

確率論の基礎と応用

－ 経済系への確率論から数理統計まで－

西岡 國雄
中央大学商学部

2007 年 12 月 19 日

目次

第 I 部	確率論	1
第 1 章	確率とは何か	3
1.1	確率論の始まり—ラプラスによる確率の定義	3
1.2	有限から無限事象へ	4
1.3	近代確率論へ	7
第 2 章	確率空間と条件付き確率	9
2.1	確率空間	9
2.2	確率空間の例	10
2.3	条件付き確率と独立	12
2.4	ベイズの公式の応用例	14
2.5	練習問題	19
第 3 章	確率変数と平均	21
3.1	確率変数と確率分布	21
3.2	平均とその応用	23
第 4 章	確率変数の分散と相関係数	29
4.1	確率変数の分散	29
4.2	共分散と相関係数	33

第 I 部

確率論

第 1 章

確率とは何か

確率論を一口で言うと、‘偶然量の法則性を調べる数学’となる。ところが偶然というのは、‘本来予測できない’という意味だから、‘その法則性’という説明は自己矛盾とも考えられる。

しかし実際には、‘粒子の存在確率’という偶然性の概念が量子力学の基礎となっており、それを取り除くと幾つもの物理現象の説明は大変難しくなる^{*1}。また近年金融の基礎となりつつある「数理ファイナンス」は、株価の変動をある法則性を持つ確率現象と捉え、‘オプション’と呼ばれる金融派生商品の適正価格を決定する理論である。これらは‘偶然量の法則’が実社会に意味のある研究対象であることを示す例と言える。

それではどういう‘偶然量’が確率論の対象になるかという、何回でも繰り返すと何らかの安定性のある量が現れるような‘偶然量’だけを対象としている。この安定性は‘極限定理’という言葉で厳密に定義されているが、良く知られている‘極限定理’としては、「大数の法則」「ポアソンの小数法則」「中心極限定理」などが有る。これらの極限定理どれも保険業務や統計などの基礎理論であり、社会と密接に関連している。

1.1 確率論の始まり—ラプラスによる確率の定義

1.1.1 黎明期

数学史で確率論の発端といわれるのは、パスカルが父親から、「ゲームを途中で止めたときに掛け金をどう配分したらよいか?」^{*2}という質問を受けた事とされている。掛け金は「賭の勝率の期待値」に従って配分するので、これは確率計算の問題になる。

これ以前にもケプラーやガリレオなども確率の計算結果を発表したり、本を書いたりしている。また零の発見国として実績のあるインドやサラセンなどでも確率の計算はなされていたようである。

しかし、パスカルがフェルマーと確率について書簡を往復させ、その後ベルヌーイが「大数の法則」（これは確率論初期からのブレイクスルーである）を発見し、やがてドゥ・モワブルの「中心

^{*1} 有名なアインシュタインは、当初‘偶然性’を物理学に持ち込むことに反対で、「神はサイコロを振らない」という言葉を残した。しかし晩年には、その立場を変えた。

^{*2} A, B 二人が、先に N 勝した方が勝つという競技をしている。それぞれのゲームで A の勝つ確率は p , B の勝つ確率は $q \equiv 1 - p$ 。既に、A は $N - m$ 勝, B は $N - n$ 勝している。A が競技に勝つ確率 Q を求めよ。

答えは $Q = \sum_{k=m}^{m+n-1} \binom{k}{m} p^m q^{k-m}$ 。

「極限定理」へと発展したこと考えると、やはり パスカルの始祖説も妥当と思える。

ここで、確率論初期に登場した人達を挙げてみる：

フェルマー (P. Fermat)	1601–1665
パスカル (B. Pascal)	1623–1662
ベルヌーイ (J. Bernoulli)	1654–1703
ドゥ・モワブル (de Moivre)	1667–1754
ラプラス (P.S. Laplace)	1749–1827

この頃の確率論の問題は、「サイコロの問題」とかごく簡単なものが対象なので、「確率」という概念も厳密な吟味を必要としなかった。例えばラプラスは、確率の定義を

定義 1.1.1 (ラプラス). 全体が同じように起こりうる n 個の場合から成る. 事象 A がそのうちの k 個の場合から成るとき、事象 A の起こる確率 $P[A] = k/n$ である。

としているが、これで特に不都合は無かったようである^{*3}。

1.1.2 ベルヌーイの発見-大数の法則

17 世紀までの確率論では、有限な事象に対して組み合わせ論的な計算で確率を求めることが主な研究だったから、現在でいえば「例題の解決」といったものだった。

この段階から本質的な飛躍が起こったのがベルヌーイの「大数の法則」の発見である。「大数の法則」は後で述べる^{*4}が、その直感的イメージは次のようなものである：

「大数の法則」のイメージ： 公平なサイコロを n 回振るとき、1 の目がでる回数を n_1 とする。すると $n \rightarrow \infty$ のとき、 $n_1/n \rightarrow 1/6$ となる。

これは

独立で同じ成功確率^{*5} $1/6$ をもつ n 個のランダムな量の成功回数の算術平均は、 n を大きくするとその確率 $1/6$ に収束する。

ということである。

別の言い方をすると、「無限個のランダムな量に関する法則性が発見された」と見る事ができ、確率論と実社会との接点となる発見であった。

1.2 有限から無限事象へ

ベルヌーイの大数の法則がそれまでの確率論からの“本質的な飛躍”と言う意味は、無限個のランダムな量を扱う故である。無限個の量を扱う場合には、「ラプラスによる確率の定義 1.1.1」自体の正当性を再検討する必要がある。

^{*3} 高校数学で学ぶ確率論でも、「確率の定義」として 定義 1.1.1 を採用している。

^{*4} 第 ??, 定理 ??.

^{*5} 1 の目が出るとき「成功」とする。

実際に、ラプラスの定義 1.1.1 では不十分な例を挙げてみよう。
例えば、‘起こりうる場合が数えられない’ とする；

練習問題 1.2.1. ランダムな量 X の値は区間 $[0, 1]$ の中にあり、どの値も同じようにとる、このとき、 X が区間 $[a, b]$ で値をとる確率 p を求めよ。

[解答] ‘ X は区間 $[0, 1]$ のどの値も同じようにとる’ のだから、

$$(1.2.1) \quad p = [a, b] \text{ が } [0, 1] \text{ で占める割合}$$

とするのが ‘自然’ な拡張であろう。よって

$$p = \frac{b-a}{1} = b-a. \quad \square$$

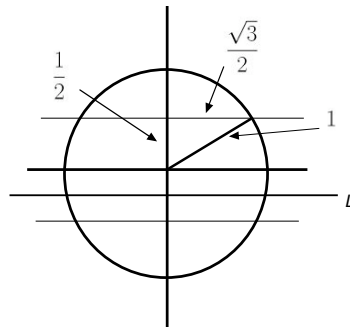
確かに、(1.2.1) による ‘確率の拡張’ は自然で妥当な感じがする。しかし、ランダムな量 X の値が、区間 $[a, b]$ よりもずっと複雑な場合もこの ‘自然’ な拡張は可能なのだろうか？

そういう複雑な場合の例として、次の「ベルトランドのパラドクス」がある。

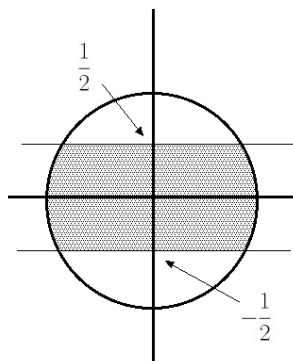
練習問題 1.2.2 (ベルトランドのパラドクス). 半径 1 の円周 C がある。これと交差するように直線をかたより無く引く。このとき、円周と直線との交点で出来る弦の長さが $\sqrt{3}$ 以上である確率 q を求めよ。

いまこの問題に、二通りの異なった解答を提出する。さてどちらが正解だろうか？ それとも正解は別に有るのだろうか？

[解答 1] 直線はかたより無く引かれているので、直線の方角を回転してよい。つまり、全ての直線は x -軸に平行としてよい。



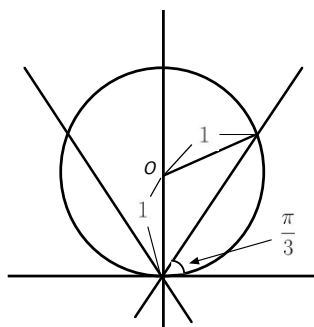
この直線 L と円周 C との 2 交点の距離が $\sqrt{3}$ 以上である事と、 L と y -軸との交点が円の中から $1/2$ 以内である事は必要十分。直線 L はかたより無く引かれているから、確率は “2 点 $(0, 1)$, $(0, -1)$ の間の距離” にたいする “2 点 $(0, 1/2)$, $(0, -1/2)$ の距離” の比になる。



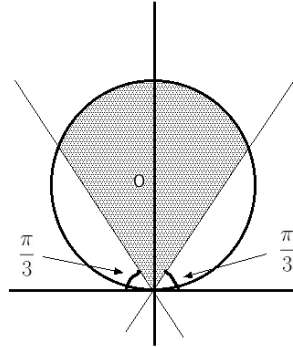
つまり, $q = 1/2$. $\square$

[解答 2] 直線はかたより無く引かれるから, 任意に指定した点 O を通るとしてよい.

O を通って円周 C に接する直線を x -軸とする.



O を通る直線 L と円周 C の 2 交点の距離が $\sqrt{3}$ 以上であるためには, L と x -軸とのなす角度が, $\pi/3$ 以上であることになる. 直線 L が円周 C と 2 点で交わるために可能な角度は $\pi/3$ 度. 直線はかたより無く引かれているから, 確率は “ $\pi/3$ と π の比” になる.



つまり, $q = 1/3$. □

1.3 近代確率論へ

1.3.1 「ベルトランドのパラドクス」への解答

まず 問題 1.2.2 の解答を考えてみよう. どちらも妥当に思える 2 つの解答が提出され, 異なる結論

$$q = 1/2, \quad q = 1/3$$

が得られた. こうなった原因は, 問題文にある「直線をかたより無く引く」の解釈として

(1.3.1a) 直線は x 軸に平行で, 区間 $[-1, 1]$ のどの値も同じように取る,

(1.3.1b) 直線は点 O を通り, x 軸となす角度 $[0, \pi]$ のどの値も同じように取る,

2 つが可能であったからである. そこで, 問題 1.2.2 に対しては「問題文に曖昧さがあり, 解答不能」が正解といえる.

あるいは, 「直線をかたより無く引く」を削除し, (1.3.1a) もしくは (1.3.1b) を問題文に書き加えれば 解答 1, 2 のどちらかが正解となる. □

1.3.2 近代確率論の立場 – 確率空間

近代の確率論では「無限個の値をとる偶然量」が主たる対象である. そこでは「ベルトランドのパラドクス」と同じように, 「確率の解釈での曖昧さ」が出現するので

まず確率を定義してから議論を開始する

との立場が共通認識となっている. より詳しくいうと, 近代確率論では

- (a) 観測する対象の全体: 「事象空間」と呼ばれ, Ω と表記されることが多い.
- (b) 数学的な操作ができる対象のクラス: 「加算加法族」と呼ばれ, $\mathcal{F}$ 等と表記される. 数理ファイナンスや確率論の専門家には重要なことだが, 本書の段階では気にしなくて良い.

(c) 確率: 事象空間 Ω に属する全てのものにたいし, それが起こる確率 $\mathbf{P}$ を定義する. $\mathbf{P}$ は確率測度と呼ばれる.

の三つ組み $(\Omega, \mathcal{F}, \mathbf{P})$ を予め用意して, 議論を開始する. つまり

確率は予め宣言するものであり, 観測や経験から導かれるものではない

が「確率とはなにか」への回答と要約できる.

注意 1.3.1. (i) この三つ組み $(\Omega, \mathcal{F}, \mathbf{P})$ は **確率空間** と呼ばれ, これを準備することが近代確率論の特徴である. $\mathcal{F}$ や $\mathbf{P}$ は幾つかの条件を満たすことが要請されているが, その概略は §2.1 で述べる.

(ii) ただし, 本書で主に扱う ‘ Ω が有限個の場合’ には, 前記 (i) の厳密な準備は必要が無い場合が多く, 「ラプラスによる確率の定義 1.1.1」でもそれほど不都合は生じない.

(iii) 事象空間 Ω が無限個からなるとき, 上記 (c) (事象空間 Ω に属する全てのものにたいし, それが起こる確率 $\mathbf{P}$ を定義する) を実行することは, 決して簡単ではない.

20 世紀初頭に, ロシア人の数学者 A. N. コルモゴロフ が (iii) を実行する一般的な方法*6を提案し, 近代確率論への道を切り開いた.

*6 「コルモゴロフの拡張定理」と呼ばれ, 証明には長い準備と厳密な議論を要する.

第 2 章

確率空間と条件付き確率

前章で述べた「ベルトランドのパラドクス」は、`確率の定義自体`に曖昧さがあったため、パラドクスが生じた。そうした事態を避けるため、現代の確率論では

まず適当な確率空間 $(\Omega, \mathcal{F}, \mathbf{P})$ を定義してから、議論を始める

という方法をとる。私達も現代確率論の文法に従って議論を進めよう。

2.1 確率空間

私たちは、三つ組み $(\Omega, \mathcal{F}, \mathbf{P})$ を **確率空間** probability space と呼ぶ。その各項について説明しよう。

(i) Ω は n 個の元からなる集合

$$\Omega = \{w_1, w_2, \dots, w_n\}$$

であり、`事象空間 sample space` と呼ばれ^{*1}、その元 $w_k \in \Omega$ や 部分集合 $A \subset \Omega$ は `事象 sample` と呼ばれる。

(ii) $\mathcal{F}$ は Ω の中集合族 (=すべての部分集合の集まり) power sets である。つまり

$$\mathcal{F} = \{\emptyset, w_1, \dots, w_n, w_1 \cup w_2, \dots, w_1 \cap w_2 \cap w_3, \dots, \Omega\}$$

という 2^n 個の部分集合の集まりである。

(iii) $\mathbf{P}$ は、すべての $A \in \mathcal{F}$ 毎に数値 $\mathbf{P}[A]$ が対応する関数 (集合関数) で `確率測度 probability measure` と呼ばれる。ただし次の性質を満たさなければならない:

$$(2.1.1a) \quad 0 \leq \mathbf{P}[A] \leq 1: \quad \text{とくに} \quad \mathbf{P}[\Omega] = 1, \quad \mathbf{P}[\emptyset] = 0,$$

$$(2.1.1b) \quad A_1, A_2 \in \mathcal{F} \text{ かつ } A_1 \cap A_2 = \emptyset \Rightarrow \mathbf{P}[A_1 \cup A_2] = \mathbf{P}[A_1] + \mathbf{P}[A_2].$$

注意 2.1.1. (i) ここで強調したいことは、「確率 $\mathbf{P}$ は予め与えられており、観測などから自動的に導かれるものではない」ことである。

^{*1} しばしば、標本空間とも呼ばれる

(ii) 上記 (i) と異なり, ‘事象空間 Ω が無限個の元からなる場合’ も取り扱う必要が生じる. その場合, $\mathcal{F}$ として何を考えるかに選択と工夫がある. しかし私達が主に扱う ‘ Ω が有限個の場合’ には, そうした選択の余地が無い.

そこで今後は特別な場合を除き, $\mathcal{F}$ を省いた 2 つ組 $(\Omega, \mathbf{P})$ で確率空間を表記する. $\diamond$

今後, 必要となる確率測度の性質を列挙する.

補題 2.1.2. 確率空間 $(\Omega, \mathbf{P})$ で次が成立する.

$$(i) \mathbf{P}[A^c] = 1 - \mathbf{P}[A], \quad (ii) A \subset B \Rightarrow \mathbf{P}[A] \leq \mathbf{P}[B]$$

(iii) 集合の集まり $\{A_1, A_2, \dots, A_m\}$ は次を満たしている:

$$(2.1.2) \quad i \neq j \text{ であるとの } A_i, A_j \text{ にたいしても } A_i \cap A_j = \emptyset$$

$$\Rightarrow \mathbf{P}[A_1 \cup A_2 \cup \dots \cup A_m] = \mathbf{P}[A_1] + \mathbf{P}[A_2] + \dots + \mathbf{P}[A_m]. \quad \diamond$$

[証明] (i) $\Omega = A \cup A^c$ かつ $A \cap A^c = \emptyset$ だから, (2.1.1a) と (2.1.1b) より

$$1 = \mathbf{P}[\Omega] = \mathbf{P}[A \cup A^c] = \mathbf{P}[A] + \mathbf{P}[A^c] \Rightarrow \mathbf{P}[A^c] = 1 - \mathbf{P}[A].$$

(ii) $A \subset B$ なら $B = A \cup (A^c \cap B)$. 当然 $A \cap (A^c \cap B) = \emptyset$ だから, (2.1.1b) より

$$\mathbf{P}[B] = \mathbf{P}[A \cup (A^c \cap B)] = \mathbf{P}[A] + \mathbf{P}[A^c \cap B].$$

ここで (2.1.1a) より $\mathbf{P}[A^c \cap B] \geq 0$ となるので

$$\mathbf{P}[B] = \mathbf{P}[A] + \mathbf{P}[A^c \cap B] \geq \mathbf{P}[A].$$

(iii) 数学的帰納法を使う. $m = 2$ の場合は (2.1.1b) より明らかに (iii) は成立. つぎに $m = k$ で (iii) が成立すると仮定し, $m = k + 1$ でも成立することを示す.

まず $C \equiv A_1 \cup \dots \cup A_k$ とおく. 仮定 (2.1.2) より

$$C_1 \cap A_{k+1} = (A_1 \cap A_{k+1}) \cup (A_2 \cap A_{k+1}) \cup \dots \cup (A_k \cap A_{k+1}) = \emptyset \cup \emptyset \cup \dots \cup \emptyset = \emptyset$$

となるので, (2.1.1b) より

$$\mathbf{P}[A_1 \cup \dots \cup A_k \cup A_{k+1}] = \mathbf{P}[C \cup A_{k+1}] = \mathbf{P}[C] + \mathbf{P}[A_{k+1}]$$

ここで $m = k$ では (iii) が成立しているから

$$\mathbf{P}[C] = \mathbf{P}[A_1] + \mathbf{P}[A_2 \cup \dots \cup A_k]$$

である. よって (iii) は $m = k + 1$ で成立した. $\square$

2.2 確率空間の例

今後の議論で良く登場する確率空間を列挙する.

例題 2.2.1 (1 回のコイン投げ). ‘不公正なコインを 1 回投げる’ 場合の確率空間を構成しよう. 起こりえる事象は

$$w_1 \equiv \{ \text{表がでる} \}, \quad w_2 \equiv \{ \text{裏が出る} \}$$

の二つだから 事象空間 は $\Omega = \{w_1, w_2\}$. またコインは不公正なので, 実数 p ($0 < p < 1$) が

$$\mathbf{P}[w_1] = p, \quad \mathbf{P}[w_2] = 1 - p$$

である.

例題 2.2.2 (1 回のサイコロ投げ). サイコロを 1 回投げる. 起こり得る事象は

$$w_1 \equiv \{1 \text{ の目が出る} \}, \quad w_2 \equiv \{2 \text{ の目が出る} \}, \dots, w_6 \equiv \{6 \text{ の目が出る} \}$$

の 6 通りであり, 事象空間は $\Omega = \{w_1, \dots, w_6\}$ である*².

また, このサイコロが公正なものとすると, 確率測度 $\mathbf{P}$ は

$$\mathbf{P}[w_1] = \frac{1}{6}, \quad \mathbf{P}[w_2] = \frac{1}{6}, \quad \dots, \mathbf{P}[w_6] = \frac{1}{6}$$

と定義する. これで公平なサイコロを 1 回投げる場合の確率空間が得られた. $\diamond$

例題 2.2.3 (2 回のコイン投げ). コインを 2 回投げる. 起こり得る事象は

$$\begin{aligned} w_1 &\equiv \{1 \text{ 回目=表, } 2 \text{ 回目=表} \}, w_2 \equiv \{1 \text{ 回目=裏, } 2 \text{ 回目=表} \}, \\ w_3 &\equiv \{1 \text{ 回目=表, } 2 \text{ 回目=裏} \}, w_4 \equiv \{1 \text{ 回目=裏, } 2 \text{ 回目=裏} \}, \end{aligned}$$

の 4 通りであり, 事象空間は $\Omega = \{w_1, \dots, w_4\}$ となる, 次に

(2.2.1) 投げたコインは必ずしも公正ではなく, 表が出る確率が p , $0 < p < 1$,

とする. このとき 確率測度 $\mathbf{P}$ を

$$\mathbf{P}[w_1] = p^2, \quad \mathbf{P}[w_2] = \mathbf{P}[w_3] = p(1 - p), \quad \mathbf{P}[w_4] = (1 - p)^2$$

と定義すれば ‘不公正なコインを 2 回投げる’ 場合の確率空間が得られた. $\diamond$

次の例の様に, 事象の数が多いときに, 確率空間をきちんと定義することは容易ではない. 20 世紀初頭にロシア人の数学者コロモゴロフ Kolmogorov が,

例題 2.2.4 で $n \rightarrow \infty$ の場合にも, 確率空間を定義できる方法 (= Kolmogorov の拡張定理)

を証明し, 近代確率論が誕生した.

最後に, 何回もコインを投げる場合を考えよう.

*² 念のため述べると, $\mathcal{F}$ は Ω の中集合族

$$\mathcal{F} \equiv \{\emptyset, \Omega, w_1, \dots, w_6, w_1 \cup w_2, \dots, w_5 \cup w_6, w_1 \cup w_2 \cup w_3, \dots\}$$

で $2^6 = 64$ 個の集合から成っている.

例題 2.2.4 (n 回のコイン投げ). (2.2.1) で述べた不公正なコインを n 回投げる. 起こり得る事象は

$$\begin{aligned}
 w_1 &\equiv \{1 \text{ 回目=表}, 2 \text{ 回目=表}, 3 \text{ 回目=表}, \dots, n \text{ 回目=表}\}, \\
 w_2 &\equiv \{1 \text{ 回目=裏}, 2 \text{ 回目=表}, 3 \text{ 回目=表}, \dots, n \text{ 回目=表}\}, \\
 (2.2.2) \quad w_3 &\equiv \{1 \text{ 回目=表}, 2 \text{ 回目=裏}, 3 \text{ 回目=表}, \dots, n \text{ 回目=表}\}, \\
 &\vdots \\
 w_N &\equiv \{1 \text{ 回目=裏}, 2 \text{ 回目=裏}, 3 \text{ 回目=裏}, \dots, n \text{ 回目=裏}\},
 \end{aligned}$$

の $N = 2^n$ 通りであり, 事象空間は $\Omega = \{w_1, \dots, w_N\}$ となる.

このように極めて多数の元からなる Ω を扱う場合には, (2.2.2) にたいし次の新しい表記法を導入した方が分かり易くなる.

定義 2.2.5 (新しい表記法). 1 が表, 0 が裏を表すとして,

$$\begin{aligned}
 (2.2.3) \quad w \in \Omega \text{ を } n \text{ 次元ベクトル } w = (w^{(1)}, w^{(2)}, \dots, w^{(n)}) \text{ と考える.} \\
 \text{ただし } w \text{ の各成分 } w^{(k)} \text{ は } 1=\text{表} \text{ もしくは } 0=\text{裏} \text{ である. } \heartsuit
 \end{aligned}$$

この表記法を使うと,

$$w_1 = 1 \text{ 回目=表}, 2 \text{ 回目=表}, 3 \text{ 回目=表}, \dots, n \text{ 回目=表} \Leftrightarrow w_1 = (1, 1, 1, \dots, 1)$$

となる. つまり (2.2.2) は

$$w_1 = (1, 1, \dots, 1), w_2 = (0, 1, \dots, 1), w_3 = (1, 0, 1, \dots, 1), \dots, w_N = (0, 0, \dots, 0)$$

と簡明に表記できる. さらに (2.2.3) で表記された $w \in \Omega$ にたいし

$$|w| \equiv w^{(1)} + w^{(2)} + \dots + w^{(n)} = \text{表がでた回数}$$

と定義すると, 確率測度 $\mathbf{P}$ は

$$(2.2.4) \quad w \in \Omega \text{ にたいし } \mathbf{P}[w] = p^{|w|} (1-p)^{n-|w|}$$

と簡単に定義できる. $\diamond$

2.3 条件付き確率と独立

確率空間 $(\Omega, \mathbf{P})$ を固定する.

定義 2.3.1. $\mathbf{P}[B] \neq 0$ である事象 A, B にたいし,

事象 B が起こるとするときに, 事象 A が起こる確率

を **条件付き確率** conditional probability $\mathbf{P}[A/B]$ と呼び,

$$(2.3.1) \quad \mathbf{P}[A/B] \equiv \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} \quad (\text{ベイズの公式})$$

で定義する*3. ◇

定義 2.3.2. ‘事象 A と B が独立 independent’ とは

$$(2.3.2) \quad \mathbf{P}[A \cap B] = \mathbf{P}[A] \cdot \mathbf{P}[B]$$

が成立することである. ◇

ベイズの公式を使うと, ‘独立’ が ‘事象 A と B は無関係’ と同値であることが判る.

定理 2.3.3. 事象 A と B が独立. $\Leftrightarrow \mathbf{P}[A/B] = \mathbf{P}[A]$. ◇

[証明] まず $\Rightarrow$ を示す. ベイズの公式 (2.3.1) と 独立の定義 (2.3.2) より

$$\mathbf{P}[A/B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} = \frac{\mathbf{P}[A] \cdot \mathbf{P}[B]}{\mathbf{P}[B]} = \mathbf{P}[A].$$

次に $\Leftarrow$ を示す. 定理の仮定と ベイズの公式 (2.3.1) より

$$\frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} = \mathbf{P}[A/B] = \mathbf{P}[A]$$

となる. 両辺に $\mathbf{P}[B]$ を乗じて, ‘独立’ の定義 (2.3.2) が得られる. □

3 個以上の事象にたいしては, 独立 を定義すること自体が煩雑な作業である.

定義 2.3.4. (i) ‘3 個の事象 A, B, C が互いに独立’ とは, 次の四つの等式が成立することである:

$$(2.3.3a) \quad \begin{aligned} \mathbf{P}[A \cap B] &= \mathbf{P}[A] \cdot \mathbf{P}[B], & \mathbf{P}[B \cap C] &= \mathbf{P}[B] \cdot \mathbf{P}[C], \\ \mathbf{P}[C \cap A] &= \mathbf{P}[C] \cdot \mathbf{P}[A], \end{aligned}$$

$$(2.3.3b) \quad \mathbf{P}[A \cap B \cap C] = \mathbf{P}[A] \cdot \mathbf{P}[B] \cdot \mathbf{P}[C].$$

(ii) (興味を持つ人のために) ‘ k 個の事象 $A_1, A_2, \dots, A_k$ が互いに独立’ とは, $A_1, A_2, \dots, A_k$ から取り出した任意の j 個の事象 $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ にたいし, 次の等式が成立することである:

$$\mathbf{P}[A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}] = \mathbf{P}[A_{i_1}] \cdot \mathbf{P}[A_{i_2}] \cdot \dots \cdot \mathbf{P}[A_{i_k}]. \quad \diamond$$

注意 2.3.5. (i) 事象 A, B, C にたいし (2.3.3a) が成立していても (2.3.3b) が成立するとは限らない.

反例: 4 枚のカード a, b, c, d があり, 「 a には 2」, 「 b には 3」, 「 c には 5」, 「 d には 30」と数字が書かれている. いまランダムにカードを選ぶとし, 事象 A, B, C を

事象 A : カードの数字が 2 で割り切れる,
 事象 B : カードの数字が 3 で割り切れる,
 事象 C : カードの数字が 5 で割り切れる,

*3 T. Bayes は英国長老派教会の牧師でアマチュア数学者. ベイズの公式 (2.3.1) は彼の死後に発表された. 条件付き確率 $\mathbf{P}[A/B]$ は ‘事後確率’ とも呼ばれる.

とする. このとき (2.3.3a) は成立しているが, (2.3.3b) は不成立である.

[反例の証明] カード a, b, c, d が選ばれる確率は各々 $1/4$ なので,

$$\mathbf{P}[A \cap B] = \mathbf{P}[\text{カード } d \text{ を選ぶ}] = \frac{1}{4}$$

だが

$$\mathbf{P}[A] = \mathbf{P}[\text{カード } a \text{ または } d \text{ を選ぶ}] = \frac{1}{2}, \quad \mathbf{P}[B] = \mathbf{P}[\text{カード } b \text{ または } d \text{ を選ぶ}] = \frac{1}{2}$$

となり, (2.3.3a) の第一式が成立している. 第二, 第三式も同様の計算で成立が示される. 一方,

$$\mathbf{P}[A \cap B \cap C] = \mathbf{P}[\text{カード } d \text{ を選ぶ}] = \frac{1}{4},$$

$$\mathbf{P}[A] \cdot \mathbf{P}[B] \cdot \mathbf{P}[C] = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

だから, (2.3.3b) は不成立である.

(ii) また逆に, 事象 A, B, C にたいし (2.3.3b) が成立していても (2.3.3a) が成立するとは限らない. (反例は 練習問題 2.5.4 を見よ.) $\diamond$

2.4 ベイズの公式の応用例

近年, 条件付き確率とベイズの公式 (2.3.1) がマーケティング理論や人工知能などに適用され, 確率論を社会現象へ応用する強力なツールとして, 注目を集めている. これは

‘条件付き確率 $\mathbf{P}[A/H]$ ’ と ‘ H の起こる確率 $\mathbf{P}[H]$ ’ がデータとして得られる

場合, ‘未知の確率 $\mathbf{P}[A \cap H]$ ’ がベイズの公式 (2.3.1) から

$$(2.4.1) \quad \mathbf{P}[A \cap H] = \mathbf{P}[H] \cdot \mathbf{P}[A/H]$$

として計算出来るからである.

2.4.1 ベイズの定理

定理 2.4.1 (ベイズの定理). $\mathbf{P}[H] \neq 0$, $\mathbf{P}[A] \neq 0$ である事象 H, A にたいし, 次の等式が成立する:

$$(2.4.2) \quad \mathbf{P}[H/A] = \frac{\mathbf{P}[H] \cdot \mathbf{P}[A/H]}{\mathbf{P}[H] \cdot \mathbf{P}[A/H] + \mathbf{P}[H^c] \cdot \mathbf{P}[A/H^c]}.$$

ただし, $\mathbf{P}[H^c] = 0$ の場合, $\mathbf{P}[H^c] \cdot \mathbf{P}[A/H^c] = 0$ と約束する. $\diamond$

[証明] (2.4.1) という形でベイズの公式を使う. $\mathbf{P}[A \cap H] + \mathbf{P}[A \cap H^c] = \mathbf{P}[A]$ になることに注意して,

$$(2.4.2) \text{ の右辺} = \frac{\mathbf{P}[A \cap H]}{\mathbf{P}[A \cap H] + \mathbf{P}[A \cap H^c]} = \frac{\mathbf{P}[A \cap H]}{\mathbf{P}[A]} = (2.4.2) \text{ の左辺. } \square$$

系 2.4.2 (ベイズの定理 2). 事象 $H_1, \dots, H_m$ を

$$H_1 \cup H_2 \cup \dots \cup H_m = \Omega; \quad i \neq j \text{ なら } H_i \cap H_j = \emptyset$$

すべての $k = 1, \dots, m$ にたいし $\mathbf{P}[H_k] \neq 0$,

とする. このとき $\mathbf{P}[A] \neq 0$ である 事象 A にたいし, 以下の等式が成立する:

$$(2.4.3) \quad \mathbf{P}[H_1/A] = \frac{\mathbf{P}[H_1] \cdot \mathbf{P}[A/H_1]}{\mathbf{P}[H_1] \cdot \mathbf{P}[A/H_1] + \dots + \mathbf{P}[H_m] \cdot \mathbf{P}[A/H_m]}. \quad \diamond$$

下は ‘得られたデータ (白玉が出た)’ から ‘途中経過 (どちらの箱から玉を出したか)’ を推測する例である.

例題 2.4.3. 箱 a には 白玉 8 個 と 黒玉 2 個, 箱 b には 白玉 20 個 と 黒玉 20 個が入っている. ランダムに箱を選んで 1 個の玉を取り出したら 白 であった. この白玉は 箱 a と 箱 b のどちらから取り出されたか? 確率の多い箱を決定せよ.

[解答] 事象 A, B, W を次の通りとする.

$$A = \text{箱 } a \text{ から玉を取り出す}, \quad B = \text{箱 } b \text{ から玉を取り出す}, \\ W = \text{箱から取り出した玉が白}.$$

計算すべき確率は $\mathbf{P}[A/W], \mathbf{P}[B/W]$ である. ベイズの公式 (2.3.1) より

$$(2.4.4) \quad \mathbf{P}[A/W] = \frac{\mathbf{P}[A \cap W]}{\mathbf{P}[W]}, \quad \mathbf{P}[B/W] = \frac{\mathbf{P}[B \cap W]}{\mathbf{P}[W]}$$

となる. 分母, 分子の各項を計算しよう. どちらの箱も選ぶ確率は $1/2$ だから, ベイズの公式を (2.4.1) のように変形すると

$$\mathbf{P}[A \cap W] = \mathbf{P}[A] \cdot \mathbf{P}[W/A] = \frac{1}{2} \cdot \frac{8}{10} = \frac{8}{20}, \\ \mathbf{P}[B \cap W] = \mathbf{P}[B] \cdot \mathbf{P}[W/B] = \frac{1}{2} \cdot \frac{20}{40} = \frac{5}{20}.$$

ここで白玉は必ず 箱 a もしくは 箱 b から取り出すので, $W = (A \cap W) \cup (B \cap W)$. しかも $(A \cap W) \cap (B \cap W) = \emptyset$ だから, 確率の性質 (2.1.1b) より

$$\mathbf{P}[W] = \mathbf{P}[(A \cap W) \cup (B \cap W)] = \mathbf{P}[A \cap W] + \mathbf{P}[B \cap W] = \frac{8}{20} + \frac{5}{20} = \frac{13}{20}.$$

以上の計算結果を (2.4.4) に代入して,

$$\mathbf{P}[A/W] = \frac{8/20}{13/20} = \frac{8}{13} \simeq 0.62, \quad \mathbf{P}[B/W] = \frac{5/20}{13/20} = \frac{5}{13} \simeq 0.38.$$

よって, 取り出された白玉は 箱 a から出た可能性が高い. $\square$

2.4.2 ベイズの定理の応用例

条件付き確率とベイズの定理は, 数学クイズやパラドクスにも使われている.

例題 2.4.4 (王に姉妹はいるか?). 先王には子供が二人おり, その内の一人が新たに王として即位した. 新王に姉妹がいる確率を求めよ. ただし, 男女の比率は $1:1$ とする. $\diamond$

[解答] “ $B \equiv$ 子供が男, $G \equiv$ 子供が女” という記号を使う. 左の記号が年長者を表すとする, 起こりえる事象は

$$w_1 \equiv (B, B), \quad w_2 \equiv (B, G), \quad w_3 \equiv (G, B), \quad w_4 \equiv (G, G)$$

であるから, 確率空間は

$$(2.4.5) \quad \Omega \equiv \{w_1, w_2, w_3, w_4\}, \quad \mathbf{P}[w_k] = \frac{1}{4}, \quad k = 1, \dots, 4$$

となる.

$$U \equiv \text{子供の一人が男} = w_1 \cup w_2 \cup w_3,$$

$$V \equiv \text{子供の一人が女} = w_2 \cup w_3 \cup w_4$$

という記号を使う. (2.4.5) を考慮して

$$\mathbf{P}[\text{王に姉妹がいる}] = \mathbf{P}[V/U] = \frac{\mathbf{P}[V \cap U]}{\mathbf{P}[U]} = \frac{\mathbf{P}[w_2 \cup w_3]}{\mathbf{P}[w_2 \cup w_3 \cup w_4]} = \frac{2/4}{3/4} = \frac{2}{3}. \quad \square$$

例題 2.4.5 (囚人のパラドクス). 3 人の囚人 a, b, c がおり, ‘そのうちの一人が死刑, 残りの二人は釈放’ との決定がなされた. 囚人 a が看守に ‘ b, c のどちらが釈放か教えてくれ’ と頼んだ. ところが看守は

(2.4.6a) 今のままならば a が死刑になる確率は $1/3$ である.

(2.4.6b) b, c のどちらが釈放か教えると a が死刑になる確率は $1/2$ になる.

の理由により, ‘気の毒だから教えない’ と答えた. いま

(2.4.7a) 死刑になる確率は a, b, c それぞれ $1/3$,

(2.4.7b) 死刑が a の場合, 看守は $1/2$ の確率で ‘ b は保釈’ という,

(2.4.7c) 死刑が c の場合, 看守は必ず ‘ b は保釈’ という,

(2.4.7d) 看守は嘘をつかない,

としたとき, 看守の主張する確率計算 (2.4.6b) は正しいのか?

[解答] 事象 $A, \dots, D$ を以下の通り定める:

$$A \equiv \{a \text{ が死刑}\}, \quad B \equiv \{b \text{ が死刑}\}, \quad C \equiv \{c \text{ が死刑}\},$$

$$D \equiv \{\text{看守が ‘} b \text{ は保釈’ という}\}$$

看守が主張した (2.4.6b) は

$$\mathbf{P}[A/D] = \frac{1}{2}$$

と同値である. 果たしてこれが正しいか確かめよう.

まず 確率データを整理する. ‘一人が死刑, 残りの二人は釈放’ だから, Ω を全事象とすると

$$A \cup B \cup C = \Omega.$$

また (2.4.7a), (2.4.7b), (2.4.7c) より

$$\mathbf{P}[A] = \mathbf{P}[B] = \mathbf{P}[C] = \frac{1}{3}, \quad \mathbf{P}[D/A] = \frac{1}{2}, \quad \mathbf{P}[D/C] = 1.$$

以上と ベイズの公式 (2.3.1) より

$$\begin{aligned} \mathbf{P}[D \cap A] &= \mathbf{P}[D/A] \cdot \mathbf{P}[A] = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}, \\ \mathbf{P}[D \cap C] &= \mathbf{P}[D/C] \cdot \mathbf{P}[C] = 1 \cdot \frac{1}{3} = \frac{1}{3}. \end{aligned}$$

$A \cup B \cup C = \Omega$ だから, (2.4.7d) も考えると

$$\mathbf{P}[D] = \mathbf{P}[D \cap A] + \mathbf{P}[D \cap B] + \mathbf{P}[D \cap C] = \frac{1}{6} + 0 + \frac{1}{3} = \frac{1}{2}$$

つまり $\mathbf{P}[A/D] = \frac{\mathbf{P}[D \cap A]}{\mathbf{P}[D]} = \frac{1/6}{1/2} = \frac{1}{3}$ となり, 看守の主張 (2.4.6b) は間違っている. $\square$

日本女性の約 1~5 % は乳癌を罹患すると言われている. ところが, 早期発見により多くの乳癌は治癒するので, 「全女性が乳癌の検査を受ける」ことが推奨されている. 果たして, この政策が正しいか否かを検証しよう.

例題 2.4.6 (医学検査の有効性). ‘乳癌の検査を全女性に行う’ことは合理的か? $\diamond$

[解答] *Step 1.* 医学検査は完全ではなく, 常にエラーを伴う. 2007 年の時点で, 乳癌の検診でエラーの起こる確率は概ね

$$(2.4.8a) \quad \mathbf{P}[\text{検査結果が白/実際には乳癌が有る}] = 0.15,$$

$$(2.4.8b) \quad \mathbf{P}[\text{検査結果が黒/実際には乳癌が無い}] = 0.15,$$

である.

ここで, 事象 A, B を

$$(2.4.9) \quad A \equiv \{ \text{乳癌が有る} \}, \quad B \equiv \{ \text{検査結果が黒} \}$$

とおくと, (2.4.8a), (2.4.8b) は

$$(2.4.10) \quad \mathbf{P}[B^c/A] = 0.15, \quad \mathbf{P}[B/A^c] = 0.15$$

と表すことができる.

Step 2. ベイズの公式 (2.3.1) から

$$\begin{aligned} \mathbf{P}[B/A] + \mathbf{P}[B^c/A] &= \frac{\mathbf{P}[B \cap A]}{\mathbf{P}[A]} + \frac{\mathbf{P}[B^c \cap A]}{\mathbf{P}[A]} = \frac{\mathbf{P}[B \cap A] + \mathbf{P}[B^c \cap A]}{\mathbf{P}[A]} \\ &= \frac{\mathbf{P}[A]}{\mathbf{P}[A]} = 1. \end{aligned}$$

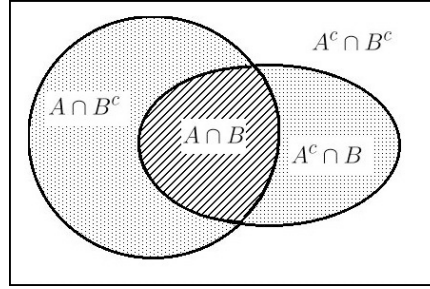


図 2.4.1 検査/乳癌による全女性の分類

これより

$$(2.4.11) \quad \mathbf{P}[B/A] = 1 - \mathbf{P}[B^c/A] = 1 - 0.15 = 0.85.$$

Step 3. 次に、日本女性の乳癌罹患率を x とする ($0 < x < 1$):

$$\mathbf{P}[A] = x.$$

これと Step 2 の計算から、 $\mathbf{P}[B]$ を求めよう. ベイズの公式 (2.3.1) を使うと、

$$\begin{aligned} 0.85 &= \mathbf{P}[B/A] = \frac{\mathbf{P}[B \cap A]}{\mathbf{P}[A]} = \frac{\mathbf{P}[B \cap A]}{x} \Rightarrow \mathbf{P}[B \cap A] = 0.85x. \\ 0.15 &= \mathbf{P}[B/A^c] = \frac{\mathbf{P}[B \cap A^c]}{\mathbf{P}[A^c]} = \frac{\mathbf{P}[B \cap A^c]}{1-x} \Rightarrow \mathbf{P}[B \cap A^c] = 0.15(1-x). \end{aligned}$$

つまり

$$\mathbf{P}[B] = \mathbf{P}[B \cap A] + \mathbf{P}[B \cap A^c] = 0.85x + 0.15(1-x) = 0.15 + 0.7x.$$

Step 4. 検査の有効性 $Q(x)$ は、

$$Q(x) = \mathbf{P}[A/B] = \text{検査結果が黒の中で本当の乳癌罹患者がいる確率}$$

だから、その値を計算する:

$$Q(x) = \mathbf{P}[A/B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} = \frac{0.85x}{0.15 + 0.7x} = \frac{17x}{3 + 14x}.$$

x にいろいろな値を入れて、 $Q(x)$ がどうなるを調べる:

x	1%	2%	3%	4%	5%	10%	15%	20%	25%
$Q(x)$	5.4%	10.4%	14.9%	19.1%	22.9%	38.6%	50%	58.6%	65.3%

上の計算結果より次の結論が得られる. :

- (i) 乳癌の罹患率 $x \leq 4\%$ の場合、検査の有効性 $Q(x) \leq 20\%$ となり、効果/費用の観点から有効とはいえない. また ‘誤診が 80 % 以上ある’ ことを被験者に告知すべきである.
- (ii) 罹患率 $x \geq 15\%$ の場合、 $Q(x) \geq 50\%$ となり、効果/費用の観点から検査は有効である. ただしこの場合も、‘誤診が 50 % 程度ある’ ことを被験者に告知すべきである. □

電子メールの普及に伴い、多数送りつけられる迷惑メールの防止策が必要となった。防止には「迷惑メール (Spam と呼ぶ) と通常のメール (Ham と呼ぶ) を分別するフィルター」を設置しなければならない。このフィルターの作動原理には「ベイズの定理」が使われることが多く、それらのフィルターは「ベ이지アン・フィルター」と呼ばれている。

例題 2.4.7 (ベ이지アン・フィルター). 迷惑メールの防止フィルターを、件名に「援助交際」という言葉が含まれているか否かで設定する。事象 H, A を

$$\begin{aligned} H &= \{ \text{迷惑メールである} \}, \\ A &= \{ \text{文中に「援助交際」の言葉が含まれる} \}, \end{aligned}$$

とおく。過去の統計データによれば

$$(2.4.12) \quad \mathbf{P}[A/H] = 0.3, \quad \mathbf{P}[A/H^c] = 0.01, \quad \frac{\mathbf{P}[H^c]}{\mathbf{P}[H]} = 2$$

であった。この迷惑メールの判定フィルターの有効確率 $\mathbf{P}[H/A]$ を求めよ。

[解答] ベイズの定理 から

$$\begin{aligned} \mathbf{P}[H/A] &= \frac{\mathbf{P}[H] \cdot \mathbf{P}[A/H]}{\mathbf{P}[H] \cdot \mathbf{P}[A/H] + \mathbf{P}[H^c] \cdot \mathbf{P}[A/H^c]} \\ &= \frac{\mathbf{P}[A/H]}{\mathbf{P}[H/A] + (\mathbf{P}[H^c] / \mathbf{P}[H]) \cdot \mathbf{P}[A/H^c]}. \end{aligned}$$

この右辺に統計データを代入して

$$\mathbf{P}[H/A] = \frac{0.3}{0.3 + 2 \times 0.01} \simeq 0.94.$$

なお、統計データ (2.4.12) を時間と共に更新することで、時間経過で進化する「学習型のベ이지アン・フィルター」が得られる。□

2.5 練習問題

練習問題 2.5.1. 例題 2.2.2 の確率空間 $(\Omega, \mathbf{P})$ で、

$$A \equiv \text{サイコロの目が奇数} = \{w_1, w_3, w_5\}, \quad B \equiv \text{サイコロの目が4以上} = \{w_4, w_5, w_6\}$$

とおく。次の確率を計算せよ。

$$(i) \mathbf{P}[A \cap B], \quad (ii) \mathbf{P}[A \cup B], \quad (iii) \mathbf{P}[A] + \mathbf{P}[B] - \mathbf{P}[A \cap B]. \quad \sharp$$

練習問題 2.5.2. $(\Omega, \mathbf{P})$ を 例題 2.2.2 の確率空間とする。

- (i) 事象 $A \equiv$ “偶数の目がでる” を $w_1, \dots, w_6$ を使って表せ。
- (ii) 条件付き確率 $\mathbf{P}[w_2/A]$ を計算せよ。
- (iii) 事象 A と w_2 は独立か?

練習問題 2.5.3. $(\Omega, \mathbf{P})$ を 例題 2.2.3 の確率空間とする。

- (i) 事象 $A \equiv$ “1 回目のコイン投げが表” を $w_1, \dots, w_4$ を使って表せ。

(ii) 事象 $B \equiv$ “2 回目のコイン投げが表” を $w_1, \dots, w_4$ を使って表せ.

(iii) 事象 A と B は独立か?

練習問題 2.5.4. 24 枚のカードがあり, 各カードには数字が一つ書かれている. ‘書かれた数字’ と ‘その数字が書いてあるカードの枚数’ は以下の通り:

カードに書いてある数字	カードの枚数
2, 3, 5	それぞれ 5 枚ずつ
6, 10, 15	それぞれ 2 枚ずつ
30	3 枚

いまランダムにカードを選ぶとし, 事象 A, B, C を

事象 A : カードの数字が 2 で割り切れる,

事象 B : カードの数字が 3 で割り切れる,

事象 C : カードの数字が 5 で割り切れる,

とする.

(i) (2.3.3b) が成立していることを示せ.

(ii) (2.3.3a) は成立しないことを示せ.

練習問題 2.5.5. 6 枚は赤, 4 枚は黒の印を裏側に付けたカードがある. その 10 枚のカードを表側を上にして並べ, その中からランダムに選んだ 3 枚のカードを順に裏返す.

$A \equiv$ {最初に裏返したカードが赤印}, $B \equiv$ {2 番目に裏返したカードが黒印},

$C \equiv$ {3 番目に裏返したカードが赤印},

とおくとき, 次の確率を計算せよ.

(i) $\mathbf{P}[A \cap B]$, (ii) $\mathbf{P}[B]$, (iii) $\mathbf{P}[B \cap C]$, (iv) $\mathbf{P}[C]$.

練習問題 2.5.6 (2006 年公認会計士-試験問題). C 社の配当は, 海外親企業の業績に依存して, 以下の表の各行の確率で金額が決まる.

		配当金			計
		1 円	5 円	10 円	
親企業の業績	最高 (E_1)	0.1	0.2	0.7	1.0
	好調 (E_2)	0.2	0.6	0.2	1.0
	停滞 (E_3)	0.6	0.3	0.1	1.0

例えば業績が最高のとき配当が 5 円である確率は $\mathbf{P}[5 \text{ 円}/E_1] = 0.2$, 業績が停滞のとき配当が 10 円である確率は $\mathbf{P}[10 \text{ 円}/E_3] = 0.1$ である. また業界情報に依れば, 親企業の業績に関して以下のような観測がなされている.

$$\mathbf{P}[E_1] = 0.2, \quad \mathbf{P}[E_2] = 0.3, \quad \mathbf{P}[E_3] = 0.5.$$

C 社の配当が 10 円だったとき, 親企業の業績が停滞であった確率 $\mathbf{P}[E_3/10 \text{ 円}]$ はいくらか, 求めなさい. なお計算過程も示しなさい.

第 3 章

確率変数と平均

数学の応用は, ‘関数’ という概念が導入されてから大幅に広がった. つまり関数とは, ある要因とその結果との関係を厳密に記述することだからである.

確率論では, 各事象ごとにどのような結果が起こるかを厳密に定義したものが ‘関数’ に当たる. ただ歴史的経緯から, この関数は確率変数と呼ばれている.

3.1 確率変数と確率分布

確率空間 $(\Omega, \mathbf{P})$ を固定する.

定義 3.1.1. ‘全ての事象 $w \in \Omega$ にたいし実数を与える関数’ $X : \Omega \rightarrow \mathbf{R}$ を **確率変数 random variable** と呼ぶ. $\diamond$

定義 3.1.2. 2つの確率変数 $X(w), Y(w)$ があり,

$$(3.1.1) \quad \begin{aligned} X(w) &\text{ は } a_1, a_2, \dots, a_n \text{ の値,} \\ Y(w) &\text{ は } b_1, b_2, \dots, b_m \text{ の値} \end{aligned}$$

をとる.

(i) (3.1.1) である確率変数 $X(w)$ にたいし,

$$(3.1.2) \quad p_1 \equiv \mathbf{P}[X(w) = a_1], \quad p_2 \equiv \mathbf{P}[X(w) = a_2], \dots, p_n \equiv \mathbf{P}[X(w) = a_n]$$

とおき, 等式の組 (3.1.2) を $X(w)$ の **確率分布** distribution of $X(w)$ と呼ぶ.

(ii) 以下の等式が成立するとき, $X(w)$ と $Y(w)$ が **独立** independent という:

$$(3.1.3) \quad \begin{aligned} &\text{すべての } 1 \leq j \leq n, \quad 1 \leq k \leq m \text{ にたいし} \\ &\mathbf{P}[X(w) = a_j, Y(w) = b_k] = \mathbf{P}[X(w) = a_j] \cdot \mathbf{P}[Y(w) = b_k]. \quad \diamond \end{aligned}$$

つぎの補題は, 補題 2.1.2 で述べた確率測度の性質から簡単に導かれる:

補題 3.1.3. 確率分布 (3.1.2) にたいし, 以下が成立:

$$\text{すべての } k = 1, \dots, n \text{ にたいし } p_k \geq 0; \quad \sum_{k=1}^n p_k = 1. \quad \diamond$$

例題 3.1.4. 例 2.2.1 の ‘1 回コインを投げる場合の確率空間 $(\Omega, \mathbf{P})$ ’ で 確率変数 $X(w)$ を次のように定義する:

$$(3.1.4) \quad X(w) \equiv \begin{cases} 1 & \text{コイン投げで表がでる} \\ -1 & \text{コイン投げで裏がでる} \end{cases}$$

すると,

$$\mathbf{P}[X(w) = 1] = \mathbf{P}[w_1] = p, \quad \mathbf{P}[X(w) = -1] = \mathbf{P}[w_2] = 1 - p$$

という確率分布が得られる.

例題 3.1.5. (i) 例 2.2.3 の ‘2 回コインを投げる場合の確率空間 $(\Omega, \mathbf{P})$ ’ で 確率変数 $X_1(w), X_2(w)$ を次のように定義する:

$$(3.1.5) \quad \begin{aligned} X_1(w) &\equiv \begin{cases} 1 & \text{1 回目のコイン投げが表} \\ -1 & \text{1 回目のコイン投げが裏} \end{cases} \\ X_2(w) &\equiv \begin{cases} 1 & \text{2 回目のコイン投げが表} \\ -1 & \text{2 回目のコイン投げが裏} \end{cases} \end{aligned}$$

すると,

$$\begin{aligned} \mathbf{P}[X_1(w) = 1] &= \mathbf{P}[w_1 \cup w_3] = \mathbf{P}[w_1] + \mathbf{P}[w_3] = p^2 + p(1-p) = p, \\ \mathbf{P}[X_1(w) = -1] &= \mathbf{P}[w_2 \cup w_4] = \mathbf{P}[w_2] + \mathbf{P}[w_4] \\ &= p(1-p) + (1-p)^2 = 1-p, \\ \mathbf{P}[X_2(w) = 1] &= \mathbf{P}[w_1 \cup w_2] = \mathbf{P}[w_1] + \mathbf{P}[w_2] = p^2 + p(1-p) = p, \\ \mathbf{P}[X_2(w) = -1] &= \mathbf{P}[w_3 \cup w_4] = \mathbf{P}[w_3] + \mathbf{P}[w_4] \\ &= p(1-p) + (1-p)^2 = 1-p \end{aligned}$$

との確率分布が得られる. $\diamond$

(ii) 上の確率分布から

$$\begin{aligned} &\mathbf{P}[X_1(w) = 1, X_2(w) = 1], \quad \mathbf{P}[X_1(w) = 1, X_2(w) = -1], \\ &\mathbf{P}[X_1(w) = -1, X_2(w) = 1], \quad \mathbf{P}[X_1(w) = -1, X_2(w) = -1], \end{aligned}$$

を計算すると, ‘ $X_1(w)$ と $X_2(w)$ が独立’ であることが判る. $\diamond$

練習問題 3.1.6. 上の計算を実行し, $X_1(w)$ と $X_2(w)$ が独立であることを確かめよ.

練習問題 3.1.7. 例 2.2.3 で述べた ‘2 回コインを投げる場合の確率空間 $(\Omega, \mathbf{P})$ ’ で

$$(3.1.6) \quad Z(w) \equiv \text{出た表の数} \times 100 \text{ 円}$$

だけ賞金がもらえるゲームを行う.

(i) この確率変数 $Z(w)$ を (3.1.5) の $X_1(w), X_2(w)$ を使って表せ.

(ii) 次の確率を計算せよ.

$$(a) \mathbf{P}[Z(w) = 200], \quad (b) \mathbf{P}[Z(w) = 100], \quad (c) \mathbf{P}[Z(w) = 0].$$

例題 3.1.8. つぎに, 例 2.2.4 の ‘ n 回コインを投げる 確率空間 $(\Omega, \mathbf{P})$ ’ で, 確率変数 $X_k(w), k = 1, 2, \dots, n$, を次で定義する.

$$(3.1.7) \quad X_k(w) \equiv \begin{cases} 1 & k \text{ 回目のコイン投げが表} \\ -1 & k \text{ 回目のコイン投げが裏} \end{cases}$$

この場合に $\mathbf{P}[X_k(w) = \pm 1]$ の確率を計算してみよう: (2.2.3) の表記法を使う. 数学的帰納法を使えば, $X_k(w)$ の確率分布が

$$\begin{aligned} \mathbf{P}[X_k(w) = 1] &= \mathbf{P}[\{w \in \Omega : w^{(k)} = 1\}] = \sum_{w \in \Omega, w^{(k)}=1} \mathbf{P}[w] = p \\ \mathbf{P}[X_k(w) = -1] &= \mathbf{P}[\{w \in \Omega : w^{(k)} = 0\}] = \sum_{w \in \Omega, w^{(k)}=0} \mathbf{P}[w] = 1 - p \end{aligned}$$

となることが簡単に証明できる.

また確率変数 $X_1, X_2, \dots, X_n$ は互いに独立である. $\diamond$

定義 3.1.9 (ランダム・ウォーク). (3.1.7) の確率変数 $X_k(w), k = 1, 2, \dots, n$ にたいし, 新しい確率変数を

$$(3.1.8) \quad S_k(w) \equiv \begin{cases} 0 & k = 0 \\ X_1(w) + \dots + X_k(w) & 1 \leq k \leq n \end{cases}$$

とおく (n 回コイン投げを行った場合の得失の累積になる).

この確率変数は**標準ランダム・ウォーク** random walk と呼ばれ, 応用上, 重要なものである. $\diamond$

命題 3.1.10. (3.1.8) の $S_k(w), k = 0, 1, \dots, n$, の確率分布は, 以下の通りである:

$$(3.1.9) \quad \mathbf{P}[S_k(w) = j] = {}_k C_{(k+j)/2} p^{(k+j)/2} (1-p)^{(k-j)/2},$$

ここで $-k \leq j \leq k$ かつ $k+j$ は偶数. $\diamond$

定義 3.1.11 (2 項分布). ((3.1.9) の右辺はもっとも基礎となる確率分布の一つで, **2 項分布** binomial distribution と呼ばれる. $\diamond$

3.2 平均とその応用

確率分布から確率変数を特徴づけるいろいろな数値が計算できる. 中でも ‘平均と分散’ が特に重要である.

定義 3.2.1. (i) $a_1, \dots, a_n$ の値をとる確率変数 $X(w)$ にたいして,

$$\mathbf{E}[X(w)] \equiv a_1 \mathbf{P}[X(w) = a_1] + a_2 \mathbf{P}[X(w) = a_2] + \dots + a_n \mathbf{P}[X(w) = a_n]$$

を ‘ $X(w)$ の**平均** mean^{*1} ’ と呼び, 上記の記号で表す. $\diamond$

*1 しばしば, **期待値** expectation とも呼ばれる.

例題 3.2.2. 例題 3.1.4 で定義された確率変数 $X(w)$ の平均を求める. (3.1.4) を使うと

$$\begin{aligned}\mathbf{E}[X(w)] &= 1 \cdot \mathbf{P}[X(w) = 1] + (-1) \cdot \mathbf{P}[X(w) = -1] \\ &= 1 \cdot p + (-1) \cdot (1 - p) = 2p - 1. \quad \diamond\end{aligned}$$

例題 3.2.3. 次の賭 G を行う:

G : サイコロを 1 回投げ, 出た目 $\times 100$ 円の賞金を獲得

このとき, 賭 G の参加料を幾らに設定すれば*2公平か? $\diamond$

[解答] 例題 2.2.2 の確率空間 $(\Omega, \mathbf{P})$ を考えると, 賭 G で得られる賞金の額は

$$w = w_k \text{ のとき } Z(w) = k \cdot 100, \quad k = 1, 2, \dots, 6$$

と表現できる. §5 で述べる ‘Kolmogorov の強大数の法則’ より

n を十分大きな数とする. 賭 G が n 回実施されたとき, 賞金の支払い総額は $n \cdot \mathbf{E}[Z(w)]$ に近づく

が判っている. これより ‘賭 G の参加料’ は

$$\begin{aligned}\mathbf{E}[Z(w)] &= 100 \cdot \mathbf{P}[Z(w) = 100] + \dots + 600 \cdot \mathbf{P}[Z(w) = 600] \\ &= 100 \cdot \frac{1}{6} + \dots + 600 \cdot \frac{1}{6} = \frac{2100}{6} = 350 \text{ 円}\end{aligned}$$

に設定すれば, 主催者/参加者の両者に対し公平である. $\square$

一般に, 平均の計算は易しくない. その計算を少しでも容易にするために, 次の補題がある.

補題 3.2.4. $X(w), Y(w)$ を確率変数, a, b を定数とする.

- (i) $\mathbf{E}[aX(w) + bY(w)] = a\mathbf{E}[X(w)] + b\mathbf{E}[Y(w)]$.
- (ii) 定数 a にたいしては, $\mathbf{E}[a] = a$.
- (iii) $X(w)$ と $Y(w)$ が独立なら $\mathbf{E}[X(w) \cdot Y(w)] = \mathbf{E}[X(w)] \cdot \mathbf{E}[Y(w)]$. $\diamond$

定義 3.2.5. 成功と失敗の 2 つの結果しかない試行を **ベルヌイ試行** と呼ぶ. $\diamond$

ベルヌイ試行は大変単純だが, 確率変数の基本であり広く応用されている. 実際, 例 2.2.4 で述べた ‘ n 回のコイン投げ’ もベルヌイ試行を n 回繰り返したものである. 他にも以下の応用例がある.

例題 3.2.6. 有る工場は注射薬を生産している. この注射薬は全製品の検査が必要だが, 検査費用が高額なので検査回数を減らしたい. どのような方法があるか?

[解答] 注射薬の生産数を N 個とする. まず

方法 A: 全製品から個別に検査用のサンプルを採取し, それを検査する つまり検査回数 $= N$.

*2 運営費用は考えない.

一方,

- 方法 B: (i) $N = k \cdot r$ とし, 全製品をそれぞれ k 個ずつからなる r 個のグループ $G_1, G_2, \dots, G_r$ に分ける.
(ii) G_1 に属する製品から検査用のサンプル k 個を採取し, その k 個を混ぜ合わせて, あたらしい検査用のサンプル S_1 を作る.
(iii) S_1 が検査に合格 $\Rightarrow G_1$ に属する製品は合格, つまり検査回数は 1.
 S_1 が不合格 $\Rightarrow G_1$ に属する全製品を再検査, 検査回数は都合 $1 + k$.
(iv) グループ $G_2, \dots, G_r$ についても上記 (ii)+(iii) の手順を繰り返す.

Step 1. 発見的的思考: 定性的な状況を把握するためには, 極端な例を考えることがしばしば有効である. 極端な二つの場合を考えてみよう.

Case 1: 全製品が合格するとき,

方法 A $\Rightarrow$ 検査回数 $= N$

方法 B $\Rightarrow$ 検査回数 $= k$

となり, “方法 A の回数 $>$ 方法 B の回数” で, 方法 B が優れている.

Case 2: 全製品が不合格なとき,

方法 A $\Rightarrow$ 回数 $= N$

方法 B $\Rightarrow$ 回数 $= r + r \cdot k = r + N$

となり, “方法 A の回数 $<$ 方法 B の回数” で, 方法 A が優れている.

Step 2. 不良品の発生率を q , $0 < q < 1$, とし, 前述の Case 1 と 2 のどちらが起こりやすいか計算しよう.

グループ G_1 の検査回数を $Y_1(w)$ とすると,

$$Y_1(w) = \begin{cases} 1 & \text{サンプル } S_1 \text{ が合格} \\ 1 + k & \text{サンプル } S_1 \text{ が不合格のとき} \end{cases}$$

ここで, サンプル S_1 が合格するには, k 個の製品がすべて良品でなければならないので

$$\text{サンプル } S_1 \text{ が合格する確率} = (1 - q)^k.$$

また

$$\text{サンプル } S_1 \text{ が不合格となる確率} = 1 - (1 - q)^k$$

である. 確率変数 $Y_1(w)$ は 1 か $1 + k$ の 2 つの値しかとらないベルヌイ試行で,

$$\mathbf{E}[Y_1(w)] = 1 \times (1 - q)^k + (1 + k) \times \{1 - (1 - q)^k\} = 1 + k - k(1 - q)^k.$$

Step 3. グループ G_2 の検査回数を $Y_2(w)$, $\dots$, グループ G_r の検査回数を $Y_r(w)$ とする. $Y_1(w), \dots, Y_r(w)$ はそれぞれ独立で同分布の確率変数なので,

$$R(q, k) \equiv \text{方法 B での平均検査回数}$$

とおくと

$$\begin{aligned} R(q, k) &= \mathbf{E}[Y_1(w)] + \cdots + \mathbf{E}[Y_r(w)] = r \{1 + k - k(1 - q)^k\} \\ &= r + r \cdot k - r \cdot k(1 - q)^k = \frac{N}{k} + N - N(1 - q)^k = N \left(\frac{1}{k} + 1 - (1 - q)^k \right). \end{aligned}$$

つまり

$$\frac{\text{方法 B での回数}}{\text{方法 A での平均回数}} = \frac{R(q, k)}{N} = \left\{ \frac{1}{k} + 1 - (1 - q)^k \right\}$$

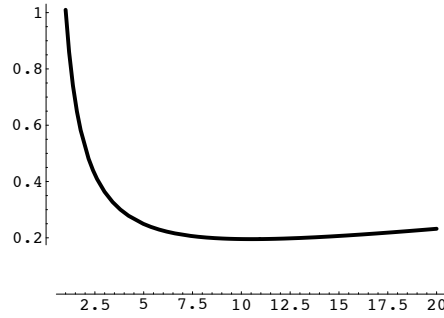
となるが, この比が 1 以下かどうかを調べればよい.

Step 4.

不良品の発生率 $q = 1/100$ として,

$$(3.2.1) \quad Q(k) \equiv \frac{1}{k} + 1 - (1 - q)^k$$

を数値計算すると, 下図の通りである.



それぞれの値は

k	2	5	8	10	12	15	20
$Q(k)$	0.519	0.249	0.202	0.1956	0.197	0.207	0.232

で, $Q(11) = 0.19557$ が最小値である.

結論: 不良品発生率 $q = 1/100$ の場合, “ $2 \leq k \leq 30$ とした検査方法 B” は “検査方法 A” より少ない検査回数で済む.

とくに $k = 11$ のとき, “検査方法 B” での検査回数は “検査方法 A” の 20% ですむ.

Step 5. [興味を持つ人のために] 数値計算によらずに (3.2.1) で与えられた $Q(k)$ の性質を調べよう.

$$0 = \frac{d}{dk} Q(k) = -\frac{1}{k^2} - (1 - q)^k \log(1 - q)$$

となる k が $Q(k)$ の極値だから, 上式をみたす k を求めればよい. ここで, 不良品発生率 q が小さな数で有ることを利用し, 基礎数学/解析学で学ぶ「テイラー展開」を使うと

$$(1 - q)^k \simeq 1 - kq, \quad \log(1 - q) \simeq -q$$

となる. これらを $\frac{d}{dk} Q(k)$ の右辺に代入し, q^2 以上の項を無視すると

$$0 = \frac{dQ(k)}{dk} = -\frac{1}{k^2} - (1 - kq)(-q) \simeq -\frac{1}{k^2} + q \Leftrightarrow k = \frac{1}{\sqrt{q}}$$

この値を

$$Q(k) = \frac{1}{k} + 1 - (1 - q)^k \simeq \frac{1}{k} + 1 - (1 - kq) = \frac{1}{k} + kq$$

に代入して, $Q(1/\sqrt{q}) \simeq 2\sqrt{q}$ が最小値であることが判る.

実際 $q = 1/100$ なら, $k \sim 1/10$ 近辺で最小値 0.2 をとる. $\square$

例題 3.2.7. 成功確率 p のベルヌイ試行を成功するまで繰り返す.

- (i) “ k 回以内に成功する確率 $F(k, p)$ ” を計算せよ.
 (ii) “成功するまでに要する回数 $Y(w)$ ” の平均を求めよ. $\diamond$

[解答] *Step 1.* $q \equiv 1 - p$ とおくと,

$$\begin{aligned} & \text{1 回目の成功する確率: } \mathbf{P}[Y(w) = 1] = p, \\ & \text{2 回目に成功する確率: } \mathbf{P}[Y(w) = 2] = qp, \\ & \vdots \\ (3.2.2) \quad & \text{\(k\) 回目に成功する確率: } \mathbf{P}[Y(w) = k] = q^{k-1}p, \\ & \vdots \end{aligned}$$

となるので,

$$F(k, p) = p + qp + q^2p + \cdots + q^{k-1}p = p \cdot \frac{1 - q^k}{1 - q} = 1 - (1 - p)^k.$$

ここで, いくつかの p にたいし $F(k, p)$ の値を計算してみる:

成功確率 p	0.1	0.2	0.3	0.5
3 回以内に成功する確率 $F(3, p)$	0.27	0.49	0.66	0.88
5 回以内に成功する確率 $F(5, p)$	0.41	0.67	0.83	0.97
7 回以内に成功する確率 $F(7, p)$	0.52	0.79	0.92	0.99
10 回以内に成功する確率 $F(10, p)$	0.65	0.89	0.97	~ 1

教訓: 同じ試行を 10 回繰り返して, 一度も成功しない.

$\Leftrightarrow$ 成功確率が 10% 以下の試行を繰り返している. 別の試行に切り替えろ.

Step 2. $Y(w)$ の平均値を求める. (3.2.2) と平均値の定義より,

$$\begin{aligned} \mathbf{E}[Y(w)] &= 1 \cdot p + 2 \cdot qp + 3 \cdot q^2p + \cdots + k \cdot q^{k-1}p + \cdots \\ &= p(1 + 2q + 3q^2 + \cdots + kq^{k-1} + \cdots) = \frac{p}{(1 - q)^2} = \frac{1}{p}. \end{aligned}$$

注: 次の級数の計算を確かめる事 (「基礎数学 I」 「基礎数学 II」):

$$1 + 2q + 3q^2 + \cdots + kq^{k-1} \cdots = \frac{1}{(1 - q)^2}. \quad \square$$

例題 3.2.8. ガムの景品にカードを付ける. カードは4種類あり, すべての種類を揃えると, 別の景品と交換できる. $Z(w)$ 個のガムを買うと, 4種類のカードがすべて揃ったとする. 確率変数 $Z(w)$ の平均をもとめよ. ◇

[解答] (i) まずガムを1個買うと, ある種類 a のカードが手に入る.

(ii) つぎにガムを買ったとき, a 以外のカードを手に入れる確率は $p = 3/4$. つまり“成功確率 $3/4$ のベルヌイ試行”を成功するまで繰り返すことになる. その試行の回数を $Y_1(w)$ とする.

(iii) 3種類目のカードを手に入れるには, “成功確率 $2/4$ のベルヌイ試行”を成功するまで繰り返すことになる. その回数を $Y_2(w)$ とする.

(iv) 4種類目のカードを手に入れるには, “成功確率 $1/4$ のベルヌイ試行”を成功するまで繰り返すことになる. その回数を $Y_3(w)$ とする.

(v) 結局 $1 + Y_1(w) + Y_2(w) + Y_3(w)$ 個のガムを買うと, 4種類すべてのカードが揃う. ここで例題 3.2.7 を使うと, $Y_k(w)$ の平均が計算できるので,

$$\begin{aligned} \mathbf{E}[Z(w)] &= \mathbf{E}[1 + Y_1(w) + \cdots + Y_3(w)] = 1 + \mathbf{E}[Y_1(w)] + \cdots + \mathbf{E}[Y_3(w)] \\ &= 1 + \frac{1}{3/4} + \frac{1}{2/4} + \frac{1}{1/4} = \frac{25}{3}. \end{aligned}$$

第 4 章

確率変数の分散と相関係数

確率変数の“散らばり具合”を特徴づける数値として、**分散**が重要である。また“二つの確率変数の独立性の度合い”を量る量として **相関係数**がある。

4.1 確率変数の分散

確率変数 $X(w)$ にたいし、平均 $m \equiv \mathbf{E}[X(w)]$ を定義 3.2.1 で与えた。

定義 4.1.1. $a_1, \dots, a_n$ の値をとる確率変数 $X(w)$ にたいして、

$$V[X(w)] \equiv \mathbf{E}[(X(w) - m)^2] = \sum_{k=0}^n (a_k - m)^2 \mathbf{P}[X(w) = a_k]$$

を“ $X(w)$ の**分散** variance”と呼ぶ。◇

注意 4.1.2. 次の 2 種類の確率変数 $X(w)$ と $Y(w)$ を考えてみよう。

$$\begin{aligned} \mathbf{P}[X(w) = 1] &= \frac{1}{2} = \mathbf{P}[X(w) = -1], \\ \mathbf{P}[Y(w) = 100] &= \frac{1}{2} = \mathbf{P}[Y(w) = -100]. \end{aligned}$$

どちらの確率変数も

$$\mathbf{E}[X(w)] = 1 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2} = 0, \quad \mathbf{E}[Y(w)] = 100 \cdot \frac{1}{2} - 100 \cdot \frac{1}{2} = 0,$$

と平均は 0 である。ところが分散を比べると

$$\begin{aligned} V[X(w)] &= \mathbf{E}[(X(w) - 0)^2] = 1^2 \cdot \frac{1}{2} + (-1)^2 \cdot \frac{1}{2} = 1 \\ V[Y(w)] &= \mathbf{E}[(Y(w) - 0)^2] = (100)^2 \cdot \frac{1}{2} + (-100)^2 \cdot \frac{1}{2} = 10000 \end{aligned}$$

となり大きく異なる。 $X(w)$ は出入り 1 円、 $Y(w)$ は出入り 100 円の公平な賭と考えると、分散は“賭の危険度を表す”と解釈できる*1。◇

*1 株式市場では、分散は“ボラタリティー”と呼ばれ、“相場変動の激しさ”をボラタリティーの大きさに判定している。

例題 4.1.3 (卵の運搬). 2つの卵を籠に入れて運搬する. 2種類の運搬方法

方法 A: 2つの卵を1つの籠に入れて, 運ぶ.

方法 B: 卵を1つずつ別の籠に入れ, 別々に運ぶ.

それぞれのリスクを計算せよ. ◇

[解答] 籠を落として卵を割る確率を q とし, 方法 A で届く卵の個数を $X(w)$, 方法 B での個数を $Y(w)$ とおく.

Step 1. $X(w), Y(w)$ それぞれの平均値を計算する.

$$\begin{aligned} E[X(w)] &= 2 \cdot (1-q) + 0 \cdot q = 2(1-q), \\ E[Y(w)] &= 2 \cdot (1-q)^2 + 1 \cdot (1-q)q + 1 \cdot q(1-q) + 0 \cdot q^2 = 2(1-q). \end{aligned}$$

つまり“無事に届く卵”の平均値は等しい.

Step 2. 次に分散を計算し, リスクを評価する. 次の補題 4.1.6, (iii) を使う:

$$\begin{aligned} V[X(w)] &= 2^2 \cdot (1-q) - (2(1-q))^2 = 4q(1-q), \\ V[Y(w)] &= 2^2 \cdot (1-q)^2 + 1^2 \cdot (1-q)q + 1^2 \cdot q(1-q) - (2(1-q))^2 = 2q(1-q) \\ &= \frac{V[X(w)]}{2} \end{aligned}$$

である.

つまり, 方法 B での運搬リスクは, 方法 A の $1/2$ と計算され, 実感と一致する.

注意 4.1.4. “すべての卵を一つの籠で運ぶ方法”がハイリスクであることは以下の事件で実証されている:

- (i) 1949 年, イタリアのサッカーリーグ, セリエ A で 4 連覇を果たした AC トリノの選手 18 人全員がポルトガルとの親善試合からの帰国中の飛行機事故で全員が命を落とした. 彼らはイタリア A 代表の主要メンバーでもあった.
- (ii) 1958 年, 英国プレミアリーグの強豪であるマンチェスターユナイテッドは, 欧州カップの準準決勝で勝利したが, 帰国途中の飛行機事故で 3 名を残して全員死亡. チーム再建に 10 年を要した.
- (iii) 1961 年, USA フィギュアスケート代表チームを乗せた飛行機がブリュッセルで墜落. 代表チーム全員を含む乗員乗客 73 名が全員死亡.

この例に見るとおり, 全員が 1 台の飛行機を利用する方法では, チーム全滅のリスクが高いので, “別々の籠で卵を運ぶ方法”は実際に行われている (=遠征に出かける選手を半数ずつに分け, 別々の飛行機を使う.) □

練習問題 4.1.5. 公平なサイコロを一回振り, 出た目により賞金を受け取る. 2種類の賭 X, Y の賞金は以下の通りである.

$$X \text{ の賞金} = \begin{cases} 50 \text{ 円} & \text{奇数の目} \\ 10 \text{ 円} & \text{偶数の目,} \end{cases} \quad Y \text{ の賞金} = \begin{cases} 60 \text{ 円} & 1 \text{ の目} \\ 40 \text{ 円} & 2 \text{ か } 3 \text{ の目} \\ 20 \text{ 円} & 4 \text{ か } 5 \text{ の目} \\ 0 \text{ 円} & 6 \text{ の目.} \end{cases}$$

- (i) 賭 X, Y の参加料をそれぞれ 30 円とする. X と Y は公平な賭か?
 (ii) X と Y ではどちらがよりハイリスク・ハイリターン型か?

一般に, 平均や分散の計算は易しくない. その計算を少しでも容易にするために, 次の補題がある.

補題 4.1.6. $X(w), Y(w)$ を確率変数, c, c' を定数とする.

- (i) $V[cX(w) + c'] = c^2 V[X(w)]$.
 (ii) $V[c] = 0$. 逆に分散 $V[X(w)] = 0$ となる 確率変数 $X(w)$ は定数である.
 (iii) $V[X(w)] = \mathbf{E}[(X(w))^2] - (\mathbf{E}[X(w)])^2$.
 (iv) $X(w)$ と $Y(w)$ が独立なら

$$V[X(w) + Y(w)] = V[X(w)] + V[Y(w)]. \quad \diamond$$

練習問題 4.1.7. (3.1.5) で与えられた 確率変数 $X_1(w)$ および $X_2(w)$ の平均と分散をもとめよ.

‡

練習問題 4.1.8. 二社の株式 X, Y があり, それぞれの分布は

$$\begin{aligned} \mathbf{P}[X = 2] &= \mathbf{P}[X = -2] = \frac{1}{2}, \\ \mathbf{P}[Y = 3] &= \mathbf{P}[Y = -3] = \frac{1}{6}, \quad \mathbf{P}[Y = 1] = \mathbf{P}[Y = -1] = \frac{1}{3}. \end{aligned}$$

どちらの株式投資がハイリスク・ハイリターン型か.

例題 4.1.9. 前出の 例題 3.2.7 で与えられた確率変数 $Y_k(w), k = 1, \dots, 4$, (ベルヌイ試行で成功するまでに要する回数) の分散を求めよ. $\diamond$

[解答] 例題 3.2.7 の解答から

$$\mathbf{E}[Y(w)] = \frac{1}{p}.$$

である.

つぎに 補題 4.1.6, (iii) を使うと,

$$\begin{aligned} V[Y(w)] &= \mathbf{E}[(Y(w))^2] - (\mathbf{E}[Y(w)])^2 \\ &= \left(1^2 \cdot p + 2^2 \cdot qp + 3^2 \cdot q^2 p + \dots + k^2 \cdot q^{k-1} p + \dots\right) - \left(\frac{1}{p}\right)^2 \\ &= p(1^2 + 2^2 q + 3^2 q^2 + \dots + k^2 q^{k-1} + \dots) - \frac{1}{p^2} \\ &= \frac{p(1+q)}{(1-q)^3} - \frac{1}{p^2} = \frac{q}{p^2} = \frac{1}{p} \left(\frac{1}{p} - 1\right). \end{aligned}$$

さらに 補題 4.1.6, (iv) を使うと

$$\begin{aligned} V[Z(w)] &= V[1 + Y_1(w) + \dots + Y_4(w)] \\ &= 1 + \frac{1}{3/4} \left(\frac{1}{3/4} - 1\right) + \frac{1}{2/4} \left(\frac{1}{2/4} - 1\right) + \frac{1}{1/4} \left(\frac{1}{1/4} - 1\right) = 19. \end{aligned}$$

$Z(w)$ の分散は大きいので、平均値 $25/3$ 個のガムでカード 4 枚を揃えられない者が多数いる。
□

注: 次の数列計算を確かめる事 (「基礎数学 I」「基礎数学 II」):

$$1^2 + 2^2 q + 3^2 q^2 + \cdots + k^2 q^{k-1} + \cdots = \frac{1+q}{(1-q)^3}. \quad \square$$

命題 4.1.10 (興味を持つ人に). 確率変数 $X_1(w), X_2(w), \dots, X_n(w)$ はそれぞれ独立である. 各 $X_k(w)$ の分散が存在し,

$$\mathbf{E}[X_k(w)] = m_k, \quad \mathbf{V}[X_k] = c_k, \quad k = 1, 2, \dots, n$$

とかく. このとき

$$\begin{aligned} \mathbf{E}[X_1(w) + X_2(w) + \cdots + X_n(w)] &= m_1 + m_2 + \cdots + m_n, \\ \mathbf{V}[X_1 + X_2 + \cdots + X_n] &= c_1 + c_2 + \cdots + c_n \end{aligned}$$

である. ◇

練習問題 4.1.11 (興味を持つ人に). 命題 4.1.10 を証明せよ.

例題 4.1.12 (ペテルスブルグ・パラドクス). 次のゲームを行う:

裏が出るまでコイントスを続ける. そして, j 回目のコイントスを行う度に, 2^j の賞金を受け取る.

このゲームは ‘ペテルスブルグ・ゲーム’ と呼ばれるが, このゲームで表が出続けた回数を Y とすると, 明らかに,

$$\mathbf{P}[Y(w) = j] = \frac{1}{2^j}, \quad j = 1, 2, \dots$$

受け取る賞金額 $X(w)$ は

$$(4.1.1) \quad \mathbf{E}[X(w)] = \sum_{j=1}^{\infty} 2^j \mathbf{P}[Y(w) = j] = \sum_{j=1}^{\infty} 1 = \infty$$

となる. ペテルスブルグ・ゲームが胴元と参加者双方に公平であるためには, ゲームの参加料 Q を無限大にするべきか?

[解答] *Step 1.* k 人目がペテルスブルグ・ゲームで受け取る賞金を X_k とする. すると $X_1, X_2, \dots$ は “独立同分布” であるが, (??) を満たさず, 後述の “大数の法則” (定理 ??) も成立しない. つまり参加料を平均値に設定する理由が無くなる.

Step 2. ではペテルスブルグ・ゲームで行うコイントスの回数に上限 N を設けてみる. すると, 受け取る賞金額 $X^{(N)}(w)$ の平均は

$$\mathbf{E}[X^{(N)}(w)] = \sum_{j=1}^N 2^j \mathbf{P}[Y(w) = j] = \sum_{j=1}^N 1 = N$$

で, 定理 (大数の法則) (??) が成立しているから, 参加料は N に設定すべきである.

つぎにペテルスブルグ・ゲームのリスクを計算してみよう. $X^{(N)}(w)$ の分散 $V[X^{(N)}(w)]$ は

$$V[X^{(N)}(w)] = \mathbf{E}[(X^{(N)}(w))^2] - N^2 = \sum_{j=1}^N (2^j)^2 \frac{1}{2^j} = 2^N - 1 - N^2.$$

であり,

N	5	10	15	20	25	30
$V[X^{(N)}(w)]$	6	923	3254	1.04×10^6	3.6^7	10^9

N が大きい場合は, ”極端にハイリスク・ハイリターン” 型のゲームとなる. $\square$

4.2 共分散と相関係数

2つの確率変数 $X(w)$ と $Y(w)$ が独立ということは, (3.1.3) で定義した. これより

$$X(w) \text{ と } Y(w) \text{ が独立} \Rightarrow \mathbf{E}[X(w) Y(w)] = \mathbf{E}[X(w)] \cdot \mathbf{E}[Y(w)]$$

である.

一方, 独立でない確率変数 $X(w)$ と $Y(w)$ の関係の程度を数値化する手段として, 相関係数がある.

定義 4.2.1. 平均がそれぞれ m_X, m_Y である確率変数 $X(w), Y(w)$ にたいし,

$$(4.2.1) \quad \begin{aligned} \text{Cov}[X, Y] &\equiv \mathbf{E}[(X(w) - m_X) \cdot (Y(w) - m_Y)], \\ \rho[X, Y] &\equiv \frac{\text{Cov}[X, Y]}{\sqrt{V[X]} \sqrt{V[Y]}} \end{aligned}$$

とおく. $\text{Cov}[X, Y]$ は **共分散** Covariance, $\rho[X, Y]$ は **相関係数** correlation coefficient と呼ばれる. $\diamond$

補題 4.2.2. $X(w), Y(w), Z(w)$ を確率変数, a, b を定数とする.

- (i) $\text{Cov}[X, X] = V[X]$.
- (ii) $\text{Cov}[aX + bY, Z] = a \text{Cov}[X, Z] + b \text{Cov}[Y, Z]$.
- (iii) $V[X(w) + Y(w)] = V[X(w)] + 2 \text{Cov}[X, Y] + V[Y(w)]$.
- (iv) $X(w)$ と $Y(w)$ が独立なら, $\text{Cov}[X, Y] = 0$. $\diamond$

補題 4.2.3. (i) 分散が存在するすべての確率変数 $X(w), Y(w)$ にたいし, $|\rho[X, Y]| \leq 1$.

(ii) $X(w)$ と $Y(w)$ が独立なら, $\rho[X, Y] = 0$. (ただし, $\rho[X, Y] = 0$ でも独立とは限らない*2.)

(iii) ある定数 a, b があり, $Y(w) = aX(w) + b$ ならば, $|\rho[X, Y]| = 1$. 逆に, $|\rho[X, Y]| = 1$ なら,

$$\frac{X(w) - \mathbf{E}[X(w)]}{\sqrt{V[X]}} = \frac{Y(w) - \mathbf{E}[Y(w)]}{\sqrt{V[Y]}}. \quad \diamond$$

*2 例題 4.2.4 を参照せよ.

たとえ共分散や相関係数が0であっても独立とは限らない。以下がその反例である。

例題 4.2.4 (反例). 確率変数 $X(w)$, $Y(w)$ を

$$\begin{aligned} \mathbf{P}[X(w) = 1] &= \frac{1}{4}, \quad \mathbf{P}[X(w) = 0] = \frac{1}{2}, \quad \mathbf{P}[X(w) = -1] = \frac{1}{4}, \\ Y(w) &\equiv (X(w))^2 \end{aligned}$$

とする。このとき

$$\begin{aligned} \mathbf{P}[X(w) = 1] \cdot \mathbf{P}[Y(w) = 1] &= \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}, \\ \mathbf{P}[X(w) = 1, Y(w) = 1] &= \mathbf{P}[X(w) = 1] = \frac{1}{4} \end{aligned}$$

だから, $\mathbf{P}[X(w) = 1] \cdot \mathbf{P}[Y(w) = 1] \neq \mathbf{P}[X(w) = 1, Y(w) = 1]$ となり, $X(w)$ と $Y(w)$ は独立ではない, cf. (3.1.3).

一方, $\mathbf{E}[X(w)] = 0$, $\mathbf{E}[Y(w)] = 1/2$ だから,

$$\begin{aligned} \text{Cov}[X, Y] &= \mathbf{E}\left[X(w) \cdot \left(Y(w) - \frac{1}{2}\right)\right] = \mathbf{E}[(X(w))^3] - \frac{1}{2} \cdot \mathbf{E}[X(w)] \\ &= 1^3 \cdot \frac{1}{4} + 0^3 \cdot \frac{1}{2} + (-1)^3 \cdot \frac{1}{4} - 0 = 0. \end{aligned}$$

また $V[X(w)] = 1/2$, $V[Y(w)] = 1/4$ となるので,

$$\rho[X, Y] = \frac{0}{\sqrt{1/2} \cdot \sqrt{1/4}} = 0$$

となり, $X(w)$ と $Y(w)$ の共分散, 相関係数ともに0である. $\diamond$

例題 4.2.5. A社の株 $X(w)$ と B社の株 $Y(w)$ がある。それぞれの間には,

$$V[X(w)] = 1, \quad V[Y(w)] = 2, \quad \text{Cov}[X, Y] = 0.4$$

の関係がある。A社の株を x , B社の株を $1-x$ 保持する時, x をどう決めれば最も安全か? $\diamond$

[解答] A社の株を x , B社の株を $1-x$ 保持したときの分散 $F(x)$ を計算する。補題 4.2.2 を使って,

$$\begin{aligned} F(x) &= V[xX(w) + (1-x)Y(w)] \\ &= V[xX(w)] + 2\text{Cov}[xX, (1-x)Y] + V[(1-x)Y(w)] \\ &= x^2 V[X(w)] + 2x(1-x) \text{Cov}[X, Y] + (1-x)^2 V[Y(w)] \\ &= x^2 + 0.8x(1-x) + 2(1-x)^2 \\ &= 2.2x^2 - 3.2x + 2 = 2.2\left(x - \frac{1.6}{2.2}\right)^2 + 2 - \frac{1.6^2}{2.2}. \end{aligned}$$

よって, $x = 1.6/2.2 = 8/11$ のときに分散が最小となり, もっとも安全. $\square$

練習問題 4.2.6. 確率変数 $X(w)$, $Y(w)$ が

$$\mathbf{E}[X(w)] = 0, \quad V[X(w)] = 1, \quad \mathbf{E}[Y(w)] = 1, \quad V[Y(w)] = 1, \quad \text{Cov}[X, Y] = 1$$

であるとする。このとき, 次を計算せよ。

- (i) $\mathbf{E}[2X(w) + Y(w)]$, (ii) $V[X(w) + Y(w)]$, (iii) $\rho[X, Y]$,
(iv) $\text{Cov}[X + Y, 3X + Y]$, (v) $\rho[X - 1, X + 2Y]$ ‡