HYPOTHESIS TEST ON A MIXTURE FORWARD–INCUBATION-TIME EPIDEMIC MODEL WITH APPLICATION TO COVID-19 OUTBREAK

Chunlin Wang[1], Pengfei Li[2], Yukun Liu[3], Xiao-Hua Zhou[4] and Jing Qin[5]

[1]*Xiamen University,* [2]*University of Waterloo,* [3]*East China Normal University,* [4]*Peking University and* [5]*National Institutes of Health*

Supplementary Material

This is a supplement to the corresponding paper submitted to *Statistica Sinica*. It contains seven sections. In Section S1, we provide a derivation of the form of $g(t)$ that appears in Section 1 of the main paper. In Section S2, we conduct a goodness-of-fit test for model (1.2) in the main paper. In Section S3, we provide regularity conditions needed in Theorems 2–4. Finally, in Sections S4–S7, we prove Theorems 1–4 respectively.

S1 Derivation of the Form of $g(t)$

To write the probability density function (pdf) of V in the form of $g(t)$, we refer to Chapter 2 of Qin (2017) using renewal process results. Here, we understand V in the way of Linton et al. (2020).

Let A be the elapsed time between exposure to the disease and departure from Wuhan. Recall that we use Y for the incubation time, i.e., from infection onset to symptom onset. We assume A and Y are independent.

By the data-collection criteria, only those individuals with $Y > A$ are included in our cohort. Moreover, we cannot observe A, but only $V = Y - A$.

Basically, we have truncated data $(A, V)|Y > A$, so the pdf of the observed V should be conditional on $Y > A$.

Consider the conditional cumulative distribution function (cdf) of V given $Y > A$:

$$P(V \le t|Y > A) = \frac{P(V \le t, Y > A)}{P(Y > A)}.$$

For the time being, we assume that A follows a uniform distribution on $[0, c]$ for some positive constant c. We further assume that Y has cdf $F(t)$ and pdf $f(t)$. Then, conditioning on A, we have

$$\begin{aligned} P(V \le t, Y > A) &= \int_0^c \frac{1}{c} P(V \le t, Y > a|A = a)da \\ &= \int_0^c \frac{1}{c} P(Y - A \le t, Y > a|A = a)da \\ &= \int_0^c \frac{1}{c} P(Y \le t + a, Y > a|A = a)da \\ &= \int_0^c \frac{1}{c} P(a < Y \le t + a)da \\ &= c^{-1} \int_0^c \{F(t + a) - F(a)\}da, \end{aligned}$$

where the second-last step follows from the assumption that A and Y are independent. Similarly,

$$\begin{aligned} P(Y > A) &= \int_0^c \frac{1}{c} P(Y > a|A = a)da \\ &= \int_0^c \frac{1}{c} P(Y > a)da \\ &= c^{-1} \int_0^c \{1 - F(a)\}da. \end{aligned}$$

As a consequence,

$$P(V \le t|Y > A) = \frac{\int_0^c \{F(t + a) - F(a)\}da}{\int_0^c \{1 - F(a)\}da}.$$

Hence, the pdf of V conditional on $Y > A$ is

$$\frac{\int_0^c f(t+a)da}{\int_0^c \{1-F(a)\}da}.$$

Letting $c \to \infty$, $\int_0^c f(t+a)da = F(t+c) - F(t) \to 1 - F(t)$ and the limit of the pdf of V conditional on $Y > A$ tends to

$$\frac{1-F(t)}{\int_0^\infty \{1-F(a)\}da}, \tag{S1.1}$$

which is precisely the forward time distribution in the renewal process when reaching equilibrium status. In other words, the pdf of the observed V is approximately equal to (S1.1).

As

$$\int_0^\infty \{1-F(a)\}da = \int_0^\infty af(a)da,$$

the pdf of V conditional on $Y > A$ is given approximately by

$$g(t) = \frac{\int_t^\infty f(a)da}{\int_0^\infty af(a)da} \quad \text{for } t > 0.$$

The above derivation holds true even if Y is bounded. The uniform distribution assumption on A is reasonable because there seemed to be no general trend in the number of people departing Wuhan per day in the early stage of the outbreak.

Regarding the equilibrium assumption in our real data, 1211 cases were collected as of February 15, 2020, and their travel dates of leaving Wuhan were between January 19 and January 23. This enabled us to have an average follow-up time for symptom onset of as long as 25 days. With an adequate long run, the renewal process would reach the equilibrium status. In summary, the forward time distribution $g(t)$ in the renewal process is a good approximation to the truncation distribution of V.

S2 Goodness-of-fit Test of Model (1.2)

In this section, we review the goodness-of-fit test of Deng et al. (2021) for model (1.2) in the main paper, then we apply it to check whether model (1.2) is suitable for the data analyzed in Section 5 of the main paper.

Recall that model (1.2) posited that $t_1, \ldots, t_n$ are n iid observations from

$$h(t; \lambda, \alpha, p) = pf(t; \lambda, \alpha) + (1-p)g(t; \lambda, \alpha), \quad t > 0,$$

with $f(t; \lambda, \alpha)$ being the pdf of a pre-specified distribution such as a Weibull distribution, and

$$g(t; \lambda, \alpha) = \frac{\int_t^\infty f(y; \lambda, \alpha) dy}{\int_0^\infty y f(y; \lambda, \alpha) dy}.$$

The idea of the goodness-of-fit test of Deng et al. (2021) is to divide the non-negative real line into k disjoint and adjacent intervals, whereupon the goodness-of-fit statistic is defined as

$$G_n = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i},$$

where O_i is the observed number of cases in the ith interval, E_i is the expected number of cases in the ith interval based on $h(t; \hat{\lambda}, \hat{\alpha}, \hat{p})$, and k is chosen such that $E_i \geq 5$ for each interval. The asymptotic null distribution of G_n is known to be a chi-squared distribution with $k - 3 - 1$ degrees of freedom because there are three parameters in total in model (1.2).

For the data in Section 5 of the main paper, Deng et al. (2021) first partitioned the non-negative real line into $k = 17$ intervals: [0,0.5), $[i - 0.5, i + 0.5)$ for $i = 1, \ldots, 15$, and $[15.5, \infty)$. When $f(t; \lambda, \alpha)$ is the pdf of a Weibull distribution, the observed value of G_n is 14.09 with an asymptotic p-value of 0.37, calibrated by the χ^2_{13} distribution. Hence, we do not have strong evidence for rejecting model (1.2) with $f(t;\lambda,\alpha)$ being the pdf of a Weibull distribution for the duration-time data in Section 5 of the main paper.

S3 Regularity Conditions

Our asymptotic results about R_n in Theorems 2–4 rely on the following regularity conditions, in which the expectation is taken with respect to the null model.

C1. (i) For sufficiently small $\epsilon > 0$, $\mathbb{E}[\log\{1 + f_\epsilon(T)\}] < \infty$ and $\mathbb{E}[\log\{1 + g_\epsilon(T)\}] < \infty$, where

$$f_\epsilon(t) = \sup_{(\lambda-\lambda_0)^2+(\alpha-\alpha_0)^2<\epsilon^2} f(t;\lambda,\alpha)$$

and $g_\epsilon(t)$ is similarly defined; (ii) for sufficiently large $r > 0$, $\mathbb{E}[\log\{1+\varphi_{f,r}(T)\}] < \infty$ and $\mathbb{E}[\log\{1 + \varphi_{g,r}(T)\}] < \infty$, where

$$\varphi_{f,r}(t) = \sup_{\lambda^2+\alpha^2\geq r^2} f(t;\lambda,\alpha)$$

and $\varphi_{g,r}(t)$ is similarly defined; (iii) $f(t;\lambda,\alpha) \to 0$ and $g(t;\lambda,\alpha) \to 0$ as $\lambda^2 + \alpha^2 \to \infty$.

C2. The parameters λ and α are identifiable.

C3. $f(t;\lambda,\alpha)$ has common support and continuous third-order partial derivatives with respect to λ and α.

C4. $\mathbf{B}$ is positive definite.

C5. For two non-negative integers h and l such that $h + l \leq 2$, there exists a function $G(t)$ with $\mathbb{E}\{G(T)\} < \infty$ such that

$$\left|\frac{\partial^{h+l} f(t;\lambda_0,\alpha_0)/\partial\lambda^h\partial\alpha^l}{f(t;\lambda_0,\alpha_0)}\right|^3 \leq G(t) \text{ and } \left|\frac{\partial^{h+l} g(t;\lambda_0,\alpha_0)/\partial\lambda^h\partial\alpha^l}{g(t;\lambda_0,\alpha_0)}\right|^3 \leq G(t).$$

Moreover, there exists a positive ϵ_0 such that for $h+l=3$,

$$\sup_{(\lambda-\lambda_0)^2+(\alpha-\alpha_0)^2\le\epsilon_0^2}\left|\frac{\partial^{h+l}f(t;\lambda,\alpha)/\partial\lambda^h\partial\alpha^l}{f(t;\lambda_0,\alpha_0)}\right|^3\le G(t)$$

and

$$\sup_{(\lambda-\lambda_0)^2+(\alpha-\alpha_0)^2\le\epsilon_0^2}\left|\frac{\partial^{h+l}g(t;\lambda,\alpha)/\partial\lambda^h\partial\alpha^l}{g(t;\lambda_0,\alpha_0)}\right|^3\le G(t).$$

S4 Proof of Theorem 1

Recall that $F(t;\lambda,\alpha)$ is the cumulative distribution function corresponding to $f(t;\lambda,\alpha)$ and

$$h(t;\lambda,\alpha,p)=pf(t;\lambda,\alpha)+(1-p)g(t;\lambda,\alpha),\quad t>0,$$

where

$$g(t;\lambda,\alpha)=\frac{1-F(t;\lambda,\alpha)}{\mu(\lambda,\alpha)}\quad\text{with}\quad\mu(\lambda,\alpha)=\int_0^\infty tf(t;\lambda,\alpha)dt.$$

Then $h(t;\lambda,\alpha,p)$ can be rewritten as

$$h(t;\lambda,\alpha,p)=pf(t;\lambda,\alpha)+(1-p)\frac{1-F(t;\lambda,\alpha)}{\mu(\lambda,\alpha)},\quad t>0.$$

For (a). We concentrate on the case in which

$$A(\lambda_1,\alpha_1)=\lim_{t\to\infty}\frac{f(t;\lambda_1,\alpha_1)}{1-F(t;\lambda_1,\alpha_1)}=0. \tag{S4.2}$$

The proof for the case in which $A(\lambda_1,\alpha_1)=\infty$ is similar.

We first argue that $(\lambda_1,\alpha_1)=(\lambda_2,\alpha_2)$ when $h(t;\lambda_1,\alpha_1,p_1)=h(t;\lambda_2,\alpha_2,p_2)$ for all $t>0$.

If $(\lambda_1,\alpha_1)\ne(\lambda_2,\alpha_2)$, then using Condition A2 and L'Hospital's rule, we have

$$\lim_{t\to\infty}\frac{1-F(t;\lambda_1\alpha_1)}{1-F(t;\lambda_2,\alpha_2)}=\lim_{t\to\infty}\frac{f(t;\lambda_1,\alpha_1)}{f(t;\lambda_2,\alpha_2)}=0\ \text{ or }\ \infty. \tag{S4.3}$$

We further consider two different scenarios: $p_1 = 1$ and $p_1 \neq 1$.

Scenario I: $p_1 \neq 1$.

Dividing by $1-F(t;\lambda_1,\alpha_1)$ on both sides of $h(t;\lambda_1,\alpha_1,p_1) = h(t;\lambda_2,\alpha_2,p_2)$, we have

$$\frac{p_1 f(t;\lambda_1,\alpha_1)}{1-F(t;\lambda_1,\alpha_1)} + \frac{(1-p_1)}{\mu(\lambda_1,\alpha_1)} = \frac{p_2 f(t;\lambda_2,\alpha_2)}{1-F(t;\lambda_1,\alpha_1)} + \frac{(1-p_2)\{1-F(t;\lambda_2,\alpha_2)\}}{\mu(\lambda_2,\alpha_2)\{1-F(t;\lambda_1,\alpha_1)\}}. \tag{S4.4}$$

When $t \to \infty$ in (S4.4), by (S4.2)–(S4.3) and Condition A3, the left-hand side becomes a positive number $(1-p_1)/\mu(\lambda_1,\alpha_1)$, whereas the right-hand side becomes either 0 or ∞, which is a contradiction.

Scenario II: $p_1 = 1$.

When $p_1 = 1$, $h(t;\lambda_1,\alpha_1,p_1) = h(t;\lambda_2,\alpha_2,p_2)$ implies that

$$f(t;\lambda_1,\alpha_1) = p_2 f(t;\lambda_2,\alpha_2) + \frac{(1-p_2)\{1-F(t;\lambda_2,\alpha_2)\}}{\mu(\lambda_2,\alpha_2)}.$$

Dividing by $f(t;\lambda_1,\alpha_1)$ on both sides of the above equation gives

$$1 = \frac{p_2 f(t;\lambda_2,\alpha_2)}{f(t;\lambda_1,\alpha_1)} + \frac{(1-p_2)\{1-F(t;\lambda_2,\alpha_2)\}}{\mu(\lambda_2,\alpha_2) f(t;\lambda_1,\alpha_1)}. \tag{S4.5}$$

When $t \to \infty$ in (S4.5), by Conditions A2 and A3, the right-hand side is equal to either 0 or ∞, whereas the left-hand side is equal to 1, which is a contradiction.

In summary, if $h(t;\lambda_1,\alpha_1,p_1) = h(t;\lambda_2,\alpha_2,p_2)$ for all $t > 0$ and $A(\lambda_1,\alpha_1) = 0$, then under Conditions A1–A3, we must have

$$(\lambda_1,\alpha_1) = (\lambda_2,\alpha_2).$$

This, together with $h(t;\lambda_1,\alpha_1,p_1) = h(t;\lambda_2,\alpha_2,p_2)$, implies that

$$p_1 - p_2 = (p_1 - p_2)\frac{\mu(\lambda_1,\alpha_1) f(t;\lambda_1,\alpha_1)}{1-F(t;\lambda_1,\alpha_1)}$$

for all $t > 0$. Letting $t \to \infty$ in the above equation and noting that

$A(\lambda_1, \alpha_1) = 0$, we obtain $p_1 = p_2$. Hence $(\lambda_1, \alpha_1, p_1) = (\lambda_2, \alpha_2, p_2)$, as claimed in (a).

For (b). We first argue that $(\lambda_1, \alpha_1) = (\lambda_2, \alpha_2)$ when $0 < A(\lambda_1, \alpha_1) < \infty$ and $h(t; \lambda_1, \alpha_1, p_1) = h(t; \lambda_2, \alpha_2, p_2)$ for all $t > 0$.

If $(\lambda_1, \alpha_1) \neq (\lambda_2, \alpha_2)$, when $t \to \infty$ in (S4.4), by (S4.3) and Condition A3, the left-hand side of (S4.4) becomes $p_1 A(\lambda_1, \alpha_1) + \frac{(1-p_1)}{\mu(\lambda_1, \alpha_1)}$, which is finite and positive, while the right-hand side of (S4.4) is equal to either 0 or ∞, which is a contradiction. Hence we must have $(\lambda_1, \alpha_1) = (\lambda_2, \alpha_2)$. This completes the first part of (b).

Recall that $(\lambda_1, \alpha_1) = (\lambda_2, \alpha_2)$ implies that

$$p_1 - p_2 = (p_1 - p_2) \frac{\mu(\lambda_1, \alpha_1) f(t; \lambda_1, \alpha_1)}{1 - F(t; \lambda_1, \alpha_1)}$$

for all $t > 0$. If $\frac{f(t;\lambda_1,\alpha_1)}{1-F(t;\lambda_1,\alpha_1)}$ is not a constant function of t, then we must have $p_1 = p_2$. If $\frac{f(t;\lambda_1,\alpha_1)}{1-F(t;\lambda_1,\alpha_1)}$ is a constant function of t, then $\frac{\mu(\lambda_1,\alpha_1) f(t;\lambda_1,\alpha_1)}{1-F(t;\lambda_1,\alpha_1)}$ must equal 1 for all $t > 0$ because both $f(t; \lambda_1, \alpha_1)$ and $\frac{1-F(t;\lambda_1,\alpha_1)}{\mu(\lambda_1,\alpha_1)}$ are pdfs. In this case, p_1 and p_2 need not be equal. This completes the second part of (b).

S5 Proof of Theorem 2

S5.1 Two Technical Lemmas

We first establish two technical lemmas. Lemma 1 establishes the consistency of the maximum likelihood estimator (MLE) under the null model; this is the first step in the proof of Theorem 2. The lemma claims that any estimator of (λ, α, p) with a large likelihood value is consistent for λ and α under the null model. Recall that the true values of λ and α under the null model are λ_0 and α_0, respectively.

Lemma 1. *Assume the conditions of Theorem 2. Let* $(\bar{\lambda}, \bar{\alpha}, \bar{p})$ *be any estimator of* (λ, α, p) *such that*

$$l_n(\bar{\lambda}, \bar{\alpha}, \bar{p}) - l_n(\lambda_0, \alpha_0, 1) > c > -\infty \tag{S5.6}$$

for some constant c *for all* n. *Then under the null model,* $\bar{\lambda} - \lambda_0 = o_p(1)$ *and* $\bar{\alpha} - \alpha_0 = o_p(1)$.

Proof. Under Condition C2, both λ and α are identifiable under the null hypothesis, although p is not. The proof then follows by using techniques similar to those in Lemma 1 of Li et al. (2009) and Wald (1949). □

In the next lemma, we strengthen the conclusion of Lemma 1 by providing an order assessment of the estimators. Recall that

$$\begin{aligned} X_i &= \frac{\partial f(t_i; \lambda_0, \alpha_0)/\partial \lambda}{f(t_i; \lambda_0, \alpha_0)}, \\ Y_{i1} &= \frac{\partial f(t_i; \lambda_0, \alpha_0)/\partial \alpha}{f(t_i; \lambda_0, \alpha_0)}, \\ Y_{i2} &= \frac{\partial g(t_i; \lambda_0, \alpha_0)/\partial \alpha}{g(t_i; \lambda_0, \alpha_0)}. \end{aligned}$$

Note that under Condition C0,

$$h(t_i; \lambda_0, \alpha_0, 1) = f(t_i; \lambda_0, \alpha_0) = g(t_i; \lambda_0, \alpha_0) \quad \text{and} \quad \frac{\partial g(t_i; \lambda_0, \alpha_0)/\partial \lambda}{g(t_i; \lambda_0, \alpha_0)} = X_i.$$

Define $\mathbf{b}_i = (X_i, Y_{i1}, Y_{i2})^\top$. Then $\mathbb{E}(\mathbf{b}_i) = \mathbf{0}$ and we denote the variance-covariance matrix

$$\mathbf{B} = \mathbb{V}\text{ar}(\mathbf{b}_i) = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix}, \tag{S5.7}$$

where the expectation and variance are taken with respect to the null model $f(t; \lambda_0, \alpha_0)$.

Lemma 2. *Assume the conditions of Lemma 1. Then under the null model,* $\bar\lambda - \lambda_0 = O_p(n^{-1/2})$ *and* $\bar\alpha - \alpha_0 = O_p(n^{-1/2})$.

Proof. In the following, we first derive an upper bound for $\ell_n(\bar\lambda, \bar\alpha, \bar p) - \ell_n(\alpha_0, \lambda_0, 1)$. Then together with the lower bound c, we obtain the order assessment of $\bar\lambda$ and $\bar\alpha$. Write

$$\ell_n(\bar\lambda, \bar\alpha, \bar p) - \ell_n(\alpha_0, \lambda_0, 1) = \sum_{i=1}^n \log\{1 + \delta_i(\bar\lambda, \bar\alpha, \bar p)\}$$

with

$$\begin{aligned}\delta_i(\bar\lambda, \bar\alpha, \bar p) &= \frac{\bar p f(t_i; \bar\lambda, \bar\alpha) + (1-\bar p) g(t_i; \bar\lambda, \bar\alpha)}{h(t_i; \alpha_0, \lambda_0, 1)} - 1 \\ &= \bar p \frac{f(t_i; \bar\lambda, \bar\alpha) - f(t_i; \lambda_0, \alpha_0)}{f(t_i; \lambda_0, \alpha_0)} + (1-\bar p) \frac{g(t_i; \bar\lambda, \bar\alpha) - g(t_i; \lambda_0, \alpha_0)}{g(t_i; \lambda_0, \alpha_0)}.\end{aligned}$$

By the inequality $\log(1+x) \le x - x^2/2 + x^3/3$, we have

$$\ell_n(\bar\lambda, \bar\alpha, \bar p) - \ell_n(\alpha_0, \lambda_0, 1) \le \sum_{i=1}^n \delta_i(\bar\lambda, \bar\alpha, \bar p) - \sum_{i=1}^n \delta_i^2(\bar\lambda, \bar\alpha, \bar p)/2 + \sum_{i=1}^n \delta_i^3(\bar\lambda, \bar\alpha, \bar p)/3. \tag{S5.8}$$

From Lemma 1, we have the consistency results $\bar\lambda - \lambda_0 = o_p(1)$ and $\bar\alpha - \alpha_0 = o_p(1)$. Applying a first-order Taylor expansion to $f(t_i; \bar\lambda, \bar\alpha)$ and $g(t_i; \bar\lambda, \bar\alpha)$, we find that

$$\delta_i(\bar\lambda, \bar\alpha, \bar p) = (\bar\lambda - \lambda_0) X_i + \bar p(\bar\alpha - \alpha_0) Y_{i1} + (1-\bar p)(\bar\alpha - \alpha_0) Y_{i2} + \varepsilon_{in},$$

and the remainder term $\varepsilon_n = \sum_{i=1}^n \varepsilon_{in}$ satisfies

$$\varepsilon_n = O_p(n^{1/2}) \left\{ (\bar\lambda - \lambda_0)^2 + (\bar\alpha - \alpha_0)^2 \right\}.$$

Let $\bar s_1 = \bar\lambda - \lambda_0$, $\bar s_2 = \bar p(\bar\alpha - \alpha_0)$, $\bar s_3 = (1-\bar p)(\bar\alpha - \alpha_0)$, and $\bar{\mathbf{s}} = (\bar s_1, \bar s_2, \bar s_3)^\top$. Then

$$\delta_i(\bar\lambda, \bar\alpha, \bar p) = \bar{\mathbf{s}}^\top \mathbf{b}_i + \varepsilon_{in}$$

and

$$\varepsilon_n = O_p(n^{1/2})\bar{\mathbf{s}}^\top\bar{\mathbf{s}} = o_p(n)\bar{\mathbf{s}}^\top\bar{\mathbf{s}}. \tag{S5.9}$$

Therefore, for the linear term in (S5.8), we have

$$\sum_{i=1}^n \delta_i(\bar{\lambda}, \bar{\alpha}, \bar{p}) = \bar{\mathbf{s}}^\top \sum_{i=1}^n \mathbf{b}_i + \varepsilon_n, \tag{S5.10}$$

where the order of ε_n is assessed in (S5.9).

After some work, we can further show that

$$\begin{aligned} \sum_{i=1}^n \delta_i^2(\bar{\lambda}, \bar{\alpha}, \bar{p}) &= \sum_{i=1}^n (\bar{\mathbf{s}}^\top\mathbf{b}_i)^2 + O_p(\varepsilon_n), \\ \sum_{i=1}^n \delta_i^3(\bar{\lambda}, \bar{\alpha}, \bar{p}) &= \sum_{i=1}^n (\bar{\mathbf{s}}^\top\mathbf{b}_i)^3 + O_p(\varepsilon_n). \end{aligned}$$

By the strong law of large numbers and Condition C4 that $\mathbf{B}$ is positive definite, we further have

$$\sum_{i=1}^n \delta_i^2(\bar{\lambda}, \bar{\alpha}, \bar{p}) = n\bar{\mathbf{s}}^\top\mathbf{B}\bar{\mathbf{s}} + o_p(n)\bar{\mathbf{s}}^\top\bar{\mathbf{s}}, \tag{S5.11}$$

$$\sum_{i=1}^n \delta_i^3(\bar{\lambda}, \bar{\alpha}, \bar{p}) = o_p(n)\bar{\mathbf{s}}^\top\bar{\mathbf{s}}. \tag{S5.12}$$

Combining (S5.8)–(S5.12), we obtain the refined upper bound for $\ell_n(\bar{\lambda}, \bar{\alpha}, \bar{p}) - \ell_n(\alpha_0, \lambda_0, 1)$ as follows:

$$\ell_n(\bar{\lambda}, \bar{\alpha}, \bar{p}) - \ell_n(\alpha_0, \lambda_0, 1) \le \bar{\mathbf{s}}^\top \sum_{i=1}^n \mathbf{b}_i - 0.5n\bar{\mathbf{s}}^\top\mathbf{B}\bar{\mathbf{s}}\{1 + o_p(1)\}. \tag{S5.13}$$

Because $\mathbf{B}$ is positive definite, the upper bound in (S5.13) is of order $O_p(1)$. Together with the lower bound c, this implies that

$$\bar{\mathbf{s}} = O_p(n^{-1/2}).$$

Any values of $\bar{\mathbf{s}}$ outside this range will violate the inequality. Note that $\bar{\mathbf{s}}$ implies that $\bar{\lambda} - \lambda_0 = O_p(n^{-1/2})$ and $\bar{\alpha} - \alpha_0 = O_p(n^{-1/2})$. This completes the proof. □

S5.2 Proof of Theorem 2

Note that

$$R_n = 2\left\{\ell_n(\hat{\lambda}, \hat{\alpha}, \hat{p}) - \ell_n(\hat{\lambda}_0, \alpha_0, 1)\right\} = R_{1n} - R_{2n}, \tag{S5.14}$$

where

$$R_{1n} = 2\left\{\ell_n(\hat{\lambda}, \hat{\alpha}, \hat{p}) - \ell_n(\lambda_0, \alpha_0, 1)\right\}, \;\; R_{2n} = 2\left\{\ell_n(\hat{\lambda}_0, \alpha_0, 1) - \ell_n(\lambda_0, \alpha_0, 1)\right\}.$$

Applying some of the classical results for regular models (Serfling, 1980), we have

$$R_{2n} = \frac{\left(n^{-1/2}\sum_{i=1}^n X_i\right)^2}{B_{11}} + o_p(1). \tag{S5.15}$$

Next, we use a sandwich method to find the approximation of R_{1n}. We proceed in two steps. In step 1, we derive an upper bound for R_{1n}, and in step 2 we argue that the upper bound is achievable.

Let $(\hat{\lambda}_p, \hat{\alpha}_p) = \arg\max_{\lambda,p} \ell_n(\lambda, \alpha, p)$ be the constrained MLE of (λ, p) for given p. Define $R_{1n}(p) = 2\left\{\ell_n(\hat{\lambda}_p, \hat{\alpha}_p, p) - \ell_n(\lambda_0, \alpha_0, 1)\right\}$. Then $R_{1n} = \sup_p R_{1n}(p)$. By the definition of $(\hat{\lambda}_p, \hat{\alpha}_p)$, we have $\ell_n(\hat{\lambda}_p, \hat{\alpha}_p, p) - \ell_n(\lambda_0, \alpha_0, 1) \geq 0$. Hence, Condition (S5.6) is satisfied. Then applying the results in Lemma 2 and (S5.13), we obtain

$$R_{1n}(p) \leq 2\hat{\mathbf{s}}^\top(p)\sum_{i=1}^n \mathbf{b}_i - n\hat{\mathbf{s}}^\top(p)\mathbf{B}\hat{\mathbf{s}}(p) + o_p(1),$$

where $\hat{\mathbf{s}}(p)$ is defined similarly to $\bar{\mathbf{s}}$ with $(\hat{\lambda}_p, \hat{\alpha}_p, p)$ in place of $(\bar{\lambda}, \bar{\alpha}, \bar{p})$.

Define

$$\hat{\mathbf{t}}(p) = \left(\hat{t}_1(p), \hat{t}_2(p)\right)^\top = \left(\hat{\lambda}_p - \lambda_0, \hat{\alpha}_p - \alpha_0\right)^\top, \;\; \mathbf{c}_i(p) = \left(X_i, Y_i(p)\right)^\top$$

with $Y_i(p) = pY_{i1} + (1-p)Y_{i2}$, and $\mathbf{C}(p) = \mathbb{V}\mathrm{ar}\{\mathbf{c}_i(p)\}$. Then after some algebra, we obtain a refined upper bound for $R_{1n}(p)$ as

$$R_{1n}(p) \leq 2\hat{\mathbf{t}}^\top(p)\sum_{i=1}^n \mathbf{c}_i(p) - n\hat{\mathbf{t}}^\top(p)\mathbf{C}(p)\hat{\mathbf{t}}(p) + o_p(1). \tag{S5.16}$$

To further simplify the upper bound in (S5.16), let

$$a(p) = p\frac{B_{12}}{B_{11}} + (1-p)\frac{B_{13}}{B_{11}}, \quad \hat{t}_1^*(p) = \hat{\lambda}_p - \lambda_0 + a(p)(\hat{\alpha}_p - \alpha_0),$$

and

$$Z_i(p) = Y_i(p) - a(p)X_i. \tag{S5.17}$$

It can be verified that $\mathbb{C}\mathrm{ov}\{X_i, Z_i(p)\} = 0$ and $\mathbb{V}\mathrm{ar}\{Z_i(p)\} = \sigma(p,p)$, where $\sigma(\cdot,\cdot)$ is defined in (3.3) of the main paper. Then the upper bound in (S5.16) becomes

$$\begin{aligned} R_{1n}(p) &\leq 2\hat{t}_1^*(p)\sum_{i=1}^n X_i - nB_{11}\{\hat{t}_1^*(p)\}^2 \\ &\quad +2\hat{t}_2(p)\sum_{i=1}^n Z_i(p) - n\sigma(p,p)\{\hat{t}_2(p)\}^2 + o_p(1) \\ &\leq \frac{\left(n^{-1/2}\sum_{i=1}^n X_i\right)^2}{B_{11}} + \left\{\frac{n^{-1/2}\sum_{i=1}^n Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1). \end{aligned} \tag{S5.18}$$

Next, we show that the upper bound in (S5.18) for $R_{1n}(p)$ is achievable. Let $(\tilde{\lambda}_p, \tilde{\alpha}_p)$ be determined by

$$\tilde{\lambda}_p - \lambda_0 + a(p)(\tilde{\alpha}_p - \alpha_0) = n^{-1/2}\sum_{i=1}^n X_i/B_{11}, \quad \tilde{\alpha}_p - \alpha_0 = \frac{n^{-1/2}\sum_{i=1}^n Z_i(p)}{\sqrt{\sigma(p,p)}}.$$

Note that it is easy to verify that $(\tilde{\lambda}_p, \tilde{\alpha}_p)$ exists and

$$\tilde{\lambda}_p - \lambda_0 = O_p(n^{-1/2}), \quad \tilde{\alpha}_p - \alpha_0 = O_p(n^{-1/2})$$

uniformly over p. With this order assessment and applying a second-order Taylor expansion, we have

$$\begin{aligned} R_{1n}(p) &\geq 2\left\{\ell_n(\tilde{\lambda}_p, \tilde{\alpha}_p, p) - \ell_n(\lambda_0, \alpha_0, 1)\right\} \\ &= \frac{\left(n^{-1/2}\sum_{i=1}^n X_i\right)^2}{B_{11}} + \left\{\frac{n^{-1/2}\sum_{i=1}^n Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1). \end{aligned} \tag{S5.19}$$

Combining (S5.18) and (S5.19) leads to

$$R_{1n}(p) = \frac{\left(n^{-1/2}\sum_{i=1}^{n} X_i\right)^2}{B_{11}} + \left\{\frac{n^{-1/2}\sum_{i=1}^{n} Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1).$$

Hence

$$R_{1n} = \sup_p R_{1n}(p) = \frac{\left(n^{-1/2}\sum_{i=1}^{n} X_i\right)^2}{B_{11}} + \sup_p \left\{\frac{n^{-1/2}\sum_{i=1}^{n} Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1), \tag{S5.20}$$

which together with (S5.15) gives

$$R_n = \sup_p \left\{\frac{n^{-1/2}\sum_{i=1}^{n} Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1).$$

Recall the form of $Z_i(p)$ in (S5.17). We can rewrite it as

$$Z_i(p) = Z_{1i} + pZ_{2i}$$

with $Z_{i1} = Y_{i2} - (B_{13}/B_{11})X_i$ and

$$Z_{i2} = \{Y_{i1} - (B_{12}/B_{11})X_i\} - \{Y_{i2} - (B_{13}/B_{11})X_i\}.$$

It can be verified that $\mathbb{E}(Z_{i1}) = \mathbb{E}(Z_{i2}) = 0$ and

$$\mathbb{V}\text{ar}(Z_{1i}) = \sigma_{11}, \quad \mathbb{V}\text{ar}(Z_{2i}) = \sigma_{22}, \quad \mathbb{C}\text{ov}(Z_{1i}, Z_{2i}) = \sigma_{12}.$$

Hence

$$R_n = \sup_p \left\{\frac{n^{-1/2}\sum_{i=1}^{n} Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1) \xrightarrow{d} R = \sup_p Z^2(p),$$

where $Z(p) = (Z_1 + pZ_2)/\sqrt{\sigma(p,p)}$ with

$$(Z_1, Z_2)^\top \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}\right). \tag{S5.21}$$

It can be verified that the process $Z(p)$ is a Gaussian process with zero mean, unit variance, and covariance function

$$\mathbb{C}\mathrm{ov}\big\{Z(p_1), Z(p_2)\big\} = \frac{\sigma(p_1, p_2)}{\sqrt{\sigma(p_1, p_1)\sigma(p_2, p_2)}}.$$

This completes the proof. □

S6 Proof of Theorem 3

Recall that $Z(p) = (Z_1 + pZ_2)/\sqrt{\sigma(p,p)}$ with the joint distribution of $(Z_1, Z_2)^\top$ provided in (S5.21). Let

$$W_1 = \left(Z_1 - \frac{\sigma_{12}}{\sigma_{22}} Z_2\right)/a_1, \quad W_2 = Z_2/a_2,$$

where

$$a_1 = \sqrt{\sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}}, \quad a_2 = \sqrt{\sigma_{22}}.$$

By construction, it can be verified that W_1 and W_2 are two independent $N(0,1)$ random variables, and

$$Z(p) = \frac{a_1 W_1 + \left(p + \frac{\sigma_{12}}{\sigma_{22}}\right) a_2 W_2}{\sqrt{\sigma(p,p)}}.$$

To find a simpler form for $Z(p)$, we consider two polar transformations. The first one is defined in the main paper:

$$(\cos\theta, \sin\theta) = \big(c_1(p), c_2(p)\big),$$

where

$$c_1(p) = \frac{a_1}{\sqrt{\sigma(p,p)}} \quad \text{and} \quad c_2(p) = \frac{(p + \sigma_{12}/\sigma_{22})a_2}{\sqrt{\sigma(p,p)}}.$$

By Condition C6, we have

$$\big\{\big(c_1(p), c_2(p)\big) : 0 \le p \le 1\big\} = \{(\cos\theta, \sin\theta) : \Delta_1 \le \theta \le \Delta_2\}.$$

The second polar transformation is

$$(W_1, W_2) = (\rho \cos \eta, \rho \sin \eta),$$

where ρ^2 with $\rho > 0$ and η are two independent random variables with ρ^2 from a χ^2_2 distribution and η from a uniform distribution on $[-\pi, \pi]$. Then

$$Z(p) = \rho \cos \eta \cos \theta + \rho \sin \eta \sin \theta = \rho \cos(\theta - \eta)$$

and

$$\sup_p Z^2(p) = \sup_{\theta \in [\Delta_1, \Delta_2]} \rho^2 \cos^2(\theta - \eta).$$

After some algebra, we can check that

$$\sup_{\theta \in [\Delta_1, \Delta_2]} \rho^2 \cos^2(\theta - \eta) = \rho^2 \{I(\eta \in A_1) + I(\eta \in A_2) \cos^2(\eta - \Delta_2) + I(\eta \in A_3) \cos^2(\eta - \Delta_1)\}.$$

This completes the proof. □

S7 Proof of Theorem 4

We proceed in two steps. In the first step, we show that the models under the local alternatives

$$H_a^n : \lambda = \lambda_0, p = p_0, \alpha = \alpha_0 + \delta n^{-1/2} \tag{S7.22}$$

are contiguous to the null model (Le Cam, 1953). In the second step, we find the asymptotic distribution of R_n under H_a^n by using Le Cam's first and third lemmas (van der Vaart, 1998).

Let

$$\Lambda_n = \ell_n(\lambda_0, \alpha, p_0) - \ell_n(\lambda_0, \alpha_0, 1).$$

Using the second-order Taylor expansion, under the null model, we have

$$\Lambda_n = \sum_{i=1}^n Y_i(p_0)(\delta n^{-1/2}) - \frac{1}{2}\delta^2 \mathbb{V}\mathrm{ar}\{Y_i(p_0)\} + o_p(1).$$

By the central limit theorem, we have

$$\Lambda_n \to N(-0.5d_0^2, d_0^2)$$

in distribution under the null model, where $d_0 = \delta^2 \mathbb{V}\text{ar}\{Y_i(p_0)\}$. Therefore, the models under the local alternatives H_a^n in (S7.22) are contiguous to the null model (Le Cam, 1953). This completes step 1.

Next, we move on to step 2. Recall that under the null model,

$$R_n = \sup_p \left\{ \frac{n^{-1/2} \sum_{i=1}^n Z_i(p)}{\sqrt{\sigma(p,p)}} \right\}^2 + o_p(1).$$

By Le Cam's contiguity theory, the limiting distribution of R_n under the local alternatives H_a^n is determined by the joint limiting distribution of $\{n\sigma(p,p)\}^{-1/2} \sum_{i=1}^n Z_i(p)$ and Λ_n under the null model.

By the central limit theorem and Slutsky's theorem, the joint limiting distribution of $\{n\sigma(p,p)\}^{-1/2} \sum_{i=1}^n Z_i(p)$ and Λ_n under the null model is bivariate normal

$$N\left(\left(\begin{array}{c} 0 \\ -0.5d_0^2 \end{array} \right), \left(\begin{array}{cc} 1 & \omega(p, p_0) \\ \omega(p, p_0) & d_0^2 \end{array} \right) \right),$$

where

$$\begin{aligned} \omega(p, p_0) &= \mathbb{C}\text{ov}\left(\{\sigma(p,p)\}^{-1/2} Z_i(p), \delta Y_i(p_0)\right) \\ &= \mathbb{C}\text{ov}\left(\{\sigma(p,p)\}^{-1/2} Z_i(p), \delta Z_i(p_0)\right) \\ &= \frac{\delta\sigma(p,p_0)}{\sqrt{\sigma(p,p)}}. \end{aligned}$$

Note that in the second equation, we have used the fact that $\mathbb{C}\text{ov}\,\{Z_i(p), X_i\} = 0$ and the definition of $Z_i(p)$ in (S5.17).

By Le Cam's third lemma (van der Vaart, 1998), under the local alternatives H_a^n,

$$\frac{n^{-1/2} \sum_{i=1}^n Z_i(p)}{\sqrt{\sigma(p,p)}} \to N\left(\omega(p, p_0), 1\right)$$

in distribution, which implies that

$$\left\{\frac{n^{-1/2}\sum_{i=1}^{n} Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 \to \{Z(p)+\omega(p,p_0)\}^2$$

in distribution under H_a^n.

Because

$$R_n = \sup_p \left\{\frac{n^{-1/2}\sum_{i=1}^{n} Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1)$$

under the null model, by applying Le Cam's first lemma (van der Vaart, 1998), we have that

$$R_n = \sup_p \left\{\frac{n^{-1/2}\sum_{i=1}^{n} Z_i(p)}{\sqrt{\sigma(p,p)}}\right\}^2 + o_p(1)$$

holds also under the local alternatives H_a^n. Therefore, the asymptotic distribution of R_n under the local alternatives H_a^n is

$$\sup_p \left[\{Z(p)+\omega(p,p_0)\}^2\right].$$

This completes the proof. □

References

Deng, Y., You, C., Liu, Y., Qin, J., and Zhou, X.-H. (2021). Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for COVID-19 outbreak in China. *Biometrics*, 77:929–941.

Le Cam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related bayes estimates. *University of California Publications in Statistics*, 1:277–330.

Li, P., Chen, J., and Marriott, P. (2009). Non-finite Fisher information and homogeneity: an EM approach. *Biometrika*, 96:411–426.

Linton, N. M., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A. R., Jung, S.-m., Yuan, B., Kinoshita, R., and Nishiura, H. (2020). Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *Journal of Clinical Medicine*, 9:538.

Qin, J. (2017). *Biased Sampling, Over-identified Parameter Problems and Beyond.* Springer, Singapore.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, New York.

Wald, A. (1949). Note on the consistency of the maximum likelihood estimate. *Annals of Mathematical Statistics*, 20:595–601.