(a) Let $w_e$ and $w_u$ be the observed total waiting time in state $e$ respectively $u$ and let $\delta_e$ and $\delta_u$ be the observed total transitions out of state $e$ respectively $u$. Then we know from the lecture notes that the likelihood is given by

$$L(\sigma, \nu) = \sigma^{\delta_e} \eta^{\delta_u} e^{-\sigma w_e} e^{-\eta w_u}. $$

Hence the log-likelihood is given by

$$\ell(\sigma, \eta) = \log L(\sigma, \nu) = \delta_e \log \sigma - \sigma w_e + \delta_u \log \eta - \eta w_u. $$

The score function $\vec{U} = (U_\sigma, U_\eta)$ is

$$U_\sigma(\sigma) = \frac{\partial \ell}{\partial \sigma}(\sigma, \nu) = \frac{\delta_e}{\sigma} - w_e, $$

$$U_\eta(\eta) = \frac{\partial \ell}{\partial \eta}(\sigma, \nu) = \frac{\delta_u}{\eta} - w_u. $$

The unique root of $U_\sigma$ and $U_\eta$ are given by $\hat{\sigma} = \delta_e / w_e$ respectively $\hat{\eta} = \delta_u / w_u$. As $U'_\sigma(\hat{\sigma}) < 0$ and $U'_\eta(\hat{\eta}) < 0$, it follows that $\hat{\sigma}$ and $\hat{\eta}$ are (global) maximisers of $\ell_\sigma(\sigma) := \delta_e \log \sigma - \sigma w_e$ respectively $\ell_\eta(\eta) := \delta_u \log \eta - \eta w_u$. Hence as

$$\ell(\sigma, \eta) = \ell_\sigma(\sigma) + \ell_\eta(\eta), $$

it follows that $(\hat{\sigma}, \hat{\eta})$ is a maximiser of $\ell$ and thus $\hat{\sigma}$ is the mle of $\sigma$ and $\hat{\eta}$ is the mle of $\eta$.

(b) From the data, we get $\hat{\sigma} = 100 / 9600 = 1 / 96$ and $\hat{\eta} = 20 / 1200 = 1 / 60$. Denote by $\{X_t : t \geq 0\}$ the underlying MJP and by $R_0$ the residual holding time after 0. Recall that $p_{\mu}(t) := \Pr(X_u = i \forall u \in [0, t] | X_0 = i) = \Pr(R_0 > t | X_0 = i)$.

(i) We need to estimate the quantity $\Pr(R_0 > 1 | X_0 = e) = p_{\mu}(1)$. We know from the lecture notes that

$$\Pr(R_0 \geq 1 | X_0 = e) = \Pr(e_{\mu} \geq 1) = e^{-\sigma}, $$

where $e_{\mu}$ is an exponentially distributed random variable with parameter $\mu_e$. Hence an estimate of this probability is given by $e^{-\hat{\sigma}} = e^{-1 / 96} = 0.9896$.

(ii) We need to estimate the quantity $p_{ee}(1) = \Pr(X_1 = e | X_0 = e)$. We have calculated in the lecture notes (see Section 4.3) that

$$p_{ee}(1) = \frac{\sigma}{\sigma + \eta} e^{-(\sigma + \eta)t} + \frac{\eta}{\sigma + \eta}. $$

Hence an estimate of this probability is given by

$$\frac{\hat{\sigma}}{\hat{\sigma} + \hat{\eta}} e^{-(\hat{\sigma} + \hat{\eta})} + \frac{\hat{\eta}}{\hat{\sigma} + \hat{\eta}} = \frac{1 / 96}{1 / 96 + 1 / 60} e^{-1(1 / 96 + 1 / 60)} + \frac{1 / 60}{1 / 96 + 1 / 60} = 0.9897.
Answer to 2

(a) Let \( w_a \) be the observed total time the pair has been married, \( w_b \) be the observed total time the husband has lived while the wife is dead and \( w_c \) be the observed total time the wife has lived while the husband is dead. Further, define \( \delta_{ij} = 1 \) if a transition from state \( i \) to state \( j \) occurs (within the time interval under consideration), where \( i, j \in \{a, b, c, d\} \).

Then from the lecture notes, we know that the likelihood is given by

\[
L(\nu_1, \mu_1, \mu_2, \nu_2) = \nu_1^{\delta_{ab}} \mu_1^{\delta_{ac}} \mu_2^{\delta_{bd}} \nu_2^{\delta_{cd}} \exp(- (\mu_1 + \nu_1)w_a - \mu_2 w_b - \nu_2 w_c).
\]

Then, similarly as in Exercise 1, one can deduce that the mles of \( \nu_1, \nu_2, \mu_1, \mu_2 \) are given by

\[
\hat{\nu}_1 = \frac{\delta_{ab}}{w_a}, \quad \hat{\nu}_2 = \frac{\delta_{cd}}{w_c}, \quad \hat{\mu}_1 = \frac{\delta_{ac}}{w_a}, \quad \hat{\mu}_2 = \frac{\delta_{bd}}{w_b}.
\]

(b) There are now only two parameters to estimate since it is assumed that \( \mu_1 = \mu_2 \) and \( \nu_1 = \nu_2 \). The likelihood is now given by

\[
L(\nu_1, \mu_1) = \nu_1^{\delta_{ab} + \delta_{cd}} \mu_1^{\delta_{ac} + \delta_{ad}} \exp(- \mu_1(w_a + w_b) - \nu_1(w_a + w_c)).
\]

Proceeding as before, we get that in this case the maximum likelihood estimators are given by

\[
\hat{\nu}_1 = \hat{\nu}_2 = \frac{\delta_{ab} + \delta_{cd}}{w_a + w_c}, \quad \hat{\mu}_1 = \hat{\mu}_2 = \frac{\delta_{ac} + \delta_{ad}}{w_a + w_b}.
\]

Answer to 3

(a) With the usual meaning for \( w_a, \delta_{ab}, \delta_{ac} \) and \( \delta_{ad} \), we have \( \hat{\alpha} = \delta_{ab}/w_a = 180/750 = 0.24 \), \( \hat{\beta} = \delta_{ac}/w_a = 175/750 = 0.233 \) and \( \hat{\mu} = \delta_{ad}/w_a = 18/750 = 0.024 \).

(b) Since the mle \( \hat{\alpha} \) is asymptotically normal with mean \( \alpha \) and (asymptotic) variance given by the first diagonal element of the inverse of the observed Fisher information matrix, i.e.

\[
aVar(\hat{\alpha}) = \hat{\Gamma}_{11}^{-1} = \frac{\delta_{ab}}{w_a^2} = 0.00032,
\]

we have that a 95% asymptotic confidence interval for \( \alpha \) is given by

\[
\hat{\alpha} \pm z_{0.025} \times \sqrt{aVar(\hat{\alpha})} = [0.205, 0.275],
\]

where recall \( z_{0.025} = 1.960 \) is the number such that the area under the standard normal density function to the right of \( z(0.025) \) equals 0.025. Similarly, a 95% asymptotic confidence interval for \( \beta \) is given by

\[
\hat{\beta} \pm z_{0.025} \times \sqrt{\frac{\delta_{ac}}{w_a}} = [0.199, 0.268],
\]

and for \( \mu \) is given by

\[
\hat{\mu} \pm z_{0.025} \times \sqrt{\frac{\delta_{ad}}{w_a}} = [0.0129, 0.0351].
\]
(c) Denote by \( \{X_t : t \geq 0\} \) the underlying MJP and by \( R_0 \) the residual holding time after 0. As return to state \( a \) is impossible once you leave state \( a \), we have
\[
p_{aa}(1) = p_{aa}(1) = \Pr(e_{\mu a} > 1) = e^{-(\alpha + \beta + \mu)}.
\]
Hence an estimate for \( p_{aa}(1) \) is given by \( e^{-(\hat{\alpha} + \hat{\beta} + \hat{\mu})} \approx 0.608 \).

For \( p_{ab}(0.5) \), note that once you leave state \( a \) and end up in state \( b \) you remain in state \( b \) forever. Therefore \( p_{ab}(0.5) = \Pr(R_0 \leq 0.5, X_{R_0} = b|X_0 = a) \). Then by Theorem 5.6 or Theorem 5.7 in the notes,
\[
p_{ab}(0.5) = \frac{\alpha}{\alpha + \beta + \mu}(1 - e^{-0.5(\alpha + \beta + \mu)}).
\]
One can also use the forward equations instead to deduce above result. Hence an estimate of \( p_{ab}(0.5) \) is given by
\[
\frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta} + \hat{\mu}}(1 - e^{-0.5(\hat{\alpha} + \hat{\beta} + \hat{\mu})}) = 0.1062.
\]

**Answer to 4**

(a) We have \( W_j = \int_0^1 \mathbf{1}_{\{X_u = j\}} du \) and so since \( X_0 = i \),
\[
\mathbb{E}[W_j] = \mathbb{E} \left[ \int_0^1 \mathbf{1}_{\{X_u = j\}} du \bigg| X_0 = i \right]
= \int_0^1 \mathbb{E} \left[ \mathbf{1}_{\{X_u = j\}} \bigg| X_0 = i \right] du
= \int_0^1 \Pr(X_u = j|X_0 = i) du
= \int_0^1 p_{ij}(u) du,
\]
where the switching of the expectation and the integral can be justified by Fubini (cf. Appendix A.1 in the notes) and in the third equality we used that the expectation of the indicator of an event equals the probability of that event.

(b) We have \( \Delta_{jk}^{(m)} = \sum_{n=1}^m \mathbf{1}_{\{y_{n-1}^{(m)} = j, y_n^{(m)} = k\}} \), so that
\[
\mathbb{E}[\Delta_{jk}^{(m)}] = \mathbb{E} \left[ \sum_{n=1}^m \mathbf{1}_{\{y_{n-1}^{(m)} = j, y_n^{(m)} = k\}} \bigg| X_0 = i \right]
= \sum_{n=1}^m \mathbb{E} \left[ \mathbf{1}_{\{y_{n-1}^{(m)} = j, y_n^{(m)} = k\}} \bigg| X_0 = i \right]
= \sum_{n=1}^m \Pr(Y_n^{(m)} = j, Y_{n-1}^{(m)} = k|X_0 = i).
By the Markov property of $X$, we have that $Y^{(m)}$ is a discrete time Markov chain. The transition probabilities of $Y^{(m)}$ are given by $\Pr(Y_{n}^{(m)} = j | X_0 = i) = p_{ij}(n/m)$ and so by Theorem 2.2 in the notes,

$$
\Pr(Y_{n-1}^{(m)} = j, Y_{n}^{(m)} = k | X_0 = i) = p_{ij} \left( \frac{n-1}{m} \right) p_{jk}(1/m).
$$

We end up with

$$
\mathbb{E}[\Delta^{(m)}_{jk}] = \sum_{n=1}^{m} p_{ij} \left( \frac{n-1}{m} \right) p_{jk}(1/m).
$$

(c) Since the sample paths of the MJP are right-continuous we have that $\Delta_{jk} = \lim_{m \to \infty} \Delta^{(m)}_{jk}$ (with probability one) and so

$$
\mathbb{E}[\Delta_{jk}] = \mathbb{E} \left[ \lim_{m \to \infty} \Delta^{(m)}_{jk} \right] = \lim_{m \to \infty} \mathbb{E} \left[ \Delta^{(m)}_{jk} \right] = \lim_{m \to \infty} \sum_{n=1}^{m} p_{ij} \left( \frac{n-1}{m} \right) p_{jk}(1/m)
$$

$$
= \left( \lim_{m \to \infty} \frac{p_{jk}(1/m)}{1/m} \right) \left( \lim_{m \to \infty} \sum_{n=1}^{m} p_{ij} \left( \frac{n-1}{m} \right) \ast 1/m \right),
$$

where the switching of the limit and expectation in the second equality can be justified (in a non-obvious way) by the dominated convergence theorem. From Equation (4.4) in the notes we have that $\lim_{m \to \infty} \frac{p_{jk}(1/m)}{1/m} = \mu_{jk}$ for $j \neq k$. Further we recognise $\lim_{m \to \infty} \sum_{n=1}^{m} p_{ij} \left( \frac{n-1}{m} \right) \ast 1/m$ as a limit of Riemann sums corresponding to the function $p_{ij}(\cdot)$ over the interval $[0, 1]$ and thus this limit equals the (Riemann) integral of $p_{ij}(\cdot)$ over the interval $[0, 1]$, i.e.

$$
\lim_{m \to \infty} \sum_{n=1}^{m} p_{ij} \left( \frac{n-1}{m} \right) \ast 1/m = \int_{0}^{1} p_{ij}(u)du.
$$

We conclude that $\mathbb{E}[\Delta_{jk}] = \mu_{jk} \int_{0}^{1} p_{ij}(u)du$. 

4