2024年度九州大学　現代数学入門　公開講座
ランダムウォーク入門

角田　謙吉（大学院数理学研究院　准教授）

はじめに

このノートは 2024 年度九州大学現代数学入門公開講座のための予稿として作成したものです. 公開講座ではこのノートの内容に沿って説明を行う予定でありますが, ノートの内容全ての説明は行わず, 内容は取捨選択するつもりです. このノートを事前に読み込む必要はないので, 公開講座が終わった後に興味が湧いたところを読んでもらえればと思います.

大まかに, 以下の内容について説明する予定です.

1. 入試問題を例として, マルコフ連鎖・ランダムウォークの導入
2. 確率論入門
3. 大数の法則・ランダムウォークの性質

先にノートを読んで挫折しないように, 必要な数学の知識についても述べておきます. 1 では数列や二項間漸化式, それに加えて等比数列の極限を知っていれば十分です. 一方 2 と 3 では数式は多数登場するものの, 基本的に前提となる知識は何もいりませんが, 高校で学ぶ集合の記法については多々登場します. そのため一部数式の記述が難しく感じるかも知れませんが, 分からないところは適当に流してもらってかまいません. 興味のある読者のために一部の証明は 4 に書いています. 一部は大学で学ぶ数学を用いていますが, 意欲のある高校生ならばある程度読めるように証明をつけたつもりです.[1] 強調しておきたいことは何かというと, 分からなくても問題なく, 雰囲気を掴み取って興味をもってもらえれば幸いということです.

1　大学入試問題の例

本講座の題はランダムウォーク入門ですが, 入門とすべく大学入試の問題を例にマルコフ連鎖・ランダムウォークとよばれるものの導入をしたいと思います.

[1] 一部の証明は深く入りたくないので, 敢えてぼやかしているところもあります.

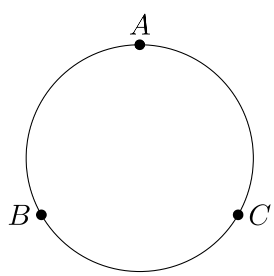
1.1 大阪大学の入試問題

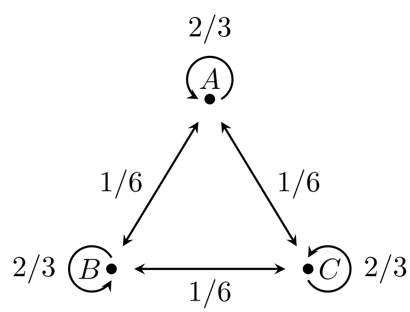
次は実際の入試問題で, 令和二年度の大阪大学の入試問題・数学 (文系) から引用しました.[2]

> 円周を 3 等分する点を時計回りに A, B, C とおく. 点 Q は A から出発し, A, B, C を以下のように移動する. 1 個のさいころを投げて, 1 の目が出た場合は時計回りに隣の点に移動し, 2 の目が出た場合は反時計回りに隣の点に移動し, その他の目が出た場合は移動しない. さいころを n 回投げたあとに Q が A に位置する確率を p_n とする. 以下の手順で p_n を求めよ.
>
> 1. p_1 を求めよ.
> 2. p_{n+1} を p_n を用いて表せ.
> 3. p_n を求めよ.

確率の問題ですが, 数列の一般項を求める問題であり大学入試で頻出する問題です. このような問題が大学入試に出題されるのは, 問題を解くための考えの理解度や習熟度を図る目的もありますが, 確率論におけるこのような問題の重要性も背景にあると思います. この問題を具体例として背後に潜む確率論の問題について説明したいと思います.

1.2 解答





図はそれぞれの点から別の点に移る確率を示したものです. Q は A から出発するので, さいころを 1 回投げたあとに A にいる確率 p_1 は 2/3 です. 次に, p_{n+1} を p_n で表すためには, 次の二通りに場合わけをする必要があります.

(a) n 回さいころを投げたあとに Q が A にいる

(b) n 回さいころを投げたあとに Q が B か C にいる

[2] 1 は「p_2 を求めよ」から「p_1 を求めよ」に変更しています.

(a) のときには次のさいころを投げたあと Q が A にいる確率は $2/3$ です. (b) のときには次のさいころを投げたあと Q が A にいる確率は, Q が B にいても C にいても, $1/6$ です. また (a) になるのは確率 p_n であり, (b) になるのは確率 $1-p_n$ です. よって

$$\begin{aligned} p_{n+1} &= (n \text{ 回後に } A \text{ にいる確率}) \times \frac{2}{3} + (n \text{ 回後に } A \text{ にいない確率}) \times \frac{1}{6} \\ &= p_n \times \frac{2}{3} + (1-p_n) \times \frac{1}{6} = \frac{p_n}{2} + \frac{1}{6}. \end{aligned} \tag{1}$$

あとはこの漸化式を解けばよいことになります. この漸化式を解くために p_{n+1} と p_n を α に変えると

$$\alpha = \frac{\alpha}{2} + \frac{1}{6} \quad \Leftrightarrow \quad \alpha = \frac{1}{3}, \tag{2}$$

になります. よって (1) から (2) を引くと,

$$p_{n+1} - \frac{1}{3} = \frac{1}{2}\left(p_n - \frac{1}{3}\right),$$

になります. これは数列 $\{p_n - 1/3\}$ が初項 $p_1 - 1/3 = 2/3 - 1/3 = 1/3$, 公比 $1/2$ の等比数列であるということなので,

$$p_n - \frac{1}{3} = \frac{1}{3}\left(\frac{1}{2}\right)^{n-1} \quad \Leftrightarrow \quad p_n = \frac{1}{3} + \frac{1}{3}\left(\frac{1}{2}\right)^{n-1},$$

となり, p_n が求められました.

1.3 マルコフ連鎖

この問題で注目したいことは以下の二つです.

(i) 点 Q が n 回投げたあとに次の点に移動する確率は, 点 Q がどこにいるかだけで決まる.

(ii) さいころを投げるという行動は重要ではなく, 点 Q が次にどの点に移動するかという確率が決まっている.

(i) についてですが, この性質は**マルコフ性**とよばれる性質になります. 例えばある時刻で A にいるとしましょう. このあとどのように移動するかは, それまでにどのように動いてきたかには無関係です. 例えば時刻 3 で Q が A にいたとします. Q が A から出発して $A \to B \to C \to A$ と移動してきたり, $A \to C \to B \to A$ と移動してきても, この先どのように動くかには無関係です. このようにこの先どのように動くかは, 過去の行動に無関係であるという性質は, **マルコフ性**とよばれます.

次に (ii) についてですが, 問題文を読み解くと Q が次にどの点に移動するかの確率が与えられています. 今の場合には, 確率 $2/3$ で今いる場所から移動せず, 隣のどちらかの点に確率

1/6 で移動します. このような次の点に移動する確率のことを, **推移確率**といいます. 元々の問題では “n 回さいころを投げる” というように問題が与えられていますが, 点が動くという観点からこれを “時間” とも捉えます.

この問題では A, B, C 上をランダムに移動する点を考えていて, 更にマルコフ性とよばれる性質をもっています. このようにマルコフ性をもち, 時間ごとにランダムに変化するものは**マルコフ連鎖**とよばれます. 標題のランダムウォークもマルコフ連鎖の一種です.

1.4 例 1: ある国の天気

問題を読み替えているだけですが, 先の問題を次のように言い換えます.

> 天気が晴れ, くもり, 雨の三種類しかない国がある. その国の天気は 1 日ごとに等確率 1/6 で別の天気になり, 確率 2/3 で変わらない. その国のある日の天気は晴れであった. その日から数えて n 日後の天気が晴れになる確率はいくつか.

この問題は先の入試問題を言い換えているだけなので, 答えは

$$\frac{1}{3}+\frac{1}{3}\left(\frac{1}{2}\right)^{n-1},$$

になります. 今の場合には天気がくもりであることと雨であることは対称的なので, n 日後にくもりになる確率は

$$\frac{1}{2}\times(n\text{ 日後の天気が晴れがでない})=\frac{1}{3}-\frac{1}{6}\left(\frac{1}{2}\right)^{n-1},$$

になります. このように天気が変わる国はおそらくないでしょうが, この天気の例はマルコフ連鎖の例としてよくあげられるものです. 点が円周上の 3 点を移動する問題に比べれば, 幾分実用的な問題になっているのではないでしょうか. 勿論このようなマルコフ連鎖に基づいて天気予報はしていないでしょうが, 「マルコフ連鎖がどうなるか」という問題は, 多くの分野で考えられている問題になります.

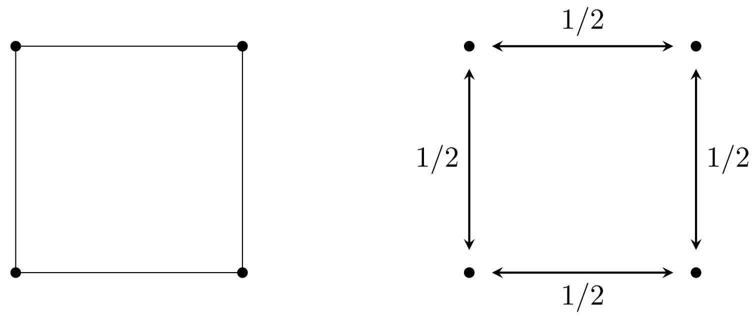
1.5 例 2: 多角形上のランダムウォーク

マルコフ連鎖の例としてここでは多角形上のランダムウォークについて説明します. 初めに四角形の場合を考えます.[3] 図では正方形でかいていますが, 別に正方形である必要はありません. 長さは重要でなく, 頂点が決まっていることと, どこに辺があるか, ということが重要です. 自然数 $n=1,2,3,\ldots$ に対して, 時刻 n から時刻 $n+1$ における点の移動の仕方を次で与えます.

> どの頂点からも隣の頂点に確率 1/2 で移動する.

[3] 少し推移確率は違いますが大阪大学の入試問題もランダムウォークの一つです.

この推移確率を図示すると次のようになります.



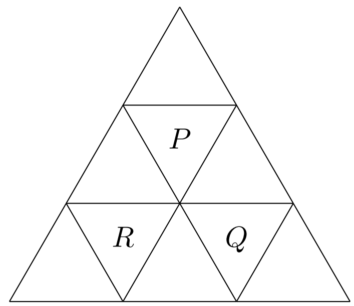
同じように五角形, 六角形, ... 上のランダムウォークを考えることができます. 後ではランダムウォークを多角形の上ではなく, 数直線の上で考えます.

1.6 大学の入試問題より別の例

もう一つ大学の入試問題から例を挙げておきます. 次の問題は 2012 年度東京大学の入試問題, 文科第三問・理科第二問からの引用です.[4] 先の問題よりも複雑ではありますが, 大阪大学の入試問題と同じ考えで解くことができます. 興味のある人は考えてみてください.

> 図のように, 正三角形を 9 つの部屋に辺で区切り, 部屋 P, Q, R を定める. 1 つの球が部屋 P を出発し, 1 秒毎に, そのままその部屋にとどまることなく, 辺を共有する隣の部屋に等確率で移動する. 球が n 秒後に部屋 Q にある確率を, 以下の手順で求めよ.
>
> 1. 時刻 2 で球が P, Q, R にいる確率をそれぞれ求めよ.
> 2. n を自然数として, 時刻 $2n-1$ に球が P, Q, R のいずれかにある確率を求めよ.
> 3. p_n を時刻 $2n$ で球が Q にいる確率とする. 時刻 $2n$ で球が Q, R にいる確率をそれぞれ p_n を用いて表せ.
> 4. 時刻 $2n$ だけを考えると, 球は P, Q, R の上を移動するマルコフ連鎖である. 小節 2.2 の図のように推移確率を図示せよ.
> 5. p_{n+1} を p_n を用いて表し, その漸化式を解くことにより p_n を求めよ.

[4] 解きやすいように誘導をつけています.



1.7 数 III からの知識

大阪大学の入試問題に潜む背景を説明するために, 等比級数の極限について説明をしておきます. 数 III で習う知識になるので, わからなくとも流してもらって大丈夫です.

初項 1, 公比 $r > 0$ の等比数列 a_n を考えます. a_n の一般項は

$$a_n = r^{n-1},$$

で与えられます. 公比がそれぞれ $1/2, 1, 2$ である場合の最初の 10 項を具体的に列挙すると,

$$\begin{aligned} r = \frac{1}{2}: &\quad 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \dots \\ r = 1: &\quad 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, \dots \\ r = 2: &\quad 1, 2, 4, 8, 16, 32, 64, 128, 256, 1024, \dots \end{aligned}$$

になります. 具体的にみて分かるように, n が大きくなるにつれ $r = 1/2$ の場合には a_n はどんどん小さくなり, 0 に近づいていきます. よって

$$\lim_{n\to\infty} \left(\frac{1}{2}\right)^{n-1} = 0,$$

になります. 上の例からの類推で, r が 1 よりも真に大きい (つまり $r > 1$) と, a_n は無限大に近づいていき, $r = 1$ の場合には a_n は常に 1 になります. 一方 r が 1 よりも真に小さい場合 (つまり $r < 1$) にも, $r = 1/2$ の場合と同じで, n が大きくなるにつれて a_n は 0 に近づきます. よって r を $0 < r < 1$ なる実数とすると,

$$\lim_{n\to\infty} r^n = 0,$$

になります.

さて $0 < r < R < 1$ なる実数 r, R を考えます. どちらも 1 より小さいので,

$$\lim_{n\to\infty} r^n = \lim_{n\to\infty} R^n = 0,$$

になります. なので r^n と R^n どちらも 0 に近づきます. ここではどちらがより速く 0 に近づくかを考えます. 再び具体例を考えます. $r = 1/1000$ と $R = 1/2$ を考えると,

$$\begin{aligned} r = 1/1000: &\quad \frac{1}{1000}, \frac{1}{1000000}, \frac{1}{1000000000}, \frac{1}{1000000000000}, \cdots \\ R = 1/2: &\quad \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \cdots \end{aligned}$$

となります. この例では r は R よりも小さいので, r^n と R^n を比べると r^n の方が 0 に速く近くことがわかります. 一般の場合にも同じことが成立します, つまり, $0 < r < R < 1$ なる場合には, r^n と R^n では r^n の方が速く 0 に近づきます. まとめると,

二つの 0 に収束する等比数列は, 公比の小さい方が速く 0 に収束する

ということになります.

1.8 定常分布と収束の速さ

さて大阪大学の問題の例に戻ります. 一般項 p_n は,

$$p_n = \frac{1}{3} + \frac{1}{3}\left(\frac{1}{2}\right)^{n-1},$$

で与えられました. このとき p_n の $n \to \infty$ での極限は

$$\lim_{n\to\infty} p_n = \lim_{n\to\infty} \frac{1}{3} + \frac{1}{3}\left(\frac{1}{2}\right)^{n-1} = \frac{1}{3},$$

になります. 極限 $\lim_{n\to\infty} p_n$ 極限のことを大雑把にいうと,

n が大きいときの p_n の値

に他ならないので, 今の場合には

時刻が十分経ったあとに点 Q が A にいる確率は 1/3

と考えられます.

さらにこの 1/3 はある特別な意味をもちます. そのため次の漸化式があったことを思い出しておきます.

$$p_{n+1} = \frac{p_n}{2} + \frac{1}{6}.$$

仮定として, 時刻 n で点 Q が A にいる確率が 1/3 であったとしましょう. このとき, 次の時刻で点 Q が A にいる確率 p_{n+1} は上の漸化式より

$$p_{n+1} = \frac{p_n}{2} + \frac{1}{6} = \frac{1}{3}\cdot\frac{1}{2} + \frac{1}{6} = \frac{1}{3},$$

となり, 次の時刻でも変わらず 1/3 です. これは振り返ってみると当たり前で, 何故かというと 1/3 は方程式[5]

$$\alpha = \frac{\alpha}{2} + \frac{1}{6},$$

の解であったからです. これは点 Q が B または C にいる確率についても同じです.

これらを天気の例でまとめると以下のようになります.

> ある日の晴れ, 雨, くもりになる確率が全て 1/3 であったとすると, 次の日の天気が晴れ, 雨, くもりになる確率も再び全て 1/3.

もともとの問題では初期時刻で点 Q は A にいるとしていたので, A, B, C にいる確率は時刻 n ごとに変化していきます. しかしながら問題を「点 Q は A, B, C を等確率 1/3 で選び, その選んだ点から出発する」と問題を変えると, 時刻 n で点 Q が A にいる確率は n に依らず 1/3 になります. このような時刻によって変わらない確率 (A, B, C 等確率 1/3) を, マルコフ連鎖の**定常分布**といいます. 実はマルコフ連鎖が (唯一の) 定常分布に収束するとき, その定常分布は時刻 0 でどうなっているかということに無関係であることが証明されます.

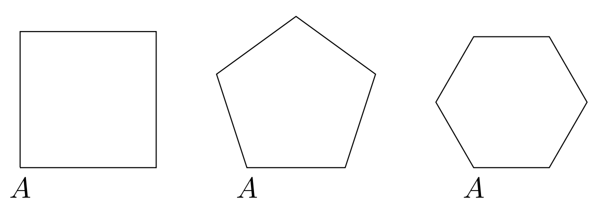
一方 1/2 は何かというと, p_n の一般項より

$$p_n - \frac{1}{3} = \frac{1}{3}\left(\frac{1}{2}\right)^{n-1} \overset{n\to\infty}{\longrightarrow} 0,$$

となっているので, 1/2 は定常分布への収束の速さを表しています. どちらについても次の小節でもうすこしみてみます.

1.9 多角形上のランダムウォーク

大阪大学の問題の他に, 定常分布と収束の速さが分かる具体例を考えてみたいと思います. ここでは単純に同じマルコフ連鎖を四角形, 五角形, 六角形上のものを考えます. つまり, 頂点 A を出発して, 次の時刻では今いる点に確率 2/3 で滞在して, 1/6 で両隣のどちらかの点に移動する, という推移確率を考えます.[6]



同じ点には 2/3 の確率, 両隣には 1/6 の確率で移動.

[5]特性方程式ともよばれます.

[6]2/3 であることは (ある意味) 重要ではないです. 大阪大学の問題に倣っているだけです.

点 A を出発して時刻 n で点 A にいる確率を p_n とすると, p_n の一般項はそれぞれ次で与えられます.

$$\begin{aligned}
\text{四角形：}\quad p_n &= \frac{1}{4} + \frac{1}{4}\left(\frac{1}{3}\right)^n + \frac{1}{2}\left(\frac{2}{3}\right)^n, \\
\text{五角形：}\quad p_n &= \frac{1}{5} + \frac{2}{5}\left(\frac{7-\sqrt{5}}{12}\right)^n + \frac{2}{5}\left(\frac{7+\sqrt{5}}{12}\right)^n, \\
\text{六角形：}\quad p_n &= \frac{1}{6} + \frac{1}{6}\left(\frac{1}{3}\right)^n + \frac{1}{3}\left(\frac{1}{2}\right)^n + \frac{1}{3}\left(\frac{5}{6}\right)^n.
\end{aligned}$$

高校のときに学ぶ数学の知識ではこれらを計算することはできないので, ここでは受け入れて頂きたいです. 大学生で学ぶ行列に関する理論を用いると, 一般項が計算できるようになります.[7] この一般項で注目したいところを赤と青で示しています. 四角形の場合について説明すると, 等比数列の極限が 0 になることを思い出せば

$$\lim_{n\to\infty} p_n = \frac{1}{4},$$

になります. これの意味するところは,

十分時間が経ったら, A にいる確率は大体 $1/4$

ということです. 頂点が四つあるのでいずれ等確率でどこかにいる, というのは直感的に納得できるのではないでしょうか.

一方四角形の場合の一般項では $(1/3)^n$ と $(2/3)^n$ があります. 小節 2.7 で考察したように, $(1/3)^n$ と $(2/3)^n$ では $(1/3)^n$ のほうが速く 0 に収束します. なので p_n が $1/4$ に収束する速さは $(1/3)^n$ に比べて 0 に収束する速さが遅い $(2/3)^n$ により決まります. 同様に五角形, 六角形の場合にも青で示しているところが収束の速さを決めています. 収束の速さを決める問題はマルコフ連鎖の研究上重要な問題であり, 現在も多くの具体的なマルコフ連鎖で研究されています.

1.10　マルコフ連鎖の応用例

これまでみた定常分布や定常分布への収束の速さを決定することは, 多くの問題で基本的かつ重要な問題です。具体的にそのような問題をみることはここではしませんが, Wikipedia のページに記載されているマルコフ連鎖の応用例を箇条書きにしておきます.

・物理学, 統計力学　・統計学, 待ち行列理論　・情報理論

・強化学習　・Google の PageRank　・声認識, バイオインフォマティクス

このように多くの分野でマルコフ連鎖は考えられています. 身近なあたりでは Google の検索技術に使われているあたりからも感じられるのではないでしょうか. そのような応用が背後に潜んでいることが, 実は大学入試の問題の背景にあります.

[7]ただ筆者は手計算ではなくてマテマティカに計算をさせています.

2　確率論の基礎概念

2.1　確率論について

確率が数学的に研究されたのは, 16 世紀のカルダーノや 17 世紀のパスカルとフェルマーによるものと考えられています. それ以前にも確率現象というものは認知され, 古くから考えられてきたと思います. 参考に例えば Wikipdea の「確率の歴史」のページをご覧ください. 紀元前に幾何学や代数学が考えられていることに比べると, 数学における確率論の歴史は比較的浅いように思えます. 現代まで研究されている確率論が昔からの延長で考えられているかというと, そうではありません. 初期に行われた確率論の研究は, **古典的確率論**とよばれます. Wikipedia より抜粋すると,

> 確率論は 16 世紀から 17 世紀にかけてカルダーノ、パスカル、フェルマー、ホイヘンス等によって数学の一分野としての端緒が開かれた。イタリアのカルダーノは賭博師でもあり、1560 年代に『さいころあそびについて』（羅: Liber de ludo aleae）を執筆して初めて系統的に確率論を論じた。その書は彼の死後の 1663 年に出版された。18 世紀から 19 世紀にかけて、ラプラスはそれまでの確率論を統合する研究を行い、1814 年 2 月に『確率の哲学的試論』を著し、古典的確率論と呼ばれる理論にまとめた。

これらの確率論では, 基本的に中学・高校で学ぶように, 確率が数学的に何かということを正確には定義しないで, 具体的な現象について考えます. 一方現代数学では, 初めに確率とは何かということを定義し, 集合論, 測度論, ルベーグ積分論とよばれる数学の理論に基づき確率論を展開します. このような確率論はロシアの数学者コルモゴロフにより基礎が与えられ, 現在では**公理的確率論**とよばれます. 再び Wikipedia より抜粋すると,

> 現代数学の確率論は、アンドレイ・コルモゴロフの『確率論の基礎概念』（1933 年）に始まる公理的確率論である。この確率論では「確率」が直接的に何を意味しているのかという問題は取り扱わず、「確率」が満たすべき最低限の性質をいくつか規定し、その性質から導くことのできる定理を突き詰めていく学問である。

と書かれています. 確率を定義するということの具体例を次の小節でみてみます.

2.2　さいころ投げの例

高校数学でよく考える設定のように, さいころを一回投げる試行を考えます. でる目の確率が同様に確からしいさいころであれば, どの目がでる確率も $1/6$ になります. 一般に事象 A が起こる確率は,

$$\frac{\text{事象 } A \text{ が起こる場合の数}}{\text{考えられる全ての場合の数}} \tag{3}$$

で与えられます. 事象とは起こりうる結果の集まりのことでした. 例えば偶数がでるという事象は, $2, 4, 6$ のいずれかがでるということに他なりません. なので偶数がでるという確率は, $3/6 = 1/2$ となります.

これらは既に学んでいる人にとっては慣れ親しんだものだと思いますが, 例えば, 「同様に確からしい」とは何か, 「確率」とは何か, といったことについては具体的には述べていません. これを集合を用いると次のように述べることができます.

集合 Ω を $\Omega = \{1,2,3,4,5,6\}$ とします. Ω は起こりうる結果全体の集合であり, **標本空間**とよばれます. つまり, Ω は単に 1 から 6 がでる, ということを述べています. Ω の部分集合は**事象**とよばれます. Ω の部分集合とは, Ω の一部の要素からなる集合のことでした. 例えば, $\{1,3,5\}$ や $\{2,4,6\}$ のことです. 前者は奇数がでる, 後者は偶数がでる事象とそれぞれ考えられます. また事象全体のことを $\mathcal{F}$ とかくことにします. 強調しておくと, 確率を定めることのできる事象が与えられており, それら全体を $\mathcal{F}$ とかくということです. さて事象 A の要素の数を $|A|$ とかくことにします. 例えば $|\Omega| = 6$ ですし, $|\{1,2\}| = 2, |\{2,4,6\}| = 3$ です. 最後に事象 A がでる確率を定義します. それを $P(A)$ とかくことにして, 次で定義します.

$$P(A) = \frac{|A|}{|\Omega|}. \tag{4}$$

(3) で与えられる確率と, (4) で与えられる確率が同じものを表しています. 例えば偶数がでる事象, つまり $\{2,4,6\}$ を A とすれば, $P(A) = |A|/|\Omega| = 3/6 = 1/2$ になります.

次に, 確率がもつ極めて重要な性質について注意しておきます. (4) で定義される確率 P は, **有限加法性**とよばれる次の性質を満たします. 二つの事象 A と B が同時に起こらない, 集合論の言葉でかくと $A \cap B = \emptyset$ であるとすると, 事象 A または B が起こる確率は, それぞれの確率の和で与えられます. 式でかくと次になります.

$$P(A \cup B) = P(A) + P(B). \tag{5}$$

高校数学でよく出てくる文言として, 「同様に確からしい」という言葉があると思います. これは確率を (4) のように定義する, ということに他なりません. しかしながら数学的に考えるのであれば, 1 から 6 がでる確率が同じでなくても構いません. $p_1, p_2, p_3, p_4, p_5, p_6$ を全て 0 から 1 の間の数として,

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1,$$

となるものとします. 確率 P の定義 (4) を次のように変更します. 事象 A に対して,

$$P(A) = \sum_{k=1}^{6} p_k \mathbf{1}\{k \text{ が } A \text{ に含まれている }\}. \tag{6}$$

和の記号 $\sum$ を用いてかかれているので分かりにくいかもしれませんが, 単に 1 から 6 がでる確率をそれぞれ $p_1, \ldots, p_6$ として, それらの確率を足しているだけです. 例えば,

$$P(\{2,4,6\}) = p_2 + p_4 + p_6,$$

になります. (6) のように確率を定義したとしても, (5) で定義される確率が有限加法性の性質を満たすことを注意しておきます.

2.3 確率空間

前の小節ではさいころ投げの試行を集合の言葉で書き直しました. 重要なことをまとめると,

- 起こりうる結果全体の集合 Ω が与えられている
- 確率を定義することのできる事象の集まり $\mathcal{F}$ が与えられている
- 事象 A に対して確率 $P(A)$ が定義されている

ということです. この三つの組み $(\Omega, \mathcal{F}, P)$ のことを**確率空間**とよびます.[8] 標本空間の元のことを**標本**とよびます. 標本はしばしば ω(小文字のオメガ) で表されます.

さいころをなげる試行であれば, 先に見たようにそこまで複雑にはなりません. 基本的には確率現象は確率空間を具体的に考えることにより, 数学的に確率現象を考えることができます. しかしながらこのように考えたからといって高校数学の確率分野がよく理解できるかというと, そうではありません.

実は確率論では有限加法性だけでは色々と不十分なので, 有限加法性ではなく**加算加法性**とよばれる次の性質が必要になります. 事象の列 $A_1, A_2, A_3, \ldots$ は互いに排反だとする, つまり任意の $i \neq j$ に対して $A_i \cap A_j = \emptyset$ とすると, 次が成立する.

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \tag{7}$$

(5) の右辺が有限個の足し算であるのに比べて, (7) の右辺は無限個の足し算になっています. この差は些細なようにも思うかもしれないですが, このことがものごとを非常に上手く展開させます. 更に正確に述べると, 事象の集まり $\mathcal{F}$ は σ 加法族とよばれるものになっていることが必要ですが, ここでは説明を割愛します. 公理的確率論は古典的確率論では考えられない数学を研究されるために発展してきたともいえます. 一方素朴に確率現象を考えれば, 公理的確率論が必要になる場面は多くはありません. 実際, 確率空間の細かい数学的定義を考えることなく, 高校数学で学んだように色々な確率を考えることもできます.

このように公理的確率論を数学的に考えるには, 測度論というものを用いる必要があり, 正確に理解することは容易ではありません. この測度論と呼ばれる理論は数学科の三年生頃に学ぶもので, 中々に抽象的で難解な理論です. ここではこのような理論について数学科で学ぶということのみを, 触れるだけにしておきます.

2.4 確率 = 面積

一度確率を上のように抽象化して考えてみると, 確率と面積は同じであるものと捉えることができます. 面積を例えば次のように考えてみます.

- 1 辺が 1 の正方形

[8]数学では集合に何かしらの構造を加えたものを空間とよびます. 線形空間、位相空間など, 数学では多くの空間が考えられます.

- 正方形内の部分図形[9]
- 正方形内の部分図形を A として, $P(A)$ を面積とする

図形に対して面積を返すという関数は, 確かに有限加法性を満たしています. このように考えると, 面積は確率に他ならない, と考えられます. 確率現象として何を捉えているかというと, 次のように考えられます.

- 1 辺が 1 の正方形から点を一様ランダムに選ぶ. このとき A を正方形内の部分図形とすると, $P(A)$ は一様ランダムに選んだ点が A に入っている確率である.

このことに基づいて, 正方形からランダムに点を取ってくるという操作により円周率を計算することができます. 実際, 1 辺が 1 の正方形内の部分図形 A として, 正方形に内接する円の内部を考えると, A の面積は $P(A)$ は

$$P(A) = \pi \times \frac{1}{2} \times \frac{1}{2} = \frac{\pi}{4}$$

です. 例えば後で述べるモンテカルロ法により, 円周率 π を計算することができます.

2.5 ベルトランのパラドックス

確率空間を正確に設定しないとパラドックスが起こってしまうことについて注意します.[10] これは出てくる結果が有限個の場合には問題になりませんが, 起こりうる結果が無限個の場合にとくに注意が必要になります. 以下の問題を考えてみます.

- 円に内接する正三角形を考える. その円の弦を一本無作為に選ぶ. その弦が正三角形の辺よりも長くなる確率はいくつか?

この問題に対して, 例えば次の三つの回答が考えられます.

(解答 1) 弦の一つの端点を頂点とする正三角形を考える. このとき選んだ弦が正三角形の辺よりも長くなるのは, 弦のもう一つの端点が向かい合う辺の円弧上にあるときなので, 求める確率は 1/3.

(解答 2) 正三角形の一つの辺 ℓ_1 が選んだ弦に並行になるように正三角形を考える. またその弦に並行であって円の中心を通るもの ℓ_2 を考えると, 求める確率は選んだ弦が ℓ_1 と ℓ_2 の間に位置するときなので, 求める確率は $1/2$.

(解答 3) 選ばれた弦の中点を P とする. 元々の円の半径の $1/2$ の円を考えると. 求める確率は P が小さい円の内部にあるときなので, 求める確率は $1/4$ (小さい円と大きい円の面積の比は半径に無関係に $1:4$).

[9] 実は正方形内の部分図形全てに面積を割り当てることはできません. そのような事情を数学に考えるために, 測度論が必要になりますが, ここではぼやかしておきます.

[10] パラドックスとは日本語では逆説といいます. 数学のパラドックスで有名なものではバナッハ・タルスキーのパラドックスやラッセルのパラドックスがあります.

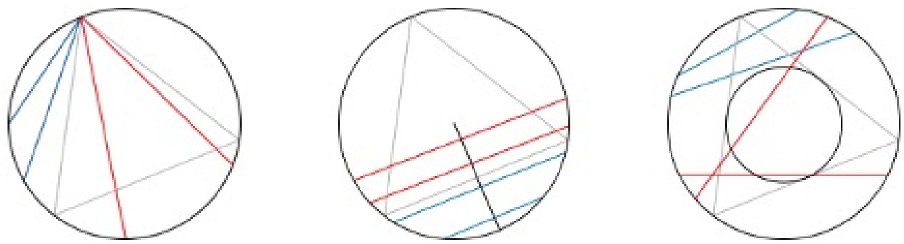


Figure 1: 左から順に解答 1–3 に対応. 青は正三角形の辺より短く, 赤は長い.

これらの解答のどれが正しいでしょうか? どれも正しくみえますが, 答えの確率は違っています. この例は**ベルトランのパラドックス**とよばれ, 確率の定義を正確に行わないことに生じるものです. 問題では「円の弦を一本無作為に選ぶ」ありますが, (解答 1) では無作為に端点を, (解答 2) では無作為に端点を半径を (ℓ_2 に対応), (解答 3) では無作為に弦の中点を, それぞれ選んでいます. このように「無作為に選ぶ」ということは一通りの方法ではなく, 色々なものが考えられます. それに応じて結果の確率が変わってきます. このことは確率空間を正確に定めていないことから生じてきます. この例から示唆されるように, 初めに確率とは何かということを正確に定義しておかないと, その後の結果自体が変わってきてしまうことがわかります.

2.6 独立性

確率論で重要な性質である独立性について説明します. 確率において独立とはそもそもなんだったでしょうか. 独立が何かということは説明できなくても, 感覚的には勿論知っていると思います. 具体例として, でる目が同様に確からしいさいころを三回投げる試行を考えます. 今後は確率空間を用いて確率をかくときもありますし, 特に不必要なときには細かいことにはあまりこだわらなくて問題ありません. 皆さんはご存知のとおり, 「独立な試行の確率はそれぞれの確率の積で与えられる」という事実を高校で学びます. 例えば一回目が偶数かつ二回目は 1 か 2 がでる確率は,

$$\begin{aligned}
&P\,(\text{一回目が偶数かつ二回目は 1 か 2}) \\
&\quad = P\,(\text{一回目が偶数}) \times P\,(\text{二回目は 1 か 2}) \\
&\quad = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6},
\end{aligned}$$

になります. 他には一回目, 二回目, 三回目全てが偶数になる確率は,

$$\begin{aligned}
&P\,(\text{一回目が偶数かつ二回目が偶数かつ三回目が偶数}) \\
&\quad = P\,(\text{一回目が偶数}) \times P\,(\text{二回目が偶数}) \times P\,(\text{三回目が偶数}) \\
&\quad = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8},
\end{aligned}$$

になります.

ここまで勿体ぶって話していますが, 二つの事象が独立であるとは, 上の性質が成立することに他なりません. つまり, 事象 A と事象 B が**独立**であるとは,

$$P(A \cap B) = P(A)P(B),$$

が成立することをいいます. たったこれだけのことですが, このことが色々なことに作用して, 豊かな数学の定理を生み出します.

二つの事象ではなくて, 三つ, 四つなどのの事象については定義がもう少し複雑になります. n 個の事象 $A_1, A_2, \dots, A_n$ が独立であるとは, 任意の $1 \le i_1 \le i_2 \cdots \le i_k \le n$ に対して,

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}),$$

が成立することをいいます. 急によく分からなくなったかもしれないので, $n = 3$ の場合を具体的に書き下してみます. 事象 A, B, C が独立であるとは, 以下が成立することです.

$$\begin{aligned} P(A \cap B \cap C) &= P(A)P(B)P(C), \\ P(A \cap B) &= P(A)P(B), \\ P(B \cap C) &= P(B)P(C), \\ P(A \cap C) &= P(A)P(C), \end{aligned}$$

二個目から四個目の式については, A, B, C 三つの中から取り出した二つの事象がそれぞれ独立であることを言っています. これに加えて最初の式が成立する必要があります. 一般的には二個目から四個目が成立したとしても, 一個目の式が成立するとは限りません.

2.7 確率変数（さいころの例）

何かしらの試行を行った際, 結果として何か数字が得られるとき, その結果の数字を X とかくことにします. X は**確率変数**とよばれます. 確率変数とはこのようにランダムに出る数字のことですが, 数学的には非常に曖昧なもののように感じられます. 数学的に確率変数が何かを述べる前に具体例をみます.

例によって 1 から 6 の目がかかれたさいころを一回投げる試行を考えます. 勿論出る目の確率は同様に確からしいとします. このとき出た目の数字を X とすると, X は確率変数です. 驚くほど簡単な例ですが, X が具体的にどんな数字になっているかはさいころを投げない限り具体的には定まりません. しかしながらこの確率変数によりいろいろな確率が定まることに注意です. 例えば, $\{X$ は $2, 4, 6$ のいずれか $\}$ とすると, これはさいころの出目が偶数である事象に他なりません. これを略記して $\{X = 2, 4, 6\}$ とかくと,

$$P(X = 2, 4, 6) = \frac{1}{3},$$

になります. 他には任意の $i = 1, 2, 3, 4, 5, 6$ に対して, でる目は等確率としたので

$$P(X = i) = \frac{1}{6}, \tag{8}$$

になります. 一般の事象の確率は, (8) に確率の加法性を組み合わせることにより求めることができます.

少し奇妙に思うかもしれませんが, $\{X = 100\}$ という事象を考えても問題はありません. 勿論このさいころをなげて出目が 100 になることはないので,

$$P(X = 100) = 0,$$

です. 何も言っていないトートロジーですが, そのような事象であったり, そのようになる確率を考えてもよいということは注意したい点です.

2.8 確率変数 (再考)

さいころの例ではなんなので, 確率変数が何かということを別の例とともに再考してみます. 次の例を考えます.

- 袋の中に 1 から 100 がかかれたくじが入っている. この中から一つのくじを無作為に選び, 引いたくじにかかれている数字に応じて以下で与えられる賞金がもらえる.
  - 数字が 100 以外の 10 の倍数 $\cdots$ 100 円.
  - 数字が 100 $\cdots$ $10,000$ 円.
  - それ以外 $\cdots$ 1 円.

  このときもらえる賞金を X 円とすると, X は確率変数である.

上の X が確率変数であることは理解できると思います. これを数学的に正確に考えようとすると, 標本空間上の関数として捉えられます. 上の例に沿うために標本空間 Ω を $\Omega = \{1, 2, 3, \dots, 100\}$ とし, ω を Ω の元とします. 先の例では ω は引いたくじに書かれている数字に対応します. もらえる賞金は ω によって変わるので, もらえる賞金を $X(\omega)$ とかくことにすると, $X(\omega)$ はそれぞれ次で与えられます.

$$X(\omega) = \begin{cases} 100, & \omega \text{ が } 100 \text{ 以外の } 10 \text{ の倍数}, \\ 10000, & \omega \text{ が } 100, \\ 1, & \text{それ以外}. \end{cases}$$

このように標本 ω ごとに, 何かしらの数字 $X(\omega)$ が定められています. このようなものは関数とよばれました. また確率論では何が起こるかは試行を行わない限り分からない, ということもあるので, しばしば標本が何であるかということを考えません. この理由もあるので $X(\omega)$ のことを単に X と通例かきます. 確率変数とは単にこの X のことです. なので数学的に確率変数は何かというと, 標本空間上の関数, として定義されます.[11]

[11] 説明は省きますが, 実際は可測な関数である必要があります.

2.9 離散型確率変数

他にも幾つか確率変数の例もあげておきます.

- でる目の確率が等しいさいころを独立に二回投げる試行を考える. 出た二つの数字をかけた数を X とする. X は確率変数.
- ある日に流れ星の数を観測する. 合計で観測できた数を Y とすると Y は確率変数.[12]
- 0 から 1 の数字を一様にサンプルする. サンプルした数字を Z とすると Z は確率変数.

X と Y の結果は飛び飛びの値にしかならないことに注意です. 実際, どうなっているかの確率は抜きにして, X は高々1 から 36 個の値しかとりませんし, Y は $0, 1, 2, 3 \ldots$ と, 確かに飛び飛びの値にしかなりません. このようなタイプの確率変数を**離散型確率変数**といいます. 一方, Z は 0 から 1 の値しかとりませんが, $0, 1, 2, \ldots$ のように数えることはできません. このような確率変数を**連続型確率変数**といいます. 連続型の確率変数は重要な対象ですが, 記述を簡単にするためにこのノートでは扱いません. 以降は確率変数と単によんだら離散型の確率変数を指すことにします.

2.10 確率変数の期待値

再び 1 から 6 の目がかかれたさいころを一回投げる試行を考えます. 勿論出る目の確率は等しいとします. このとき出る目の**期待値 (平均値)** とは,

1 (それぞれの出た目の数字)×(それが出る確率) をそれぞれ計算して,

2 それら全てを足して得られる値,

のことでした. 今の場合には 1 から 6 は等確率で出るので,

$$1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2}, \tag{9}$$

となります. 例えばさいころの出目だけお金がもらえるとすると, この試行を行えば平均的には 7/2 円ほどもらえることが期待されます.

このことを確率変数の言葉でかいてみます. 出た目の数字を X とします. このとき期待値は和の記号 Σ を用いて次のようにかくことができます.

$$\sum_{k=1}^{6} kP(X = k). \tag{10}$$

勿論これは (9) と同じものを表しています.

[12] この例は不正確で, 例えば流れ星が観測される頻度を与えていないため, Y がどうなっているかという確率は計算できません. まあ雑に考えて下さい

一般の場合の離散型確率変数に対しても, 期待値は (10) のように定義されます. X を離散型確率変数とします. つまり, X は $a_1, a_2, a_3, \ldots$ の値になる可能性があり, X がそれぞれの値になる確率は $p_1, p_2, p_3, \ldots$ で与えられているものとします. $a_1, a_2, a_3, \ldots$ はそれぞれ異なる実数ですが, $p_1, p_2, p_3, \ldots$ は全て 0 から 1 の数であって, 次を満たしています. (全ての確率を足すと 1 ということ)

$$\sum_{k=1}^{\infty} p_k = 1.$$

このとき確率変数 X の**期待値**もしくは**平均値** $E[X]$ とは次のことを指します.

$$E[X] = \sum_{k=1}^{\infty} a_k P(X = a_k) = \sum_{k=1}^{\infty} a_k p_k.$$

また期待値は線型性とよばれる次の性質を満たします.

線型性 a, b を実数, X_1, X_2 を確率変数とすると $E[aX_1 + bX_2] = aE[X_1] + bE[X_2]$.
特に, $a = b = 1$ とすると $E[X_1 + X_2] = E[X_1] + E[X_2]$.

2.11 サンクトペテルブルクのパラドックス

期待値はランダムに出る数字が平均的にどの程度なのかということを感覚的に表しています. なのでさいころの出た目だけ賞金がもらえるとすると, 平均的に $7/2 = 3.5$ 円もらえると考えられます. それでは期待値が参加料よりも大きい賭けがあったとして, 必ず儲けることが出来るでしょうか. この問題に対して**サンクトペテルブルクのパラドックス**とよばれる直感的にどうすればよいのか分からない問題があります. それは以下のようになります.

> 表裏が出る確率が等しいコインがある. 裏が出続けるまでコインを投げて, 裏が出るまでに出た表が n 回でたとすると, 賞金として 2^{n+1} もらえる. 例えば, 3 回目に裏が初めて出たとすると, 表が 2 回続けて出るので $2^{2+1} = 2^3 = 8$ 円, 5 回目に裏初めて出たとすると, 表が 4 回続けて出るので $2^{4+1} = 2^5 = 32$ 円もらえることになる (最初に裏がでたら賞金は 2 円とする). 言い換えると, 表が出るごとに賞金が 2 倍になるゲームである. このゲームの参加費がいくらであれば行う方が得か.

表が n 回連続出た後に裏がでる確率は,

$$\left(\frac{1}{2}\right)^n \times \left(\frac{1}{2}\right) = \frac{1}{2^{n+1}},$$

です. このときは賞金として 2^{n+1} 円もらえます. よってもらえる賞金を X とすると

$$P(X = 2^{n+1}) = \frac{1}{2^{n+1}}, \quad n = 0, 1, 2, \ldots,$$

になります. よって X の期待値 $E[X]$ は

$$E[X] = \sum_{n=0}^{\infty} \left(2^{n+1} \times \frac{1}{2^{n+1}} \right) = \sum_{n=0}^{\infty} 1 = \infty,$$

となります.

このゲームで得られる賞金の期待値はいくらでも大きくなるので, 期待値だけでいうといくらお金を払ってもこのゲームには参加すべきです. しかしながら参加費が例えば百万円だとしてこのゲームをやりたいかと言われると, 多くの人はやらないのではないかと思います. このようにゲームに参加して得られる賞金の妥当性を期待値だけで決めるのは現実的におかしいように思われます. このことはこのゲームの特性上, 「大きな賞金を得ることが非常に小さい確率で起こる」ことに起因しています. このゲームにいくら払えば得かという問いには答えてはいませんが, 期待値の捉え方については一概になんとも言えないことはこの例からわかります.

2.12 確率変数の独立性

確率変数の期待値になれるために以下の問題を考えます.

> さいころを順に二回投げる試行を考える. ただし一回目と二回目の出目は独立とする. 一回目のさいころの出目を X_1, 二回目のさいころの出目を X_2 とする. X_1 と X_2 はそれぞれ確率変数であり, 小節 2.10 で考えたように, $E[X_1] = E[X_2] = 7/2$ である. このとき次をそれぞれ求めよ.
>
> - 確率変数 $Y = X_1 + X_2$ の期待値 $E[Y]$.
> - 確率変数 $Z = X_1 X_2$ の期待値 $E[Z]$.

X_1, X_2 の値に応じて Y, Z がどのようになるか表に書いておきました. 表のどのパターンになるのも確率 $1/36$ であることを考えると, $E[Y]$ は左の表の数字を全て足して, 36 で割れば求められます. しかしながら期待値の線型性を用いると次のように簡単に求められます. 期待値の線型性から,

$$E[Y] = E[X_1 + X_2] = E[X_1] + E[X_2] = \frac{7}{2} + \frac{7}{2} = 7,$$

となります. 一般に n 回さいころを投げたとして, n 回目の出目をそれぞれ X_n とし, S_n を n 回目までのさいころの出目の総数とすると, 同様に

$$E[S_n] = E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] = \frac{7}{2} + \cdots + \frac{7}{2} = \frac{7}{2}n,$$

となります.

$X_1 \backslash X_2$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$X_1 \backslash X_2$	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

Table 1: 左から順に $Y = X_1 + X_2$, $Z = X_1X_2$ の表.

一方線型性では確率変数の積については何も述べていないので, Z の期待値を同様には計算できません. おとなしく表の数字を全て足して, 36 で割ることにより, 次がわかります.

$$E[Z] = 441/36 = 49/4.$$

$49/4 = (7/2)^2$ ということと, $E[X_1] = E[X_2] = 7/2$ であったので, 今の場合には

$$E[Z] = E[X_1X_2] = E[X_1]E[X_2],$$

が成立しています. このような「二つの確率変数の積の期待値が確率変数の期待値の積に等しい」ということは一般には正しくないのですが, 今の場合には成立しています. このことは一回目と二回目の出目が独立である, ということから成立します.

一回目と二回目の出目が独立とういうのは, 感覚的にはどちらの行動にも関係がない, と捉えると思います. 実際細かいことを気にしなければその理解で何も問題ないです. ただ数学的に確率変数が独立であることを考えようとすると, 事象のときと同じようになります. 正確な定義を述べると, 確率変数 X_1 と確率変数 X_2 が**独立である**とは,

$$P(X_1 \in A_1, X_2 \in A_2) = P(X_1 \in A_1)P(X_2 \in A_2), \tag{11}$$

が全ての A_1, A_2 に対して成立することを指します. ただし A_1, A_2 はそれぞれ X_1, X_2 が取りうる値の集合です.

例えば先のさいころ投げの例では, $A_1 = \{1, 2\}, A_2 = \{3, 4, 5, 6\}$ とすると,

$$P(X_1 = 1, 2, X_2 = 3, 4, 5, 6) = P(X_1 = 1, 2)P(X_2 = 3, 4, 5, 6),$$

が成立しています. しかし $A_1 = \{1, 2\}, A_2 = \{3, 4, 5, 6\}$ 以外の場合にも (11) は成立しています. このように X_1 と X_2 が独立であれば (11) が全ての A_1, A_2 に対して成立しているので, これをもって確率変数の独立性を定義するわけです.

また証明はしませんが, 確率変数 X_1, X_2 が独立であれば先の事実「二つの確率変数の積の期待値が確率変数の期待値の積に等しい」は確かに成立していることが示されます. つまり,

$$E[X_1X_2] = E[X_1]E[X_2],$$

が成立しています. 他にも例えば,

$$E[X_1^2X_2^2] = E[X_1^2]E[X_2^2],$$

なんかも成立します. より一般に自然数 n, m に対して

$$E[X_1^nX_2^m] = E[X_1^n]E[X_2^m],$$

が成立します.

確率変数が三個, 四個と増えても確率変数が独立であることが定義されますが, 事象のときにみたように書き方より複雑になるので, ここではみないことにします.[13]

2.13　確率変数の分散・モーメント

X をしばらく確率変数とします. 確率変数 X の期待値を m とかくことにします, つまり, $m = E[X]$ です.[14] 確率変数 X から平均を引いた数は $X - m$ となりますが, この数を二乗した $(X-m)^2$ もまた確率変数になります. 確率変数であればその期待値を考えることができますので, $(X-m)^2$ の期待値を考えることが出来ます.[15] $(X-m)^2$ の期待値のことを確率変数 X の**分散**といいます. 確率変数 X の分散を $V[X]$ とかくことにします.[16]

先と同じように X^2 を考えると, これもまた確率変数になります. 以降同じように $X^3, X^4, \dots$ が考えられますが, これらもまた確率変数になります. X^2 や X^3 はそれぞれ確率変数であるので, それらの期待値を考えることができます. X^2 の期待値 $E[X^2]$ は **2 次モーメント**, X^3 の期待値 $E[X^3]$ は **3 次モーメント**とそれぞれよばれます. 以下同様に自然数 $n = 1, 2, 3, \dots$ に対して確率変数 X^n の期待値 $E[X^n]$ は ***n* 次モーメント**とよばれます.

分散 $E[(X-m)^2]$ について少し考えてみます. $(X-m)^2$ を展開すると,

$$(X-m)^2 = X^2 - 2mX + m^2,$$

となります. $m = E[X]$ であったことと, 期待値の線型性に気を付けると,

$$\begin{aligned} E[(X-m)^2] &= E[X^2 - 2mX + m^2] \\ &= E[X^2] - 2mE[X] + E[m^2] \\ &= E[X^2] - 2mm + m^2 \\ &= E[X^2] - m^2, \end{aligned}$$

---

[13]勿論期待値が積に分かれることは一般的に成立します.

[14]期待値は expectation ですが, 平均値 mean ともよばれるのでしばしば期待値は m で表されます.

[15]一般的には確率変数に対してその期待値を定義できない場合があります. サンクトペテルブルクのパラドックスのように期待値が無限大になることは許されますが, そもそも期待値が定義されない確率変数も往々にあります.

[16]分散は variance なので V としています. 文献によっては期待値や分散を $E(X), V(X)$ など丸括弧を用いて表します.

になります. 再び $m = E[X]$ であったことを思い出すと, 確率変数 X の分散は

$$E[(X-m)^2] = E[X^2] - (E[X])^2, \tag{12}$$

で与えられます. 右辺が少し奇妙に思えるかもしれませんが, 「$E[X^2]$ は確率変数 X^2 の期待値」であって, 「$(E[X])^2$ は確率変数 X の期待値の二乗」であることに注意です. (12) より, 分散は 2 次モーメントと期待値を計算することにより求められます.

また (12) の左辺は, $(X-m)^2$ という 0 よりも大きい数を足し算した形になっているので, それらの合計である $E[(X-m)^2]$ も 0 より大きいです. 結果として (12) の右辺は 0 よりも大きくなります. よって,

$$0 \leq E[X^2] - (E[X])^2 \quad \Leftrightarrow \quad (E[X])^2 \leq E[X^2], \tag{13}$$

が一般的に成立することが分かります. (13) は実はコーシー・シュワルツの不等式の特殊な場合です.[17]

2.14 分散・モーメントの計算

前の小節では分散・モーメントとよばれる量を紹介しましたが, 何かしらの確率変数の期待値なので, 計算することは同じです. 例によってさいころの例で考えてみます. 出る目の確率が等しいさいころを投げ, その出目を X とします. 以前計算したように X の期待値は $E[X] = 7/2$ となります.

ここでは X の 2 次モーメントと分散を求めることにします. X の 2 次モーメントは X^2 の期待値のことでした. 今 X^2 は $1, 4, 9, 16, 25, 36$ の値を取る可能性があり, いずれになる確率も $1/6$ です. よって X の 2 次モーメント $E[X^2]$ は

$$E[X^2] = 1 \times \frac{1}{6} + 4 \times \frac{1}{6} + 9 \times \frac{1}{6} + 16 \times \frac{1}{6} + 25 \times \frac{1}{6} + 36 \times \frac{1}{6} = \frac{91}{6},$$

になります. 分散を計算するためには確率変数 $(X - E[X])^2$ の期待値を同じように計算すればよいです. ただ今は既に 2 次モーメントと平均が分かっているので, (12) を使うと簡単です. $E[X] = 7/2, E[X^2] = 91/6$ なので, 確率変数 X の分散は

$$E[X^2] - (E[X])^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12},$$

になります.

2.15 独立な確率変数の和の分散

確率変数 X の分散を $V[X]$ とかきました, つまり

$$V[X] = E[(X - E[X])^2],$$

[17] コーシー・シュワルツの不等式とは $|\vec{a} \cdot \vec{b}| \leq |\vec{a}||\vec{b}|$ のことでした.

です. X_1, X_2 を独立な確率変数とすると, 期待値の線型性より

$$E[X_1 + X_2] = E[X_1] + E[X_2]$$

が成立します. 更に, 分散についても次が成立します.

命題 2.1 X_1, X_2 を独立な確率変数とする. このとき次が成立する.

$$V[X_1 + X_2] = V[X_1] + V[X_2], \tag{14}$$

証明は後回しにして, 小節 4.1 で行います. ここで注意したいのは, (14) は期待値の線型性に似ているようにみえますが, 本質的に異なる性質であることです. 実際, 分散は二乗を計算する必要があるため, (14) は二つの確率変数が独立でない等の一般の場合には成立しません.

(14) は確率変数が独立でありさえすれば, 三つの和, 四つの和と確率変数が増えたとしても成立します. 例えば確率変数 X_1, X_2, X_3 が独立の場合には,

$$V[X_1 + X_2 + X_3] = V[X_1 + X_2] + V[X_3] = V[X_1] + V[X_2] + V[X_3],$$

となります. 一つ目の等号では $X_1 + X_2$ と X_3 に (14) を用いて, その次に再び (14) を用いています.

例としてさいころを独立に n 回投げることを考えます. 1 回目から n 回目の出目を順に $X_1, X_2, \ldots, X_n$ とかくことにします. また出目の総数を S_n, つまり,

$$S_n = X_1 + \cdots + X_n,$$

とします. また $X_1, \ldots, X_n$ の分散は全て等しく 35/12 なので, S_n の分散 $V[S_n]$ は

$$V[S_n] = V[X_1] + \cdots + V[X_n] = \frac{35n}{12},$$

になります.

2.16 確率変数の分布

確率変数の具体例として次の二つの例を考えたいと思います.

- 1 から 6 が書かれた出目の確率が等しいさいころを投げて出目を X とする.
- 1 から 6 が書かれた数字が出る確率が等しいルーレットを回して出た数字を Y とする.

X と Y は行なっている試行が違うので, 確率変数としては異なるものです. しかしながらどちらの確率変数も, 1 から 6 が出る確率は全て等しく 1/6 となっていて, どの数字が出るかという確率は全て等しいです. このように確率変数そのものよりも, その結果がどのようになるかの確率のみに興味があることがしばしばあります. このように「確率がどのようになっているか」ということを指して, 確率変数の**分布**や**法則**といいます. 離散型の確率変数であれば確率を全て列挙することによって分布を表現することができます. 上の例であれば X の分布は

$$P(X = i) = \frac{1}{6}, \quad i = 1, 2, 3, 4, 5, 6,$$

で与えられます. 勿論 X を Y に変えても同じなので, これをもって X の分布と Y の分布は等しい, といいます.

2.17　離散型確率分布の例

離散型の確率分布の有名な例を紹介します. 以下では X は確率変数とします.

- p を 0 から 1 の間の実数とする. パラメータ p の**ベルヌーイ分布**とは,[18]

$$P(X=1)=p, \quad P(X=0)=1-p.$$

- p を 0 から 1 の間の実数, n を自然数とする. パラメータ p, n の**二項分布**とは,[19]

$$P(X=k)={}_nC_k p^k(1-p)^{n-k}, \quad k=0,1,2,\ldots,n.$$

- p を 0 から 1 の間の実数とする. パラメータ p の**幾何分布**とは,

$$P(X=k)=p(1-p)^{k-1}, \quad k=1,2,3,\ldots$$

- n を自然数, $a_1, a_2, \ldots, a_n$ を全て異なる実数とする. **一様分布**とは,

$$P(X=a_k)=\frac{1}{n}, \quad k=1,2,\ldots,n.$$

急に色々と紹介されても分からないと思うので, 対応する試行の例もあげます.

- p を 0 から 1 の間の実数として, 表が出る確率が p のコインを一回投げる (裏が出る確率は $1-p$). 表が出たら $X=1$, 裏が出たら $X=0$ とおくと, X の分布はベルヌーイ分布.
- 上のコインを n 回独立に投げ, n 回目までに出た表の回数を S_n とすると, S_n の分布は二項分布.
- 上のコインを表が出るまで投げて, 初めて表がでるまでに投げたコインの数を H とする. H の分布は幾何分布.
- a_1 から a_n が書かれたくじがあり, それを一枚無作為にとる. かかれている数字を U とすると, U の分布は一様分布.

これらの分布は比較的単純ではありますが, 現象としても単純であるが故に, 多くの物事を説明するために使われます. 実際にはこれらが合わさった分布なんかもよく現れますが, 現象の複雑さに応じて分布の具体系は最早求められないくらいに複雑な形になっていきます.

[18] ここでは X がとる値を $0,1$ としていますが, 二つの値であればなんでもよいです.
[19] ${}_nC_k$ は二項係数 $n!/[k!(n-k)!]$.

2.18 独立同分布確率変数列

ここまで確率変数について色々触れてきましたが, 確率変数について色々知りたいと思ったときあまりに一般的なものを考えると, それでは物事の対象が大きすぎて具体的に性質を調べることはできません. そのために標題の「独立同分布確率変数」とよばれるものを中心に, 確率論は研究されてきました. 名前が少し長いですが細かく分ければどういうものを指すか理解できます. つまり**独立同分布確率変数**とはその名のとおり

- 「独立」かつ「同分布」な「確率変数 (の列)」

ということです. 例によってまずは具体例を考えます.

具体例として, 同じコイン (偏りがあってもよい) を独立に「無限に」投げ続ける試行を考えます. n 回目のコイン投げで表ならば $X_n = 1$, 裏ならば $X_n = 0$ とすると, ランダムな数字の列 $X_1, X_2, X_3, \ldots$ が得られます. このようなに確率変数の列が得られますが, これらは問題の設定上全て独立です. さらに同じコインを投げているので表が出る確率はどの回も変わりません。同分布とはこのことを指しています. 一つの具体例としてはこのような列 $X_1, X_2, X_3, \ldots$ のことになります.

以上の例からも分かるように独立同分布確率変数列を簡単に述べるとすると,

- 「独立に同じ試行を繰り返し行なった結果, 得られるランダムな数字の列」

といえます. 英語では independent and identically distributed random variable なので, 単に i.i.d. r.v. とか IID 列とかかかれます. このように同じ試行を何度も繰り返すだけですが, 同じ試行であれば何を行なってもよいということに注意して下さい. 例えば, コインを投げ続け, さいころを投げ続ける, (少し理想的に考えなければいけませんが) 宝くじを毎日引くなど, 同じ試行を繰り返す限りは何を行なっても問題はありません. このような確率変数の列に対して, 公理的確率論では大数の法則や中心極限定理といったかなり一般的な定理を示すことができます. また大数の法則なんかはモンテ・カルロ法とよばれる数値計算の手法とも密接に関係していて, IID 列に対する研究は純粋数学の枠を遥かに超えて応用されています.

詳しくは取り扱いませんが実はここで数学的に大きな問題に当たっています. それは何かというと,

- 「独立同分布確率変数の列は存在するか？」

ということです. 確率を考えるためには確率空間を設定する必要があると説明しました. このことはとりうる結果が有限の場合には問題がありません. しかしながら今考えたいのは「無限回の」コイン投げの試行です. 何も気にしなければそんなものは出来て, 結果を $X_1, X_2, X_3, \ldots$ とかいてそれらは独立かつ同分布, とすればいいわけですが, 数学的に本当にそのような列が存在するかは示す必要があります. このような列の存在は高校数学の知識では残念ながら示せないですし, 古典的確率論では (恐らく) 解決できない話です.[20] 公理的確率論はそのような問題も正確に考えるために発展してきたという側面もあります. 長々かきましたが, こういう点が実は問題がある, という程度の認識でよくて, 無限回のさいころを投げることは数学的にできるという感覚で勿論大丈夫です.

[20] コイン投げ程度であればそこまで難しくなくできますが.

3　大数の法則とランダムウォーク

3.1　大数の法則

大数の法則とよばれるものを考えますが, 一般的な場合について述べる前に, いつものようにさいころ投げの場合で考えたいと思います. 1 から 6 の目がかかれているさいころを順に投げ, 出た目を順に $X_1, X_2, X_3, \ldots$ とかくことにします. また各回の出目は独立とします. 先に紹介した言葉でいうと, $X_1, X_2, X_3, \ldots$ は独立同分布確率変数列になっています. $k = 1, 2, 3, \ldots$ に対して X_k はさいころを一回投げたときの出目なので, $E[X_k] = 7/2$ です.

イメージを沸かせるためにさいころを百回投げてみます. さいころの出目は順に以下のようになりました.[21]

$$\begin{aligned}&4,3,1,1,1,4,1,6,4,6,3,3,6,3,5,3,5,2,5,5,6,2,6,5,5,3,5,1,6,4,\\&5,3,6,5,3,5,2,5,3,5,3,1,5,2,2,1,5,1,3,4,3,3,6,4,5,6,4,3,5,3,\\&5,6,2,4,3,3,5,2,4,5,2,2,5,2,2,4,5,4,5,2,2,6,3,5,4,1,2,5,4,1,\\&3,6,2,1,1,3,4,4,1,5.\end{aligned}$$

それではここで期待値 $E[X_1] = 7/2 = 3.5$ について考えてみます. 出た目だけお金がもらえるとすると, 一回では大体 3.5 だけもらえると思えます. しかしながら実際には 1 から 6 の整数値しかもらえないので, この考えはあまりにも感覚的です. これを正確に考えるために, 10 回, 50 回, 100 回目までに出た数字の平均を考えてみます. つまり, n 回目までの出目の総数を S_n,

$$S_n = X_1 + \cdots + X_n,$$

とし, S_n を n で割った S_n/n を $n = 10, 50, 100$ に対して上の場合に書いてます. 結果は順に,

$$\frac{31}{10} = 3.1, \quad \frac{181}{50} = 3.62, \quad \frac{358}{100} = 3.58,$$

となっていて, n が増えるにつれどんどん 3.5 に近づいているように見えます. この 3.5 に近づいているという現象は, 実際に数学の定理として示す事ができます. より正確に述べるとすると, 確率 1 で次が成立します.

$$\lim_{n\to\infty} \frac{S_n}{n} = E[X_1]. \tag{15}$$

これは**大数の強法則**とよばれる確率論における定理の一つになります.

(15) に対する注意を幾つか述べます. はじめに, S_n/n はさいころを実際に n 回投げない限り具体的な値が定まらないランダムな数です. 一方極限の $E[X_1]$ は 3.5 という決まった数になっています. 次の注意ですが, (15) はさいころ投げに限らず, 期待値が定まる独立同分布な確率変数列であれば成立する, という一般的な点です. 大数の強法則というくらいなので, 別に**大数の弱法則**とよばれるものも存在します.

[21] 実際はさいころを投げずにコンピュータで 1 から 6 の一様乱数を百個発生させています.

定理 3.1 独立同分布な確率変数列 $X_1, X_2, \ldots$ に対して, $S_n = X_1 + \cdots + X_n$ とおく.[22] このとき任意の $\varepsilon > 0$ に対して次が成立する.

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - E[X_1]\right| \geq \varepsilon\right) = 0. \tag{16}$$

大分難しくなりますが, 示すことは一応可能なので証明を小節 4.2 に書いておきました. 「弱法則」, 「強法則」と字からもわかるように強法則の方が強い結果になります.[23] 大数の強法則は大数の弱法則を導くため, 大数の弱法則に意味がないかというと, そうではありません. 大数の弱法則の証明では, 次の不等式を示します.

$$P\left(\left|\frac{S_n}{n} - E[X_1]\right| \geq \varepsilon\right) \leq \frac{V[X_1]}{n\varepsilon^2}. \tag{17}$$

(17) をみてみると, 左辺は計算したい確率ではありますが, 一般的には複雑すぎて n の式で具体的に表すことは基本的に無理です. しかしながら (17) は左辺の確率がどのくらい小さいかということを具体的に教えてくれます. このような確率の評価はしばしば有用であり, 数値計算などでは近似の精度を与えるものであることから, 理論的にも応用的にも意味のあるものになっています.

3.2 モンテカルロ法

モンテカルロ法とはシミュレーションや数値計算を乱数を用いて行う手法の総称です. モンテカルロとはモナコ公国のある地域のことで, この地域ではカジノが有名であることからこの名前がつけられています. モンテカルロ法について具体的な例を通じて考えていきます.

モンテカルロ法の典型的な例として, 円周率の近似的な計算があります. その計算のために, 一辺の長さが 2 の正方形 S とそれに内接する一辺の長さが 1 の円 C を考えます. ここで正方形 S の中から一様に点をサンプルします. 図ではサンプルされた点が円の内部にあれば赤で, 円の外部にあれば青で示しています. S からのサンプルを独立に何度も繰り返します. 図からもわかるように, サンプルする点の数が増えれば増えるほど円の中の点はどんどん密集していきます.

さてここで確率変数 $X_1, X_2, X_3, \ldots$ を次のように定義します. はじめに, 得られた点を順に $P_1, P_2, P_3, \ldots$ とかくことにします. $k = 1, 2, 3, \ldots$ 回目のサンプル P_k に対して,

$$X_k = \begin{cases} 1, & P_k \text{ が円 C の内部にあるとき}, \\ 0, & P_k \text{ が円 C の外部にあるとき}, \end{cases}$$

とします. いつものように

$$S_n = X_1 + \cdots + X_n,$$

[22] $V[X_1] < \infty$ であることは仮定しておきます.

[23] 数学でいう「強い」とは, 前者は後者から演繹できるという意味です.

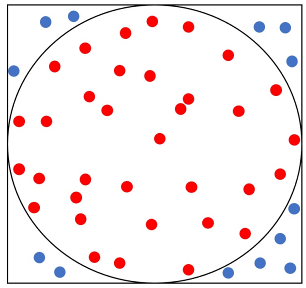


Figure 2: 赤は円の内部からのサンプル. 青は円の外部からのサンプル.

とおくと, S_n は n 回目までのサンプルで得られた赤の点の総数になっています.

$X_1, X_2, \ldots$ は独立同分布な確率変数の列です. 先に説明した大数の強法則から, 確率 1 で次が成立します.

$$\lim_{n\to\infty} \frac{S_n}{n} = E[X_1]. \tag{18}$$

注意したいところは, 左辺の S_n/n はランダムではありますが, 実際に S からの一様サンプルを行うことにより計算できるところです.

まだ $E[X_1]$ を求めていないのでそれを計算します. X_1 は 1 と 0 の値しかとらないことと期待値の定義から,

$$\begin{aligned} E[X_1] &= 1 \times P(P_1 \text{ が円 C の内部にある}) + 0 \times P(P_1 \text{ が円 C の外部にある}) \\ &= P(P_1 \text{ が円 C の内部にある}) = \frac{\text{円 C の面積}}{\text{正方形 S の面積}} = \frac{\pi}{4}. \end{aligned} \tag{19}$$

三つ目の等号では P_1 が正方形 S からの一様サンプルであることを用いています. 以上の計算 (18), (19) をまとめると, n が十分に大きければ高確率で

$$\frac{4S_n}{n} \fallingdotseq \pi,$$

となっていて, 先にも注意したように左辺の量は具体的に計算できるため, 円周率 π の値を近似的に求めることができたことになります.

3.3 ランダムウォーク

やっと標題の**ランダムウォーク**とよばれる, 確率論の代表的な問題を考えます. 知っている人もいると思いますが, 次のように非常に単純な問題です.

- $0 \le p \le 1$ を固定し, 時刻 0 で原点を出発する点を考える. 表が出る確率が p, 裏が出る確率が $1-p$ のコインを投げ, 表ならば $+1$ だけ動き, 裏ならば -1 だけ動く. コインを

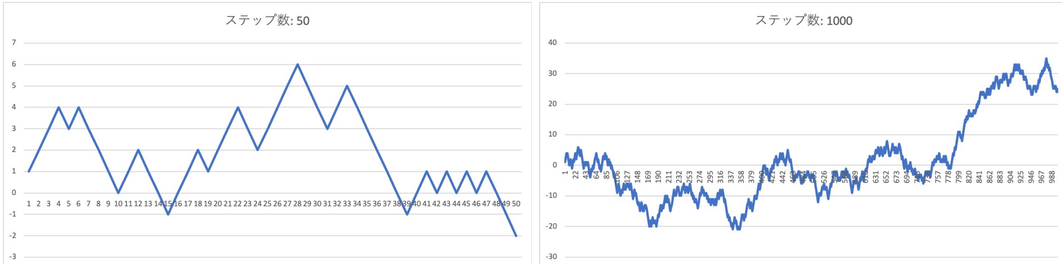


Figure 3: $p = 1/2$ のランダムウォーク

投げる試行を各回独立に行い, 同じ手順に沿って毎回点を動かす. 時刻 $n = 0, 1, 2, \ldots$ における点の位置を S_n とかく. S_n は**単純ランダムウォーク**とよばれる. $p = 1/2$ のときランダムウォークは**対称**であるとよばれる.

ランダムウォークは日本語では**乱歩**, もしくは**酔歩**とよばれます. ランダムウォークは非常に簡単なモデルでありますが, それゆえに色々な問題が交差している重要なものになります. 例えば物理学では点の移動は粒子の運動と考えられ, 経済学では株価の価格の変動などと考えられます. またランダムウォークはマルコフ性をもっているため, 整数格子 $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ 上のマルコフ連鎖になっています. その推移確率は任意の整数 x に対して,

- x から $x + 1$ に移動する確率 ··· p,
- x から $x - 1$ に移動する確率 ··· $1 - p$,

により与えられます. また単純とは $\mathbb{Z}$ の隣の点のみに移動することを指します. 単にランダムウォークというと単純でない場合も含みますが, ここでは単純なもののみを考えます.

ランダムウォークは $\mathbb{Z}$ 上のマルコフ連鎖であるだけでなく, 独立同分布な確率変数列の和になっています. 実際, $X_1, X_2, \ldots$ を独立同分布な確率変数で

$$P(X_1 = +1) = p, \quad P(X_1 = -1) = 1 - p,$$

となるものとします. このとき $S_0 = 0, n = 1, 2, \ldots$ に対して

$$S_n = X_1 + \cdots + X_n,$$

とおくと, S_n は単純ランダムウォークに他なりません. これにより大数の法則が成立します. 今 X_1 は次のように計算できます.

$$E[X_1] = 1 \times p + (-1) \times (1 - p) = 2p - 1.$$

よってランダムウォーク S_n に対して次が成立します.

- (大数の強法則) 確率 1 で次が成立する.

$$\lim_{n\to\infty} \frac{S_n}{n} = 2p - 1.$$

大数の法則から n が大きければランダムウォークは大体

$$S_n \fallingdotseq (2p-1)n,$$

の位置にいることが分かります. よって時刻 n が経つにつれて, $p < 1/2$ のときは $E[X_1] < 0$ なので S_n は負の方向に進み, $p > 1/2$ のときは $E[X_1] > 0$ なので S_n は正の方向に進むことがわかります. $p = 1/2$ のときは $E[X_1] = 0$ なので大数の法則ではあまり情報が得られません.

3.4 再帰性・非再帰性

引き続きランダムウォーク S_n を考えます. 天下りになりますが, 次の問題を考えます.

- ランダムウォーク S_n は原点 0 に無限回戻ってくるか?

問題が突飛すぎる気もしますが, ランダムウォークを賭けのモデル (表ならば 1 円もらえ, 裏ならば 1 円支払う) と考えると, 原点に戻ってくるかどうか (破産するかどうか) ということは気になる問題です. また株価の変動などにおいても, ある決めた価格を下回るかどうかということは, 考えたい基本的な問題です.

ランダムウォークが原点に無限回戻ってくるという事象を, R でかくことにします (R は return の頭文字), つまり,

$$R = \{S_n = 0 \text{ となる } n = 0, 1, 2, \dots \text{ が無限個存在する}\},$$

ということです. このように R をおくと, 問題は R の確率 $P(R)$ の値はいくつか, ということになります. ここで再帰的・非再帰的ということを次で定義します.

- $P(R) = 1$ のとき, ランダムウォークは**再帰的**であるという.
- $P(R) < 1$ のとき, ランダムウォークは**非再帰的**であるという.[24]

この言葉を用いるとランダムウォークに対して次の問題が考えられます.

- ランダムウォーク S_n は (p がどの場合に) 再帰的・非再帰的のいずれかであるか.

この問題に対して次の定理が知られています.

定理 3.2 $p \neq 1/2$ のとき非再帰的であり, $p = 1/2$ のとき再帰的である.

[24]非再帰的は**過渡的**ともよばれます

証明は大分難しいですが, 興味ある人のために小節 4.3 に書いておきました. 実は既にみているように, $p \neq 1/2$ の場合には大数の法則からランダムウォークは非再帰的であることが分かります. 実際, 大数の法則から n が大きければ

$$S_n \fallingdotseq (2p-1)n,$$

であるので, S_n は大体, $p < 1/2$ のときは直線的に負の方向に進み, $p > 1/2$ の直線的に正の方向に進みます. よって, 確率 1 で n が大きければ原点に戻ってくることはないので, 特に $P(R) = 0$ が分かり, $p \neq 1/2$ のときには非再帰的であることが分かります. 一方 $p = 1/2$ のときは大数の法則からは再帰的か非再帰的か判別できません. そのため小節 4.3 で行っているような詳細な評価を行う必要があります.

3.5 二次元・三次元ランダムウォークの再帰性

これまで整数格子 $\mathbb{Z}$ 上のランダムウォークを考えましたが, 最も近い整数格子点に等確率で移動する, というマルコフ連鎖を考えることにより, 二次元整数格子 $\mathbb{Z}^2$ や三次元整数格子 $\mathbb{Z}^3$ 上でも同じ問題が考えられます.[25] ここで, 整数格子はそれぞれ

$$\mathbb{Z}^2 = \{(x,y); x, y \text{ は整数 }\},$$
$$\mathbb{Z}^3 = \{(x,y,z); x, y, z \text{ は整数 }\},$$

で定義されます. それぞれ (x,y), (x,y,z) から次の点への移動は

- $(x \pm 1, y), (x, y \pm 1)$ の四つを等確率 $1/4$,
- $(x \pm 1, y, z), (x, y \pm 1, z), (x, y, z \pm 1)$ の六つを等確率 $1/6$,

で選び, 選ばれた点に移動することで定義されます. $\mathbb{Z}$ の場合と同様に $\mathbb{Z}^2, \mathbb{Z}^3$ の場合にも原点に無限回戻ってくるかどうかで再帰性が定義されます. この問題では二次元の場合には再帰的であり三次元の場合には非再帰的であることが知られています.[26]

4 補遺：命題の証明

4.1 命題 2.1 の証明

分散の定義に従って計算をすると,

$$\begin{aligned} V[X_1 + X_2] &= E\left[(X_1 + X_2 - E[X_1 + X_2])^2\right] \\ &= E\left[((X_1 - E[X_1]) + (X_2 - E[X_2]))^2\right] \\ &= E[(X_1 - E[X_1])^2] + 2E[(X_1 - E[X_1])(X_2 - E[X_2])] + E[(X_2 - E[X_2])^2]. \end{aligned}$$

[25]勿論それ以上の次元の整数格子上でも考えらます.

[26]次元が 3 以上であれば, 対称単純ランダムウォークは非再帰的であることが知られています.

複雑な式変形をしているように見えるかもしれませんが, 最初の等号は分散の定義, 二つ目の等号は期待値の中身を後のために整理して, 三つ目の等号は $(a+b)^2 = a^2 + 2ab + b^2$ を行なっているだけです. 分散の定義を思い出すと, 最後の式の一項目は $V[X_1]$, 三項目は $V[X_2]$ であるので, 残りは (14) を示すためには二項目が 0 になっていることを見ればよいです. 実際, X_1 と X_2 が独立であったので, $X_1 - E[X_1]$ と $X_2 - E[X_2]$ の期待値の積はそれぞれの期待値の積にわかれます. よって

$$E[(X_1 - E[X_1])(X_2 - E[X_2])] = E[(X_1 - E[X_1])] \times E[(X_2 - E[X_2])], \tag{20}$$

となりますが, 期待値の線型性から

$$E[(X_1 - E[X_1])] = E[X_1] - E[E[X_1]] = E[X_1] - E[X_1] = 0,$$

となります. ただし二つ目の等号では $E[E[X_1]] = E[X_1]$ であることを用いています. 以上により示したかった二項目が 0 であることがわかるので, (14) が示されました.

4.2 定理 3.1 の証明

この小節では定理 3.1 を示します. そのためには**チェビシェフの不等式**とよばれる次の不等式が必要なので, それを最初に示します.

命題 4.1 確率変数 X と $a > 0$ に対して次の不等式が成立する.

$$P\left(|X - E[X]| \geq a\right) \leq \frac{V[X]}{a^2}. \tag{21}$$

証明に入る前にまずは意味することを幾らか読み解きます. まず (21) の左辺が何を意味しているかみてみます. $E[X]$ とは確率変数 X の期待値でした. なので X は大体 $E[X]$ くらいだろうと (平均的には) 考えているわけです. なのでここではどのくらい平均値からずれているかを見ています. P の中身の $|X - E[X]|$ とはまさしくそのずれのことです. 以上を踏まえると (21) の左辺は

- 「確率変数 X が期待値 $E[X]$ から a だけずれている確率」

と読むことができます.[27] 一方右辺は X の分散 $V[X]$ を a^2 で割った数になっています. こちらにはあまり意味はないので, ひとまずそういう数だとそのまま受け入れて欲しいです.

それではチェビシェフの不等式を証明します.

チェビシェフの不等式の証明

[27] 高校生の時にも慣れなかったと思いますが, 絶対値の記号 $|a - b|$ は a と b の距離を表しています.

はじめに確率変数 $Y, Z, \widetilde{Z}$ をそれぞれ次のように定義します.

$$
\begin{aligned}
Y &= \begin{cases} 1, & |X - E[X]| \geq a \text{ のとき}, \\ 0, & \text{そうでないとき}, \end{cases} \\
Z &= \begin{cases} \dfrac{|X - E[X]|^2}{a^2}, & |X - E[X]| \geq a \text{ のとき}, \\ 0, & \text{そうでないとき}, \end{cases} \\
\widetilde{Z} &= \frac{|X - E[X]|^2}{a^2}.
\end{aligned}
$$

このように $Y, Z, \widetilde{Z}$ を定義して, 次のことに注意します.

1. $E[Y] = P(|X - E[X]| \geq a)$.
2. $|X - E[X]| \geq a \Leftrightarrow \dfrac{|X - E[X]|}{a} \geq 1 \Leftrightarrow \dfrac{|X - E[X]|^2}{a^2} \geq 1$.
3. 2 からとくに $Y \leq Z$. また明らかに $Z \leq \widetilde{Z}$. よって $Y \leq \widetilde{Z}$.

3 と期待値の単調性から, $E[Y] \leq E[\widetilde{Z}]$ が成立します. よって 1 から,

$$
\begin{aligned}
P\left(|X - E[X]| \geq a\right) &= E[Y] \\
&\leq E[\widetilde{Z}] \\
&= \frac{E[|X - E[X]|^2]}{a^2} = \frac{V[X]}{a^2},
\end{aligned}
$$

となりチェビシェフの不等式が示されました.

最後にチェビシェフの不等式をを用いて大数の弱法則を証明したいと思います.

大数の弱法則の証明

$X_1, X_2, \ldots$ を独立同分布確率変数列で, X_1 の分散 $V[X_1]$ が有限であるものとします. またいつものように $m = E[X_1]$,

$$S_n = X_1 + \cdots + X_n,$$

とおきます. このとき次を示します.

$$\lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) = 0. \tag{22}$$

(22) を示すために後では,

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{V[X_1]}{n\varepsilon^2}, \tag{23}$$

を示します.[28] (23) が一旦示されると, 左辺が 0 以上であるので $n \to \infty$ とすると右辺が 0 に収束するので, はさみうちの原理から (22) が示されます.

それでは (23) を示します. 初めに, P の中身の式に n をかけることにより

$$\left|\frac{S_n}{n} - m\right| \geq \varepsilon \quad \Leftrightarrow \quad |S_n - nm| \geq n\varepsilon,$$

であることに注意します. また,

$$E[S_n] = E[X_1] + \cdots + E[X_n] = nm,$$

です, ここでは $X_1, X_2, \ldots$ が同分布であることから $k = 1, 2, 3, \ldots$ に対して $E[X_k] = E[X_1] = m$ となることを用いています. よって (22) の確率は

$$P(|S_n - E[S_n]| \geq n\varepsilon),$$

とかけることが分かります. よってチェビシェフの不等式を $X = S_n, a = n\varepsilon$ の場合に用いると

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{V[S_n]}{n^2\varepsilon^2}, \tag{24}$$

がわかります. 残るは右辺の分散 $V[S_n]$ の計算ですが, $X_1, X_2, \ldots$ が独立であることから S_n の分散は和の分散にわかれます, つまり,

$$V[S_n] = V[X_1] + \cdots + V[X_n] = nV[X_1], \tag{25}$$

となります. 再び $X_1, X_2, \ldots$ が同分布であることから $k = 1, 2, 3, \ldots$ に対して $V[X_k] = V[X_1]$ となることを用いています. よって (24), (25) から (23) が成立することが分かります. これで定理 3.1 が示されました.

4.3 $p = 1/2$ のときの再帰性

再帰性を考えるために, 時刻 0 以降に初めて S_n が原点に戻る時刻を T_1 とかくことにします. T_1 はランダムな時刻であることに注意です. この段階では原点を出発した後に一度も戻ってこないこともあり得るので, T_1 は ∞ になることもあり得ます. なので $P(T_1 < \infty)$ は原点に少なくとも一度戻ってくる確率を表しており, とくに**再帰確率**とよばれます. さらに $n = 2, 3, \ldots$ に対しても n 回目にランダムウォークが 0 に戻る時刻を T_n とかくことにします.

初めに次の命題を示します.

命題 4.2 次の三条件は全て同値.

[28] この式からも分かるように実は $V[X_1] < \infty$ はここでは仮定しています. これを仮定せずとも示せますが, 証明は遥かに難しくなるので, ここではこの仮定の下で示しています.

(1) $P(T_1 < \infty) = 1$ (確率 1 で少なくとも一度は原点に戻る).

(2) $P(R) = 1$ (確率 1 で無限回原点に戻る, つまり, 再帰的).

(3) $\displaystyle\sum_{n=1}^{\infty} P(S_n = 0) = \infty$.

命題 4.2 の証明
はじめにマルコフ性よりランダムウォークが原点に戻ってくると, そのあとは再び独立にランダムウォークにより動くことに注意します. よって

$$\begin{aligned} P(T_n < \infty) &= P(1\text{ 回原点に戻ってきた後に } n-1 \text{ 回原点に戻ってくる}) \\ &= P(1\text{ 回原点に戻ってくる})P(n-1\text{ 回原点に戻ってくる}) \\ &= P(T_1 < \infty)P(T_{n-1} < \infty) \end{aligned}$$

が成立します.[29] これを繰り返し用いることにより,

$$P(T_n < \infty) = P(T_1 < \infty)^n \tag{26}$$

が成立することがわかります.

次にランダムウォークが原点に戻ってくる回数 N に注目します. $n = 1, 2, \ldots$ に対して, 確率変数を次のように二つ定めます.

$$\mathbf{1}\{S_n = 0\} = \begin{cases} 1, & S_n = 0 \text{ のとき}, \\ 0, & \text{そうでないとき}, \end{cases}$$

$$\mathbf{1}\{T_n < \infty\} = \begin{cases} 1, & T_n < \infty \text{ のとき}, \\ 0, & \text{そうでないとき}. \end{cases}$$

これらの確率変数を用いると, ランダムウォークが原点に戻ってくる回数 N は次のようにかけます.

$$N = \sum_{n=0}^{\infty} \mathbf{1}\{S_n = 0\} = \sum_{n=0}^{\infty} \mathbf{1}\{T_n < \infty\}$$

これらの式で期待値をとると

$$E[N] = \sum_{n=0}^{\infty} P(S_n = 0) = \sum_{n=0}^{\infty} P(T_n < \infty) = \frac{1}{1 - P(T_1 < \infty)} \tag{27}$$

となります. ただし最後の等式では (26) を用いています.

[29]実はここでは強マルコフ性とよばれる, 中々に難しい性質を用います.

(26) と (27) より命題が従います. 実際, (1) であれば (26) より任意の $n = 1, 2, \ldots$ に対して $P(T_n < \infty) = 1$ となり, 確率 1 でランダムウォークは無限回原点に戻ります. 次に (2) とすると N は確率 1 で無限大なので (27) の最左辺は無限大になり (3) がわかります. 最後に (3) であれば再び (27) より $P(T_1 < \infty) = 1$ がわかります. これで命題 4.2 が示されました.

命題 4.2 をまとめると

- $P(T_1 < \infty) = 1 \quad \Leftrightarrow \quad$ 再帰的 $\quad \Leftrightarrow \quad \displaystyle\sum_{n=1}^{\infty} P(S_n = 0) = \infty$,
- $P(T_1 = \infty) > 0 \quad \Leftrightarrow \quad$ 非再帰的 $\quad \Leftrightarrow \quad \displaystyle\sum_{n=1}^{\infty} P(S_n = 0) < \infty$,

ということになります. つまり再帰的であるか非再帰的であるかは最後の和が収束しているか発散しているかで決まります.

再帰的か非再帰的かを判定するために, 最後の和を計算します. $P(S_n = 0)$ は時刻 n で原点にいる確率なので, 具体的に求まります. n が奇数のときには常に原点にいないのでこのときには $P(S_n = 0) = 0$ です. n が偶数, つまり $n = 2k$ のときには, $S_n = 0$ となるのは, $2k$ 回コインを投げて表が k 回裏が k 回でるときなので,

$$P(S_{2k} = 0) = \frac{(2k)!}{(k!)(k!)} \times \left(\frac{1}{2}\right)^k \times \left(\frac{1}{2}\right)^k = \frac{(2k)!}{(k!)(k!)} \left(\frac{1}{4}\right)^k,$$

になります. よってスターリングの公式[30]

$$\lim_{k \to \infty} \frac{\sqrt{2\pi k} k^k e^{-k}}{k!} = 1,$$

を用いると,

$$\begin{aligned} \sum_{n=1}^{\infty} P(S_n = 0) &= \sum_{k=1}^{\infty} P(S_{2k} = 0) = \sum_{k=1}^{\infty} \frac{(2k)!}{(k!)(k!)} \left(\frac{1}{4}\right)^k \\ &\fallingdotseq \sum_{k=1}^{\infty} \frac{\sqrt{4\pi k}(2k)^{2k} e^{-2k}}{2\pi k k^{2k} e^{-2k}} \left(\frac{1}{4}\right)^k = \frac{1}{\sqrt{\pi}} \sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} = \infty, \end{aligned}$$

であることが分かります. 上の計算で $\fallingdotseq$ のところは幾分雑ですが, 真面目にやればいいだけで何も問題ではありません. 最後の和が発散することは, 面積を考えることにより

$$\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} \geq \int_1^{\infty} \frac{1}{\sqrt{x}} dx = \left[2x^{1/2}\right]_1^{\infty} = \infty,$$

であることから従います. 以上により $p = 1/2$ のとき対称単純ランダムウォークが再帰的であることが示されました.

[30] これはつまり $k!$ は k が大きいと $\sqrt{2\pi k} k^k e^{-k}$ くらいということをいっています.