Joint Distribution and Marginal Distribution Methods for Generalized Linear Model

Supplementary Material

By Junyi Dong$^a$ and Qiqing Yu$^a$

\textit{a. Department of Mathematical Sciences, SUNY, Binghamton, NY 13902}

Corresponding author’s email address: qyu@math.binghamton.edu

In this supplementary material, we prove Lemmas and Theorems and also derive the asymptotic variance of $T_2$ in Section 3.2 (Dong and Yu 2018).

1. Proofs

**Lemma 1.** Let $(X, \mathcal{F}, P)$ be a probability space. Let $\mu_n(t, \omega)$, $t \in \mathbb{R}$ and $\omega \in X$, be a sequence of measure. Let $f_n$ and $g_n$ be measurable functions,

\[
\Omega_a = \{ \omega \in X : \mu_n(\cdot, \omega) \rightarrow \mu(\cdot, \omega) \text{ set-wisely} \},
\]

\[
\Omega_b = \{ \omega \in X : f_n(t, \omega) \rightarrow f(t, \omega) \text{ point-wisely in } t \}, \text{ and}
\]

\[
\Omega_c = \{ \omega \in X : g_n(t, \omega) \rightarrow g(t, \omega) \text{ point-wisely in } t \}.
\]

If $P(\Omega_a \cap \Omega_b \cap \Omega_c) = 1$, $|f_n| \leq g_n$, and $\int g_n d\mu_n \xrightarrow{a.s.} \int g d\mu < \infty$, then $\int f_n d\mu_n \xrightarrow{a.s.} \int f d\mu$.

**Proof.** Let $\Omega = \Omega_a \cap \Omega_b \cap \Omega_c$, then $P(\Omega) = 1$. For each $\omega \in \Omega$, $\mu_n(\cdot, \omega) \rightarrow \mu(\cdot, \omega)$ set-wisely, $f_n(t, \omega) \rightarrow f(t, \omega)$ point-wisely in $t$, and $f_n(t, \omega) \rightarrow f(t, \omega)$ point-wisely in $t$. Since $|f_n| \leq g_n$ and $\int g_n d\mu_n \xrightarrow{a.s.} \int g d\mu < \infty$, by the General Convergence Theorem (R1988), $\lim \int f_n(t, \omega) d\mu_n(t, \omega) = \int f(t, \omega) d\mu(t, \omega)$. Since $P(\Omega) = 1$, $\int f_n d\mu_n \xrightarrow{a.s.} \int f d\mu$.

**Remark 2.** Let $\Omega_0$ be the event that $\tilde{F}_{Y,Z}(t, z) = \frac{1}{n} \sum_{i=1}^{n} 1(Y_i \leq t, Z_i \leq z) \rightarrow F_{Y,Z}(t, z)$. Let $\Omega_1$ be the event that $\tilde{F}_Y(t) = \frac{1}{n} \sum_{i=1}^{n} 1(Y_i \leq t) \rightarrow F_Y(t)$. Then, by the strong law of large number (SLLN), $P(\Omega_0) = 1$ and $P(\Omega_1) = 1$.

**Proof of Lemma 8 (Dong and Yu 2018).** We first prove Statement 1. Notice that the proof for the Normal GLM is proved in Remark 9 (Dong and Yu 2018). Let $\mathcal{L}(\beta, \phi)$ and $\tilde{\mathcal{L}}_n(\beta, \phi, Y)$ be defined as in (10) (Dong and Yu 2018).

By the SLLN and assumption (C6), $\tilde{\mathcal{L}}_n(\beta, \phi) \rightarrow \mathcal{L}(\beta, \phi)$ almost surely for each $(\beta, \phi) \in \mathbb{R}^{p+1}$. By assumption (C1), $B = \{(\beta_0, \phi_0) : (\beta_0, \phi_0) = \arg\sup_{\beta, \phi} \mathcal{L}(\theta, \phi)\}$ is a singleton set, thus $(\beta_0, \phi_0) = \arg\sup_{\beta \in \mathbb{R}^{p}, \phi \in \mathbb{R}} \mathcal{L}(\theta, \phi)$ is uniquely determined. $(\hat{\beta}_n, \hat{\phi}_n) = (\hat{\beta}, \hat{\phi}) = \arg\sup_{\beta \in \mathbb{R}^{p}, \phi \in \mathbb{R}} \tilde{\mathcal{L}}_n(\beta, \phi)$.
yields $\mathcal{L}(\beta_0, \phi_0) \geq \mathcal{L}(\beta, \phi)$ and $\tilde{\mathcal{L}}_n(\hat{\beta}, \hat{\phi}) \geq \tilde{\mathcal{L}}_n(\beta, \phi)$ for any $(\beta, \phi) \in \mathbb{R}^{p+1}$. Since $\tilde{\mathcal{L}}_n(\hat{\beta}, \hat{\phi}) \geq \tilde{\mathcal{L}}_n(\beta_0, \phi_0)$,

$$\lim_{n \to \infty} \tilde{\mathcal{L}}_n(\hat{\beta}, \hat{\phi}) \geq \lim_{n \to \infty} \tilde{\mathcal{L}}_n(\beta_0, \phi_0) = \mathcal{L}(\beta_0, \phi_0) \text{ a.s.} \quad (1)$$

For each $\omega$ in $\Omega_0$ (see Remark 2), let $(\beta^*, \phi^*)$ be a limiting point of $(\hat{\beta}, \hat{\phi})$ such that there exists a subsequence $(\hat{\beta}_{n_l}(\omega), \hat{\phi}_{n_l}(\omega))$ converges to $(\beta^*, \phi^*)$, that is, $(\hat{\beta}_{n_l}(\omega), \hat{\phi}_{n_l}(\omega)) \to (\beta^*, \phi^*)$. Let $f_n(\hat{\beta}_{n_l}(\omega), \hat{\phi}_{n_l}(\omega)) = f_n(\omega) = \frac{Y h(\beta_{n_l}(\omega)^T Z) - k(\beta_{n_l}(\omega)^T Z)}{a(\phi_{n_l}(\omega))} + c(Y, \phi_{n_l}(\omega))$ and $f(\beta^*, \phi^*) = f(\omega) = \frac{Y h(\beta^*(\omega)^T Z) - k(\beta^*(\omega)^T Z)}{a(\phi^*)} + c(Y, \phi^*)$. Then $f_n(\omega) \to f(\omega)$. We shall show that $\exists$ a function $g_n(\hat{\beta}_{n_l}(\omega), \hat{\phi}_{n_l}(\omega)) = g_n(\omega)$ such that

(a) $|f_n(\omega)| \leq g_n(\omega)$, (b) $g_n(\omega) \to g(\omega)$, and (c) $\int g_n(\omega) d\tilde{F}_Y Z(t, z) \to \int g(\omega) dF_Y Z(t, z) < \infty$, \quad (2)

then by Lemma 1 $\int f_n(\omega) d\tilde{F}_Y Z(t, z) \to \int f(\omega) dF_Y Z(t, z)$, that is, $\tilde{\mathcal{L}}_n(\hat{\beta}_{n_l}(\omega), \hat{\phi}_{n_l}(\omega)) \to \mathcal{L}(\beta^*, \phi^*)$. Since $\lim \tilde{\mathcal{L}}_n(\hat{\beta}_{n_l}(\omega), \hat{\phi}_{n_l}(\omega)) \geq \lim \tilde{\mathcal{L}}_n(\beta_0, \phi_0)$, we have $\mathcal{L}(\beta^*(\omega), \phi^*(\omega)) \geq \mathcal{L}(\beta_0, \phi_0)$. Then $\mathcal{L}(\beta^*(\omega), \phi^*(\omega)) = \mathcal{L}(\beta_0, \phi_0)$ which implies $(\beta^*(\omega), \phi^*(\omega)) = (\beta_0, \phi_0)$ as $B$ is a singleton set. Since every convergent subsequence of $(\hat{\beta}(\omega), \hat{\phi}(\omega))$ converges to $(\beta_0, \phi_0)$ for all $\omega \in \Omega_1$ (see Remark 2 and $P(\Omega_1) = 1$, we have $(\hat{\beta}, \hat{\phi}) \overset{a.s.}{\to} (\beta_0, \phi_0)$.

We now prove the existence of $g_n(\omega)$ satisfying (2) and under the Poisson, Binomial and Gamma with their canonical link functions separately as there is no unified proof.

**Poisson GLMs.** In the Poisson GLM with mean $\mu$, $\phi = 1$, $a(\phi) = 1$, $\theta = \ln \mu$, $b(\theta) = \exp(\theta)$ and $c(y, \phi) = -\ln y!$. The canonical link function is $g(t) = \ln t$. Then $f_n(\omega) = Y(\hat{\beta}_{n_l}(\omega)^T Z) - \exp(\hat{\beta}_{n_l}(\omega)^T Z) - \ln Y!$. Under assumptions (C2) and (C4), $\hat{\beta}_{n_l}(\omega)$ is bounded and $Z$ is bounded. It follows that $||\hat{\beta}_{n_l}(\omega)^T Z|| \leq K$. Then $|f_n(\omega)| \leq K|Y| + e^K + \ln Y! = g_n = g$, then $g_n \to g$ and $\int g_n(\omega) d\tilde{F}_Y Z(t, z) \to \int g(\omega) dF_Y Z(t, z) = K\mathbb{E}[|Y|] + e^K + \mathbb{E}[\ln Y!] < \infty$ by the assumption that all expectations exist (see (C6)).

**Binomial GLMs.** In the Binomial GLM, $\text{Binom}(m, \mu_i)/m$, $\phi = 1$, $a(\phi) = 1/m$, $\theta = \ln \frac{\mu}{1-\mu}$, $b(\theta) = -\ln(1 + \exp(\theta))$ and $c(y, \phi) = \ln \binom{m}{\mu y}$. The canonical link function is $g(t) = \ln \frac{t}{1-t}$. Let $f_n(\omega) = \left[Y(\hat{\beta}_{n_l}(\omega)^T Z) - \ln(1 + \exp(\hat{\beta}_{n_l}(\omega)^T Z))\right]m + \ln \binom{m}{\mu Y}$. By assumptions (C2) and (C4), we can assume that $||\hat{\beta}_{n_l}(\omega)^T Z|| \leq K$, and then $|f_n(\omega)| \leq [K|Y| + \ln(1 + e^K)]m + \ln \binom{m}{\mu Y} = g_n = g$. Then $g_n \to$
Lemma 8 in Dong and Yu (Dong and Yu 2018), expectation exist due to the fact that \( Y \) and \( m \) are bounded under the binomial distribution.

**Gamma GLMs.** In the Gamma GLM, Gamma(\( \alpha, \mu / \alpha \)), \( \phi = \alpha, a(\phi) = 1 / \phi, \theta = -1 / \mu, b(\theta) = -\ln(-\theta) \) and \( c(y, \phi) = -\ln(\Gamma(\alpha) + (\alpha - 1)\ln y + a \ln \alpha) \). The canonical link function is \( g(\phi) = -1 / \phi \). Let \( f_n(\omega) = \alpha \mathbb{E} \left[ (\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z}) + \ln(- (\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z})) \right] - \ln(\Gamma(\alpha) + (\alpha - 1)\ln y + a \ln \alpha) \).

Since \( \mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z} \) is bounded above and bounded below by some assumptions (C5), there exists \( M_1 \) and \( M_2 \) such that \( |(\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z})| < M_1 \) and \( |\ln(- (\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z}))| < M_2 \).

Thus \( |f_n(\omega)| \leq \alpha M_1 |Y| + M_2 - \ln(\Gamma(\alpha) + (\alpha - 1)\ln y + a \ln \alpha) = g_n = g, g_n \rightarrow g \) and \( \int g_n(\omega) d\hat{F}_{Y,Z}(t, z) \rightarrow \int g(\omega) dF_Y(t, z) \). Let \( \mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z} \) be as defined in Proof of Lemma 8 in Dong and Yu (Dong and Yu 2018).

To prove (12) in Dong and Yu (Dong and Yu 2018), let \( f_n(\omega) = \mathbb{E} \left[ (\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z}) + \ln(- (\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z})) \right] - \ln(\Gamma(\alpha) + (\alpha - 1)\ln y + a \ln \alpha) \). Since \( \mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z} \) is bounded above and bounded below by some assumptions (C5), there exists \( M_1 \) and \( M_2 \) such that \( |(\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z})| < M_1 \) and \( |\ln(- (\mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z}))| < M_2 \).

Thus \( |f_n(\omega)| \leq \alpha M_1 |Y| + M_2 - \ln(\Gamma(\alpha) + (\alpha - 1)\ln y + a \ln \alpha) = g_n = g, g_n \rightarrow g \) and \( \int g_n(\omega) d\hat{F}_{Y,Z}(t, z) \rightarrow \int g(\omega) dF_Y(t, z) \). Let \( \mathbf{\hat{b}}_{nl}(\omega)^T \mathbf{Z} \) be as defined in Proof of Lemma 8 in Dong and Yu (Dong and Yu 2018).
$g_n = g = 1$ and $\int g dF_Z = 1$, then by Lemma 1 $\int f_n(\omega) d\hat{F}_Z(\omega) \rightarrow \int f dF_Z$. Since $P(\Omega_1 \cap \Omega_2 \cap \Omega_3) = 1$, (12) in Dong and Yu (Dong and Yu 2018) follows.

The convergence in (13) follows (12) and Theorem 5 in Dong and Yu (Dong and Yu 2018).

2. Asymptotic Variance of $T_2$ (Dong and Yu 2018)

Let $Y_i \in \{M_1 < M_2 < ... < M_k\}$, $i = 1, ..., n$. $T_2 = \int [\hat{F}(t) - \hat{F}^*(t)] d\hat{F}(t) = \sum_{i=1}^{k} \hat{F}(M_i) \hat{f}(M_i) - \sum_{i=1}^{k} \hat{F}^*(M_i) \hat{f}(M_i)$

When $Y$ is discrete (as in Binomial or Poisson model):

$$V[T_2] = V[\sum_{i=1}^{k} \hat{F}(M_i) \hat{f}(M_i)] + V[\sum_{i=1}^{k} \hat{F}^*(M_i) \hat{f}(M_i)]$$

$$- 2Cov[\sum_{i=1}^{k} \hat{F}(M_i) \hat{f}(M_i), \sum_{i=1}^{k} \hat{F}^*(M_i) \hat{f}(M_i)]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} Cov[\hat{F}(M_i) \hat{f}(M_i), \hat{F}(M_j) \hat{f}(M_j)] = V_1$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} Cov[\hat{F}^*(M_i) \hat{f}(M_i), \hat{F}^*(M_j) \hat{f}(M_j)] = V_2$$

$$- 2 \sum_{i=1}^{n} \sum_{j=1}^{n} Cov[\hat{F}(M_i) \hat{f}(M_i), \hat{F}^*(M_j) \hat{f}(M_j)] = V_3$$

Estimate $V_1$.

Let $\hat{F}_i = \hat{F}(M_i)$ and $\hat{F}_j = \hat{F}(M_j)$, then $\hat{f}(M_i) = \hat{F}_i - \hat{F}_{i-1}$ and $\hat{f}(M_j) = \hat{F}_j - \hat{F}_{j-1}$.

$$Cov[\hat{F}(M_i) \hat{f}(M_i), \hat{F}(M_j) \hat{f}(M_j)]$$

$$= Cov[\hat{F}(M_i) (\hat{F}_i - \hat{F}_{i-1}), \hat{F}(M_j) (\hat{F}_j - \hat{F}_{j-1})]$$

$$= Cov[\hat{F}_i^2, \hat{F}_j^2] - Cov[\hat{F}_i \hat{F}_{i-1}, \hat{F}_j \hat{F}_{j-1}] - Cov[\hat{F}_i^2, \hat{F}_j \hat{F}_{j-1}] + Cov[\hat{F}_i \hat{F}_{i-1}, \hat{F}_j \hat{F}_{j-1}]$$
Each term can be estimated as follows.

\[
\text{COV}[\hat{F}^2_i, \hat{F}^2_j] = E[\hat{F}^2_i \hat{F}^2_j] - E[\hat{F}^2_i]E[\hat{F}^2_j] \\
E[\hat{F}^2_i \hat{F}^2_j] = E[(\frac{1}{n} \sum_{k=1}^{n} 1(Y_k \leq M_i))^2(\frac{1}{n} \sum_{l=1}^{n} 1(Y_l \leq M_j))^2] \\
= \frac{1}{n^4} \sum_{k=1}^{n} \sum_{p=1}^{n} \sum_{l=1}^{n} \sum_{q=1}^{n} \mathbb{P}(Y_k \leq M_i, Y_p \leq M_i, Y_l \leq M_j, Y_q \leq M_j) \\
= \frac{1}{n^4} [(F_i F_j^2 + 4 \min(F_i, F_j) F_i F_j + F_i^2 F_j^2)n(n-1)(n-2) + 2 \min(F_i, F_j) F_i \\
+ 2 \min(F_i, F_j) F_j n(n-1) + n \min(F_i, F_j) + F_i^2 F_j^2 n(n-1)(n-2)(n-3) \\
+ (F_i F_j + 2 \min(F_i, F_j)^2) n(n-1)] \\
E[\hat{F}^2_i] = E[(\frac{1}{n} \sum_{k=1}^{n} 1(Y_k \leq M_i))^2] \\
= \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbb{P}(Y_k \leq M_i, Y_l \leq M_i) \\
= \frac{1}{n^2} [nF_i + n(n-1)F_i^2] \\
E[\hat{F}^2_j] = E[(\frac{1}{n} \sum_{k=1}^{n} 1(Y_k \leq M_j))^2] \\
= \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbb{P}(Y_k \leq M_j, Y_l \leq M_j) \\
= \frac{1}{n^2} [nF_j + n(n-1)F_j^2] \\
\text{COV}[\hat{F}^2_i, \hat{F}^2_j] \approx \frac{1}{n^4} [(\hat{F}_i \hat{F}_j^2 + 4 \min(\hat{F}_i, \hat{F}_j) \hat{F}_i \hat{F}_j + \hat{F}_i^2 \hat{F}_j^2)n(n-1)(n-2) + 2 \min(\hat{F}_i, \hat{F}_j) \hat{F}_i \\
+ 2 \min(\hat{F}_i, \hat{F}_j) \hat{F}_j n(n-1) + n \min(\hat{F}_i, \hat{F}_j) + \hat{F}_i^2 \hat{F}_j^2 n(n-1)(n-2)(n-3) \\
+ (\hat{F}_i \hat{F}_j + 2 \min(\hat{F}_i, \hat{F}_j)^2) n(n-1) - (n\hat{F}_i + n(n-1)\hat{F}_i^2)(n\hat{F}_j + n(n-1)\hat{F}_j^2)]
\[
\text{COV}[\hat{F}_i \hat{F}_{i-1}, \hat{F}_j^2] = E[\hat{F}_i \hat{F}_{i-1} \hat{F}_j^2] - E[\hat{F}_i \hat{F}_{i-1}]E[\hat{F}_j^2]
\]

\[
E[\hat{F}_i \hat{F}_{i-1} \hat{F}_j^2] = E[\frac{1}{n} \sum_{k=1}^{n} 1(Y_k \leq M_i) \frac{1}{n} \sum_{l=1}^{n} 1(Y_l \leq M_{i-1}) (\frac{1}{n} \sum_{p=1}^{n} 1(Y_p \leq M_j))^2]
\]

\[
= \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} P(Y_k \leq M_i, Y_l \leq M_{i-1}, Y_p \leq M_j, Y_q \leq M_j)
\]

\[
= \frac{1}{n^4} [(F_{i-1} F_j^2 + 2 \min(F_i, F_j) F_{i-1} F_j + 2 \min(F_{i-1}, F_j) F_j F_{i-1} F_j) n(n-1)(n-2) + (2 \min(F_{i-1}, F_j) F_j + \min(F_i, F_j) F_{i-1} + \min(F_{i-1}, F_j) F_i)
\]

\[
E[\hat{F}_i] = E[(\frac{1}{n} \sum_{k=1}^{n} 1(Y_k \leq M_i))^2]
\]

\[
= \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} P(Y_k \leq M_j, Y_l \leq M_j)
\]

\[
= \frac{1}{n^2} [n F_i + n(n-1) F_i^2]
\]

\[
\text{COV}[\hat{F}_i \hat{F}_{i-1}, \hat{F}_j^2] \approx \frac{1}{n^4} [(\hat{F}_{i-1} \hat{F}_j^2 + 2 \min(\hat{F}_i, \hat{F}_j) \hat{F}_{i-1} \hat{F}_j + 2 \min(\hat{F}_{i-1}, \hat{F}_j) \hat{F}_i \hat{F}_j + \hat{F}_i \hat{F}_{i-1} \hat{F}_j)
\]

\[
n(n-1)(n-2) + (2 \min(\hat{F}_{i-1}, \hat{F}_j) \hat{F}_j + \min(\hat{F}_i, \hat{F}_j) \hat{F}_{i-1} + \min(\hat{F}_{i-1}, \hat{F}_j) \hat{F}_i)
\]

\[
n(n-1) + \min(\hat{F}_{i-1}, \hat{F}_j) n + \hat{F}_i \hat{F}_{i-1} \hat{F}_j n(n-1)(n-2)(n-3) +
\]

\[
(2 \min(\hat{F}_{i-1}, \hat{F}_j) \min(\hat{F}_i, \hat{F}_j) + \min(\hat{F}_{i-1}, \hat{F}_i) \hat{F}_j) n(n-1) -
\]

\[
(n \min(\hat{F}_{i-1}, \hat{F}_i) + n(n-1) \hat{F}_i \hat{F}_{i-1}) (n \hat{F}_j + n(n-1) \hat{F}_j^2)]
\]
\[ \text{COV}[\hat{F}_j, \hat{F}_{j-1}, \hat{F}_i^2] \approx \frac{1}{n^4} \left[ (\hat{F}_{j-1} \hat{F}_i^2 + 2 \min(\hat{F}_j, \hat{F}_i) \hat{F}_{j-1} \hat{F}_i + 2 \min(\hat{F}_{j-1}, \hat{F}_i) \hat{F}_j \hat{F}_i + \hat{F}_j \hat{F}_{j-1} \hat{F}_i) \\
(n(n-1)(n-2) + (2 \min(\hat{F}_{j-1}, \hat{F}_i) \hat{F}_i + \min(\hat{F}_j, \hat{F}_i) \hat{F}_{j-1} + \min(\hat{F}_{j-1}, \hat{F}_i) \hat{F}_j) \\
n(n-1) + \min(\hat{F}_{j-1}, \hat{F}_i)n + \hat{F}_j \hat{F}_{j-1} \hat{F}_i^2 n(n-1)(n-2)(n-3) + \\
(2 \min(\hat{F}_{j-1}, \hat{F}_i) \min(\hat{F}_j, \hat{F}_i) + \min(\hat{F}_{j-1}, \hat{F}_j) \hat{F}_i)n(n-1) \\
- (n \min(\hat{F}_{j-1}, \hat{F}_j) + n(n-1) \hat{F}_j \hat{F}_{j-1})(n\hat{F}_i + n(n-1)\hat{F}_i^2) \right] \]
$$\text{COV}[\hat{F}_i, \hat{F}_{j-1}, \hat{F}_j] = \mathbb{E}[\hat{F}_i \hat{F}_{j-1}] - \mathbb{E}[\hat{F}_i] \mathbb{E}[\hat{F}_{j-1}]$$

$$\mathbb{E}[\hat{F}_i \hat{F}_{j-1}] = \frac{1}{n^4} \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} \mathbb{P}(Y_k \leq M_i, Y_l \leq M_i, Y_p \leq M_j, Y_q \leq M_j)$$

$$= \frac{1}{n^2} \left[ n(n-1)(n-2)(F_{i-1} F_{j-1} + \min(F_i, F_j) F_{i-1} F_{j-1} + \min(F_{i-1}, F_{j-1}) F_i F_j + F_{j-1} F_i F_{i-1} + n(n-1)(\min(F_{i-1}, F_j) F_{j-1} + \min(F_{i-1}, F_{j-1}) F_j + \min(F_{j-1}, F_i) F_{i-1} + \min(F_{i-1}, F_{j-1}) F_i) + n \min(F_{i-1}, F_{j-1}) + n(n-1)(F_{i-1} F_{j-1} + \min(F_i, F_j) \min(F_{i-1}, F_{j-1}) + \min(F_{i-1}, F_j) \min(F_{j-1}, F_i)) \right]$$

$$\mathbb{E}[\hat{F}_i] = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbb{P}(Y_k \leq M_i, Y_l \leq M_i)$$

$$= \frac{1}{n^2} \left[ n \min(F_{i-1}, F_i) + n(n-1) F_{i-1} \right]$$

$$\mathbb{E}[\hat{F}_j] = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbb{P}(Y_k \leq M_j, Y_l \leq M_j)$$

$$= \frac{1}{n^2} \left[ n \min(F_{j-1}, F_j) + n(n-1) F_{j-1} \right]$$

$$\text{COV}[\hat{F}_i, \hat{F}_{j-1}, \hat{F}_j] \approx \frac{1}{n^4} \left[ n(n-1)(n-2)(\hat{F}_{i-1} \hat{F}_j \hat{F}_{j-1} + \min(\hat{F}_i, \hat{F}_j) \hat{F}_{i-1} \hat{F}_{j-1} + \min(\hat{F}_{i-1}, \hat{F}_j) \hat{F}_i \hat{F}_{j-1} + \hat{F}_{j-1} \hat{F}_i \hat{F}_{i-1} + n(n-1)(\min(\hat{F}_{i-1}, \hat{F}_j) \hat{F}_{j-1} + \min(\hat{F}_{i-1}, \hat{F}_{j-1}) \hat{F}_j + \min(\hat{F}_{j-1}, \hat{F}_i) \hat{F}_{i-1} + \min(\hat{F}_{i-1}, \hat{F}_{j-1}) \hat{F}_i) + n \min(\hat{F}_{i-1}, \hat{F}_{j-1}) + n(n-1)(\hat{F}_{i-1} \hat{F}_{j-1} + \min(\hat{F}_i, \hat{F}_j) \min(\hat{F}_{i-1}, \hat{F}_{j-1}) + \min(\hat{F}_i, \hat{F}_j) \min(\hat{F}_{i-1}, \hat{F}_{j-1}) + \min(\hat{F}_{i-1}, \hat{F}_j) \min(\hat{F}_i, \hat{F}_{j-1})) \right]$$

Estimate $\mathbb{V}_2$.

Let $\hat{F}_i = \hat{F}^*(M_i)$ and $\hat{F}_j = \hat{F}^*(M_j)$, then $\hat{f}(M_i) = \hat{F}_i - \hat{F}_{i-1}$ and $\hat{f}(M_j) = \hat{F}_j - \hat{F}_{j-1}$. 
\[ \text{COV}[\hat{F}^*(M_i), \hat{F}(M_j)] \]
\[ = \text{COV}[\hat{F}^*(M_i)(\hat{F}_i - \hat{F}_{i-1}), \hat{F}^*(M_j)(\hat{F}_j - \hat{F}_{j-1})] \]
\[ = \text{COV}[\hat{F}^*_i \hat{F}_i, \hat{F}^*_j \hat{F}_j] - \text{COV}[\hat{F}^*_i \hat{F}_{i-1}, \hat{F}^*_j \hat{F}_j] - \text{COV}[\hat{F}^*_i \hat{F}_i, \hat{F}^*_j \hat{F}_{j-1}] + \text{COV}[\hat{F}^*_i \hat{F}_{i-1}, \hat{F}^*_j \hat{F}_{j-1}] \]

Each term can be estimated as follows.

\[ \text{COV}[(\hat{F}_i^*)^2, (\hat{F}_j^*)^2] = \text{COV}[(\frac{1}{k} \sum_{p=1}^{k} \hat{f}^*(M_p)1(M_p \leq M_i))((\frac{1}{k} \sum_{q=1}^{k} 1(M_q \leq M_i))], \]
\[ = (\frac{1}{k} \sum_{m=1}^{k} \hat{f}^*(M_m)1(M_m \leq M_j))((\frac{1}{k} \sum_{l=1}^{k} 1(M_l \leq M_j))] \]
\[ = \frac{1}{k^2} \sum_{p=1}^{k} \sum_{q=1}^{k} \sum_{m=1}^{k} \sum_{l=1}^{k} \text{COV}[\hat{f}^*(M_p)1(M_p \leq M_i, M_q \leq M_l)], \]
\[ \hat{f}^*(M_m)1(M_m \leq M_j, M_l \leq M_j)] \]

Let \( Z_i \in \{W_1, ..., W_{kk}\}, i=1, ..., n. \) Let \( p_j = \hat{f}^*_Z(W_j) = \frac{1}{n} \sum_{i=1}^{n} 1(Z_i = W_j). \) Then

\[ \hat{f}^*(M_p) = \frac{1}{kk} \sum_{l=1}^{kk} \hat{f}^*(M_p | W_l)p_l \]
\[ \text{COV}[\hat{f}^*(M_p)1(M_p \leq M_i, M_q \leq M_l), \hat{f}^*(M_m)1(M_m \leq M_j, M_l \leq M_j)] \]
\[ = \frac{1}{kk} \sum_{u=1}^{kk} \sum_{v=1}^{kk} \text{COV}[\hat{f}^*(M_p | W_u)p_u1(M_p \leq M_i, M_q \leq M_l), \hat{f}^*(M_m | W_v)p_v1(M_m \leq M_j, M_l \leq M_j)] \]

Each term can be estimated as follows.

Let \( \Sigma = \text{COV}[\hat{\beta}]_{p \times p}, \) and let \( g_1(\beta) = \hat{f}^*(M_p | W_u)p_u1(M_p \leq M_i, M_q \leq M_l), g_2(\beta) = \hat{f}^*(M_m | W_v)p_v1(M_m \leq M_j, M_l \leq M_j) \) and \( g(\beta) = (g_1(\beta), g_2(\beta))^T. \) Let \( \nabla g = (\frac{\partial}{\partial \beta} g_1(\beta), \frac{\partial}{\partial \beta} g_2(\beta))^T, \) then

\[ \text{COV}[\hat{f}^*(M_p | W_u)p_u1(M_p \leq M_i, M_q \leq M_l), \hat{f}^*(M_m | W_v)p_v1(M_m \leq M_j, M_l \leq M_j)] \]
\[ \approx [\nabla g^T \Sigma \nabla g]_{1,2} \]
Example 2.1. If one wants to test $H_0$ : the data set is from a Poisson GLM with identity link ($\mu_Y = \beta^T Z$), then

$$g_1(\beta) = \frac{e^{-\beta^T W_u (\beta^T W_u)^{M_p} 1(M_p \leq M_i, M_q \leq M_j)}}{M_p!}$$

$$g_2(\beta) = \frac{e^{-\beta^T W_v (\beta^T W_v)^{M_m} 1(M_m \leq M_j, M_l \leq M_j)}}{M_m!}$$

$$\frac{\partial}{\partial \beta} g_1(\beta) = \frac{1}{M_p!} \left[ (-e^{-\beta^T W_u (\beta^T W_u)^{M_p} W_u} + e^{-\beta^T W_u M_p (\beta^T W_u)^{M_p-1}} 1(M_p \leq M_i, M_q \leq M_j)) \right]$$

$$\frac{\partial}{\partial \beta} g_2(\beta) = \frac{1}{M_m!} \left[ (-e^{-\beta^T W_v (\beta^T W_v)^{M_m} W_v} + e^{-\beta^T W_v M_m (\beta^T W_v)^{M_m-1}} 1(M_m \leq M_j, M_l \leq M_j)) \right]$$

$$\text{COV}[\hat{F}_{i-1}, \hat{F}_j] = \text{COV}[\frac{1}{k} \sum_{p=1}^k \hat{f}^*(M_p) 1(M_p \leq M_i), \frac{1}{k} \sum_{q=1}^k 1(M_q \leq M_{i-1})]$$

$$= \frac{1}{k^4} \sum_{p=1}^k \sum_{q=1}^k \sum_{m=1}^k \sum_{l=1}^k \text{COV}[\hat{f}^*(M_p) 1(M_p \leq M_i, M_q \leq M_{i-1}), \hat{f}^*(M_m) 1(M_m \leq M_j, M_l \leq M_j)]$$

Then

$$\text{COV}[\hat{f}^*(M_p) 1(M_p \leq M_i, M_q \leq M_{i-1}), \hat{f}^*(M_m) 1(M_m \leq M_j, M_l \leq M_j)]$$

$$= \frac{1}{kk^2} \sum_{u=1}^{kk} \sum_{v=1}^{kk} \text{COV}[\hat{f}^*(M_p | W_u) p_u 1(M_p \leq M_i, M_q \leq M_{i-1}), \hat{f}^*(M_m | W_v) p_v 1(M_m \leq M_j, M_l \leq M_j)]$$

Each term can be estimated using delta method as in the first term.
\[
\text{COV}[\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j, \hat{F}_{j-1}] = \text{COV}\left[\frac{1}{k} \sum_{p=1}^{k} \hat{f}^*(M_p)1(M_p \leq M_i) \frac{1}{k} \sum_{q=1}^{k} 1(M_q \leq M_{i-1})\right] \\
= \frac{1}{k^2} \sum_{p=1}^{k} \sum_{q=1}^{k} \sum_{m=1}^{k} \sum_{l=1}^{k} \text{COV}[\hat{f}^*(M_p)1(M_p \leq M_i, M_q \leq M_{i-1}), \hat{f}^*(M_m)1(M_m \leq M_j, M_l \leq M_{j-1})]
\]

Then

\[
\text{COV}[\hat{f}^*(M_p)1(M_p \leq M_i, M_q \leq M_{i-1}), \hat{f}^*(M_m)1(M_m \leq M_j, M_l \leq M_{j-1})] \\
= \frac{1}{k^2} \sum_{u=1}^{k} \sum_{v=1}^{k} \text{COV}[\hat{f}^*(M_p|W_u)p_u1(M_p \leq M_i, M_q \leq M_{i-1}), \hat{f}^*(M_m|W_v)p_v1(M_m \leq M_j, M_l \leq M_{j-1})]
\]

Each term can be estimated using delta method as in the first term.
Estimate $V_3$.

$$
\text{COV}\left[\sum_{i=1}^{k} \hat{F}(M_i), \hat{F}(M_i), \sum_{i=1}^{k} \hat{F}^*(M_i), \hat{F}(M_i)\right] = \sum_{i=1}^{k} \sum_{j=1}^{k} \text{COV}[\hat{F}(M_i), \hat{F}^*(M_j)] 
$$

$$
\text{COV}[\hat{F}(M_i), \hat{F}(M_j), \hat{F}^*(M_j)\hat{F}(M_j)]
= \text{COV}[\hat{F}_i, \hat{F}_i - \hat{F}_i, \hat{F}_j - \hat{F}_j]
= \text{COV}[\hat{F}_i, \hat{F}_i, \hat{F}_j] - \text{COV}[\hat{F}_i, \hat{F}_i, \hat{F}_j] - \text{COV}[\hat{F}_i, \hat{F}_j, \hat{F}_j - \hat{F}_j] + \text{COV}[\hat{F}_i, \hat{F}_j, \hat{F}_j - \hat{F}_j]
$$

Each term can be estimated as follows.

$$
\text{COV}[\hat{F}_i, \hat{F}_j] = \frac{1}{k} \sum_{p=1}^{k} \text{COV}[\hat{F}_i, \hat{F}_j, \hat{F}_j] 1(M_p \leq M_j) \hat{F}_j
$$

$$
= \frac{1}{k} \sum_{p=1}^{k} \sum_{u=1}^{k} \text{COV}[\hat{F}_i, \hat{F}_j, \hat{F}_j] 1(M_p \leq M_j) \hat{F}_j
$$

$$
\text{COV}[\hat{F}_i, \hat{F}_j, \hat{F}_j] 1(M_p \leq M_j) \hat{F}_j \approx [\nabla g^T \Sigma \nabla g]_{1,2}
$$

where $g_1(\hat{F}_i, \hat{F}_j, \beta) = \hat{F}_i^2$ and $g_2(\hat{F}_i, \hat{F}_j, \beta) = \hat{F}_j^2 (M_p \leq M_j) \hat{F}_j$, $g(\hat{F}_i, \hat{F}_j, \beta) = (g_1, g_2)^T$ and

$$
\nabla g = \begin{bmatrix}
\frac{\partial}{\partial \hat{F}_i} g_1 & \frac{\partial}{\partial \hat{F}_j} g_1 & \frac{\partial}{\partial \beta} g_1 \\
\frac{\partial}{\partial \hat{F}_i} g_2 & \frac{\partial}{\partial \hat{F}_j} g_2 & \frac{\partial}{\partial \beta} g_2
\end{bmatrix}
$$

$$
\Sigma = \begin{bmatrix}
\text{COV}[\hat{F}_i, \hat{F}_j]_{2 \times 2} & \begin{bmatrix} 0_{2 \times p} \end{bmatrix} \\
\begin{bmatrix} 0_{p \times 2} \end{bmatrix} & \text{COV}[\beta]_{p \times p}
\end{bmatrix}
$$
\[
\text{COV} \{\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j \} = \frac{1}{k} \sum_{p=1}^{k} \text{COV} \{\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j \} \approx \frac{1}{k} \sum_{p=1}^{k} \text{COV} \{\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j \} \approx [\nabla g^T \Sigma \nabla g]_{1,2}
\]

where \( g_1(\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j, \beta) = \hat{F}_i \hat{F}_{i-1} \) and \( g_2(\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j, \beta) = \hat{f}^*(M_p|Z_u) p_u 1(M_p \leq M_j) \hat{F}_j \), \( g(\hat{F}_i, \hat{F}_j, \beta) = (g_1, g_2)^T \) and

\[
\nabla g = \begin{bmatrix}
\frac{\partial}{\partial \hat{F}_i} g_1 & \frac{\partial}{\partial \hat{F}_{i-1}} g_1 & \frac{\partial}{\partial \hat{F}_j} g_1 & \frac{\partial}{\partial \beta} g_1 \\
\frac{\partial}{\partial \hat{F}_i} g_2 & \frac{\partial}{\partial \hat{F}_{i-1}} g_2 & \frac{\partial}{\partial \hat{F}_j} g_2 & \frac{\partial}{\partial \beta} g_2 \\
\end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix}
\text{COV} \{\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j \}_{3 \times 3} & \mathbf{0}_{3 \times p} \\
\mathbf{0}_{p \times 3} & \text{COV} \{\beta \}_{p \times p}
\end{bmatrix}
\]

\[
\text{COV} \{\hat{F}_i^2, \hat{F}_j^2 \} = \frac{1}{k} \sum_{p=1}^{k} \text{COV} \{\hat{F}_i^2, \hat{F}_j^2 \} \approx \frac{1}{k} \sum_{p=1}^{k} \text{COV} \{\hat{F}_i^2, \hat{F}_j^2 \} \approx [\nabla g^T \Sigma \nabla g]_{1,2}
\]
where \( g_1(\hat{F}_i, \hat{F}_{j-1}, \beta) = \hat{F}_i^2 \) and \( g_2(\hat{F}_i, \hat{F}_{j-1}, \beta) = \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j)\hat{F}_{j-1} \), \( g(\hat{F}_i, \hat{F}_{j-1}, \beta) = (g_1, g_2)^T \) and

\[
\nabla g = \begin{bmatrix}
\frac{\partial}{\partial \hat{F}_i} g_1 & \frac{\partial}{\partial \hat{F}_{j-1}} g_1 & \frac{\partial}{\partial \beta} g_1 \\
\frac{\partial}{\partial \hat{F}_i} g_2 & \frac{\partial}{\partial \hat{F}_{j-1}} g_2 & \frac{\partial}{\partial \beta} g_2 \\
\end{bmatrix}
= \begin{bmatrix}
2\hat{F}_i & 0 & 0 & \vdots & 0_{1	imes p} \\
0 & \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j) & \frac{\partial}{\partial \beta} \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j) & \vdots & 0_{2	imes p} \\
\end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix}
\text{COV}[\hat{F}_i, \hat{F}_{j-1}] & 0_{2	imes p} \\
0_{p	imes 2} & \text{COV}[\beta]_{p	imes p} \\
\end{bmatrix}
\]

\[
\text{COV}[\hat{F}_i, \hat{F}_{j-1}, \hat{F}_j] = \frac{1}{k} \sum_{p=1}^k \text{COV}[\hat{F}_i, \hat{F}_{j-1}, \hat{f}_p^*(M_p \leq M_j)\hat{F}_{j-1}] \\
= \frac{1}{k} \sum_{p=1}^k \sum_{u=1}^{k_k} \text{COV}[\hat{F}_i, \hat{F}_{j-1}, \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j)\hat{F}_{j-1}] \\
\]

\[
\text{COV}[\hat{F}_i, \hat{F}_{j-1}, \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j)\hat{F}_{j-1}] \approx [\nabla g^T \Sigma \nabla g]_{1,2}
\]

where \( g_1(\hat{F}_i, \hat{F}_{j-1}, \hat{F}_j, \hat{F}_{j-1}, \beta) = \hat{F}_i\hat{F}_{j-1} \) and \( g_2(\hat{F}_i, \hat{F}_{j-1}, \hat{F}_j, \hat{F}_{j-1}, \beta) = \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j)\hat{F}_{j-1} \), \( g(\hat{F}_i, \hat{F}_{j-1}, \hat{F}_j, \hat{F}_{j-1}, \beta) = (g_1, g_2)^T \) and

\[
\nabla g = \begin{bmatrix}
\frac{\partial}{\partial \hat{F}_i} g_1 & \frac{\partial}{\partial \hat{F}_{j-1}} g_1 & \frac{\partial}{\partial \beta} g_1 \\
\frac{\partial}{\partial \hat{F}_i} g_2 & \frac{\partial}{\partial \hat{F}_{j-1}} g_2 & \frac{\partial}{\partial \beta} g_2 \\
\end{bmatrix}
= \begin{bmatrix}
\hat{F}_{j-1} & \hat{F}_i & 0 & 0 & \vdots & 0_{1	imes p} \\
0 & \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j) & \frac{\partial}{\partial \beta} \hat{f}^*(M_p|Z_u)p_u1(M_p \leq M_j) & \vdots & 0_{2	imes p} \\
\end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix}
\text{COV}[\hat{F}_i, \hat{F}_{j-1}, \hat{F}_j, \hat{F}_{j-1}]_{4\times 4} & 0_{4	imes p} \\
0_{p\times 4} & \text{COV}[\beta]_{p\times p} \\
\end{bmatrix}
\]
When Y is continuous (as in Normal or Gamma model):

\[ T_2 = \int [\hat{F}(t) - \hat{F}^*(t)]d\hat{F}(t) = \frac{1}{n} \sum_{i=1}^{n} \hat{F}(M_i) - \frac{1}{n} \sum_{i=1}^{n} \hat{F}^*(M_i) = \frac{n+1}{2} - \frac{1}{n} \sum_{i=1}^{n} \hat{F}^*(M_i) \]

\[ n^2 \mathbb{V}[T_2] = \sum_{i=1}^{k} \sum_{j=1}^{k} \text{COV}[\hat{F}^*(M_i), \hat{F}^*(M_j)] \]

\[ \text{COV}[\hat{F}^*, \hat{F}^*_j] = \text{COV}[\frac{1}{k} \sum_{p=1}^{k} \hat{f}^*(M_p)1(M_p \leq M_i), \frac{1}{k} \sum_{m=1}^{k} \hat{f}^*(M_m)1(M_m \leq M_j)] \]

\[ = \frac{1}{k^2} \sum_{p=1}^{k} \sum_{m=1}^{k} \text{COV}[\hat{f}^*(M_p)1(M_p \leq M_i), \hat{f}^*(M_m)1(M_m \leq M_j)] \]

\[ \text{COV}[\hat{f}^*(M_p)1(M_p \leq M_i), \hat{f}^*(M_m)1(M_m \leq M_j)] = \frac{1}{kk} \sum_{u=1}^{kk} \sum_{v=1}^{kk} \text{COV}[\hat{f}^*(M_p|W_u)p_u1(M_p \leq M_i), \hat{f}^*(M_m|W_v)p_v1(M_m \leq M_j)] \]

\[ \text{COV}[\hat{f}^*(M_p|W_u)p_u1(M_p \leq M_i), \hat{f}^*(M_m|W_v)p_v1(M_m \leq M_j)] \text{ can be estimated using delta method as in the discrete case.} \]

Bibliography