

# ロボットの確率・統計

—製作・競技・知能研究で役立つ考え方と計算法—

上田 隆一 著

コロナ社

# まえがき

## 企画の意図

本書は、ロボットを作って動かすときに必要な確率・統計の知識をまとめようということで企画されました。確率ロボティクス<sup>1)~3)†</sup>や機械学習<sup>4)</sup>の書籍に登場する各アルゴリズムを理解するためには、確率・統計の基礎知識が助けになることは当然ですから、勉強したい人も多いものと考えています。しかし、ロボティクスには広大な範囲の知識が必要とされるため、後手に回ってしまった人も多いのではないのでしょうか。

学部、学科によっては、そうならないように、大学に入るとすぐに統計学の講義が用意されていることもあります。しかし、このような講義で学ぶ統計学は、確率ロボティクスや機械学習で使われるものと流儀や範囲が思いのほか異なっています。また、講義で使われる教科書は、定義などをしっかり正確に説明する役割を背負っているので、読者に強い動機を与えるものではありません。

そこで、もっとロボットや機械の話に寄せることで確率・統計を学ぶ動機を与える本、いわば教科書を読むための入門書を作ろうということになりました。ロボットを組み立てる、制御やセンサー処理をプログラミングする、性能を評価する、ロボコンで作戦を立てるなど、ロボットだけに話を限定しても、さまざまな統計の知識、確率の計算が必要になります。そこで、話題をロボットに限定して、「ロボット好き」を自認する好き者を誘い込んで、確率・統計の話に一通り触れてもらおうと考えました。

ロボットの世界には、数学が苦手でも天才的な仕事をする人がいます。筆者は、小難しい統計用語満載の怪しいプレゼン資料を作る人よりも、手が動く人

---

† 肩付数字は巻末の引用・参考文献の番号を表す。

たちのほうが好きなのですが、そういう職人氣質の人たちにも、とりあえずは一度読んでみてもらえると幸いです。

## 本書の内容

ということで、なんでもロボットに話をからめていった結果、本書はつぎに挙げる9章で構成されることになりました。

- 1章：代表値——統計の知識がないと喧嘩になる** 小学生の素朴なロボット競技での言い争いから、統計の使い方を考える。
- 2章：確率——機械が動くという奇跡** すぐリタイアするけど無反省なロボコンチームを題材に、確率の計算方法や使い方を考える。
- 3章：期待値——運を神頼みにしない** 筆者のRoboCup（ロボットサッカーの世界大会）でのインチキを題材に、期待値の求め方や使い方を確認する。
- 4章：連続値と多変量——ロボットは空間を動く** ロボットの動きの再現性に関する実験を題材に、多次元、連続な空間での確率の扱い方を理解する。
- 5章：試行回数と信頼性——実験で教員に叱られないために** ロボットを研究している学生の「何回実験すればいいのか」という疑問から、ベイズの定理の導出方法と使い方を確認する。
- 6章：動く確率分布——ロボットは時空を進む** ロボットの動きの予測を例題に、動く確率分布の計算原理と近似計算の方法を理解する。
- 7章：センシングと推定——ロボットは実世界を観察する** 人間の連想ゲームをベイズの定理から解釈し、ロボットのセンサー情報処理に利用する方法を考える。
- 8章：機械学習——ロボットはぼんやり実世界を観察する** 人間や動物の認知機能に対する考察から、学習の原理およびそこでの確率・統計の役割を把握する。
- 9章：意思決定と制御——ロボットは自律する** 確率的に不確かな将来に対して人間やロボットがどう行動を決めるか、なんのたけに行動するのかについて、ベルマン方程式から考察する。

<b>1章 代表値</b> <ul style="list-style-type: none"> <li>• 要約統計量</li> <li>• 統計的探求プロセス</li> <li>• 統計的リテラシー</li> </ul>	<b>2章 確率</b> <ul style="list-style-type: none"> <li>• 乗法/加法定理</li> <li>• 確率変数</li> <li>• 確率分布</li> </ul>	<b>3章 期待値</b> <ul style="list-style-type: none"> <li>• 意思決定への利用</li> <li>• 計算への利用</li> <li>• シミュレーションによる計算</li> </ul>
<b>4章 連続値と多変量</b> <ul style="list-style-type: none"> <li>• 連続型確率変数</li> <li>• モンテカルロ法と密度</li> <li>• 確率密度関数</li> </ul>	<b>5章 試行回数と信頼性</b> <ul style="list-style-type: none"> <li>• 確率分布の分布</li> <li>• ベイズの定理</li> <li>• 共役性</li> </ul>	<b>6章 動く確率分布</b> <ul style="list-style-type: none"> <li>• 状態遷移</li> <li>• 線形/非線形性</li> <li>• 確率分布の近似計算</li> </ul>
<b>7章 センシングと推定</b> <ul style="list-style-type: none"> <li>• ベイズフィルタ</li> <li>• カルマンフィルタ</li> <li>• パーティクルフィルタ</li> </ul>	<b>8章 機械学習</b> <ul style="list-style-type: none"> <li>• 損失関数と勾配</li> <li>• 教師なし(ベイズ)学習</li> <li>• 教師あり学習</li> </ul>	<b>9章 意思決定と制御</b> <ul style="list-style-type: none"> <li>• ベルマン方程式</li> <li>• 最適化/最適制御</li> <li>• 強化学習</li> </ul>

図 1 各章のキーワード

図 1 に、各章でキーワードになる語句を書き出してみました。大雑把にいうと、前半の 3 章くらいまでが微積分の出でこない基礎の基礎です。そして、後半の 7 章以降が「知能」を考える応用的な内容になり、ロボットだけではなく、人間はどうしているだろうかという考察も交えて話を進めます。それらの間の章では、具体例を用いて、基礎から応用へ、1 つずつ知識をつけていきます。特に(順番と関係なく)結びつきの強い章を挙げておくと、4, 6, 7, 9 章は確率ロボティクス、5, 8 章は機械学習の講義で扱われる内容です。また、1, 3, 9 章は、統計を自身の行動や世の中の理解に活かすことを考える章になっています。1 章で「統計を使って他人のことを評論するのは悪趣味」というような記述があり、9 章で評論してないで自分はどうするのか決めると迫ります。9 章は確率・統計では通常扱われない内容ですが、筆者が読んださまざまな教科書は客観的すぎて当事者意識が希薄なのが気になっていましたので、頭でっかちになることを防ぐために必須と考えています。ただ、バランス的にあまりページを割くわけにはいかなかったので、ちょっと急ぎ足でわかりにくいかもしれません。8 章も難しいわりには駆け足です。これらの章については、「わからないぞ！」と抗議の声を上げていただければ、また別の機会にまとめます。

人間の話をするときには筆者の個人的な見解が含まれてしまうので、そういう

意味で、この本は「教科書」とはいえませんが、本文中で突然世間話をするこゝともありますし、脚注で愚痴をたくさん書いています。ただ、人間の行動を眺めると、興味深いことがいろいろ見えてくるので、読者の興味を掻き立てられるのではないかと考えました。数式を追うときに世間話で気が散ったり、脚注で「嘘じゃないか?」と思ったりすることがあるかもしれません。そうならないように素直に書くこともできたのですが、それで読者が満足して終わらないよう、少し「毒」を盛ってみました。毒による疑問や違和感が、論文や、具体的な研究事例を紹介している書籍 5) を手にとる動機に繋がれば幸いです。

### 本書の性質と対象となる読者、推奨する読み方

原稿を書いているときは

- 工学系や情報系の学部に入學した大学1年生
- ロボット用や機械学習用のソフトウェアをカスタマイズしようと思っても原理がわからなくて踏み込めない人
- 自分でソフトウェアを触っているわけではないが、確率・統計の考え方や、自ら考えて行動するロボット（自律ロボット）の動作原理を理解したい人

を頭に思い浮かべて、その人たちがどう読んでいくかを考えました。また、すぐに何かに使える内容を志向するのではなく、思考の土台を整えることを心がけて執筆しました。

「土台」の一部をなす数式については、数学が苦手、あるいは学習してこなかった人に配慮して極力避けることも最初は考えたのですが、結局、容赦なく使いました。これは数学が苦手な人はお断りということではなく、むしろ逆に、**数式アレルギーの人の症状を和らげよう**くらいに考えています。世の中の多くの人が、理不尽に学校で問題を解かされて数式アレルギーになっているわけですが、よくよく考えてみたら、受験したり、自分でいきなりコードを書いたりということでもなければ、教科書の数式は解かなくても暗記しなくてもよく、読めるだけでいいわけです。ということで、各数式に対し、（限界はありますが）

その数式の読み方や意味をしっかりと言葉で補足しました。

また、筆者のような人間がすらすらと数式のある教科書や論文を読めていると思うなら、それは誤解です。しつこく時間をかけて読んでいるだけです<sup>†1</sup>。例えば、筆者は機械学習の教科書 4) をもう十年くらい、何かあったら開いていますが、まだたぶん3分の2くらいしか理解していません。「数式のわからない文系の人にも～」のような本は半日で読めてしまうこともあります。それと同じと考えてはいけません。数式というのは、言葉で説明しようとする膨大な字数がかかることを短縮して書くためのものなので、数式のある本というのは、見かけの文字数より断然内容が濃く、得られる知見も大きいわけですね<sup>†2</sup>。読破できなくても、ちょっとでも自身の理解が進めば、数千円の価値はあったと考えたほうがよいでしょう。

ですので、本書もおそらく、何年もかけて読むたぐいの本だと考えています。ただ、1ページ目から順に何年もかけて読むのではなく、わからないところはすっ飛ばして、何度も通読する方法をおすすめします。後ろのほうは、最初はほぼ全部すっ飛ばすことになります。おそらく多くの人にとってまだ習っていない数学も出てくるので、「こういうふうに確率論や線形代数、微積分を使うんだ」と理解するだけで、今後の数学の勉強に良い影響があると思います。勉強したら本書に戻って、またわからないところを確認して、と繰り返すうちに理解できるかもしれませんし、そのように時間をかけて理解していった経験は、その後の生き方にも良い影響を与えることでしょう。

だんだん世の中が理屈っぽくなっていく中で、数式という言葉が使えることの重要度は増しています。いままで数式から逃げ回ってきた人には、ここでお縄になっていただければ幸いです。また、若干後ろ向きですが、本書の中には、「数字や理屈を振りかざして『論破！』と叫びながら迫ってくる人たちをどうやってのらりくらりとかわすか」の指南もありますので、そこまで（そこだけ？）

---

<sup>†1</sup> 本書をちょっとだけ読んで、「難しい」と Amazon のレビューに書くのはやめましょう。

<sup>†2</sup> ですから、本書をちょっとだけ読んで、「高い」と Amazon のレビューに書くのはやめましょう。（しつこい）

は頑張って読んでいただきたいです。

過去に筆者の著書を読んだことがある方は、「今回はコードの掲載がない」と思うかもしれませんが。これまでの筆者の著作は、プログラミングしながら手で理解しましょうというものばかりだったのですが、本書については布団の中や、学校、職場への行き帰りの電車で思考にふけりながら読んでいただきたいと思い、コードの掲載はやめました。一方、この手の本は自分で数式を書いたり変形したりしながら読むと理解が速くなります。手を動かせる方は、キーボードを紙と鉛筆に持ち替えて取り組んでいただければ幸いです。

## 謝 辞

本書は、2022年の5月頃に筆者が当時のTwitterで「書きたい」とつぶやいたところを、コロナ社の編集の方に拾っていただいて実現したものです。いつもの筆者は短期決戦型なのですが、今回は「論文を書くから」と4回くらい執筆を中断して、脱稿まで1年以上かかってしまいました。たいへんご迷惑をおかけしました。

また、公式や数式の展開、その他記載内容に間違いがないかどうかの確認には、教科書のほか、ウェブサイトも参考にしました。本文中で挙げたサイトのほかには、特に「高校数学の美しい物語」(<https://manabitimes.jp/math>)とWikipedia (<https://ja.wikipedia.org/>)をよく閲覧しました。運営されている皆様に感謝申し上げます。もちろん(学生さんも本書を読むので明記しておきますが)、軽い確認以上に依存してしまった際は、少し間を置いて筆者自身の解釈を経た上で、自分の言葉で改めて記述しています。Wikipediaの利用については、記載内容を信頼しつつ、一次資料(Wikipediaでいうところの「信頼できる公刊された情報源」)まで目を通して、一次資料のほうを引用しています。

査読については、たいへんお忙しいところを東北大学の田村雄介先生、軍司健太さん、奈良貴明さん、千葉工業大学の藤井浩光先生に快諾いただき、さまざまな指摘をいただきました。軍司さん、奈良さんは、東北大学の犬野和則先生からご紹介いただきました。なお、筆者の個人的な主張やほやき、与太話に

については、査読いただいた皆様いずれも、見て見ぬ振りをされていたものと思われまます。それらも含め、本書の記述内容の責任の所在は筆者にあることをご理解ください。

また、毎度のことながら、千葉工業大学先進工学部未来ロボティクス学科の教員の皆様、学生の皆様、そして筆者が所属する研究室の皆様には、本書執筆中の筆者の反応の悪さに辛抱強くお付き合いいただき、誠にありがとうございました。家族については、全員勉強か趣味に没頭していて、筆者からは（悪態をつくわけではなく、本書の9章を踏まえて）好きにやってくれ俺は知らんという感じで、感謝したりされたりする気にはならないのですが、そのままいいと思います。

#### カラー画像

以下よりカラー図面がダウンロード可能です。

<https://www.coronasha.co.jp/static/04687/color.pdf>



- ※ 表紙のイラストは、パブリックドメインの写真 ([https://en.wikipedia.org/wiki/Marrus\\_orthocanna#/media/File:Marrus\\_orthocanna\\_crop.jpg](https://en.wikipedia.org/wiki/Marrus_orthocanna#/media/File:Marrus_orthocanna_crop.jpg)) を筆者が加工して作ったものです。ヒノオビクラゲという、複数の個体が集まって群体を作るクダクラゲの一種です。群体を作るクラゲについては、9章で少し話題にしました。「なんでロボットの本なのにクラゲなんだよ」という質問につきましては、「筆者がよく思考の出発点にするから」と答えることとなりますが、そんなに深い話でもありません。

2024年1月

上田 隆一



# 目 次

## 1. 代表値——統計の知識がないと喧嘩になる

1.1 代表値で比較する	2
1.1.1 試行と事象	2
1.1.2 代表値の選択と計算	3
1.1.3 判断材料にすることで代表値は意味を持つ	6
1.2 他の基本的な代表値とロボティクスでの利用	7
1.2.1 中央値	7
1.2.2 最頻値, モード	10
1.3 データのばらつきを表す指標	12
1.3.1 分散	12
1.3.2 標準偏差	15
1.4 データの偏りと外れ値	16
1.4.1 データの偏り	16
1.4.2 外れ値	18
1.5 まとめ	19

## 2. 確率——機械が動くという奇跡

2.1 動く可能性が掛け算で減っていくロボット	23
2.1.1 確率の導入	23
2.1.2 ロボットの起動率の計算	25

2.1.3	乗法定理と独立	27
2.2	冗長化されたシステム	30
2.2.1	加法定理	31
2.2.2	冗長化された $A_1$ の起動率	32
2.3	部品が互いに影響を与える場合の計算	35
2.3.1	乗法定理による計算	36
2.3.2	隠れた条件と加法定理	39
2.4	確率変数, 確率質量関数と確率分布	42
2.4.1	確率変数	42
2.4.2	確率質量関数	43
2.4.3	確率分布	44
2.4.4	確率分布と事象の関係	45
2.4.5	同時確率質量関数, 同時確率分布	46
2.5	補遺 (「確率的リテラシー」のようなもの)	47
2.5.1	なぜ自動車や飛行機が大丈夫なのかを不真面目に考える	47
2.5.2	信頼性の改善方法に関する考察	49
2.5.3	冗長化に関する考察	51
2.6	ま と め	52

### 3. 期待値——運を神頼みにしない

3.1	期待値の計算で最大の成果を挙げる	54
3.1.1	期待値の考え方	55
3.1.2	期待値の線形性	57
3.1.3	期待値の他の性質	60
3.2	シミュレーションによる期待値の算出	61
3.2.1	シミュレーションを繰り返して平均値をとる	61

3.2.2	シミュレーションと確率	63
3.3	期待値と確率分布の平均、分散	64
3.3.1	確率分布の平均値	64
3.3.2	確率分布の分散	65
3.3.3	2つの確率変数の分散と共分散	66
3.3.4	不偏分散と確率分布の分散の関係	67
3.4	ま と め	69

#### 4. 連続値と多変量——ロボットは空間を動く

4.1	連続型確率変数の離散化	73
4.1.1	連続型確率変数	73
4.1.2	離散化による確率質量関数の導出	74
4.1.3	連続型の多変量確率変数の離散化	76
4.1.4	適切な解像度	77
4.2	モンテカルロ法と密度	78
4.2.1	囲って数える	79
4.2.2	密度の導入	81
4.3	ガウス分布	84
4.3.1	確率分布へ式を当てはめる	84
4.3.2	多変量ガウス分布	95
4.3.3	多変量ガウス分布の演算	102
4.4	ま と め	106

#### 5. 試行回数と信頼性——実験で教員に叱られないために

5.1	確率の確率分布を考える	109
5.1.1	実験から求めた結果の不確かさ	109

5.1.2	確率の確率分布の計算	111
5.1.3	一般的な事前, 事後分布の計算 (ベイズの定理)	112
5.2	離散化による事後分布の近似計算と比較	117
5.2.1	試行結果を反映した事後分布の計算	117
5.2.2	改良前後のソフトウェアの (確率的な) 完走率の比較	118
5.2.3	試行回数を増やす効果と限界	120
5.3	共役事前分布	122
5.3.1	ベルヌーイ試行とベータ分布	122
5.3.2	共 役 性	124
5.4	ま と め	124

## 6. 動く確率分布——ロボットは時空を進む

6.1	動く物体と確率	127
6.1.1	時間と座標系	127
6.1.2	ロボットの意思と, 意思どおりにならない現実	128
6.1.3	ロボットの動きの確率的な表現	130
6.1.4	ある時刻の位置の分布	131
6.2	「線形」なロボットの位置の予測	132
6.2.1	問題の定式化	132
6.2.2	解 き 方	134
6.3	物体が「非線形」に移動する場合の位置の予測	134
6.3.1	ロボット座標系の準備と移動量の座標変換	135
6.3.2	座標変換の線形化	137
6.3.3	移動後の分布の算出	141
6.4	モンテカルロ法による予測	143
6.4.1	モンテカルロ法による予測の例	143

6.4.2 数式での表現 ..... 144

6.5 ヒストグラム状の離散化を用いた予測 ..... 147

6.6 ま と め ..... 148

## 7. センシングと推定——ロボットは実世界を観察する

7.1 情報のフィルタとしてのベイズの定理 ..... 150

7.1.1 ベイズの定理と尤度関数による情報の変換 ..... 150

7.1.2 人間の考えの偏りに対するベイズ的な解釈 ..... 152

7.2 ベイズフィルタ ..... 153

7.2.1 ロボットの信念 ..... 153

7.2.2 観測に関する信念分布の変形 ..... 155

7.2.3 移動に関する信念分布の変形 ..... 155

7.2.4 導出されたベイズフィルタ ..... 156

7.3 ベイズフィルタの実装例 I: カルマンフィルタ ..... 157

7.3.1 線形な場合の観測の扱い ..... 157

7.3.2 非線形な場合の観測の扱い ..... 162

7.4 ベイズフィルタの実装例 II: パーティクルフィルタ ..... 167

7.4.1 パーティクルの再定義 ..... 168

7.4.2 パーティクルの更新 ..... 168

7.4.3 応 用 例 ..... 170

7.5 補 遺 ..... 173

7.6 ま と め ..... 175

## 8. 機械学習——ロボットはぼんやり実世界を観察する

8.1 バラバラなデータから像を見る I: ベイズ線形回帰 ..... 179

8.1.1 最小二乗法による直線の当てはめ ..... 180

8.1.2	ベイズ線形回帰	184
8.2	バラバラなデータから像を見る II：混合ガウス分布の推定	192
8.2.1	「確率の確率」による問題の定式化	194
8.2.2	変分推論による解き方	197
8.2.3	混合ガウス分布の推論	201
8.3	関数を自在に生成する	208
8.3.1	教師あり学習が扱う問題	208
8.3.2	情報をとりあえず全部入力してみる	209
8.3.3	訓練データ	210
8.3.4	損失関数の設定	211
8.3.5	深層学習と利用の例	212
8.3.6	教師あり学習の「学習」についての考察	216
8.4	まとめ	218

## 9. 意思決定と制御——ロボットは自律する

9.1	決めるということはどういうことか I	222
9.1.1	数式による表現	223
9.1.2	自己位置推定チャレンジの再考	227
9.1.3	最適化問題	230
9.2	決めるということはどういうことか II	231
9.2.1	多段の意思決定	231
9.2.2	状態の価値と計算方法	234
9.2.3	最適性とベルマン方程式	236
9.2.4	議論（着眼大局，損して得取れ）	237
9.3	ベルマン方程式と制御	239
9.3.1	最適制御の導出	241

9.3.2	$V^*$ の 形	244
9.3.3	議論：制御と意思、そして物理の関係について	245
9.4	強 化 学 習	246
9.4.1	学 習 の 原 理	246
9.4.2	議論：Q 学習からいえること、考えると面白いこと	249
9.5	さらに難しい問題	252
9.5.1	他者がいる系	253
9.5.2	POMDP	254
9.5.3	自由エネルギー原理と能動的推論	255
9.6	ま と め	256
	引用・参考文献	260
	索 引	267

# 第 1 章

## 代表値

### 統計の知識がないと喧嘩になる

最初の章では、まず統計というものを考える上で基礎となる代表値について考えを深めていきましょう。多くの人は統計を数学の一種だと認識していると思うのですが、よほど高度な研究をしない限りは、統計は「人と会話をするための道具」だと考えたほうがよい、というのが筆者の考えです。統計の基礎概念である代表値は、何かの実験結果や営業成績、あるいはロボコンの試合結果や途中の経過について、あれやこれやとチームで話すとき、知らないうちに使っているはずです。

本章で取り上げる代表値は、平均値など、小学生でも計算できるような簡単なものばかりです。ただ、計算ができることと使いこなすことの間には高い壁があります。例えば、学生が何か実験結果をまとめたレポートや論文を持ってくると、(まだ修行中なのでオッケーなのですが) たいていの場合、統計の使い方がなんとなく変です。そして、それが大変困ったことであると理解しないまま卒業したであろう大人は、SNS で変な統計を持ち出して他人をやり込めようと喧嘩をし、この世に地獄を造成しています。「たかが平均値」と侮らず、地獄からの救いを見出すために、基礎を確認していきましょう。

**例 1.1** (ある小学生の喧嘩) まず、代表値のない世界に行ってみましょう。そんな世界はないのですが、小学校低学年に戻ればそういう状態になります。小学生の A さんと B 君が、それぞれのロボットに周回コースを走らせ、コースアウトせず何周できるかという遊びをしている状況<sup>†</sup>を考え

<sup>†</sup> 小学校低学年がロボットで遊ぶかという話ではありますが、最近はそのような教材もあるようです。



## 2 1. 代表値——統計の知識がないと喧嘩になる

ましよう。2人は何回挑戦するかは決めずになんとか遊んでいましたが、各回でロボットが何周したかはきっちり記録していました。記録は、つぎのようなものでした。

- Aさん（10回挑戦）：2, 3, 1, 5, 0, 3, 7, 2, 4, 1
- B君（7回挑戦）：5, 3, 6, 4, 2, 5, 4

そして、遊び終わったあと、「ああ面白かった。また明日！」となればよかったのですが、「うちのロボット、どっちのほうが『つよい』<sup>†</sup>んだろう？」という話になってしまいました。Aさんが「私のロボットは7周できた！」と主張し、B君が「でもAさんのロボットはすぐにコースアウトすることが多いし、そもそも自分のほうが3回少ないし…」と反論し始めて、だんだん収拾がつかなくなって、最後は小学生らしい悪口の言い合いになりました。

### 1.1 代表値で比較する

小学生にはちょっと酷な評論になってしまいますが、いまの例のAさんとB君には、「得られた結果（数字）について会話する方法を知らない」という点で問題があります。このときに助けになる唯一の道具が統計の知識です。また、知識があって計算できるだけではなく、相手に勘違いさせないように、正しい言葉で伝えることが必要です。Aさん、B君の遊びを例として、初歩的な統計の考え方、用語を確認していきましょう。

#### 1.1.1 試行と事象

まず、ロボットの周回数「2, 3, 1, …」のように、何かの結果が偶然出てくるような行為を、統計学や確率論では**試行**と呼びます。また、試行の結果として起こる現象を**事象**と呼びます。「Aさんが10回挑戦」の「挑戦」は、統計の話

<sup>†</sup> 特に定義があるわけではなく、小学生的な漠然とした「つよさ」。

をするときは「試行」と表現するほうが話を通じます。また、「Aさんの1回目の試行の結果は2だった」は、「Aさんの1回目の試行の結果、事象として2が観測された」と堅苦しくいうこともできます。ただ、「試行の結果は2だった」で通じる場合は、特に事象という言葉を使う必要はありません。事象には、「コインを投げる試行をしたら表（裏）が出た」など、数字以外のもの<sup>†1</sup>もあり得ますが、しばらくは結果が数だとして話を進めます。また、本書では「試行の結果としての事象」を「試行結果」と表記することがあります。

試行と事象について、数学的な表現を導入しておきましょう。「Xさん（君）の*i*回目の試行の結果観測される事象（の値）」を $x_i$ と表すことにします。 $x_i$ の $x$ にはAさん、B君にそれぞれ対応する $a$ か $b$ を入れることとしましょう。このような記号を導入するのは、いちいち「Xさん（君）の*i*回目の試行の結果観測される事象の値」と書くと同面倒だからであって、数学が嫌いな人に対する嫌がらせではありません。 $x_i$ という記号を見ると「数学だ」と身構える人がいますが、数式を解けといわれたとき以外は、単なる国語の読解問題です。慣れようと思えばなんとかなります<sup>†2</sup>。

もう1つ、よく使われる記号を導入しておきます。「Xさん（君）の*i*回目から*j*回目の試行の結果観測される事象」を $x_{i:j}$ と表すことにします。これで、Aさん、B君の試行結果の事象は、「2,3,1,5,⋯」と具体的に数値を出さずに、それぞれ $a_{1:10}$ ,  $b_{1:7}$ と書くことができます。また、このような事象の並びや寄せ集め（数学の用語では列や集合）をデータと呼ぶことがあります。

### 1.1.2 代表値の選択と計算

試行や事象の定義や表記の説明は単なる前座で、ここからが本題です。データに対して、代表値というものを考えてみます。代表値というのは、「2, 3, 1,

---

<sup>†1</sup> 次章で出てきます。

<sup>†2</sup> ということで、「まえがき」でも述べたとおり、この本はためらわずに数学の記号を使います。その都度ちゃんと説明を入れますので、慣れていきましょう。一度慣れると一生お得です。ただ、「ちゃんと」にも限界はありますので、他の資料（教科書やインターネット上の文章や動画）も併用することをおすすめします。

5, 0, 3, 7, 2, 4, 1] というような、たくさん数字が並んだデータを、(乱暴に) 1つの数字で説明するためのもので、以下の値がよく使われます。

(1) **最大値, 最小値** 例えば, Aさんの「私のロボットは7周できた!」の7は、**最大値**と呼ばれる種類の代表値です。事象を数値で表したり置き換えたりできる場合に、結果の中で最大のものを指します。数学の記号を使うと、データ  $x_{1:N}$  の最大値を

$$\max_{i=1,2,\dots,N} x_i \quad \text{あるいは} \quad \max x_{1:N} \quad (1.1)$$

などと表せます<sup>†</sup>。上式の右の表記を使うと  $\max a_{1:10} = 7$  となり、B君の結果の最大値  $\max b_{1:7} = 6$  を上回ります。つまり、この点ではAさんのほうが優れていることになります。

最大値があれば、**最小値**

$$\min_{i=1,2,\dots,N} x_i \quad \text{あるいは} \quad \min x_{1:N} \quad (1.2)$$

もあります。最小値を比較すると、 $\min a_{1:10} = 0$ 、 $\min b_{1:7} = 2$  となり、この点ではB君のほうが優れていることになります。ということで、最大値、最小値で比較すると、どちらが優れていたかはわかりません。B君があと3回挑戦していたら、B君の最大値、最小値は変わっていた可能性もあります。また、最大値、最小値だけでは、それ以外の値だった試行に関する情報が抜けてしまいます。

(2) **平均値** では、特定の試行だけではなく、各試行をまんべんなく反映した代表値はないのでしょうか。後述のように万能な代表値は存在しないのですが、無難なものとして**平均値**があります。平均値は、試行結果をすべて足して、試行の回数で割った数で、試行結果  $x_{1:N}$  に対して

$$\bar{x}_{1:N} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i \quad (1.3)$$

<sup>†</sup> 左の表記法の場合、 $\max$  の下に書いてあるのが範囲で、この場合は「 $i$  が 1 から  $N$  のもののうち、最大の  $x_i$  の値」という意味になります。

で定義できます<sup>†1</sup>。この式の  $\bar{x}_{1:N}$  のように、本書では記号の上の横棒を、データに対する平均値の印とします。

A さん、B 君の全試行について、それぞれ平均値を計算してみましょう。

- $\bar{a}_{1:10} = \frac{2+3+1+5+0+3+7+2+4+1}{10} = 2.8$
- $\bar{b}_{1:7} = \frac{5+3+6+4+2+5+4}{7} \simeq 4.1$

となります<sup>†2</sup>。

この 2.8, 4.1 は、どの試行の値を変えても変化するので、全部の試行の結果を反映しているといえ、最大値、最小値のように 1 つの試行だけに焦点を当てた代表値よりも、全試行の傾向を良く表しているように思われます。また、試行の回数が多くても少なくても、有利不利は出ないように思われます。B 君の「でも A さんのロボットはすぐにコースアウトすることが多いし…」という主張は、おそらく「平均値で考えたときに自分のほうが優れている」という思考に基づくものでしょう。この思考には一定の合理性があります。

しかし、例えば走り高跳びや棒高跳びのように、世の中には最大値で優勝が決まる競技もあるので、B 君の主張が 100% 通るわけでもありません。平均値はほとんどの人が知っている代表値かもしれませんが、それですべてが語れるわけではありません。

<sup>†1</sup>  $\sum$  も嫌がる人がいますが、「 $x_1 + x_2 + \dots + x_N$ 」とか「 $i = 1$  から  $N$  まで  $x_i$  を全部足す」とか書くのが面倒なので、 $\sum_{i=1}^N x_i$  と省略して書いているにすぎません。

<sup>†2</sup> 「 $\simeq$ 」は、左辺と右辺がまったく同じではなく、右辺が近似した数になっているという意味で使っています。B 君の平均値は 4.1428... ですが、工学では小さい桁の数値を表示すると、「その小さい桁になんらかの意義があるから無視するな」という意味になります。例えば 4.1428 まで書いておくと、「小数第 4 位まで意味がある」ということを堂々と宣言することになります。しかし、この例では試行の回数が少なすぎて、小さい桁には意味がないでしょう。ここでは 4.1 と書いていますが、たぶん 0.1 もあまり意味がなく、4 と書くほうがより適切でしょう。意味のある桁数は、有効数字と呼ばれます。この桁数をどうするかという細かい話がありますが、とりあえず自分が意味があると考えた桁まで書いておくのがよいでしょう。最近、表計算ソフトで算出した桁の多い小数を、そのまま論文やスライドに貼る人が（プロの研究者でも）多いので、ついつい説明がしつこくなりました。

# 索引

## 【あ】

アウトライア 18

## 【い】

意思決定 221

一次変換 102

位置ベクトル 72

一様分布 64

## 【う】

ウイシャート分布 196

ウォッチドッグ 50

## 【え】

エージェント 113, 225

エントロピー 198

## 【お】

驚き → サプライズ

重みつき平均 57

## 【か】

回転行列 99

ガウス-ウイシャート分布

196

ガウス-ガンマ分布 188

ガウス分布 86

過学習 218

学習 177

学習率 247

拡張カルマンフィルタ 157

確率 24

確率質量関数 43

確率分布 44

確率変数 43

確率密度関数 82

確率モデル 64

学力偏差値 92

偏り 17

価値 234

価値反復 237

加法定理 31

カルバック-ライブラー

情報量 197

カルマンゲイン 161

カルマンフィルタ 87, 157

観測 154

ガンマ関数 206

ガンマ分布 187

## 【き】

機械学習 177

期待値 53, 54, 56

期待値の線形性 58

逆強化学習 252

キャリブレーション 17

強化学習 246

教示 210

教師あり学習 208

教師なし学習 208

共分散 66

共分散行列 96

共役勾配法 183

共役事前分布 124

共役性 124

行列式 96

## 【く】

空間 74

空事象 31

クラスタ 179

クラスタリング 192

グリーディ方策 249

群知能 254

訓練 211

訓練データ 211

訓練データセット 211

## 【け】

計画 221

結合分布 199

決定論の方策 228

ゲーム理論 254

## 【こ】

交差エントロピー 198

校正 17

行動 223

行動価値関数 236

行動列 233

勾配ベクトル 182

誤差 181

誤差楕円 101

根元事象 25

コンコルド効果 238

コントローラ 225

## 【さ】

最急降下法 183

最小値 4

最小二乗法	180, 181
再生性	87
最大値	4
最適化問題	230
最適状態値関数	237
最適制御	239
最適方策	236
最適レギュレータ	245
最頻値	10
座標系	136
サプライズ	115
サンクコスト効果	238
サンプリング	64
サンプル	64, 145
<b>【し】</b>	
時間	127
自己位置推定	153
試行	2
時刻	127
事後分布	111
姿勢	72
事前分布	111
従う	45
シックスシグマ	93
実数	74
四分位偏差	18
シミュレーション	61
ジャイロ	17
シャノンのサプライズ	
→ サプライズ	
自由エネルギー原理	256
集合	3, 47
集合の分配法則	39
収束	65
従属	30
終端状態	233
周辺化	47
周辺確率分布	47
周辺尤度	115
条件つき確率	28
条件つき確率質量関数	47
条件つき確率分布	47

状態	129
状態値関数	234
状態遷移関数	129
状態遷移分布	130
状態遷移モデル	130
状態方程式	129
冗長化	31
情報	150
乗法定理	28
初期状態	234
人工ニューラルネット	
ワーク	212
深層学習	177, 213
深層ニューラルネット	
ワーク	213
信念	154
信念分布	154
<b>【す】</b>	
随意運動	226
スカラー	96
<b>【せ】</b>	
正解	211
正規化定数	93
正規分布	86
正規分布の再生性	87
制御	221
制御器	225
制御指令	128
正定値行列	96
精度	94
制約条件	231
世界座標系	136
線形化	138
線形近似	148
線形計画法	231
線形計画問題	231
線形変換	102
潜在変数	195
全事象	25
全体主義	253
占有格子地図	171

**【そ】**

測域センサー	170
ソフトウェアテスト	50
損失	224
損失関数	181

**【た】**

代表値	1, 3
多変量ガウス分布	95
多変量確率質量関数	46
多変量確率分布	46
多変量確率変数	76
多変量正規分布	95
単一障害点	35

**【ち】**

知覚の見せかけ問題	174
中央値	7
直列システム	24

**【て】**

ディガンマ関数	206
定数	57
ディープニューラルネット	
ワーク	213
ディリクレ分布	196
定量的	13
敵対的生成ネットワーク	216
データ	3
転置	95
点ランドマーク	163

**【と】**

統計量	19
同時確率質量関数	46
同時確率分布	46
特徴量	209
独立	30
独立同分布	45
度数分布表	75
ドリフト	17
トレース	196

ドロー	64	不偏分散	14	問題の切り分け	49
ドローイング	64	分散	14	モンテカルロ積分	80
<b>【に】</b>		分散共分散行列	96	モンテカルロ法	80
任意の	57	<b>【へ】</b>		<b>【や】</b>	
<b>【ね】</b>		平均値	4	ヤコビ行列	139
ネイピア数	85	平均二乗誤差	181	<b>【ゆ】</b>	
<b>【の】</b>		ベイズ線形回帰	184	誘拐ロボット問題	174
能動的推論	256	ベイズの定理	113	有限マルコフ決定過程	232
<b>【は】</b>		ベイズフィルタ	156	有効数字	5
バイアス	17	平方完成	89	尤度	113
排反	32	並列システム	31	尤度関数	113
バスタブ曲線	52	ベータ分布	122	尤度場	172
外れ値	18	ベルヌーイ試行	122	有理数	74
パーティクル	145	ベルヌーイ分布	45	<b>【よ】</b>	
ハミルトン-ヤコビ-ベルマン		ベルマン最適方程式	237	要約統計量	19
方程式	239	ベルマン方程式	237	余事象	34
ハミルトン-ヤコビ方程式	243	偏微分	139	<b>【ら】</b>	
ばらつき	12	変分事後分布	202	ライダーオドメトリ	170
半正定値行列	96	変分自由エネルギー	256	ラベル	211
<b>【ひ】</b>		変分推論	197	ランドマーク	163
ヒストグラム	75	変分ベイズ	197	<b>【り】</b>	
ヒストグラムフィルタ	147	<b>【ほ】</b>		離散化	75
非線形	138	方策	223	離散型確率変数	74
評価関数	226, 231	方策改善	237	離散時間系	127
標準偏差	15	方策評価	235	リサンプリング	169
標本	64	報酬	224	<b>【る】</b>	
標本抽出	64	報酬関数	224	ルベーグ積分	84
標本分散	14	<b>【ま】</b>		<b>【れ】</b>	
頻度	11	マハラノビス距離	92	列	3
<b>【ふ】</b>		マルコフ決定過程	232	連続型確率変数	74
フィルタ	156	マルコフ性	129	<b>【ろ】</b>	
不随意運動	226	マルチパス	159	ログ	50
負担率	201	<b>【み】</b>		ロバスト	21
部分観測マルコフ決定過程	255	密度	78, 81	ロボット座標系	136
不偏共分散	97	<b>【も】</b>			
		目的関数	231		
		尤もらしい	114		
		モード	11		

<b>[B]</b>		<b>[K]</b>	<b>[Q]</b>
belief MDP	254	KL (Kullback–Leibler) 情報量 → カルバック–ライブラー 情報量	Q 学習 246
<b>[D]</b>		<b>[L]</b>	<b>[S]</b>
diag	99	LiDAR 170	SLAM 219
draw	64		<b>[T]</b>
drawing	64		tr 196
<b>[F]</b>		<b>[M]</b>	~~~~~
FastSLAM	220	Markov decision process 232	<b>【数字・記号】</b>
finite MDP	232	MDP 232	3 シグマ 93
<b>[G]</b>		<b>[N]</b>	* 119
GAN	216	$\mathcal{N}$ 87	$\cap$ 27
GNSS	157	$n$ シグマ範囲 92	$\cup$ 25
GPS	158		$\implies$ 27
graph-based SLAM	219	<b>[P]</b>	$\wedge$ 155
<b>[I]</b>		POMDP 255	$\langle f(x) \rangle_{P(x)}$ 56
iid	45	$\Pr\{ \}$ 24	$\infty$ 93
			$\sim$ 45
			$\emptyset$ 31



—— 著者略歴 ——

2001年 東京大学工学部精密機械工学科卒業  
2003年 東京大学大学院工学系研究科修士課程修了（精密機械工学専攻）  
2004年 東京大学大学院工学系研究科博士課程中退  
2004年 東京大学助手  
2007年 博士（工学）（東京大学）  
2008年 東京大学助教  
2009年 有限会社 USP 研究所技術研究員  
2013年 産業技術大学院大学助教  
2015年 千葉工業大学准教授  
現在に至る

## ロボットの確率・統計

–製作・競技・知能研究で役立つ考え方と計算法–

Introduction to Probability and Statistics for Robotics

–Basis for Thinking and Mathematical Manipulation on Creation, Competition, and Research–

© Ryuichi Ueda 2024

2024年3月21日 初版第1刷発行



検印省略

著者 うえだ りゅういち  
発行者 株式会社 コロナ社  
代表者 牛来真也  
印刷所 三美印刷株式会社  
製本所 有限会社 愛千製本所

112-0011 東京都文京区千石 4-46-10  
発行所 株式会社 コロナ社  
CORONA PUBLISHING CO., LTD.  
Tokyo Japan

振替 00140-8-14844・電話(03)3941-3131(代)  
ホームページ <https://www.coronasha.co.jp>

ISBN 978-4-339-04687-8 C3053 Printed in Japan

(新宅) G



<出版者著作権管理機構 委託出版物>

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつと事前に、出版者著作権管理機構（電話 03-5244-5088, FAX 03-5244-5089, e-mail: info@jcopy.or.jp）の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。