システム工学基礎
System Engineering
待ち行列理論

電子情報工学科
伊庭　斉志

待ちの性質

- 客の到着間隔
  - まちまち
  - 混むときはどっときて、しばらく誰も来ない
- サービス時間も変動する

- 待ちの発生原因
- ⇒　到着間隔やサービス時間が客ごとに大幅に変動する

定義

- 窓口（扱者）
- サービス
- 客　または　呼(call)
- 到着分布
- サービス時間分布

指数分布

- 到着間隔（x）の分布

$$f(x) = \lambda e^{-\lambda x}$$

$$(x > 0)$$

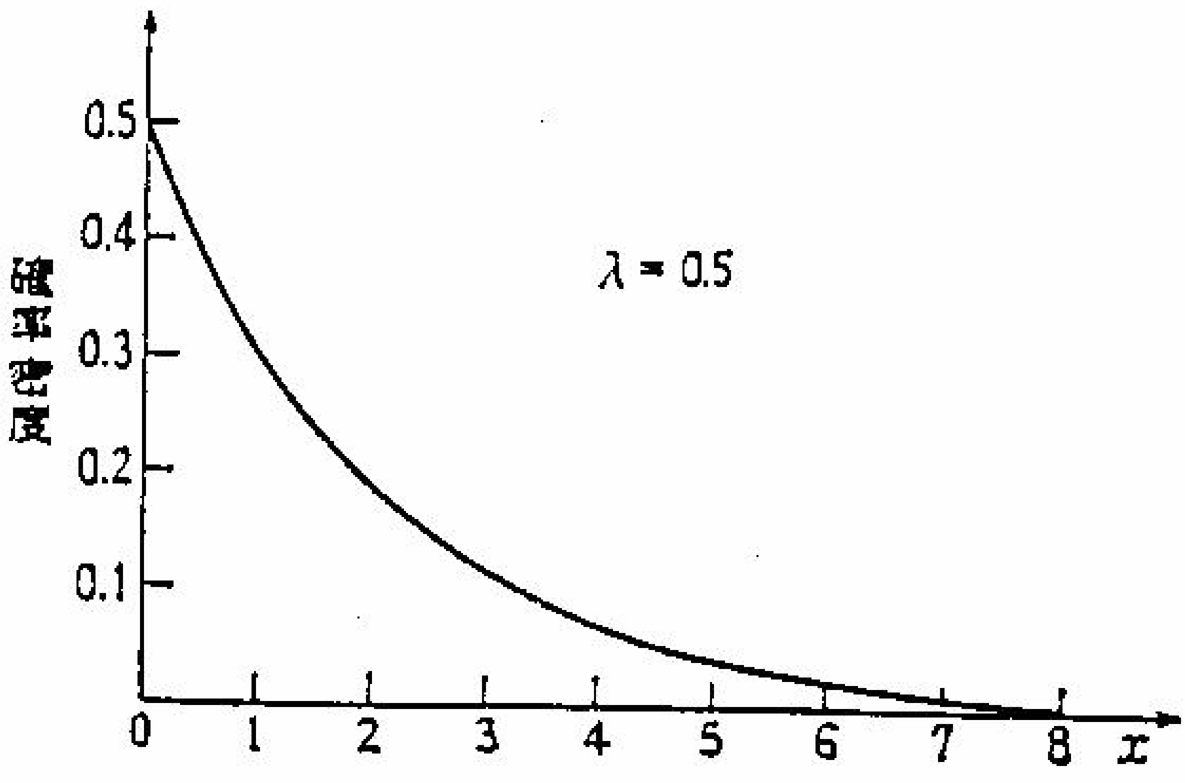


図 8.4　指数分布

•平均：1/λ

•標準偏差：1/λ

指数分布の特徴

- 客が来て次に来るまでの時間の分布
- ０のところがおかしい
- 来るときは来て、来ないときは来ないという状況にうまく合っている

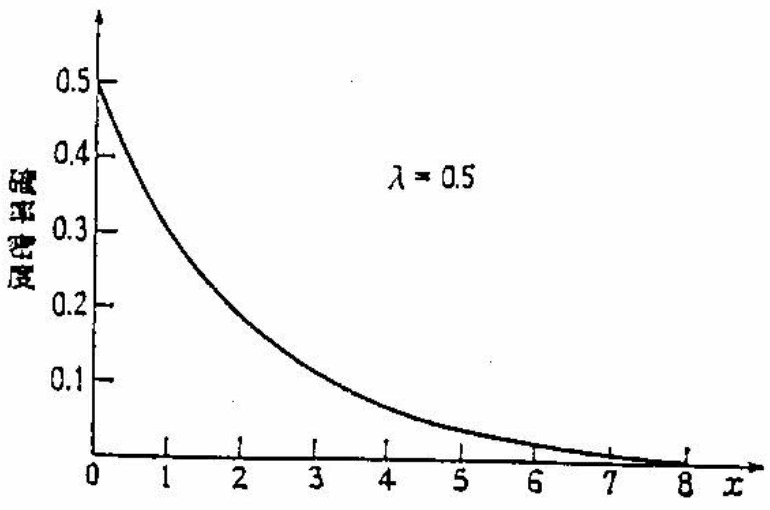


図 8.4　指数分布

ポアソン分布

$$P(N=k)=\frac{\theta^k}{k!}e^{-\theta} \qquad (k=0,1,2,\ldots)$$

$$\begin{cases} \text{平均} & :\theta \\ \text{分散} & :\theta \end{cases}$$

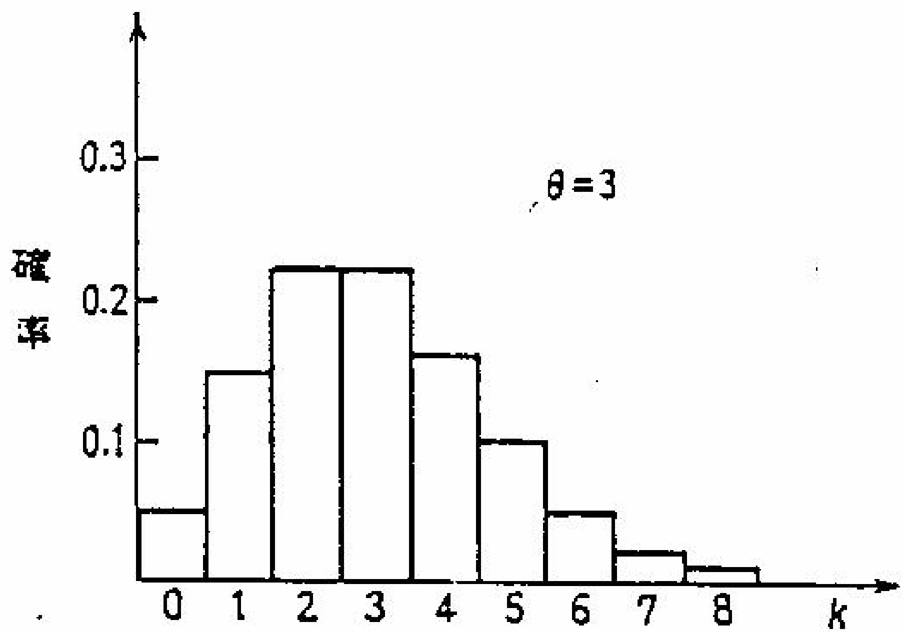
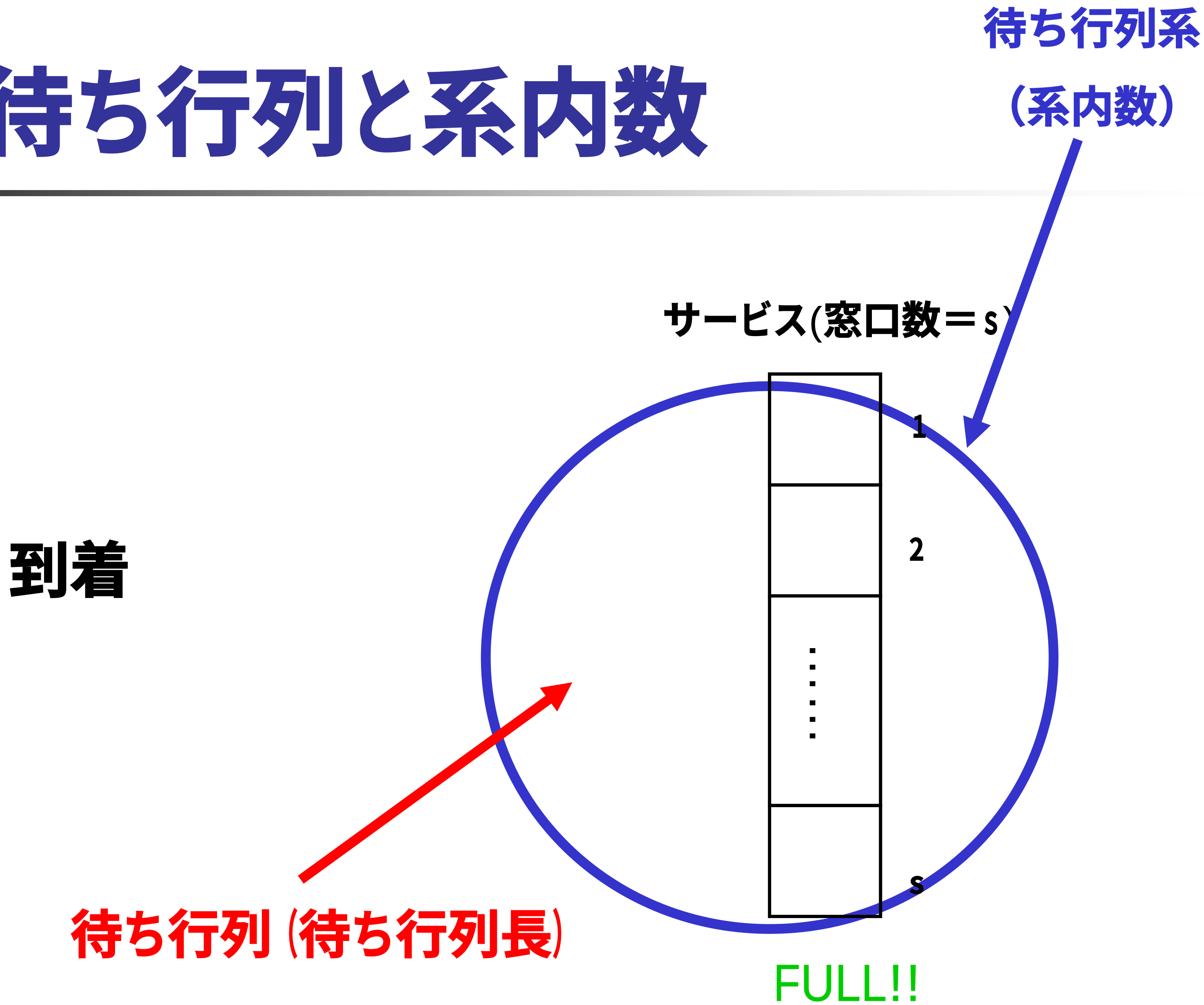


図 8.5　ポアソン分布

到着人数*k*の分布

待ち行列と系内数



待ち行列と系内数

- （狭い意味で）待ち行列＝サービスを受ける前に待たされている客

- 窓口＋待ち行列＝待ち行列系
- サービス中の客の数＋待ち行列中の客の数＝系内数
- 系内数＞s（窓口数）なら
  - 系内数ーs＝待ち行列長（待ち行列の大きさ）
- 系内数≦s（窓口数）なら
  - 待ち行列長は0

もっともでたらめな客

- **独立性**
  - ある時間内に到着する人数はそれ以前に到着した人数によらない
- **定常性**
  - ある時間内にｋ人到着する確率はその時間がいつはじまるかによらない
- **二値性**
  - 微小時間内に２人来ることはない（確率は０）

サイコロの確率

例．さいころn回投げて、１の目がk回出る確率

$$ {}_nC_k\left(\frac{1}{6}\right)^k\left(1-\frac{1}{6}\right)^{n-k} $$

ベルヌーイ試行

客の到着確率

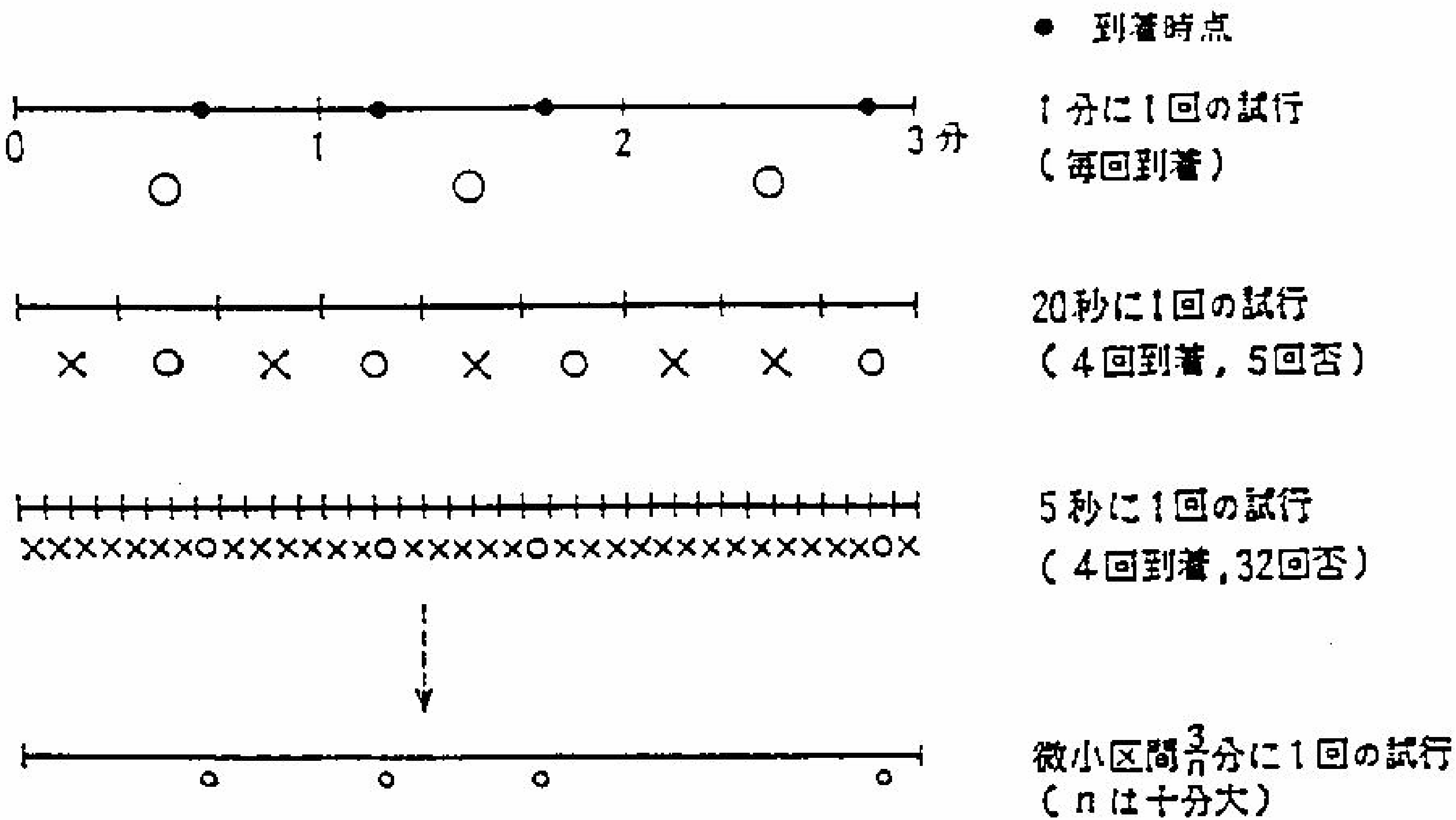

● 到着時点
0
1
2
3分
1分に1回の試行
(毎回到着)
20秒に1回の試行
(4回到着，5回否)
5秒に1回の試行
(4回到着，32回否)
微小区間$\frac{3}{n}$分に1回の試行
(nは十分大)

図 8.8

客の到着確率

ベルヌーイ試行で考える

平均到着率λ　→　単位時間内に到着する人数

（単位：人数／時間）

$\frac{t}{n} = \Delta t$に客が来る確率は$\lambda \Delta t$

全時間tをn等分する

p, qを以下のようにする。

$p = \lambda \Delta t$

$q = 1 - p = 1 - \lambda \Delta t$

$(p + q = 1)$

全時間t内にk人が来る確率

$\boxed{{}_nC_k p^k q^{n-k}}$

ポアソン分布

$$P(N=k)=\frac{(\lambda t)^k}{k!}\times\left(1-\frac{1}{n}\right)\times\left(1-\frac{2}{n}\right)\times\cdots\times\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda t}{n}\right)^n\left(1-\frac{\lambda t}{n}\right)^{-k}$$

$n\to\infty$のとき　　の部分は１に近づく！！

また$\left(1-\dfrac{\lambda t}{n}\right)^n = \left[(1+u)^{\frac{1}{u}}\right]^{-\lambda t} \to e^{-\lambda t}$

$u=-\lambda t/n$　　　$n\to\infty, u\to 0$

ポアソン分布

よって、

$$P(N=k)=\frac{(\lambda t)^k}{k!}e^{-\lambda t} \qquad (n\to\infty)$$

平均が λ *t* のポアソン分布

時間間隔が *t* の間に *k* 人が到着する確率

ポアソン分布と指数分布

$$p(N=k)=\frac{(\lambda t)^k}{k!}e^{-\lambda t}$$:t時間内にk人が来る確率

$$f(t)=\lambda e^{-\lambda t}$$ 平均$\frac{1}{\lambda}$の指数分布

$\frac{1}{\lambda}$: 平均到着間隔（時間／人）

λ：平均到着率（人／時間）

ケンドールの記法

Ａ／Ｂ／ｓ（ｎ）

Ａ：到着分布

Ｂ：サービス分布

ｓ：窓口数

ｎ：許される行列長　　∞もある：そのときは省略

ケンドールの記法

例： M／M／s

M : 指数分布時間到着：平均 $\frac{1}{\lambda}$(時間／人）　　$f(x) = \lambda e^{-\lambda x}$

$\updownarrow$

ポアソン到着：平均　λ(人／時間）

人数　　　λ：到着率

M：サービス時間分布　　平均サービス時間 : $\frac{1}{\mu}$(時間／人）

ケンドールの記法

例：　M／D／s

D：一定時間　　（最も規則的）

$\updownarrow$

M：指数分布　　（最もランダム）

M／M型待ち行列の利点

- “待ち”を多めに見積もる
- “安全側”の答を与える
- 解きやすい
- ポアソン到着は実際的
- ３つの仮定に注意
  - 朝昼夜では電話を状況はかなり異なるため、定常性は成り立たない

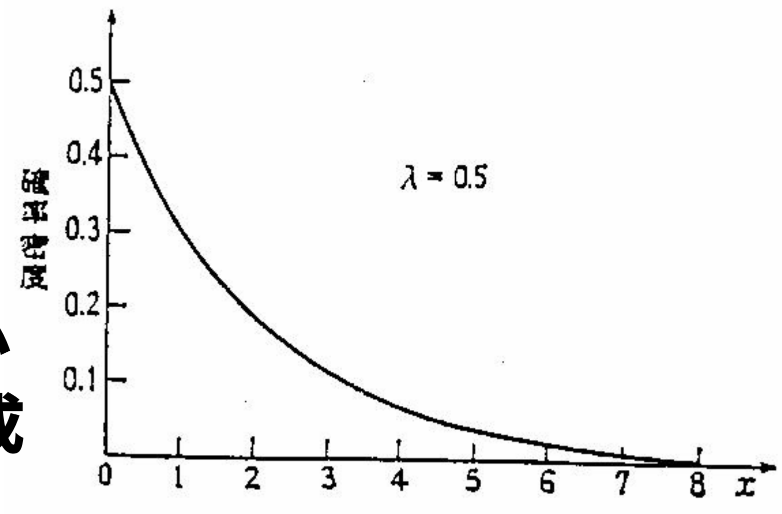


図 8.4　指数分布

アーラン分布

$$f(x) = \frac{(\lambda k)^k x^{k-1}}{(k-1)!} e^{-\lambda k x} \qquad (x > 0)$$

平均 $:\dfrac{1}{\lambda}$ 　　分散 $\dfrac{1}{k\lambda^2}$

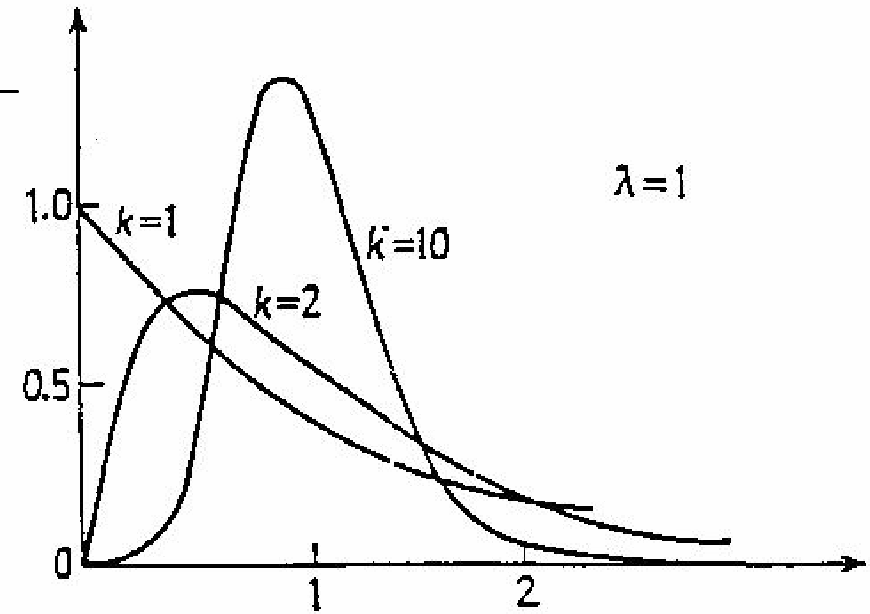


図 8.6　アーラン分布

$k = 1$ のとき指数分布

$k \to \infty$ のときD（一定分布）

例：$M/E_k/1$　（∞）

アーラン分布の特徴

E_k　次数 k のアーラン分布

- 指数分布と異なり、山を持つ分布
- サービス時間がゼロに近いところの確率に配慮
  - 指数分布→大きすぎる

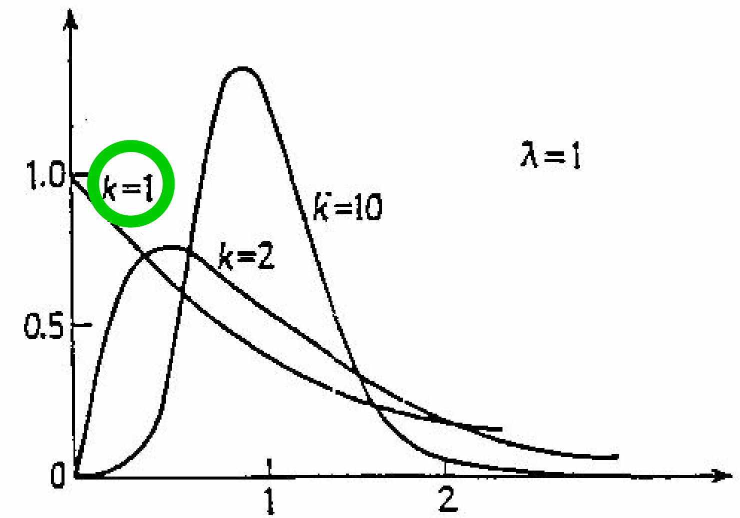
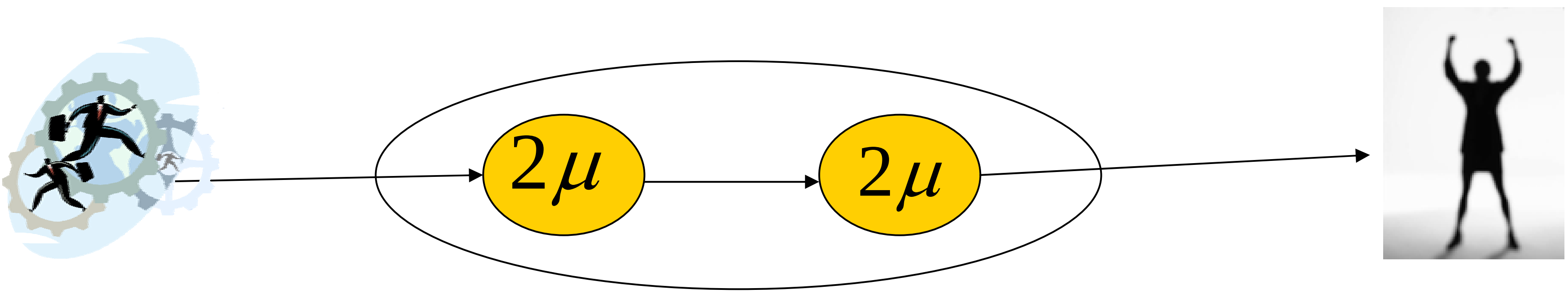


図 8.6　アーラン分布

アーラン分布の例

次数 2のアーラン分布　E_2



サービス施設

変動係数

$$変動係数 = \frac{標準偏差}{平均}$$

例１：指数分布（M）→1　：最もばらつきが大きい

例２：$D \to 0$:ばらつき無し

例３：　$E_k \to \frac{1}{\sqrt{k}}$: これらの中間となる

M／M／s（∞）

- 到着間隔：指数分布
  - ポアソン到着
- サービス時間：指数分布型
- 窓口数s
- 許される行列長：制限無し

M／M／s (∞)

λ：到着率（人／時間）

$\frac{1}{\lambda}$: 平均到着時間

$f(x) = \lambda e^{-\lambda x}$

μ：サービス率（人／時間）

$\frac{1}{\mu}$: 平均サービス時間（時間／人）

$f(y) = \mu e^{-\mu y}$

ポアソン到着と指数分布

M：最もランダム ⇔ D：最も規則的（一定時間）

ポアソン到着 ⇔ レギュラー到着

$$P(N=k)=\frac{(\lambda t)^k}{k}e^{-\lambda t}$$: t時間内に到着する人数がk人の確率

$$P(N=0)=e^{-\lambda t}$$: t時間内に一人も来ない確率 ⇐ kに0を代入

$$\therefore P(X>t)=e^{-\lambda t}$$

到着間隔Xの分布は指数分布になる！！

ポアソン到着とでたらめさ

微小時間h 以内に客が一人到着する確率
$= \lambda h + o(h) \quad \lambda h$

$$P(t < X < t+h) = e^{-\lambda t} - e^{-\lambda(t+h)} \quad ①$$

$$P(X > t) = e^{-\lambda t} \quad ②$$

$$P(t < X < t+h \quad | \quad X > t) : \text{条件つき確率}$$

$$= ① \div ②$$

$$= 1 - e^{-\lambda h}$$

$$= \lambda h + 0(h)$$

系内数のマルコフ連鎖

$Q(t) =$ 時刻tでの系内数

独立性条件＋hが微少量 $\Rightarrow$ 来るのも出るのもせいぜい一人

$Q(t+h)-Q(t)$　は、+1,0,–1のいずれかの値をとる

$\therefore Q(t)$　は$\boxed{\text{マルコフ連鎖}}$　　→**Mの意味！！**

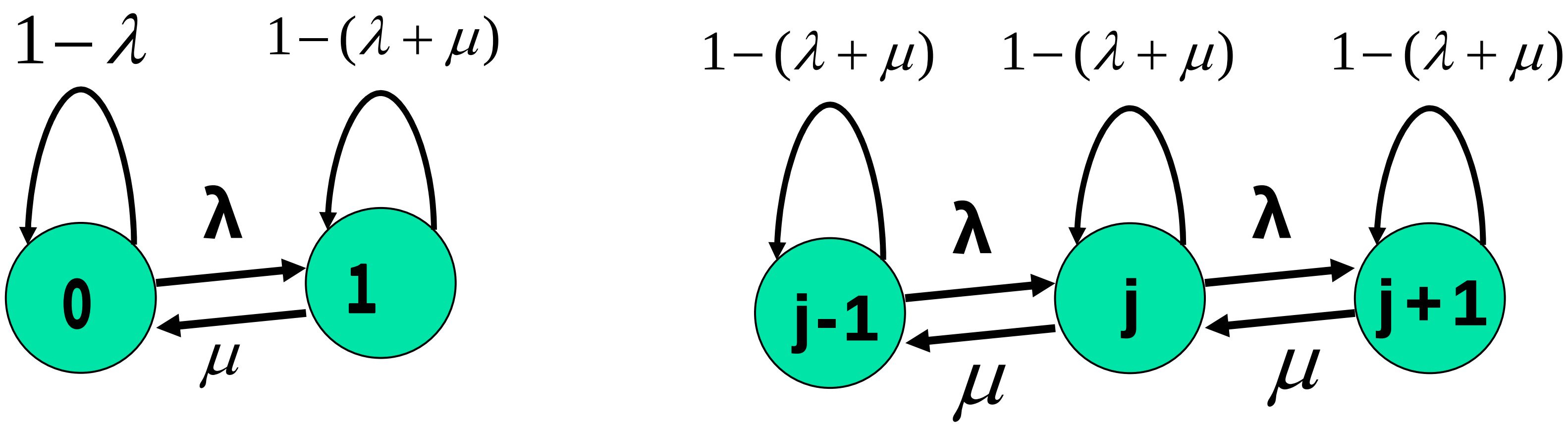
系内数のマルコフ連鎖

長さ　h　の時間区間内に１人到着する確率は
$\boxed{\lambda h + o(h)}$ である。
同様に、サービス中の客が長さ　h　の時間区間内にサービスを終了する確率は $\boxed{\mu h + o(h)}$ となる。

これらは互いに独立であり、到着とサービス終了の両方が起こる確率は $\boxed{\lambda\mu h^2 + o(h^2)}$ となり、hが小さければ無視できるほど小さい。

系内数のマルコフ連鎖

$$\left.\begin{aligned}
&P[Q(t+h)=j+1 \mid Q(t)=j]=\lambda h+o(h)\\
&P[Q(t+h)=j \mid Q(t)=j]=1-(\lambda+\mu)h+o(h)\\
&P[Q(t+h)=j-1 \mid Q(t)=j]=\mu h+o(h)
\end{aligned}\right\}$$



遷移図

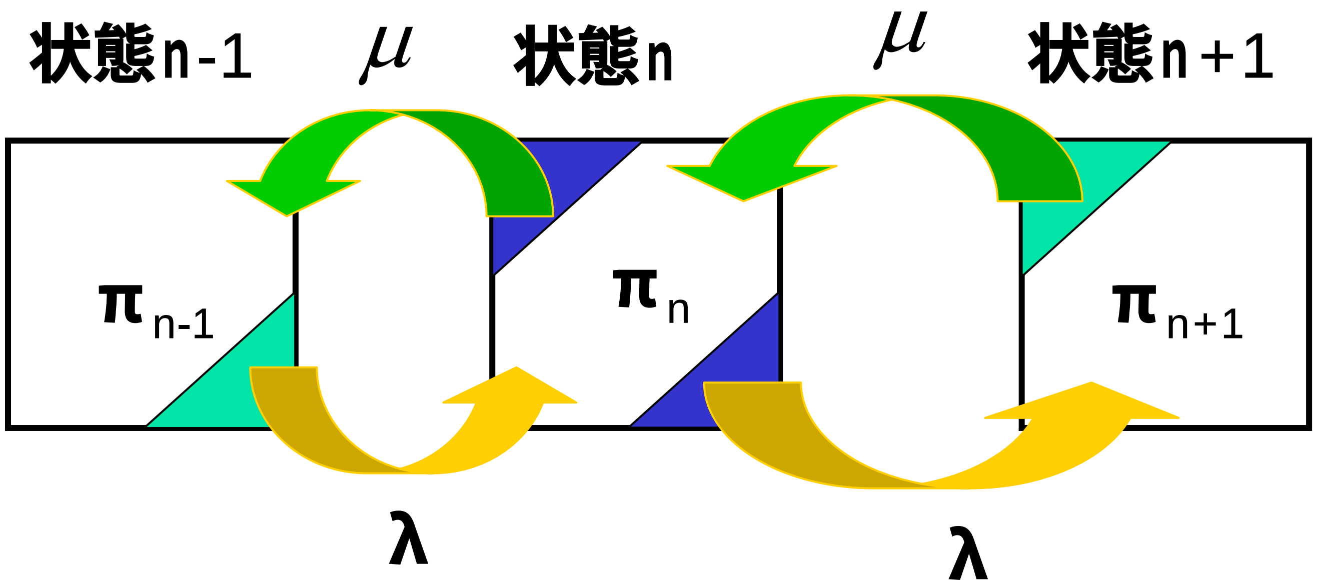
定常分布

$$\left.\begin{aligned} &\lambda u_0 = \mu u_1 \\ &(\lambda + \mu) u_j = \lambda u_{j-1} + \mu u_{j+1} \quad (j \geq 1) \end{aligned}\right\} \text{が成り立つ}$$

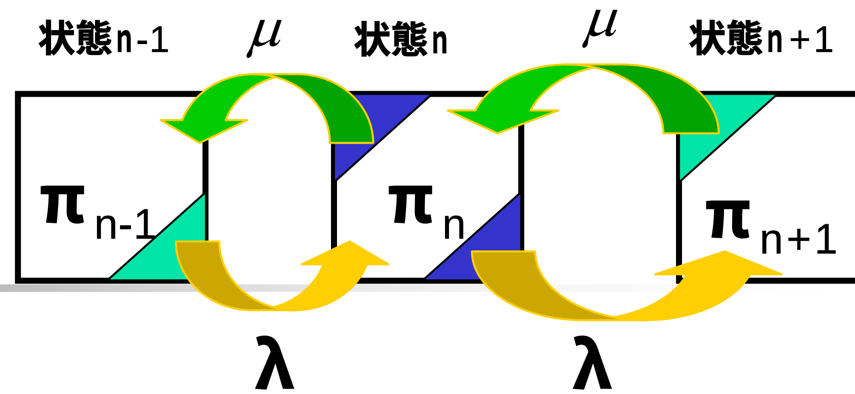
これを解いて平衡状態の分布を求める

平衡方程式の導出

- 統計的平衡による平衡方程式の導出
- 状態間の推移があっても状態確率は変わらない



統計的平衡



各状態（系内数）　$n-1,\quad n,\quad n+1$　がある確率を
それぞれ　$\pi_{n-1}, \pi_n, \pi_{n+1}$　とする。

陰の部分　　$(\lambda+\mu)\pi_n$
横線の部分　$\lambda\pi_{n-1}+\mu\pi_{n+1}$　　が等しければ、
推移によって各状態にいる待ち行列の数には変化がない。

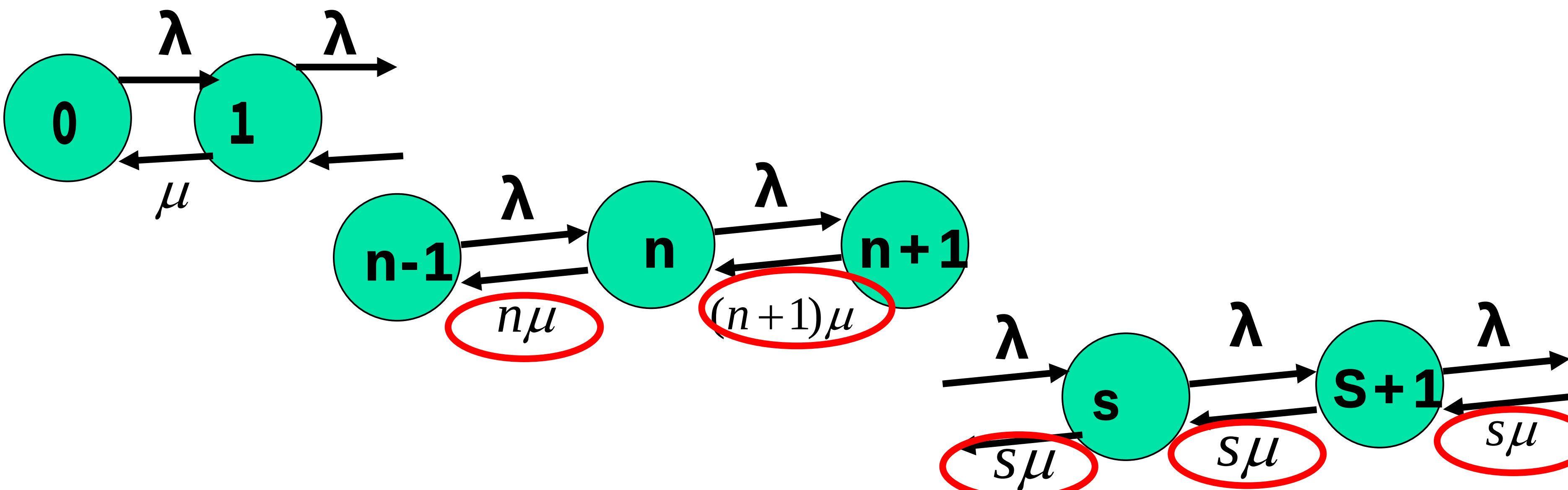
よって、

$$\boxed{\begin{aligned}&(\lambda+\mu)\pi_n = \lambda\pi_{n-1}+\mu\pi_{n+1} \quad (n \geq 1)\\ &\lambda\pi_0 = \mu\pi_1\end{aligned}}$$

統計的平衡

- 定常分布の式＝平衡方程式
- 定常分布＝平衡分布

【例題1】M/M/s型待ち行列モデルの平衡方程式



M／M／1 (∞) について

π_n: 定常分布状態　（系内にn人いる確率）

平衡方程式

$$\mu\pi_{i+1} + \lambda\pi_{i-1} = \lambda\pi_i + \mu\pi_i \quad (i \geq 1)$$

$$\lambda\pi_0 = \mu\pi_1$$

$$\Sigma\pi_i = 1$$

$0 < \dfrac{\lambda}{\mu} < 1$のとき平衡分布が存在

これを解くと、

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right) \quad (n \geq 0)$$

M／M／1 (∞) について

定義　$\frac{\lambda}{\mu}=\rho$：負荷率$(load \quad factor)$、トラフィック密度

平衡分布が存在するには$\rho<1$でなくてはならない

平衡条件、エルゴード条件

$$\pi_n=\rho^n(1-\rho)$$

$$\pi_0=1-\rho \qquad\qquad 1-\pi_0=\rho$$

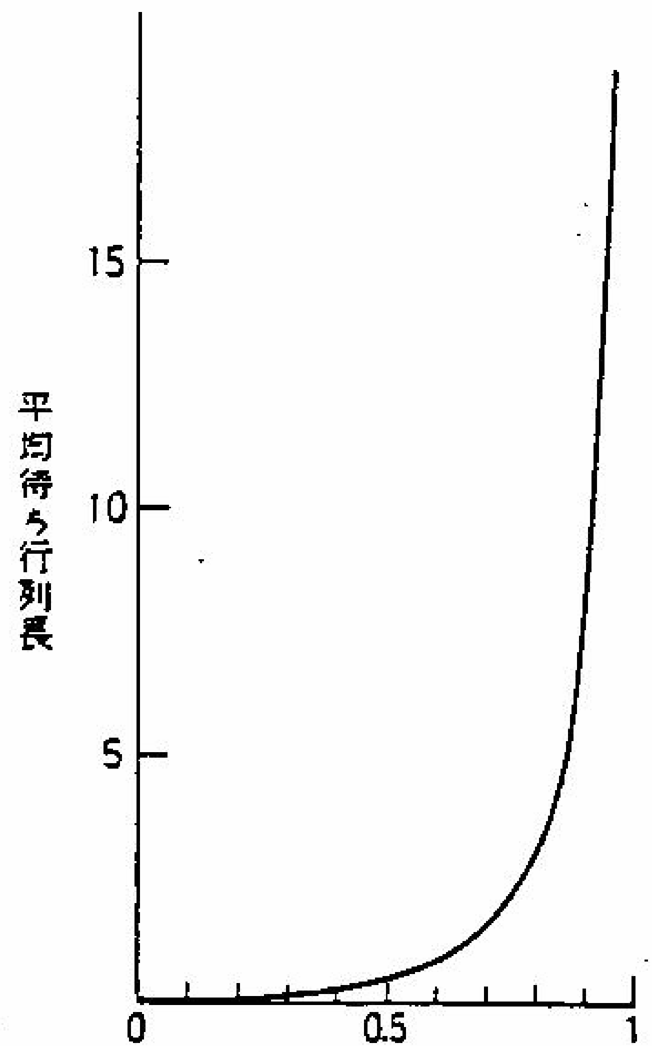
ひま,ゆとり　　　　　窓口がふさがれている：利用率

正でなければ平衡分布は存在しない！！

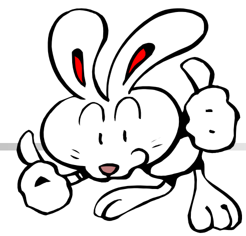
Ｍ／Ｍ／１の平均待ち行列長

平均待ち行列長

$$= \sum_{n=1}^{\infty} (n-1)\pi_n = \frac{\rho^2}{1-\rho}$$

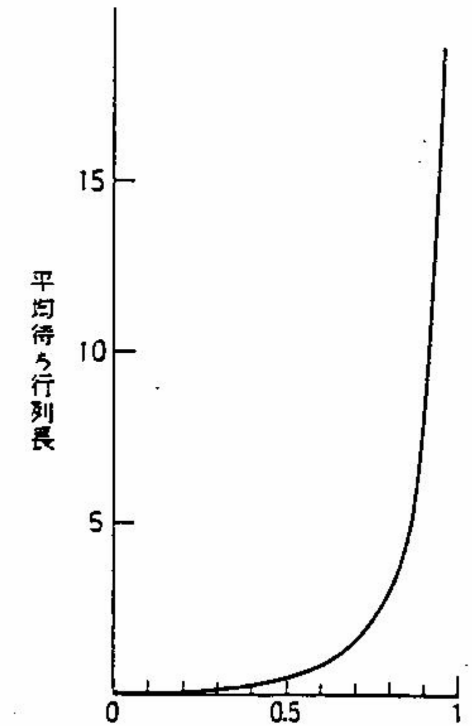


７０%の法則

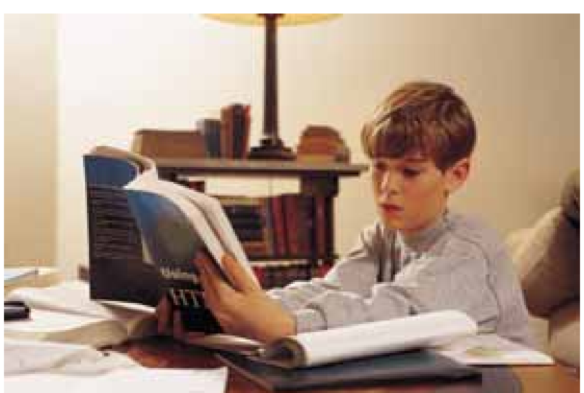


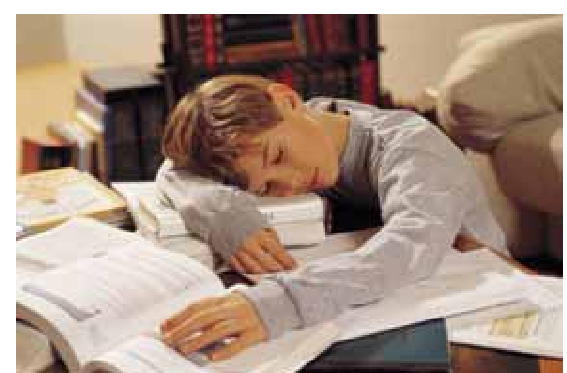
負荷率 $\dfrac{\lambda}{\mu} = \rho$

が７０%を超えたら要注意！！



例：レポート提出





待ち時間分布

$M/M/1$系での客の待ち時間＝γ

$$P(\gamma > t) = \rho e^{-(1-\rho)\mu t}$$

待ち時間分布

客の待ち時間　γ　は先客　k　人のサービス時間の和に等しい。
結局　γ　は　k　人分のサービス時間の和ということになる。

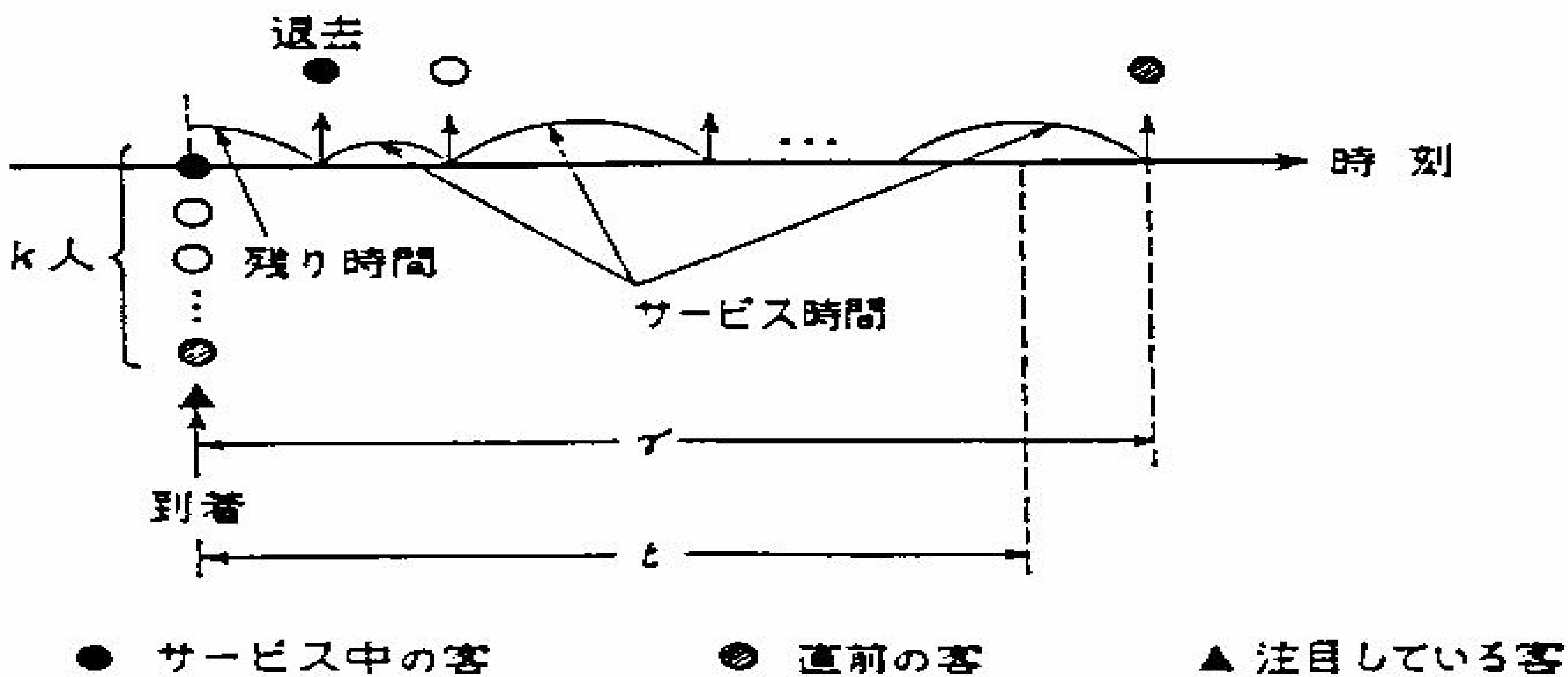


図 8.13

待ち時間分布

$$P(\gamma > t) = \sum_{k=1}^{\infty} \pi_k \sum_{\gamma=0}^{k-1} e^{-\mu t} \frac{(\mu t)^{\gamma}}{\gamma !} = e^{-\mu t}(1-\rho) \sum_{k=1}^{\infty} \rho^k \sum_{\gamma=0}^{k-1} \frac{(\mu t)^{\gamma}}{\gamma !}$$

$$= e^{-\mu t}(1-\rho) \sum_{\gamma=0}^{\infty} \frac{(\mu t)^{\gamma}}{\gamma !} \sum_{k=\gamma+1}^{\infty} \rho^k = e^{-\mu t}(1-\rho) \sum_{\gamma=0}^{\infty} \frac{(\mu t)^{\gamma}}{\gamma !} \cdot \frac{\rho^{\gamma+1}}{1-\rho}$$

$$= e^{-\mu t} \rho \sum_{\gamma=0}^{\infty} \frac{(\rho \mu t)^{\gamma}}{\gamma !} = \rho e^{-\mu t} e^{\rho \mu t} = \rho e^{-(1-\rho)\mu t}$$

M／M／s　の平衡分布

$$a = \frac{\lambda}{\mu}, \qquad \boxed{\rho = \frac{a}{s} < 1}\text{として}$$

平衡条件

$$\left.\begin{aligned}\pi_n &= \frac{a^n}{n!}\pi_0 \quad (0 \le n \le s) \\ &= \frac{a^n}{s!\,s^{n-s}}\pi_0 \,(s \le n)\end{aligned}\right\}$$

$$\pi_0^{-1} = \sum_{n=0}^{s-1}\frac{a^n}{n!} + \frac{a^s}{(s-1)!\,(s-a)}$$

プリントを見てみよう

M／M／s　の平衡分布

待ち確率：窓口がすべて塞がっている確率

$$\Pi = \sum_{n=s}^{\infty} \pi_n = \frac{s^s \rho^s \pi_0}{s!(1-\rho)}$$

Ｍ／Ｍ／ｓ と Ｍ／Ｍ／１

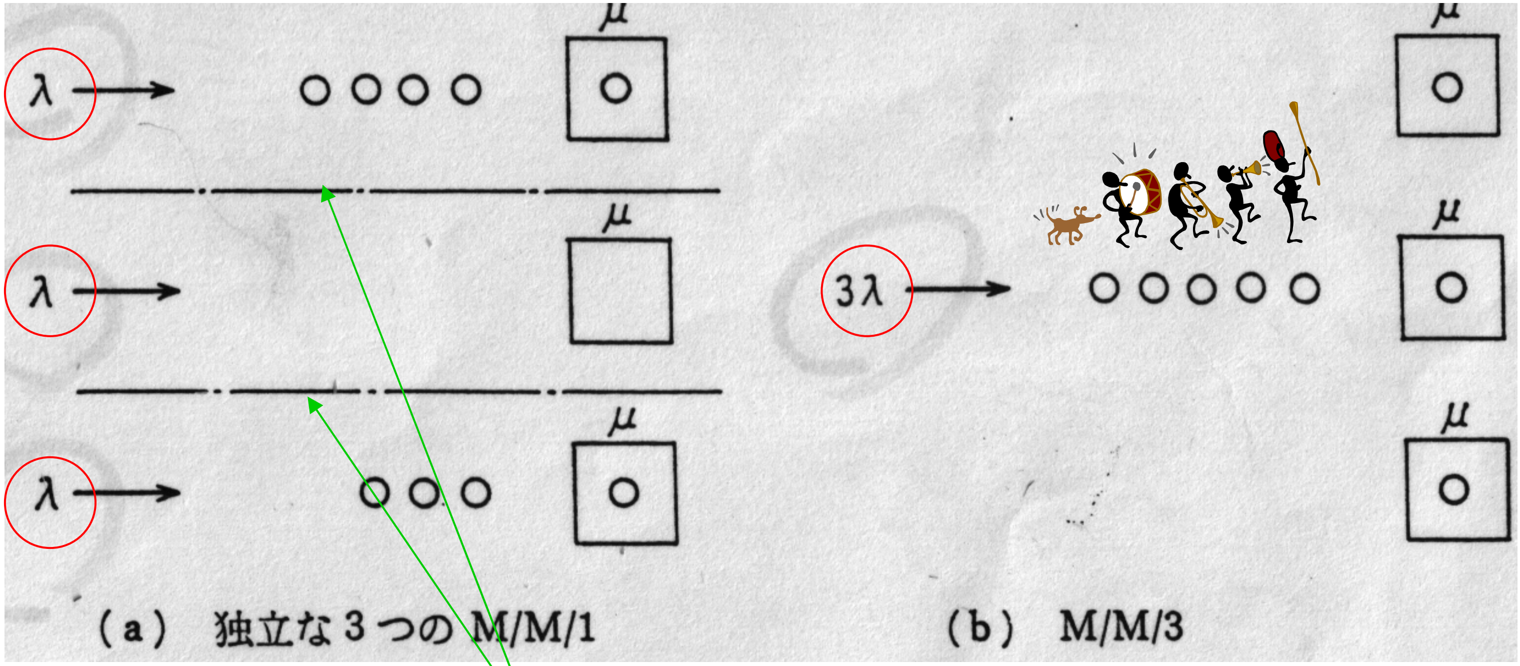
７０％の法則
　ρ＞0.7になると
システムはやがて破綻する

- $s=1$のとき　　$\rho=\dfrac{\lambda}{\mu}$: 負荷率

$\rho>1$なら行列は∞

- $s\geq 1$　　$\dfrac{\rho}{s}=\dfrac{\lambda}{s\mu}\geq 1$ なら行列は∞

3つのM／M／1 と M／M／3



横入り出来ない

複数窓口の効用

$M/M/s$型待ち行列系

Πの値＝待ちの確率（窓口が塞がっている確率）

$s=3$　のとき

$$\Pi = \frac{9\rho^3}{2+4\rho+3\rho^2}$$

$s=1$　のときは

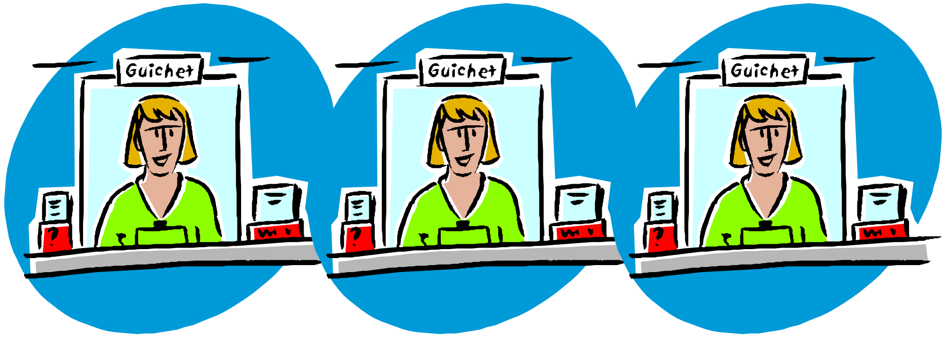
$$\Pi = \rho$$

ただし、　$\rho = \dfrac{\lambda}{s\mu}$

$a = \dfrac{\lambda}{\mu}$　は窓の利用率

複数窓口の効用

- 一般に、どんなρ、sについても、sが大きいほど$1-\Pi$は大きくなる
- 客の数に比例して窓口を増やすほど待ちは減る
- ∴窓口は共通に使うと効率があがる



呼損系： M／M／s （N）

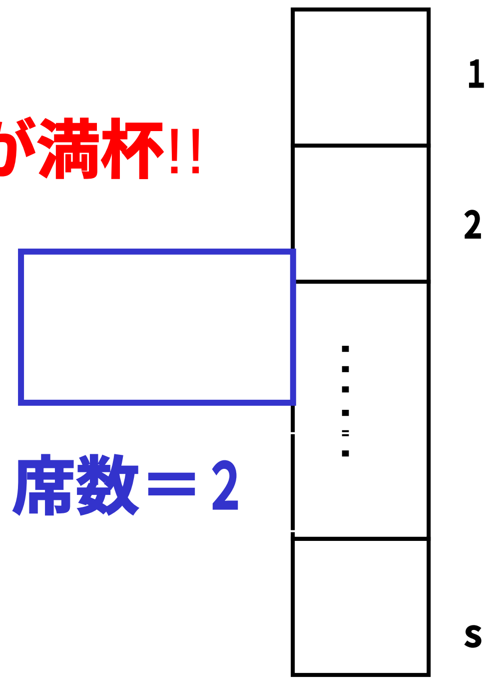
待ち行列が有限なとき

例： M／M／s （s＋2）

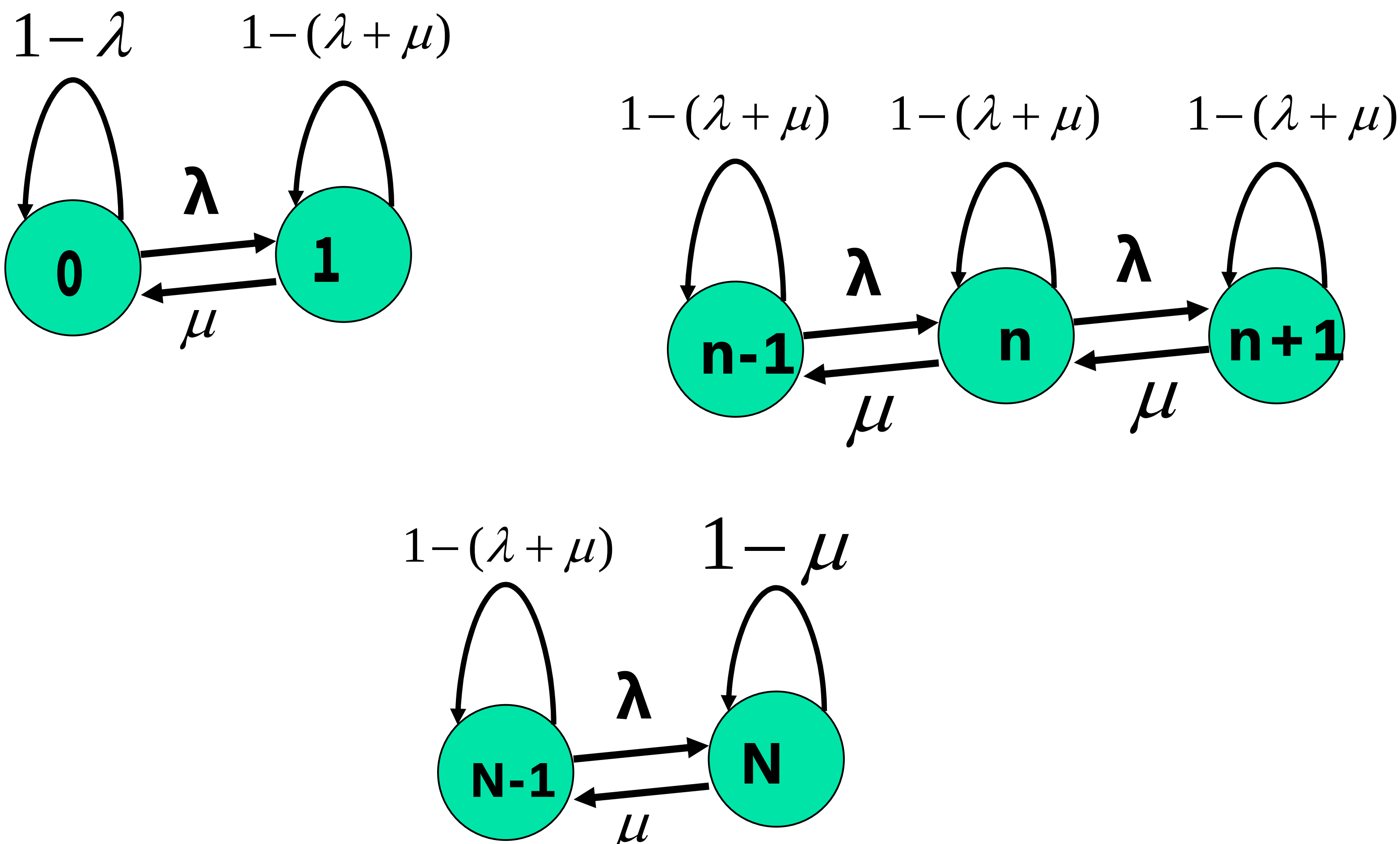
行列が満杯!!

到着

サービス(窓口数＝s)



M／M／1 (N)



$1-\lambda$
$1-(\lambda+\mu)$
λ
0
1
μ
$1-(\lambda+\mu)$
$1-(\lambda+\mu)$
$1-(\lambda+\mu)$
λ
λ
n-1
n
n+1
μ
μ
$1-(\lambda+\mu)$
$1-\mu$
λ
N-1
N
μ

M／M／1 （N）

$$\begin{cases} \lambda P_0 = \mu P_1 \\ (\lambda + \mu) P_n = \mu P_{n+1} + \lambda P_{n-1} \\ \mu P_N = \lambda P_{N-1} \end{cases} \qquad \sum_{i=0}^{N} P_i = 1$$

$$\Rightarrow \quad P_n = \frac{\rho^n (1-\rho)}{1-\rho^{N+1}} \qquad (1 \le n \le N)$$

M／M／s (N)

系の中にn人いる確率

系が空である確率

$$P_n = \frac{\rho^n}{n!} P_0 \quad (1 \le n \le s)$$

$$P_n = \frac{\rho^n}{s!\, s^{n-s}} P_0 \quad (s \le n \le N)$$

M／M／s（N）：呼損率

サービスを受けずに立ち去る確率

$$P_N = \frac{\rho^N}{s!\, s^{N-s}} P_0$$

$N=s$ のとき、 M／M／s（s） ： 即時系

プリントの記号

- $E(n)$: 系（窓口＋行列）内の平均人数　$E(m)=\sum_{n=0}^{\infty} nP_n$
- $E(m)$: 平均行列長　$E(m)=\sum_{n=s}^{\infty}(n-s)P_n$
- $E(v)$: 到着した人が系内で費やす平均時間　$E(u)=\dfrac{1}{\lambda}E(n)$
- $E(w)$: 到着したものの平均待ち時間　$E(w)=\dfrac{1}{\lambda}E(m)$
- $\pi=P(n\geq s)=\sum_{n=s}^{\infty}P_n$: 閉塞率　窓口が全てふさがっている確率

プリントを見てみよう

系内数と待ち行列数

一般に、

$$\underbrace{E(m)}_{\text{平均行列人数}} = \underbrace{E(n)}_{\text{系内平均人数}} - \rho$$

利用率：$\rho = \dfrac{\lambda}{\mu}$

M／G／s

- 到着分布はポアソン分布に近いことは多い
- サービス時間は変動の少ない分布が多い

- M/Mでないと系内数は**マルコフ連鎖**とならない

- M/G/s　は解析が難しい

ポラチェック･ヒンチンの公式

$$<M/G/1のE(w)> = \boxed{\frac{\lambda}{\mu(\mu-\lambda)}} \cdot \frac{1+C^2}{2}$$

$$= \boxed{<M/M/1のE(w)>} \cdot \frac{1+C^2}{2}$$

E(w) ： **平均待ち時間**

C ： **分布Gの変動係数**

ポラチェック･ヒンチンの公式

$$< M/G/1\text{の}E(w) >=< M/M/1\text{の}E(w) > \cdot \frac{1+C^2}{2}$$

$$< M/G/s\text{の}E(w) >=< M/M/s\text{の}E(w) > \cdot \frac{1+C^2}{2}$$

E(w) ： **平均待ち時間**

C ： **変動係数**

リトルの公式

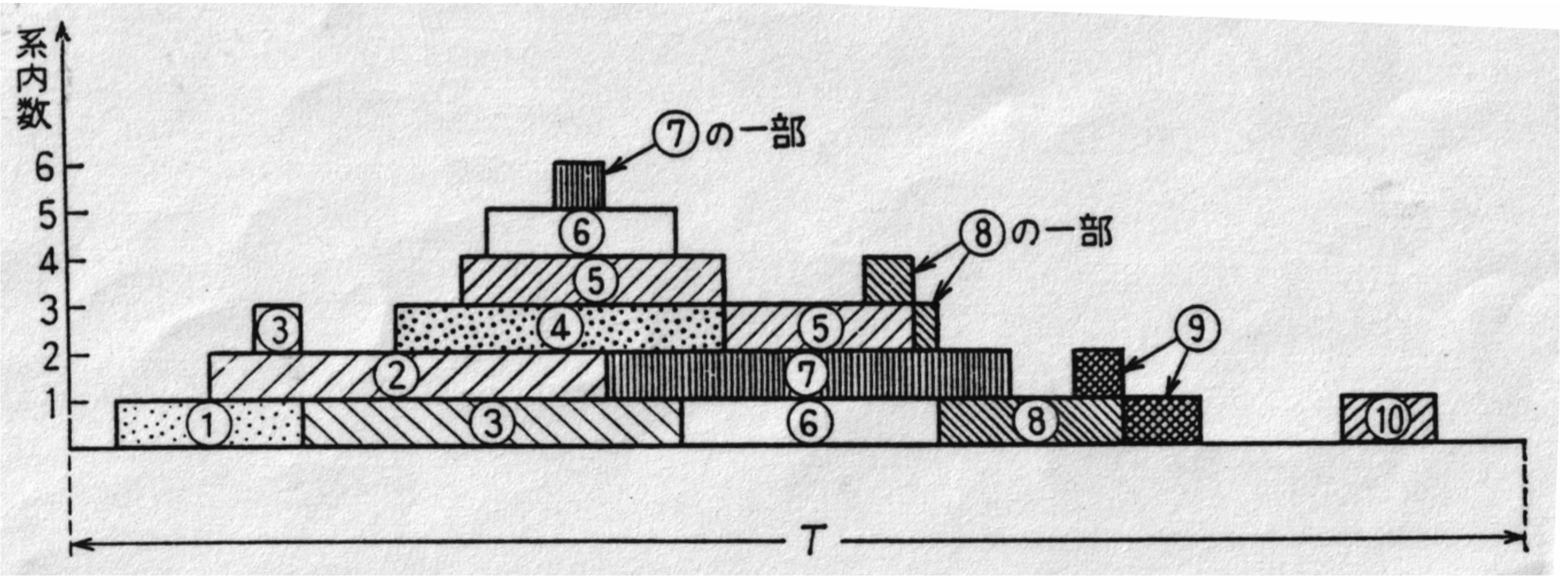
$$\underline{E(\upsilon)} = \frac{1}{\lambda}\underline{E(n)}$$

系内平均時間　　系内平均人数

$$\underline{E(w)} = \frac{1}{\lambda}\underline{E(m)}$$

平均待ち時間　　平均行列人数

リトルの公式



同じ番号の長方形の和＝その番号の客の系内時間

高さ＝1, 横幅＝系内時間

∴面積＝系内時間

リトルの公式：平均値の法則

$$\underline{E(\upsilon)} = \frac{1}{\lambda}\underline{E(n)}$$

系内平均時間　　系内平均人数

$$\underline{E(w)} = \frac{1}{\lambda}\underline{E(m)}$$

平均待ち時間　　平均行列人数

この導出には確率分布（M／Mなど）を仮定していない。

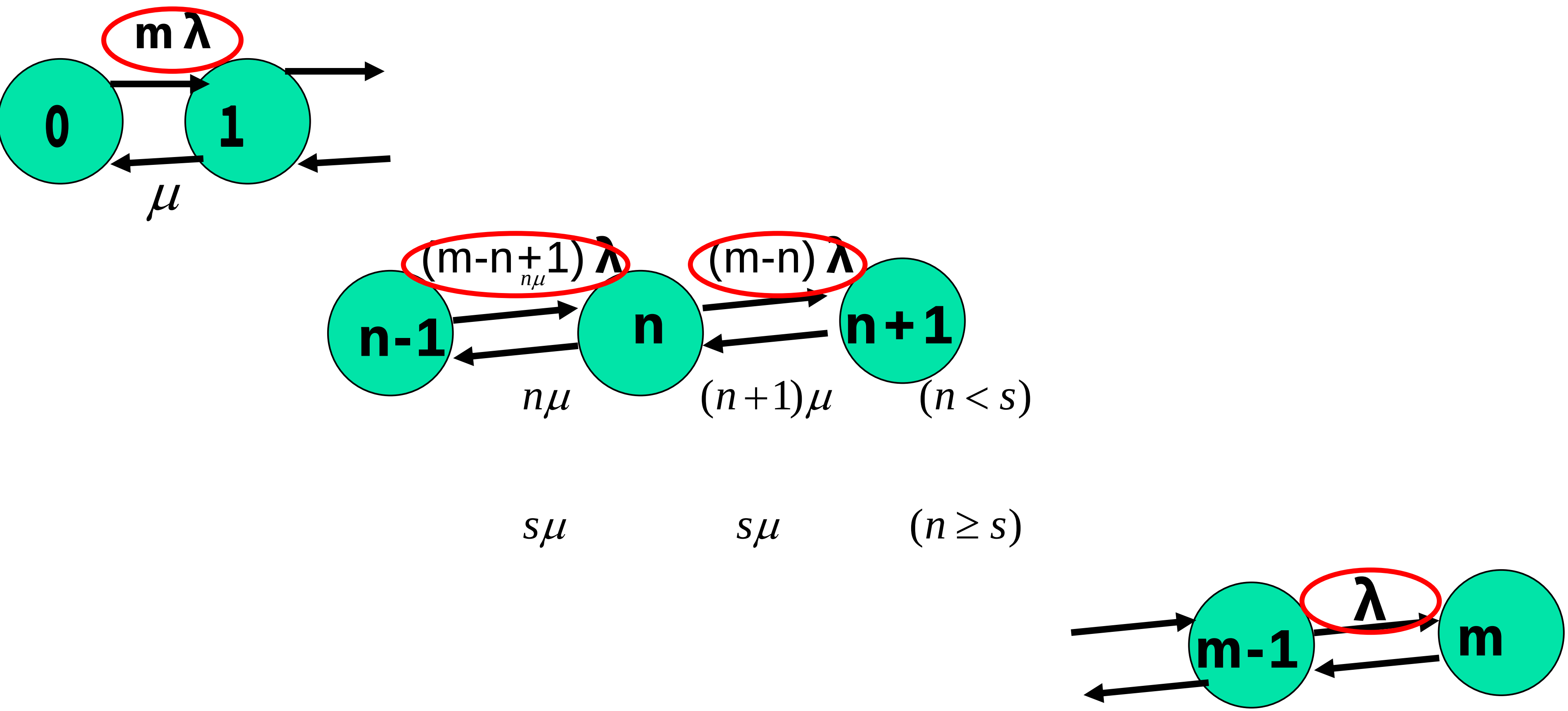
M／G／sでOK

有限呼源

- 今までは、客は無尽蔵にいて、次から次へと到着して来るものと想定していた。
- 客が元々極端に少ないと、その有限性を無視出来なくなる。
- 例：m台のコンピュータ、s人のシスアド（管理者）
  - mが小さいと到着確率は故障中の台数により変わる
  - mが十分大きければこれまでのモデルでよい

有限呼源

m台の計算機、s人のシスアド



有限呼源

平衡方程式

$$m\lambda\pi_0 = \mu\pi_1$$

$$[(m-n)\lambda + n\mu]\pi_n = (m-n+1)\lambda\pi_{n-1} + (n+1)\mu\pi_{n+1}$$

$$(1 \le n < s)$$

$$[(m-n)\lambda + s\mu]\pi_n = (m-n+1)\lambda\pi_{n-1} + s\mu\pi_{n+1}$$

$$(s \le n \le m)$$