことを、対数差収益率という\*3。

$$r_t = \log P_t - \log P_{t-1} = \log \frac{P_t}{P_{t-1}} \quad (1.4)$$

**こいし**：あ、 $\log$  中の分数は変化率になってる。

**こころ**：そう。対数を使えば、引き算を分数にすることができるわけ。

**アリス**：ここから、TOPIX の対数差収益率を見てください……って、TOPIX の対数差収益率って表現、ちょっとおかしいわよね。まあ、ここは時系列解析の練習のためだと思って、大目に見てね。R では、隣同士の階差は `diff` コマンドで求めることができるのよ。TOPIX の値は `dataset01_01` の

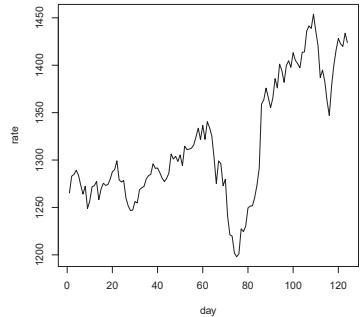


図 1.1 2014 年 7 月～12 月の TOPIX のプロット

\*3 横内、青木、前掲 p.43