Joint Distribution and Marginal Distribution Methods for Generalized Linear Model Supplementary Material

By Junyi Dong[a] and Qiqing Yu[a]

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 $(X, \mathscr{F}, P)$ 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,*

$\Omega_a = \{\omega \in X: \ \mu_n(\cdot,\omega) \to \mu(\cdot,\omega)$ *set-wisely*$\}$,

$\Omega_b = \{\omega \in X: \ f_n(t,\omega) \to f(t,\omega)$ *point-wisely in* $t\}$, *and*

$\Omega_c = \{\omega \in X: \ g_n(t,\omega) \to g(t,\omega)$ *point-wisely in* $t\}$.

If $P(\Omega_a \cap \Omega_b \cap \Omega_c) = 1$, $|f_n| \le g_n$, 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 $\omega \in \Omega$, $\mu_n(\cdot,\omega) \to \mu(\cdot,\omega)$ set-wisely, $f_n(t,\omega) \to f(t,\omega)$ point-wisely in t, and $f_n(t,\omega) \to f(t,\omega)$ point-wisely in t. Since $|f_n| \le g_n$ 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 Ω_0 be the event that $\hat{F}_{Y,\boldsymbol{Z}}(t,\boldsymbol{z}) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(Y_i \le t, \boldsymbol{Z} \le \boldsymbol{z}) \to F_{Y,\boldsymbol{Z}}(t,\boldsymbol{z})$. Let Ω_1 be the event that $\hat{F}_Y(t) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(Y_i \le t) \to F_Y(t)$. Then, by the strong law of large number (SLLN), $\mathbb{P}(\Omega_0) = 1$ and $\mathbb{P}(\Omega_1) = 1$.*

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 $\mathscr{L}(\boldsymbol{\beta},\phi)$ and $\tilde{\mathscr{L}}_n(\boldsymbol{\beta},\phi,\mathbf{Y})$ be defined as in (10) (Dong and Yu 2018).

By the SLLN and assumption (C6), $\tilde{\mathscr{L}}_n(\boldsymbol{\beta},\phi) \to \mathscr{L}(\boldsymbol{\beta},\phi)$ almost surely for each $(\boldsymbol{\beta},\phi) \in \mathbb{R}^{p+1}$. By assumption (C1), $B = \{(\boldsymbol{\beta}_0,\phi_0) : (\boldsymbol{\beta}_0,\phi_0) = \arg\sup_{\boldsymbol{\beta},\phi} \mathscr{L}(\boldsymbol{\theta},\phi)\}$ is a singleton set, thus $(\boldsymbol{\beta}_0,\phi_0) = \arg\sup_{\boldsymbol{\beta}\in\mathbb{R}^p,\phi\in\mathbb{R}} \mathscr{L}(\boldsymbol{\theta},\phi)$ is uniquely determined. $(\hat{\boldsymbol{\beta}}_n,\hat{\phi}_n) = (\hat{\boldsymbol{\beta}},\hat{\phi}) = \arg\sup_{\boldsymbol{\beta}\in\mathbb{R}^p,\phi\in\mathbb{R}} \tilde{\mathscr{L}}_n(\boldsymbol{\beta},\phi)$

yields $\mathscr{L}(\boldsymbol{\beta}_0,\phi_0) \geq \mathscr{L}(\boldsymbol{\beta},\phi)$ and $\tilde{\mathscr{L}}_n(\hat{\boldsymbol{\beta}},\hat{\phi}) \geq \tilde{\mathscr{L}}_n(\boldsymbol{\beta},\phi)$ for any $(\boldsymbol{\beta},\phi) \in \mathbb{R}^{p+1}$. Since $\tilde{\mathscr{L}}_n(\hat{\boldsymbol{\beta}},\hat{\phi}) \geq \tilde{\mathscr{L}}_n(\boldsymbol{\beta}_0,\phi_0)$,

$$\varliminf_{n\to\infty} \tilde{\mathscr{L}}_n(\hat{\boldsymbol{\beta}},\hat{\phi}) \geq \varliminf_{n\to\infty} \tilde{\mathscr{L}}_n(\boldsymbol{\beta}_0,\phi_0) = \mathscr{L}(\boldsymbol{\beta}_0,\phi_0) \quad \text{a.s.} \tag{1}$$

For each ω in Ω_0 (see Remark 2), let $(\boldsymbol{\beta}^*,\phi^*)$ be a limiting point of $(\hat{\boldsymbol{\beta}},\hat{\phi})$ such that there exists a subsequence $(\hat{\boldsymbol{\beta}}_{nl},\hat{\phi}_{nl})(\omega)$ converges to $(\boldsymbol{\beta}^*,\phi^*)$, that is, $(\hat{\boldsymbol{\beta}}_{nl},\hat{\phi}_{nl})(\omega) \to (\boldsymbol{\beta}^*,\phi^*)$. Let $f_n(\hat{\boldsymbol{\beta}}_{nl}(\omega),\hat{\phi}_{nl}(\omega)) = f_n(\omega) = \frac{Yh(\boldsymbol{\beta}_{nl}(\omega)^T\mathbf{Z})-k(\boldsymbol{\beta}_{nl}(\omega)^T\mathbf{Z})}{a(\phi_{nl}(\omega))} + c(Y,\phi_{nl}(\omega))$ and $f(\boldsymbol{\beta}^*,\phi^*) = f(\omega) = \frac{Yh((\boldsymbol{\beta}^*)^T\mathbf{Z})-k((\boldsymbol{\beta}^*)^T\mathbf{Z})}{a(\phi^*)} + c(Y,\phi^*)$. Then $f_n(\omega) \to f(\omega)$. We shall show that $\exists$ a function $g_n(\hat{\boldsymbol{\beta}}_{nl}(\omega),\hat{\phi}_{nl}(\omega)) = g_n(\omega)$ such that

$$\text{(a) } |f_n(\omega)| \leq g_n(\omega), \text{ (b) } g_n(\omega) \to g(\omega), \text{ and (c) } \int g_n(\omega)\,d\hat{F}_{Y,\mathbf{Z}}(t,\mathbf{z}) \to \int g(\omega)\,dF_{Y,\mathbf{Z}}(t,\mathbf{z}) < \infty, \tag{2}$$

then by Lemma 1, $\int f_n(\omega)\,d\hat{F}_{Y,\mathbf{Z}}(t,\mathbf{z}) \to \int f(\omega)\,dF_{Y,\mathbf{Z}}(t,\mathbf{z})$, that is, $\tilde{\mathscr{L}}_n(\hat{\boldsymbol{\beta}}_{nl}(\omega),\hat{\phi}_{nl}(\omega)) \to \mathscr{L}(\boldsymbol{\beta}^*,\phi^*)$. Since $\lim \tilde{\mathscr{L}}_n(\hat{\boldsymbol{\beta}}_{nl}(\omega),\hat{\phi}_{nl}(\omega)) \geq \varliminf_{n\to\infty} \tilde{\mathscr{L}}_n(\hat{\boldsymbol{\beta}}(\omega),\hat{\phi}(\omega)) \geq \mathscr{L}(\boldsymbol{\beta}_0,\phi_0)$, we have $\mathscr{L}(\boldsymbol{\beta}^*(\omega),\phi^*(\omega)) \geq \mathscr{L}(\boldsymbol{\beta}_0,\phi_0)$. Then $\mathscr{L}(\boldsymbol{\beta}^*(\omega),\phi^*(\omega)) = \mathscr{L}(\boldsymbol{\beta}_0,\phi_0)$ which implies $(\boldsymbol{\beta}^*(\omega),\phi^*(\omega)) = (\boldsymbol{\beta}_0,\phi_0)$ as B is a singleton set. Since every convergent subsequence of $(\hat{\boldsymbol{\beta}}(\omega),\hat{\phi}(\omega))$ converges to $(\boldsymbol{\beta}_0,\phi_0)$ for all $\omega \in \Omega_1$ (see Remark 2) and $\mathbb{P}(\Omega_1) = 1$, we have $(\hat{\boldsymbol{\beta}},\hat{\phi}) \xrightarrow{\text{a.s.}} (\boldsymbol{\beta}_0,\phi_0)$.

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 μ, $\phi = 1$, $a(\phi) = 1$, $\theta = \ln\mu$, $b(\theta) = \exp(\theta)$ and $c(y,\phi) = -\ln y!$. The canonical link function is $g(t) = \ln t$. Then $f_n(\omega) = Y(\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}) - \exp(\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}) - \ln Y!$. Under assumptions (C2) and (C4), $\hat{\boldsymbol{\beta}}_{nl}(\omega)$ is bounded and $\mathbf{Z}$ is bounded. It follows that $||\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}|| \leq K$. Then $|f_n(\omega)| \leq K|Y| + e^K + \ln Y! = g_n = g$, then $g_n \to g$ and $\int g_n(\omega)\,d\hat{F}_{Y,\mathbf{Z}}(t,\mathbf{z}) \to \int g(\omega)\,dF_{Y,\mathbf{Z}}(t,\mathbf{z}) = K\mathbb{E}[|Y|] + e^K + \mathbb{E}[\ln Y!] < \infty$ 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 $c(y,\phi) = \ln\binom{m}{my}$. The canonical link function is g(t) = $\ln\frac{t}{1-t}$. Let $f_n(\omega) = [Y(\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}) - \ln(1+\exp(\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}))]m + \ln\binom{m}{mY}$. By assumptions (C2) and (C4), we can assume that $||\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}|| \leq K$, and then $|f_n(\omega)| \leq [K|Y| + \ln(1+e^K)]m + \ln\binom{m}{mY} = g_n = g$. Then $g_n \to$

g and $\int g_n(\omega)d\hat{F}_{Y,\mathbf{Z}}(t,\mathbf{z}) \to \int g(\omega)dF_{Y,\mathbf{Z}}(t,\mathbf{z}) = [K\mathbb{E}[|Y|] + \ln(1+e^K)]m + \mathbb{E}[\ln\binom{m}{mY}] < \infty$, as all expectations 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 $c(y,\phi) = -\ln\Gamma(\alpha) + (\alpha-1)\ln y + \alpha\ln\alpha$. The canonical link function is $g(t) = -1/t$. Let $f_n(\omega) = \alpha[Y(\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}) + \ln(-(\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}))] - \ln\Gamma(\alpha) + (\alpha-1)\ln Y + \alpha\ln\alpha$. Since $\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}$ is bounded above and bounded below by assumption (C5), there exists some M_1 and M_2 such that $|\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z})| < M_1$ and $|\ln(-(\hat{\boldsymbol{\beta}}_{nl}(\omega)^T\mathbf{Z}))| < M_2$. Thus $|f_n(\omega)| \le \alpha M_1|Y| + M_2 - \ln\Gamma(\alpha) + |\alpha-1||\ln Y| + \alpha\ln\alpha = g_n = g$, $g_n \to g$ and $\int g_n(\omega)d\hat{F}_{Y,\mathbf{Z}}(t,\mathbf{z}) \to \int g(\omega)dF_{Y,\mathbf{Z}}(t,\mathbf{z}) < \infty$ by the assumption that all expectations exist.

Proof of Statement 2. Let $\omega \in \Omega_1$ (see Remark 2), $f_n = f = \mathbf{1}(Y \le t)f(Y,\boldsymbol{\beta},\phi|\mathbf{z})$, then $|f_n| \le f(Y,\boldsymbol{\beta},\phi|\mathbf{z}) = g_n = g$. Since $\int g d\hat{F}_Y(\omega) = 1 < \infty$, by Lemma 1, since $\mathbb{P}(\Omega_1) = 1$, $\hat{F}_{Y^*|\mathbf{Z}}(t;\boldsymbol{\beta},\phi|\mathbf{z}) = \int f_n d\hat{F}_Y(\omega) \to \int f dF_Y = F_{Y^*|\mathbf{Z}}(t;\boldsymbol{\beta},\phi|\mathbf{z})$ almost surely.

Proof of Statement 3. Let Ω_2 be the event that $(\hat{\boldsymbol{\beta}},\hat{\phi}) \to (\boldsymbol{\beta},\phi)$, then by Proof of Statement 1, $\mathbb{P}(\Omega_2) = 1$. Notice that $\hat{F}_{Y^*|\mathbf{Z}}(t;\hat{\boldsymbol{\beta}},\hat{\phi}|\mathbf{z}) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(Y_i \le t)\hat{f}_{Y^*|\mathbf{Z}}(Y_i;\hat{\boldsymbol{\beta}},\hat{\phi}|\mathbf{z})$. Let $\omega \in \Omega_1 \cap \Omega_2$, let $f_n(\omega) = \mathbf{1}(Y \le t)\hat{f}_{Y^*|\mathbf{Z}}(Y;\hat{\boldsymbol{\beta}},\hat{\phi}|\mathbf{z})(\omega)$ and $f = \mathbf{1}(Y \le t)f_{Y^*|\mathbf{Z}}(Y;\boldsymbol{\beta},\phi|\mathbf{z})$. By the similar argument as in the Proof of Statement 1, there exists $g_n(\omega)$ such that $|f_n(\omega)| \le g_n(\omega)$, $g_n(\omega) \to g$ and $\int g d\hat{F}_Y < \infty$.

Since $\hat{f}_{Y^*|\mathbf{Z}}(Y;\hat{\boldsymbol{\beta}},\hat{\phi})(\omega)$ is a continuous function of $(\hat{\boldsymbol{\beta}},\hat{\phi})(\omega)$ and we've shown that $(\hat{\boldsymbol{\beta}},\hat{\phi})(\omega) \to (\boldsymbol{\beta}_0,\phi_0)$, then $\hat{f}_{Y^*|\mathbf{Z}}(Y;\hat{\boldsymbol{\beta}},\hat{\phi}|\mathbf{z})(\omega)$ converges to $f_{Y^*|\mathbf{Z}}(Y;\boldsymbol{\beta},\phi|\mathbf{z})$ and $f_n \to f$. By Lemma 1, since $\mathbb{P}(\Omega_1 \cap \Omega_2) = 1$, $\hat{F}_{Y^*|\mathbf{Z}}(t;\hat{\boldsymbol{\beta}},\hat{\phi}|\mathbf{z}) \xrightarrow{\text{a.s.}} F_{Y^*|\mathbf{Z}}(t;\boldsymbol{\beta},\phi|\mathbf{z})$.

Proof of Theorem 10 in Dong and Yu (Dong and Yu 2018).

Let $F_{\mathbf{Z}}(\mathbf{s}) = \mathbb{P}(\mathbf{Z} \le \mathbf{s})$ be the CDF of $\mathbf{Z}$, let $\hat{F}_{\mathbf{Z}}(\mathbf{s}) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(\mathbf{Z}_i \le \mathbf{s})$ be the empirical CDF, and let Ω_3 be the event such that $\hat{F}_{\mathbf{Z}}(\mathbf{s}) \to F_{\mathbf{Z}}(\mathbf{s})$, then by the SLLN, $\mathbb{P}(\Omega_3) = 1$. Let $\omega \in \Omega_1 \cap \Omega_2 \cap \Omega_3$, where Ω_2 is as defined in Proof of Lemma 8 in Dong and Yu (Dong and Yu 2018).

To prove (11) in Dong and Yu (Dong and Yu 2018), let $f_n(\omega) = \hat{F}_{Y^*|\mathbf{Z}}(Y;\boldsymbol{\beta},\phi|\mathbf{Z})(\omega)$ and $f = F_{Y^*|\mathbf{Z}}(Y;\boldsymbol{\beta},\phi|\mathbf{Z})$, by Lemma 8, $f_n(\omega) \to f$. Since $|f_n| \le g_n = g = 1$ and $\int g dF_{\mathbf{Z}} = 1$, then by Lemma 8 in Dong and Yu (Dong and Yu 2018), $\int f_n(\omega)d\hat{F}_{\mathbf{Z}}(\omega) \to \int f dF_{\mathbf{Z}}$. Since $\mathbb{P}(\Omega_1 \cap \Omega_2 \cap \Omega_3) = 1$, (11) in Dong and Yu (Dong and Yu 2018) follows.

To prove (12) in Dong and Yu (Dong and Yu 2018), let $f_n(\omega) = \hat{F}_{Y^*|\mathbf{Z}}(Y;\hat{\boldsymbol{\beta}},\hat{\phi}|\mathbf{Z})(\omega)$ and $f = F_{Y^*|\mathbf{Z}}(Y;\boldsymbol{\beta},\phi|\mathbf{Z})$, by Lemma 8 in Dong and Yu (Dong and Yu 2018), $f_n(\omega) \to f$. Since $|f_n| \le$

$g_n = g = 1$ and $\int g dF_{\mathbf{Z}} = 1$, then by Lemma 1, $\int f_n(\omega) d\hat{F}_{\mathbf{Z}}(\omega) \to \int f dF_{\mathbf{Z}}$. Since $\mathbb{P}(\Omega_1 \cap \Omega_2 \cap \Omega_3) = 1$, (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):

$$\begin{aligned}
\mathbb{V}[T_2] = \mathbb{V}[\sum_{i=1}^k \hat{F}(M_i)\hat{f}(M_i)] + \mathbb{V}[\sum_{i=1}^k \hat{F}^*(M_i)\hat{f}(M_i)] \\
- 2\mathbb{COV}[\sum_{i=1}^k \hat{F}(M_i)\hat{f}(M_i), \sum_{i=1}^k \hat{F}^*(M_i)\hat{f}(M_i)] \\
= \sum_{i=1}^n \sum_{j=1}^n \mathbb{COV}[\hat{F}(M_i)\hat{f}(M_i), \hat{F}(M_j)\hat{f}(M_j)] \qquad (=\mathbb{V}_1) \\
+ \sum_{i=1}^n \sum_{j=1}^n \mathbb{COV}[\hat{F}^*(M_i)\hat{f}(M_i), \hat{F}^*(M_j)\hat{f}(M_j)] \qquad (=\mathbb{V}_2) \\
- 2\sum_{i=1}^n \sum_{j=1}^n \mathbb{COV}[\hat{F}(M_i)\hat{f}(M_i), \hat{F}^*(M_j)\hat{f}(M_j)] \qquad (=\mathbb{V}_3)
\end{aligned}$$

Estimate $\mathbb{V}_1$.

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 $\hat{f}(M_j) = \hat{F}_j - \hat{F}_{j-1}$.

$$\begin{aligned}
&\mathbb{COV}[\hat{F}(M_i)\hat{f}(M_i), \hat{F}(M_j)\hat{f}(M_j)] \\
&= \mathbb{COV}[\hat{F}(M_i)(\hat{F}_i - \hat{F}_{i-1}), \hat{F}(M_j)(\hat{F}_j - \hat{F}_{j-1})] \\
&= \mathbb{COV}[\hat{F}_i^2, \hat{F}_j^2] - \mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{F}_j^2] - \mathbb{COV}[\hat{F}_i^2, \hat{F}_j\hat{F}_{j-1}] + \mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{F}_j\hat{F}_{j-1}]
\end{aligned}$$

Each term can be estimated as follows.

$$
\begin{aligned}
\mathbb{COV}[\hat{F}_i^2, \hat{F}_j^2] &= \mathbb{E}[\hat{F}_i^2 \hat{F}_j^2] - \mathbb{E}[\hat{F}_i^2]\mathbb{E}[\hat{F}_j^2] \\
\mathbb{E}[\hat{F}_i^2 \hat{F}_j^2] &= \mathbb{E}[(\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}(Y_k \le M_i))^2 (\frac{1}{n}\sum_{l=1}^{n} \mathbf{1}(Y_l \le M_i))^2] \\
&= \frac{1}{n^4}\sum_{k=1}^{n}\sum_{p=1}^{n}\sum_{l=1}^{n}\sum_{q=1}^{n} \mathbb{P}(Y_k \le M_i, Y_p \le M_i, Y_l \le M_j, Y_q \le M_j) \\
&= \frac{1}{n^4}[(F_i F_j^2 + 4\min(F_i, F_j)F_i F_j + F_i^2 F_j)n(n-1)(n-2) + (2\min(F_i, F_j)F_i \\
&\quad + 2\min(F_i, F_j)F_j)n(n-1) + n\min(F_i, F_j) + F_i^2 F_j^2 n(n-1)(n-2)(n-3) \\
&\quad + (F_i F_j + 2\min(F_i, F_j)^2)n(n-1)] \\
\mathbb{E}[\hat{F}_i^2] &= \mathbb{E}[(\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}(Y_k \le M_i))^2] \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} \mathbb{P}(Y_k \le M_i, Y_l \le M_i) \\
&= \frac{1}{n^2}[nF_i + n(n-1)F_i^2]
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\hat{F}_j^2] &= \mathbb{E}[(\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}(Y_k \le M_j))^2] \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} \mathbb{P}(Y_k \le M_j, Y_l \le M_j) \\
&= \frac{1}{n^2}[nF_j + n(n-1)F_j^2] \\
\mathbb{COV}[\hat{F}_i^2, \hat{F}_j^2] &\approx \frac{1}{n^4}[\hat{F}_i \hat{F}_j^2 + 4\min(\hat{F}_i, \hat{F}_j)\hat{F}_i \hat{F}_j + \hat{F}_i^2 \hat{F}_j)n(n-1)(n-2) + (2\min(\hat{F}_i, \hat{F}_j)\hat{F}_i \\
&\quad + 2\min(\hat{F}_i, \hat{F}_j)\hat{F}_j)n(n-1) + n\min(\hat{F}_i, \hat{F}_j) + \hat{F}_i^2 \hat{F}_j^2 n(n-1)(n-2)(n-3) \\
&\quad + (\hat{F}_i \hat{F}_j + 2\min(\hat{F}_i, \hat{F}_j)^2)n(n-1) - (n\hat{F}_i + n(n-1)\hat{F}_i^2)(n\hat{F}_j + n(n-1)\hat{F}_j^2)]
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{F}_j^2] &= \mathbb{E}[\hat{F}_i\hat{F}_{i-1}\hat{F}_j^2] - \mathbb{E}[\hat{F}_i\hat{F}_{i-1}]\mathbb{E}[\hat{F}_j^2] \\
\mathbb{E}[\hat{F}_i\hat{F}_{i-1}\hat{F}_j^2] &= \mathbb{E}[\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}(Y_k \le M_i)\frac{1}{n}\sum_{l=1}^{n}\mathbf{1}(Y_l \le M_{i-1})(\frac{1}{n}\sum_{p=1}^{n}\mathbf{1}(Y_p \le M_j))^2] \\
&= \frac{1}{n^4}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{p=1}^{n}\sum_{q=1}^{n}\mathbb{P}(Y_k \le M_i, Y_l \le M_{i-1}, Y_p \le M_j, Y_q \le M_j) \\
&= \frac{1}{n^4}[(F_{i-1}F_j^2 + 2\min(F_i, F_j)F_{i-1}F_j + 2\min(F_{i-1}, F_j)F_iF_j + F_iF_{i-1}F_j) \\
&\quad n(n-1)(n-2) + (2\min(F_{i-1}, F_j)F_j + \min(F_i, F_j)F_{i-1} + \min(F_{i-1}, F_j)F_i) \\
&\quad n(n-1) + \min(F_{i-1}, F_j)n + F_iF_{i-1}F_j^2 n(n-1)(n-2)(n-3) + \\
&\quad (2\min(F_{i-1}, F_j)\min(F_i, F_j) + \min(F_{i-1}, F_i)F_j)n(n-1)]
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\hat{F}_i\hat{F}_{i-1}] &= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathbb{P}(Y_k \le M_i, Y_l \le M_{i-1}) \\
&= \frac{1}{n^2}[n\min(F_{i-1}, F_i) + n(n-1)F_iF_{i-1}] \\
\mathbb{E}[\hat{F}_j^2] &= \mathbb{E}[(\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}(Y_k \le M_j))^2] \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathbb{P}(Y_k \le M_j, Y_l \le M_j) \\
&= \frac{1}{n^2}[nF_j + n(n-1)F_j^2] \\
\mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{F}_j^2] &\approx \frac{1}{n^4}[(\hat{F}_{i-1}\hat{F}_j^2 + 2\min(\hat{F}_i, \hat{F}_j)\hat{F}_{i-1}\hat{F}_j + 2\min(\hat{F}_{i-1}, \hat{F}_j)\hat{F}_i\hat{F}_j + \hat{F}_i\hat{F}_{i-1}\hat{F}_j) \\
&\quad n(n-1)(n-2) + (2\min(\hat{F}_{i-1}, \hat{F}_j)\hat{F}_j + \min(\hat{F}_i, \hat{F}_j)\hat{F}_{i-1} + \min(\hat{F}_{i-1}, \hat{F}_j)\hat{F}_i) \\
&\quad n(n-1) + \min(\hat{F}_{i-1}, \hat{F}_j)n + \hat{F}_i\hat{F}_{i-1}\hat{F}_j^2 n(n-1)(n-2)(n-3) + \\
&\quad (2\min(\hat{F}_{i-1}, \hat{F}_j)\min(\hat{F}_i, \hat{F}_j) + \min(\hat{F}_{i-1}, \hat{F}_i)\hat{F}_j)n(n-1) - \\
&\quad (n\min(\hat{F}_{i-1}, \hat{F}_i) + n(n-1)\hat{F}_i\hat{F}_{i-1})(n\hat{F}_j + n(n-1)\hat{F}_j^2)]
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{COV}[\hat{F}_j\hat{F}_{j-1},\hat{F}_i^2] \approx \frac{1}{n^4}[(&\hat{F}_{j-1}\hat{F}_i^2 + 2\min(\hat{F}_j,\hat{F}_i)\hat{F}_{j-1}\hat{F}_i + 2\min(\hat{F}_{j-1},\hat{F}_i)\hat{F}_j\hat{F}_i + \hat{F}_j\hat{F}_{j-1}\hat{F}_i) \\
&n(n-1)(n-2) + (2\min(\hat{F}_{j-1},\hat{F}_i)\hat{F}_i + \min(\hat{F}_j,\hat{F}_i)\hat{F}_{j-1} + \min(\hat{F}_{j-1},\hat{F}_i)\hat{F}_j) \\
&n(n-1) + \min(\hat{F}_{j-1},\hat{F}_i)n + \hat{F}_j\hat{F}_{j-1}\hat{F}_i^2 n(n-1)(n-2)(n-3)+ \\
&(2\min(\hat{F}_{j-1},\hat{F}_i)\min(\hat{F}_j,\hat{F}_i) + \min(\hat{F}_{j-1},\hat{F}_j)\hat{F}_i)n(n-1) \\
&-(n\min(\hat{F}_{j-1},\hat{F}_j) + n(n-1)\hat{F}_j\hat{F}_{j-1})(n\hat{F}_i + n(n-1)\hat{F}_i^2)]
\end{aligned}
$$

$$\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{F}_j\hat{F}_{j-1}] = \mathbb{E}[\hat{F}_i\hat{F}_{i-1}\hat{F}_j\hat{F}_{j-1}] - \mathbb{E}[\hat{F}_i\hat{F}_{i-1}]\mathbb{E}[\hat{F}_j\hat{F}_{j-1}]$$

$$\begin{aligned}
\mathbb{E}[\hat{F}_i\hat{F}_{i-1}\hat{F}_j\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 \le M_i, Y_l \le M_{i-1}, Y_p \le M_j, Y_q \le M_{j-1}) \\
&= \frac{1}{n^4}[n(n-1)(n-2)(F_{i-1}F_jF_{j-1} + \min(F_i,F_j)F_{i-1}F_{j-1} + \\
&\min(F_i,F_{j-1})F_{i-1}F_j + \min(F_{i-1},F_j)F_iF_{j-1} + \min(F_{i-1},F_{j-1})F_iF_j + \\
&F_{j-1}F_iF_{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)(n-2)(n-3)F_iF_{i-1}F_jF_{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,F_{j-1})\min(F_{i-1},F_j))] \\
\mathbb{E}[\hat{F}_i\hat{F}_{i-1}] &= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathbb{P}(Y_k \le M_i, Y_l \le M_{i-1}) \\
&= \frac{1}{n^2}[n\min(F_{i-1},F_i) + n(n-1)F_iF_{i-1}] \\
\mathbb{E}[\hat{F}_j\hat{F}_{j-1}] &= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathbb{P}(Y_k \le M_j, Y_l \le M_{j-1}) \\
&= \frac{1}{n^2}[n\min(F_{j-1},F_j) + n(n-1)F_jF_{j-1}] \\
\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{F}_j\hat{F}_{j-1}] &\approx \frac{1}{n^4}[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,\hat{F}_{j-1})\hat{F}_{i-1}\hat{F}_j + \min(\hat{F}_{i-1},\hat{F}_j)\hat{F}_i\hat{F}_{j-1} + \min(\hat{F}_{i-1},\hat{F}_{j-1})\hat{F}_i\hat{F}_j + \\
&\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)(n-2)(n-3)\hat{F}_i\hat{F}_{i-1}\hat{F}_j\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-1})\min(\hat{F}_{i-1},\hat{F}_j)) \\
&-(n\min(\hat{F}_{i-1},\hat{F}_i) + n(n-1)\hat{F}_i\hat{F}_{i-1})(n\min(\hat{F}_{j-1},\hat{F}_j) + n(n-1)\hat{F}_j\hat{F}_{j-1})]
\end{aligned}$$

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 $\hat{f}(M_j) = \hat{F}_j - \hat{F}_{j-1}$.

$$
\begin{aligned}
&\mathbb{COV}[\hat{F}^*(M_i)\hat{f}(M_i), \hat{F}^*(M_j)\hat{f}(M_j)] \\
&= \mathbb{COV}[\hat{F}^*(M_i)(\hat{F}_i - \hat{F}_{i-1}), \hat{F}^*(M_j)(\hat{F}_j - \hat{F}_{j-1})] \\
&= \mathbb{COV}[\hat{F}_i^*\hat{F}_i, \hat{F}_j^*\hat{F}_j] - \mathbb{COV}[\hat{F}_i^*\hat{F}_{i-1}, \hat{F}_j^*\hat{F}_j] - \mathbb{COV}[\hat{F}_i^*\hat{F}_i, \hat{F}_j^*\hat{F}_{j-1}] + \mathbb{COV}[\hat{F}_i^*\hat{F}_{i-1}, \hat{F}_j^*\hat{F}_{j-1}]
\end{aligned}
$$

Each term can be estimated as follows.

$$
\begin{aligned}
\mathbb{COV}[(\hat{F}_i^*)^2, (\hat{F}_j^*)^2] &= \mathbb{COV}[(\frac{1}{k}\sum_{p=1}^{k}\hat{f}^*(M_p)\mathbf{1}(M_p \le M_i))(\frac{1}{k}\sum_{q=1}^{k}\mathbf{1}(M_q \le M_i)), \\
&\qquad (\frac{1}{k}\sum_{m=1}^{k}\hat{f}^*(M_m)\mathbf{1}(M_m \le M_j))(\frac{1}{k}\sum_{l=1}^{k}\mathbf{1}(M_l \le M_j))] \\
&= \frac{1}{k^4}\sum_{p=1}^{k}\sum_{q=1}^{k}\sum_{m=1}^{k}\sum_{l=1}^{k}\mathbb{COV}[\hat{f}^*(M_p)\mathbf{1}(M_p \le M_i, M_q \le M_i), \\
&\qquad \hat{f}^*(M_m)\mathbf{1}(M_m \le M_j, M_l \le M_j)]
\end{aligned}
$$

Let $\mathbf{Z}_i \in \{\mathbf{W}_1, ..., \mathbf{W}_{kk}\}$, i=1,...,n. Let $p_j = \hat{f}^*_{\mathbf{Z}}(\mathbf{W}_j) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(\mathbf{Z}_i = \mathbf{W}_j)$. Then

$$
\begin{aligned}
&\hat{f}^*(M_p) = \frac{1}{kk}\sum_{l=1}^{kk}\hat{f}^*(M_p|\mathbf{W}_l)p_l \\
&\mathbb{COV}[\hat{f}^*(M_p)\mathbf{1}(M_p \le M_i, M_q \le M_i), \hat{f}^*(M_m)\mathbf{1}(M_m \le M_j, M_l \le M_j)] \\
&= \frac{1}{kk^2}\sum_{u=1}^{kk}\sum_{v=1}^{kk}\mathbb{COV}[\hat{f}^*(M_p|\mathbf{W}_u)p_u\mathbf{1}(M_p \le M_i, M_q \le M_i), \\
&\qquad \hat{f}^*(M_m|\mathbf{W}_v)p_v\mathbf{1}(M_m \le M_j, M_l \le M_j)]
\end{aligned}
$$

Each term can be estimated as follows.

Let $\Sigma = \mathbb{COV}[\hat{\boldsymbol{\beta}}]_{p\times p}$, and let $g_1(\boldsymbol{\beta}) = \hat{f}^*(M_p|\mathbf{W}_u)p_u\mathbf{1}(M_p \le M_i, M_q \le M_i)$, $g_2(\boldsymbol{\beta}) = \hat{f}^*(M_m|\mathbf{W}_v)p_v\mathbf{1}(M_m \le M_j, M_l \le M_j)$ and $g(\boldsymbol{\beta}) = (g_1(\boldsymbol{\beta}), g_2(\boldsymbol{\beta}))^T$. Let $\bigtriangledown g = (\frac{\partial}{\partial\boldsymbol{\beta}}g_1(\boldsymbol{\beta}), \frac{\partial}{\partial\boldsymbol{\beta}}g_2(\boldsymbol{\beta}))^T$, then

$$
\begin{aligned}
&\mathbb{COV}[\hat{f}^*(M_p|\mathbf{W}_u)p_u\mathbf{1}(M_p \le M_i, M_q \le M_i), \hat{f}^*(M_m|\mathbf{W}_v)p_v\mathbf{1}(M_m \le M_j, M_l \le M_j)] \\
&\approx [\bigtriangledown g^T \Sigma \bigtriangledown g]_{1,2}
\end{aligned}
$$

Example 2.1. *If one wants to test H_0 : the data set is from a Poisson GLM with identity link ($\mu_Y = \boldsymbol{\beta}^T \boldsymbol{Z}$), then*

$$
\begin{aligned}
g_1(\boldsymbol{\beta}) &= \frac{e^{-\boldsymbol{\beta}^T \boldsymbol{W}_u} (\boldsymbol{\beta}^T \boldsymbol{W}_u)^{M_p} \mathbf{1}(M_p \le M_i, M_q \le M_i)}{M_p!} \\
g_2(\boldsymbol{\beta}) &= \frac{e^{-\boldsymbol{\beta}^T \boldsymbol{W}_v} (\boldsymbol{\beta}^T \boldsymbol{W}_v)^{M_m} \mathbf{1}(M_m \le M_j, M_l \le M_j)}{M_m!} \\
\frac{\partial}{\partial \boldsymbol{\beta}} g_1(\boldsymbol{\beta}) &= \frac{1}{M_p!}[(-e^{-\boldsymbol{\beta}^T \boldsymbol{W}_u} (\boldsymbol{\beta}^T \boldsymbol{W}_u)^{M_p} \boldsymbol{W}_u + e^{-\boldsymbol{\beta}^T \boldsymbol{W}_u} M_p (\boldsymbol{\beta}^T \boldsymbol{W}_u)^{M_p - 1}) \mathbf{1}(M_p \le M_i, M_q \le M_i)] \\
\frac{\partial}{\partial \boldsymbol{\beta}} g_2(\boldsymbol{\beta}) &= \frac{1}{M_m!}[(-e^{-\boldsymbol{\beta}^T \boldsymbol{W}_v} (\boldsymbol{\beta}^T \boldsymbol{W}_v)^{M_m} \boldsymbol{W}_v + e^{-\boldsymbol{\beta}^T \boldsymbol{W}_v} M_m (\boldsymbol{\beta}^T \boldsymbol{W}_v)^{M_m - 1}) \mathbf{1}(M_m \le M_j, M_l \le M_j)]
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{COV}[\hat{F}_i^* \hat{F}_{i-1}, \hat{F}_j^* \hat{F}_j] &= \mathbb{COV}[\frac{1}{k} \sum_{p=1}^{k} \hat{f}^*(M_p) \mathbf{1}(M_p \le M_i) \frac{1}{k} \sum_{q=1}^{k} \mathbf{1}(M_q \le M_{i-1}), \\
&\qquad \frac{1}{k} \sum_{m=1}^{k} \hat{f}^*(M_m) \mathbf{1}(M_m \le M_j) \frac{1}{k} \sum_{l=1}^{k} \mathbf{1}(M_l \le M_j)] \\
&= \frac{1}{k^4} \sum_{p=1}^{k} \sum_{q=1}^{k} \sum_{m=1}^{k} \sum_{l=1}^{k} \mathbb{COV}[\hat{f}^*(M_p) \mathbf{1}(M_p \le M_i, M_q \le M_{i-1}), \\
&\qquad \hat{f}^*(M_m) \mathbf{1}(M_m \le M_j, M_l \le M_j)]
\end{aligned}
$$

Then

$$
\begin{aligned}
&\mathbb{COV}[\hat{f}^*(M_p) \mathbf{1}(M_p \le M_i, M_q \le M_{i-1}), \hat{f}^*(M_m) \mathbf{1}(M_m \le M_j, M_l \le M_j)] \\
&= \frac{1}{kk^2} \sum_{u=1}^{kk} \sum_{v=1}^{kk} \mathbb{COV}[\hat{f}^*(M_p | \mathbf{W}_u) p_u \mathbf{1}(M_p \le M_i, M_q \le M_{i-1}), \\
&\qquad \hat{f}^*(M_m | \mathbf{W}_v) p_v \mathbf{1}(M_m \le M_j, M_l \le M_j)]
\end{aligned}
$$

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

$$\mathbb{COV}[\hat{F}_i^* \hat{F}_{i-1}, \hat{F}_j^* \hat{F}_{j-1}] = \mathbb{COV}[\frac{1}{k}\sum_{p=1}^{k} \hat{f}^*(M_p)\mathbf{1}(M_p \le M_i)\frac{1}{k}\sum_{q=1}^{k}\mathbf{1}(M_q \le M_{i-1}),$$
$$\frac{1}{k}\sum_{m=1}^{k} \hat{f}^*(M_m)\mathbf{1}(M_m \le M_i)\frac{1}{k}\sum_{l=1}^{k}\mathbf{1}(M_l \le M_{j-1})]$$
$$= \frac{1}{k^4}\sum_{p=1}^{k}\sum_{q=1}^{k}\sum_{m=1}^{k}\sum_{l=1}^{k}\mathbb{COV}[\hat{f}^*(M_p)\mathbf{1}(M_p \le M_i, M_q \le M_{i-1}),$$
$$\hat{f}^*(M_m)\mathbf{1}(M_m \le M_j, M_l \le M_{j-1})]$$

Then

$$\mathbb{COV}[\hat{f}^*(M_p)\mathbf{1}(M_p \le M_i, M_q \le M_{i-1}), \hat{f}^*(M_m)\mathbf{1}(M_m \le M_j, M_l \le M_{j-1})]$$
$$= \frac{1}{kk^2}\sum_{u=1}^{kk}\sum_{v=1}^{kk}\mathbb{COV}[\hat{f}^*(M_p|\mathbf{W}_u)p_u\mathbf{1}(M_p \le M_i, M_q \le M_{i-1}),$$
$$\hat{f}^*(M_m|\mathbf{W}_v)p_v\mathbf{1}(M_m \le M_j, M_l \le M_{j-1})]$$

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

Estimate $\mathbb{V}_3$.

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

$$\begin{aligned}
&\mathbb{COV}[\hat{F}(M_i)\hat{f}(M_i),\hat{F}^*(M_j)\hat{f}(M_j)]\\
&=\mathbb{COV}[\hat{F}_i(\hat{F}_i-\hat{F}_{i-1}),\hat{F}_j^*(\hat{F}_j-\hat{F}_{j-1})]\\
&=\mathbb{COV}[\hat{F}_i^2,\hat{F}_j^*\hat{F}_j]-\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{F}_j^*\hat{F}_j]-\mathbb{COV}[\hat{F}_i^2,\hat{F}_j^*\hat{F}_{j-1}]+\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{F}_j^*\hat{F}_{j-1}]
\end{aligned}$$

Each term can be estimated as follows.

$$\begin{aligned}
\mathbb{COV}[\hat{F}_i^2,\hat{F}_j^*\hat{F}_j]&=\frac{1}{k}\sum_{p=1}^{k}\mathbb{COV}[\hat{F}_i^2,\hat{f}_p^*\mathbf{1}(M_p\le M_j)\hat{F}_j]\\
&=\frac{1}{k}\sum_{p=1}^{k}\sum_{u=1}^{kk}\mathbb{COV}[\hat{F}_i^2,\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j]
\end{aligned}$$

$$\mathbb{COV}[\hat{F}_i^2,\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j]\approx[\bigtriangledown g^T\Sigma\bigtriangledown g]_{1,2}$$

where $g_1(\hat{F}_i,\hat{F}_j,\boldsymbol{\beta})=\hat{F}_i^2$ and $g_2(\hat{F}_i,\hat{F}_j,\boldsymbol{\beta})=\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j$, $g(\hat{F}_i,\hat{F}_j,\boldsymbol{\beta})=(g_1,g_2)^T$ and

$$\begin{aligned}
\bigtriangledown g&=\begin{bmatrix}\frac{\partial}{\partial\hat{F}_i}g_1 & \frac{\partial}{\partial\hat{F}_j}g_1 & \frac{\partial}{\partial\hat{\boldsymbol{\beta}}}g_1\\ \frac{\partial}{\partial\hat{F}_i}g_2 & \frac{\partial}{\partial\hat{F}_j}g_2 & \frac{\partial}{\partial\hat{\boldsymbol{\beta}}}g_2\end{bmatrix}\\
&=\begin{bmatrix}2\hat{F}_i & 0 & \mathbf{0}_{1\times p}\\ 0 & \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j) & \frac{\partial}{\partial\hat{\boldsymbol{\beta}}}\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j\end{bmatrix}\\
\Sigma&=\begin{bmatrix}\mathbb{COV}[\hat{F}_i,\hat{F}_j]_{2\times2} & \mathbf{0}_{2\times p}\\ \mathbf{0}_{p\times2} & \mathbb{COV}[\boldsymbol{\beta}]_{p\times p}\end{bmatrix}
\end{aligned}$$

$$\begin{aligned}\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{F}_j^*\hat{F}_j] &= \frac{1}{k}\sum_{p=1}^{k}\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{f}_p^*\mathbf{1}(M_p\le M_j)\hat{F}_j]\\ &= \frac{1}{k}\sum_{p=1}^{k}\sum_{u=1}^{kk}\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j]\end{aligned}$$

$$\mathbb{COV}[\hat{F}_i\hat{F}_{i-1},\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j] \approx [\bigtriangledown g^T\Sigma\bigtriangledown g]_{1,2}$$

where $g_1(\hat{F}_i,\hat{F}_{i-1},\hat{F}_j,\boldsymbol{\beta}) = \hat{F}_i\hat{F}_{i-1}$ and $g_2(\hat{F}_i,\hat{F}_{i-1},\hat{F}_j,\boldsymbol{\beta}) = \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j$, $g(\hat{F}_i,\hat{F}_j,\boldsymbol{\beta}) = (g_1,g_2)^T$ and

$$\begin{aligned}\bigtriangledown g &= \begin{bmatrix}\frac{\partial}{\partial\hat{F}_i}g_1 & \frac{\partial}{\partial\hat{F}_{i-1}}g_1 & \frac{\partial}{\partial\hat{F}_j}g_1 & \frac{\partial}{\partial\hat{\boldsymbol{\beta}}}g_1\\ \frac{\partial}{\partial\hat{F}_i}g_2 & \frac{\partial}{\partial\hat{F}_{i-1}}g_2 & \frac{\partial}{\partial\hat{F}_j}g_2 & \frac{\partial}{\partial\hat{\boldsymbol{\beta}}}g_2\end{bmatrix}\\ &= \begin{bmatrix}\hat{F}_{i-1} & \hat{F}_i & 0 & \mathbf{0}_{1\times p}\\ 0 & 0 & \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j) & \frac{\partial}{\partial\hat{\boldsymbol{\beta}}}\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_j\end{bmatrix}\\ \Sigma &= \begin{bmatrix}\mathbb{COV}[\hat{F}_i,\hat{F}_{i-1},\hat{F}_j]_{3\times 3} & \mathbf{0}_{3\times p}\\ \mathbf{0}_{p\times 3} & \mathbb{COV}[\boldsymbol{\beta}]_{p\times p}\end{bmatrix}\end{aligned}$$

$$\begin{aligned}\mathbb{COV}[\hat{F}_i^2,\hat{F}_j^*\hat{F}_{j-1}] &= \frac{1}{k}\sum_{p=1}^{k}\mathbb{COV}[\hat{F}_i^2,\hat{f}_p^*\mathbf{1}(M_p\le M_j)\hat{F}_{j-1}]\\ &= \frac{1}{k}\sum_{p=1}^{k}\sum_{u=1}^{kk}\mathbb{COV}[\hat{F}_i^2,\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_{j-1}]\end{aligned}$$

$$\mathbb{COV}[\hat{F}_i^2,\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p\le M_j)\hat{F}_{j-1}] \approx [\bigtriangledown g^T\Sigma\bigtriangledown g]_{1,2}$$

where $g_1(\hat{F}_i, \hat{F}_{j-1}, \boldsymbol{\beta}) = \hat{F}_i^2$ and $g_2(\hat{F}_i, \hat{F}_{j-1}, \boldsymbol{\beta}) = \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j)\hat{F}_{j-1}$, $g(\hat{F}_i, \hat{F}_j, \boldsymbol{\beta}) = (g_1, g_2)^T$ and

$$\begin{aligned}
\nabla g &= \begin{bmatrix} \frac{\partial}{\partial \hat{F}_i} g_1 & \frac{\partial}{\partial \hat{F}_{j-1}} g_1 & \frac{\partial}{\partial \hat{\boldsymbol{\beta}}} g_1 \\ \frac{\partial}{\partial \hat{F}_i} g_2 & \frac{\partial}{\partial \hat{F}_{j-1}} g_2 & \frac{\partial}{\partial \hat{\boldsymbol{\beta}}} g_2 \end{bmatrix} \\
&= \begin{bmatrix} 2\hat{F}_i & 0 & \mathbf{0}_{1\times p} \\ 0 & \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j) & \frac{\partial}{\partial \hat{\boldsymbol{\beta}}}\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j)\hat{F}_{j-1} \end{bmatrix} \\
\Sigma &= \begin{bmatrix} \mathbb{COV}[\hat{F}_i, \hat{F}_{j-1}]_{2\times 2} & \mathbf{0}_{2\times p} \\ \mathbf{0}_{p\times 2} & \mathbb{COV}[\boldsymbol{\beta}]_{p\times p} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{F}_j^*\hat{F}_{j-1}] &= \frac{1}{k}\sum_{p=1}^{k}\mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{f}_p^*\mathbf{1}(M_p \le M_j)\hat{F}_{j-1}] \\
&= \frac{1}{k}\sum_{p=1}^{k}\sum_{u=1}^{kk}\mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j)\hat{F}_{j-1}]
\end{aligned}$$

$$\mathbb{COV}[\hat{F}_i\hat{F}_{i-1}, \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j)\hat{F}_{j-1}] \approx [\nabla g^T \Sigma \nabla g]_{1,2}$$

where $g_1(\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j, \hat{F}_{j-1}, \boldsymbol{\beta}) = \hat{F}_i\hat{F}_{i-1}$ and $g_2(\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j, \hat{F}_{j-1}, \boldsymbol{\beta}) = \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j)\hat{F}_{j-1}$, $g(\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j, \hat{F}_{j-1}, \boldsymbol{\beta}) = (g_1, g_2)^T$ and

$$\begin{aligned}
\nabla g &= \begin{bmatrix} \frac{\partial}{\partial \hat{F}_i} g_1 & \frac{\partial}{\partial \hat{F}_{i-1}} g_1 & \frac{\partial}{\partial \hat{F}_j} g_1 & \frac{\partial}{\partial \hat{F}_{j-1}} g_1 & \frac{\partial}{\partial \hat{\boldsymbol{\beta}}} g_1 \\ \frac{\partial}{\partial \hat{F}_i} g_2 & \frac{\partial}{\partial \hat{F}_{i-1}} g_2 & \frac{\partial}{\partial \hat{F}_j} g_2 & \frac{\partial}{\partial \hat{F}_{j-1}} g_2 & \frac{\partial}{\partial \hat{\boldsymbol{\beta}}} g_2 \end{bmatrix} \\
&= \begin{bmatrix} \hat{F}_{i-1} & \hat{F}_i & 0 & \mathbf{0}_{1\times p} & \\ 0 & 0 & \hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j) & \frac{\partial}{\partial \hat{\boldsymbol{\beta}}}\hat{f}^*(M_p|\mathbf{Z}_u)p_u\mathbf{1}(M_p \le M_j) & \frac{\partial}{\partial \hat{F}_{i-1}} g_1 \end{bmatrix} \\
\Sigma &= \begin{bmatrix} \mathbb{COV}[\hat{F}_i, \hat{F}_{i-1}, \hat{F}_j, \hat{F}_{j-1}]_{4\times 4} & \mathbf{0}_{4\times p} \\ \mathbf{0}_{p\times 4} & \mathbb{COV}[\boldsymbol{\beta}]_{p\times p} \end{bmatrix}
\end{aligned}$$

When Y is continuous (as in Normal or Gamma model):

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

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

$$\begin{aligned}\mathbb{COV}[\hat{F}_i^*, \hat{F}_j^*] &= \mathbb{COV}[\frac{1}{k}\sum_{p=1}^{k} \hat{f}^*(M_p)\mathbf{1}(M_p \le M_i), \frac{1}{k}\sum_{m=1}^{k} \hat{f}^*(M_m)\mathbf{1}(M_m \le M_j)] \\ &= \frac{1}{k^2}\sum_{p=1}^{k}\sum_{m=1}^{k} \mathbb{COV}[\hat{f}^*(M_p)\mathbf{1}(M_p \le M_i), \hat{f}^*(M_m)\mathbf{1}(M_m \le M_j)]\end{aligned}$$

$$\begin{aligned}&\mathbb{COV}[\hat{f}^*(M_p)\mathbf{1}(M_p \le M_i), \hat{f}^*(M_m)\mathbf{1}(M_m \le M_j) \\ &= \frac{1}{kk^2}\sum_{u=1}^{kk}\sum_{v=1}^{kk} \mathbb{COV}[\hat{f}^*(M_p|\mathbf{W}_u)p_u\mathbf{1}(M_p \le M_i), \hat{f}^*(M_m|\mathbf{W}_v)p_v\mathbf{1}(M_m \le M_j)]\end{aligned}$$

$\mathbb{COV}[\hat{f}^*(M_p|\mathbf{W}_u)p_u\mathbf{1}(M_p \le M_i), \hat{f}^*(M_m|\mathbf{W}_v)p_v\mathbf{1}(M_m \le M_j)]$ can be estimated using delta method as in the discrete case.

Bibliography

Dong, J., and Q. Yu. 2018. “Joint Distribution and Marginal Distribution Methodsfor Generalized Linear Model.”