Recall the following notations
\[ \ell(\beta) = \sum_{i=1}^{n_r} c_i + \log \{ h(w_i; \beta) \}, \]
\[ m(\beta) = \frac{\partial \ell(\beta)}{\partial \beta}, \]
\[ m(w, \beta) = \frac{\partial \log \{ h(w; \beta) \}}{\partial \beta}. \]
where \( c_i \) is a constant independent of \( \beta \), for \( i = 1, \ldots, n_r \).

As mentioned in the paper, the penalized log-likelihood estimator (PMLE) \( \hat{\beta} \) satisfies the following score equation
\[ m(\beta) - P(\gamma) \beta = 0_{p(1+q)}. \] (1)

We now define
\[ \phi(\beta) = \frac{\partial^2 \ell(\beta)}{\partial \beta \partial \beta^T}, \quad \phi(w, \beta) = \frac{\partial^2 \log \{ h(w; \beta) \}}{\partial \beta \partial \beta^T}. \]

Based on Assumption (A2), we prove the following two lemmas that will streamline the proof of the first part of Theorem 1, i.e., the weak consistency of \( \hat{\beta} \).

**Lemma 1** Let \( h(w; \beta) \) be continuously differentiable a.e. for \( \beta \in B \). If \( \int \sup_{\beta \in B} \| \partial h(w; \beta)/\partial \beta \| \, dw < \infty \), then for \( \beta \in B \):

1. \( \int h(w; \beta) \, dw \) is continuously differentiable.
2. \( \int \partial h(w; \beta)/\partial \beta \, dw = \partial \int h(w; \beta) \, dw / \partial \beta. \)
1. AUXILIARY LEMMAS AND PROOFS

Proof See Newey and McFadden (1994, Lemma 3.6).

Lemma 2 If (A2) holds, then

\[ E \{m(W, \beta_0)\} = 0, \quad -E \{\phi(W, \beta_0)\} = i(\beta_0). \]

Proof By Lemma 1, it follows that

\[ E \{m(W, \beta_0)\} = \int m(w, \beta_0)h(w; \beta_0) \, dw = \int \left. \frac{\partial h(w; \beta)}{\partial \beta} \right|_{\beta = \beta_0} \, dw = \left. \frac{\partial h(w; \beta)}{\partial \beta} \right|_{\beta = \beta_0} = 0. \]

Now, using (A2) (parts 3 and 4) we have that

\[ E \{\phi(W, \beta_0)\} = \int \phi(w, \beta_0)h(w; \beta_0) \, dw \]
\[ = \int \left. \frac{\partial m(w, \beta)h(w; \beta)}{\partial \beta} \right|_{\beta = \beta_0} h(w; \beta_0) \, dw - \int m(w, \beta_0)m(w, \beta_0)^{T} h(w; \beta_0) \, dw \]
\[ = -\int m(w, \beta_0)m(w, \beta_0)^{T} h(w; \beta_0) \, dw. \]

We now prove the first part of Theorem 1 by proving the following lemma.

Lemma 3 Let \( Q_a = \{ \beta \in \mathbb{B} : \beta = \beta_0 + n_{t}^{-1/2}a \} \) be the surface of the sphere around \( \beta_0 \) with radius \( n_{t}^{-1/2}||a|| \). Then, for every \( \varepsilon > 0 \), there exists \( a \) such that

\[ \Pr \left\{ \sup_{\beta \in Q_a} \ell(\beta, y) < \ell(\beta_0, y) \right\} \geq 1 - \varepsilon. \]

for \( n_{t} \) large enough.

Proof Let \( \beta \in Q_a \), i.e., there exists \( a \) such that \( \beta = \beta_0 + n_{t}^{-1/2}a \). Applying a second-order Taylor expansion around \( \beta_0 \) of the penalized log-likelihood \( \ell(\beta, y) \), we have

\[
\ell(\beta, y) - \ell(\beta_0, y) = \ell(\beta) - \ell(\beta_0) - \frac{1}{2} \left\{ \beta^T P(y) \beta - \beta_0^T P(y) \beta_0 \right\} \\
= n_t^{-1/2}m(\beta_0)^T a + \frac{n_t^{-1}}{2} a^T \phi(\beta_0) a - n_t^{-1} a^T P(y) a - n_t^{-1/2} \beta_0^T P(y) a \\
\quad + \frac{n_t^{-3/2}}{2} \sum_q \sum_r \sum_s a_q a_r a_s \frac{\partial^3 \ell(\beta)}{\partial \beta_q r s} \bigg|_{\beta = \beta_0} \quad (2)
\]

with \( \beta^* \) in the interior of \( Q_a \). The terms involving the penalty matrix \( P(y) \) converge in probability to 0 due to the vanishing penalty from (A1). By the central limit theorem (based on a Lindeberg-type condition), Lemma 2, and (A2), we have that \( n_{t}^{-1/2}m(\beta_0) \xrightarrow{d} N(0, i(\beta_0)) \) implying that \( |n_{t}^{-1/2}m(\beta_0)^T a| = O_p(1)||a|| \). By the law of large numbers and Lemma 2, we have that \( n_{t}^{-1}\phi(\beta_0) \xrightarrow{p} -i(\beta_0) \). Hence, applying the continuous mapping theorem, we end up with \( a^T n_t^{-1} \phi(\beta_0) a/2 \xrightarrow{p} -a^T i(\beta_0) a/2 \leq ||a||^2 \lambda_{\min}/2 \), where \( \lambda_{\min} > 0 \) is the smallest eigenvalue of \( i(\beta_0) \). Finally, Assumption (A2) (last part) implies that the terms \( \partial^3 \ell(\beta)/\partial \beta_q r s |_{\beta = \beta^*} < \infty \) and that, by Cauchy–Schwartz inequality, the remainder
2. EXTREME TEMPERATURE ANALYSIS

The term in (2) vanishes in probability (is $O_p(n_r^{-1/2})$). Leaving out the terms vanishing in probability, (2) yields

$\ell(\beta, y) - \ell(\beta_0, y) \leq O_p(1)\|a\| - \|a\|^2 \lambda_{\min}/2 = T, \quad \beta \in Q_a,$

for $n_r$ large enough. Thus,

$$\Pr \left\{ \sup_{\beta \in Q_a} \ell(\beta, y) < \ell(\beta_0, y) \right\} \geq \Pr(T < 0),$$

implying that for every $\varepsilon > 0$, there exists an $a$ such that $\Pr(\sup_{\beta \in Q_a} \ell(\beta, y) < \ell(\beta_0, y)) \geq 1 - \varepsilon$. Hence, with probability tending to 1, the penalized log-likelihood $\ell(\beta, y)$ has a local maximum $\hat{\beta}$ in the interior of a sphere around $\beta_0$.

Lemma 3 yields the first part of Theorem 1. Now we move to the second part of Theorem 1. The asymptotic normality of the PMLE $\hat{\beta}$ is derived from a second-order Taylor expansion of the score equation (1) around the true parameter $\beta_0$. The Taylor expansion of (1) yields

$$m(\beta_0) - P(y)\beta_0 + \{\phi(\beta_0) - P(y)\}(\hat{\beta} - \beta_0) + r = 0_{p(1+q \hat{d})},$$

where

$$r = \frac{1}{2} (\hat{\beta} - \beta_0)^\top \frac{\partial^2 m(\beta)}{\partial \beta \partial \beta^\top} |_{\beta = \beta_0} (\hat{\beta} - \beta_0),$$

(3)

and $\beta^*$ is such that $\|\beta^* - \beta_0\| \leq \|\hat{\beta} - \beta_0\|$. Dividing (3) by $n_r^{1/2}$, we obtain

$$\frac{1}{n_r} \{\phi(\beta_0) - P(y) + \tilde{r}\} n_r^{1/2}(\hat{\beta} - \beta_0) = n_r^{-1/2} \{P(y)\beta_0 - m(\beta_0)\},$$

(4)

where

$$\tilde{r} = \frac{1}{2} (\hat{\beta} - \beta_0)^\top \frac{\partial^2 m(\beta)}{\partial \beta \partial \beta^\top} |_{\beta = \beta^*}.$$  

The consistency of $\hat{\beta}$ and the assumption on the third order derivative of $\log(h(w; \beta))$ in (A2) implies that $\tilde{r} \overset{p}{\rightarrow} 0$. Assumption (A1) implies that the terms involving $P(y)$ vanish in probability. Since $n_r^{-1/2} m(\beta_0) \overset{d}{\rightarrow} N(0, i(\beta_0))$ and $n_r^{-1} \phi(\beta_0) \overset{d}{\rightarrow} -i(\beta_0)$ (see above), Slutsky’s theorem implies that $n_r^{-1/2}(\hat{\beta} - \beta_0) \overset{d}{\rightarrow} N(0, i(\beta_0)^{-1})$ and proves hence the second part of Theorem 1.

2 | EXTREME TEMPERATURE ANALYSIS

This section supplements Section 5.2 in the paper.

Dependence of extreme high winter temperatures

The Dirichlet model of Table 2 is fitted to the pseudo-sample of extreme high temperatures where the angular observations corresponding to a radial component exceeding its 90%, 93%, and 97% quantiles, are considered. Figure
Dependence of extreme low winter temperatures

The Dirichlet model of Table 3 is fitted to the pseudo-sample of extreme low temperatures where the angular observations corresponding to a radial component exceeding its 90%, 93%, and 97% quantiles, are considered. Figure 2 displays the fitted smooth effects of time and day in season on the extremal coefficient, along with their associated 95% confidence intervals.

References

FIGURE 1  Fitted smooth effects for the extremal coefficient under the Dirichlet model of Table 2 along with their associated 95% (pointwise) asymptotic confidence bands. Different radial thresholds are considered: the 90% quantile (top), the 93% quantile (middle), and the 97% quantile (bottom).
FIGURE 2  Fitted smooth effects for the extremal coefficient under the Dirichlet model of Table 3 along with their associated 95% (pointwise) asymptotic confidence bands. Different radial thresholds are considered: the 90% quantile (top), the 93% quantile (middle), and the 97% quantile (bottom).