数理解析特論

服部哲弥　1995.4–1995.6

1　バスはなぜ来ないか　–　平均（期待値）について　（４月２０日）

たいていのバスは時刻通り来ない．いくつかの大都市では，バス停に電気仕掛けで次のバスがどれくらい近くにいるかを表示している．時刻表通り来るならばこのような表示はいらないはずである．それほどバスは時刻通り来ない.（頑張っているところもある．アメリカワシントン州シアトルのバスはちゃんと来た.）遅れや進みが極端になってバスが 2 台以上つながって来ることすら多い.（先に来たバスに乗客が多く乗るので，一度バスの間隔が等間隔でなくなるとその傾向はどんどんひどくなる，という面もあるが，この問題には立ち入らない.）

場所によっては停留所に，時刻でなく，例えば「10 分に 1 本」という書き方をしている．ちょっと頭をひねって「10 分に 1 本ならば，平均すると 5 分待てば次のバスが来る」と考えて待つと，ちっとも来ない.（この計算の根拠はすぐ次に示す.）遅れたり早くなったりするのは，交通事情もあってやむを得ないとも思う．しかし，単に時刻通り来ないだけでなく，時刻表よりもバスの来る量が少ない気がすることも多いのはどういうわけだろう．時刻表より実際の運行回数が少ない「間引き」が行われていると即断していいのだろうか？

100 分の間に合計 10 本のバスを走らせる路線があったとする．平均的には 10 分間隔で運行していることになる．平均運行間隔 10 分である．乗客はこの停留所に勝手な時刻にやって来てバスを待つとする．このとき，この停留所で待つ乗客は平均何分待てば乗れるだろうか？

まず，バスが等間隔に来る場合を考える．10 分ごとに 1 本バスが来る．乗客が停留所に到着してから次のバスが来るまでに待つ時間 Z は 0 分（次のバスの来る直前＝幸運な場合）から 10 分（前のバスが行った直後＝悪夢の場合）までの可能性がある．Z のばらつきは乗客には予測・制御のできない，確率変数であると考える．Z の値は $0 < Z < 10$ の範囲に分布することになる．この分布（乗客のバス停到着時刻の分布）が一様（等確率）とすると，平均待ち時間 $\mathrm{E}[\,Z\,]$（確率変数 Z の期待値を $\mathrm{E}[\,Z\,]$ と書く）は

$$\mathrm{E}[\,Z\,] = \int_0^{10} t\,dt/10 = 5$$

即ち，

平均運行間隔 10 分のバス停において，乗客の到着時刻の分布が一様分布のとき，この乗客の平均待ち時間 $\mathrm{E}[\,Z\,]$ は 5 分である．

次にバスが等間隔に来ない場合を考える．先ほどと同様に，平均運行間隔 10 分の路線を考える．交通事情や乗客の具合で，ある停留所に来たときには等間隔でないとする．それでも，平均運行間隔 10 分，つまり 100 分の時間幅の間に最初から最後まで 10 台のバスが通るとする．このときこの停留所で待つ乗客は平均 5 分待てば乗れるだろうか？

答えは，一般には平均 5 分より長く待たないといけない．これを理解するために，極端な場合を考える．10 本のバスが全部一斉に 100 分目に到着したとしよう．平均運行間隔は $100/10 = 10$ 分である．乗客は t 分目に到着したとすると，待ち時間は $100 - t$ である．t は 0 から 100 の間のどの時刻も同じ確率でとるとすると，平均は

$$\int_0^{100} (100-t)\,dt/100 = 50$$

つまり，

バスが 100 分ごとに 10 台数珠繋ぎになって来るバス停においては，平均運行間隔は 10 分だが，到着時刻の分布が一様分布の乗客の平均待ち時間は $\mathrm{E}[\,Z\,] = 50$ 分である．

バスは平均 10 分に 1 本走っているのに，乗客は平均 50 分も待たないといけない！

なぜ，平均運行間隔が等しくても平均待ち時間が 5 分から 50 分まで違うのか？バスの運転間隔が狭いと一つ前のバスを逃してもすぐ次が来る．しかし，間隔が狭いということは偶然その時間帯に乗客が停留所に到着する可能性も低いということである．より多くの場合，乗客は間隔のあいた時間帯に到着するから，平均より長く待つ可能性の方が高いことになる．

平均運行間隔が同じ二つのバス路線を比べた場合，（大雑把に言うと）等間隔に近い走り方をしているほど平均待ち時間は短く，等間隔からずれているほど平均待ち時間は長い．

さていよいよ，バスが全くでたらめに運行している場合を考える．でたらめ，というのは，運行間隔（あるバスが通ったあと次のバスが来るまで）T が（道路事情などで）ばらつくので（正の値を取る）確率変数であるということである．どんなふうにでたらめか，ということを定義する必要があるが，数学的に最も「単純」とされるのは次のように定義される．（書くと「単純」に見えないかも知れないが，いろいろな計算が可能なことが知られているので実用上も非常に利用される．）

「運行間隔を表す確率変数 T の分布の密度が指数分布

$$\lambda \exp(-\lambda t), \quad t > 0$$

で与えられる．λ（ラムダ）は平均運行間隔の逆数．しかも，T は注目したバスより前のバス達がどんな間隔で通ったかということと独立である.」（このとき，時刻 t におけるバスの累積到着台数（時刻 t までに何台通ったか）$X(t)$ が平均運行間隔 $1/\lambda$ の Poisson（ポワッソン）確率過程になっているという．確率過程という言葉は講義が進んだ段階で説明したい.）

分布の密度という言葉の意味は，t_0 と t_1 を $0 < t_0 < t_1$ を満たす定数とするとき $t_0 < T < t_1$ となる確率 $\mathrm{Prob}[\, t_0 < T < t_1 \,]$ が

$$\mathrm{Prob}[\, t_0 < T < t_1 \,] = \int_{t_0}^{t_1} \lambda \exp(-\lambda t)\, dt\,, \tag{1.1}$$

で計算されるということである．

（二つ以上の確率変数が）独立というのは，どういう値を取るかが互いに関係がないことを表す確率論の用語である．確率変数 $X_1, X_2, \cdots,$ が独立とは $Prob[X_1 \in A_1,\ X_2 \in A_2,\ \cdots] = \prod_i Prob[X_i \in A_i]$ ということ．

今の例では平均運行間隔が 10 分であったから $\lambda = 1/10$．このとき，乗客の平均待ち時間は $1/\lambda = 10$ となる（計算は後述）．即ち平均運行間隔に等しい．

平均運行間隔 10 分で，累積到着台数が Poisson 確率過程になっているバスに対して，乗客の平均待ち時間は $\mathrm{E}[\, Z \,] = 10$ 分である．

このことを逆に見ると，

もし，乗客の平均待ち時間と平均運行間隔が一致しないならば，バスの到着は Poisson 確率過程になっていない．それは，バスの運行間隔が「全くのでたらめではない」ことを意味する．

先ほどの例と合わせて考えると，平均待ち時間が短い場合は，バスが「でたらめに来る」場合（到着が Poisson 確率過程）に比べて等間隔に近い（つまりバスの定時運行努力が実っている）ことになり，平均待ち時間が長い場合は，バスが数珠繋ぎになったり間隔が空いたりが極端になっていると予想される（利用が多く，乗降ののべ時間が長い場合にはバスが団子状態になりやすい，と考えられる）．

定義式 (1.1) から，乗客の平均待ち時間が $1/\lambda$ になることを導いておく．

乗客の平均待ち時間 $= 1/\lambda$ になること の証明． ある特定のバスに乗れるケースに注目する．それは直前のバスが通り過ぎてから問題のバスが到着する直前までに乗客が停留所に来た場合である．仮に乗客は直前のバスが通り過ぎてから時間 s たって停留所に着いたとしよう．ということは，問題のバスと直前のバスは間隔 T が s 以上だということである．この条件の下で待ち時間 $T - s$ の期待値 $\mathrm{E}[\, T - s \mid T > s \,]$ を計算すればよい．（$\mid T > s$ は $T > s$ という条件付きで期待値を取ることを表す.）T の分布の密度 (1.1) から

$$\mathrm{E}[\, T - s \mid T > s \,] = \frac{\int_s^\infty (t-s)\lambda \exp(-\lambda t)\, dt}{\int_s^\infty \lambda \exp(-\lambda t)\, dt}\,.$$

（条件 $T>s$ がついているので積分範囲が $t>s$ となる.）これを計算するには分母分子とも $t=u+s$ という積分変数変換をするのがよい．すると定数 $\lambda\exp(-s\lambda)$ が分母分子で打ち消して

$$\mathrm{E}[\,T-s\mid T>s\,]=\frac{\int_0^\infty u\exp(-\lambda u)\,du}{\int_0^\infty \exp(-\lambda u)\,du}$$

と変形され，これを計算すると

$$\mathrm{E}[\,T-s\mid T>s\,]=1/\lambda$$

を得る．この値はどのバスに乗れたかにも s がいくらかにもよらないので，結局この式を導いた条件に関係なく常に平均待ち時間は $1/\lambda$ になる． □

最後の正方形は「証明終わり」を意味する．

また，大文字小文字の使い分けは，確率変数は T のように大文字で書くことが多いという習慣による．分布の密度を用いて期待値の積分を実行するときは確率変数 T が t という値を取る，即ち，$T=t$ での密度という意味で，小文字の t になっている．

今回の話では平均運行間隔という平均値が一定でも平均待ち時間が変わりうることを示した．直感的には「１時間に４本なら１５分くらい待てば１本来るだろう」と考えたいところだが（そしてバスが「でたらめ」に来ればその通りなのだが），一般には平均の取り方によって答えが違う．もちろん，これは取り上げた「バスの到着」という例が見かけに反して（確率過程という）高度な問題だったからであるが，確率論の教えるところには安易な直感の誤りを指摘するものが他にもある．誤った直感を排して，正しい推論に導くために現代確率論は，一見直感的でない用語や定義を用いる．それらは取っつきにくく，この講義でもどれだけ触れられるか分からないが，確率論の問題は根本に立ち返って考えないといけない，という注意は学んでほしい．次回は分散という，平均値同様これも良く知っているはずの量について，誤った直感を体験したことを講義したい．

参考：　Poisson 確率過程 $X(t)$ は，バスの到着だけでなく，時間的にでたらめに１個単位で発生する確率現象のモデルとして実用上も重要である．例えば，レジや案内カウンターやチケット売場などのサービスカウンターへの客の累積到着数，また，通信回線や画像などの（ビットやピクセル単位の）ノイズ，等のモデルに利用される.（実用上相互に無関係に見える問題に適用できるということが Posson 確率過程の重要性を表している.）

Poisson 確率過程 $X(t)$ は次の性質を持つ．

(1) 標本が階段関数，即ち，ときどき（バスが到着するごとに） 1 ずつ増え，それ以外のときは時間的に一定の値をとる．
(2) 加法的，即ち時間 $(t_0,t_1]$ の間の $X(t)$ の増分 $X(t_1)-X(t_0)$ （何台バスが来たか）が t_0 以前の $X(t)$ と独立である．
(3) 時間的一様，即ち，時間幅 $(s,t]$ の間に到着するバスの台数 $X(t)-X(s)$ の分布は $t-s$ だけで決まる．

逆に以上の性質を持つ確率過程は Poisson 確率過程である.（これが本来の定義.）以上の事実を用いると，いろいろな量を式変形で計算することができる．例えば，(1.1) に出てくる λ が実際に平均運行間隔の逆数（即ち単位時間当たり通るバスの台数の平均）$\mathrm{E}[\,X(a+1)-X(a)\,]$ に等しいことなども示すことができる．また，平均待ち時間が $1/\lambda$ に等しいことを先ほど証明したが，この証明も実は加法性と時間的一様性が本質的な役割を果たしている.（チャンスがあれば確率過程の話に進んだときに触れたい.）

数理解析特論（服部）　レポート問題１

講義（資料）を参照して，次の問１または問２の少なくとも一つに解答せよ．**提出期限：4月２７日（木）**．（講義・レポートの難易・分量などの感想や内容についての希望なども自由に書いて下さい.）

問１．平均運行間隔が定数 a で与えられるバスのバス停において，乗客の到着時刻の分布が一様分布のとき，この乗客の平均待ち時間を次の各場合に求めよ：

(1) バスが a ごとに等間隔に到着する場合.
(2) バスが到着間隔 $3a/2$ と $a/2$ を交互に繰り返して到着する場合（つまり，到着時刻が，$3a/2$, $2a$, $7a/2$, $4a$, $\cdots$, のとき.）
(3) $0 < r < 1$ を定数とするとき到着間隔 $r\,a$ と $(1-r)\,a$ を交互に繰り返して到着する場合
(4) バスの累積到着台数が平均運行間隔 a の Poisson 確率過程になっている場合.

問２．あなたはバス会社（その他どんな顧客サービス会社でも当てはまりうるが）の顧客相談窓口に配属されているとする．乗客から次のような苦情が来た.「停留所の時刻表では１０分に１本となっているのに２０分待たされることもざらだ．平均を取ったら１５分待たされている．間引き運転しているのではないか？」あなたの会社が間引き運転をしていない良心的な会社の場合，この苦情に正しく説明するにはどう返事をすればよいかを考えよ.（回答の鍵は講義（資料）に全て書かれているが，客は確率論の専門家でもないし，この講義も聞いていない，という前提で何とか分かるように説明せよ.）

数理解析特論

服部哲弥　1995.4–1995.6

2　切符売り場の行列　–　分散について　（４月２７日）

銀行のキャッシュコーナー，電話ボックス，飛行機のチェックインカウンター，など，窓口（またはボックス）が複数ある場所で，それぞれの窓口の前に別々に並ぶのではなく，一列に並んで空いた窓口を順に利用する「一列待ち」が数年前に新聞などでも話題になった（別紙図参照）．昔は，たいていの場合，スーパーのレジのように全てボックスや窓口の前に別々に並んでいた.「一列待ち」が話題になってからしばらくは「一列待ち」がいろいろなところで試みられたと思う．銀行のキャッシュコーナー，特にスペースのゆったり取れる利用客が多い支店，では「一列待ちが」定着したように思うが，それ以外のところでは必ずしも定着していないように思う.（話題になったとき「欧米ではそうなっている」ということだったが,たしかに，アメリカの飛行場のチェックインカウンターはロープをひっぱて「一列待ち」にしている．日本のはどうなっているのだろう.）

ＪＲのＹ駅の指定券前売り窓口は，昔は,「並列待ち」だったのを，３年ほど前のことだが,「一列待ち」が話題になってしばらく後,「一列待ち」に変えた．ところがＹ駅は一旦開始した「一列待ち」をまもなくやめて,「並列待ち」に戻した．意見箱を利用して理由を問い合わせたところ，Ｙ駅の意見は

(1) 平均時間は変わらない,
(2) 列が長く見えてＹ駅は混んでいると思われてしまう,
(3) 誘導人員が確保できない,

ということであった．第２点は心理学的な問題，第３点は経済的な問題，が中心だが，第１点は平均値という数学的（しかも確率論的）な問題である.「平均が変わらないから一列待ちには意味がない」という主張にみえるが，これは適切な理由とは言えない．今回の講義では

待ち時間の平均は変わらなくても,「一列待ち」には意味がある

ことを論じよう．

最初に，平均値に関するＹ駅の主張は正しいことを確認しておく．

分かりやすくするため，かつ，議論に曖昧さが出ないように，いくつか条件をおく．

(1) 窓口の数を定数 M とし，待っている人数を N，即ち自分が N 人目（自分が到着する直前 $N-1$ 人が待っていた）とする．窓口 M に比べて人数 N がいつも十分大きいとする．特に，窓口に空き時間はない（全ての窓口がいつも仕事をしている）とする．
(2) 客 i $(i=1,2,\cdots,N)$ の処理時間 S_i は確率変数であり，$\{S_i\}$ は独立（客の間に相関はない）とする．窓口は全て同じ処理能力で，各々の客の処理時間の分布は等しい（見ただけでは時間がかかりそうかどうか区別できない）とする．特に平均処理時間 $\tau = \mathrm{E}[\,S_i\,]$ は客によらず一定である．実際に測定してもらえれば直ちに分かるが，客によって券の購入に必要な時間はばらばらである．このことを S_i が確率変数であるとしてとらえる．確率変数 $X_1, X_2, \cdots,$ が独立とは，$\mathrm{Prob}[\,X_1 \in A_1,\ X_2 \in A_2,\ \cdots\,] = \prod_i \mathrm{Prob}[\,X_i \in A_i\,]$ ということ．
(3) 「一列待ち」と「並列ならび」は並び方以外の条件は変わらないとする．例えば，どの窓口があいたか判断して窓口まで歩くのに要する時間は無視する．

上記条件から従うことだが,「並列待ち」の場合どの窓口も均等に並ぶとする．どの窓口が早いか分からないときは人々は自然に均等に並ぶ傾向がある．即ち「並列待ち」は M 個の窓口それぞれに N/M 人ずつ客が並び，ある窓口が空けばその窓口に並んでいた客が順に処理を受け,「一列待ち」は N 人の客が長い一列を作り，M 個の窓口のどれかが空き次第，順に客が空いた窓口へ行って処理を受ける．別紙図では

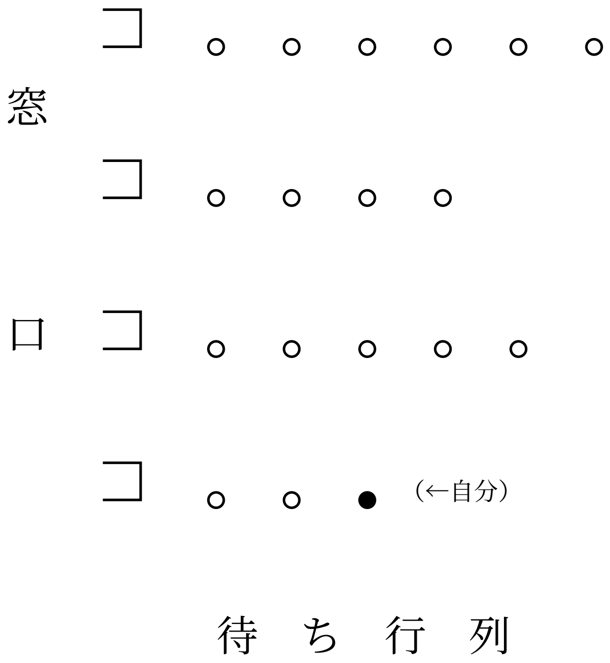


図 2.1: 窓口で並列に並んで待つ場合 (窓口数 $M = 4$)

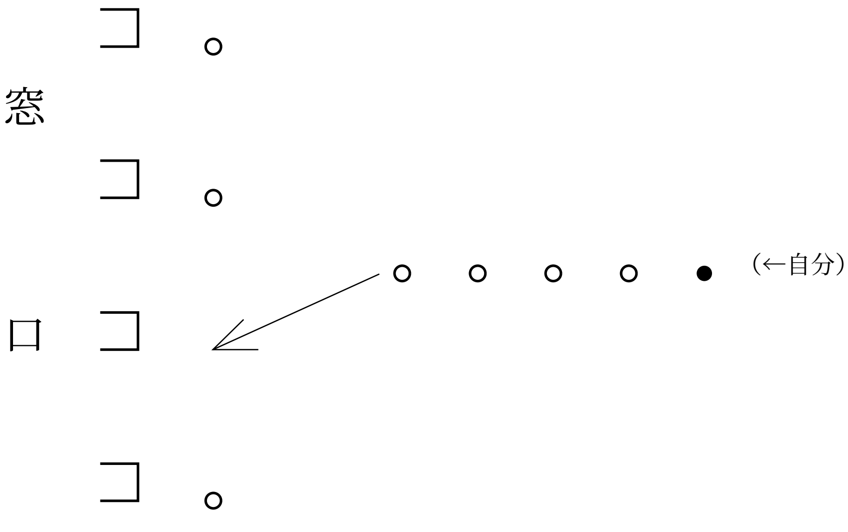


図 2.2: 一列待ち ($M = 4$)

でこぼこしているが，このようなことは考えない．本題とは関係ないが，レジのようにうまい下手があり，また，時間がかかるかどうか買い物かごの中身からある程度予測がつくときは，ある程度でこぼこに並んだ方がいい場合がある．私の観測では，毎日回に来る人々は驚くほど正確に，待ち時間が一番短い列に並ぶ．この問題についてはここでは扱わない．

さて，一人の客の処理をするのに平均 τ（タウ，ギリシャ文字の一つ．t や T は時刻を表すために頻繁に使われるので，文字が足りなくなると τ もよく使われる）の処理時間を要するとする ($\tau = \mathrm{E}[\, S_i \,]$)．自分が N 人目だとすると，自分の分が終わるまでに窓口がしなければならない仕事の総量（処理時間の合計，のべ時間）は平均 $N\tau$ である．窓口が一つしかかなければ平均 $N\tau$ だけの時間がかかるという意味.「並列待ち」でも「一列待ち」でも常時 M 個の窓口が常に処理を続けているから，

自分の番が終わるまでの自分の平均待ち時間 t_1 は $t_1 = N\tau/M$ となり，並び方によらない．即ち平均値に関してはＹ駅の主張は正しい．

しかし,「一列待ち」と「並列待ち」は実際に体験してみると,（体験するまでもなく）たしかに「何かが違う」．どこに違いが出るかを示すために，次の点に注目する．

前売り券を買いに来る人は，次の約束までの時間や仕事の休み時間を利用して窓口に来るか，あるいは，乗る直前に来て発車までの時間に前売り券を買いたい，と考えている.（前売り券を買うためにわざわざ一日中並ぶ覚悟，というのは特殊な場合であろう.）そこで，電車の発車までにあるいは次の約束・用事までに前売り券が買えるかという問題を考える．次の用事までに t_0 の時間的余裕があるとき，それが平均待ち時間 t_1 に比べて長ければ ($t_0 > t_1$)，間に合いそうだと判断して，前売り券を買うために並ぶ.（t_0 と t_1 の大小は判断できることは仮定する.）実際の待ち時間 T は平均 t_1 の周りにばらつく確率変数である．たまたま前の人が長引いて結局買えない場合もあるし，無事買えることもある．$T < t_0$ が実現すれば券を買えるが，窓口が自分より前の客に手間取って $T > t_0$ となってしまったならば待っている間に時間切れとなって券を買わずに次の仕事や約束に向かわなければならない.（発車に間に合わなかった場合は計画を変更しないといけない.）確率 $\mathrm{Prob}[\, T > t_0 \,]$ で前売り券を買い損なうので，この確率が小さいほど「望ましい窓口」ということになる．図 2.3 のように，平均待ち時間 t_1 が同じでも待ち時間 T の分布の形によって「前の客に手間取ったための不運な買い損ない」$\mathrm{Prob}[\, T > t_0 \,]$ の大きさが変わりうる．この値は，客の処理時間 S_i の分布の具体形が分からないと計算できないが，目安として分布の **分散** を考えることができる．図 2.3 の 2 つの分布を比べると，分布が広がっているほど $\mathrm{Prob}[\, T > t_0 \,]$ は大きい（但し $t_0 > t_1$）．分布の広がりの目安が分散であるから，分散が小さいほど買い損ないが少なく望ましい，と想像される．

注意． この想像は一般的には成り立たない．分布の形が違いすぎると $\mathrm{Prob}[\, T > t_0 \,]$ の大小と分散の大小が一致しなくなる場合がある．本当は $\mathrm{Prob}[\, T > t_0 \,]$ を比較すべきである．今回は計算のしやすい分散を用いて話をする．正規分布などの具体的な分布を仮定すれば $\mathrm{Prob}[\, T > t_0 \,]$ も計算可能になる．

「並列待ち」と「一列待ち」の待ち時間をそれぞれ $T^{(1)}$ および $T^{(2)}$ と置くと，

$$T^{(1)} = \sum_{i=1}^{N/M} S_{j_i}\,,\ \text{（並列待ちの待ち時間）},$$
$$T^{(2)} = \frac{1}{M}\sum_{i=1}^{N} S_i\,,\ \text{（一列待ちの待ち時間）},$$

と書ける.（S_i は客 i の処理時間．$j_1, j_2, \cdots, j_{N/M}$, は並列並びについて自分と同じ列に並ぶ客.）既に見たように，$\mathrm{E}[\, T^{(1)} \,] = \mathrm{E}[\, T^{(2)} \,] = t_1$，即ち待ち時間の期待値（平均値）は等しい．

客の処理時間の分布は独立でかつ客によらない（独立同分布）という仮定から，客一人当たり窓口当たりの処理時間の分散 $\mathrm{V}[\, S_i \,]$ は客によらず一定なので，これを $\mathrm{V}[\, S_i \,] = \sigma^2$ とおく．σ はシグマと読む．これもギリシャ文字で，標準偏差（分散の平方根）などを表すのに良く用いる．

分散についての次の公式を利用する．

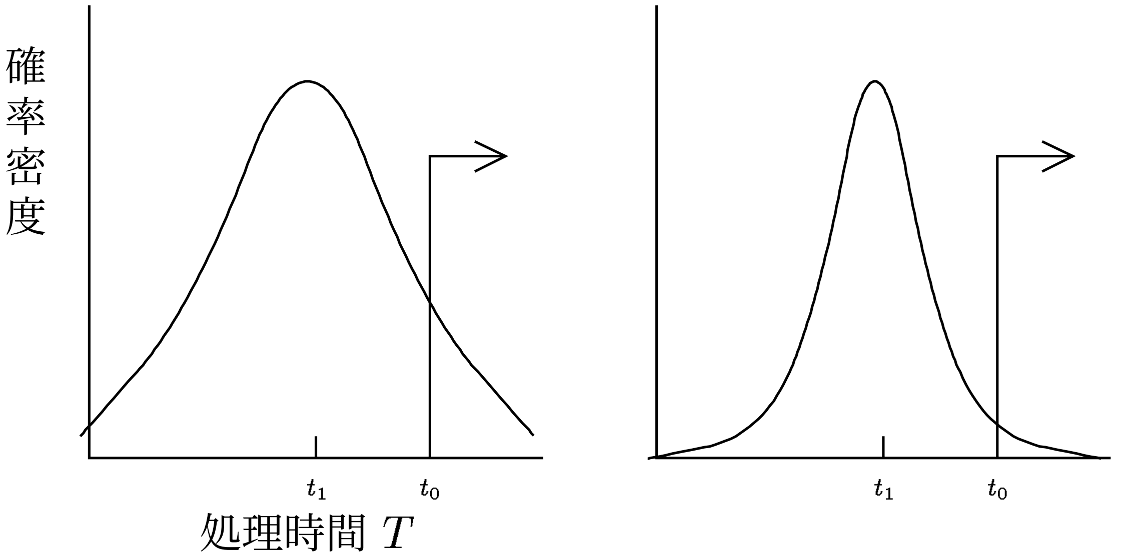


図 2.3: 処理が間に合わない確率（$T > t_0$ の面積）

(1) $\{S_i\}$ が独立ならば，$\mathrm{V}[\, S_1 + S_2 + \cdots + S_n \,] = \mathrm{V}[\, S_1 \,] + \mathrm{V}[\, S_2 \,] + \cdots + \mathrm{V}[\, S_n \,]$.
(2) a が定数ならば（確率変数でないならば）$\mathrm{V}[\, a\,X \,] = a^2\, \mathrm{V}[\, X \,]$.

二つ目の公式で右辺 a の 2 乗が出ることが鍵となるが，この公式の導出は別の機会にまわす．これらの公式と $T^{(1)}, T^{(2)}$ の定義から

$$\mathrm{V}[\, T^{(1)} \,] = \frac{N}{M}\, \sigma^2\,, \text{（並列待ちの場合）},$$
$$\mathrm{V}[\, T^{(2)} \,] = \frac{N}{M^2}\, \sigma^2\,, \text{（一列待ちの場合）}.$$

$\mathrm{V}[\, T^{(2)} \,]$ について証明しておこう．

$\mathrm{V}[\, T^{(2)} \,] = \dfrac{N}{M^2}\, \sigma^2$ **の証明．** 分散についての二つの公式を順に用いて

$$\mathrm{V}[\, T^{(2)} \,] = \mathrm{V}[\, M^{-1}\, S_1 + M^{-1}\, S_2 + \cdots + M^{-1}\, S_N \,]$$
$$= \mathrm{V}[\, M^{-1}\, S_1 \,] + \mathrm{V}[\, M^{-1}\, S_2 \,] + \cdots + \mathrm{V}[\, M^{-1}\, S_N \,]$$
$$= M^{-2}\, (\mathrm{V}[\, S_1 \,] + \cdots + \mathrm{V}[\, S_N \,])$$
$$= \frac{N}{M^2} \sigma^2\,.$$

□

$M > 1$ だと $\mathrm{V}[\, T^{(1)} \,] > \mathrm{V}[\, T^{(2)} \,]$ となるから，図 2.3 で言えば，並列待ちが分散の大きい左側の図，一列待ちが右側の図に対応する．一列待ちの方が並列待ちに比べて，待ち時間（並び始めてから買えるまでの時間）の分散，即ち散らばり具合が小さい．それ故一列待ちのほうが処理し損なう確率（前売り券を買い損なう確率）が低いと思われる．(実際この結論は正しい.)

従って次の約束までの余裕（並んでいられる時間）が限られているときに，買い損なう恐れが小さい．

前回と今回で，安易な直感は誤りに陥る場合があること，確率論には（他の自然科学同様に），根本まで立ち返って考えることによって正しい結論を目指す役割があること，を示した．一見直感に反することを正しいと言うことや正しい直感を示すことは簡単ではなく，一見直感の利かない準備を必要とする．前回と今回は確率論の結果を利用してきたのに対して，結果を導く根拠をこれから講義していく．

来週５月４日は休みです．次回（とレポート提出期限）は再来週５月１１日．

数理解析特論（服部）　レポート問題２

講義（資料）を参照して，次の問１または問２の少なくとも一つにに解答せよ．**提出期限：５月１１日（木）**．

ＪＲのＹ駅の指定券前売り窓口では一旦開始した「一列待ち」をまもなくやめた．問い合わせに対する回答によると，やめた理由は

(1) 平均時間は変わらない,
(2) 列が長く見えてＹ駅は混んでいると思われてしまう,
(3) 誘導人員が確保できない,

ということであった.

問１．第１点については講義でみたように適切な理由とは言えない．この観点から,「一列待ち」に戻すよう駅の担当者を説得する文書を作れ.

問２．第２点は前売り券購入客の理解が必要である．この点もふまえて，ＪＲは「一列待ち」について客の理解を得るためにどうＰＲすればよいか，ＪＲの広報担当官の立場になって，広告説明文を提案せよ.

駅の担当者及び前売り券購入客（一般の人たち）は，確率論の専門家ではないし，この講義を聞いてもいない．その前提で分かりやすく，しかし誤解のないように説明せよ.

数理解析特論

服部哲弥　1995.4–1995.6

3　確率，密度，一様分布，$\cdots$　—　確率論の基礎　（５月１１日）

確率論は，曖昧な言い方をすれば「ばらつく量」を扱う数学である．数学ということは，矛盾なく推論及び計算ができるということである．ばらつく量と言ったのは，例えば測定実験で測定のたびに値が異なることなどを念頭に置いている．測定実験の場合は正しい値は一つのはずだが，制御できない撹乱要因に由来する誤差のために，値が散らばる．天気予報のように観測能力や予測能力の限界から予測値（雨の量など）が不確実な場合もある．大雑把に言えば，値に散らばり（不規則性）を生じる量を**確率変数** **(random variable)** と呼び，各々の値が実現する可能性を**分布**と呼ぶ．

切符売り場の「一列待ち」で既に見たように，不規則性の分析はシステムのあり方を決める上で積極的な意味を持つ．有名な分布の例に偏差値がある．能力に分布がなければ能力測定型の入学試験は無意味になる．分布が入学試験という社会現象の原因になっている．実験や天気予報では値が散らばることは望ましくないので，測定装置を良くするなど制御能力を増す努力が行われる．しかし，誤差は決してなくならない．賭事で百発百中は無理な相談だ．そこで，不規則性を不可避として，なおかつ規則性を見い出そうというのが確率論の役割になる．

不規則性が当然に持っていると考えられている性質を矛盾なく導くための数学が **確率論** である．

さいころや硬貨投げに関連する確率の計算は高校時代までにも習ったかも知れない．直感通りと思った諸君もいるかも知れないが，それは正しい．さいころや硬貨投げのように確率変数の取りうる値（さいころの目や硬貨の表裏で決まる値）が有限個しかない場合は，確率の計算は直感通りであると言ってほぼよい．基本的には

ある事象の起きる確率 $=$ その事象の起こる場合の数 $\div$ 全ての場合の数

で全て計算できる．場合の数が有限だから,「数えればいい」のである．

取りうる値が無数にある場合は気を付けないといけない．取りうる値が無数にある簡単な分布の例として区間 $[0,1)$ 上の**一様分布**がある（大学学部の講義で習ったとかもしれない.）この分布は A を区間 $[0,1)=\{x \mid 0\le x<1\}$ の部分集合（事象）とするとき，事象 A が起きる確率 $P[A]$ が区間 A の長さに等しいものを言う．確率変数という言葉を使うと，確率変数 X が一様分布に従うとは，A を区間 $[0,1)$ の部分集合とするとき，$X\in A$ が起きる確率 $\mathrm{Prob}[\,X\in A\,]$ が区間 A の長さに等しいことを言う．なお，ここでは区間を $0\le x<1$ と 0 を含んで 1 を含まない定義にしているが，これは趣味の問題で両端を含むか含まないかは本質的な意味はない．

長さ 1cm の線分を針でランダム（でたらめ）に刺したとき，刺した場所が左端から何 cm にあるかを測ってその値を X としたとき，X の分布が一様分布である．例えば $\mathrm{Prob}[\,0\le X<0.1\,]=0.1$ などとなる．$0\le x<1$ を満たす任意の実数 x を値として取りうるので，取りうる値に無数の可能性がある．

一様分布を例にとって，安易な計算 (?) が矛盾を引き起こすことを示そう．$[0,1)$ の上の一様分布 P では１点 x $(0\le x<1)$ が生じる確率 $P[\{x\}]$ は，１点の長さは 0 だから，$P[\{x\}]=0$ のはずである．点 x と点 y が異なるならば，事象 $\{x\}$ と事象 $\{y\}$ は**排反事象だから，和事象の確率は個々の事象の確率の和**なので，$P[\{x,y\}]=0$ である．これをとことん突き詰めると，$[0,1)$ は $0\le x<1$ を満たす点の和集合であるから，つぎのように考えたくなる：

$$P[\bigcup_{x\in[0,1)}\{x\}]=\sum_{x\in[0,1)}P[\{x\}]=\sum_{x\in[0,1)}0=0 \quad (?)$$

ところが，$\displaystyle\bigcup_{x\in[0,1)}\{x\}=[0,1)$ だから，

$$0=P[\bigcup_{x\in[0,1)}\{x\}]=P[[0,1)]=1$$

となり，$0=1$ となって矛盾する．1 点の確率 $P[\{x\}]=a>0$ としても同様の理由から $1=\infty$ となってやはり矛盾する．

この矛盾は，排反事象の和事象の確率は個々の事象の確率の和である，という規則を無限個の和事象 $\displaystyle\bigcup_{x\in[0,1)}\{x\}$ に適用するときに論理的限界があることを示している．実数区間の点は $n=1,2,3,\cdots$ のように番号を付けて並べることができない（自然数と対応がつかない）ことが知られている．このことを実数区間の点の**濃度が非可算**であるという．濃度とは個数という言葉を無限個の場合に拡張した概念．確率の定義では**自然数で番号づけられるときに限り**，排反事象の和の確率 = 各事象の確率の和，という公式を正しいと認めている．自然数で番号づけられる「無限個」を（濃度が）**可算** (countable) であるという．

$$P[\bigcup_{x\in[0,1)}\{x\}]=\sum_{x\in[0,1)}P[\{x\}]\quad(?)$$

は，区間 $[0,1)$ の中の点 x の個数が非可算なので，**許されない**変形である．

現代確率論は抽象的な準備の定義から始まるが，これは有限の場合には素朴な定義に合うように，また，無限の場合にも矛盾が起きないように，そして，矛盾が起きない限りできるだけ広いケースに確率が定義できるように，ぎりぎりまで一般化した結果である．過去の数学者の工夫の結果，現代での決定版となった確率の定義は初めて見ると意味が分かりにくいかもしれないが，人智の一つの到達点として一見に値する．

Ω（オメガ，ギリシャ文字）を集合とする（考察対象の空間）．Ω の集合族（部分集合からいくつか選んだ集まり）$\mathcal{F}$ がσ-加法族であるとは以下を満たすことと定義する．

(1) $\Omega\in\mathcal{F}$,（全体集合は必ず事象,）
(2) $A\in\mathcal{F}\ \Rightarrow\ A^c\in\mathcal{F}$,（$A^c$ は補集合を表す，集合 A が選ばれていればその補集合も事象,）
(3) $A_i\in\mathcal{F}$, $i=1,2,3,\cdots\ \Rightarrow\ \displaystyle\bigcup_{i=1}^{\infty}A_i\in\mathcal{F}$,（事象の**可算個の和集合**も事象.）

組 $(\Omega,\mathcal{F})$ を可測空間と呼ぶ.

$$P\ :\ \mathcal{F}\to\mathbf{R}\ \text{（$\mathcal{F}$ に属する集合を与えると実数値が決まる関数，という意味）},$$

即ち，σ-加法族 $\mathcal{F}$ の上で定義された関数（つまり，Ω の部分集合が与えられるごとに値が決まる関数）P が確率であるとは以下を満たすことと定義する．

(1) 全ての $A\in\mathcal{F}$ に対して $P[A]\geq 0$,（正値性，確率は負にならない,）
(2) $P[\Omega]=1$,（全事象，即ち全体集合は確率 1,）
(3) $A_i\in\mathcal{F}, i=1,2,3,\cdots,$ が互いに素 $(A_k\cap A_n=\phi,\ k\neq n)$ のとき，$P[\displaystyle\bigcup_{n=1}^{\infty}A_n]=\sum_{n=1}^{\infty}P[A_n]$,（可算加法性，即ち，排反事象の和の確率 = 各事象の確率の和，が**可算個**の和集合の場合に成立.）

$\mathcal{F}$ の要素 $A\in\mathcal{F}$ を**事象**，$P[A]$ を事象 A の**確率**という．組 $(\Omega,\mathcal{F},P)$ を確率空間と呼ぶ.

全体集合 Ω とは，制御できない撹乱の結果の全体である．Ω の要素は見本と呼ばれ,「さいころを振ったとき出た個々の目」を表す．一様分布の反例で見たように，Ω が無限集合の時は Ω の各要素（見本）の起きる確率だけでは事象の確率は求められない．そこで，見本（要素）ではなく，事象（部分集合）ごとに確率を定める，と考える．ところが，全ての部分集合に確率を定義しようとすると，無限集合の場合に,

確率が持っていてほしい性質に矛盾することが起こりうる（上述の反例より少しこみいってくる）．勝手な部分集合を全て事象とは呼べない（実際に利用する場合もそんなに奇妙な部分集合を考察の対象にはしないので必要がない）．事象の集まりを σ-加法族（シグマかほうぞく） と呼んで定義し，確率を σ-加法族の上の関数と考えるのはこのような考察に基づく．

確率という言葉の定義はこれだけ（つまり,「どんなさいころについても共通に成り立つ性質」が全て出てくると考えている）である．この定義を満たす限り矛盾は生じないことが分かっている（可算加法性, 即ち，排反事象の和の確率 = 各事象の確率の和，が**可算個**の和集合の場合にのみ保証されているところがみそ）．できるだけ多くの場合を統一的に扱い，かつ，できるだけ短い言葉で定義する，という一般化と抽象化のために一見「確率らしさ」が見えないかも知れないが，これだけの定義からもいろいろな確率らしい性質が証明できる．

他方，具体的な計算をするためには具体的な確率が必要である．確率論の一般論の成果を利用するためには，具体例は以上の定義を満たしている必要がある．素朴な（さいころや硬貨投げのような）確率は上記の定義を満たしているので，今まで（高校までで習った）通り確率を計算してよい．一様分布のように実数（あるいはその区間）の上の確率を考えるときは積分で定義すれば自動的に上記の定義を満たしていることが分かっている．実は上記の確率の定義は，現代的な積分の一般論の定義と本質的に同じである．だから，確率が「積分」で定義されていれば自動的に上記の確率論の定義を満たす． 一般に ρ （ロー） を，負にならない $(\rho(x) \geq 0)$ 関数であって，$\displaystyle\int_\Omega \rho(x)\,dx = 1$ を満たすものとするとき，

$$P[A] = \int_A \rho(x)\,dx$$

で定義される確率を考えることができる．$\int_A$ は積分範囲が集合 A であるという意味．ρ をこの確率の **密度** と呼ぶ．全体集合 Ω は実数全体や，区間 $[0,1)$ など実数の部分集合が考えうる．

例えば，区間 $[0,1)$ 上の一様分布は，$P[A]$ が集合 A の長さに等しいと定義したが，これは積分

$$P[A] = \int_A 1\,dx\,, \quad A \subset [0,1),$$

で確率を定義したということと同じである．つまり密度関数 $\rho(x) = 1$, $x \in [0,1)$, で定義される分布である．

また，多次元積分を考えることで，多次元空間上の分布も考えることができる．例えば，それぞれ $[0,1)$ の上の一様分布に従う独立な 2 つの確率変数 X と Y があったとき，$X+Y \leq 1$ となる確率は

$$\mathrm{Prob}[\ X+Y \leq 1\ ] = \int\!\!\int_{x+y\leq 1,\ 0\leq x,y<1} 1\,dx\,dy = \int_0^1 (\int_0^{1-y}\ dx)dy = \int_0^1 (1-y)\,dy = \frac{1}{2}$$

などと積分で計算できる．それぞれ密度 ρ_X および ρ_Y で定義される分布に従う，互いに独立な確率変数 X と Y の結合分布は密度 $\rho(x,y)$ が，積 $\rho(x,y) = \rho_X(x)\,\rho_Y(y)$ で書ける分布に従う．一様分布の場合は $\rho(x) = 1$ なので結合分布の密度も 1 である．

結局,「さいころは場合の数を数えればよく，密度を持つ分布では積分で計算すればよい」．話が進んだとき，確率過程論を理解するためには抽象化された確率論による理解が（本当は）必要である．

確率変数という言葉の現代的定義を述べずに使ってきた．次回は確率変数の定義から始める．

数理解析特論（服部）　レポート問題３

講義（資料）を参照して，次の問１または問２の少なくとも一つに解答せよ．**提出期限：５月１８日（木）**
.

$(\Omega, \mathcal{F}, P)$ を確率空間とする．確率の定義から直ちに導かれる性質には次のようなものがある．

(1) $P[\phi] = 0$,

(2) 有限加法性：$A_i \in \mathcal{F},\ i = 1, 2, 3, \cdots, N,$ が互いに素のとき $P[\bigcup_{n=1}^{N} A_n] = \sum_{n=1}^{N} P[A_n]$,

(3) $P[A^c] + P[A] = 1$,

(4) 単調性：$A_1 \subset A_2$ のとき $P[A_1] \le P[A_2]$,

(5) $P[A] \le 1$,

(6) 劣加法性：$P[\bigcup_{n=1}^{\infty} A_n] \le \sum_{n=1}^{\infty} P[A_n]$,

(7) $A_1 \subset A_2 \subset A_3 \cdots$ のとき $\lim_{n\to\infty} P[A_n] = P[\bigcup_{n=1}^{\infty} A_n]$,

(8) 確率の連続性：$A_1 \supset A_2 \supset A_3 \cdots$ のとき $\lim_{n\to\infty} P[A_n] = P[\bigcap_{n=1}^{\infty} A_n]$.

問１．講義中に述べた確率の定義から上記最初の性質 $P[\phi] = 0$ を証明せよ．(まず σ-加法族の最初の２つの定義から，空集合の確率が定義できること，即ち $\phi \in \mathcal{F}$ を言う．次に確率の定義のうち，可算加法性において，全ての A_n を空集合 ϕ においてみよ.)

問２．$[0,1) \supset A_1 \supset A_2 \supset A_3 \cdots$ を満たす実数の集合の列 $A_1,\ A_2,\ \cdots,$ を次のように定義する．A_1 は $[0,1)$ を３等分して真ん中の区間だけを取り除いたもの，即ち $A_1 = [0, 1/3) \cup [2/3, 1)$ で，２つの区間の和集合．A_2 は A_1 の２つの区間を各々３等分して，各々真ん中の区間だけを取り除いたもの，即ち $A_2 = [0, 1/9) \cup [2/9, 3/9) \cup [6/9, 7/9) \cup [8/9, 1)$ で，４つの区間の和集合．以下同様に，A_n は A_{n-1} の 2^{n-1} 個の（各長さ 3^{-n+1} の）区間を各々３等分して，各々真ん中の区間だけを取り除いた，2^n 個の長さ 3^{-n} の区間の和集合とする．A_n の長さを求めよ．

$A = \bigcap_{n=1}^{\infty} A_n$ （全ての A_n に含まれる点の集合）で定義される集合をカントール集合と呼ぶ．カントール集合 A の長さ（区間 $[0,1)$ の上の一様分布に従うときの $P[A]$）を上記列挙した最後の性質（確率の連続性）を利用して求めよ．(A は非可算個の点からなる集合である，つまり,「実数と同じくらい個数が多い」ことが知られている．カントール集合は長さ０の非可算集合の最も簡単な例として有名である.)

数理解析特論

服部哲弥　1995.4–1995.6

4　確率変数，期待値，分散，独立性，$\cdots$　（５月１８日）

確率変数とは本質的には関数のことである．確率空間 $(\Omega, \mathcal{F}, P)$（前回講義参照）が与えられたとき，Ω の上の関数を確率変数と呼ぶ．正確には「確率の計算できる」関数だが，通常考えるような関数は全て許される．この講義では実数値関数のみを考える：

$$X:\ \Omega \to \mathbf{R} \quad （\Omega\ 上の実数値関数，という意味．\mathbf{R}\ は実数全部の集合）．$$

確率変数の定義の「気持ち」は，制御不能な要因 Ω によって値が変わりうる量，である．関数をわざわざ確率変数と呼ぶのは，集合を事象，積分を確率，と呼ぶのと同様に，これらの基本的な数学がそのまま確率論の「気持ち」を表しているからである．

確率変数 X は関数だから，実数の部分集合 $A \subset \mathbf{R}$ に対して，X の値が A に含まれる確率が考えられる．つまり，事象 $\{\omega \in \Omega \mid X(\omega) \in A\}$ の確率 $P[\{\omega \in \Omega \mid X(\omega) \in A\}]$ である．これを $Q[A]$ とおくと，Q は実数の集合 A を与えるごとに値が決まる．$Q[A] = P[\{\omega \in \Omega \mid X(\omega) \in A\}]$ を，**X が A に入る確率** といい,（実数空間 $\mathbf{R}$ 上の）確率 Q を確率変数 X の **分布**，または**法則** と呼ぶ．例えば，X が正規分布に従うと言うときは，上記の Q が正規分布であるという意味で使う．

確率を実際に計算するときは分布の具体形が問題であって，確率変数という概念はあまり意識しない．これに対して，確率変数は，現実に調べたい問題を，確率の問題として自然に定式化 (formulate)（式に書き下すなどして，曖昧さなく問題を明らかにこと）するのに有効なことが多い．次回以降見るように，複数のばらつく量を調べたいとき確率変数は重要である．

前回注意したように,（Ω が非可算無限集合の場合には）勝手な集合の確率が全て存在するわけではない．確率を定義できる集合，つまり事象，の集まりとして σ-加法族を定義した．同様に，勝手な関数 X を全て考慮の対象にできない．X が確率変数である，というときの「気持ち」には，確率が計算できる（X の分布 Q が存在する）ことも期待されている．そこで，現代確率論では，確率変数の定義にも若干の条件をおく．確率変数の正しい定義を見ておこう．

実数値関数が取る値は実数なので，確率変数の分布は実数上の確率（$\Omega = \mathbf{R}$ のときの確率）になる．実数の場合，区間の確率，即ち，例えば $0.1 \le X < 0.3$ となる確率，が当然ほしいから，どんな σ-加法族でもいいのではなく，区間 $[a, b)$ は全て $(a < b)$ 含まれるような σ-加法族 $\mathcal{F}$ が必要である．区間を全て含む最小の σ-加法族を 1 次元 Borel 集合族と呼び，記号で（$\mathcal{F}$ の代わりに）$\mathcal{B}$ と書く．（σ-加法族の定義から，$\mathcal{B}$ は開区間 (a, b) や閉区間 $[a, b]$ も含むことが証明できるので，片側開区間 $[a, b)$ を用いて $\mathcal{B}$ を定義したのは単なる趣味.）

$(\Omega, \mathcal{F}, P)$ を確率空間とする．$X:\ \Omega \to \mathbf{R}$ が**確率変数**であるとは

$$A \in \mathcal{B} \Rightarrow X^{-1}(A) \in \mathcal{F}\,,$$

を満たすことと定義する．ここで，X^{-1} は X の逆関数を表す．即ち，$X^{-1}(A)$ とは $X(\omega) \in A$ となるような点（見本）$\omega \in \Omega$ の集合である．

$A \in \mathcal{B}$ に対して $Q[A] = P[X \in A] = P[X^{-1}(A)]$ で $\mathcal{B}$ 上の関数 $Q = P \circ X^{-1}$ を定義すれば，Q は $(\mathbf{R}, \mathcal{B})$ 上の確率になる．Q を X の分布（法則）という．

通常考えるような実数の集合は $\mathcal{B}$ に含まれるし，通常考えるような関数は，確率変数の定義を満たす．定義の一つの重要な数学的意味は，この定義の下で**確率 P が測度（長さの拡張概念）となり**，期待値 $\mathrm{E}[\ X\ ]$ が積分として定義できることである．

数学的には確率論は測度論（積分論）である．

前回までに使った $\mathrm{E}[\,X\,]$ と $\mathrm{V}[\,X\,]$ の定義及びいくつかの性質を，復習をかねて取り上げる．

$(\Omega, \mathcal{F}, P)$ を確率空間とし，X, Y を（$(\Omega, \mathcal{F}, P)$ の上の）確率変数とする．確率 P が密度 ρ_P を持っているときには,「普通の」積分で期待値が定義できる：

$$\mathrm{E}[\,X\,] \;=\; \int_\Omega X(\omega)\,\rho_P(\omega)\,d\omega\,, \tag{4.1}$$

$$\mathrm{E}[\,X; A\,] \;=\; \int_A X(\omega)\,\rho_P(\omega)\,d\omega = \mathrm{E}[\,X\,\chi_A\,]\,. \tag{4.2}$$

式 (4.2) は，集合 A の上での期待値，の定義である．χ_A は集合 $A \in \mathcal{F}$ の定義関数（$\omega \in A$ のとき $\chi_A(\omega) = 1$，$\omega \notin A$ のとき $\chi_A(\omega) = 0$ となる関数）．

積分の線型性から，**期待値の線型性**，即ち，定数 a, b と確率変数 X, Y に対して

$$\mathrm{E}[\,aX + bY\,] = a\mathrm{E}[\,X\,] + b\mathrm{E}[\,Y\,] \tag{4.3}$$

が成り立つ.

$A \in \mathcal{B}$ に対して $Q[A] = P[X \in A] = P[X^{-1}(A)]$ で定義された Q が $(\mathbf{R}, \mathcal{B})$ 上の確率になり，X の分布（法則）と呼ぶ，と先ほど説明した．Q は X の値の出る頻度を与えるのだから，X の期待値は Q で計算できるはずである．実際，Q の密度が ρ_Q のとき，

$$\mathrm{E}[\,X\,] = \int_{\mathbf{R}} x\,\rho_Q(x)\,dx\,, \tag{4.4}$$

が定義と一般論から導かれる（最後の「参考」のような議論をする）.

Q は X の値の分布だから，例えば X が整数値しか取らない確率変数ならば，密度は持たない．その場合でも一般論は変わらない．例えば n が整数のとき $P[X = n] = q_n$ $(\sum_n q_n = 1)$ とすると，同様に次式を得る：

$$\mathrm{E}[\,X\,] = \sum_{n=-\infty}^{\infty} n\,q_n\,, \tag{4.5}$$

X の期待値が元の確率空間 $(\Omega, \mathcal{F}, P)$ に関係なく，値の空間（X の**状態空間** と呼ぶ）の上の確率 $(\mathbf{R}, \mathcal{B}, Q)$ だけで決まることにも注目してほしい．例えば，測定実験を行うことを考えると，実験データ X を集めることで，計器の値の分布 Q が実験的に求められ，これに基づいて次の実験の予測が可能になる．データがばらつく本当の原因 $(\Omega, \mathcal{F}, P)$ は必要ない.

確率変数の分散は

$$\mathrm{V}[\,X\,] = \mathrm{E}[\,(X - \mathrm{E}[\,X\,])^2\,]$$

で定義される．期待値の線型性 (4.3) から直ちに次の性質が分かる．

(1) $\mathrm{V}[\,X\,] = \mathrm{E}[\,X^2\,] - \mathrm{E}[\,X\,]^2$,
(2) a が定数のとき $\mathrm{V}[\,aX\,] = a^2\mathrm{V}[\,X\,]$,
(3) $\mathrm{V}[\,X + Y\,] = \mathrm{V}[\,X\,] + \mathrm{V}[\,Y\,] + 2\mathrm{E}[\,(X - \mathrm{E}[\,X\,])(Y - \mathrm{E}[\,Y\,])\,]$,

第 1 の性質を証明するとき，期待値 $\mathrm{E}[\,X\,]$ は定数（期待値を取ると確率変数ではなくなる）であること，従って，線型性から

$$\mathrm{E}[\,\mathrm{E}[\,X\,]\,] = \mathrm{E}[\,X\,]\,\mathrm{E}[\,1\,] = \mathrm{E}[\,X\,]$$

となること，を用いる．第 2 の性質は第 2 回の講義で問題解決の鍵になった性質である．最後の性質は，分散には加法性が一般にはないことを表す．分散が加法性を持つのは，X と Y が独立な場合である．

n 個の確率変数 $X_i,\ i = 1, \cdots, n,$ が**独立**とは,

$$P(X_1 \in A_1, \cdots, X_n \in A_n) = \prod_{i=1}^{n} P(X_i \in A_i), \qquad A_i \in \mathcal{B},\ i = 1, 2, \cdots, n,$$

が成り立つことと定義する．

分布の言葉でいうと次のようになる．$(X_1,\cdots,X_n)$ の結合分布（まとめて n 次元実数値確率変数としてみたときの n 次元空間上の分布）を Q，X_i の分布を Q_i とすると，$X_i,\ i=1,\cdots,n,$ が独立とは，

$$Q[A_1\times\cdots\times A_n]=Q_1[A_1]\,Q_2[A_2]\,\cdots\,Q_n[A_n]$$

が全ての $A_i\in\mathcal{B},\ i=1,2,\cdots,n,$ に対して成り立つこと．ここで，左辺は直積集合の確率，右辺は確率の積である．

期待値との関連では，次のことが言える．$X_i,\ i=1,\cdots,n,$ が独立で，$f_i,\ i=1,\cdots,n,$ が関数のとき，

$$\mathrm{E}[\,f_1(X_1)\,f_2(X_2)\,\cdots\,f_n(X_n)\,]=\mathrm{E}[\,f_1(X_1)\,]\,\mathrm{E}[\,f_2(X_2)\,]\,\cdots\,\mathrm{E}[\,f_n(X_n)\,] \tag{4.6}$$

即ち，**独立な確率変数の積の期待値は期待値の積に等しい．**

特に，このことから，X と Y が独立ならば，

$$\mathrm{V}[\,X+Y\,]=\mathrm{V}[\,X\,]+\mathrm{V}[\,Y\,]+2\,\mathrm{E}[\,X-\mathrm{E}[\,X\,]\,]\,\mathrm{E}[\,Y-\mathrm{E}[\,Y\,]\,]=\mathrm{V}[\,X\,]+\mathrm{V}[\,Y\,]$$

を得る．

即ち，**独立な確率変数の和の分散は分散の和に等しい．**または，**独立な確率変数に対しては分散は加法的である．**

確率変数の無限列，即ち確率変数が無限個 $X_i,\ i=1,2,\cdots,$ ある場合には，$\{X_i\}$ が独立とは，任意のの有限部分列が独立，即ち，任意の n に対して，$X_i,\ i=1,\cdots,n,$ が独立ということ，と定義する．確率変数列が独立ならば任意の部分列は独立になる．これは元の列が有限列でも無限列でも正しい．しかし，有限列の場合，X と Y が独立，Y と Z が独立，Z と X が独立で，X,Y,Z が独立でない例が知られている．無限列の場合はその有限部分列が全て独立であることで全体の独立性を**定義する**.

次回は，確率変数という概念の一つの役割として，確率変数を無限個考える問題（確率過程）に触れたい．

参考： 確率 P が与えられたときに，積分即ち期待値

$$\mathrm{E}[\,X\,]=\int_\Omega X(\omega)dP(\omega)$$

を**密度に頼らずに**定義する現代的な積分論の粗筋を述べておく．

先ず，X が集合 $A\in\mathcal{F}$ の定義関数 χ_A （$\omega\in A$ のとき $\chi_A(\omega)=1$，$\omega\notin A$ のとき $\chi_A(\omega)=0$ となる関数），のとき（ちょうど A の上でだけ 1 になるから），$\mathrm{E}[\,\chi_A\,]=\displaystyle\int_A dP(\omega)=P[A]$ で定義する.（P は確率だから $P[A]$ は存在する.）

次に，X が階段関数，即ち，複数の集合の定義関数の和 $X=\displaystyle\sum_{i=1}^N a_i\chi_{A_i}$ （a_i は定数，$A_i\in\mathcal{F}$）のとき（定義関数の重ね合わせだから），$\mathrm{E}[\,X\,]=\displaystyle\sum_{i=1}^N a_i\int_{A_i}dP(\omega)=\sum_{i=1}^N a_iP[A_i]$ で定義する．

最後に，一般の確率変数 X の場合，階段関数の列 $g_n,\ n=1,2,3,\cdots,$ であって，X に収束する $(\displaystyle\lim_{n\to\infty}g_n(\omega)=X(\omega))$ 増大する $(g_1(\omega)\le g_2(\omega)\le\cdots)$ 列が取れるので，

$$\mathrm{E}[\,X\,]=\lim_{n\to\infty}\mathrm{E}[\,g_n\,]$$

で定義する．確率変数の定義から，このような階段関数（$A_i\in\mathcal{F}$）の近似列がとれ，極限が近似列の取り方によらないことが保証される．

確率 P が密度 ρ を持っているときは,「普通の」積分で期待値が書ける：

$$\mathrm{E}[\,X\,]=\int_\Omega X(\omega)dP(\omega)=\int_\Omega X(\omega)\rho(\omega)\,dx\,.$$

数理解析特論（服部）　レポート問題４

講義（資料）を参照して，次の問に解答せよ．**提出期限：５月２５日（木）**．

講義では分散についていくつかの結果を書いた．例えば,

(1) $\mathrm{V}[\,X\,] = \mathrm{E}[\,X^2\,] - \mathrm{E}[\,X\,]^2$,
(2) a が定数のとき $\mathrm{V}[\,aX\,] = a^2\mathrm{V}[\,X\,]$,
(3) X と Y が独立ならば $\mathrm{V}[\,X+Y\,] = \mathrm{V}[\,X\,] + \mathrm{V}[\,Y\,]$.

問．分散の定義と期待値の線型性 (4.3) （と，必要ならば，独立性の期待値による表現 (4.6)）を用いて，上の３つの性質のうち少なくとも一つを証明せよ．

数理解析特論

服部哲弥　1995.4–1995.6

5　少しだけ，確率過程　－　１．離散時間の場合　（５月２５日）

人生は偶然の連続，という考え方がある.「人生万事塞翁が馬」という言い回しは，人生いつまでたってもその先が予測不能，という意味である．必然性のないようにみえる突然の出来事の連続で，不運（幸運）にも人生がころころ変わる．

ファミコンの RPG (role playing game) は各場面ごとに選択肢があって，一人の主人公についていろいろな冒険が可能である．（ファミコンや RPG は少なくとも 1990 年代前半までは流行していました．もう死語になっていたら，ごめんなさい.）各選択場面でさいころを振って選択肢を決めるとすると，ある時刻 t における RPG の主人公の状態（例えば存在位置）を X_t と書いたとき， X_t が各 t で確率変数になっている，といえる．

実数 $t \in \mathbf{R}$ で番号づけられた確率変数の族（集まり） $\{X_t\}_{t\in\mathbf{R}}$ を 確率過程 と呼ぶ．
（これは正確な定義である.）t は実数全体 $\mathbf{R}$ を動かなくてもよい．例えば区間 $[0,1)$ や非負実数区間 $[0,\infty) = \{t \mid t \geq 0\}$ などがよく用いられる．整数値で（離散的に）番号づけられている場合も確率過程と呼ぶ．この場合は添字を n （には限らないが）にして，X_n, $n = 0,1,2,\cdots$, と記すことが多い．添字（番号）t あるいは n を，time parameter （時刻を表す変数）と呼ぶのが普通だが，必ずしも時刻を表さなくてもよい．例えば : (i) 実験データ X_n, $n = 1,2,\cdots$. 大勢の人が同じ実験をやったとき，制御不能な撹乱要因（測定誤差）のために値が異なる場合，これを確率変数とみてデータ処理を行うことが考えられる．n は誰が行った実験かを区別する．(ii) （１次元）画像データ X_x, $x \in [a,b)$. データに望まない信号（ノイズ）が加わって本来の画像がランダムに乱される場合，各点 x 毎の値（例えば白黒画像ならば輝度）X_x が確率変数と考えられる．x は空間的な位置を表す変数である．

時間変数の値 t を一つ決める毎に X_t は定義によって確率変数 $X_t :\ \Omega \to \mathbf{R}$ である．今まで確率変数は実数値関数としてきたが，実数値でない場合に拡張できる．X_t がある空間 S を値域とするとき，

$$X_t :\ \Omega \to S, \tag{5.1}$$

S をこの確率過程の **状態空間** と呼ぶ．

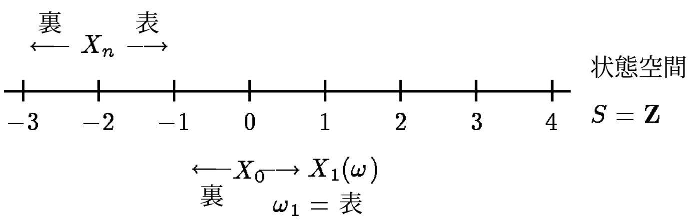
硬貨を投げて表裏に応じて左右に一こまずつ進むゲームを簡単な例にとって，具体的に確率過程を説明する．確率変数の講義（第４回）で示唆したように，元になる確率空間 $(\Omega, \mathcal{F}, P)$ を明示しなくても，X_t の間の関係だけで確率過程 X_t は定まるが，練習のために確率空間を指定するところから始める（適切な確率空間の指定の仕方は一通りではない．簡単な一例をあげる．確率空間の用語については第３回の講義を参照）．

$\omega = (\omega_1, \omega_2, \cdots)$ を，各項が $\omega_n \in \{\pm 1\}$，即ち，± 1 のどちらかであるような数列とし，空間 Ω をそのような ± 1 からなる数列の集合とする．事象の集まり $\mathcal{F}$ は次のようにとる．最初の n 項以内の値を指定した数列の集合

$$A = \{\omega \in \Omega \mid \omega_{k_1} = a_1, \omega_{k_2} = a_2, \cdots, \omega_{k_j} = a_j\}, \quad k_1 \leq n,\ k_2 \leq n,\ \cdots,\ \ a_1,\ a_2,\ \cdots, \in \{\pm 1\} \tag{5.2}$$

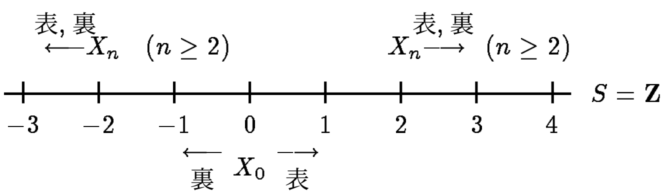
を全て含む最小の σ-加法族を $\mathcal{F}_n$ とおく．例えば $\{\omega \in \Omega \mid \omega_1 = 1\} \in \mathcal{F}_1$ であり，$\{\omega \in \Omega \mid \omega_2 = -1\} \in \mathcal{F}_2 \setminus \mathcal{F}_1$ である（$A \setminus B$ は A に含まれて B に含まれない要素の集合）．定義から

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots . \tag{5.3}$$



裏
表
X_n
状態空間
$S = \mathbf{Z}$
-3 -2 -1 0 1 2 3 4
X_0
$X_1(\omega)$
裏
$\omega_1 =$ 表

図 5.1: 例 1



表, 裏
X_n $(n \geq 2)$
表, 裏
X_n $(n \geq 2)$
$S = \mathbf{Z}$
-3 -2 -1 0 1 2 3 4
裏
X_0
表

図 5.2: 例 2

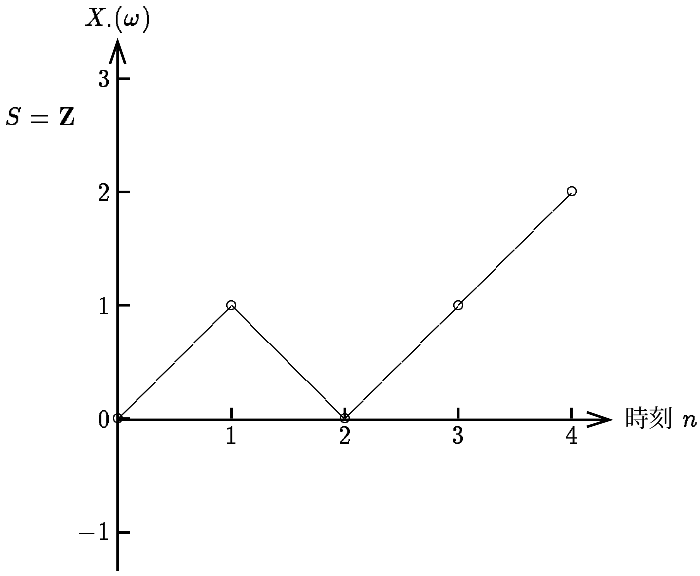


図 5.3: 関数 $X.(+-++\cdots)$; 例 1

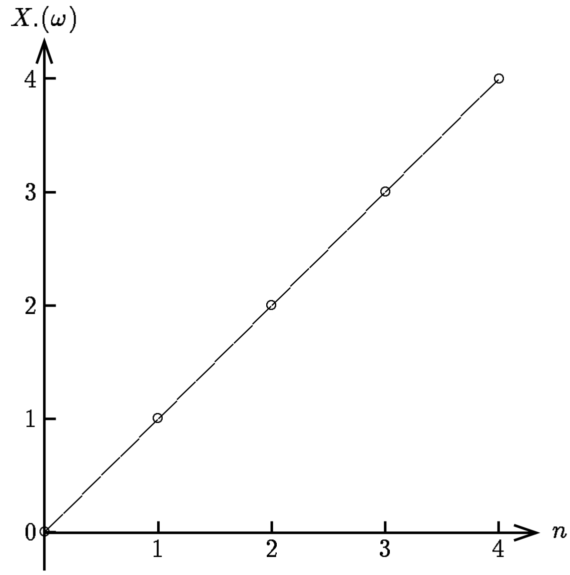


図 5.4: 関数 $X.(+-++\cdots)$; 例 2

この増大する σ-加法族の極限 $\mathcal{F}=\bigcup_{n=1}^{\infty}\mathcal{F}_n$ を事象の集まりとする可測空間 $(\Omega,\mathcal{F})$ を考える．式 (5.2) の集合 A に対して

$$P[A]=2^{-j} \tag{5.4}$$

で確率 P を定義する．（確率空間の定義（第３回）から，(5.2) の形の集合 A について $P[A]$ を定義すれば，全ての $\mathcal{F}$ の要素，即ち事象，に対して確率が定義できる．）数列の，どれでもいいから j 個の項の値を指定すると，それらの値がその通り生じる確率は 2^{-j} になる，という意味である．例えば

$$P[\{\omega \mid \omega_n=1\}]=P[\{\omega \mid \omega_n=-1\}]=1/2 \tag{5.5}$$

が全ての n について成り立つ．

これは,（いかさまでない，まともな）硬貨を（無限回）投げたとき，第 n 回目に表が出れば $w_n=1$, 裏が出れば $w_n=-1$ とおいたことに相当する．

（原理的にはこれで $(\Omega,\mathcal{F},P)$ 上の任意の確率変数の期待値が計算できることになる（第４回講義最後の「参考」)．しかし，P が密度を持たないために，第４回の公式 (4.1) は使えない．他方公式 (4.5) も，無限個の確率変数，確率過程，を扱う場合には万能公式ではない．確率過程の具体的計算は，通常，有限個の確率変数ほど素朴には行かない．この話もいつか触れたい.）

この空間の上の確率変数の無限列，即ち確率過程の簡単な例を２つ定義しよう．簡単のために，時間変数が非負整数をとる場合を考える; $X_n,\ n\in\mathbf{Z}_+=\{0,1,2,3,\cdots\}$. また，状態空間（$X_t$ の値域）を整数値 $\mathbf{Z}=\{0,\pm1,\pm2,\cdots\}$ としよう；$X_n:\ \Omega\to\mathbf{Z}$. 出発点（時刻 $n=0$）は原点とする；

$$X_0(\omega)=0,\quad \omega\in\Omega\,. \tag{5.6}$$

（全ての $\omega\in\Omega$ に対して $X_0(\omega)=0$ が成り立つことは，$P[X_0=0]=1$ と同値である．上記の条件を $P[X_0=0]=1$ と書いてもよい.）

例１ X_n を次のように定義する（図 5.1)；

$$X_n(\omega)=X_{n-1}(\omega)+\omega_n\,,\quad n=1,2,3,\cdots,\ \omega=(\omega_1,\omega_2,\cdots)\in\Omega\,. \tag{5.7}$$

例２ X_n を次のように定義する（図 5.2)；

$$X_1(\omega) = \omega_1\,. \tag{5.8}$$

$$X_n(\omega) = X_{n-1}(\omega)+\mathrm{sgn}(\,X_{n-1}(\omega)\,)\,,\quad n=2,3,\cdots,\ \omega\in\Omega\,. \tag{5.9}$$

ここで $\mathrm{sgn}(\,x\,)=x/|x|$，即ち x の符号を与える関数とする．（すぐ示すように $\mathrm{sgn}(\,0\,)$ は使わないので，定義しないまま放っておく.）

例１のように定義すれば X_n は全て整数値をとる確率変数になる．実際，X_0 は (5.6) で定義される確率変数である，つまり，Ω から整数への 0 という定数関数．X_{n-1} まで定義されたとしよう．ω を定める毎にその第 n 成分 ω_n は定まって，±1 のどちらかの値である．他方 X_{n-1} が確率変数として整数値をとるように定義されていると，それに ±1 の正負どちらを足しても整数値だから，(5.7) によって X_n も Ω から $\mathbf{Z}$ への関数，即ち確率変数になる．従って，全ての X_n 即ち確率過程が帰納的に定義できる．例２も確率過程になっていること，即ち，X_n が $\mathbf{Z}$ に値をとる確率変数になっていること，が帰納的に示せることは例１と同様.

確率空間 Ω の要素を見本 (sample) と呼ぶ，と第３回講義で述べた．今の場合，sample は数列である．Sample を一つ，例えば $\omega=(+-++\cdots)$ と固定しよう（簡単のため符号だけ書いた)．各 n 毎に $X_n(\omega)$ が定まる．これを $n\in\mathbf{Z}_+$ の関数と見ることができる．n を横軸に，$X_n(\omega)$ を縦軸にとればグラフに書ける．例えば $\omega=(+-++\cdots)$ の場合に例１及び例２についてこのようなグラフを書くと，それぞれ図

5.3 及び図 5.4 になる．**（各自定義 (5.6), (5.7), (5.8), (5.9) を用いてこれらのグラフを確かめよ.）** 関数 $X.(\omega)$ は各時刻毎にゲームのコマがどこにいるかをたどっていく．つまり，コマの軌跡（道のり）を表す. (‘·’ は，そこが関数の変数を代入する場所であることを表す便法)．$X.(\omega)$ を ω に対する**道 (sample path, sample function)** と呼ぶ．即ち，確率過程には二つの見方があることになる；

(1) $n \in \mathbf{Z}_+$ を固定したとき，$X_n(\cdot):\ \Omega \to \mathbf{Z}$ （確率変数），
(2) $\omega \in \Omega$ を固定したとき，$X.(\omega):\ \mathbf{Z}_+ \to \mathbf{Z}$ （時間の関数，即ち道）．

第１の見方が確率過程の定義であり，世の中は各時刻 n 毎にランダムで制御しきれない確率変数である，という「人の立場」と言えるかもしれない．すると，第２の見方は賽の目 ω を全て知っている「神の立場」ということになろうか．第２の見方で眺めると，Ω から関数の集まり（関数空間）への関数 $X.$ があることになる．Ω には確率が定義されているので，関数の集合が事象であるような確率空間を考えることになる，とも言える．

第２の見方は，特に上記例２では極めて分かりやすい見方になる．（正直に言うと，例２は，第２の見方を分かりやすく説明するために無理矢理作ったもので，確率過程らしくない確率過程です.）というのは，例２は実は次のように書けるからである；

$$X_n(\omega) = \omega_1\, n, \quad n = 0, 1, 2, \cdots.$$

実際，$n = 0$ のときは (5.6) からこれが正しいことが分かり，$n = 1$ のときは (5.8) からやはり正しいことが直ちに分かる．$\omega_1 = 1$ の場合を考えると，$X_1 = 1 > 0$ だから (5.9) から $n = 2$ のとき $X_2 = X_1 + 1 = 2 > 0$，従って (5.9) で $n = 3$ とおくと $X_3 = X_2 + 1 = 3 > 0$，以下同様である．$\omega_1 = -1$ の場合も同様に順に考えていくと $X_n(\omega) = -n$ が全ての n で成り立つことが分かる．

つまり，例２は本質的に２つの path，$X(n) = n$, $n = 0, 1, 2, \cdots$, 即ち，n 歩目に n にいるから，原点を出発して右にまっすぐ進む path と，これと対称的な左にまっすぐ進む path，だけからなり，$\omega_1 = 1$ ならば前者，$\omega_1 = -1$ ならば後者が実現する．定義 (5.5) から $P[\omega_1 = 1] = P[\omega_1 = -1] = 1/2$ だから，各々が確率 1/2 で現れる．例２のルールでは，(5.9) の見かけの複雑さにも関わらず，最初のさいころの一振りで全てが決まってしまう，といってもよい．

例１では，異なる ω には異なる path $X.(\omega)$ が対応し，無数の path が実現しうる．これが「確率過程らしい」確率過程である．この例は **１次元酔歩 (one-dimensional simple random walk)** と呼ばれる．一本道を一歩毎にでたらめに行ったり来たりする，という描像である．例１は単に酔歩（酔っぱらいの行動）やすごろくを表すだけではない．ω_n は各時刻 n 毎にランダムに値をとり，しかも，時刻間で独立なので（P の定義と独立の定義（第４回）から $\omega_1, \omega_2, \cdots$, が独立であることが示せる），本来 0 になるべき実験データ（あるいはゼロ信号）に加わったランダムな誤差（ノイズ）とみることができる.（比較的）簡単だが大変重要な，典型的な例である．

ω_n を実験データとみなすと，平均 $n^{-1} X_n$ に興味がある．実はこの例１で $n^{-1} X_n(\omega)$ を計算すると，$n \to \infty$ の極限では確率 1 で 0 になる；

$$n^{-1} X_n \to 0, \quad n \to \infty, \quad \text{（確率 1 で）}. \tag{5.10}$$

この結果は **大数の法則** と呼ばれる定理の具体例になっている．

例１では，時間がたつほどいろいろな道が可能になり，$n \to \infty$ を考えることが難しく見える．実際やさしくはないけれども，いろいろなことが分かる．一般に n が大きくなっていくときの確率過程の傾向を与える結果は極限定理と呼ばれる．大数の法則も極限定理の一つ．例１に対して成り立つ，大数の法則より一段精密な極限定理に **中心極限定理** がある；$n^{-1/2} X_n$ の分布は $n \to \infty$ のとき正規分布 $N(0, 1)$ に収束する．以上の話については後日触れたい．

例１，例２とも時刻 n の状態 X_n は X_{n-1} だけで（分布が）定まり，それ以前の状態（過去の履歴）には全くよらない（独立になる）．このような性質を持つ確率過程を **マルコフ過程** と呼ぶ．学部で **分枝過程**

を習ったかも知れない．これは時間変数が離散的なマルコフ過程の典型例の一つで，詳しい研究がある． 非マルコフ過程はマルコフ過程はに比べて，一般に極めて計算が困難である（過去の履歴を全て調べないと次の状態が決まらないので計算が大変）． このため，今日まで多くの原理的研究及び実用的応用がマルコフ過程であった.（研究者としては残念な状況である.）次回は時間変数が連続な場合について，マルコフ過程から典型的な例を紹介する．

数理解析特論（服部）　レポート問題５

講義（資料）を参照して，次の問１または問２の少なくとも一つに解答せよ．**提出期限：６月１日（木）**．

問１．　講義で取り上げた確率過程の例（硬貨を投げて表裏に応じて左右に一こまずつ進むゲーム）の例１及び例２それぞれについて，時刻 $n \in \mathbf{Z}_+$ における期待値 $\mathrm{E}[\, X_n \,]$ を計算せよ．特に $n = 0, 1, 2, \cdots$ の値によって期待値が変わるか否かを指摘せよ．

問２．　講義で取り上げた１次元酔歩の例（例１）について，計算機で ± 1 を確率 $1/2$ ずつで発生する乱数列を作り，定義 (5.6) と (5.7) を用いて見本 (sample path) $X_1, X_2, \cdots$ を発生させ，図示して見よ．横軸に $n \in \mathbf{Z}_+$ 縦軸に $X_n \in \mathbf{Z}$ をとれ．n が大きくなるとどのような傾向があるか．乱数の初期値を変えて，見本を何個か図示して見よ．

（計算機による計算についてはプログラムもつけること．Mathematica などの高級ソフトの利用も歓迎するが，どの場合でも，プログラムまたはソースをつけること.）

数理解析特論

服部哲弥　1995.4–1995.6

6　少しだけ，確率過程　−　２．Poisson 過程　（６月１日）

第５回で確率過程を定義して，時刻変数 n が離散（整数）的な場合について，簡単な例で説明した．時刻変数が離散的な場合でも，極限定理を中心に，問題はたくさんある．応用上も，多くの社会的な活動は，RPG やすごろくのように，ランダムな選択が離散的に起こる．計算機や digital 通信は，下部構造そのものが離散時間的である．

他方，自然現象は時間が連続であり，いつでも「ノイズ」が生じうる．第１回のバスの到着問題や，第２回の窓口の処理待ち問題は，連続時間確率過程の問題である．基礎事項の詳細には立ち入れないが，連続時間の場合，有限の時間間隔でも無限個の確率変数があり，sample path $X_t(\omega)$ の t に関する連続性のような，短い時間での変化を問う問題が生じる．今回は，**時間変数が連続で，状態空間（値域，X_t の取りうる値）が離散的** な確率過程の代表例として，Poisson process （ポワッソン過程）を紹介する．

Poisson 過程は，１個単位で時間的にでたらめに発生して累積していく量の数学的モデルである．第１回講義のバスの例では，特定の停留所で見張っていて，時刻 t までに到着（発車）したバスの累積台数を X_t とおいた（図 6.1）．その性質が Poisson 過程でよく表されることが分かれば,「でたらめに」到着していると判断できる．応用上重要な例として，電話交換台が処理する電話の累積回数やサービスカウンター（レジ，案内カウンター，チケット売場など）への客の累積到着数，通信回線や画像などのビットやピクセル単位のノイズの累積個数，等がある．

時間が連続な確率過程 $X_t,\ t \geq 0$, が Poisson 確率過程であるとは，次の条件を満たすことである．

(1) 状態空間（X_t の値域）が非負整数で，sample path （t の関数としてみた $X_t(\omega)$，第５回講義参照）の時間的変化は（上記例のように） 1 ずつ増えるだけである（図 6.1）.

(2) **加法的**，即ち， $t_1 < t_2 < \cdots < t_n$ とするとき，増分 $X_{t_n} - X_{t_{n-1}}$ （時間 $(t_{n-1}, t_n]$ に何件到着したか）が，それ以前の時刻の増分 $X_{t_{n-1}} - X_{t_{n-2}}, \cdots, X_{t_1} - X_{t_0}$, と（確率変数として）独立である．

(3) **時間的一様**，即ち，確率変数 $X_t - X_s$ の分布は $t - s$ だけで決まり，s そのものにはよらない．

(4) $X_0 = 0$. （測定開始時 $t = 0$ は累積ゼロ件という意味で，本質的でなく，省いてもよい.）

$X_t - X_s$ の分布は明示していないことに注意．定義上は分布の具体形によらず，上記４条件を満たせば（即ち，幅 1 で増大し，加法的かつ時間的一様な連続時間の確率過程を）Poisson 過程と呼ぶ．ところが，次の事実が証明できる．

定理 6.1. 上記４条件を満たせば，$X_t - X_s$ の分布は，平均値が $t - s$ に比例する *Poisson* 分布になる．

Poisson 分布 とは，非負整数 $n = 0, 1, 2, \cdots$ 上の分布であって値 n をとる確率が

$$Q[\{n\}] = \exp(-\mu)\,\frac{1}{n!}\,\mu^n \tag{6.1}$$

で表されるものを言う．正定数 μ は Poisson 分布の平均値である；

$$\sum_{n=0}^{\infty} nQ[\{n\}] = \mu\,. \tag{6.2}$$

確率変数 Z が平均 μ の Poisson 分布に従うとは，$\mathrm{Prob}[\ Z = n\ ] = Q[\{n\}]$ ということであった（第４回講義参照）．定理 6.1 の言うことは，X_t が Poisson 過程ならば，任意の $s < t$ に対して $X_t - X_s$ （時間 $(s, t]$ 内の到着数）が

$$P[X_t - X_s = n] = \exp(-\lambda\,(t-s))\,\frac{1}{n!}\,\{\lambda\,(t-s)\}^n\,, \quad n = 0, 1, 2, \cdots, \tag{6.3}$$

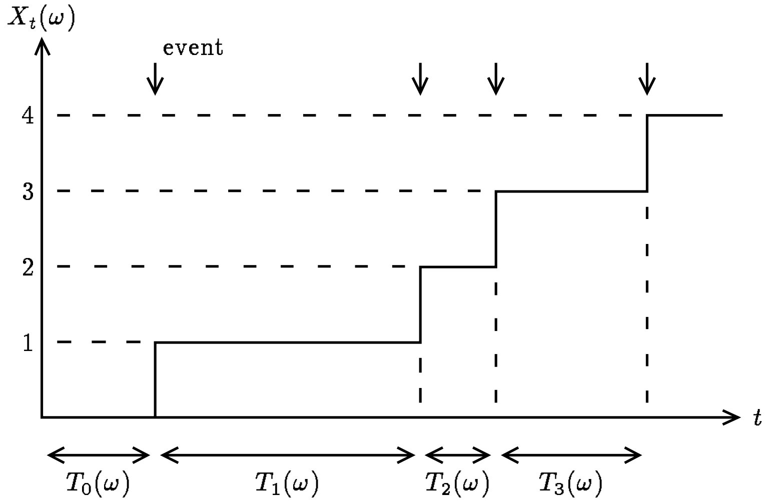


図 6.1: Poisson 過程の sample path $X.(\omega)$

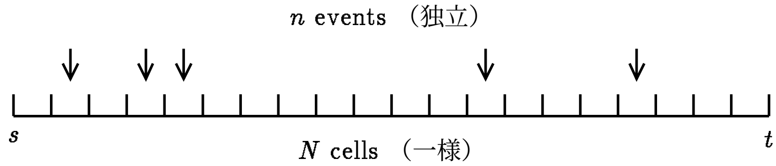


図 6.2: 式 (6.3) の導出

X_t Poisson 過程 $\left(\begin{array}{l}\text{sample path は階段関数,} \\ \text{加法的, 時間一様}\end{array}\right)$ $\underset{\text{定理 6.3}}{\overset{\text{定理 6.2}}{\rightleftarrows}}$ $\{T_n\}$ i.i.d., $\mathrm{E}[\,T_n\,] = 1/\lambda$

$\Big\downarrow$ 定理 6.1

$X_t - X_s$ Poisson 分布, $\lambda = \mathrm{E}[\,X_1 - X_0\,]$

図 6.3: Poisson 過程の定義・諸性質の関係

に従う，ということである．ここで正定数 λ は単位時間幅 1 内の到着数の期待値であることは Poisson 分布の説明から明らかであろう．Poisson 過程という名前はこの事実に由来する．一見計算の役に立たないように見える Poisson 過程の定義（４つの条件）から，具体的な公式 (6.3) が得られることに注意．

定理 *6.1* **の証明．** 定理の証明をきちんと行う余裕はないが，一般原理（加法性，一様性）から (6.3) という具体形が出る「気持ち」を示す．

$P[X_t - X_s = n]$，即ち時間幅 $(s, t]$ に n 回 event が発生する（X が n だけ増える）確率，を計算するために，区間 $(s, t]$ を N 等分して，ひと区画当たり時間間隔 $(t-s)/N$ にする（図 6.2）．N を n に比べて十分大きくとれば，短い時間間隔 $(t-s)/N$ に event 発生が 2 回以上起こる可能性は極めて小さくなる．(正確には，N が大きいときこの可能性が小さいことを，加法性と一様性を用いて言うのだが，省略する．) Event が N 個の小区画のうちの n 箇所で起こる．単位時間当たり平均 λ 回 event が発生するとすると，時間間隔 $(t-s)/N$ では $p = \lambda(t-s)/N$ の確率で１回 event が発生する．小区画あたり確率 p で起こることが N 個の小区画のうちの n 箇所で起こる確率は（高校で習ったように），${}_N C_n \, p^n \, (1-p)^{N-n}$ である（この式を導くところで加法性と一様性を使っている）．但し，これは１つの小区画の中で 2 回以上 event が発生するケースを除外して導いたので，この式は $N \to \infty$ で初めて $P[X_t - X_s = n]$ に等しくなる；

$$P[X_t - X_s = n] = \lim_{N\to\infty} {}_N C_n \left(\frac{\lambda\,(t-s)}{N}\right)^n \left(1 - \frac{\lambda\,(t-s)}{N}\right)^{N-n} .$$

後は標準的な計算で証明できるが，参考までに右辺の計算方法の方針を示しておく．${}_N C_n = N!/(n!\,(N-n)!)$ と分数にして，さらに分母と分子をそれぞれ $\sqrt{2\pi N}\, N^N \, \exp(-N)$ で割っておく．公式 $\displaystyle\lim_{N\to\infty} \frac{N!}{\sqrt{2\pi N}\, N^N \exp(-N)} = 1$ を $N!$ と $(N-n)!$ に適用する．$(\frac{\lambda\,(t-s)}{N})^n$ からも分母に N^n が出ることに注意すると，右辺は

$$\lim_{N\to\infty} \frac{1}{\sqrt{1-n/N}\,(1-n/N)^{N-n}\,\exp(n)\,n!} \,(\lambda\,(t-s))^n \left(1 - \frac{\lambda\,(t-s)}{N}\right)^{N-n}$$

と変形される．a, b が N によらないとき成り立つ公式 $\displaystyle\lim_{N\to\infty}(1-a/N)^N = \exp(-a)$ 及び $\displaystyle\lim_{N\to\infty}(1-a/N)^b = 1$ を用いれば (6.3) を得る． □

X_t が変化する（即ち 1 だけ増える）ことを，event が発生した，ということがある．この表現の気持ちは明らかであろう．(時刻ゼロから測定開始して）初めて event が発生した時刻を T_0, 以下 n 件目と $n+1$ 件目の event 発生の **時間間隔** を T_n $(n \ge 1)$ とおく（図 6.1）．T_n たちは $\{X_t\}$ で定まる確率変数である．これについて次の性質が証明できる．

定理 6.2. *X_t が Poisson 過程ならば，$T_0, T_1, T_2, \cdots,$ は独立で，全て同じ分布を持つ（この性質を* **独立同分布** *であると言い，i.i.d. と書く）．T_0 の分布は平均 $\mathrm{E}[\,T_0\,] = 1/\lambda$ の指数分布になる（どの T_n でも同じ）．ここで λ は (6.3) の λ，即ち，単位時間内の到着数の期待値 $\mathrm{E}[\,X_1 - X_0\,]$ である．*

（平均 $1/\lambda$ の）指数分布とは $t \ge 0$ 上の分布で，密度が

$$\rho_\lambda(t) = \lambda \exp(-\lambda\, t) \tag{6.4}$$

で与えられるものを言う．例えば，時間間隔 T_n が a 以上である確率は

$$P[T_n \ge a] = \int_a^\infty \rho_\lambda(t)\, dt \tag{6.5}$$

で計算される．

定理 *6.2* **の証明．** $\{T_n\}$ が独立同分布であることは，Poisson 過程の加法性と一様性から導かれるのだが，省略する（以下の証明にかなり含まれている）．T_n が平均 $1/\lambda$ の指数分布に従うことだけ証明しよう．定

理 6.1 により，Poisson 過程の定義から (6.3) が導かれることが（証明は完全にはしなかったが）分かっているので，(6.3) から (6.4) を得ることができればよい．$a \geq 0$ を定数とする．$T_0, T_1, \cdots, T_{n-1}$ をそれぞれある値に固定するという条件の下で $T_n > a$ となる条件付き確率（これを $P[T_n > a \mid T_0, T_1, \cdots, T_{n-1}]$ と書く）を計算する．$S = \sum_{i=0}^{n-1} T_i$ とおいておく．T_n の定義から，求める確率は，時間間隔 $(S\,,\ S+a]$ の間に event が発生しない確率，即ち，この時間に X_t が変化しない確率に等しい．従って，

$$P[T_n > a \mid T_0, T_1, \cdots, T_{n-1}] = P[X_{S+a} = X_S \mid S] = \exp(-a\,\lambda)\,.$$

最後の等号は (6.3) において $n = 0$, $s = S$, $t = S + a$ とおいて得られる．右辺は S によらないから，T_0, $T_1, \cdots, T_{n-1}$ の値によらない（即ち，これらの変数と T_n は独立である）．だから，条件をはずしても等号が成り立つ；$P[T_n > a] = \exp(-a\,\lambda)$. この式と，(6.5) を見比べれば，$T_n$ の分布の密度は $\exp(-a\,\lambda)$ を a で微分して符号を変えたものになることが分かる．これは（a を t と書き換えれば）(6.4) の右辺に他ならない． □

定理 6.2 の逆も成り立つ．

定理 6.3. *$\{T_n\}$, $n = 0, 1, 2, \cdots$, が独立同分布で，平均 $1/\lambda$ の指数分布ならば，*

$$X_t = \min\{n \mid \sum_{i=0}^{n} T_i > t\} \tag{6.6}$$

で定義される確率過程 X_t, $t \geq 0$, は Poisson 過程になる．

定義 (6.6) の右辺は時刻 t までに発生した event の総数を表している（各自確認せよ）．定理 6.3 の証明は省略する．X_t の加法性と独立性を証明すればいいのだが，$\{T_n\}$ が独立同分布であることだけでなく，指数分布に従うことも用いる必要がある．

以上の事実を用いると，いろいろな量を計算できる．既に定理の中で，λ が単位時間当たり通る車の台数の平均 $\mathrm{E}[\,X_{t+1} - X_t\,]$ に等しく，さらに，平均通過時間間隔の逆数 $1/\mathrm{E}[\,T_n\,]$ に等しいことを示した．これが，平均待ち時間の逆数に等しいことは，第１回講義で（以上の結果を先取りして）証明した．平均待ち時間と言ったのは，ある時刻 s に停留所に到着した乗客がバスがくるまで待つ時間の平均値 $\mathrm{E}[\,\sum_{i=0}^{X_s} T_i - s\,]$ のことである（この式がそういう意味を持つことは各自確認せよ）．第１回講義で述べたようにこの値は，$T_n > s'$ なる条件下での $T_n - s'$ の期待値（条件付き期待値）$\mathrm{E}[\,T_n - s' \mid T_n > s'\,]$ に等しい（この主張を式変形で証明するには，条件付き期待値について説明する必要がある．いつかこの問題にも触れたい）．そしてそれが $1/\lambda$ に等しいことは，第１回講義で計算したとおりである．

参考： 定理 6.3 は event 発生の累積観測データを得たとき，それが Poisson 過程になっている（つまりでたらめな event）かどうかを確かめるのに便利である．(i) まず，累積到着数 sample path $X_t(\omega)$ を多数集める．１時間毎に（毎日，など）データを異なる sample path とみなす.（t は 0 から１時間（１日）しか動かないことになる.）(ii) それぞれ，最初の event 発生時刻を $T_0(\omega)$ とし，以下順に発生間隔を $T_n(\omega)$ に割り振る．(iii) 各 n ごとに T_n の sample として割り振られたデータのヒストグラム（度数分布）または，分布関数，を作って，それが指数分布になることを確認する．(iv) 異なる n の間の T_n が確率変数として独立（第４回講義参照）なことを確認する．

以上を満たせば，定理 6.3 より，この観測データは Poisson 過程になっていると結論される．Poisson 過程の定義（事実上，加法性と一様性）をチェックする方法も良い．

最後に，極めて重要，かつ驚くべきことは，このように，計算しやすい性質をいくつも持つ確率過程（図 6.3）が本当に存在するという事実である．講義では触れなかったが，Poisson 過程の定義を満たす確率過程が，適当な確率空間の上の確率過程（確率変数の集まり）として，存在することが証明されている．

今回は時間変数が連続で，状態空間が離散的な代表例としての Poisson 過程を取り上げた．次回は時間変数も状態空間を連続な代表例として Wiener 過程を取り上げる．

数理解析特論（服部）　レポート問題６

講義（資料）を参照して，次の問１，問２，または問３の少なくとも一つに解答せよ．**提出期限：６月８日（木）**．なお，計算機による計算を含む場合はプログラムもつけること．Mathematica などの高級ソフトの利用も歓迎するが，どの場合でも，プログラムまたはソースをつけること．

問１．　Poisson 分布の定義 (6.1) を用いて (6.2) を証明せよ．さらに，定義 (6.1) に出てくる（n によらない）定数 $\exp(-\mu)$ は全事象の確率が（確率の定義通り） 1 になるように決まっているはずである；

$$\sum_{n=0}^{\infty} Q[\{n\}] = 1\,.$$

この式を証明せよ．

また，指数分布の密度 (6.4) についても，全事象の確率 $\int_0^\infty \rho_\lambda(t) = 1$ であること，及び，平均が $1/\lambda$ になること，を証明せよ．最後に，平均 $1/\lambda$ の指数分布に従う確率変数 T が，定数 a 以上になる確率 $P[T \ge a]$ を計算せよ．

問２．　計算機の乱数を利用して Poisson 過程 X_t の sample path を発生して見よ．乱数の初期値を変えていろいろ発生させ，いくつかを，X_t を縦軸 t を横軸にして図示せよ．Sample を多数作って，適当な s, t をいくつかとって，時間 $(s, t]$ の間の到着数 $X_t - X_s$ の統計分布（0 の sample がいくつ，といったふうに）をとってみよ．また，Event 間隔 T の統計分布も図示して見よ．

Sample の発生には，event 発生間隔 T が指数分布に (6.4) に従うことを使うのが便利かも知れない．(計算機で，区間 $[0, 1)$ の一様分布に従う乱数 w から，指数分布に従う乱数 u を作るには，$u = -\log(w)$ と置けばよい．)

問３．　レポート問題５で発生させた１次元酔歩について，いろいろな乱数の初期値をとって，１次元酔歩の長い（N 歩目までとるとする）sample をたくさん（T 本とるとする）用意せよ．第 i $(1 \le i \le T)$ sample の n 歩目の位置を $X_n(i)$ とする．このとき，X_n の分散の sample 平均

$$V_n = \frac{1}{T}\sum_{i=1}^{T} X_n(i)^2$$

が n のどのような関数になているか，図示してみよ．(ヒント．N よりもむしろ T を大きくとる必要がある．乱数は直前の sample を作るのに使った最後の値を続けて使えばよい．十分 sample 数 T が大きければ $V_n = n$ に近づくはず．)

数理解析特論

服部哲弥　1995.4–1995.6

7　少しだけ，確率過程　―　３．Wiener 過程　（６月８日）

Poisson 過程は，時間変数が連続で状態空間が離散的な確率過程の代表例であった．今回は，時間変数も状態空間も連続な確率過程の代表例，**Wiener（ウィーナー）過程**，を取り上げる．Wiener 過程は，確率過程論の出発点になった代表例中の代表例であり，多くの興味深い，特徴的な性質を持つ．応用上の重要性も格別である．自然現象は時間的にも空間的にも連続なものが多いから，当然かも知れない．

1827 年，花粉の中の微小な粉が水中で不規則な運動を行うことが顕微鏡観察で発見された．Wiener 過程の研究の出発点となったこの運動は，発見した植物学者にちなみ，Brown（ブラウン）運動と呼ばれる．Einstein（アインシュタイン）等の研究により，不規則な運動の原因は水分子の熱運動による撹乱であることが確立した．１９世紀末に「観測できないものは存在しない」という哲学から，見ることのできない分子や原子を否定した物理学者・科学哲学者 Mach（マッハ）が，Brown 運動を顕微鏡で見せられて分子の存在を認めた，という科学史上のエピソードもある．Wiener 過程は，Brown 運動の数学モデルであると同時に，通信回線等の（アナログ）電気回路の熱雑音のモデルでもある．名前は，研究の端緒を与えた電気工学者 Wiener に由来する．

時間が連続な確率過程 $X_t,\ t \geq 0$, が Wiener 過程であるとは，次の条件を満たすことである．

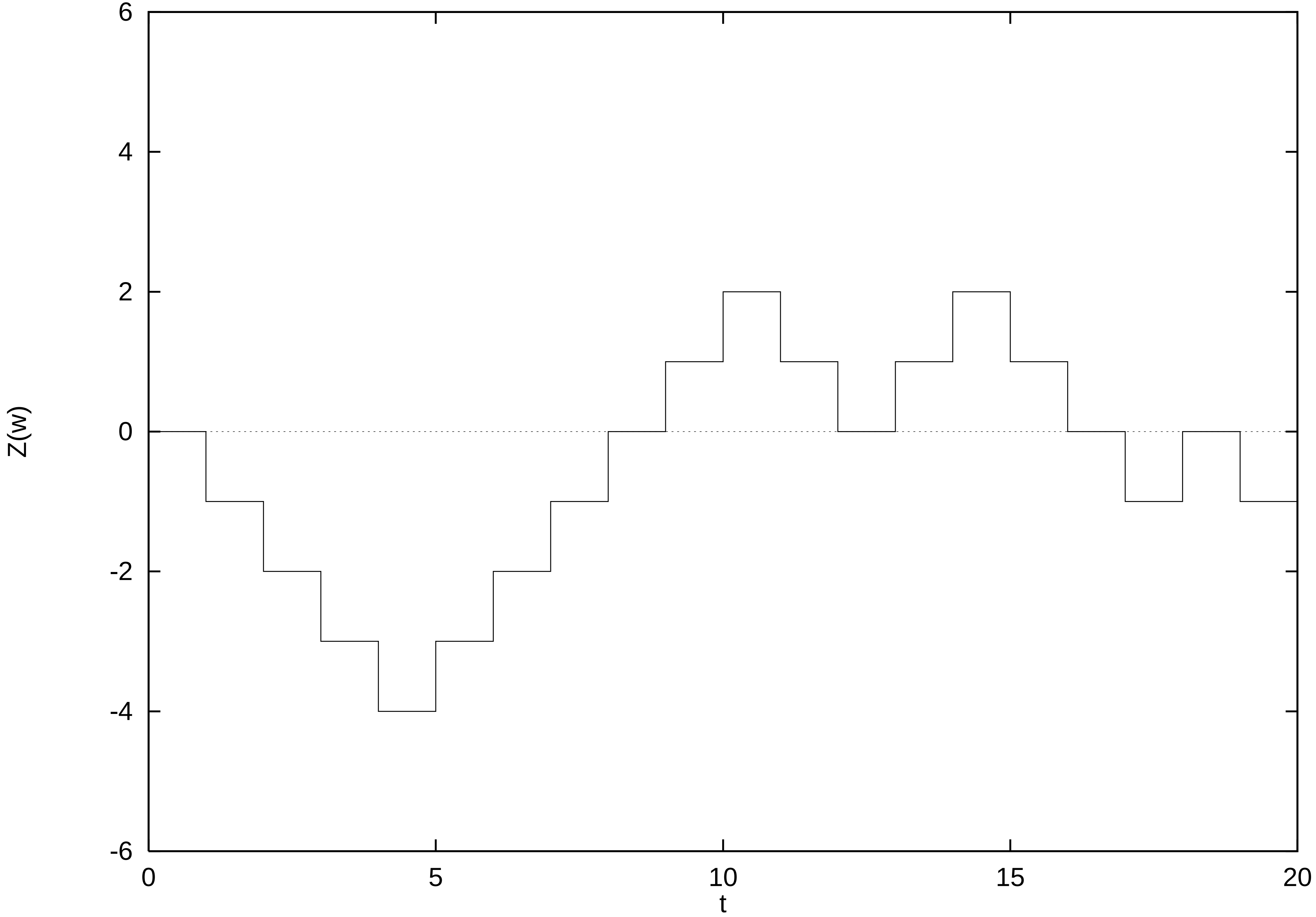
(1) 状態空間（X_t の値域）が実数 $\mathbf{R}$ で，sample path （時間 t の関数としての $X_t(\omega)$，第５回講義参照）は連続関数である．
(2) 加法性（$t_1 < t_2 < \cdots < t_n$ とするとき，$X_{t_n} - X_{t_{n-1}}$ が，$X_{t_{n-1}} - X_{t_{n-2}}, \cdots, X_{t_1} - X_{t_0}$, と独立なこと，第６回講義参照）．
(3) 時間的一様性（$X_t - X_s$ の分布が $t - s$ だけで決まること，第６回講義参照）．
(4) $\mathrm{E}[\,X_1 - X_0\,] = 0$，かつ，$\mathrm{V}[\,X_1 - X_0\,] = 1$.
(5) $X_0 = 0$.

以上の性質を持つ確率過程が存在することは証明されている．Wiener 過程のように種々の強力な公式が成り立つ確率過程が存在すること自体が，深い内容であることは，Poisson 過程（第６回講義）でも注意したとおり．（計算しやすいように安易にいろんな性質を仮定すると，そんな話のうまい確率過程は存在しない，ということになりかねない.）

条件 (5) は原点を決めるだけで，重大な内容はない．すぐ述べるように，条件 (4) も本質的ではないので，Wiener 過程の定義のうちで本質的な条件は (1)(2)(3) （連続性，加法性，時間的一様性）である．条件 (2) 及び (3) は Poisson 過程と共通である（第６回講義）．確率過程の分類学上，Wiener 過程と Poisson 過程はともに，実数 $\mathbf{R}$ 上の，時間的にも空間的に一様な，マルコフ過程である．全く異なる性質を持つが，ある観点からは「親類」である．マルコフ過程については改めて紹介したい．条件 (4) が本質的ではない根拠を述べる．条件 (1)(2)(3)(5) を満たす確率過程 X_t が与えられたとし，$m = \mathrm{E}[\,X_1\,]$, 及び $v = \mathrm{V}[\,X_1\,]$, とおく．このとき，$Y_t = X_{t/v} - m\,t/v$ とおくと，Y_t は条件を全て満たし，Wiener 過程になる（(4) 以外を満たすことは明らか）．逆に言えば，X_t が条件 (4) 意外を満たせば，Wiener 過程 Y_t を用いて $X_t = Y_{v\,t} + m\,t$ と書ける．これは次の定理による（証明はしない）．

定理 7.1.　確率過程 X_t が条件 *(1)(2)(3)(5)* を満たせば，ある実数 m 及び $v \geq 0$ がとれて，任意の $t > s$ に対して $X_t - X_s$ は，平均 $\mathrm{E}[\,X_t - X_s\,] = m\,(t - s)$，分散 $\mathrm{V}[\,X_t - X_s\,] = v\,(t - s)$ の正規分布に従う．特に X_t が *Wiener* 過程ならば（*(4)* も満たせば），$m = 0,\ v = 1$ である ($t = 1,\ s = 0$ と (4) から)．

Fig. 7.1: simple random walk: sample path Z(w)



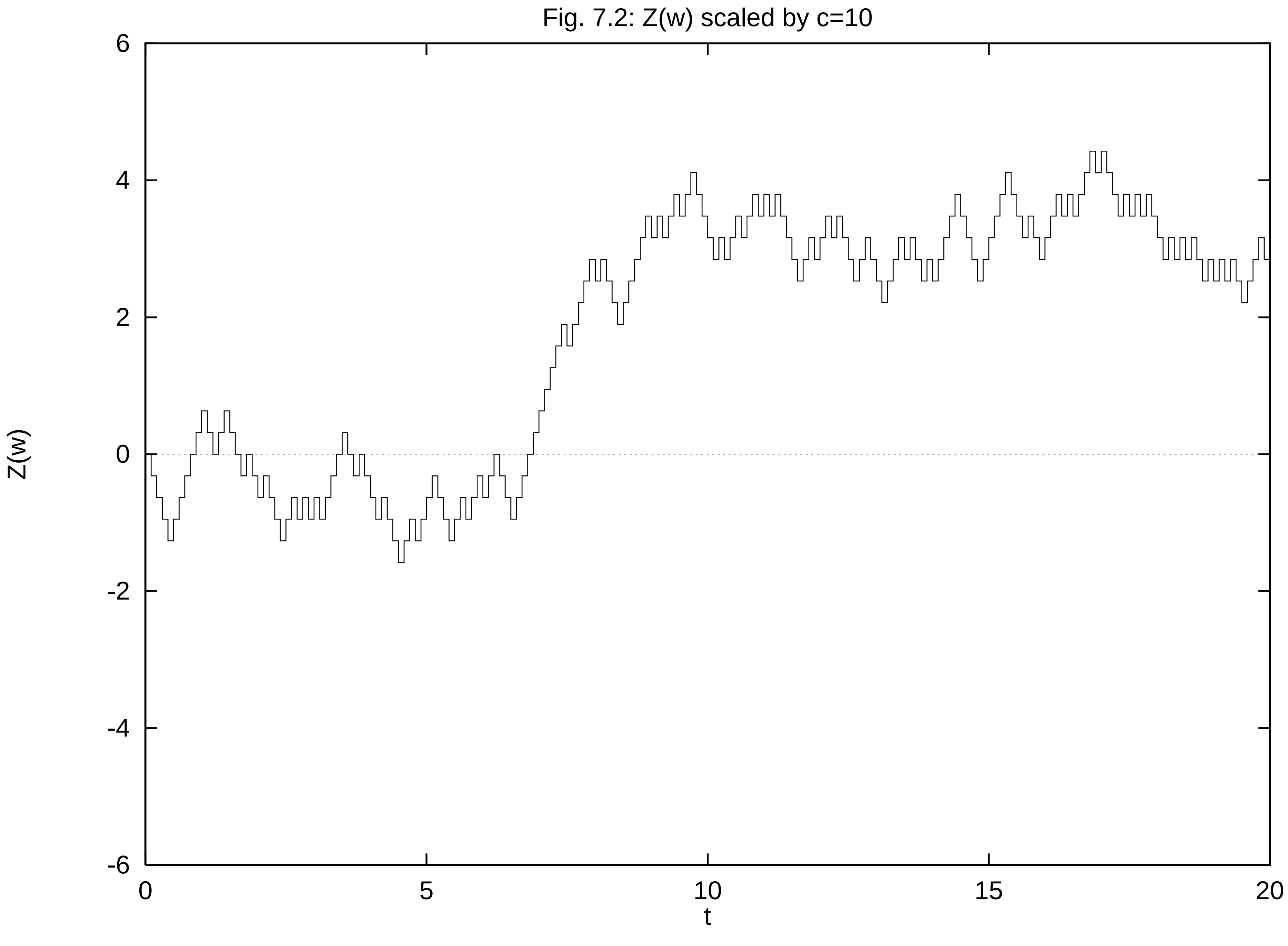


Fig. 7.2: Z(w) scaled by c=10
Z(w)
t
6
4
2
0
-2
-4
-6
0
5
10
15
20

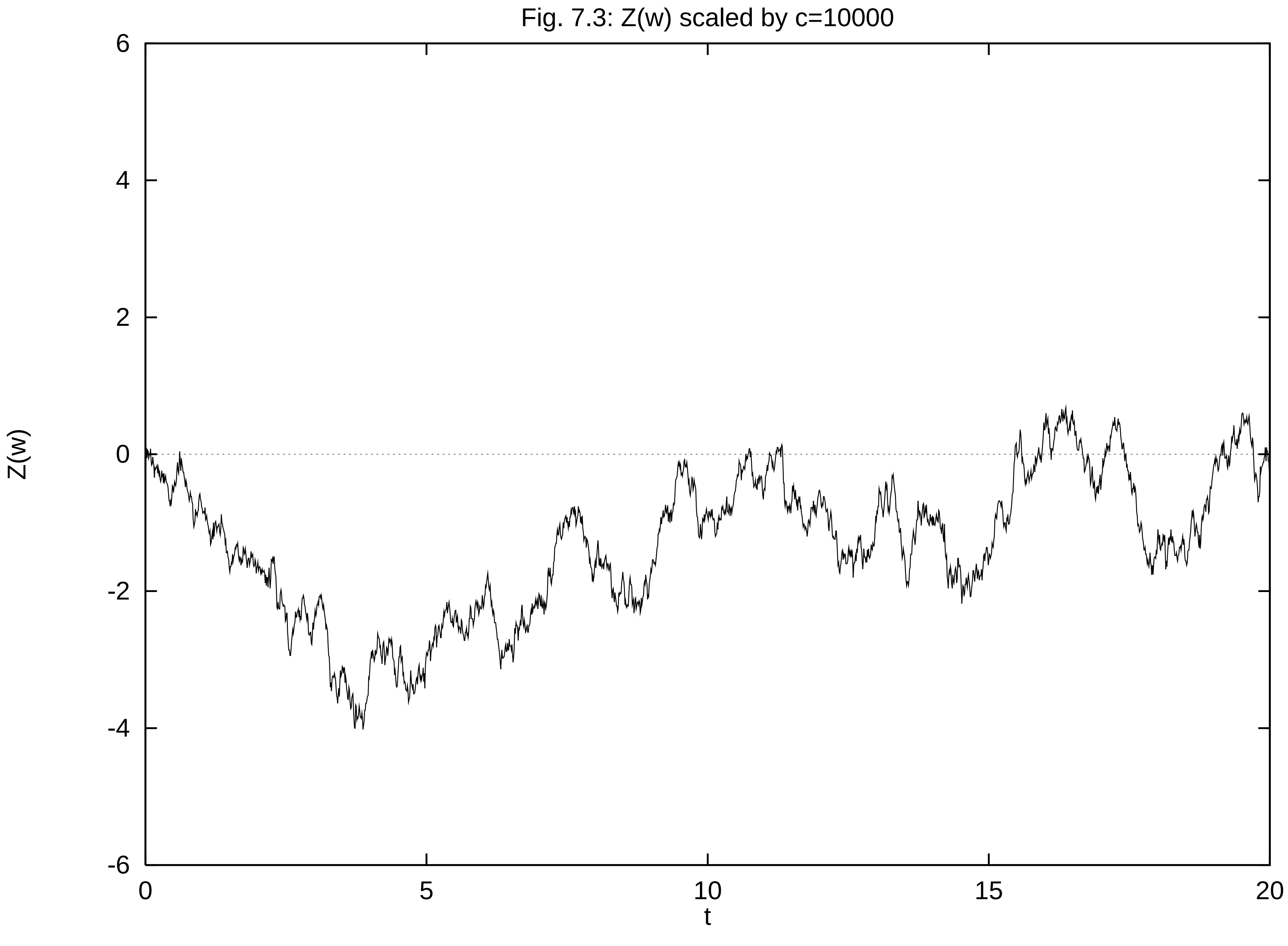


Fig. 7.3: Z(w) scaled by c=10000
6
4
2
0
-2
-4
-6
0
5
10
15
20
t
Z(w)

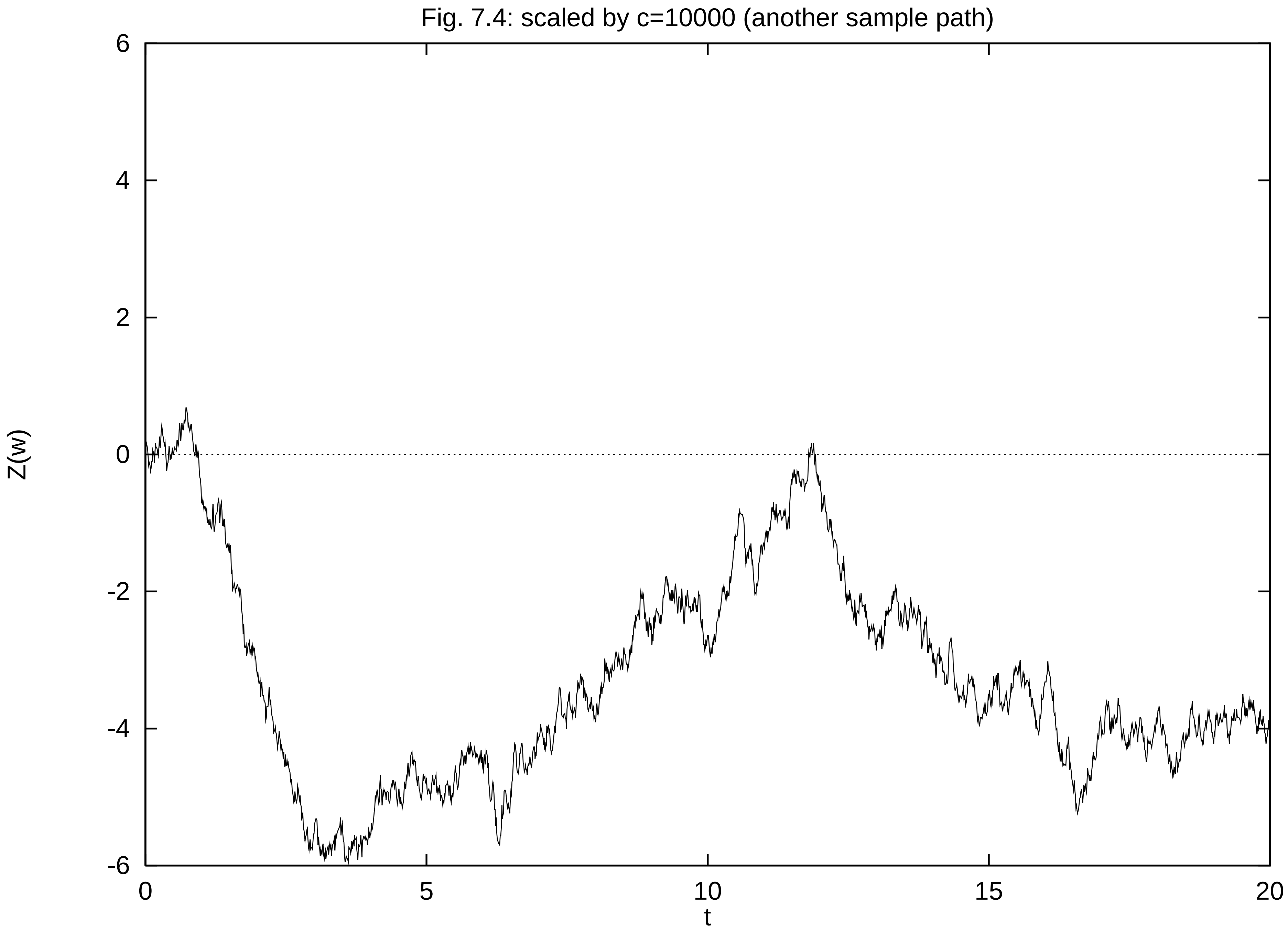


Fig. 7.4: scaled by c=10000 (another sample path)
Z(w)
t
6
4
2
0
-2
-4
-6
0
5
10
15
20

平均 m，分散 v の正規分布とは，実数上の確率であって，密度 $\rho_{m,v}(x)$ が

$$\rho_{m,v}(x) = \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{(x-m)^2}{2\,v}\right) \tag{7.1}$$

で与えられるものを言う．

Wiener 過程の種々の確率や期待値などは (7.1) を用いて計算できる．Poisson 過程と同様，一見抽象的に見えるかも知れない定義から，定理 7.1 及び (7.1) という具体的な公式が出る．例えば，$t > s$，かつ，$A \subset \mathbf{R}$ （正しくは $A \in \mathcal{B}$）のとき，$X_s = y$ という条件下で $X_t \in A$ となる確率は，

$$P[X_t \in A \mid X_s = y] = \int_{x\in A} \rho_{0,t-s}(x-y)\,dx \tag{7.2}$$

となる．ここで，事象 $A, B \in \mathcal{F}$ に対して，記号 $P[A \mid B]$ は条件付き確率

$$P[A \mid B] = P[A \cap B] \,/\, P[B] \tag{7.3}$$

を表す．B が起こったという前提の下で A の起こる確率，である．（(7.2) の導出．$X_s = X_s - X_0$ なので加法性から，左辺は $P[X_t - X_s \in A - y \mid X_s - X_0 = y] = P[X_t - X_s \in A - y]$．ここで，$A - y = \{z \mid z + y \in A\}$．これに定理 7.1 を用いれば (7.2) を得る.）特に $s = 0,\ y = 0$ とおけば，定義より $X_0 = 0$ なので，$P[X_0 = 0] = 1$．これと (7.3) から，$P[X_t \in A] = P[X_t \in A \mid X_0 = 0]$．よって，(7.2) から

$$P[X_t \in A] = \int_{x\in A} \rho_{0,t}(x)\,dx\,, \quad \rho_{0,t}(x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2\,t}\right). \tag{7.4}$$

この式は，時刻 0 に原点 0 から出発 $(X_0 = 0)$ した Wiener 過程の t 時間後の位置 X_t が平均 0 分散 t の正規分布に従うことを意味する（「確率変数の従う分布」の定義は第４回講義参照）．少し平たく言えば，時刻 t に点 x にいる相対的な確率（確率密度）が，正規分布で表される．

水の中に一滴のインキを落とすと，次第に薄まりながら広がっていく．金属棒や板の一箇所を一瞬暖めると，次第に熱は広がって全体が一様な温度になる．これらの現象は，分子の熱運動という不規則な運動の，巨視的な効果である．微視的には（インキ粒子一つは） Wiener 過程で記述される不規則な運動 (図 7.3, 図 7.4) だが，多数の粒子が集まってインキ一滴となると，微視的な運動の生じる確率に応じた割合で，密度分布が実現する（これは粒子間の運動が独立であるという仮定の下で正しい）．例えば，長細い容器に入った水にインキを落とし，落とした点を原点 0，時刻を 0 とする．t 時間後点 x でのインキの密度（薄さ） $u(t,x)$ は (7.4) の確率密度に比例する；$u(x,t) \propto \rho_{0,t}(x)$．時間がたつにつれて山が低くなる正規分布となる．微視的 (sample path) には不規則な運動でも，巨視的（密度）にはなめらかな関数になる．確率過程のおもしろさも難しさもここにある．

Wiener 過程は存在する，と書いたが,「作り方」もいろいろある．理論的にはもっとも単純とは言えないが，計算機上で作りやすく，またある観点からは極めて興味深い,「作り方」を説明する．

第５回講義で１次元酔歩 (simple random walk) という離散時間確率過程を定義した．定義を要約すると，整数時刻の整数値確率過程 Z_n で，sample path が

$$Z_n(\omega) = \sum_{m=1}^{n} \omega_m\,, \quad n = 0, 1, 2, \cdots,$$

と書けるものであった（図 7.1．X ばかり使うと紛らわしいので，simple random walk は Z で表した）．ここで m 回目の硬貨投げで，表ならば $\omega_m = 1$，裏ならば $\omega_m = -1$，とおく（確率 1/2 で ± 1 の値をとり，$\{\omega_m\}$ は独立，ということ）．

定理 7.2. 離散時空確率過程 $c^{-1/2}\,Z_{[c\,t]}$ は $c \to \infty$ のとき *Wiener* 過程 X_t に法則収束する.（Z の添字の $[c\,t]$ は，実数 $c\,t$ の整数部分をとる *Gauss* の記号.）

ここで，確率変数列 Y_c が $c \to \infty$ のとき確率変数 W に **法則収束する** とは，直感的には Y_c の従う分布 Q_c が W の従う分布 R に近づくという意味である．詳しく言うと，分布 Q_c が R に **弱収束** する，という意味である．これは，$R[\partial A] = 0$ を満たす任意の集合（正しくは，$A \in \mathcal{F}$）に対して，$\lim_{c\to\infty} Q_c[A] = R[A]$ が成り立つ，という意味である．$R[\partial A] = 0$ とは A の境界集合の，R で測ったときの確率が 0 という意味である．この一見ややこしい留保条件は,「分布が近づく」という言葉の直感的意味に近い形に弱収束を定義するための条件だが，詳しい説明は，定理 7.2 の証明とともに，割愛する．

定理 7.2 では，X_t を確率変数の集まり（確率過程）としてではなく，$\Omega \ni \omega$ から sample path （時間の関数）$X_\cdot(\omega)$ への関数（一つの関数値確率変数）と見て法則収束を論じている．次の 定理 7.3 でもその見方をしたときの確率（分布）を論じている．確率過程の２種類の見方については第５回講義も参照せよ．

定理 7.2 は，simple random walk という，計算機でもシミュレーションしやすく，数学的にも直感的にも比較的簡単な確率過程の極限として，Wiener 過程を「作る」ことができることを表している（図 7.1–7.4）．時空ともに離散的な確率過程が時空ともに連続的な Wiener 過程に近づく，というのも興味深いが，上記の定理は，次の定理の意味で，Wiener 過程の「自己相似性」を表すので，非常に興味深い．

定理 7.3. *Wiener* 過程 X_t の分布は，任意の正の実数 a に対して *Wiener* 過程を時空についてスケール変換（拡大・縮小）した確率過程 $a^{-1/2} X_{[a\,t]}$ の分布に等しい．

証明. 定理 7.2 において，c を $c\,a$ と置き換えて $c \to \infty$ としても（極限をとる変数の変数変換だから）構わない．即ち，$(c\,a)^{-1/2} Z_{[c\,a\,t]}$ は $c \to \infty$ のとき Wiener 過程 X_t に法則収束する．$(c\,a)^{-1/2} Z_{[c\,a\,t]} = a^{-1/2} c^{-1/2} Z_{[c\,(a\,t)]}$ だから，定理 7.2 において $t \to a\,t$, $Z \to a^{-1/2} Z$ と置き換えたものになっている．よって，$c \to \infty$ で $a^{-1/2} X_{a\,t}$ にも法則収束する．故に，$a^{-1/2} X_{a\,t}$ と X_t の分布は等しい． □

定理 7.2 は X_t の時間と空間をうまく呼応させてスケール変換（拡大・縮小）した後，極限をとることになっている．このような極限の取り方を **scaling limit** と呼ぶ．他方，定理 7.3 は Wiener 過程が，時間と空間をうまく呼応させてスケール変換すると，再び Wiener 過程に戻る（分布が不変である）こと，を表す．これを分布の **スケール不変性 (scale invariance)** という．物理用語では**統計的自己相似性 (statistical self-similarity)** ともいう.「統計的」という修飾語は，一つ一つの見本関数 sample path の関数としての自己相似性ではなく，分布（確率）の自己相似性に言及していることからつく．

大雑把に言うと，Brown 運動を顕微鏡で 3 倍拡大してビデオに録画し，$1/9 = 1/3^2$ 倍のスロー再生で再生画像を見ると,「元の Brown 運動らしく」見える，ということである．拡大と早回し（時空のスケール変換）の関係は古典力学でもおなじみである．怪獣映画やアクション映画で，ビルが爆破されて破片が飛び散る（放物運動）シーンを，ミニチュア撮影ですませるには，スローモーションで再生する必要がある．違いは，放物運動の場合は拡大が 3 倍ならばスピードは $1/\sqrt{3}$ 倍のスロー再生，ということである．2 乗か 1/2 乗かの違いがある．

「統計的」という修飾語にもう一つの特徴がある．放物運動は時空の適切なスケール変換で正確に元の運動に一致する（自己相似である）．Wiener 過程の sample path （例えば図 7.3, 図 7.4）は自己相似ではない．不規則な運動が拡大・スロー再生でぴたっと一致したら，かえっておかしい．あくまで「典型的な振る舞いが元の path と同じ」に過ぎない．しかし，そもそも，Wiener 過程の典型的な sample path の振る舞いとはなにか？これは直感的にすら容易ではない．図 7.3 と図 7.4 に共通するものを直感したならば，それが Wiener 過程に対する諸君の直感ということになろうか．

Wiener 過程は自然現象（例えば，分子の大きさの熱運動に由来する現象を扱う統計力学の分野）を表すには単純過ぎて，十分当てはまらない現象が多いことが分かっている．しかし，Wiener 過程とは一致しなくても，統計的自己相似性には関連すると見られる現象が普遍的に知られている．この意味で，定理 7.3 は Wiener 過程を越えて興味ある結果であり，この性質に直接結びついた Wiener 過程の「作り方」定理 7.2 は，数学的にも重要になってきている．

連続時間確率過程の実例としてあげた Wiener 過程と Poisson 過程は両方ともマルコフ性を持つ．マルコフ性は理論上も応用上も大変本質的な役割を果たす性質である．次回はこの言葉について説明したい．

数理解析特論（服部）　レポート問題７

講義（資料）を参照して，次の問１，問２，または問３の少なくとも一つに解答せよ．**提出期限：６月１５日（木）**．計算機による計算を含む場合はプログラムもつけること．Mathematica などの高級ソフトの利用も歓迎するが，どの場合でも，プログラムまたはソースをつけること．

問１．　正規分布の密度 (7.1) について，次のような公式が成り立つことが知られている．この式が成り立つように，右辺の M と V を左辺の m_1, m_2, v_1, v_2, を用いて表せ．

$$\int_{-\infty}^{\infty} \rho_{m_1,v_1}(x-z)\,\rho_{m_2,v_2}(z-y)\,dz = \rho_{M,V}(x-y)\,.$$

特に，$m_1 = m_2 = 0$ かつ $v_1 = t-u$, $v_2 = u-s$ のときの結果を，Wiener 過程の言葉で解釈せよ．

公式 $\displaystyle\int_{-\infty}^{\infty} \exp(-a\,(z-b)^2)\,dz = \sqrt{\frac{\pi}{a}}$ を用いてもよい．

問２．　定理 7.2 を用いれば第５回講義で導入した simple random walk Z_n を用いて Wiener 過程の性質を数値計算で調べることができる．適当に大きな定数 c を用いて $c^{-1/2}\,Z_{[c\,t]}$ で近似するのである．計算機で sample path を多数発生（乱数の初期値をいろいろ変えて計算）させて，固定した t （例えば $t=1$）における $c^{-1/2}\,Z_{[c\,t]}$ の値の統計（ヒストグラム）をとることにより，Wiener 過程の遷移確率密度 $\rho_{0,1}(x)$ の概形を求めてみよ．理論的結果（(7.4)）と比べてどうか？c は計算機のスピードと相談して様子を見ながら決めればよい．

参考：図 7.1–7.4 を作るのに使ったプログラムの主要部．

```
      implicit real*8(a-h,o-z)
（中略）
      parameter (ir30=1073741824, irm=48828125, r31=2147483648d0)
      data irr/1000001/
      dx=dsqrt(dt)
（中略）
1     continue
      it=it+1
      irr=irr*irm
      if (irr .lt. 0) irr=(irr+ir30)+ir30
      r=dfloat(irr)/r31
      if (r .le. 5d-1) then
      ix=ix+1
      else
      ix=ix-1
      endif
      t=dfloat(it)*dt
      x=dfloat(ix)*dx
（中略）
      if (it .eq. ilast)     goto 9999
      goto 1
（後略）
```

問３．　Wiener 過程の「作り方」には 定理 7.2 以外にもある．例えば，$\{Y_k\}$, $k=0,1,2,\cdots$, を独立同分布（第６回講義参照）な確率変数列で，各々正規分布 $N(0,1)$ （平均 0 分散 1）に従うとする．このとき

$$X_{N,t} = \frac{t}{\sqrt{\pi}}\,Y_0 + \sqrt{\frac{2}{\pi}}\sum_{n=1}^{N}\frac{\sin(n\,t)}{n}\,Y_n\,,\quad t\in[0,\pi]\,,$$

とおくと，見本関数 $X_{N,t}(\omega)$ は（確率１で） t について一様に収束し，極限 $X_t = \lim_{N\to\infty} X_{N,t}$ は Wiener 過程になる．

正規分布に従う乱数列を計算機で発生させて，$X_{N,t}$ の sample path （$0\le t\le\pi$）を生成してみよ．N を増やしていくときの変化を，定理 7.2 に基づいて生成した図 7.1–7.4 と比較して見よ．

数理解析特論

服部哲弥　1995.4–1995.6

8　Simple random walk と self-avoiding walk　ー　マルコフ性（６月１５日）

連続時間確率過程の実例としてあげた，simple random walk（第５回），Poisson 過程（第６回），Wiener 過程（第７回），はいずれも **マルコフ性** を持つ．マルコフ性を持つ確率過程（マルコフ過程）は，非マルコフ過程に比べて，一般に格段に計算が容易である．確率論以外の数学の分野との深い関連がはっきりしていて，それを利用していっそう詳しい計算ができる．そのため，応用上も広く用いられる．

いつもの通り，確率空間 $(\Omega, \mathcal{F}, P)$ と，その上の実数値確率過程 X_t が与えられているとする．即ち，X_t の状態空間は $S = \mathbf{R}$ である.（用語は，第３回から第５回あたりまでを参照.）

X_t が **マルコフ過程** であるとは，任意の自然数 n と，任意の $n+1$ 個の時刻 $s_1 < s_2 < \cdots < s_n < t$, 及び，任意の（$S$ の）事象 $A \in \mathcal{B}$，任意の n 個の S の点（つまり，値）$y_1, y_2, \cdots, y_n\ (\in S)$，に対して

$$P[X_t \in A \mid X_{s_1} = y_1,\ X_{s_2} = y_2,\ \cdots,\ X_{s_n} = y_n] = P[X_t \in A \mid X_{s_n} = y_n] \tag{8.1}$$

が成り立つことをいう．ここで，$P[A \mid B] = P[A \cap B] \,/\, P[B]$ は，B が起こったという前提の下で A の起こる確率，即ち，条件付き確率である（第７回講義）．

大雑把にいうと，マルコフ過程とは，時刻 t における位置（値）の分布 X_t は直前の時刻 s_n での位置 X_{s_n} だけで決まり，それ以前の時刻 $s_1, \cdots, s_{n-1}$ にどこにいたかによらない，ということである．第５回講義のゲームの例は，X_n の分布が直前の位置 X_{n-1} だけで定まり，それ以前の状態（過去の履歴）によらない（独立になる）．離散時間のマルコフ過程の例である．第６回と第７回で Poisson 過程と Wiener 過程の例を挙げた．両者とも定義によって加法性を満たすが,（$X_0 = 0$ を用いれば）加法性からマルコフ性が導けるので，Poisson 過程と Wiener 過程は連続時間マルコフ過程の例である．

非マルコフ的な，即ち，直前の時刻の位置だけで決まらない，というのは，小説や劇，ドラマの伏線という概念に例えられる．主人公が，直前の状況からは，突拍子もなくみえる行動に走るシーンがある．その行動の理由を説明する背景が最初の方にさりげなく描かれていた，というストーリーの構成法である．突拍子もない行動とは，そのような行動が起こる確率が小さい，ということであろう．直前の状態の下では確率の小さい行動が，過去のある条件の下で高い確率になる，即ち過去に依存する非マルコフ性がある．確率過程が過去の「記憶」を持っている，とも言えるだろう．

式 (8.1) の右辺の形の条件付き確率を遷移確率（推移確率）と言い，

$$P(s, y, t, A) = P[X_t \in A \mid X_s = y]\,, \quad (s < t, \quad y \in S, \quad A \in \mathcal{F})\,, \tag{8.2}$$

と書く．マルコフ過程は，遷移確率 $P(s, y, t, A)$ と初期分布（X_0 の分布）だけで決まることが知られている．

$$P(s, y, t, A) = P[X_t \in A \mid X_s = y] = \int_{x \in A} \rho(s, y, t, x)\, dx \tag{8.3}$$

と書けるとき，ρ を **遷移確率密度** という．第７回講義の式 (7.2) から，Wiener 過程の遷移確率密度は式 (7.1) の正規分布の密度 $\rho_{m,v}$ を用いて，

$$\rho(s, y, t, x) = \rho_{0,t-s}(x - y) \left(= \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(x-y)^2}{2\,(t-s)} \right) \right)\,, \tag{8.4}$$

と書けることが分かる．

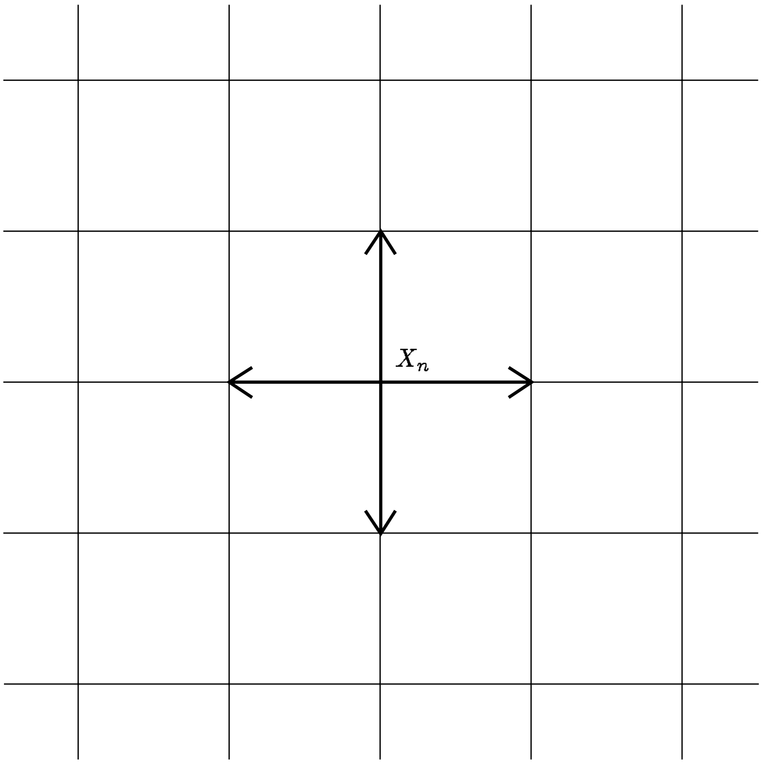


図 8.1: $d = 2$ simple random walk （矢印の各方向が各々確率 1/4）

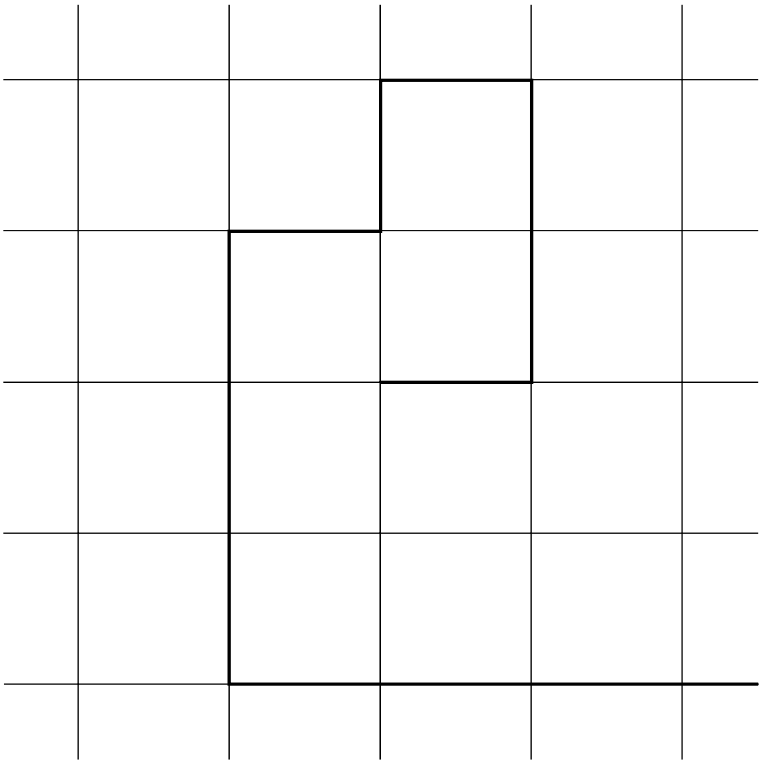


図 8.2: self-avoiding walk

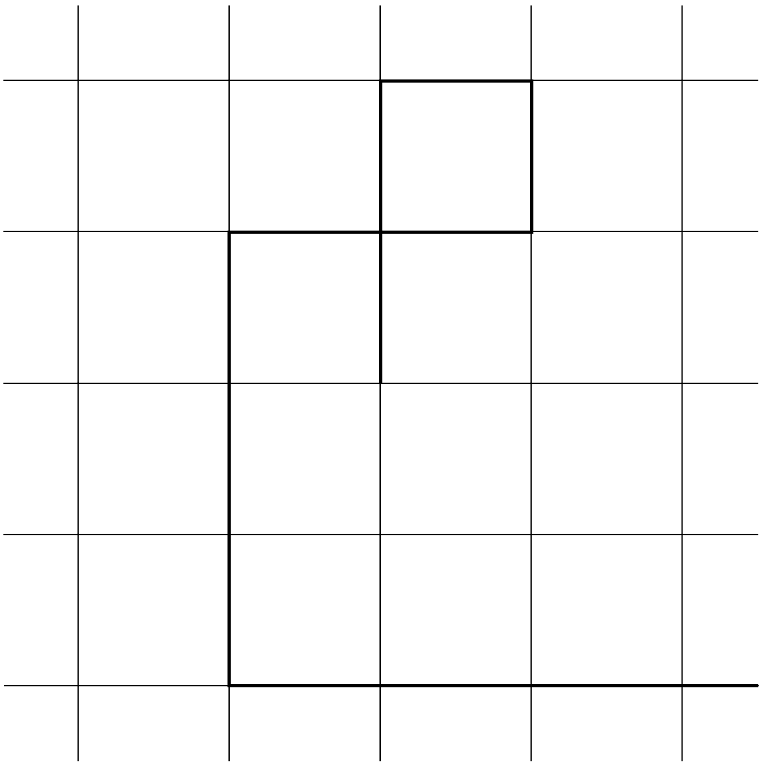


図 8.3: self-avoiding ではない walk

Wiener 過程は sample path が（確率 1 で）連続関数である．一般に sample path が（確率 1 で）連続関数になる連続時間マルコフ過程を**拡散過程** と呼ぶ．遷移確率密度が (8.4) と書ける拡散過程を **Brown 運動** という．Brown 運動で初期状態を $X_0 = 0$ としたのが第 7 回講義の Wiener 過程である．

Poisson 過程は状態空間 $S = \mathbf{Z}$ が離散的なので，遷移確率密度はないが，遷移確率が簡単に書ける．X_n が Poisson 過程の場合，第 6 回の (6.3) から直ちに次を得る（(6.3) では増分を n とおいた）．

$$P_{Poisson}(s,m,t,n) = P[X_t = n \mid X_s = m] = e^{-\lambda\,(t-s)}\,\frac{\{\lambda\,(t-s)\}^{(n-m)}}{(n-m)!}\,,\quad s<t,\ \ m\le n\,.$$

参考：現代のマルコフ過程の定義では，状態空間 S は局所コンパクト Haudorff 空間で第 2 可算公理を満たすものならばよい．また，消滅時刻 $0<\zeta\le\infty$ と消滅点 ∂ を考えて，$X_t=\partial,\ t\ge\zeta,$ をマルコフ過程の定義に加える．$\zeta=\infty$ が非消滅に対応.

マルコフ性については話がつきないのでここまでとし，Wiener 過程に戻る．あとの準備もかねて，Wiener 過程と simple random walk の状態空間 S を d 次元空間 $\mathbf{R}^d$ に拡張する．d を空間の次元を表す自然数とする.（第 5 − 7 回講義の確率過程の例は $d=1$ に対応する.）$\mathbf{R}^d$ 上の連続時間確率過程を考える．d 個の実数値確率過程を成分に持つベクトル値確率過程 $X_t = (X_{1,t}, X_{2,t}, \cdots, X_{d,t})$ ということである（数学の本ではベクトルでも $\vec{X}_t$ とせず，X_t などと書くことが多い）.

X_t が d 次元 Wiener 過程であるとは，各成分 $X_{i,t}$ が Wiener 過程であって，成分 $X_{1,\cdot}, X_{2,\cdot}, \cdots, X_{d,\cdot}$ が独立なものを言う．各成分毎に独立に計算すればよいので，第 7 回の (1 次元)Wiener 過程の公式 (7.1) 以下を用いることができる．例えば d 次元 Wiener 過程の遷移確率密度は

$$\rho(s,y,t,x) = \frac{1}{\sqrt{2\pi\,(t-s)}^{\,d}}\exp\left(-\frac{1}{2\,(t-s)}\sum_{i=1}^{d}(x_i-y_i)^2\right),\quad x=(x_1,x_2,\cdots,x_d),\ \ y=(y_1,y_2,\cdots,y_d), \tag{8.5}$$

と書けることは，成分間の独立性と (7.2) (7.1) から証明できる．d 次元 Wiener 過程もマルコフ過程である（拡散過程でもある）.

d 次元 Wiener 過程に対応して，1 次元酔歩 (simple random walk) （第 5 回講義）を d 次元 simple random walk に拡張する．1 次元 simple random walk （第 7 回講義に定義要約）は，原点から出発する整数時刻の整数値確率過程 Z_n で，sample path が $Z_n(\omega) = \sum_{m=1}^{n}\omega_m,\ n=0,1,2,\cdots,$ と書けるものであった（図 5.1，図 7.1）．ここで，$\{\omega_m\},\ m=1,2,\cdots,$ は独立で，それぞれ確率 1/2 で $+1$ または -1 となる．d 次元 simple random walk は，状態空間が d 次元格子空間 $\mathbf{Z^d}$ の場合である．即ち，Z_n が d 成分ベクトルで，各成分が整数値をとる；$Z_n = (Z_{1,n}, Z_{2,n}, \cdots, Z_{d,n})$．そして，各時刻毎に次に，現在の位置 Z_n の（$2d$ 個の）隣の点のどれかに各確率 $1/(2d)$ で移る（図 8.1．ゲーム版が 1 次元から d 次元に変わったと思えばよい）．言い換えれば，$Z_n(\omega) = \sum_{m=1}^{n}\omega_m$ であって，$\{\omega_m\},\ m=1,2,\cdots,$ は独立，かつ，各 ω_m は，それぞれ確率 $1/(2d)$ で $2d$ 個のベクトル $(+1,0,0,\cdots,0)$, $(-1,0,0,\cdots,0)$, $(0,+1,0,\cdots,0)$, $\cdots$, $(0,0,0,\cdots,+1)$, $(0,0,0,\cdots,-1)$, になる．種々の性質は 1 次元と同様である（計算は煩雑になる場合がある）．例えば，d 次元 simple random walk と d 次元 Wiener 過程の間にも第 7 回講義で紹介した scaling limit の関係がある.

定理 8.1. *d 次元離散時空確率過程 $c^{-1/2}\,Z_{[c\,t]}$ は $c\to\infty$ のとき d 次元 Wiener 過程 X_t に法則収束する.*

定理 7.2 と，c のべきの値 $-1/2$ まで含めて，全く同じ定理が全ての d で成り立つ.

確率過程の典型例が全てマルコフ過程であったように，これまでの確率過程論の研究は（応用も）マルコフ過程中心であった．マルコフ性の成り立たない例として self-avoiding walk （自己回避道）を取り上げる．Self-avoiding walk とは空間上の道であって同じ点に二度と戻らないものである.

d 次元 simple random walk で N 歩目までの道のりを考えると，原点から出発して隣の点へ移りながら N 歩でたどれる可能な全ての歩き方が等しい確率（確率 $(2d)^{-N}$）で出現する．w が d 次元格子空間 $\mathbf{Z^d}$

上の長さ N の walk であるとは $\mathbf{Z^d}$ の列 $w = (w(0), w(1), \cdots, w(N))$ であって，

$$|w(i) - w(i+1)| = 1,\ i = 0, 1, \cdots, N-1, \tag{8.6}$$

を満たすものである，と定義する．$|\cdot|$ は通常のユークリッド距離である．条件 (8.6) は，一歩毎に隣の格子点に移る，ということを表す．すると，simple random walk という確率過程の定義は，原点から出発する ($w(0) = 0$) 長さ N の walk がどれも，等しい確率 $(2d)^{-N}$ を持つ確率空間（が全ての N で定義されていること），と見ることができる．(これは第５回で説明した２つの見方のうち,「神の立場」と言えよう.）これに対して，w が d 次元格子空間 $\mathbf{Z^d}$ 上の長さ N の **self-avoiding walk** とは長さ N の walk $w = (w(0), w(1), \cdots, w(N))$ であって，

$$w(j) \neq w(i), \quad j \neq i, \tag{8.7}$$

を満たすものと定義する．W_N を原点から出発する ($w(0) = 0$) 長さ N の self-avoiding walk の集合とする．この集合の各要素に（simple random walk のときのように）等しい確率を与えたのが d 次元 self-avoiding walk の確率空間である．条件 (8.7) は一度通った点は２度と通らないという意味である（図 8.2, 図 8.3）．過去の状態（通った点）を全て覚えていないと次の動きが決まらないから，非マルコフ性を持つ典型的な例である．これが self-avoiding （自己回避的）という名前の由来である．

$$\mathrm{E}[\ |w(N)|^2\ ] = \frac{1}{C_N} \sum_{w \in W_N} |w(N)|^2$$

は self-avoiding walk が N 歩で出発点からどれくらい遠くまで到達できるかを測る期待値で，mean-square displacement と呼ばれる．C_N は 原点から出発する N 歩の self-avoiding walk の本数.

$$\nu = \lim_{N \to \infty} \frac{\log \mathrm{E}[\ |w(N)|^2\ ]}{2 \log N} \tag{8.8}$$

とおく．Simple random walk について同様の期待値をとって計算すると，d の値によらず $\nu = 1/2$ となる（定理 8.1 で c のべきが d によらず $-1/2$ になることに対応する）．$d = 1$ の self-avoiding walk とは直線上を（右か左に）まっすぐ進むことだから，$|w(N)| = N$ となるので，(8.8) より $\nu = 1$ である（第５回のゲームの例の例２が $d = 1$ self-avoiding walk）．$d \geq 2$ では self-avoiding walk の性質は難しい．$d \geq 5$ の self-avoiding walk では simple random walk と同じ値 $\nu = 1/2$ になることが Hara–Slade によって最近証明された．高い次元の self-avoiding walk の，N が大きいときの性質は simple random walk のそれに近い，という意味である．

$d < 4$ のときは self-avoiding walk の ν の値は 1/2 からずれる，と予想されている．予想値 (Flory の値) $\nu = \frac{3}{d+2}$ は数値的にはいい近似値になっている．低次元（小さい d）で値が simple random walk の値 1/2 からずれるということは，低次元空間上の self-avoiding walk では非マルコフ性（自己回避効果）が顕著に現れることを意味する．しかし，$d \leq 4$ の self-avoiding walk については証明はまだない.（参考書： N. Madras, G. Slade, The Self-Avoiding Walk, Birkhäuser, 1993 年.）

「少しだけ，確率過程」という題に前回までの３回の講義を割いた.「少しだけ」とは，分野の広がりに比べて少ししか講義できない，という意味である．確率過程は奥行きの深い分野で，Wiener 過程についてすら，紹介できなかった多くの研究成果がある．それでも，これからの数学研究の発展が期待される応用上の問題も多い．種々の有益な計算ができるほど詳しい性質が明らかになっている確率過程はマルコフ過程だけ，と言ってよく，今回紹介したマルコフ性を持たない self-avoiding walk は今後の重要な課題の一つである．非マルコフ過程は物理や経済への応用の機運が高い．例えば，株価変動は不規則に見えるが（新聞などの株価のグラフを見よ），これを適当な確率過程（の sample path）と見よう，というのである．

数理解析特論（服部）　レポート問題８

講義（資料）を参照して，次の問１，問２，または問３の少なくとも一つに解答せよ．**提出期限：６月２２日（木）**．なお，計算機による計算についてはプログラムもつけること．Mathematica などの高級ソフトの利用も歓迎するが，どの場合でも，プログラムまたはソースをつけること．

問１．　原点から出発する N 歩の２次元 $(d = 2)$ self-avoiding walk （$w(0) = 0$ で，１歩毎に平面内の４つの隣の格子点のどれかに移り，一度通った点は２度と通らない N 歩の歩き方）の総本数 C_N をいくつかの N について求めよ．例えば $C_1 = 4$, $C_2 = 12$.（計算機で大きな N まで数え上げればなお望ましい．ν の定義 (8.8) の右辺の，lim の中身（極限をとる前の量）が各 N でいくらになるかも求められればなおよい.）

問２．　(d=1) Wiener 過程の遷移確率密度 $\rho(s, y, t, x)$ は次の形の偏微分方程式を満たすことが，具体形 (8.4) を微分することで確かめられる．下式が成り立つように定数 A の値を求めよ．

$$\frac{\partial\rho(s, y, t, x)}{\partial t} = A\,\frac{\partial^2\rho(s, y, t, x)}{\partial x^2}\,.$$

一般の d について，d 次元 Wiener 過程の遷移確率密度 (8.5) はどんな偏微分方程式を満たすか？（ヒント：２回偏微分が d 次元ラプラシアンに一般化される．なお，これらの偏微分方程式は拡散方程式と呼ばれる.）

問３．　これまでの課題のうちで興味を持ったものを，引き続き発展させよ．例えば，(i) 証明し残した性質を証明する，(ii) より詳しい計算を行う，(iii) 数値シミュレーションのデータをはるかに増やす，など．

数理解析特論

服部哲弥　1995.4–1995.6

9　Sierpinski gasket 上の self-avoiding walk　（6月22日）

第8回で，マルコフ性を持たない確率過程の例として self-avoiding walk を取り上げた．復習をかねて要約する．Self-avoiding walk とは空間上の道であって同じ点に二度と戻らないものである．d 次元格子空間 $\mathbf{Z^d}$ 上の長さ N の walk，$w=(w(0),w(1),\cdots,w(N))$ のうちで，自己回避条件 (8.7) $w(j)\neq w(i)$, $j\neq i$, を満たすものを self-avoiding walk と呼ぶ．通った点を全て知らないと可能な動きが決まらないから，マルコフ性を持たない．

原点から出発する $(w(0)=0)$ N 歩の self-avoiding walk の集合を W_N とおき，W_N の要素の数，即ち，原点から出発する N 歩の self-avoiding walk の本数を C_N とする．W_N の各要素に等しい確率を与えたのが（N 歩の）d 次元 self-avoiding walk の確率空間である．N 歩の self-avoiding walk 1本あたり確率 $1/C_N$ ということである．N が有限のとき，全体集合が有限集合だから，原理的には場合の数を数えれば確率が計算できる（第3回講義参照）．例えば，期待値 $\mathrm{E}[\ \cdot\ ]$ は N 歩の self-avoiding walk の上の関数（確率変数） $f(w)$ に対して $\mathrm{E}[\,f\,]=\dfrac{1}{C_N}\displaystyle\sum_{w\in W_N} f(w)$ で計算される．N が有限のときは原理的問題はないから，N が大きくなるとき諸量がどのように変化していくかが問題になる．Mean-square displacement $\mathrm{E}[\,|w(N)|^2\,]$ は self-avoiding walk が N 歩で出発点からどれくらい遠くまで到達できるかを測る．その指数

$$\nu=\lim_{N\to\infty}\frac{\log \mathrm{E}[\,|w(N)|^2\,]}{2\log N} \tag{9.1}$$

は，mean-square displacement の N が大きくなるときの変化の度合いを表す量である（(8.8) と同じ式）．「気持ち」は，$2\nu\log N\sim\log \mathrm{E}[\,|w(N)|^2\,]$ だから,「典型的な self-avoiding walk」(の sample path) について $|w(N)|\sim N^{\nu}$ とおいたことになる．自己回避条件 (8.7) を落とした simple random walk，及び，$d\geq 5$ の self-avoiding walk では $\nu=1/2$ であり，$d=1$ の self-avoiding walk は直線運動と同じで $\nu=1$ である．低次元空間上の self-avoiding walk は非マルコフ性（自己回避効果）が顕著に現れて，$2\leq d<4$ のとき simple random walk の値 $\nu=1/2$ からずれる，と予想されている．しかし，非マルコフ性の強い場合を研究する方法は知られていないため，この予想の証明はない．

$d=1$ self-avoiding walk は直線運動で，やさしすぎて参考にならない．$d=2$ self-avoiding walk は既に難しすぎる．その間を埋める空間として，Sierpiński gasket（シールピンスキーガスケット）及び，3次元 Sierpiński gasket と呼ばれるフラクタル空間の上の self-avoiding walk を研究した．成果を，証明抜きで，簡単に紹介したい.（文献：K. Hattori, T. Hattori, S. Kusuoka, *Self-avoiding paths on the pre-Sierpiński gasket,* Probability Theory and Related Fields **84** (1990) 1–26.　K. Hattori, T. Hattori, *Self-avoiding process on the Sierpiński gasket,* Probability Theory and Related Fields **88** (1991) 405–428.　T. Hattori, S. Kusuoka, *The exponent for mean square displacement of self-avoiding random walk on Sierpiński gasket,* Probability Theory and Related Fields **93** (1992) 273–284.　K. Hattori, T. Hattori, S. Kusuoka, *Self-avoiding paths on the three dimensional Sierpiński gasket,* Publications of RIMS **29** (1993) 455–509.　解説記事：服部哲弥, フラクタル上の確率過程, 数理科学 (1991.10) 42–48.）

Sierpiński gasket とは正三角形を3つつないで1辺が2倍の正三角形を作る，その操作を繰り返して得られるフラクタル（自己相似図形）である（図9.1）．図で Oa_1b_1 を3つ正三角形状につないで Oa_2b_2 を作る．同様に Oa_nb_n 3つから $Oa_{n+1}b_{n+1}$ を得る．$\overline{Oa_0}=\overline{Ob_0}=1$ としておく．通常，フラクタルは自己相似的に無限に細かい構造をいれたものを考える．ここでは最小単位を1として外に大きくしていく．その（通常のフラクタル図形との）区別のために以下 pre-Sierpiński gasket と呼ぶことにする．Pre-Sierpiński gasket 上で原点 O から出発する $(w(0)=O)$ N 歩の self-avoiding walk の集合 W_N を考える．Self-avoiding

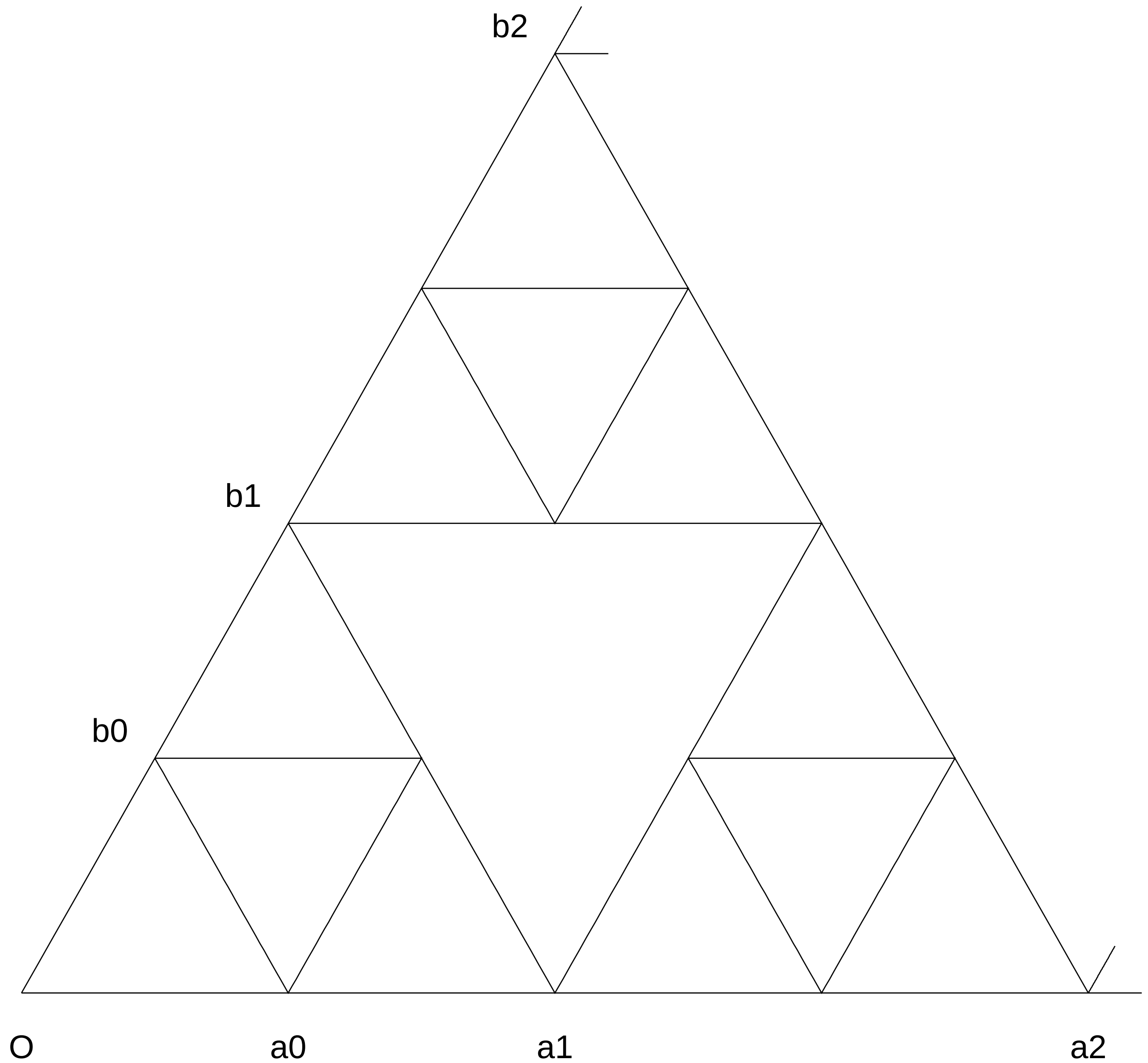


Fig. 9.1

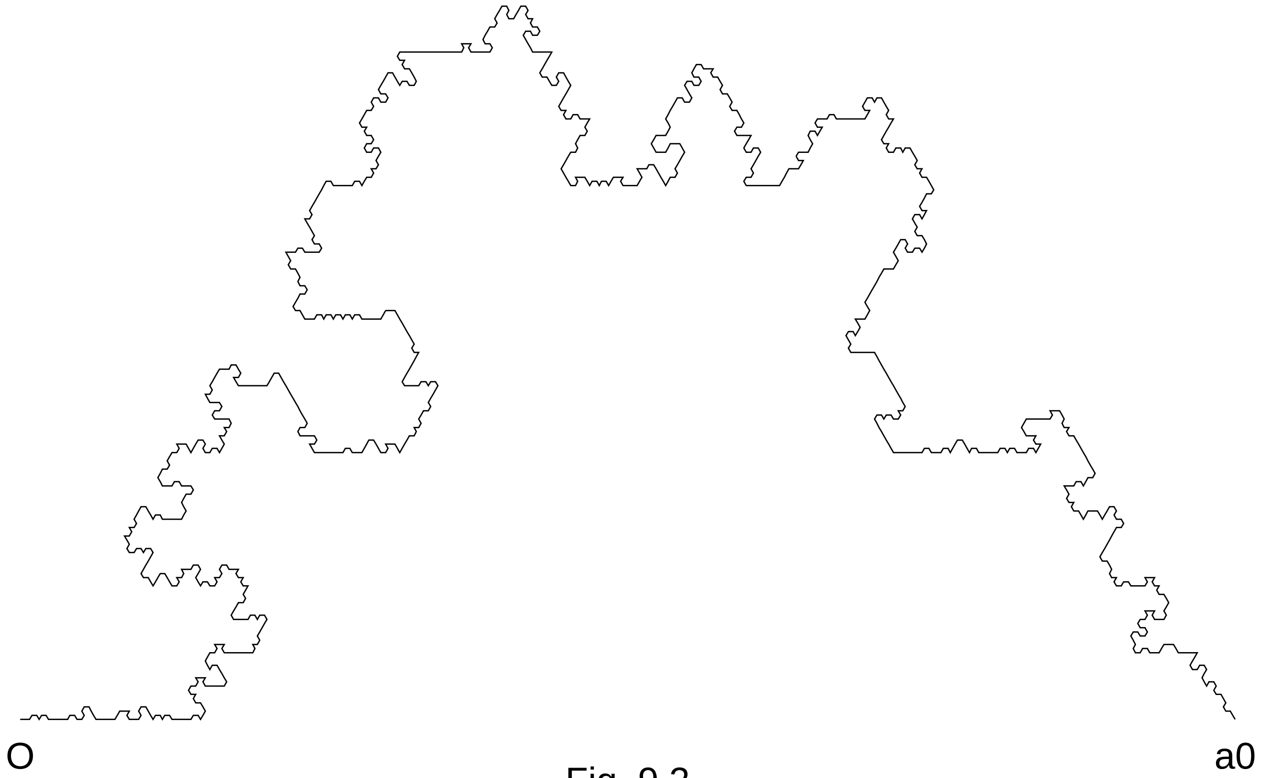


Fig. 9.2

walk の定義は $\mathbf{Z^d}$ のときと同様で，自己回避条件 $w(j) \neq w(i)$, $j \neq i$, を満たす walk である．Walk は pre-Sierpiński gasket の頂点（原点 O または，４本の線分の交点）の列 $w = (w(0), w(1), \cdots, w(N))$ であって，全ての i について隣り合う２点 $w(i)$ と $w(i+1)$ は pre-Sierpiński gasket の線分で結ばれているものである．(Pre-Sierpiński gasket の線分の上しか歩けない.）自己回避条件のために，self-avoiding walk はマルコフ性を持たない．$\mathbf{Z^d}$ の時と同様に，W_N の要素の数を C_N とし，W_N の各要素に等しい確率 $1/C_N$ を与える．Mean-square displacement $\mathrm{E}[\,|w(N)|^2\,]$ の定義も ν の定義 (9.1) も $\mathbf{Z^d}$ の時と同様である．ν は，$d = 2, 3, 4$ の格子空間 $\mathbf{Z^d}$ ではいくつになるか数学的に証明ができていない量であるが，pre-Sierpiński gasket 上の self-avoiding walk では ν の値が数学的に求まる．

定理 9.1. *Pre-Sierpiński gasket* 上の *self-avoiding walk* に対して *mean-square displacement* の指数 *(9.1)* は $\nu = \dfrac{\log 2}{\log(7-\sqrt{5})/2} = 0.79862\cdots$ である．特に，この値は *pre-Sierpiński gasket* 上の *simple random walk* の対応する指数 $\log 2/\log 5 < 0.5$ と異なる．

フラクタル上では一般に simple random walk の ν は $1/2$ からずれる．その計算は省略するが，結論は，pre-Sierpiński gasket 上では self-avoiding walk と simple random walk の ν は異なる．自己回避効果（非マルコフ性の影響）が証明できたことを意味する．

原点から出発する N 歩の self-avoiding walk の本数 C_N の N が大きいときの増大の様子（漸近的振る舞い）も，$d = 2, 3, 4$ の格子空間 $\mathbf{Z^d}$ では数学的には求められていない．Pre-Sierpiński gasket 上の self-avoiding walk ではこの漸近的振る舞いが分かる．これを説明するために，予備的結果が必要である．

n を非負整数とする．Pre-Sierpiński gasket 上で原点 O から出発する $(w(0) = O)$ self-avoiding walk のうち，b_n を通らずに，a_n にたどり着くものに注目する．何歩目で a_n にたどり着くかは歩き方によって異なるが，とにかく，O から a_n までの walk で自己回避条件を満たすもの (self-avoiding walk) の集合を $W^{(n)}$ とする．(W_N と同じ W を使ったが，添字の位置で区別した．また，$W^{(n)}$ を考える際は，W_N 上に定義した確率は忘れる.）$w \in W^{(n)}$ に対して， w の長さ，即ち，w という歩き方で O から a_n まで何歩かかったか，を $L(w)$ とする．$w(L(w)) = a_n$．そして，実数 β に対して

$$\zeta_n(\beta) = \sum_{w \in W^{(n)}} e^{-L(w)\,\beta} \tag{9.2}$$

とおく．$\zeta_n(0)$ は $W^{(n)}$ の要素の個数になるが，すぐ分かるように，個数は $n \to \infty$ で発散するので，$\lim_{n\to\infty} \zeta_n(0) = \infty$. 定義から $\zeta_n(\beta)$ は β の単調減少関数なので，$\lim_{n\to\infty} \zeta_n(\beta)$ も（極限があれば） β の単調減少関数である．従って，全ての $\beta \geq 0$ で $\lim_{n\to\infty} \zeta_n(\beta) = \infty$ であるか，さもなければ，ある有限値 $\beta_c \geq 0$ があって，$\beta < \beta_c$ では $\lim_{n\to\infty} \zeta_n(\beta)$ は発散，$\beta > \beta_c$ では $\lim_{n\to\infty} \zeta_n(\beta)$ は収束，となる（もし，極限が存在すれば）．実際後者のようになっている．

定理 9.2. 以下が成り立つ定数 β_c が存在する．

(1) $\beta < \beta_c$ ならば $f_0(\beta) = \lim_{n\to\infty} 3^{-n} \log \zeta_n(\beta) > 0$ が存在する．従って，特に $\zeta_n(\beta)$ は発散する．

(2) $\beta = \beta_c$ ならば $\lim_{n\to\infty} \zeta_n(\beta) = \dfrac{1}{2}\,(\sqrt{5}-1)$,

(3) $\beta > \beta_c$ ならば $f_1(\beta) = \lim_{n\to\infty} 2^{-n} \log \zeta_n(\beta) < 0$ が存在する．従って，特に $\lim_{n\to\infty} \zeta_n(\beta) = 0$.

$e^{\beta_c} = 2.2880\cdots(\beta_c = 0.827691\cdots)$ である．原点から出発する N 歩の self-avoiding walk の本数 C_N の漸近的振る舞いが，この β_c を用いて記述できる．

定理 9.3. $\displaystyle \lim_{N\to\infty} \frac{1}{N} \log C_N = \beta_c\,.$

即ち，$C_N \sim e^{\beta_c N} \sim 2.2880^N$ くらいの速さで総本数が増大する．１歩毎に平均 $2.2880\cdots$ 行ける方向がある，ということである．Pre-Sierpiński gasket の各交点には 4 本の線分が集まっている．自分が来た方向には戻れないのは明らかだが，残り 3 本のうち平均「約 0.7 本」の方向は既に通っていることになる．

第７回講義で，simple random walk が Wiener 過程に法則収束するという定理 7.2 (scaling limit) を紹介した．離散時間確率過程が連続時間確率過程に近づく，ということである．同様の scaling limit は pre-Sierpiński gasket 上の self-avoiding walk にも存在する．$W^{(n)}$ （pre-Sierpiński gasket 上で原点 O から出発し，b_n を通らずに，a_n にたどり着く self-avoiding walk の集合）に確率を定義する．有限集合なので，要素毎に確率を定義すればよい．$w \in W^{(n)}$ に対して，$L(w)$ を w の長さ（a_n までの歩数）としたが，これを用いて， $W^{(n)}$ の上に次式によって確率 P を定義する．

$$P[\{w\}] = \frac{1}{\zeta_n(\beta_c)}\, e^{-L(w)\,\beta_c}\,, \quad w \in W^{(n)}\,.$$

正しく $P[W^{(n)}] = 1$ となることは (9.2) から明らか. 確率空間 $(W^{(n)}, P)$ 上の確率過程 Z_N, $N = 0, 1, 2, \cdots$, を $Z_N(w) = w_N$ ($N \le L(w)$ のとき), $Z_N(w) = a_n$ ($N > L(w)$ のとき), で定義する．即ち，単に self-avoiding walk の path そのものだが，a_n にたどり着いた後はそこにとどまる，という意味である．

定理 9.4. $2^{-n}\, Z_{[\lambda^n t]}$ は，$n \to \infty$ で，ある連続時間確率過程 X_t に法則収束する．X_t の *sample path* は（確率 1 で）連続関数であり，$X_0(\omega) = O$, $X_\infty(\omega) = a_0$, を満たし，しかも，自己回避的 *(self-avoiding)* であって，さらに，*path* のハウスドルフ次元が $\dfrac{\log\lambda}{\log 2} > 1$ である．ここで $\lambda = \dfrac{1}{2}\,(7 - \sqrt{5})$.

即ち，時空ともに連続な確率過程であって，自己回避的なもの X_t が存在する．Path のハウスドルフ次元が 1 より大きいとは，軌跡がいくら細かく見てもぎざぎざしている（従って至るところ微分不可能な）ことを表す（図 9.2）．無限にぎざぎざしているのに一度通った点は２度と通らない確率過程があることを示している．

同様の結果は，正四面体 4 個からより大きな正四面体を作る操作を繰り返して得られる，３次元 pre-Sierpiński gasket の上の self-avoiding walk についても成り立つことが証明できる．しかし，これより複雑な図形では self-avoiding walk の問題は未解決である．基本的に前回から今回にかけて述べたことが，self-avoiding walk の漸近的性質について精密に分かっていることの全てである（世界最先端の研究！）．これは，Wiener 過程や simple random walk について分かっている性質に比べれば，ほんのわずかである．マルコフ性がない場合に有力な研究方法がないことが，simple random walk に比べて成果の極端に少ないことの原因である．

９回の講義で確率論の基礎から始めて，確率過程の入門的紹介を行った．典型的な確率過程というものがどのようなことを表現しうるのか，それを表現するために数学的にはどのような言葉遣いをするのか，どのような量が計算できればよいのか，どのような量は計算しやすいのか，について少しずつでも感じがつかめてもらえたならば講義は成功だと思う．

取り上げなかった題材が多い．特に，条件付き期待値とマルチンゲールは基礎的である．抽象化を避けるために σ-加法族の説明も最小限にとどめたが，確率過程を議論する上で，これが足枷になった．当初講義で予定した条件付き期待値を行えなかったのも，σ-加法族の理解という問題があったのが原因の一つである．離散時間確率過程の場合は直感的に考えて殆ど定義（定式化）を間違えることはないと思うが，連続時間確率過程の定義には事象の数学的定式化としてのσ-加法族の理解は必要である．離散時間でも，時間について極限をとるときは気をつけないといけない場合がある．Wiener 過程やマルコフ過程には講義でふれなかった多くの性質が知られている．興味と必要に応じて専門書を読んでほしい．

現在研究が進行中の話題については，非マルコフ過程以外にもいろいろある．生物種の分布と環境の関係や，DNA の突然変異が新種として集団に固定する確率を調べる集団遺伝学，素粒子の力学を扱う場の量子論，なども２０世紀後半になってから進歩が著しい問題であり，確率過程の進んだ話題として，確率場，確率超過程，状態空間が測度空間であるような確率過程，弦（紐）の確率過程，といった名前で研究が進行中である．２次元画像の問題も一般的には確率場（時間変数が２つ以上ある確率過程）の問題になる．現実は数学より難しい．

数理解析特論（服部）　レポート問題９

講義（資料）を参照して，次の問に解答せよ．**提出期限：６月２９日（木）**．なお，計算機による計算についてはプログラムもつけること．Mathematica などの高級ソフトの利用も歓迎するが，どの場合でも，プログラムまたはソースをつけること．

問．これまでの課題のうちで興味を持ったものを，引き続き発展させよ．例えば，(0) 選択しなかった問題を解く，(i) 証明し残した性質を証明する，(ii) より詳しい計算を行う，(iii) 数値シミュレーションのデータをはるかに増やす，など．

次ページのアンケートにも回答をください．

数理解析特論

情報工学科　博士前期課程　１９９５年度前期

アンケート

服部哲弥

講義に参加して下さってどうもありがとうございます．この講義を改善するために，ご意見を頂きたく思います．下記の項目に答えて下さい（項目に関わらず自由感想も歓迎します）．よろしくお願いします．

問１ 受講の仕方について，該当するものにまるをつけて下さい．

電子メール中心
手渡し中心
主に友達のをコピーしてすませた
その他（途中から切替，など）：

問２ 電子メール講義に参加された方について（手渡し混在の方を含む），この方法と通常の教室講義を比べた場合の，あなたにとって良かった点（助かった点など）と，悪かった点（困った点など）を，それぞれ挙げて下さい．

良かった点.

悪かった点.

問３ 手渡し方式を利用された方（電子メールと混在の方を含む）について，電子メールでは不都合だった理由について，該当するものにまるをつけて下さい（複数回答可）．

電子メールをよく知らない（慣れてない）
電子メールで入手した LaTeX または ps ファイルの処理方法が分からない
便利な環境にない（電子メールが届かない，プリンターが古い，等）：
手渡しの方が慣れている
（レポート提出に関して）手書きの方が早い
（提出に関して）図を postscript にする便利な方法がない（知らない）
コピーですませた
その他（コンピュータは嫌いだ，など具体的に）：

問４ 講義資料及びレポートの分量について，該当するものにまるをつけて下さい.

(i) **（ほぼ）全て１度は読み通した　半分くらいは読んだ　レポートに必要な部分くらい**
(ii) **書き方は分かりやすかった　分かりにくかった**
(iii) **内容は物足りなかった　ちょうど通常の講義程度の手応え　多すぎた**
(iv) **レポートは毎週やるには１回あたり難しすぎた　物足りないことの方がちょっと多かった**

問５ 講義資料の内容について，(a), (b) のうち該当する方にまるをつけて下さい.

(i)
(a) 数学的に基礎的な定義から順番に講義してほしい
(b) 基礎的な定義はなるべく省いて計算の仕方を中心に教えてほしい

(ii)
(a) １回あたり読む分量が増えてもいいから，今より分かりやすくしてほしい
(b) もっと説明を省いてもいいから，１回あたり読む分量を減らしてほしい

問６ 講義資料の内容について，問５ (i)(ii) 以外の点についても具体的な改善の提案があれば書いて下さい.（「別のテーマ $\cdots$ をやってほしかった」といった大枠でも,「n ページ m 行目に漢字の間違い」といった細かいことでも.）

問７ メールの送り方について，何かご意見がありましたら書いて下さい.（「確認のメールが多すぎる」,「メールは短く」,「催促はもっとして」，など）

問８ 今回はこちらにソフトもハードも知識がなかったため，もっとも原始的なメールの使い方しかしませんでした．電子メール講義の役に立ちそうなソフト，ハード，知識の提案があれば是非お願いします.

問９ その他感想や希望がありましたら自由に書いて下さい.