Scand J Statist 5: 119–122, 1978

Transition Probabilities for a Truncated Birth–Death Process

STIG I. ROSENLUND

University of Göteborg

Received January 1977

ABSTRACT. By elementary real analysis a simple expression is found for the transition probability $p_{n,j}(t)$ in the homogeneous irreducible birth–death process on a finite state space. Some conditions are obtained for determining whether $p_{n,j}$ is increasing or not.

Key words: transition probability, truncated birth–death process

1. Definitions and results

We will study the birth–death process $N(t)$ on the state space $\{0, 1, ..., K\}$, where K is a positive integer. The rates of transition are λ_n for passage from state n to state $n+1$ and μ_n for passage from n to $n-1$. Here $\lambda_n > 0$ for $0 \leq n \leq K-1$, $\lambda_K = 0$, and $\mu_n > 0$ for $1 \leq n \leq K$, $\mu_0 = 0$. For $t \geq 0$ put

$$p_{n,j}(t) = P(N(u+t) = j \mid N(u) = n);$$

$$\pi_j = \lim_{t\to\infty} p_{n,j}(t) = \left(\sum_{r=0}^{K} \prod_{i=1}^{r} \mu_i^{-1}\lambda_{i-1}\right)^{-1} \prod_{i=1}^{j} \mu_i^{-1}\lambda_{i-1},$$

with the convention $\prod_{i=k}^{k-1} = 1$.

The transition probability will be expressed in terms of two sequences of polynomials. Define these recursively

$$P_0 = 1,\ P_1 = s + \lambda_0,$$

$$P_r = (s + \lambda_{r-1} + \mu_{r-1})P_{r-1} - \lambda_{r-2}\mu_{r-1}P_{r-2},\quad 2 \leq r \leq K+1;$$

$$Q_0 = 1,\ Q_1 = s + \mu_K,$$

$$Q_r = (s + \lambda_{K+1-r} + \mu_{K+1-r})Q_{r-1} - \lambda_{K+1-r}\mu_{K+2-r}Q_{r-2},\quad 2 \leq r \leq K+1.$$

Here P_r and Q_r are of degree r with coefficient 1 for s^r. Further we set $c_{n,n} = 1$ and

$$c_{n,j} = \begin{cases} \mu_{j+1} \dots \mu_n, & j < n \\ \lambda_n \dots \lambda_{j-1}, & j > n. \end{cases}$$

Proposition 1. *Let β_r be defined by $P_{K+1}(s) = Q_{K+1}(s) = s(s+\beta_1)\dots(s+\beta_K)$, $0 < \beta_1 < \dots < \beta_K$. It holds $P'_{K+1}(-\beta_r) = -\beta_r \prod_{i \neq r}(\beta_i - \beta_r)$. Put*

$$a_{n,j,r} = \frac{c_{n,K}c_{K,j}P_n(-\beta_r)P_j(-\beta_r)}{P_K(-\beta_r)P'_{K+1}(-\beta_r)} = \frac{c_{n,j}Q_{K-\max(n,j)}(-\beta_r)P_{\min(n,j)}(-\beta_r)}{P'_{K+1}(-\beta_r)}.$$

Then for $0 \leq n \leq K$ and $0 \leq j \leq K$ it holds

$$p_{n,j}(t) = \pi_j + \sum_{r=1}^{K} a_{n,j,r} \exp(-\beta_r t). \tag{1}$$

Note that $p_{n,j}/c_{n,j} = p_{j,n}/c_{j,n}$, so what is said about $p_{n,j}$ and $p_{m,r}$ in Propositions 3 and 4 will hold for $p_{j,n}$ and $p_{r,m}$ as well.

Proposition 2. *For $n = j$ all terms in (1) are non-negative and there are as many null terms as there are zeros of P_n equal to a zero of Q_{K-n} (at most $\min(n, K-n)$).*

Proposition 3. *Let $n < j$. If $p_{n,j}$ is increasing in t, then $p_{m,r}$ is increasing for $m \leq n$ and $r \geq j$. If $p_{n,j}$ is not increasing, then $p_{m,r}$ is not increasing for $n \leq m \leq j$ and $n \leq r \leq j$.*

If either n or j is a reflecting state (0 or K) more information is available. Define the two properties (i) the function is increasing and (ii) the function has exactly one local maximum after which it decreases to its limit. Then from Keilson (1971), p. 393, Cor. 1, we can infer that $p_{n,j}$ has either property (i) or (ii) when n or j is reflecting. We shall give criteria for deciding which.

Proposition 4. *If $P_n(-\beta_1) > 0$ then $p_{n,K}$ obeys* (i), *but if $P_n(-\beta_1) < 0$ then $p_{n,K}$ obeys* (ii). *If $Q_{K-j}(-\beta_1) > 0$ then $p_{0,j}$ obeys* (i), *but if $Q_{K-j}(-\beta_1) < 0$ then $p_{0,j}$ obeys* (ii). *In particular $p_{0,K}$ is increasing. If μ_K is sufficiently small $p_{n,K}$ is increasing for $n < K$ and if λ_0 is sufficiently small $p_{0,j}$ is increasing for $j > 0$.*

For any n, $0 \leqslant n \leqslant K$, it holds

$$P_{K+1} = (s + \lambda_n + \mu_n) Q_{K-n} P_n - \lambda_{n-1} \mu_n Q_{K-n} P_{n-1} - \lambda_n \mu_{n+1} Q_{K-n-1} P_n. \quad (2)$$

We have the relation

$$\beta_1 \cdots \beta_K = \mu_1 \cdots \mu_K \sum_{r=0}^{K} \prod_{i=1}^{r} \mu_i^{-1} \lambda_{i-1} = \mu_1 \cdots \mu_K / \pi_0. \quad (3)$$

As $t \to 0$ we have asymptotically

$$p_{n,j}(t) \sim c_{n,j} t^{|n-j|} / |n-j|!. \quad (4)$$

2. Relation to previous work on the subject

The model has been studied by Ledermann & Reuter (1954), Keilson (1964, 1965, 1971) and Rosenlund (1977). Our (1) corresponds to a spectral representation in Ledermann & Reuter (1954), eq. (1.79), modified by Keilson (1964), eq. (3.13). This is written in a different form, via the $K+1$ eigenvectors of an eigenvalue equation of order $K+1$. Keilson (1964) gives the Laplace transform (4.13) which is similar to our (9), but the structure of and relations between the polynomials $h_n^{(1)}$, $h_m^{(2)}$ and c^{-1} are not made clear. It seems that our procedure is shorter, more elementary and gives a theoretically and numerically simpler and for the truncated case more informative solution. It depends on the relation (6) in Rosenlund (1977). This result is due to Ledermann & Reuter (1954), Lemma 1, p. 328, but the direct application to the passage time problem, which is thereby simplified, seems to be due to the present author. In the work of Keilson (1964, 1965, 1971) the transition probability solution yields passage time results, whereas we go the other way, the advantage of which is demonstrated for example by the simple proof of Proposition 3. The first part of Proposition 2 follows from the spectral representations already known, otherwise our results appear to be new.

3. Proofs

Proof of Proposition 1, relations (2), (3) and (4)

The polynomials P_r were introduced in Rosenlund (1977) in dealing with passage times in the birth-death process on the non-negative integers. Letting $-\theta_{r,i}$ $(i=1, ..., r)$ be the zeros of P_r it was shown (relation (6)) that these are real and with proper ordering satisfy $\theta_{r,i} < \theta_{r-1,i} < \theta_{r,i+1}$ for $1 \leqslant i \leqslant r-1$ and $r \geqslant 2$. The proof holds as well for the process truncated at K for $2 \leqslant r \leqslant K+1$. It was further shown that $\theta_{r,i} > 0$ for $r \geqslant 1$, but for the present case this holds only for $1 \leqslant r \leqslant K$, because $\lambda_{r-1} = 0$ for $r = K+1$.

We have $P_r(0) = \lambda_0 \dots \lambda_{r-1}$ (see Rosenlund (1977), eq. (11)), hence $P_{K+1}(0) = 0$. Thus

$$0 = \theta_{K+1,1} < \theta_{K,1} < \theta_{K+1,2} < \dots < \theta_{K,K} < \theta_{K+1,K+1}. \quad (5)$$

Setting $\beta_r = \theta_{K+1,r+1}$ the representation of P_{K+1} in Proposition 1 follows.

The sequence Q_r has the same structure as the sequence P_r: the recurrence relation for P_r is satisfied by Q_r if λ_n and μ_n are replaced by $\lambda_n^* = \mu_{K-n}$ and $\mu_n^* = \lambda_{K-n}$, respectively (these are the birth and death rates for the process $K - N(t)$). In particular, with $-\varphi_{r,i}$ the zeros of Q_r, it holds $\varphi_{r,i} < \varphi_{r-1,i} < \varphi_{r,i+1}$.

We now take the forward Kolmogorov differential equation as the basis for our study. This reads

$$p'_{n,j} = -(\lambda_j + \mu_j) p_{n,j} + \lambda_{j-1} p_{n,j-1} + \mu_{j+1} p_{n,j+1}, \quad 0 \leqslant j \leqslant K, \quad (6)$$

with initial condition $p_{n,j}(0) = \delta_{n,j}$ (Kronecker's delta).

Letting for $s > 0$

$$\bar{p}_{n,j}(s) = \int_0^\infty e^{-st} p_{n,j}(t)\, dt,$$

Laplace transformation of (6) gives (Churchill (1958), sec. 4)

$$s\bar{p}_{n,j} - \delta_{n,j} = -(\lambda_j + \mu_j)\bar{p}_{n,j} + \lambda_{j-1}\bar{p}_{n,j-1} + \mu_{j+1}\bar{p}_{n,j+1}. \quad (7)$$

For each n this defines a linear system with $K+1$ equations, where we note that $\bar{p}_{n,-1} = \bar{p}_{n,K+1} = 0$. Define its $K+1$ by $K+1$ matrix $\mathbf{A}$ with element $a_{j,r}$ in row j and column $r (0 \leqslant j \leqslant K, 0 \leqslant r \leqslant K)$ defined by

$$a_{j,r} = \begin{cases} -\lambda_{j-1}, & r = j-1 \\ s + \lambda_j + \mu_j, & r = j \\ -\mu_{j+1}, & r = j+1 \\ 0, \text{ elsewhere.} \end{cases}$$

Then we write (7) in the form

$$\mathbf{A}[\bar{p}_{n,0}, ..., \bar{p}_{n,K}]^T = [\delta_{n,0}, ..., \delta_{n,K}]^T. \quad (8)$$

Denote by C_r the subdeterminant obtained from A by deleting rows and columns nos. $r, ..., K$. Set $C_0 = 1$ and $C_{K+1} = A$. Expanding C_r along its last row gives a recursive relation identical with the one defining P_r, hence $C_r = P_r$, $0 \leqslant r \leqslant K+1$. Cf. Ledermann & Reuter (1954), eqs. (1.17) and (1.18). Further denote by D_r the subdeterminant of A obtained by deleting rows and columns nos. $0, ..., K-r$. Set $D_0 = 1$ and $D_{K+1} = A$. Expanding D_r along its first row gives a recursive scheme identical with the one defining Q_r, hence $D_r = Q_r$, $0 \leq r \leq K+1$. It follows that $P_{K+1} = Q_{K+1} = A$. By (5) then $A(s) > 0$ and (8) has a unique solution for $s > 0$. Expanding A along row no. n gives the relation (2). The solution of (8) by Cramer's rule is written $\bar{p}_{n,j} = B_{n,j}/A$, where $B_{n,j}$ is obtained from A by replacing column no. j with $[\delta_{n,0}, \ldots, \delta_{n,K}]^T$. Thus $B_{n,j}$ is the cofactor of the element $a_{n,j}$, and this is easily found to be $\mu_{j+1} \ldots \mu_n C_j D_{K-n}$ for $n > j$, $\lambda_n \ldots \lambda_{j-1} C_n D_{K-j}$ for $n < j$, and $C_n D_{K-n}$ for $n = j$. Hence

$$\bar{p}_{n,j} = c_{n,j} Q_{K-\max(n,j)} P_{\min(n,j)} / P_{K+1}, \quad 0 \leq n \leq K, \quad 0 \leq j \leq K. \tag{9}$$

Inversion of (9) gives (Churchill (1958), sec. 20) the representation (1) with $a_{n,j,r}$ given by the second expression, noting that the term corresponding to $r = 0$ and $\beta_0 = 0$ is the limit of $p_{n,j}(t)$ as $t \to \infty$ and hence equal to π_j. Using this and setting $n = K$ and $j = 0$ in (9) gives relation (3). The Tauberian result (4) follows from (9) completely parallel to the derivation of (9) in Rosenlund (1977). The first expression for $a_{n,j,r}$ in Proposition 1 follows from equating the second one with the coefficients of (1.79) in Ledermann & Reuter (1954). In the present notation these are

$$a_{n,j,r} = P_n(-\beta_r) P_j(-\beta_r) / \Big(c_{0,n} c_{j,0} \sum_{i=0}^{K} P_i(-\beta_r)^2 / (c_{0,i} c_{i,0})\Big).$$

Proof of Proposition 2

First assume that P_n and Q_{K-n} have no zeros in common. Let $-\alpha_{n,r}$, where $0 < \alpha_{n,1} < \alpha_{n,2} < \ldots < \alpha_{n,K}$, be the zeros of $Q_{K-n} P_n$. We shall prove

$$\alpha_{n,1} < \beta_1 < \alpha_{n,2} < \ldots < \beta_{K-1} < \alpha_{n,K} < \beta_K. \tag{10}$$

Then the numerator as well as the denominator in the second expression for $a_{n,n,r}$ will be non-zero and alternate in sign as r goes from 1 to K, hence all terms will be positive. Now (10) is equivalent to

$$(-1)^r P_{K+1}(-\alpha_{n,r}) > 0, \quad 1 \leq r \leq K. \tag{11}$$

Let us for r given assume that $\alpha_{n,r} = \theta_{n,i}$ for some i. Then

$$\varphi_{K-n,r-i} < \theta_{n,i} < \varphi_{K-n,r-i+1},$$

hence

$$(-1)^{r-i} Q_{K-n}(-\theta_{n,i}) > 0.$$

From Rosenlund (1977), relation (6″), we have

$$(-1)^{i+1} P_{n-1}(-\theta_{n,i}) > 0.$$

By (2) we get

$$(-1)^r P_{K+1}(-\alpha_{n,r}) = \lambda_{n-1}\mu_n (-1)^{r-i} Q_{K-n}(-\theta_{n,i}) \times (-1)^{i+1} P_{n-1}(-\theta_{n,i}) > 0.$$

Similarly (11) follows when $\alpha_{n,r} = \varphi_{K-n,i}$ for some i.

For the case when P_n and Q_{K-n} have common zeros, we construct a process $N^a(t)$, where λ_{n-1} is replaced by $\lambda^a_{n-1} = \lambda_{n-1} + a$, $a > 0$, while all other transition rates are unchanged. Then $P^a_n(-\theta_{n,r}) = aP_{n-1}(-\theta_{n,r}) \neq 0$. When a is small enough P^a_n and $Q^a_{K-n} = Q_{K-n}$ will have no common zeros, since $\theta^a_{n,r}$ is continuous in a. Hence

$$Q_{K-n}(-\beta_r) P_n(-\beta_r) / P'_{K+1}(-\beta_r) = \lim_{a \to 0} Q^a_{K-n}(-\beta^a_r) P^a_n(-\beta^a_r) / P^{a\prime}_{K+1}(-\beta^a_r) \geq 0.$$

For this case (10) will hold with $<$ replaced by $\leq$.

The remainder of Proposition 2 follows from eq. (2), which shows that a zero of P_n is a zero of P_{K+1} if and only if it is also a zero of Q_{K-n}, and that a zero of Q_{K-n} is a zero of P_{K+1} if and only if it is also a zero of P_n.

An example where P_n has zeros in common with Q_{K-n} is the case $K = 2M$, $n = M$ and $\mu_r = \lambda_{K-r}$. Then $P_n = Q_n = Q_{K-n}$, and there will be only $K/2 + 1$ positive terms in the representation (1) for $p_{M,M}$.

Proof of Proposition 3

Here we use eq. (5) in Rosenlund (1977), which for $n < j$ gives $\sigma_{n,j} = c_{n,j} P_n / P_j$, where $\sigma_{n,j}$ is the Laplace transform of the density of a first passage time from n to j. Likewise $\sigma_{n,j} = c_{n,j} Q_{K-n} / Q_{K-j}$ for $n > j$; this follows from considering $K - N(t)$ and using the result above. Hence from (9) we have $s\bar{p}_{m,r} = (c_{j,r}/c_{r,j}) \sigma_{r,j} \sigma_{m,n} s\bar{p}_{n,j}$ for $m \leq n < j \leq r$ (let here $\sigma_{k,k} = 1$), and so the Laplace transform of $p'_{m,r}$ will be the Laplace transform of a probability density (except for a constant factor) if $p_{n,j}$ is increasing. The second part of the proposition is simply equivalent to the first part.

Proof of Proposition 4

Whether (i) or (ii) holds for $p_{n,K}$ depends on the sign of the derivative of the first term for $r = 1$ in (1); this term dominates for $t \to \infty$ if it is non-zero. Its derivative has the same sign as $P_n(-\beta_1)$. From (10) we can infer that $\beta_1 \leq \theta_{n,2}$ for each n (define $\theta_{n,n+1} = \infty$). Thus $P_n(-\beta_1) > 0$ is equivalent to $\beta_1 < \theta_{n,1}$ and $P_n(-\beta_1) < 0$ is equivalent to $\theta_{n,1} < \beta_1 < \theta_{n,2}$. If $\mu_K < \theta_{K-1,1}$, then $P_{K+1}(-\theta_{K-1,1}) = (\mu_K - \theta_{K-1,1}) \times P_K(-\theta_{K-1,1}) > 0$ by Rosenlund (1977), relation (6′), hence $\beta_1 < \theta_{K-1,1}$ since $P_{K+1} < 0$ in $(-\beta_1, 0)$. For $\mu_K = \theta_{K-1,1}$ property (i) for $p_{K-1,K}$ follows from a passage to the limit $\mu_K \uparrow \theta_{K-1,1}$, since the limit of a sequence of increasing functions is increasing. By Proposition 3 then $p_{n,K}$ is increasing for $n<K$ if $\mu_K \leqslant \theta_{K-1,1}$. The assertion on $p_{0,j}$ follows analogously (consider $K-N(t)$). Here $\lambda_0 \leqslant \varphi_{K-1,1}$ implies $p_{0,j}$ increasing for $j>0$. Note that we can verify directly that $s\bar{p}_{0,K} = c_{0,K}/((s+\beta_1)\ \dots\ (s+\beta_K))$, which is proportional to the transform of a convolution of K exponential densities, so that $p_{0,K}$ is increasing.

Remark. One might conjecture that either (i) or (ii) would hold generally for $p_{n,j}$. This is false. Counterexample: Take $K=3$, $\lambda_1=1.1$, $\lambda_0=\lambda_2=\mu_1=\mu_2=\mu_3=1$. Then exactly $\beta_1=0.6$, $\beta_2=2$, $\beta_3=3.5$, and $p_{2,1}(t)=(5/609)(29-8e^{-0.6t}+29e^{-2t}-50e^{-3.5t})$. This function has a local maximum approximately at 1.02 and a local minimum approximately at 1.51, after which it increases to its limit.

References

Churchill, R. V. (1958). *Operational mathematics*, 2nd ed. McGraw-Hill, New York.

Keilson, J. (1964). A review of transient behaviour in regular diffusion and birth–death processes. *J. Appl. Prob.* **1**, 247–266.

Keilson, J. (1965). A review of transient behaviour in regular diffusion and birth–death processes. Part II. *J. Appl. Prob.* **2**, 405–428.

Keilson, J. (1971). Log-concavity and log-convexity in passage time densities of diffusion and birth–death processes. *J. Appl. Prob.* **8**, 391–398.

Ledermann, W. & Reuter, G. E. H. (1954). Spectral theory for the differential equations of simple birth and death processes. *Philos. Trans. Roy. Soc. London Ser. A* **246**, 321–369.

Rosenlund, S. I. (1977). Upwards passage times in the non-negative birth–death process. *Scand. J. Statist.* **4**, 90–92.

Stig I. Rosenlund
Västmannagatan 93
S-113 43 Stockholm
Sweden