入門講座

やさしい待ち行列 (2)——等間隔運転は待ちを減らす

高橋　幸雄

前回は流体モデルを使って、サービス要求量が処理能力を上回ったときにおきる待ち行列の性質について調べました。今回は、バスや電車が等間隔で運転するのとバラバラな間隔で運転するのでは、待ち時間がどう違ってくるのかを見てみましょう。

このような議論にはどうしてもランダムネスの概念が必要です。そのために確率を使わざるを得ません。確率と聞くと難しい式がいっぱい並んでいるのを思い出して、それだけで厭になってしまう人もいるかもしれませんが、ここでは平均やモーメントなどごく簡単なものしか使いませんので、ぜひお付き合いください。確率によってモデルの幅がぐっと広がることがおわかりいただけると思います。

1.　バスや電車の待ち時間

まずバスや電車での待ち時間を考えてみましょう。

等間隔で到着するバス

バスが一定の間隔、たとえば5分間隔、で到着するバス停を考えましょう (図1)。のっけからクイズです。

> **クイズ 1**
> あなたがバスの時刻表を知らずにバス停にやってきたとすると、つぎのバスがくるまでに平均、どのくらい待たされるでしょうか。

そんなに難しく考えることはありません。この場合、平均で到着間隔の半分、2.5分くらい待たされる、ということは直感的にわかるでしょう。

これを確率論的に説明すると、つぎのようになります。あなたが到着する時刻はバスの到着とは無関係に決まるわけですから、あなたの到着時刻はバスとバス

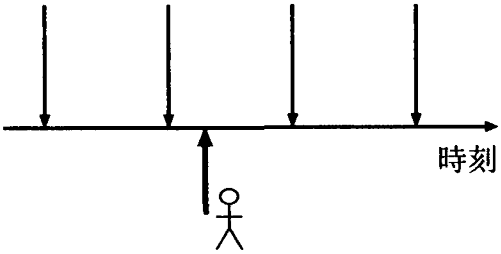

図 1: バスの発車間隔

の間で一様に分布すると考えられます。したがって、あなたの待ち時間の分布は区間 $[0,5)$ 上の一様分布となり、その平均は 2.5 分です。むろん、5 分以上待たされることは決してありません。

急行と各駅停車のある電車

では、もう少し難しい問題を考えてみましょう。ある私鉄では、ある時間帯、急行と各駅停車が交互に発車します。ただし等間隔ではなく、急行が発車して 4 分後に各駅停車が発車し、その 6 分後につぎの急行が発車します (図 2)。ではクイズ、

> **クイズ 2**
> あなたは急行でも各駅停車でも、どちらか先に発車する電車に乗るものとします。ではあなたが電車の発車時刻とは無関係に駅に着いたとすると、あなたが駅に着いてから電車が発車するまでに平均どのくらい待たされるでしょうか。

電車は平均すると 5 分に一本の割合でくるのですから、前のバスのときと同様、その間隔の半分 2.5 分だけ平均で待たされる、と考えた方がおられるかもしれません。実はもうちょっと待たされるのです。

まず、あなたは急行と各駅停車のどちらに乗る可能性が高いかを考えてみてください。図 2 から、6 : 4 で

たかはし ゆきお　東京工業大学 大学院 情報理工学研究科 数理・計算科学専攻
〒152 東京都目黒区大岡山 2-12-1

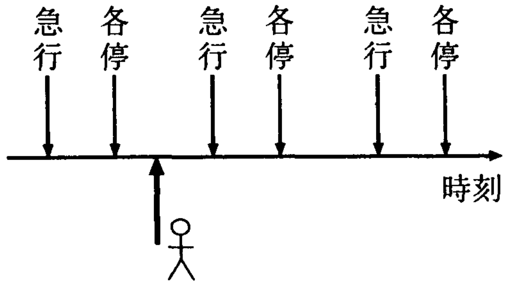

図 2: 電車の発車間隔

急行に乗る可能性が高いことがわかります (だから急行は混むんですよね)。急行に乗るときは平均で 3 分待たされるし、各駅停車に乗るときは平均で2分しか待たされないのですから、結局あなたが待たされる平均時間は

$$E\{W\} = \frac{6}{10} \times 3 + \frac{4}{10} \times 2 = 2.6\ (分)$$

となって、2.5分よりすこし長くなるのです。最大の待ち時間も 5 分ではなく 6 分です。

このように、完全な等間隔でくる場合とそれから少しずれる場合とでは、等間隔でくる場合の方が客の待ち時間は少なくなるということがわかります。ただこの場合は、その差はそれほど大きくはありません。

2. ランダムな到着

それではもっとばらつく場合として、バスの到着間隔が指数分布にしたがうときを考えてみましょう。

クイズ 3

バスの到着間隔が互いに独立で平均 5 分の指数分布にしたがう場合、あなたがバスの到着とは無関係にバス停に着いたとすると、あなたの待ち時間の平均はどのくらいでしょうか。

この場合も平均で5分に1台の割でバスがくるのですから、待ち時間はだいたい平均 2.5 分、一定間隔ではないので少し長くなってもおよそ 2.7 分か 2.8 分くらい、と考えられた方が多いのではないでしょうか。じつはもっと待たされるのです。このことをみるために、指数分布独特の性質を説明しておきましょう。

指数分布の無記憶性

指数分布の特徴は“無記憶性”とか“マルコフ性”とか呼ばれる特有の性質です。確率変数 X が任意の $x > 0$ に対して $\mathrm{P}\{X > x\} = e^{-\alpha x}$ となるとき、X の分布をパラメータ α の指数分布といいます。この分布の平均は $1/\alpha$ です。このとき任意の正数 x と y に対して、つぎの式が成り立ちます。

$$\mathrm{P}\{X > x + y \,|\, X > x\} = \mathrm{P}\{X > y\} \qquad (1)$$

実際、左辺は「条件 $X > x$ のもとで $X > x + y$ となる条件付き確率」ですが、条件付き確率の定義から

$$\begin{aligned} &\mathrm{P}\{X > x + y \,|\, X > x\} \\ &= \mathrm{P}\{X > x + y,\, X > x\} / \mathrm{P}\{X > x\} \\ &= \mathrm{P}\{X > x + y\} / \mathrm{P}\{X > x\} \\ &= e^{-\alpha(x+y)} / e^{-\alpha x} = e^{-\alpha y} \end{aligned}$$

となり、(1) の右辺と等しくなります。

(1) の性質は、公衆電話をかけている人をイメージすると、つぎのように解釈できます。

X をある人の通話時間としましょう。その人が話し始めてから x 時間たったときに次の人が電話をかけようとして、公衆電話のところにきたものとします。すると (1) の左辺は、後からきた人が y 時間以上待たされる確率を表しています。ところが右辺は、いままでに x 時間話していたことは忘れてしまって、その時点から話をはじめて、y 時間以上話をする確率です。これらが等しいのですから、“後からきた人がどれだけ待つか”ということに関しては、“それまでどれだけ話をしていたか”ということは無関係になるのです。

クイズの答えに戻りましょう。バスの到着間隔は指数分布ですから、この無記憶性から、あなたの待ち時間は、直前のバスがいつ行ってしまったかには無関係で、前のバスが発車した直後にあなたが到着したと考えたときの待ち時間、つまりバスの到着間隔、と同じ分布にしたがいます。ですからあなたの平均待ち時間はバスの平均到着間隔と同じ5分です。したがってあなたはバスが等間隔でくる場合の 2 倍も待たされるのです。

ポアソン過程

ここでの到着パターン、つまり到着間隔が互いに独立で同一の指数分布にしたがうような到着パターンはポアソン過程と呼ばれ、確率論的に最もランダムな到着パターンと考えられています。これは次のような考察からわかります。

いま α を考えているポアソン過程のパラメータ、n を大きな正整数、T を n/α として、時間区間 $(0, T)$ を

考えます。この区間の中に、n 個の点をランダムにばらまきます。つまり、ひとつひとつの点を互いに独立に $(0,T)$ 上の一様分布にしたがってとるのです。これらの点の位置を小さい順に $S_1, S_2, \cdots, S_n$ とすると、これらの点によって分割されてできた $n+1$ 個の小区間

$$(0,S_1),\ \ (S_1,S_2),\ \ \cdots\ \ ,\ \ (S_{n-1},S_n),\ \ (S_n,T)$$

の長さの分布はみな同じになります (これはちょっと不思議な性質ですが、対称性をうまく使うと証明できます)。ここで $n \to \infty$ とした極限を考えると、これらの小区間の長さは互いに独立になり、共通の分布はパラメータが α の指数分布になります。実際、最初の区間 $(0,S_1)$ の長さの分布を考えると、

$$\begin{aligned} \mathrm{P}\{S_1 > x\} &= \mathrm{P}\{n \text{ 個の点が } (x,T) \text{ に入る }\} \\ &= \left(\frac{T-x}{T}\right)^n = \left(1 - \frac{\alpha x}{n}\right)^n \\ &\to e^{-\alpha x} \quad (n \to \infty) \end{aligned}$$

と、指数分布であることがわかります。独立性の証明は少し複雑になりますが、はじめの k 個の区間の長さの同時分布を考えることによって示せます。

ランダムなときにより長く待たされる理由

バスの到着がランダムだとより長く待たされる理由を考えてみましょう。これにはバスの到着間隔 X がある確率密度関数 $f(x)$ をもつ分布にしたがうものとして、平均待ち時間を計算してみるとよいでしょう。

まず、バスとバスの間隔が長いときと短いときでは、あなたはどちらのケースに出会いやすいか考えてみてください。急行電車と各駅停車の場合には 6 分と 4 分の 2 種類の間隔があって、それらに出会う確率はそれぞれ 6/10 と 4/10 でした。このように、一般にあなたが長さ x の間隔に到着する可能性は x に比例すると考えられます (ここが一定でないことが直感を狂わせる原因になっているのです)。

ところで長さ x の区間が出現する可能性は確率密度関数 $f(x)$ で与えられますから、あなたが長さ x の間隔の間に到着する可能性は x と $f(x)$ の積に比例することになります。これを表す確率密度関数を $g(x) = Kxf(x)$ と書くと、比例定数 K は $g(x)$ の $[0,\infty)$ での積分が 1 となるように決められます。X の期待値は

$$\mathrm{E}(X) = \int_0^\infty x f(x)\,dx$$

ですから、$K = 1/\mathrm{E}(X)$ であることがわかります。

このように、あなたが長さ x の間隔に到着する確率密度関数は $g(x) = xf(x)/\mathrm{E}(X)$ で、このときの (条件付き) 平均待ち時間は $x/2$ です。あなたの平均待ち時間 W はこれらの積を $[0,\infty)$ で積分すれば求められます。

$$\begin{aligned} W &= \int_0^\infty \frac{x}{2} g(x)\,dx = \frac{1}{2\mathrm{E}(X)} \int_0^\infty x^2 f(x)\,dx \\ &= \frac{\mathrm{E}(X^2)}{2\mathrm{E}(X)} \end{aligned} \tag{2}$$

ここで $\mathrm{E}(X^2)$ は X の 2 次のモーメントと呼ばれる特性量で、分散とは $\mathrm{E}(X^2) = \mathrm{V}(X) + \{\mathrm{E}(X)\}^2$ という関係にあります。

このように、平均待ち時間 W は X の 2 次のモーメントを平均の 2 倍で割ったものになっています。ですから、待ち時間は到着間隔のばらつきが大きいほど長くなるわけです。$\mathrm{E}(X^2)$ を実際に計算してみると、到着間隔が一定の場合には $W = \mathrm{E}(X)/2$ となり、指数分布にしたがうときは $W = \mathrm{E}(X)$ となることが確かめられます。

ダンゴ運転のときの平均待ち時間

(2) から、平均到着間隔が同じならば、到着間隔が一定のときがもっとも待ち時間が小さいことがわかります。ランダムな到着のときは、平均待ち時間は一定間隔のときの 2 倍ですが、これが最大というわけではありません。意地悪な例を作ればいくらでも平均待ち時間を大きくできるのです。たとえばダンゴ (団子) 運転をしているときを想定して、大きい m に対して

$$X = \begin{cases} \dfrac{a}{m} & \text{確率 } \dfrac{m}{m+1} \text{ で} \\[2ex] ma & \text{確率 } \dfrac{1}{m+1} \text{ で} \end{cases}$$

という分布を考えてみましょう。これは平均して $m+1$ 回に 1 回、ma という長い区間が出現し、それ以外は a/m という短い区間が続く、というケースです。短い間隔が続いている間、ダンゴ状態になっていると考えます (図 3)。

この到着間隔 X の期待値は a で、2 次のモーメントは $\mathrm{E}(X^2) = (m^3+1)a^2/(m^2+m)$ です。したがって平均待ち時間は (2) から

$$W = \frac{m^3+1}{2(m^2+m)} a \approx \frac{m}{2} a \ \ (m \text{ が大きいとき})$$

と、一定間隔のときのおよそ m 倍になります。これは m が大きければ、あなたの到着はほとんどの場合

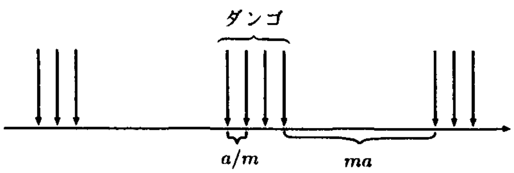

図 3: バスのダンゴ運転

ダンゴとダンゴの間に入ってしまって、つぎのダンゴがくるのを待たなければならない、という状況を反映しています。

このようにダンゴ運転では、等間隔運転に比べてはるかに長い時間待たされるのです。

3. 電車は混むと遅れる

これまで、バスや電車が等間隔で到着しないと客はより長く待たされることをみましたが、じつはそれが原因となって混雑と遅れの正のフィードバックが出現することがあります。

電車の遅れと混雑

たとえば東京の山手線や大阪の環状線を思い浮かべてください。電車が混み出すと、乗客の乗り降りによけいに時間がかかるようになります。すると電車はすこし遅れ気味になって、前の電車との間隔が少しだけ長くなります。するとつぎの駅で乗ってくるお客の数がそれだけ多くなり、これがさらなる遅れの原因となります。これを繰り返しているうちに、混雑による遅れは雪だるま式にふくれあがってしまいます。

電車遅れのモデル

この現象をごく簡単なモデルで考えてみましょう。各駅は等間隔に配置されていて、客の到着率や降車割合は皆全く同じとします。平常時には電車は等間隔で運転され、スムースに客をさばくことができます。あるとき、なにかの原因で1台の電車が遅れはじめたとしましょう。この電車の乗客数と遅れ時間がひと駅ごとに変わっていく様子を考えます。

いま、注目している電車が m 人の乗客を乗せて前の電車の s 分後にある駅を発車したものとします。この電車が次の駅を発車するときの乗客数 m' と前の電車との間隔 s' は

$$\begin{cases} m' = (1-\beta)m + rs \\ s' = s + b(\beta m + rs) - c \end{cases} \tag{3}$$

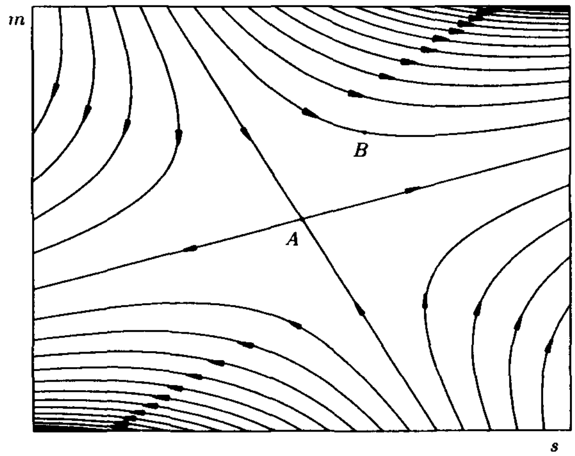

図 4: 電車の乗客数と運転間隔の推移

となると仮定しましょう。この式は次のような考えから導かれています。

次の駅では乗客 m 人のうち βm 人が降り、ホームで待っていた rs 人の客が乗り込みます。したがって新しい乗客数は $m - \beta m + rs$ となります。ここで r は単位時間当たりの客の到着率で、rs は前の電車が駅に着いてからこの電車が駅に着くまでに到着した客の数を表します。発車間隔 s' は、それまでの間隔 s から、前の電車の客の乗降に必要な時間 c (前の電車は平常に運転しているものとしてこれは定数として扱います) を引いて、この電車の乗降に必要な時間 $b(\beta m + rs)$ を加えたものになっています。ここでは、乗降に必要な時間は乗降客数に比例するものとしていますが、もっと実際に近い関数を導入することも可能です。ただ、この関数の形は以下の議論ではあまり本質的ではありませんので、ここでは簡単なものにしておきましょう。

パラメータの値を適当に決めて (3) による (m,s) から (m',s') への動きを図示すると、図4のようになります。横軸が s で、縦軸が m です。この図の中央には (不安定な) 不動点 A が1つあって、そこを通る斜め右下がりの直線の上と下では、(m,s) の動きが全く違ってきます。

この斜めの直線から右上の領域にある点からスタートすると、線に沿っていずれ斜め右上の方向へ出ていってしまいます。これは電車の間隔が開き乗客の数が大幅に増加することを意味します。逆に斜めの直線の左下からスタートすると、電車の間隔は縮まり乗客

の数も減っていきます。

平常時の電車の間隔と乗客数は、おそらく不動点のすこし左下にあるはずです。このあたりでは、放っておけば電車の間隔はだんだん狭くなるのですが、そうはならないように各駅で少しずつ発車時刻を遅らせて調整しています。

何かの原因で (m, s) がこの斜めの直線の右上にでてしまったときが問題です。放っておけば前の電車との間隔はどんどん開き、後続の電車との間隔が詰まってきます。すると後続の電車も遅れるのでその後との間隔も縮まります。このように、山手線を走っている電車はいくつかのダンゴにかたまってしまいます。ダンゴのスピードは先頭の電車のスピードでおさえられますから、単位時間に運べる客の数、という意味で、山手線全体のサービス能力が低下して、集まってくる客をさばききれなくなります。こうなると駅のホームには人があふれ、駅の外ではホームに入れない人が騒ぎだす、という最悪の事態になるわけです。

このように混雑と遅れの間に存在する正のフィードバックは、朝のラッシュアワーの電車の運行に大きな障害となります。では、

> **クイズ 4**
> 図4のBのような状態になったとき、事態をこれ以上悪化させないで平常状態に近づけるには、どのように対応策をとったらよいでしょうか。

図4から明らかなように、何とかして状態を斜線の左下の領域へ移動しなければなりません。それには m を下げるか s を小さくするかです。乗っているお客さんに降りてもらうわけにはいきませんので m を下げるのは難しいのですが、s は簡単に下げられます。

「前を走っている電車の発車を意図的に遅らせる」

のです。これはじつにうまい手でしょう。この手の効果は絶大で、電車に乗るときによく注意していると、JRでもしばしばこれを使っていることがわかります。

読者の中には、乗客の乗降にかかる時間を短くするため "尻押し" や "はぎとり" をすることを考えた方もおられるかもしれません。また、入場制限や他の私鉄や地下鉄による振り替え輸送を提案される方もおられるでしょう。ただ、これらの対応には手配などに時間がかかりますし、費用もかかります。まず、前の電車を遅らせてみて、それでもダメなときにとる手でしょう。電車のドアを広くしたり数を増やすこともそれなりに効果的ですが、これは突発的な事態に対応するというよりは山手線の輸送力そのものを増強することになります。

なお、図4の不動点は、式(3)で $m' = m, s' = s$ とおいて解くことにより、$m = c/2b\beta, s = c/2br$ のように求めることができます。不動点を通る斜線の方程式は、式(3)の右辺の変数 m, s に対する係数行列の固有値と対応する固有ベクトルに基づいて求められます。計算は少しやっかいですが、力自慢の方は挑戦してみてください。

4. エレベータのダンゴ運転

状況は少し違いますが、高層ビルのエレベータの運行にも、このような混雑と遅れのフィードバックがあります。何台かの(たとえば4台の)エレベータが組になって、上下に行ったりきたりして、お客さんを各階に運んでいるものとしましょう。

ここでは簡単のため、どのエレベータも最下階から最上階まで上行し、また最下階まで降りていくものとします。つまり途中で向きを変えて引き返すということはしないことにするのです。こうすると、各エレベータは図5のように円運動をしていると考えることができ、見通しがよくなります。必要ならば同じ階の上りから下りにはグルッとまわって時間ゼロで移動できるとすれば、途中での引き返しも扱うことができるでしょう。

もし各エレベータが追い越しをしないことにすると、このエレベータの動きは山手線とそっくりです。どんな配置から出発してもしばらくするとかならずダンゴ運転になることは容易に想像できます。

じつは、たとえ追い越しを認めたとしても、客の要求にすべて反応しようとすると、じきにダンゴ運転が始まってしまいます。前のエレベータとの間隔があくと途中で待っている客が多くなり、電車と同様の理由でその間隔はどんどん開いていってしまうのです。最近、エレベータを自動運転しているデパートがありますが、こういうところではダンゴ運転がよく見られます。みなさんもぜひシミュレーションで確かめてみてください。では、

> **クイズ 5**
> エレベータのダンゴ運転を避けるにはどのようなオペレーションをしたらよいでしょう。

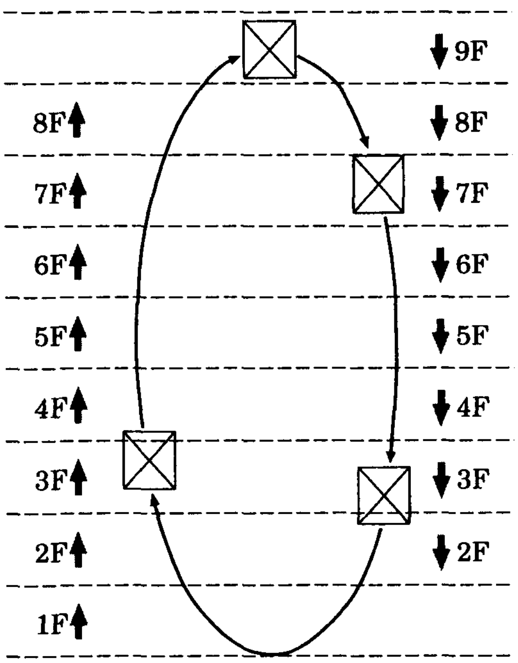

図 5: 円運動するエレベータ

電車と比べて、エレベータの場合には次のような特徴があります。まず、エレベータでは、動いている状態から停まって扉を開閉して客を乗降させ、再びスタートするのにその運行時間の大部分を消費して、階の移動や通過にはほとんど時間がかかりません。60 階ものビルの 1 階から最上階まで 40 秒足らずで到達できてしまうのですから、本来のスピードはたいへん速いのです。また、電車と違ってレールに相当する縦穴がエレベータごとに確保されています。ですから追い越しや通過も自由にできます。エレベータの定員が少ないことも特徴の一つです。降りる客がいなければ途中の階を通過することも可能です。

このように、エレベータは個々の動きの自由度が大きいので、この自由度を最大限に利用すればダンゴ運転を避けて効率的に運用することができるはずです。ポイントは、待ち行列の感覚でいえばサービス率を上げること、つまり余計なところには停まらず、目的階までなるべく早く到達することです。これにはある階で待っている客がいたら、その客を拾うエレベータを決め、他のエレベータはそこにはなるべく停まらないように仕組むのです。つまり積極的に通過させるのです。こうすればダンゴ運転にもならず、一定の時間でさばくことのできる客の数を大幅に増やすことができます。

昔のエレベータでは (あるいは今でも台数が少ないエレベータでは)、待っている客にどのエレベータが何階にいるかはっきりとわかるようになっていました。ところが最近のエレベータでは、エレベータが到着する少し前にそこのランプが点灯して客に予告するだけで、待っている客にはどのエレベータが何階にいるかいっさいわからないようになっています。これは、せっかく待っているのに目の前を通過されてしまった客を怒らせないための配慮なのだそうです。なるほど、プロはいろいろ考えるものですね。

実際のエレベータは大変複雑な管理プログラムによって制御されていて、ダンゴ運転が起きないように、さらには客の待ち時間が大きくならないように工夫されています。エキスパートシステムなどを取り入れていろいろな制御ルールを考え、その効果をシミュレーションで確かめながら管理プログラムを開発しているそうです。

混雑が増すと処理能力が落ちる現象

混みだすと処理能力が減少しどんどん混雑が激しくなる、という現象は、鉄道やエレベータばかりではありません。道路でもたくさんの車が流入してくるとその処理能力が減少してしまうことがよく知られていますし、最先端の情報通信ネットワークの中でもこのような現象が起きています。

このような現象はなかなか複雑で、まだモデル化されておらず定量的な解析ができていないケースもたくさんあります。これらの挙動をきちんと把握することは大変重要なことですので、若い人にもぜひ挑戦していただきたいと思っています。

本稿の 1 節と 2 節の内容のより詳しい取り扱いについては [1] をご参照ください。[1'] は [1] の第 1 版の翻訳です。本稿でも図は大学院生の大原久樹君にお願いしました。次回は標準的な待ち行列モデルについて解説するつもりです。

参考文献

[1] W. Feller, *An Intrtoduction to Probability Theory and Its Applications, Volume II*, 2nd ed., Wiley, 1971.

[1'] W. フェラー著, 国澤清典監訳, 「確率論とその応用 II, 上, 下」, 紀伊國屋書店, 1969.