The Distribution of the Minimum of Independent Phase Type Random Variables

Darcy Bermingham, supervised by Yoni Nazarathy.
The University of Queensland School of Mathematics and Physics

December 2015

In this short note we illustrate the well known property, that the minimum of independent phase type random variables is also a phase type random variable. We give both an algebraic and a probabilistic proof, and illustrate these graphically.

A phase type (PH) random variable, with parameters \( \alpha' \in \mathbb{R}^n \) and \( T \in \mathbb{R}^{n \times n} \), such that \( \alpha_1 = 1 \), is a hitting time \( \tau := \inf\{t > 0 : X(t) = n + 1\} \), where \( X(t) \) is Markov process on \( \{1, \ldots, n+1\} \), with generator,

\[
Q = \begin{pmatrix}
T & \eta \\
0 & 0
\end{pmatrix}.
\]

Here \( \alpha \) is the initial distribution over the states \( \{1, \ldots, n\} \), \( \eta = -T1 \) and \( 1 \) is a column vector of ones of appropriate length. It has the CDF

\[
P(\tau \leq t) = 1 - \alpha e^{tT}1
\]

where \( e^{tT} \) is the matrix exponential.

We now have the following

**Theorem:** Let \( Z_i \) for \( i = 1, 2, \ldots, k \) be independent PH\((\alpha^i, T^i)\) random variables of order \( n_i \). Then \( Z := \min(Z_1, \ldots, Z_k) \sim PH(\alpha, T) \) of order \( \Pi_{i=1}^k n_i \) where

\[
\alpha = \alpha^1 \otimes \alpha^2 \otimes \ldots \otimes \alpha^k, \quad T = T^1 \oplus T^2 \oplus \ldots \oplus T^k,
\]

where for matrices, \( A \in \mathbb{R}^{r \times s} \) and \( B \in \mathbb{R}^{p \times q} \) the Kronecker product, \( \otimes \), and Kronecker sum, \( \oplus \), are respectively defined by:

\[
A \otimes B = \begin{bmatrix}
a_{11}B & \cdots & a_{1n}B \\
\vdots & \ddots & \vdots \\
a_{m1}B & \cdots & a_{mn}B
\end{bmatrix} \in \mathbb{R}^{r \times s \times q},
\]

\[
A \oplus B = A \otimes I_p + I_r \otimes B,
\]

with \( I_s \) denoting the \( s \times s \) identity matrix.

**Proof:**

First consider the case where \( k = 2 \). Let the corresponding Markov process of \( Z_1 \) and \( Z_2 \) be \( X_1 \) and \( X_2 \), then consider a Markov process \( X \) with state space \( S = \{(i,j) | 1 \leq i \leq m, 1 \leq j \leq n\} \cup \{0\} \) where 0 is the absorbing state. We denote the transition rates \( \lambda_{i,j} = T^1_{i,j} \) and \( \mu_{i,j} = T^2_{i,j} \). \( X \) is \( X_1 \) and \( X_2 \) running simultaneously, reaching 0 once one of the processes reaches its absorbing state.

Since the processes are independent, the probability that \( X(0) \) is in state \((i,j)\) is \( \alpha^1_i \cdot \alpha^2_j \). So if the states
are ordered as in the matrix in Figure 3, it is easy to see that $\alpha^1 \otimes \alpha^2$ gives the initial distribution over states.

The only possible transitions in $X$ are

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Rate</th>
<th>Induced by</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(i,j)$</td>
<td>$(h,j)$</td>
<td>$\lambda_{i,h}$</td>
<td>$X_1$</td>
</tr>
<tr>
<td>$(i,j)$</td>
<td>$(i,k)$</td>
<td>$\mu_{j,k}$</td>
<td>$X_2$</td>
</tr>
<tr>
<td>$(i,j)$</td>
<td>$(0)$</td>
<td>$\lambda_{i,0} + \mu_{j,0}$</td>
<td>$X_1$ or $X_2$</td>
</tr>
</tbody>
</table>

and so the total rate out of state $(i,j)$, denoted, $q_{i,j}$ is $\sum_{h \neq i} \lambda_{i,h} + \sum_{k \neq j} \mu_{j,k} + \lambda_{i,0} + \mu_{j,0}$. Denote the transition rate $(i,j) \to 0$ as $\eta_{i,j}$. The process can be visualized more generally as follows:

Writing this out in matrix form yields the generator matrix on page 3, from which it is easy to see why the Kronecker sum gives rise to the generator $T$ of the minimum phase type distribution (upper left sections of the matrix). More particularly notice that

$$
\eta_{i,j} = -T1 = q_{i,j} - \sum_{(i,j)} T_{(i,j)\to(h,k)} = q_i^1 + q_j^2 - \sum_{h \neq i} \lambda_{i,h} - \sum_{k \neq j} \mu_{j,k}
$$
\[
= \sum_{h \neq i} \lambda_{i,h} + \lambda_{i,0} + \sum_{k \neq j} \mu_{j,k} + \mu_{j,0} - \sum_{h \neq i} \lambda_{i,h} - \sum_{k \neq j} \mu_{j,k} = \lambda_{i,0} + \mu_{j,0}.
\]

\[
\begin{bmatrix}
(1,1) & (1,2) & \cdots & (1,n) & (2,1) & (2,2) & \cdots & (2,n) & \cdots & (m,1) & (m,2) & \cdots & (m,n) & (0) \\
-\eta_{1,1} & \mu_{1,2} & \cdots & \mu_{1,n} & \lambda_{1,2} & \lambda_{1,2} & \cdots & \lambda_{1,2} & \cdots & \lambda_{1,m} & \lambda_{1,m} & \cdots & \lambda_{1,m} & \eta_{1,1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\lambda_{2,1} & \lambda_{2,1} & \cdots & \lambda_{2,1} & \lambda_{2,2} & \lambda_{2,2} & \cdots & \lambda_{2,2} & \cdots & \lambda_{2,m} & \lambda_{2,m} & \cdots & \lambda_{2,m} & \eta_{2,1} \\
2 & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(1,1) & (1,2) & \cdots & (1,n) & (2,1) & (2,2) & \cdots & (2,n) & \cdots & (m,1) & (m,2) & \cdots & (m,n) & (0) \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\lambda_{m,1} & \lambda_{m,1} & \cdots & \lambda_{m,1} & \lambda_{m,2} & \lambda_{m,2} & \cdots & \lambda_{m,2} & \cdots & \lambda_{m,m} & \lambda_{m,m} & \cdots & \lambda_{m,m} & \eta_{m,1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

Figure 3: Generator of the minimum distribution

Before the algebraic proof for \( k = 2 \), we will show that
\[
e^{T_{11} \otimes e^{St}} = e^{(T \otimes S)t},
\]
where \( T \) is a square matrix of order \( m \) and \( S \) is a square matrix of order \( n \). Here the mixed product property of the Kronecker product \((AB) \otimes (CD) = (A \otimes C)(B \otimes D)\), and the left-distributive property \( A \otimes (B + C) = A \otimes B + A \otimes C \) will be used. By definition of the matrix exponential,
\[
e^{T_{11} \otimes e^{St}} = \sum_{r=0}^{\infty} \frac{(T)^r}{r!} \otimes \sum_{l=0}^{\infty} \frac{(S)^l}{l!}
= \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \frac{t^{r+l}T^r \otimes S^l}{r!l!}
= \sum_{v=0}^{\infty} \frac{t^v}{v!} \sum_{l=0}^{v} \frac{T^{v-l} \otimes S^l}{(v-l)!!}
= \sum_{v=0}^{\infty} \frac{t^v}{v!} \sum_{l=0}^{v} \frac{v!T^{v-l} \otimes S^l}{(v-l)!!}
= \sum_{v=0}^{\infty} \frac{t^v}{v!} \sum_{l=0}^{v} \binom{v}{l} (T^{v-l}I_m^l) \otimes (I_n^{v-l}S^l)
= \sum_{v=0}^{\infty} \frac{t^v}{v!} \sum_{l=0}^{v} \binom{v}{l} (T^{v-l} \otimes I_n^{v-l})(I_m^l \otimes S^l)
= \sum_{v=0}^{\infty} \frac{t^v}{v!} \sum_{l=0}^{v} \binom{v}{l} (T \otimes I_n)^{v-l}(I_m \otimes S)^l
\]
by the mixed product property

\[3\]
\[
= \sum_{v=0}^{\infty} \frac{t^v}{v!} (T \otimes I_n + I_m \otimes S)^v \quad \text{by the binomial theorem}
\]
\[
= e^{(T \otimes S)t} \quad \text{by definition.}
\]

The binomial theorem can be applied because the matrices \((T \otimes I_n)^{v-l}\) and \((I_m \otimes S)^l\) are commutative, noting that
\[
T^{v-l} \otimes S^l = (I_m^l T^{v-l}) \otimes (S^l I_n^{v-l})
\]
\[
= (I_m^l \otimes S^l)(T^{v-l} \otimes I_n^{v-l})
\]
\[
= (I_m \otimes S)^l (T \otimes I_n)^{v-l}
\]
\[
= (T \otimes I_n)^{v-l} (I_m \otimes S)^l.
\]

Completing the proof,
\[
\mathbb{P}(\min(Z_1, Z_2) > t) = \mathbb{P}(Z_1 > t) \mathbb{P}(Z_2 > t) \quad \text{by independence}
\]
\[
= (\alpha_1^1 e^{T_{1}^l} 1_m)(\alpha_2^{2} e^{T_{2}^l} 1_n) \quad \text{by (1)}
\]
\[
= (\alpha_1^l e^{T_{1}^t} 1_m) \otimes (\alpha_2^{2} e^{T_{2}^t} 1_n) \quad \text{1 \times 1 matrices}
\]
\[
= (\alpha_1 \otimes \alpha_2^2)(e^{T_{1}^t} \otimes e^{T_{2}^t})(1_m \otimes 1_n) \quad \text{mixed product property}
\]
\[
= (\alpha_1 \otimes \alpha_2^2)(e^{(T_{1}^t \oplus T_{2}^t)}) 1_{mn} \quad \text{by (2)}
\]

which, by (1), is a phase type distribution with parameters \(\alpha_1 \otimes \alpha_2^2\) and \(T_{1}^t \oplus T_{2}^t\).

![Double Summation Change of Variables](image)

Figure 4: Double Summation Change of Variables

To prove for all \(k \geq 2\), assume that the theorem is true for \(k = p\). So, \(\min(Z_1, Z_2, \ldots, Z_p) \sim PH(\alpha^{\bar{p}}, T^{\bar{p}})\) where \(\alpha^{\bar{p}} = \alpha_1 \otimes \alpha_2^2 \otimes \ldots \otimes \alpha_p\) and \(T^{\bar{p}} = T_{1}^t \oplus T_{2}^t \oplus \ldots \oplus T_{p}^t\). Now \(\min(Z_1, Z_2, \ldots, Z_p, Z_{p+1}) = \min(\min(Z_1, Z_2, \ldots, Z_p), Z_{p+1})\) which, given the assumption, is distributed \(PH(\alpha^{\bar{p}} \otimes \alpha^{p+1}, T^{\bar{p}} \oplus T^{p+1})\) by the same reasoning as either of the proofs for \(k = 2\). Q.E.D.
Example 1: Minimum of Two Generalised Erlang (Hypoexponential) Random Variables

A Generalised Erlang random variable is defined as $\sum_{k=1}^{m} X_k$ where the $X_k$ are independent exponential random variables with rate $\lambda_k$. So it is $PH$ distributed with parameters $\alpha = [1 \ 0 \ldots 0] \in \mathbb{R}^m$ and

$$T = \begin{bmatrix}
-\lambda_{1,2} & \lambda_{1,2} & 0 & 0 & 0 \\
0 & -\lambda_{2,3} & \lambda_{2,3} & 0 & 0 \\
0 & 0 & \ddots & \ddots & 0 \\
0 & 0 & 0 & -\lambda_{m-1,m} & \lambda_{m-1,m} \\
0 & 0 & 0 & 0 & -\lambda_{m,0}
\end{bmatrix}$$

and can be visualised as follows:

![Figure 5: Single Hypoexponential PH distribution](image)

So let the above Markov process be $X_1$, and let $X_2$ also be a Generalised Erlang of dimension $n$ and rates $\mu_k$. Then, the two processes running at the same time can be represented as in figure 5 below.

![Figure 6: State Space and Transitions for Minimum of Two Hypoexponentials](image)
**Example 2: Minimum of Two Hyper-exponential Random Variables**

A hyper-exponential random variable is a mixture of \( n \) exponential random variables \( X_i \) for \( i = 1, \ldots, n \), with rates \( \lambda_i \) and weights \( p_i \). It is a phase type random variable and in the case when \( n = 2 \), its parameters are

\[
\alpha = \begin{bmatrix} p & 1 - p \end{bmatrix}, \quad T = \begin{bmatrix} -\lambda_1 & 0 \\ 0 & -\lambda_2 \end{bmatrix}.
\]

Take another hyper-exponential random variable, with \( n = 2 \), weights \( q \) and \( 1 - q \) and rates \( \mu_i \). By applying the theorem, the parameters of the PH distribution of the minimum of the two random variables are

\[
\alpha = \alpha^1 \otimes \alpha^2 = \begin{bmatrix} pq & p(1 - q) & (1 - p)q & (1 - p)(1 - q) \end{bmatrix}
\]

\[
T = \begin{bmatrix} (1,1) & (1,2) & (2,1) & (2,2) \\ -(\lambda_1 + \mu_1) & -(\lambda_1 + \mu_2) & -(\lambda_2 + \mu_1) & -(\lambda_2 + \mu_2) \end{bmatrix}
\]

where all empty elements of the \( T \) matrix are 0. Note that this is also a hyper-exponential distribution.

![Hyper-exponential Representation of the Minimum](image)