wq-3. M/M/S 待ち行列, アーランの即時式モデル

（待ち行列の数理）

URL: https://www.kkaneko.jp/cc/wq/index.html

金子邦彦

アウトライン

3-1 M/M/S 待ち行列

3-2 アーランの即時式モデル

3-1 M/M/S 待ち行列

ケンドール記法 M/M/S

X:　到着過程

→**ポアソン分布**のとき「M」と書く

Y：　処理時間分布

→**指数分布**のとき「M」と書く

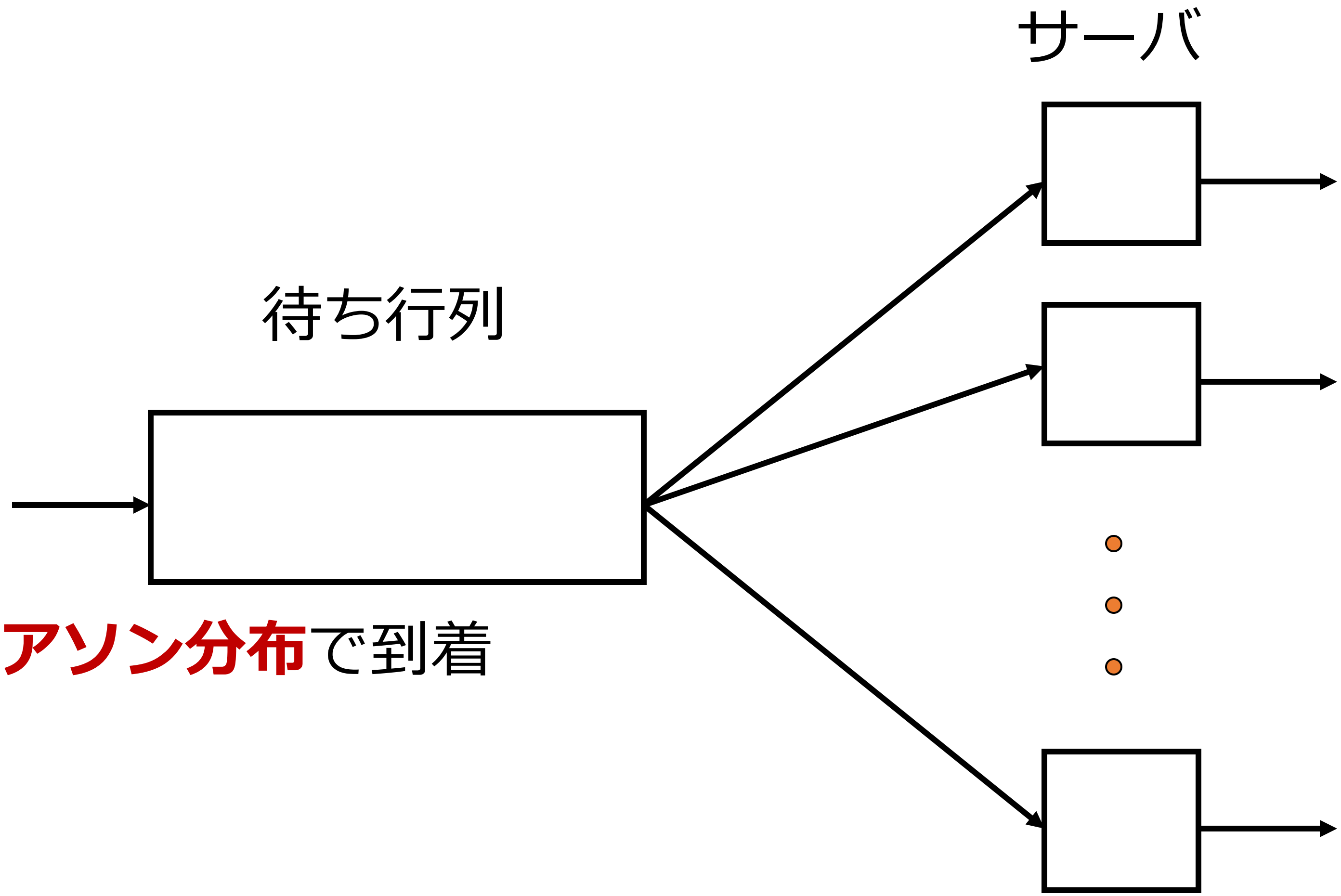
Z:　サーバ数

個数を限定しないとき「S」と書く

K：　**待ち行列の長さ**の制限

制限しないときは何も書かない

M/M/S 待ち行列モデル



ポアソン分布で到着

サーバ数：s

処理時間は**指数分布**

M/M/S 待ち行列モデルの解析

- 状態遷移
  - システム内のジョブ数を「状態」と見る
- 定常状態
  - 定常状態での、各「状態」の確率を求める
- システムの「処理率」を求める

状態

状態０：　システム内のジョブ数が　０

状態１：　システム内のジョブ数が　１

・
・
・

状態S:　システム内のジョブ数がS

・
・
・

状態遷移

- ジョブが到着：

  状態 k から状態 k＋1 に遷移

- ジョブが完了：

  状態 k から状態 k－1 に遷移

遷移確率に関する方程式

- 微小時間　⊿ｔ（限りなく０に近い）についての式
- 「ジョブの到着」と「ジョブの完了」は、「同時」には起きない
- 「ジョブの完了」の確率は、⊿ｔ とμの式
- 「ジョブの到着」の確率は、⊿ｔ とλの式

「ジョブの完了」に関する式

- k ≦S ならば（ k は状態番号）
  - サーバに空きがある

$$k\mu\Delta t\,(1-\mu\Delta t)^{k-1} \fallingdotseq k\mu\Delta t$$

- k ＞Sならば（ k は状態番号）
  - サーバは全て忙しい

$$S\mu\Delta t\,(1-\mu\Delta t)^{S-1} \fallingdotseq S\mu\Delta t$$

状態遷移

- ジョブが到着：

状態 k から状態 k＋1 に遷移

λ⊿t

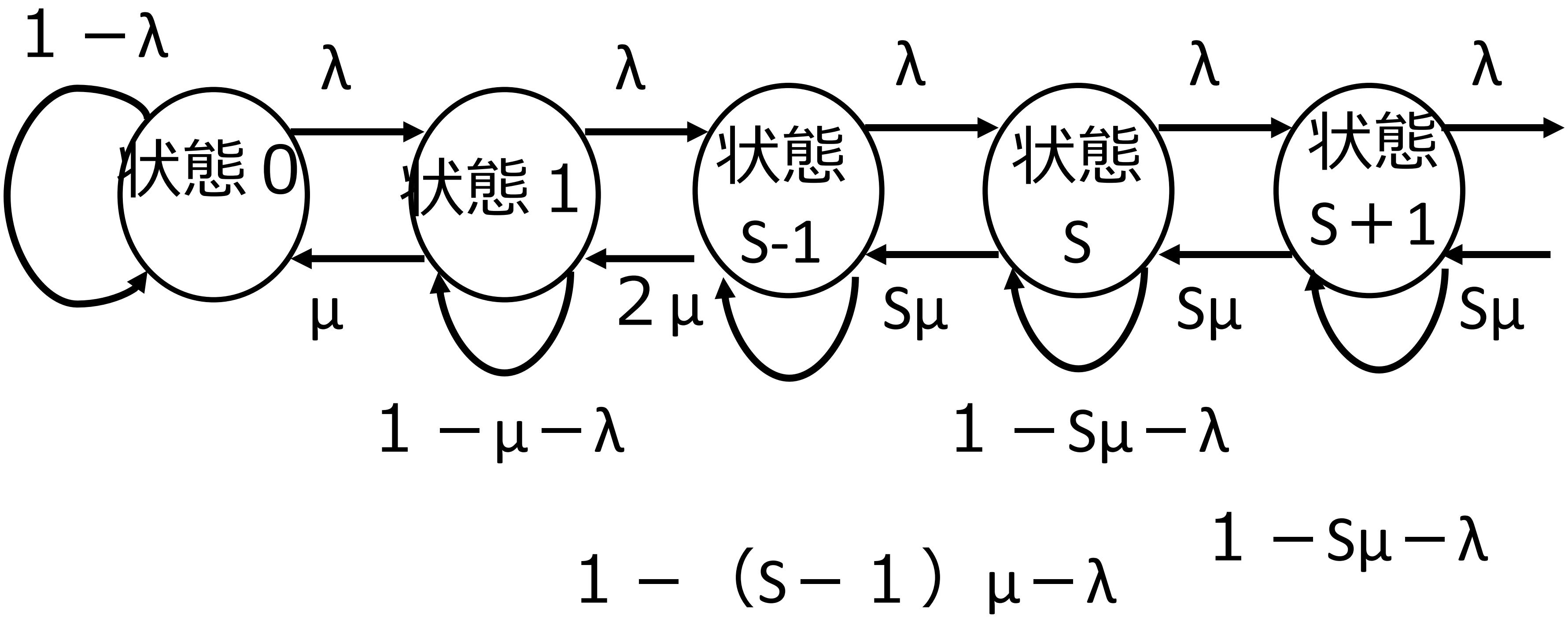
- ジョブが完了：

状態 k から状態 k − 1 に遷移

k ≦S ならば： k μ⊿t

k ＞ Sならば： Sμ⊿t

状態遷移図



定常状態方程式

$0 < n < S$ では

$$(\lambda + n\mu) P_n = \lambda P_{n-1} + (n+1)\mu P_{n+1}$$

$S \leqq n$ では

$$(\lambda + S\mu) P_n = \lambda P_{n-1} + S\mu P_{n+1}$$

$$\lambda P_0 = \mu P_1$$

定常確率をP0で表す

0 ＜n ＜S では

$$Pn = \left(\frac{\lambda}{\mu}\right)^n \frac{P0}{n!}$$

S ≦n　では

$$Pn = \left(\frac{\lambda}{\mu}\right)^n \frac{P0}{S!S^{n-S}}$$

システム処理能力ρ

ρ= λ/μ

- λ⊿ｔ： 「時間 (*t*, *t* +⊿ *t*) に到着するジョブ数」の平均
- μ⊿ｔ： 「サーバがジョブを処理中の間, ⊿ｔ 内に完了する処理数」の平均

待ち合わせが生じる確率

- システム内のジョブ数が S 以上

- Pqueue $= \sum_{n=S}^{\infty} Pn$

$$= \sum_{n=S}^{\infty} \left(\frac{\lambda}{\mu}\right)^{n} \frac{P0}{S!S^{n-S}}$$

$$= \sum_{n=S}^{\infty} P0 \frac{(S\rho)^{n}}{S!} \frac{1}{S^{n-S}}$$

3-2 アーランの即時式モデル

ケンドール記法 M/M/S/1

X:　到着過程

→**ポアソン分布**のとき「M」と書く

Y：　処理時間分布

→**指数分布**のとき「M」と書く

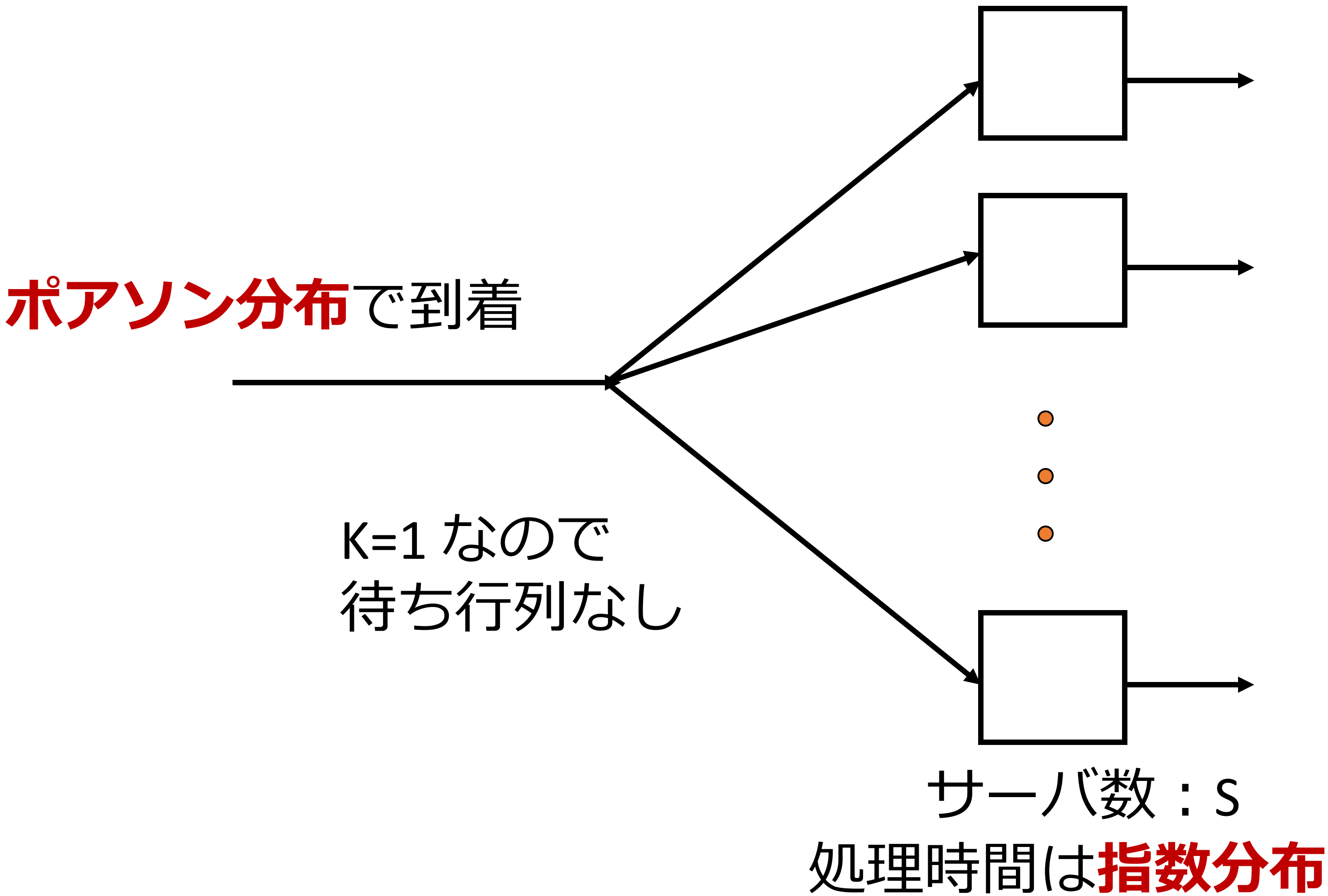
Z:　サーバ数

個数を限定しないとき「S」と書く

K：　**待ち行列の長さ**の制限

K = 1

M/M/S/1 のとき



アーランの即時式モデル

- 待ち行列モデル：　M/M/S/S
- 入力回線数：　無限
- 待ち行列長：　L＝ 0
- システム内の最大占有サーバ数：　K＝S

「即時式」の意味

K = 1 なので，次の性質を持つ

- 待ち行列なし
- 待ち行列長は 0
- システム内の占有サーバ数がS （すべてのサーバが占有されている）のとき，到着したジョブは棄却される
- サーバの空きがあれば，直ちに処理される

→　**即時式**

ジョブのモデル

- **ジョブの到着**
  - 平均λのポアソン分布

  → ある時刻に客が到着してから時間t内に次の客が到着する確率はポアソン分布に従う： $1-e^{-\lambda t}$
  - 微少時間⊿tの間にジョブが到着する確率：λ⊿ t
- **ジョブの処理時間**
  - 平均１/μの指数分布
  - 微少時間⊿tの間に処理が終わる確率： μ⊿ t

システム処理能力：ρ＝λ/μ

状態

状態０ ：　　系内のジョブの数が　０

状態１ ：　　系内のジョブの数が　１

・
・
・

状態S:　　　系内のジョブの数が　S

（状態はsまで）

状態遷移

- ジョブが到着：

  1) k ＜S のとき

  状態 k から状態 k＋1 に遷移

  2) k＝S のとき

  状態Sのまま（新しい到着は棄却される）

- ジョブの回線の占有終わり：

  状態 k から状態 k－1 に遷移

状態遷移

- ジョブが到着：

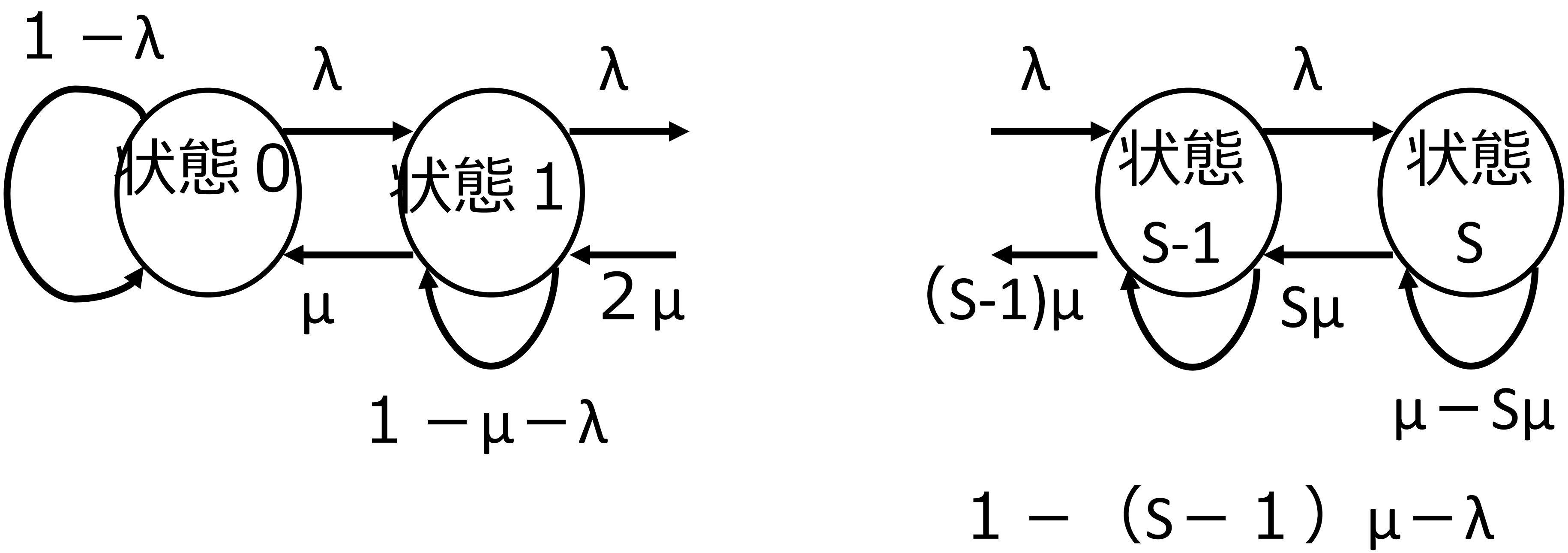
  1) k ＜S のとき

  状態 k から状態 k＋1 に遷移：　λ⊿t

- ジョブの回線の占有終わり：

  状態 k から状態 k－1 に遷移　：　k μ⊿t

状態遷移図



定常状態についての方程式

$$(\lambda + n\mu) P_n = \lambda P_{n-1} + (n+1)\mu P_{n+1}$$

$$\lambda P_0 = \mu P_1$$

$$\lambda P_{n-1} = S\mu P_n$$

棄却率

- 定常状態で，系内にn個のジョブ（0≦n≦s）がある確率を Pn とすると

$$P_n(t) = \frac{\dfrac{\rho^n}{n!}}{1 + \rho + \dfrac{\rho^2}{2} + \ldots + \dfrac{\rho^s}{s!}}$$

- 棄却率：
  - システム内の占有サーバ数が Sになる確率：Ps
  - 上の式で，n=s として計算する