

# Level Crossing 法による $GI/G/1$ 待ち行列モデルの解析

小林 香・片山 勁

(工学部電子情報工学科)

現代生活においては、ある限られた資源を共有するという状況は頻繁に発生する。そしてその際、各利用者には自分が資源を利用できるまでの待ち時間を出来るだけ短くしたいという要求がある。銀行の ATM 機を例に挙げると、以前は ATM 機毎に個別に客の行列が出来る、単一処理装置+単一待ち行列を構成単位とするシステムが複数並列するものであったが、複数処理装置+単一待ち行列というフォーク型のシステムに変化してきたことは、誰しも思い当たることであろう。様々な待ち行列システムの中で最も一般的な  $GI/G/1$  待ち行列モデルは、利用者のシステム到着間隔が独立でかつ同一の一般分布に従う到着過程、処理時間分布が一般分布、単一処理装置という構成単位を持っている。この中でも、利用者の到着過程がポアソン分布に従う  $M/G/1$  待ち行列モデルについては、任意時刻における系内利用者数の母関数を解析することで、細かい部分まで既に解明されている。しかし、本来待ち行列システムを解析する大きな目的となっている、待ち時間分布を求めるには、母関数による解法は間接的であり、その理解が難しいものとなっている。

そこで、本稿では、系内に存在する残余仕事量に着目し *Level Crossing* 法を導入することで、より直感的に  $GI/G/1$  待ち行列モデルにおける待ち時間分布を求め、この手法の有効性を論議する。

キーワード : *Level-Crossing* 法,  $GI/G/1$  待ち行列モデル, 残余仕事量, 待ち時間分布

## 1 $GI/G/1$ 待ち行列モデル

Kendall の記号を用いて  $GI/G/1$  と表記される待ち行列システムは、その処理時間分布が一般分布  $H(t)$  に従う単一の処理装置 (サーバ) と、処理装置が他の利用者により使用中である時、自分の処理の順番が廻ってくるまで利用者が待つことのできる無限長の待ち行列 (バッファ) から構成される。  $r$  番目の利用者 (客, 通信呼, パケット等) のシステムへの到着時点列を  $\{\tau_r, r = 0, 1, \dots\}$ , 到着間隔列を  $\{A_r = \tau_r - \tau_{r-1}, r = 1, 2, \dots; \tau_0 = 0\}$ , 処理間隔列を  $\{H_r, r = 1, 2, \dots\}$  と表せば、  $r = 1, 2, \dots$  に対して独立な  $A_r$  と  $H_r$  は

$$A(t) \equiv Pr\{A_r \leq t\}, \quad H(t) \equiv Pr\{H_r \leq t\}$$

各々同一の分布関数  $A(t)$ ,  $H(t)$  を定義する。この時要請されることは、利用者のシステムへの到着間隔が独立でかつ同一の一般分布に従う到着過程であることであり、その分布関数としては何ら具体的な関数形が与えられる必要はない。

$GI/G/1$  待ち行列モデルの一つである  $M/G/1$  待ち行列モデルは、利用者の到着が到着率  $\lambda$  のポアソン過程 (Poisson

process) に従うものである。ポアソン到着過程は 1) 到着する利用者間の独立性, 2) 時間的な定常性, 3) 微小時間内における希少性, という 3 つの仮定を満すランダムな到着である。この時、利用者のシステムへの到着間隔は指数分布に従う。

$$A(t) = Pr\{A_r \leq t\} = 1 - e^{-\lambda t} \quad (1)$$

この指数分布には無記憶性という性質があり、数学的な取り扱いが容易である特徴を持つ。なお式 (1) より、任意の時間帯  $(\tau, \tau + t]$  の到着利用者数  $N_r(t)$  は、ポアソン分布に従い

$$Pr\{N_r(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \dots \quad (2)$$

と表される。  $M/G/1$  モデルの常套的な解析法として、利用者の処理終了直後におけるシステム内利用者数に着目する隠れマルコフ連鎖法や、システム内利用者数の他に残余処理時間 (あるいはサービス経過時間) を状態変数として取り入れる補助変数法が知られている。待ち行列モデル  $M/G/1$  におい

て、処理時間分布  $H(t)$  の  $n$  次積率を  $E(H^n)$ ,  $n = 1, 2, \dots$ , そのラプラス・スチルチェス変換 (LST: Laplace-Stieltjes transform) を  $H^*(s)$  と表し、利用者の処理終了 (退去時点) 直後におけるシステム内利用者数の定常分布  $\{\pi_j\}$  の母関数を  $\Pi(z)$  と表す。利用率 (utilization factor)  $\rho = \lambda E(H) < 1$  のときにシステムは安定となり、時間の経過とともに定常状態 (平衡状態) へ近づく。非割り込みの処理規律 (先着順, ランダム順など) の下に、母関数  $\Pi(z)$  は

$$\begin{aligned} \Pi(z) &= \frac{\pi_0(1-z)}{1-z/H^*(\lambda-\lambda z)}, \\ \pi_0 &= 1-\rho \end{aligned} \quad (3)$$

で与えられる。先着順処理規律の下では、ある利用者  $c$  の系内時間  $\theta$  内に到着する利用者数と  $c$  の退去時点の系内利用者数は等しく且つ  $c$  の系内時間  $\theta$  と  $\theta$  内の到着利用者の計数過程  $\{N_r(t)\}$  は独立であるから、待ち時間  $W$  の定常分布  $W(x)$  と系内時間  $\theta$  の定常分布  $\Theta(x)$  の LST をそれぞれ  $W^*(s)$ ,  $\Theta^*(s)$  と表せば、 $\theta$  の間に到着する利用者数に関して

$$\begin{aligned} \Pi(z) &= \Theta^*(\lambda(1-z)) \\ &= W^*(\lambda(1-z))H^*(\lambda(1-z)) \end{aligned} \quad (4)$$

の関係が成立し、式 (3),(4) より  $W^*(s)$  について式 (5) を得る。

$$W^*(s) = \frac{(1-\rho)s}{s-\lambda+\lambda H^*(s)} \quad (5)$$

## 2 Level Crossing 法による解析

### 2.1 GI/G/1 待ち行列モデルにおける サンプルパス

今、到着順 (FIFO) に処理される GI/G/1 待ち行列モデルに対する、利用者の適当な到着状況を仮定する。利用者  $c$  がシステムに到着する毎に、その利用者に対する処理時間に相当する仕事をシステムに持ち込むことになるので、時刻  $t$  におけるシステム内に存在する全仕事量  $W_t$  の一例を描画すると図 1 のようになる。これをサンプルパスと称する。観測開始時刻  $t=0$  でのシステム内利用者数 0 を初期状態とし、1 番目、2 番目、3 番目に到着した利用者 ( $c_1, c_2, c_3$ ) の到着時刻を各々  $ta_1, ta_2, ta_3$  とする。 $c_1$  の到着時点ではシステム内利用者数 0 のため、 $c_1$  は待つことなく処理を受ける (待ち時間  $W_1 = 0$ )。  $ta_1$  時点でのシステムの仕事量

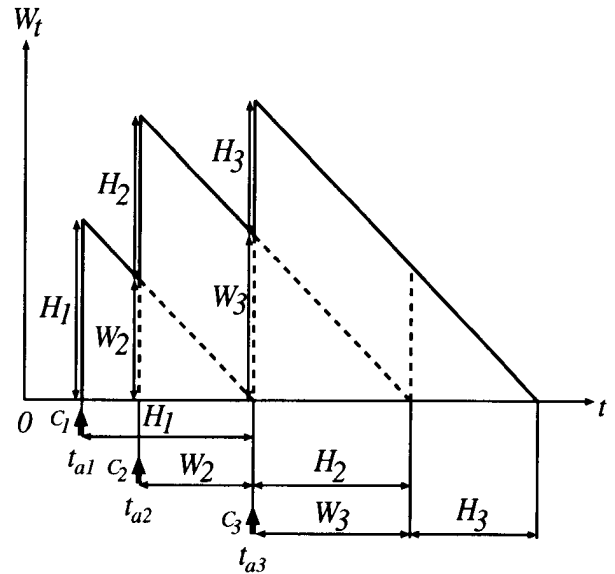


図 1: sample path

$W_t$  は  $c_1$  の持ち込んだ仕事量  $H_1$  に相当する分増加する。時間の経過とともに処理装置が処理した分、仕事量  $W_t$  は減少する。 $c_1$  の処理中に到着する  $c_2$  は時刻  $ta_2$  に  $H_2$  の仕事量をシステムにもたらし、自分に対する処理が開始されるまでに  $W_2$  待つことになる。また  $c_3$  の到着時点  $ta_3$  では、この利用者がシステムに持ち込む仕事量  $H_3$  だけ、全仕事量  $W_t$  は増加し、この時点で  $c_1$  はシステムを退去しているものの、 $c_2$  が処理を受けているのでその終了を待つ (待ち時間  $W_3$ )、 $c_3$  に対する処理が開始される。つまり、ある利用者の到着時点におけるシステムの全仕事量  $W_t$  が、その利用者が処理を受けるまでに必要な待ち時間であることが分かる。

システムが安定に存在し、利用者の待ち時間が発散しないために、処理装置の稼働状況すなわちシステムの利用率  $\rho = \lambda E(H) < 1$  であることが必要である。

### 2.2 Level Crossing 法による $W^*(s)$ の導出

図 1 のサンプルパスにおいて、任意の  $x(x > 0)$  となる直線を設定する。観測時間  $(0, T)$  においてシステムの仕事量  $W_t$  がこのレベル  $x$  の直線を、上から下に横切る回数を  $N_{down}(T)$ 、下から上に横切る回数を  $N_{up}(T)$  と表すものとする。 $N_{down}(T)$  と  $N_{up}(T)$  は高々 1 しか異なることはなく、十分長い間観測することでその時間的な平均は等しいものとなる (図 2)。

$$\lim_{T \rightarrow \infty} \frac{N_{down}(T)\Delta t}{T} = \lim_{T \rightarrow \infty} \frac{N_{up}(T)\Delta t}{T} \quad (6)$$

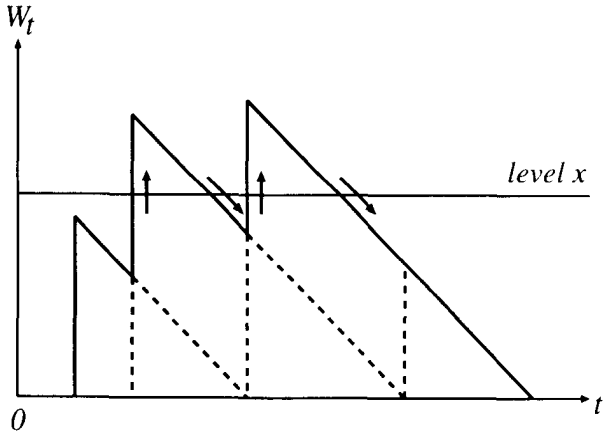


図 2:  $N_{up}(T)$  and  $N_{down}(T)$

エルゴード性により  $\lim_{T \rightarrow \infty} \{N_{down}(T)\Delta t/T\}$  及び  $\lim_{T \rightarrow \infty} \{N_{up}(T)\Delta t/T\}$  は全体の観測時間  $T$  に対して、システム内仕事量が  $x$  から  $x + \Delta x$  の間にある時間的な割合を示している。

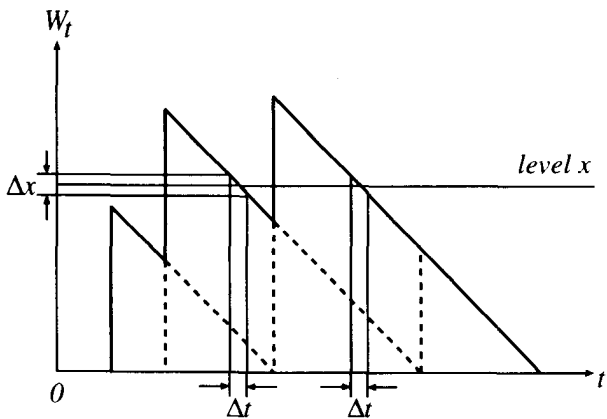


図 3: downward crossing case

観測時間  $(0, T)$  において、処理装置が利用者を処理することでシステムの仕事量  $W_t$  がこのレベル  $x$  の直線を、上から下に横切る場合、図 3 に示すように、横切る微小時間  $\Delta t$  の間に仕事量は  $\Delta x$  変化する。定義より、残余仕事量に対応する待ち時間  $W_t$  の確率密度関数  $w_t(x)$  は次のように得られる。

$$w_t(x)\Delta x \equiv Pr[x < W_t \leq x + \Delta x] \quad (7)$$

従って、

$$\lim_{T \rightarrow \infty} \frac{N_{down}(T)\Delta t}{T} = w_t(x)\Delta x \quad (8)$$

を得る。

次に観測時間  $(0, T)$  においてシステムの仕事量  $W_t$  が

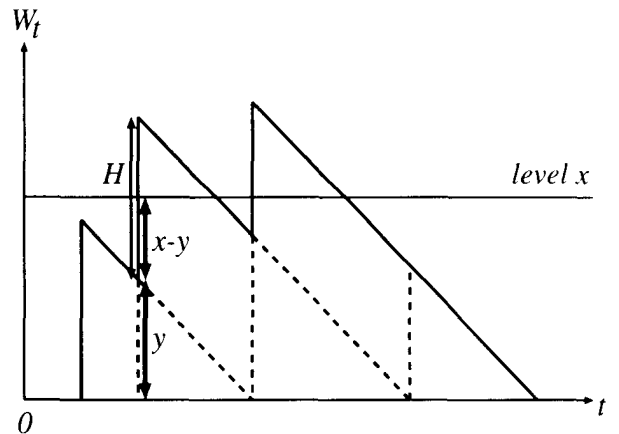


図 4: upward crossing case

このレベル  $x$  の直線を、下から上に横切る場合を考える。 $GI/G/1$  システムにおいて想定されるケースは、新たな利用者のシステム到着時の残余仕事量 (レベル  $x$  より下) と、その利用者がシステムにもたらす仕事量の和がレベル  $x$  を越える場合である。つまりある利用者の到着時の残余仕事量を  $y$  ( $y \leq x$ ) とすると、この利用者のもたらす仕事量である処理時間  $H$  が  $x - y$  より大きいと、残余仕事量のパスはレベル  $x$  を横切ることになる。ここでレベル  $x$  の値の取りうる範囲を考える。新規に利用者が到着してその処理時間に対応する仕事量をシステムにもたらすことを既に述べたが、処理時間がゼロとなる利用者の到着はありえない(そのような利用者は到着していないのと同じ扱いになる)ことから、条件  $x > 0$  を考慮する必要がある。確率  $Pr[H > (x - y)]$  は処理時間分布  $H(x)$  の補分布  $\bar{H}(x) = 1 - H(x)$  を用いて以下のように表現される。

$$Pr[H > (x - y)] = \bar{H}(x - y) \quad (9)$$

利用者の到着時点での待ち時間  $W_a$  が  $y$  と  $y + \Delta y$  の間にある確率は、確率密度関数  $w(y)$  を用いて、

$$w(y)\Delta y = Pr[y < W_a \leq y + \Delta y] = dW(y) \quad (10)$$

となる。従って観測時間  $(0, T)$  において単位時間当たり平均  $\lambda$  の利用者が到着すると、レベル  $x$  を下から上に横切る回数は

$$N_{up}(T) = \lambda T \int_0^x \bar{H}(x - y) dW(y) \quad (11)$$

であり、変形すると以下ようになる。

$$\frac{N_{up}(T)}{T} = \lambda \int_0^x \bar{H}(x - y) dW(y)$$

$$\lim_{T \rightarrow \infty} \frac{N_{up}(T)}{T} = \lambda \int_0^x \bar{H}(x-y) dW(y) \quad (x > 0) \quad (12)$$

得られた  $\lim_{T \rightarrow \infty} \frac{N_{down}(T)}{T}$  及び  $\lim_{T \rightarrow \infty} \frac{N_{up}(T)}{T}$  を式 (6) に代入する。この時、処理装置が処理を行うことによるシステムの仕事量  $W_t$  の変化率は  $\Delta x / \Delta t = 1$  であることから、

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{N_{down}(T) \Delta t}{T} &= w_t(x) \Delta x \\ \lim_{T \rightarrow \infty} \frac{N_{down}(T)}{T} &= w_t(x) \frac{\Delta x}{\Delta t} \\ \lim_{T \rightarrow \infty} \frac{N_{down}(T)}{T} &= w_t(x) \end{aligned} \quad (13)$$

と置換できる。ゆえに以下のような関係式を得る。

$$w_t(x) = \lambda \int_0^x \bar{H}(x-y) dW(y) \quad (x > 0) \quad (14)$$

式 (14) を解くために、確率分布  $W_t(x)$  の LST  $W_t^*(s)$  を定義する。

$$\begin{aligned} W_t^*(s) &\equiv \int_{0-}^{\infty} e^{-sx} dW(x) \quad (x \geq 0) \\ W_t^*(s) - W_t(0) &= \int_{0+}^{\infty} e^{-sx} dW(x) \quad (x > 0) \end{aligned} \quad (15)$$

この  $W_t(0)$  は待ち時間がゼロ、つまり利用者の到着時にはシステムが待機状態にあり、待ち合わせを受けることなく直ちに処理が開始される確率である。式 (15) に式 (14) を代入し、展開する (附録 A 参照)。

$$\begin{aligned} W_t^*(s) - W_t(0) &= \int_{0+}^{\infty} e^{-sx} dW(x) \\ &= \int_{0+}^{\infty} e^{-sx} \left[ \lambda \int_0^x \bar{H}(x-y) dW(y) \right] dx \\ &= \lambda \cdot \mathcal{L}[\bar{H}(x)] \cdot W_t^*(s) \\ &= \lambda \cdot \left( \frac{1 - H^*(s)}{s} \right) \cdot W_t^*(s) \end{aligned} \quad (16)$$

整理して式 (17) を得る。

$$W_t^*(s) - \lambda \cdot \left( \frac{1 - H^*(s)}{s} \right) \cdot W_t^*(s) = W_t(0) \quad (17)$$

式 (17) において  $s \rightarrow 0$  の極限を取り、

$$W_t^*(s)|_{s=0} - \lambda \cdot \left( \frac{1 - H^*(s)}{s} \Big|_{s=0} \right) \cdot W_t^*(s)|_{s=0} = W_t(0), \quad (18)$$

$W^*(s)|_{s=0} = 1, W_t^*(s)|_{s=0} = 1$  から得られる  $W_t(0)$

$$W_t(0) = 1 - \lambda E(H) = 1 - \rho \quad (19)$$

を式 (16) に代入して、GI/G/1 待ち行列システムの待ち時間の表式を得る。

$$W_t^*(s) - \lambda \cdot \left( \frac{1 - H^*(s)}{s} \right) \cdot W_t^*(s) = 1 - \rho \quad (20)$$

式 (20) を変形して式 (21) を得る。

$$W_t^*(s) = (1 - \rho) + \rho \left( \frac{1 - H^*(s)}{s E(H)} \right) \cdot W_t^*(s) \quad (21)$$

式 (21) の右辺第一項はシステム内利用者数ゼロに伴う処理装置の待機状態、第二項は利用者が到着した時点のシステム内の残余仕事量が現時点においてどの程度減少しているかを示している。図 5 に示すように、式 (21) の左辺は任意時刻でのシステム内残余仕事量  $W_t^*(s)$  になる。

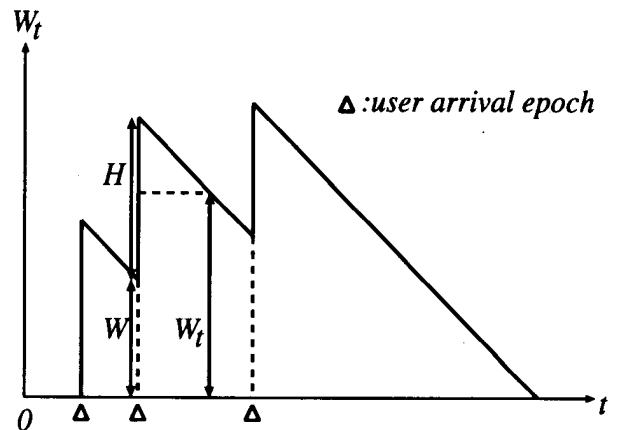


図 5: virtual waiting time and waiting time

これまで想定してきた、利用者の到着過程が一般分布である GI/G/1 システムにおいては、任意時点における待ち時間分布  $W_t(x)$ 、利用者の到着時における待ち時間分布  $W(x)$ 、利用者がシステムを退去する時点での待ち時間分布は必ずしも一致するものではない。しかし、利用者の到着過程がポアソン過程である M/G/1 システムにおいては、PASTA(Poisson Arrivals See Time Averages) の性質により、これら三者が一致することが保証されている。この性質を利用するために、以後想定するシステムを M/G/1 とする。  $W_t^*(s) = W^*(s)$  を式 (21) に代入して、式 (23) の待ち時間の分布関数の LST  $W^*(s)$  を得る。これは式 (5) と

同じである。

$$W^*(s) = \frac{s(1-\rho)}{s-\lambda+\lambda H^*(s)} \quad (22)$$

さらに待ち時間の確率密度関数に関する以下の式を得る。

$$w(x) = \lambda \int_0^x \bar{H}(x-y) dW(y) \quad (x > 0) \quad (23)$$

第1章で紹介した母関数を用いる方法では、LSTを $n$ 回微分する数学的操作で待ち時間の $n$ 次積率を得ることができるが、その確率密度関数の形を具体的に得ることはできないが、Level Crossing法では式(23)のように表現ができるのである。

### 3 Level Crossing法の適用

#### 3.1 Level Crossing法の利点

第1章で述べた様に、待ち行列を含むシステムの常套的な解析法として、ある利用者の処理終了直後における系内利用者数に着目する隠れマルコフ連鎖法や、系内利用者数の他に残余処理時間(あるいはサービス経過時間)を状態変数として取り入れる補助変数法が一般的である。しかし、これらの方法は、待ち時間という時間に関する表現を得るために、それに直接関与することのない利用者の待ち数あるいは系内利用者数に関する、物理的に解釈が困難である母関数を用いており、計算を進めれば解を得られるものの直感的な理解が難しいという問題点を持つ。

それに対してLevel Crossing法を導入する場合、連続量であるシステム内残余仕事量(残余処理時間) $W_t$ に関する平衡方程式を得ることで、定常状態における待ち時間分布のLST  $W^*(s)$ を容易に得ることができる。この際に必要な数学的手法は単純な微分及び積分であり、数学的困難な点はほとんど見られない。

Level Crossing法では、直接的に系内利用者数に関する情報を得ることは出来ないが、ほぼ最短の手間で待ち時間に関する量を得ることができるということは特筆すべきことである。また、式(23)を解析的あるいは数値的に計算することで、待ち時間の確率密度関数を得ることができ、平均待ち時間のみならず待ち時間の揺らぎを直接評価できる。実時間性を重要視するシステムにおいては、待ち時間の揺らぎを定量的に評価することで、システムの性能を数値的に表現することが可能となる。

#### 3.2 Level Crossing法適用時の注意事項

Level Crossing法の骨子は、任意時点での残余仕事量 $W_t$ が、利用者の到着時点でその利用者のもたらす仕事量つまり処理時間だけ増加し、残余仕事量がゼロでない限り処理装置により処理された仕事量だけ減少するサンプルパスにおける平衡方程式である。仕事量つまり処理時間という時間量で評価するということは、利用者数という離散量でしか測ることが出来なかった場合と比べると、精密に評価できるという利点があることになる。第2章で述べたM/G/1待ち行列モデル以外のシステムにLevel Crossing法を適用する時には、以下の注意が必要である。

1. システム内の仕事量の増加をもたらすのは「利用者」の到着時点のみであること、その仕事量は「利用者」の到着時点でのみ評価されることは重要である。M/G/1待ち行列モデルにおいて、システム内利用者数がゼロの場合、処理装置が別の処理(「休暇」と称する)を行う、M/G/1休暇付き待ち行列モデルを例に挙げる。処理装置が休暇を取ることによる仕事量の増加はあたかも「利用者」が到着したかのように見なすことで、Level Crossing法を適用することが可能である。ただし、この時の仮想利用者の平均到着間隔は何らかの手段で与える必要がある。
2. 仕事量の増加分が「利用者」の到着時点でのみ評価されるということは、利用者のシステム到着後にシステム内状態に依存して処理時間が変化するようなシステムは、そのままの形ではLevel Crossing法で取り扱うことができないということである。例えば、利用者の処理終了時点で、確率的に次の利用者の処理状況を決めるベルヌーイ式の場合は、利用者の処理終了時点の残余仕事量の変化が、利用者の到着時点で発生するようにサンプルパスを修正することで対応することができる。またトラヒック規制手段の一つとして、システム到着からある規定時間を越えた利用者をシステムから離脱させる待ち時間制御方式を取入れたシステムや到着順(FIFO)でない処理規律のシステムも、そのままではLevel Crossing法で扱うことができないケースになる。その際には、Level Crossing法で扱うことの出来るよう、サンプルパスの修正及び等価モデルへ落す作業が必要となる。

今後、様々な待ち行列システムにおいてLevel Crossing法の適用を継続的に検討し、他に配慮の必要な事項があるかを調査を進める。

### 参考文献

- [1] 藤木 正也, 雁部 穎一 (1980): 通信トラヒック理論, 丸善.
- [2] H. Takagi(1991): *Queueing Analysis, Vol.1 Vacation and Priority Systems, Part 1*, North-Holland.
- [3] B. T. Doshi (1992): *Level-crossing analysis of queues*, *Queueing and Related Models*, edited by U. Narayan Bhat and I. V. Basawa, 3-33.
- [4] T. Katayama (1999): A note on sojourn time analysis of a two-stage queueing system, *Stochastic Models*, **15**, 298-313.

この時  $H(x)$  が処理時間の分布関数であり,  $H(0) = 0$  及び

$$\lim_{X \rightarrow \infty} \left[ e^{-sX} H(X) \right] = 0 \text{ が成り立つことから,}$$

$$\mathcal{L}[H(x)] = \frac{H^*(s)}{s}. \tag{A.4}$$

式 (A.4) を用いて

$$\begin{aligned} \mathcal{L}[\overline{H}(x)] &= \mathcal{L}[1 - H(x)] \\ &= \mathcal{L}[1] - \mathcal{L}[H(x)] \\ &= \frac{1 - H^*(s)}{s} \end{aligned} \tag{A.5}$$

を得る.

### 附録 A $\mathcal{L}[\overline{H}(x)]$ の導出

確率分布  $H(x)$  が微分可能であれば, 確率密度  $h(x)$

$$h(x) = \frac{dH(x)}{dx} \tag{A.1}$$

が得られる. この時, 確率分布  $H(x)$  の LST  $H^*(s)$  は

$$\begin{aligned} H^*(s) &\equiv \int_0^\infty e^{-sx} dH(x) \\ &= \int_0^\infty e^{-sx} h(x) dx \\ &= \mathcal{L}[h(x)] \end{aligned} \tag{A.2}$$

と表現される. ラプラス変換の定義から

$$\begin{aligned} &\mathcal{L}[H(x)] \\ &= \int_0^\infty e^{-sx} H(x) dx \\ &= \lim_{X \rightarrow \infty} \int_0^X e^{-sx} H(x) dx \\ &= \lim_{X \rightarrow \infty} \left[ \frac{1}{-s} e^{-sx} H(x) \right]_0^X \\ &\quad - \lim_{X \rightarrow \infty} \int_0^X \frac{1}{-s} e^{-sx} h(x) dx \\ &= \frac{1}{s} H(0) - \frac{1}{s} \lim_{X \rightarrow \infty} \left[ e^{-sX} H(X) \right] + \frac{1}{s} \lim_{X \rightarrow \infty} \int_0^X e^{-sx} h(x) dx \\ &= \frac{1}{s} \left[ H(0) - \lim_{X \rightarrow \infty} \left[ e^{-sX} H(X) \right] \right] + \frac{1}{s} \int_0^\infty e^{-sx} h(x) dx \\ &= \frac{1}{s} \left[ H(0) - \lim_{X \rightarrow \infty} \left[ e^{-sX} H(X) \right] \right] + \frac{1}{s} H^*(s). \end{aligned} \tag{A.3}$$

# An Analysis of $GI/G/1$ Queueing Model by the Level Crossing Method

Kaori KOBAYASHI and Tsuyoshi KATAYAMA

Department of Electronics and Informatics, Faculty of Engineering

## Summary

In everyday life, a situation is frequently happened that some users commonly share some limited resources and wait while the resources are busy; e.g. automated teller machines in a bank. As one of these queueing systems, a  $GI/G/1$  queueing model (denoted by Kendall's notation;  $GI$  specifies the arrival process is a renewal arrival process,  $G$  specifies the service time distribution is general distribution, and 1 denotes a single server) is a basic queueing system, and has a fundamental and an important role in evaluation of quality of service (QoS) of queueing systems. To analyze the  $M/G/1$  (a special case of the  $GI/G/1$  queueing model), the method of the imbedded Markov chain is commonly used for finding the distribution of the number of users (or clients, customers, messages, e.t.c.) in the system, though an obtained generating function of the distribution of the number of users is expressed in mathematical form and is not easily understandable.

In this paper, taking the Level Crossing method, we have obtained the probability density function of waiting time in the  $GI/G/1$  queueing model more intuitively than using the generating function of the distribution of the number of users, and discuss this method's validity.

**keywords:** the level-crossing method,  $GI/G/1$  queueing model, virtual waiting time, distribution function of the waiting time