Supplementary Material

In this section we will prove that under a special case of the null hypothesis the test statistic (5) with scoring function $S_i(t) = \pm 1$ given in (7) has an asymptotic normal distribution. To show this we will use several theorems and definitions from Fleming and Harrington [10]. Let $\lambda_i(t)$ be the hazard rate for species i at time t. We shall assume a strong version of our null hypothesis given by

$$H_0 : \lambda_i(t) = \lambda(t) \text{ for } i = 1, ..., n \text{ and } \forall t \in R^+ \tag{22}$$

This amounts to assuming that there are no age effects (the Red Queen hypothesis) or covariate effects on species extinction. I will prove that:

$$\frac{J}{\sqrt{V}} \xrightarrow{D} N(0,1) \tag{23}$$

This will be shown by applying a martingale central limit theorem to statistic J [10]. Using the univariate case of theorem 5.3.4 of Fleming and Harrington we have the following central limit theorem:

Theorem 6.1 *Let W be a Brownian motion process and f be a measurable nonnegative function such that $\alpha(t) = \int_0^t f^2(s)ds < \infty$, $\forall t > 0$.*
Suppose,

(1) $\{N_i(t) : i = 1, ..., n\}$ is a counting process with stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t > 0\}, P)$
(2) The compensator $A_i(t)$ of $N_i(t)$ is continuous.
(3) $H_i(t)$ is a locally bounded $\mathcal{F}_t$-predictable process.

Define

$$M_i(t) \equiv N_i(t) - A_i(t), \tag{24}$$

$$U_i(t) = \int_0^t H_i(s)dM_i(s),\ U(t) = \sum_{i=1}^{n} U_i(t), \tag{25}$$

and for any $\epsilon > 0$

$$U_{i,\epsilon}(t) = \int_0^t H_i(s)I_{\{|H_i(s)| \geq \epsilon\}}dM_i(s),$$

$$U_\epsilon(t) = \sum_{i=1}^{n} U_{i,\epsilon}(t).$$

Assume for any $t \in [0, \eta]$ as $n \to \infty$

i. $\langle U, U \rangle (t) \xrightarrow{p} \int_0^t f^2(s)ds$
and
ii. $\langle U_\epsilon, U_\epsilon \rangle (t) \xrightarrow{p} 0$ *for any $\epsilon > 0$*

Then $U \xrightarrow{D} \int f dW$ in $D[0, \eta]$ as $n \to \infty$.

Proof of (23):
Before showing that *i.* and *ii.* hold, I will show conditions 1-3 for the situation of our statistic J:

(1) Show $N_i(t)$ is a counting process.
The relevant definitions are:

Definition 6.2 *A counting process is a stochastic process* $\{N(t) : t > 0\}$ *adapted to a filtration* $\{\mathcal{F}_t : t > 0\}$ *with* $N(0) = 0$ *and* $N(t) < \infty$ *a.s. and whose paths are with probability one right-continuous, piecewise constant, and have only jump discontinuities, with jumps of size* $+1$.

Definition 6.3 *A stochastic process* $\{X(t) : t \geq 0\}$ *is adapted to a filtration if, for every* $t \geq 0$, $X(t)$ *is* $\mathcal{F}_t$*-measurable, i.e.,* $\{\omega : X(t, \omega) \leq x\} \in \mathcal{F}_t$.

In our situation

$$N_i(t) = I_{\{R_i \leq t\}} = \begin{cases} 0, \, t < R_i \\ 1, \, t \geq R_i \end{cases} \tag{26}$$

where $R_i > 0$. This clearly satisfies $N(0) = 0$ and $N(t) < \infty$, and has paths which are right-continuous and piecewise constant, with a single jump of size 1. The filtration we use throughout this work is

$$\mathcal{F}_t = \sigma\left\{N_i(u), N_i^L(u) : 0 \leq u \leq t\right\}, \text{ where } N_i^L(u) = I_{\{L_i \leq u\}}. \tag{27}$$

Since $N_i(t) \in \left\{N_i(u), N_i^L(u) : 0 \leq u \leq t\right\}$, it is clear that $N_i(t)$ is $\mathcal{F}_t$-measurable.

(2) Now I will show that the continuous compensator of $N_i(t)$ is $A_i(t) \equiv \int_\gamma^t Y_i(u)\lambda_i(u)du$, that is $A_i(t)$ is increasing, continuous and $\mathcal{F}_t$-predictable and $M_i(t) = N_i(t) - A_i(t)$ is an $\mathcal{F}_t$-martingale, where

$$\lambda_i(t) = \lim_{h \to 0} \frac{1}{h} P(t < R_i < t + h | L_i < t < R_i). \tag{28}$$

To show this I will need to establish the following:
a) $A_i(t)$ is continuous, increasing, and $\mathcal{F}_t$-predictable.
b) $M_i(t)$ is adapted to $\mathcal{F}_t$.
c) $E|M_i(t)| < \infty$
d) $E(M_i(t+s)|\mathcal{F}_t) = M_i(t)$

a) Show $A_i(t)$ is increasing, continuous and $\mathcal{F}_t$-predictable.
$A_i(t)$ is continuous and increasing since it is the cumulative integral of a non-negative integrand. Define

$$\Lambda_i(t) = \int_\gamma^t \lambda_i(u)\, du\,.$$

$A_i(t)$ is adapted since

$$A_i(t) = \int_\gamma^t I(L_i < u \leq R_i)\lambda_i(u)\, du = \Lambda_i(t \wedge R_i) - \Lambda_i(t \wedge L_i)$$

and both $t \wedge R_i$ and $t \wedge L_i$ are easily seen to be adapted. Now, since $A_i(t)$ is continuous and adapted, we conclude (from Lemma 1.4.1 of Fleming and Harrington) that $A_i(t)$ is predictable.

b) Show $M_i(t)$ is adapted to $\mathcal{F}_t$.
It suffices to show that $N_i(t)$ and $A_i(t)$ are adapted to $\mathcal{F}_t$, but these were shown earlier.

c) Show $E|M_i(t)| < \infty$.

$$
\begin{aligned}
E|M_i(t)| &\leq E(N_i(t)) + E\int_{\gamma}^{t} Y_i(u)\lambda_i(u)du \\
&\leq 1 + \int_{\gamma}^{t} P(L_i < u \leq R_i)\lambda_i(u)du \\
&\leq 1 + \int_{\gamma}^{t} \lambda_i(u)du \\
&< \infty
\end{aligned}
$$

d) Show $E(M_i(t+s)|\mathcal{F}_t) = M_i(t)$ a.s. $\forall s, t \geq 0$.

$$
\begin{aligned}
E(M_i(t+s)|\mathcal{F}_t) &= E\left\{N_i(t+s) - \int_{\gamma}^{t+s} Y_i(u)\lambda_i(u)du|\mathcal{F}_t\right\} \\
&= N_i(t) - \int_{\gamma}^{t} Y_i(u)\lambda_i(u)du + E\left\{N_i(t+s) - N_i(t)|\mathcal{F}_t\right\} \\
&\quad - E\left\{\int_{t}^{t+s} Y_i(u)\lambda_i(u)du|\mathcal{F}_t\right\} \\
&= M_i(t) + E\left\{N_i(t+s) - N_i(t)|\mathcal{F}_t\right\} - E\left\{\int_{t}^{t+s} Y_i(u)\lambda_i(u)du|\mathcal{F}_t\right\}
\end{aligned}
$$

Thus it suffices to show that

$$
E\left\{N_i(t+s) - N_i(t)|\mathcal{F}_t\right\} = E\left\{\int_{t}^{t+s} Y_i(u)\lambda_i(u)du|\mathcal{F}_t\right\} \tag{29}
$$

For the remainder of the proof we will suppress the use of i in the notation. Let us start by noting that $\mathcal{F}_t = \sigma\left\{L^{(t)}, R^{(t)}\right\}$ where

$$
L^{(t)} = \begin{cases} L, \; L \leq t \\ \infty, \; L > t \end{cases}
$$

$$
R^{(t)} = \begin{cases} R, \; R \leq t \\ \infty, \; R > t \end{cases}
$$

Removing the i notation from the left hand side of the equation (29) we have:

$$
\begin{aligned}
E(N(t+s) - N(t)|\mathcal{F}_t) &= P(t < R \leq t+s|\mathcal{F}_t) \\
&= E(I_{\{t<R\leq t+s\}}|\mathcal{F}_t) \\
&= \begin{cases} 0, & \text{if } R \leq t \\ P(t < R \leq t+s|L > t), & \text{if } L > t \\ 1 - e^{-\int_t^{t+s}\lambda(u)du}, & \text{if } L \leq t < R \end{cases}
\end{aligned} \tag{30}
$$

When $L > t$ equation (30) follows from $\{\omega : L(\omega) > t\} = \left\{\omega : L^{(t)}(\omega) = R^{(t)}(\omega) = \infty\right\}$.
Now we will evaluate the right hand side of equation (29) and show that it is equal to the left.

$$E\left(\int_t^{t+s} Y(u)\lambda(u)du|\mathcal{F}_t\right) = \int_t^{t+s} E(Y(u)|\mathcal{F}_t)\lambda(u)du \tag{31}$$

For $t < u < t+s$, let's consider equation (31) on the 3 sets: $\{R \leq t\}$, $\{L > t\}$, and $\{L \leq t < R\}$. By definition we know that

$$E(Y(u)|\mathcal{F}_t) = P(L < u \leq R|\mathcal{F}_t).$$

For $\omega \in \{R \leq t\}$, $P(L < u \leq R|\mathcal{F}_t) = 0$. Thus,

$$\int_t^{t+s} E(Y(u)|\mathcal{F}_t)\lambda(u)du = 0 \tag{32}$$

on $\{R \leq t\}$. For $\omega \in \{L \leq t < R\}$,

$$P(L < u \leq R|\mathcal{F}_t) = P(R \geq u|L, R > t) = e^{-\int_t^u \lambda(z)dz}.$$

Thus, if $L \leq t < R$,

$$\begin{aligned}\int_t^{t+s} E(Y(u)|\mathcal{F}_t)\lambda(u)du &= \int_t^{t+s} e^{-\int_t^u \lambda(z)dz}\lambda(u)du \\ &= -e^{-\int_t^u \lambda(z)dz}\Big|_{u=t}^{u=t+s} \\ &= 1 - e^{-\int_t^{t+s} \lambda(z)dz} \end{aligned} \tag{33}$$

For $\omega \in \{L > t\}$,

$$\begin{aligned}E(Y(u)|\mathcal{F}_t) &= P(L < u \leq R|L > t) = P(t < L < u \leq R|L > t) \\ &= \frac{P(t < L < u \leq R)}{P(L > t)} \\ &= \frac{\int_t^u P(R \geq u|L = z)dF_L(z)}{P(L > t)} \\ &= \frac{\int_t^u e^{-\int_z^u \lambda(w)dw}dF_L(z)}{P(L > t)} \end{aligned} \tag{34}$$

Thus by equation (34) we have,

$$\begin{aligned}
\int_t^{t+s} E(Y(u)|\mathcal{F}_t)\lambda(u)du &= \int_t^{t+s} \lambda(u) \left(\frac{\int_t^u e^{-\int_z^u \lambda(w)dw} dF_L(z)}{P(L>t)} \right) du \\
&= \frac{1}{P(L>t)} \int_t^{t+s} dF_L(z) \int_z^{t+s} e^{-\int_z^u \lambda(w)dw} \lambda(u)du \\
&= \frac{\int_t^{t+s} dF_L(z)}{P(L>t)} \left(-e^{-\int_z^u \lambda(w)dw} |_z^{t+s} \right) \\
&= \frac{\int_t^{t+s} dF_L(z)}{P(L>t)} \left(1 - e^{-\int_z^{t+s} \lambda(w)dw} \right) \\
&= \frac{P(R \le t+s, L>t)}{P(L>t)} \\
&= P(R \le t+s | L>t) \\
&= P(t < R \le t+s | L>t) \qquad (35)
\end{aligned}$$

Now by equations (31), (32), (33), and (35) we have,

$$\int_t^{t+s} E(Y(u)|\mathcal{F}_t)\lambda(u)du = \begin{cases} 0, & \text{if } R \le t \\ P(t < R \le t+s | L > t), & \text{if } L > t \\ 1 - e^{-\int_t^{t+s} \lambda(z)dz}, & \text{if } L \le t < R \end{cases} \qquad (36)$$

This is equivalent to equation (30).

(3) Show that $H_i(t) = S_i(t)Y_i(t)$ is locally bounded.
We will show this for the particular score function $S_i(t)$ which takes the values $+1$ or -1 (7). This score function is actually bounded globally:

$$|H_i(t)| = |S_i(t)Y_i(t)| \le Y_i(t) \le 1 \ \forall t \ge 0 \qquad (37)$$

i. Show $\langle J, J \rangle (t) \xrightarrow{p} \int_0^t f^2(s)ds$.

Let's first break down Statistic J.

$$
\begin{aligned}
J(t) &\equiv \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dN_i(u) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u) + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) d\Lambda_i(u) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u) + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) \lambda_i(u) du \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u) + \frac{1}{\sqrt{n}} \int_{\gamma}^{t} \sum_{i=1}^{n} S_i(u) Y_i(u) \lambda_i(u) du \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u) + \frac{1}{\sqrt{n}} \int_{\gamma}^{t} \sum_{i=1}^{n} S_i(u) Y_i(u) \lambda(u) du \text{ (by } H_0) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u) \qquad (38)
\end{aligned}
$$

$$
\begin{aligned}
\langle J, J \rangle (t) &= \left\langle \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u), \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u) \right\rangle \\
&= \frac{1}{n} \sum_{i=1}^{n} \left\langle \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u), \int_{\gamma}^{t} S_i(u) Y_i(u) dM_i(u) \right\rangle \\
&= \frac{1}{n} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i^2(u) Y_i(u) d \langle M_i, M_i \rangle (u) \\
&= \frac{1}{n} \sum_{i=1}^{n} \int_{\gamma}^{t} S_i^2(u) Y_i(u) d\Lambda_i(u) \\
&= \frac{1}{n} \sum_{i=1}^{n} \int_{\gamma}^{t} Y_i(u) d\Lambda(u) \text{ by } (7) \\
&= \int_{\gamma}^{t} \frac{1}{n} \sum_{i=1}^{n} Y_i(u) \lambda(u) du
\end{aligned}
$$

Now we will show a.s. uniform convergence of $\frac{1}{n}\sum_{i=1}^{n} Y_i(u)$.

$$
\frac{1}{n} \sum_{i=1}^{n} Y_i(u) = \frac{1}{n} \sum_{i=1}^{n} I_{\{R_i \geq u\}} - \frac{1}{n} \sum_{i=1}^{n} I_{\{L_i \geq u\}} \overset{a.s.}{\rightarrow}_{uniformly} S_R(u) - S_L(u)
$$

as $n \to \infty$ by the Glivenko-Cantelli Theorem. Here we are assuming the pairs (L_i, R_i), $i = 1, ..., n$ are i.i.d. and denote the survival functions of L_i and R_i

by S_L and S_R, respectively. This implies that

$$\lambda(u)\left(\frac{1}{n}\sum_{i=1}^{n} Y_i(u)\right) \stackrel{a.s.}{\rightarrow}_{uniformly} \lambda(u)\left(S_R(u) - S_L(u)\right) \text{ as } n \rightarrow \infty. \tag{39}$$

Since

$$\sup_u \left| \left(\frac{1}{n}\sum_{i=1}^{n} Y_i(u)\right)\lambda(u) - \left(S_R(u) - S_L(u)\right)\lambda(u)\right| \stackrel{a.s.}{\rightarrow} 0,$$

integration leads to

$$\langle J, J\rangle(t) \stackrel{a.s.}{\rightarrow} \int_{\gamma}^{t} f^2(u)du$$

for all t where $f^2(u) = (S_{R(u)} - S_{L(u)})\lambda(u)$.

$$J_{\epsilon}(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_{\gamma}^{t} S_i(u)Y_i(u)I_{\left\{|S_i(u)Y_i(u)\frac{1}{\sqrt{n}}|>\epsilon\right\}}dM_i(u)$$

$$\begin{aligned}
\langle J_{\epsilon}, J_{\epsilon}\rangle(t) &= \frac{1}{n}\sum_{i=1}^{n}\int_{\gamma}^{t} S_i^2(u)I_{\left\{|S_i(u)Y_i(u)\frac{1}{\sqrt{n}}|>\epsilon\right\}}d\Lambda(u) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_{\gamma}^{t} I_{\left\{|\frac{Y_i(u)}{\sqrt{n}}|>\epsilon\right\}}d\Lambda(u) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_{\gamma}^{t} I_{\left\{Y_i(u)>\sqrt{n}\epsilon\right\}}d\Lambda(u) \\
&= 0 \text{ for } n > 1/\epsilon^2.
\end{aligned}$$