1 確率システムの理論

- 決定的：A ならば B，論理
- 確率的：サイコロの目，到着するメッセージ数。
  結果か決まっているわけではないが，推定して有用な情報を得ることは可能

1.1 離散確率

- Ω 標本の集合
  例：サイコロ１つを投げる場合：
  $\Omega = \{$ １の目が出る，２の目が出る，３の目が出る，４の目が出る，５の目が出る，６の目が出る $\}$
  サイコロ２つを投げる場合：
  $\Omega = \{$ １と１の目が出る，１と２の目が出る$, \ldots,$ ６と６の目が出る $\}$
  メッセージ数の場合：
  $\Omega = \{0, 1, 2, 3, \ldots\}$
- 事象：Ω の部分集合
  例：サイコロの目が偶数，メッセージ数が３以下
- 事象 S に対する，確率 (測度)$P(S)$ が決まっている。
  測度：(部分) 集合に対して値を対応付けるもの
  S_1 と S_2 を，Ω の部分集合で，$S_1 \cap S_2 = \phi$ となるものとする。確率 (測度) は次の条件を満たさなくてはいけない。

$$\begin{aligned} P(\phi) &= 0 \\ P(\Omega) &= 1 \\ P(S_1 \cup S_2) &= P(S_1) + P(S_2) \end{aligned}$$

- (離散) 確率変数 X：Ω から 0 以上の整数の集合 N への写像

$$X : \Omega \to N$$

- $P(x)$：確率分布：以下のように定義される。

$$P(x) = P(\{\omega \,|\, X(\omega) = x\})$$

  次の関係が成立する。

$$0 \leq P(x) \leq 1$$
$$\sum_{x=0}^{\infty} P(x) = 1$$

確率変数の関数 $g(X)$ の期待値：

$$E[g(X)] = \sum_{x=0}^{\infty} g(x)P(x)$$

平均値

$$E[X] = \sum_{x=0}^{\infty} xP(x)$$

分散

$$Var[X] = \sum_{x=0}^{\infty} (x - E[X])^2 \, P(x)$$

n 次積率 (n モーメント)

$$X^{(n)} = \sum_{x=0}^{\infty} x^n P(x)$$

次の関係が成立する。

$$Var(X) = X^{(2)} - (X^{(1)})^2$$

z 変換：$|z| < 1$ なる複素数に対して,

$$G(z) = \sum_{x=0}^{\infty} z^x P(x) = E[z^x]$$

と定義する。サイコロの場合,

$$G(z) = \frac{1}{6}(z + z^2 + z^3 + z^4 + z^5 + z^6)$$

次の関係が成立する。

$$P(x) = \lim_{z \to 0} \frac{1}{x!} \frac{d^x}{dz^x} G(z)$$

$G(z)$ は確率母関数と呼ばれる。

$$\begin{aligned} X^{(1)} &= \lim_{z \to 1} \frac{d}{dz} G(z) \\ X^{(2)} &= \lim_{z \to 1} \frac{d}{dz} \left(z \frac{d}{dz} G(z) \right) \end{aligned}$$

X_1, X_2：2つの確率変数
$P_1(x_1)$, $P_2(x_2)$：それぞれの確率分布
$P(x_1, x_2) = P(\{\omega \,|\, X_1(\omega) = x_1,\ X_2(\omega) = x_1\}$：同時 (結合) 確率分布
X_1 と X_2 が**独立**：$P(x_1, x_2) = P_1(x_1)P_2(x_2)$

2つの確率変数の和も確率変数：

$$X_3 = X_1 + X_2$$

X_3 の確率分布 $P_3(x_3) = P(\{\omega \,|\, X_1(\omega) + X_2(\omega) = x_3\})$ は,

$$P_3(x_3) = \sum_{x_1=0}^{\infty} P_1(x_1)P_1(x_3 - x_1) = \sum_{x_2=0}^{\infty} P_1(x_3 - x_1)P_1(x_2)$$

となる。P_3 の z 変換 $G_3(z)$ は，

$$
\begin{aligned}
G_3(z) &= E[z^{X_3}] = E[z^{X_1+X_2}] = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} z^{x_1+x_2} P_1(x_1)P_2(x_2) \\
&= \sum_{x_1=0}^{\infty} z^{x_1} P_1(x_1) \sum_{x_2=0}^{\infty} z^{x_2} P_2(x_2) = G_1(z)G_2(z)
\end{aligned}
$$

のように，$G_1(z)$ と $G_2(z)$ の積になる。

1.2 連続確率

Ω (標本の集合) の例：長さ，通信時間
確率変数 X：通信で使うため，X は Ω から負でない実数の集合への写像とする。
事象の例：長さが 1cm から 3cm，通信時間が 3 分以内
累積確率分布関数：

$$F(x) = P(\{\omega \,|\, X(\omega) \leq x\})$$

確率密度関数：

$$f(x) = \frac{dF}{dx}$$

確率密度関数の意味：

$$f(x)\Delta x = F(x+\Delta x) - F(x) = P(\{\omega \,|\, x < X(\omega) \leq x + \Delta x\})$$

すなわち，$f(x)\Delta x$ は，確率変数 X が $[x, x+\Delta x]$ に含まれるような事象が生じる確率である。厳密には，区間は $(x, x+\Delta x]$ であるが，確率密度関数を考えるときは，1 点における確率は 0 であるから，区間の端点が含まれるかどうかに関しては無視してよい。

次の性質が成立する。

$$
\begin{aligned}
f(x) &\geq 0 \\
\int_0^{\infty} f(x)dx &= F(\infty) - F(0) = 1
\end{aligned}
$$

関数 $g(x)$ の期待値

$$E[g(X)] = \int_0^{\infty} g(x)f(x)dx$$

平均

$$E[X] = \int_0^{\infty} xf(x)dx$$

分散

$$Var[X] = \int_0^{\infty} (x - E[X])^2 f(x)dx$$

ラプラス変換 (Laplace 変換)

$$L(s) = \int_0^{\infty} e^{-sx} f(x)dx = E[e^{-sX}]$$

$$X^{(n)} = E[X^n] = \lim_{s \to 0} (-1)^n \frac{d^n}{ds^n} F(s)$$

$X_3 = X_1 + X_2$ のとき，次式が成立する。
(f_1，f_2，f_3：確率変数 X_1，X_2，X_3 の確率密度関数)

$$\begin{aligned} f_3(x_3) &= \int_0^\infty f_1(x_1) f_2(x_3 - x_1) dx \\ L_3(s) &= L_1(s) L_2(s) \end{aligned}$$

1.3 条件付き確率

A, B：事象 (Ω の部分集合)
$P(A, B)$：同時確率：事象 A と B が同時に生じる確率

$$P(A, B) = P(\{\omega \,|\, \omega \in A,\, \omega \in B\})$$

$P(A \,|\, B)$：条件付き確率：事象 B が生じるという仮定のもとで，事象 A が生じる確率。

$$P(A \,|\, B) = \frac{P(A, B)}{P(B)} = \frac{P(\{\omega \,|\, \omega \in A,\, \omega \in B\})}{P(\{\omega \,|\, \omega \in B\})}$$

2 ポアソン分布

- Δt の間にものごとが生じる確率が $\lambda \Delta t$
- ポアソン分布の確率変数 K：
  $K = k$ は，時間 t の間にそのものごとが k 回生じることを表す。
- メッセージの到着の分布などに使われる。

t を n 個に分割する。

- n が十分大きければ，分割された時間内に 2 回生じる影響は無視できる。
- t/n の間に 1 回生じる確率は，$\lambda t / n$
- t の間に k 回生じる確率：

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

$n \to \infty$ とすれば，

$$P(k) = P(\{\omega \,|\, K(\omega) = k\}) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

となる。

$$E[K] = \lambda t \tag{1}$$
$$Var[K] = \lambda t \tag{2}$$
$$E[K^2] = (\lambda t)^2 + \lambda t \tag{3}$$

ポアソン分布の z 変換は次のようになる。

$$P(z) = e^{-\lambda t(1-z)}$$

2.1　ポアソン到着

例：時間 t の間にメッセージが k 個到着する確率。

- 独立性
- 定常性
- 微小時間内の希少性

$\Rightarrow$ ポアソン分布になる。

2.1.1　合流

合流：メッセージの発生源が 2 つあり，1 箇所に集まる。
両者のメッセージの到着は，ポアソン分布に従う。それぞれの発生する数を，確率変数を X_1 に X_2 で表す。
合流して到着するメッセージの数は,

$$K_3 = K_1 + K_2$$

となる。
z 変換で考えると，

$$P_3(z) = P_1(z)P_2(z) = e^{-\lambda_1 t(1-z)}e^{-\lambda_2 t(1-z)} = e^{-(\lambda_1+\lambda_2)t(1-z)}$$

となる。従って，$\lambda_1 + \lambda_2$ のポアソン分布になる。

2.1.2　分流

分流：メッセージの発生源は 1 つで，それを 2 カ所で受け付ける。
確率 ϵ と $1-\epsilon$ でランダムに 2 つに分ける。
$P[K_1 = k_1, K_2 = k_2 \,|\, K_1 + K_2 = k]$ で，$K_1 + K_2 = k$ という条件の元で，$K_1 = k_1$ かつ $K_2 = k_2$ となる確率を表すものとする。

$$\begin{aligned}
& P[K_1 = k_1, K_2 = k_2 \,|\, K_1 + K_2 = k] \\
= \;& P(\{\omega \,|\, K_1(\omega) = k_1, K_2(\omega) = k_2\} \,|\, \{\omega \,|\, K_1(\omega) + K_2(\omega) = k\}) \\
= \;& {}_kC_{k_1}\epsilon^{k_1}(1-\epsilon)^{k_2} = \frac{k!}{k_1!k_2!}\epsilon^{k_1}(1-\epsilon)^{k_2}
\end{aligned}$$

$$
\begin{aligned}
P[K_1 = k_1, K_2 = k_2] &= P[K_1 = k_1, K_2 = k_2 \,|\, K_1 + K_2 = k]P[K_1 + K_2 = k] \\
&= \frac{k!}{k_1!k_2!}\epsilon^{k_1}(1-\epsilon)^{k_2}\frac{(\lambda t)^k e^{-\lambda t}}{k!} \\
&= 6\frac{(\epsilon\lambda t)^{k_1} e^{-\epsilon\lambda t}}{k_1!} \cdot \frac{((1-\epsilon)\lambda t)^{k_2} e^{-(1-\epsilon)\lambda t}}{k_2!}
\end{aligned}
$$

平均 $\epsilon\lambda t$ と $(1-\epsilon)\lambda t$ のポアソン分布の同時確率分布関数の積になっている。

3 指数分布

- 放射性物質の崩壊する時間の分布
- ジョブがサーバーを占有する時間 (保留時間) の分布
- 指数分布の確率変数 T：
  $T \leq t$ で，時刻 t までに事象が生じることを表す。
- μ：Δt の間に終了する確率が $\mu\Delta t$ で与えられる。
- $H(t) = P(\{\omega \,|\, T(\omega) \leq t\})$：時刻 t までには終了する確率
- $1 - H(t)$：時刻 t までに終了していない確率
- $h(t)$：時刻 t で終了する確率密度関数

$$h(t) = \frac{dH}{dt}$$

- 指数分布の微分方程式：

$$H(t + \Delta t) - H(t) = \mu\Delta t(1 - H(t))$$

- その解：

$$H(t) = 1 - e^{-\mu t}$$

$$h(t) = \mu e^{-\mu t}$$

$$
\begin{aligned}
E(T) &= 1/\mu \\
E(T^2) &= 2(1/\mu)^2 \\
Var(T) &= (1/\mu)^2
\end{aligned}
$$

指数分布のラプラス変換は次のようになる。

$$H(s) = \mu/(s + \mu)$$

3.1 まとめ

ポアソン分布：**個数**
指数分布：**時間**

3.2 時間間隔の分布

$$
\begin{aligned}
P[T > t] &= \text{時間間隔が } t \text{ より大となる確率} \\
&= \text{ポアソン分布で，t までに } k = 0 \text{ となる確率} \\
&= \frac{(\mu t)^0 e^{-\mu t}}{0!} = e^{-\mu t}
\end{aligned}
$$

従って，

$$P[T \leq t] = 1 - P[T > t] = 1 - e^{-\mu t}$$

指数分布の累積分布関数になっている。
$\rightarrow$ 時間間隔も指数分布

3.3 無記憶性

$$P[T < t + t_0 \,|\, T > t_0] = P[T < t]$$

時刻 t_0 までに生じなかったという条件のもとで，時刻 $t + t_0$ までに生じる確率と，単に時刻 t までに生じる確率が等しい。
指数分布では，

$$
\begin{aligned}
P[T < t + t_0 \,|\, T > t_0] &= H(t + t_0 \,|\, t_0) \\
&= \frac{H(t + t_0) - H(t_0)}{1 - H(t_0)} = \frac{1 - e^{-\mu(t+t_0)} - (1 - e^{-\mu t_0})}{1 - (1 - e^{-\mu t_0})} \\
&= \frac{e^{-\mu t_0} - e^{-\mu(t+t_0)}}{e^{-\mu t_0}} = 1 - e^{-\mu t} = H(t) = P[T < t]
\end{aligned}
$$

となり，無記憶性が成立する。

3.4 2 相アーラン (Erlang) 分布

平均処理時間 $1/\mu_1$ の処理の後，平均処理時間 $1/\mu_2$ の処理行うときの，全体の時間の分布

$$T = T_1 + T_2$$

ラプラス変換

$$H(s) = H_1(s)H_2(s) = \frac{\mu_1\mu_1}{(s + \mu_1)(s + \mu_2)}$$

となる。逆ラプラス変換すれば,

$$h(t) = \begin{cases} -\frac{\mu_1\mu_2}{\mu_1-\mu_2}\left(e^{-\mu_1 t} - e^{-\mu_2 t}\right) & \mu_1 \neq \mu_2 \\ \mu_1^2 t e^{-\mu_1 t} & \mu_1 = \mu_2 \end{cases}$$

3.5 超指数分布

Hyper exponential distribution
確率 ϵ_i で，平均処理時間 $1/\mu_i$ のサーバーにふり分けるときの，処理時間の分布。

$$h(t) = \sum_{i=1}^{n} \epsilon_i \mu_i e^{-\mu_i t}$$

4 マルコフ過程

Ω：システムがとりうる状態の集合 (離散集合)
$S(t)$：時刻 t における状態を表す確率変数
$P(S(t) = a)$：t における状態が a である確率
時間的に一様 (stationary)：

- 確率的な構造が時間とともに変化しない。(ある状態から別の状態に遷移する確率)
- 個々の状態に存在する確率などは変化する。

マルコフ過程

- 次の状態の確率が現在の状態と現在の入力だけで決まる。
- 現在より過去の状態には，直接的には依存しない。

4.1 時間的に離散なマルコフ過程

S_k : $k = 0, 1, 2, \ldots$ における状態を表す確率変数
マルコフ過程

$$P[S_{n+1} = a_{n+1} \,|\, S_n = a_n, S_{n-1} = a_{n-1}, \ldots, S_0 = a_0] = P[S_{n+1} = a_{n+1} \,|\, S_n = a_n]$$

定常性は，この確率の値が n によらないこと。
状態も離散で，0, 1, 2, $\ldots, N$ と番号付けられている。そして,

$$P_{ij} = P[S_{n+1} = i \,|\, S_n = j]$$

と定義する。
いずれの状態 j にあっても，次にどこかの状態に移動するので，次式が成立する。

$$\sum_{i=0}^{N} P_{ij} = 1$$

いま，状態 j のときに，m ステップ後に状態 i になる確率を，$P_{ij}^{(m)}$ とおけば，

$$P_{ij}^{(m)} = \sum_{k=0}^{N} P_{ik}^{(m-1)} P_{kj} = \sum_{k=0}^{N} P_{ik} P_{kj}^{(m-1)}$$

が成立する。チャップマン・コルモゴロフの方程式

行列で表記する。

$$\boldsymbol{P} = \begin{pmatrix} P_{00} & \cdots & P_{0N} \\ \cdots & \cdots & \cdots \\ P_{N0} & \cdots & P_{nN} \end{pmatrix}$$

および，

$$\boldsymbol{P}^{(n)} = \begin{pmatrix} P_{00}^{(n)} & \cdots & P_{0N}^{(n)} \\ \cdots & \cdots & \cdots \\ P_{N0}^{(n)} & \cdots & P_{nN}^{(n)} \end{pmatrix}$$

とすれば，チャップマン・コルモゴロフの方程式より，

$$\boldsymbol{P}^{(n)} = \boldsymbol{P}^{(n-1)} \boldsymbol{P} = (\boldsymbol{P})^n$$

が成立する。

$p_i(n)$：時刻 n において，状態が i である確率

$$\begin{aligned} \sum_{i=1}^{N} p_i(n) &= 1 \\ \boldsymbol{p}(n) &= \begin{pmatrix} p_0(n) \\ \vdots \\ p_N(n) \end{pmatrix} \\ \boldsymbol{p}(n+1) &= \boldsymbol{P}\boldsymbol{p}(n) \qquad \Leftarrow \textbf{状態方程式} \end{aligned}$$

が成立する

- 到達可能：状態 i から状態 j へ到達可能：有限回のステップで，状態 i から状態 j に遷移する確率が 0 でない。
- 既約：すべての状態間が到達可能
- 周期性：状態 i から状態 i 遷移する (同じ状態に戻る) 確率が 0 でないステップ数が，k $(k > 1)$ の倍数のときだけの場合
- エルゴード性 (ergodic)：時間平均が平均 (アンサンブル平均，母集団平均) と等しくなる。
  有限マルコフ過程は，既約で非周期ならば，エルゴード性をもつ。
  $t \to \infty$ で定常状態に収束する

$$\lim_{n \to \infty} \boldsymbol{p}(n) = \boldsymbol{p}^*$$

となる。

$\boldsymbol{p}^*$：定常状態確率
$\boldsymbol{p}(n+1) = \boldsymbol{P}\boldsymbol{p}(n)$ で $n \to \infty$ とすれば,

$$\boldsymbol{p}^* = \boldsymbol{P}\boldsymbol{p}^*$$

が成立する。**定常状態方程式**と呼ばれる。

4.2 連続時間型マルコフ過程

$S(t)$：時刻 t における状態の確率変数
任意の時間分割

$$t_{n+1} > t_n > \cdots > t_0$$

に対して,

$$P[S(t_{n+1}) = a_{n+1} \,|\, S(t_n) = a_n, \ldots, S(t_0) = a_0] = P[S(t_{n+1}) = a_{n+1} \,|\, S(t_n) = a_n]$$

が成立する。
$P_{ij}(\tau)$：状態 j から時間 τ だけ経過したときに，状態 j に遷移している確率。

$$P_{ij}(\tau) = P[S(t+\tau) = i \,|\, S(t) = j]$$

チャップマン・コロモゴロフ方程式 (時間 t を 2 つに区分)

$$\begin{aligned} P_{ij}(t) &= \sum_{k=1}^{N} P_{ik}(\Delta\tau)P_{kj}(t-\Delta\tau) && \text{前進方程式} \\ P_{ij}(t) &= \sum_{k=1}^{N} P_{ik}(t-\Delta\tau)P_{kj}(\Delta\tau) && \text{後退方程式} \end{aligned}$$

$p_i(t)$：t において状態 i である確率

$$\begin{aligned} p_i(t) &= \sum_{j=0}^{N} P_{ij}(t)p_j(0) = \sum_{j=0}^{N}\sum_{k=0}^{N} P_{ik}(\Delta\tau)P_{kj}(t-\Delta\tau)p_j(0) \\ &= \sum_{k=0}^{N} P_{ik}(\Delta\tau)p_k(t-\Delta\tau) \end{aligned}$$

が成立する。従って,

$$\frac{p_i(t) - p_i(t-\Delta\tau)}{\Delta\tau} = \sum_{k=0}^{N}(P_{ik}(\Delta\tau) - \delta_{ik})p_k(t-\Delta\tau)$$

となる。

$$q_{ik} = \begin{cases} \lim_{\Delta\tau\to 0} \frac{P_{ik}(\Delta\tau)}{\Delta\tau} & k \neq i \\ \lim_{\Delta\tau\to 0} \frac{P_{ii}(\Delta\tau)-1}{\Delta\tau} & k = i \end{cases}$$

とおけば，次式が成立する。

$$\frac{dp_i(t)}{dt} = \sum_{k=0}^{N} q_{ik}p_k(t)$$

q_{ij}：微小変化率

$t \to \infty$ に対して，定常状態確率 p_j^* は,

$$\frac{dp_i^*(t)}{dt} = 0$$

より，

$$\sum_{k=0}^{N} q_{ik} p_k^*(t) = 0$$

を解けば求められる。ただし，$p_k^*(t)$ は確率であるので，

$$\sum_{k=0}^{N} p_k^*(t) = 1$$

を満たさなくてはいけない。

5　待ち行列モデル

ケンドル表記

到着の確率分布 / 処理の確率分布 / サーバーの数 / システムの容量

分布：

M：ポアソン分布 or 指数分布
E_k：アーラン分布
H_b：超指数分布
D：一定分布 (到着間隔，処理時間が固定されている)
G：一般の分布

システムの容量：システムに格納できる最大のジョブ数 (サーバー＋待ち行列)
例：M/M/1/K
まず，$K = \infty$ のときを考える。
$S(t)$：時刻 t における待ち行列とサーバで処理中のジョブの数
$S(t) = 0$：サーバにも待ち行列にもジョブがない。
$S(t) = 1$：サーバで 1 つのジョブが処理中で，待ち行列にはジョブがない。
$S(t) \geq 2$：サーバで 1 つのジョブが処理中で，待ち行列には $s(t) - 1$ のジョブが待機中。
λ：ジョブの到着率
μ：ジョブの終了率
t から $t + \Delta\tau$ への変化を考える。
時間が短いため，2 つ以上の変化が生じることは無視できる ($\Delta\tau$ の 2 次以上のオーダになる)。
$S(t) \geq 1$ のとき

状態変化	確率	
$S(t+\Delta\tau) = S(t) + 1$	$\lambda\Delta\tau$	：ジョブが到着
$S(t+\Delta\tau) = S(t) - 1$	$\mu\Delta\tau$	：ジョブが終了
$S(t+\Delta\tau) = S(t)$	$1 - \lambda\Delta\tau - \mu\Delta\tau$	：ジョブが到着も終了もしない

$S(t)=0$ のとき，(この場合はジョブの終了はない)

状態変化	確率	
$S(t+\Delta\tau)=S(t)+1$	$\lambda\Delta\tau$	：ジョブが到着
$S(t+\Delta\tau)=S(t)$	$1-\lambda\Delta\tau$	：ジョブが到着しない

となる。

$$\boldsymbol{P}(\Delta\tau)=\begin{pmatrix} 1-\lambda\Delta\tau & \mu\Delta\tau & 0 & 0 & \dots \\ \lambda\Delta\tau & 1-\lambda\Delta\tau-\mu\Delta\tau & \mu\Delta\tau & 0 & \ddots \\ 0 & \lambda\Delta\tau & 1-\lambda\Delta\tau-\mu\Delta\tau & \mu\Delta\tau & \ddots \\ 0 & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

となる。$\Delta\tau\to 0$ として，微分方程式を立てると，

$$\begin{aligned} \frac{dp_0(t)}{dt} &= -\lambda p_0(t)+\mu p_1(t) \\ \frac{dp_n(t)}{dt} &= \lambda p_{n-1}(t)-(\lambda+\mu)p_n(t)+\mu p_{n+1}(t) \qquad (n\geq 2) \end{aligned}$$

となる。$t\to\infty$ として，定常状態 p_n^* を考える。定常状態では時間微分が 0 になるから，

$$\begin{aligned} -\lambda p_0^*+\mu p_1^* &= 0 \\ \lambda p_{n-1}^*-(\lambda+\mu)p_n^*+\mu p_{n+1}^* &= 0 \quad (n\geq 2) \\ \sum_{n=0}^{\infty} p_n^* &= 1 \end{aligned}$$

が成立する必要がある。

上式から，

$$\begin{aligned} p_1^*-\frac{\lambda}{\mu}p_0^* &= 0 \\ p_{n+1}^*-\frac{\lambda}{\mu}p_n^* &= p_n^*-\frac{\lambda}{\mu}p_{n-1}^* \end{aligned}$$

となり，

$$p_{n+1}^*-\frac{\lambda}{\mu}p_n^*=0$$

となる。

$$\rho=\frac{\lambda}{\mu}$$

とおけば，$p_n^*=\rho^n p_0^*$ であり，

$$\sum_{n=0}^{\infty}\rho^n p_0^*=1$$

より，$\rho<1$ のときに解を持ち，次の結果が得られる。

$$p_n^*=(1-\rho)\rho^n$$

注意：$\rho = 1$：到着と処理の頻度が同じ場合は定常解がない。

z 変換

$$p(z) = \sum_{n=0} p_n^* z^n = \frac{1-\rho}{1-\rho z}$$

平均ジョブ数 (システム内のジョブ数)：

$$N = \sum_{n=0} n p_n^* = \left.\frac{dp(z)}{dz}\right|_{z=0} = \frac{\rho}{1-\rho}$$

平均ジョブ待ち数：(到着したときに処理を待っている個数)

$$\mu_W = \sum_{n=1} (n-1) p_n^* = \sum_{n=0} n p_n^* - \sum_{n=1} p_n^* = N - (1 - p_0^*) = \frac{\rho^2}{1-\rho}$$

平均システム遅延時間 D：到着から処理が終わるまでの時間

$$D = \frac{1}{\mu} \sum_{n=1} (n+1) p_n^* = \frac{N+1}{\mu} = \frac{1}{\mu} \frac{\rho + (1-\rho)}{1-\rho} = \frac{1}{\mu} \frac{1}{1-\rho}$$

リトルの公式

時間 D の間に到着したジョブ数 = 平均ジョブ数　$(D\lambda = N)$

が成立する。従って，次式が成立する。

$$D = \frac{N}{\lambda} = \frac{\rho}{\lambda(1-\rho)} = \frac{1}{\mu} \frac{1}{1-\rho}$$

注意：N/μ ではない。$p_0^* = 1 - \rho$ より，

$$\frac{N}{\lambda} = \frac{N}{\mu\rho} = \frac{N}{\mu(1-p_0^*)}$$

と書くことができる (システムにジョブが全くないときは処理をしていないため)。

5.1 待ち行列の長さが有限

$M/M/1/K$

待ち行列の長さ $K - 1$

システムがいっぱいの時は，入力を受け付けない (入力を破棄する)。

$$\begin{aligned}
p_0(t+\Delta\tau) &= (1-\lambda\Delta\tau) p_0(t) + \mu\Delta\tau p_1(t) \\
&\vdots \\
p_n(t+\Delta\tau) &= \lambda\Delta\tau p_{n-1}(t) + (1-\lambda\Delta\tau-\mu\Delta\tau) p_n(t) + \mu\Delta\tau p_{n+1}(t) \qquad (K-1 \le n \le 1) \\
&\vdots \\
p_K(t+\Delta\tau) &= \lambda\Delta\tau p_{K-1}(t) + (1-\lambda\Delta\tau) p_K(t)
\end{aligned}$$

となる。定常状態では，

$$p_n^* = \frac{1-\rho}{1-\rho^{K+1}}\rho^n$$

となる。

$S(t)=K$ のとき，入力が来ても棄却される。

p_K^* は棄却率と等しい。

平均ジョブ数：

$$N=\sum_{n=0}^{K} np_n^* = \frac{\rho(1-\rho^{K+1})-(K+1)\rho^{K+1}(1-\rho)}{(1-\rho)(1-\rho^{K+1})}$$

平均システム遅延：

$$D=\frac{1}{\lambda}\frac{N}{(1-p_K^*)}$$

- $\lambda(1-p_K^*)$ が単位時間あたりに受付られるジョブ数
  (リジェクトしたものは待ち時間の平均の計算に参入しない。)

5.2 サーバ数が複数で待ち行列の長さが有限

$M/M/S/K$

待ち行列の長さ $K-S$

サーバは空いているところへ投入する。

$$\begin{aligned}
p_0(t+\Delta\tau) &= (1-\lambda\Delta\tau)p_0(t)+\mu\Delta\tau p_1(t) \\
&\vdots \\
p_n(t+\Delta\tau) &= \lambda\Delta\tau p_{n-1}(t)+(1-\lambda\Delta\tau-n\mu\Delta\tau)p_n(t)+(n+1)\mu\Delta\tau p_{n+1}(t) \qquad (1\le n\le S-1) \\
&\vdots \\
p_n(t+\Delta\tau) &= \lambda\Delta\tau p_{n-1}(t)+(1-\lambda\Delta\tau-S\mu\Delta\tau)p_n(t)+S\mu\Delta\tau p_{n+1}(t) \qquad (S\le n\le K-1) \\
&\vdots \\
p_K(t+\Delta\tau) &= \lambda\Delta\tau p_{K-1}(t)+(1-S\lambda\Delta\tau)p_K(t)
\end{aligned}$$

となる。

$\Delta\tau$ の間に同時に 2 つ以上のジョブが終了することは無視

ただし，1 つのジョブが終了する確率はジョブを処理しているサーバの数に比例する。

$$\begin{aligned}
p_1^*-\rho p_0^* &= 0 \\
\left(p_{n+1}^*-\frac{\rho}{n+1}p_n^*\right) &= \frac{n}{n+1}\left(p_n^*-\frac{\rho}{n}p_{n-1}^*\right) \qquad 1\le n\le S-1 \\
\left(p_{n+1}^*-\frac{\rho}{S}p_n^*\right) &= \left(p_n^*-\frac{\rho}{S}p_{n-1}^*\right) \qquad S\le n\le K-1 \\
0 &= p_K^*-\frac{\rho}{S}p_{K-1}^* \\
\sum_{n=0}^{K} p_n^* &= 1
\end{aligned}$$

これを解くと，

$$p_n^* = \begin{cases} \frac{\rho^n}{n!} p_0^* & (n \leq S-1) \\ \frac{\rho^S}{S!} \left(\frac{\rho}{S}\right)^{n-S} p_0^* & (S \leq n \leq K) \end{cases}$$

ここで，p_0^* は和が 1 になる条件から求まる。

$$p_0^* = \left[\sum_{n=1}^{S-1} \frac{\rho^n}{n!} + \frac{\rho^S}{S!} \frac{1 - \left(\frac{\rho}{S}\right)^{K+1-S}}{1 - \left(\frac{\rho}{S}\right)} \right]^{-1}$$

となる。

6　通信トラフィック理論

- 呼 (call)：通信トラフィック
- トラフィック量：呼の通信時間の総和

$$\text{トラフィック密度} = \frac{\text{トラフィック量}}{\text{観測時間}}$$

- 単位：erl (アーラン)
  1 [erl]：1 時間あたり 1 本の通信回線を 1 時間使用する。
- トラフィックの推定法 (の一つ)

$$\lambda_{ij} = \alpha \frac{P_i P_j}{D_{ij}}$$

P_i, P_j：人口 D_{ij}：都市 i, j 間の距離

6.1　アーラン即時式モデル

$M/M/S/S$

- ポアソン到着
- 通信時間：指数分布
- 回線数：S
- 待ち行列：0

定常状態確率：$M/M/S/K$ で，$K = S$ とする。

$$\begin{aligned} p_0^* &= (1 - \lambda\Delta\tau) p_0^* + \mu\Delta\tau p_1^* \\ &\vdots \\ p_n^* &= \lambda\Delta\tau p_{n-1}^* + (1 - \lambda\Delta\tau - n\mu\Delta\tau) p_n^* + (n+1)\mu\Delta\tau p_{n+1}^* \qquad (1 \leq n \leq S-1) \\ &\vdots \\ p_S^* &= \lambda\Delta\tau p_{K-1}^* + (1 - S\mu\Delta\tau) p_K^* \end{aligned}$$

解は次のようになる。

$$p_0^* = \left[\sum_{n=1}^{S} \frac{\rho^n}{n!}\right]^{-1}$$
$$p_n^* = p_0^* \frac{\rho^n}{n!}$$

損失確率：回線がすべて使用されていて通信できない確率 (「棄却率」と同じ)
アーランの損失式 (アーランの B 式)

$$p_S^* = \frac{\rho^S}{S! \sum_{i=1}^{S} \frac{\rho^n}{n!}}$$

6.2 エッグセットの即時モデル

$M/M/S/S\ (K)$

- ポアソン到着
- 通信時間：指数分布
- 回線数：S
- 待ち行列：0
- 全ユーザ数が K
  通話している人の数だけ，到着する確率が減少する。
  λ_0：単位人数あたりの平均到着数 (到着率$= \lambda_0 \times$ 通信していない人の数)

定常状態では，

$$\begin{aligned}
p_0^* &= \mu\Delta\tau p_1^* + (1 - K\lambda_0\Delta\tau)p_0^* \\
&\vdots \\
p_n^* &= (K-n+1)\lambda_0\Delta\tau p_{n-1}^* + (1-(K-n)\lambda_0\Delta\tau + n\mu\Delta\tau)p_n^* + (n+1)\mu\Delta\tau p_{n+1}^* \quad (1 \le n \le S-1) \\
&\vdots \\
p_K^* &= (K-S+1)\lambda_0\Delta\tau p_{K-1}^* + (1 - S\lambda_0\Delta\tau)p_K^*
\end{aligned}$$

となる。この式から，

$$(n+1)\mu p_{n+1}^* - \lambda_0(K-n)p_n^* = n\mu p_n^* - \lambda_0(K-(n-1))p_{n-1}^*$$

が成立する。従って，

$$p_n^* = \rho\frac{K-n+1}{n}p_{n-1}^* = \rho^n\frac{(K-n+1)(K-n)\cdots K}{n(n-1)\cdots 1}p_0^* = \rho^n {}_K C_n p_0^*$$

となり，

$$p_0^* = \frac{1}{\sum_{i=0}^{S} {}_K C_i \rho^i}$$

となる。従って，

$$p_n^* = \frac{\rho^n {}_K C_n}{\sum_{i=0}^{S} {}_K C_i \rho^i}$$

となる。

6.3 パケット交換のモデル

パケットの伝達にかかる時間

可変長	指数分布	$M/M/1$
固定長	一定分布	$M/D/1$
一般	確率密度関数 $f(b)$	$M/G/1$

パケットの伝送時間を単位として，離散時間 n で考える。

(単位時間に 1 つのパケットの伝送が完了する)

離散時間といっても，等間隔とは限らない。

S_n：時刻 n においてシステム内に残っているパケット数

A_{n+1}：時刻 n から時刻 $n+1$ までに到着したパケット数

次の関係が成立する。

$$S_{n+1} = \begin{cases} A_{n+1} & (S_n = 0) \\ S_n - 1 + A_{n+1} & (S_n \geq 1) \end{cases}$$

$$I(k) = \begin{cases} 0 & (k = 0) \\ 1 & (k \geq 1) \end{cases}$$

とおけば,

$$S_{n+1} = S_n - I(S_n) + A_{n+1} \tag{4}$$

となる。$n \to \infty$ において，$S = \lim_{n\to\infty} S_n$ の平均値を求める。

$$E[S] = E[S] - E[I(S)] + E[A] \tag{5}$$

となる。また,

$$E[I(S)] = (0 \cdot P[S = 0]) + (1 \cdot P[S \geq 1]) = P[S \geq 1]$$

となる。

$\rho = P[S \geq 1]$ とおく。

式 (5) より，$E[A] = E[I(S)] = \rho$ となる。

式 (4) の両辺を 2 乗して,

$$S_{n+1}^2 = S_n^2 + I(S_n)^2 + A_{n+1}^2 - 2S_nI(S_n) - 2A_{n+1}I(S_n) + 2A_{n+1}S_n$$

を得る。これを平均して $n \to \infty$ とする。

$$E[S^2] = E[S^2] + E[I(S)^2] + E[A^2] - 2E[SI(S)] - 2E[AI(S)] + 2E[AS]$$

となる。$I(S)$ の定義から，$E[I(S)^2] = E[I(S)]$，$E[SI(S)] = E[S]$ となる。S と A は独立と仮定しているので，$E[AI(S)] = E[I(S)]E[A]$ と，$E[AS] = E[A]E[S]$ が成立する。

従って,

$$E[S] = \frac{\rho - 2\rho^2 + E[A^2]}{2(1-\rho)}$$

となる。

- 到着するパケット数の平均値を表す確率変数 A を，λB のポアソン分布の確率変数とする (B はパケットの伝送時間を表す確率変数)。
- B を固定すれば，A は単なるポアソン分布として見なされる。

$$E[A \mid B = b] = \lambda b$$

- 確率変数 B の確率密度関数を $f(b)$ とする。

$$E[A] = \lambda E[B]$$

- ポアソン分布の式 (2) より

$$E[A^2 \mid B] = (\lambda B)^2 + (\lambda B)$$

となるので (E_B は，確率変数 B に関する期待値)，

$$E[A^2] = E_B[E[A^2 \mid B]] = \lambda^2 E_B[B^2] + \lambda E_B[B] = \lambda^2 E_B[B^2] + \rho$$

となる。
- 処理時間の変動を表す C_b を，

$$\lambda^2 E[B^2] = \rho^2(1 + C_b^2)$$

で定義する。
- システムに残っている平均パケット数に関して，次のポラチェック・ヒンチンの式が成立する。

$$E[S] = \rho + \rho^2 \frac{1 + C_b^2}{2(1-\rho)}$$

- リトルの公式と $\rho = E[A] = \lambda E[B]$ より，平均処理時間は次式のようになる。

$$D = \frac{E[S]}{\lambda} = E[B]\left\{1 + \rho\frac{1 + C_b^2}{2(1-\rho)}\right\}$$