In this report, we shall give the proofs of some statements in the main paper. To be consistent with the paper, we order the sections in this report as §8 and §9. The proofs in §7 are given in §8. The other proofs are given in §9.

§8. Proofs of the lemmas in §7 related to Theorems 6.1 and 6.2.

Most proofs in §6 are given here, except for Lemma 7.1. Lemma 7.1 essentially says that the GMLE satisfies both the cdf form and the df form of the SC equation. The proof of the equivalence between the df form and the cdf form is quite similar to the proof under the interval censoring case in (Li et al. (1997)), and is skipped.

**Proof of Lemma 7.2.** Statement (1) of the lemma follows from Eq.s (7.4) and (7.5).

To prove Statement (2), it follows from (7.5) that

\[
\begin{align*}
\mathcal{R}_F(F)(x, y) &= \int_{v<x} \frac{F(x, y) - F(v, y)}{1 - F_T(v)} (1 - F_T(v))dF(v) \\
&\quad + \sum_{w \in J} \int_{v \leq x} \sum_{i \leq w} f_{C|T}(i|v) \sum_{c \in w} f_{C|T}(c|v)q(w|v) dF_T(v) \\
&= \int_{v<x} F(x, y) - F(v, y)dF_R(v) + \int_{v \leq x} \sum_{i \leq w} \sum_{w \in J} f_{C|T}(i|v)q(w|v)dF_T(v) \\
&= \int_{v<x} F(x, y) - F(v, y)dF_R(v) + \int_{v \leq x} \sum_{i \leq y} \sum_{w \in J} f_{C|T}(i|v) \sum_{i \in w} q(w|v) dF_T(v) \\
&= \int_{v<x} \int_{v \leq u \leq x, c \leq y} dF(u, c)dF_R(v) + \int_{v \leq x} \sum_{i \leq y} f_{C|T}(i|v) S_R(v-) dF_T(v) \\
&\quad (\text{as } \sum_{w \in J} q(w|x)1_{y \in w}) = \sum_{w: y \in w \cap w \in h} \sum_{w \in P_h} P(\Delta = h, R \geq x) (\text{see A2}) \\
&= \sum_{h: y \in P_h} \sum_{j} P(\Delta = h, R \geq x) = S_R(x-) \\
&= \sum_{c \leq y} f_{C}(c) \{ \int_{v<x} \int_{v \leq u \leq x} dF_{T|C}(u|c)dF_R(v) + \int_{v \leq x} \int_{v \leq r} dF_R(r)dF_{T|C}(v|c) \} \\
&= \sum_{c \leq y} f_{C}(c)F_{T|C}(x|c) = F(x, y).
\end{align*}
\]

It follows that \( F = \mathcal{R}_F(F) = \mathcal{B}_F(Q) \) by (1). Thus statement (2) holds. Verify that in the proof of \( F = \mathcal{R}_F(F), \) \( F \) is not critical, as long as \( F \in \Theta_o. \) Thus statement (3) holds too.

The proofs of statements (4), (5) and (6) are similar to those of (1), (2) and (3). \( \square \)

**Proof of Lemma 7.3.** For each \( \tilde{F} \in \Theta_o, \) by Lemma 7.2, \( H = \tilde{F} \) is a solution to \( H = \mathcal{R}_H(\tilde{F}). \) If there exists a different solution of \( H = \mathcal{R}_H(\tilde{F}) \) in \( \Theta_o, \) say \( H = G \neq \tilde{F} \) and \( S_G \supset S_{\tilde{F}}, \) then by Lemma 7.2, we have \( G = \mathcal{R}_G(G) \) as well as \( G = \mathcal{R}_G(\tilde{F}). \) That is, \( H = G \) and \( H = \tilde{F} \) are two different solutions to \( G = \mathcal{R}_G(H). \)

On the other hand, given \( H \in \Theta_o, \) by Lemma 7.2, \( \tilde{F} = H \) is a solution to \( H = \mathcal{R}_H(\tilde{F}). \) If \( \tilde{F} = G \neq H \) in \( \Theta_o \) is a different solution to \( H = \mathcal{R}_H(\tilde{F}) \) and \( S_H \supset S_G, \) then by Lemma 7.2, \( G = \mathcal{R}_G(G). \) Thus \( \tilde{F} = H \) and \( \tilde{F} = G \) are two different solutions to \( H = \mathcal{R}_H(\tilde{F}). \) \( \square \)
Proof of Lemma 7.4. It is obvious that $B_F$ is linear. Let $t_0 = -\infty$, $t_{m+1} = \infty$, $S_{F_T} \cap D_V = \{ t_1, ..., t_m \}$, where $t_1 < \cdots < t_m$. It suffices to show that $\forall g \in D_Q$, $H = B_F(g)$ satisfies (7.2), (1) in (2.2) and $S_H \subset S_{F_T}$.

Since $\hat{F}$ satisfies (2.2) and it follows from (7.4) that

$$H(x, y) = B_F(g)(x, y) = \sum_{h=1}^{J} 1(h \leq y) \{ \int_{v < x} \frac{\hat{F}_h(x) - \hat{F}_h(v)}{1 - \hat{F}_T(v)} dg(v, 0, C_r) + \int_{v \leq x, v \in S_{F_T}, w \in J} 1(h \leq w) \hat{f}_C|T(h|v) \sum_{i \in w} \hat{f}_C|T(i|v) \}.$$ (8.2)

Since $\hat{F}$ satisfies (7.2) and (1) in (2.2), in view of (8.2) $H$ also satisfies (7.2) and (1) in (2.2). Since $||S_{F_T}|| \leq (m + 1)J$, in order to show $S_H \subset S_{F_T}$, it suffices to show

$S_{H_T} \subset S_{F_T}$, i.e., if $t_j \leq x < t_{j+1}, j \leq m$, then $B_F(g)(x, \infty) - B_F(g)(t_j, \infty) = 0 \ \forall \ g \in D_Q$; and if $(t_j, c) \notin S_{F_T}, j \leq m + 1$ and $c \in C_r$, then $\mu_H(\{(t_j, c)\}) = 0$. (8.3)

Actually for each $y \in C_r$, if $t_j \leq x < t_{j+1}, j \leq m$, then

$$B_F(g)(x, y) - B_F(g)(t_j, y) = \int_{v < x} \frac{\hat{F}(x, y) - \hat{F}(v, y)}{1 - \hat{F}_T(v)} dg(v, 0, C_r) + \int_{v \leq x, v \in S_{F_T}, w \in J} \frac{\hat{f}_C|T(h|v)}{\sum_{i \in w} \hat{f}_C|T(i|v)} \sum_{h \leq y, h \in w} \hat{f}_C|T(h|v) \sum_{i \in w} \hat{f}_C|T(i|v) \}.$$ (8.2)

Moreover, if $\mu_F(\{(t_j, c)\}) = 0$, then noting $t_j \in S_{F_T},$ by (8.2)

$$\mu_H(\{(t_j, c)\}) = \int_{v < t_j} \frac{\hat{F}_c^s(t_j) - \hat{F}_c^s(t_j-1)}{1 - \hat{F}_T(v)} dg(v, 0, C_r) + \int_{v = t_j, w \in J} \frac{1(c \leq w) \hat{f}_C|T(c|v)}{\sum_{i \in w} \hat{f}_C|T(i|v)} \sum_{i \in w} \hat{f}_C|T(i|v) \}.$$ (8.2)

Thus (8.3) holds. This completes the proof of the statement related to $B_H$. Replacing $dg$ in $B_F$ in the foregoing arguments by $(H(\infty, J) - H(v, J))dF_R$ or $q(w|v)dH(v, c)$ in $R_{F_T}$, we can establish a similar version of (8.3) for $R_{F_T}$. Since $R_{F_T}$ is also linear, it completes the proof of the lemma. □

Proof of Lemma 7.5. Since (1) $R_{F_T}(\hat{F}) = \hat{F} = B_{\hat{F}}(\hat{Q})$ (by Lemma 7.1), (2) $R_{F_T}(F) = B_{\hat{F}}(Q)$ (by Lemma 7.2), $R_{F_T}(\hat{F}) - R_{F_T}(F) = B_{\hat{F}}(\hat{Q}) - B_{\hat{F}}(Q)$. Since both of the operators are linear by Lemma 7.4, $R_{F_T}(\hat{F} - F) = B_{\hat{F}}(\hat{Q} - Q)$. □

Proof of Lemma 7.6. We shall give the proof in 3 steps.
Step 1 (preliminary). Recall \( f = f_{T,C} \). The uniqueness of the solution to the equation 
\( \tilde{F} = \mathcal{R}_{\tilde{F}}(F) \) is equivalent to the uniqueness of the solution to equations (7.3) (w.r.t. the measure 
\( \mu(\cdot, \cdot) \) on \( D_V \times C_r \)). Since we are interested in the equation 
\( \tilde{F} = \mathcal{R}_{\tilde{F}}(H) \) rather than \( \tilde{F} = \mathcal{R}_{\tilde{F}}(F) \), Eq. (7.3) becomes, for \( x \in S_{\tilde{F}_T} \) and \( H \in \mathcal{D} \),

\[
\hat{f}(x, y) = \int_{v < x} \frac{\tilde{f}(x, y)(1 - HT(v))}{1 - \tilde{F}_T(v)} dF_R(v) + \sum_{w \in J} q(w|x) \frac{1(y \in w) \tilde{f}(x, y)}{\sum_{k \in w} \tilde{f}(x, k) \sum_{h \in w} h(x, h)} h(x, h); \tag{8.4}
\]

\[
\hat{f}_T(x) = \hat{f}_T(x) \int_{v < x} \frac{1 - HT(v)}{1 - \tilde{F}_T(v)} dF_R(v) + 1_{(x \in S_{\tilde{F}_T})} S_R(x-\cdot) h_T(x), \tag{8.5}
\]

where \( (H_T, h_T) \) correspond to \( (F_T, f_T) \), respectively.

Step 2 (uniqueness of the solution to (8.5)). It is easy to show that \( H = \tilde{F} \) is a solution to the aforementioned equations. In particular, \( h_T = \hat{f}_T \) is the solution of the SC equation (8.5).

By the assumption \( S_H \subset S_{\tilde{F}_T} \), \( S_{H_T} \subset S_{\tilde{F}_T} \), which is finite. If \( h_T \neq \hat{f}_T \), let \( t_1 < \cdots < t_m \) be all the points in \( S_{\tilde{F}_T} \cap D_V \). Let \( t_{i_0} \) be the smallest point in \( S_{\tilde{F}_T} \) such that \( \hat{f}_T(t_{i_0}) \neq h_T(t_{i_0}) \). Since \( \hat{f}_T(t_i) = h_T(t_i) \) for \( i < i_0 \) and \( H_T(-\infty) = \tilde{F}_T(-\infty) = 0 \) by (1) in (2.2), \( H_T(v) = \tilde{F}_T(v) \) for \( v < t_{i_0} \).

Letting \( x = t_{i_0} \), Eq. (8.5) yields

\[
\hat{f}_T(t_{i_0}) = \hat{f}_T(t_{i_0}) \int_{v < t_{i_0}} 1 dF_R(v) + S_R(t_{i_0}-) h_T(t_{i_0}),
\]

\[
\Rightarrow \hat{f}_T(t_{i_0})(1 - F_T(t_{i_0}-)) = S_R(t_{i_0}-) h_T(t_{i_0}),
\]

\[
\Rightarrow \hat{f}_T(t_{i_0}) = h_T(t_{i_0}) \ (\text{as} \ S_R(t_{i_0}-) > 0 \ \text{due to} \ t_{i_0} \in D_V),
\]

contradicting the assumption \( \hat{f}_T(t_{i_0}) \neq h_T(t_{i_0}) \). The contradiction implies that \( h_T = \hat{f}_T \) on \( D_V \).

By (7.2), \( H_T = \tilde{F}_T \) on \( [-\infty, \infty) \). Since \( H_T(\infty) = \tilde{F}_T(\infty) = 1 \), \( H_T = \tilde{F}_T \).

Step 3 (conclusion). Since \( h_T = \hat{f}_T \) by Step 2, (8.4) is equivalent to

\[
\hat{f}_{C|T}(y|x) = \hat{f}_{C|T}(y|x) F_R(x-) + \sum_{w \in J} \sum_{j \in w} h_{C|T}(j|x) \frac{q(w|x) \hat{f}_{C|T}(y|x)}{\sum_{k \in w} \hat{f}_{C|T}(k|x)} \sum_{j \in w} h_{C|T}(j|x) \frac{S_R(x-)}{\sum_{k \in w} \hat{f}_{C|T}(k|x)}, \quad (x, y) \in S_{\tilde{F}}; \tag{8.6}
\]

or

\[
\hat{f}_{C|T}(y|x) = \hat{f}_{C|T}(y|x) \sum_{w \in J} \frac{q(w|x) \mathbb{1}(y \in w)}{S_R(x-)} \sum_{j \in w} h_{C|T}(j|x) \frac{S_R(x-)}{\sum_{k \in w} \hat{f}_{C|T}(k|x)}, \quad (x, y) \in S_{\tilde{F}}. \tag{8.7}
\]

Verify that \( h_{C|T} = \hat{f}_{C|T} \) is a solution to (8.7), as the right hand side of (8.7) becomes

\[
\hat{f}_{C|T}(y|x) \sum_{w \in J} \frac{q(w|x) \mathbb{1}(y \in w)}{S_R(x-)} = \hat{f}_{C|T}(y|x) \frac{S_R(x-)}{S_R(x-)} = \hat{f}_{C|T}(y|x)
\]

by (8.1). We shall prove that

the solution \( h_{C|T}(|x) \) to (8.7) is unique a.e. in \( x \) on \( D_V \) w.r.t. \( \mu_{F_T} \).

\[
\tag{8.8}
\]

Consider two cases: (1) \( \hat{f}_{C|T}(c|x) = 0 \) for some \( c \), (2) \( \hat{f}_{C|T}(c|x) > 0 \) for each \( c \).
In case (1), since \( S_F \subset S_{F} \) (by (7.2)), by A2 \( \exists y = \{c\} \in J \) such that \( q(w|x) > 0 \). Then it follows from (8.7) that if \( y = c \), then (8.7) yields

\[
\tilde{f}_{C'|T}(c|x) = \tilde{f}_{C'|T}(c|x) \sum_{w \in J} [1_{(w = \{c\})} + 1_{(w \neq \{c\})}] \frac{q(w|x)1_{(c \in w)}}{S_R(x^-)} \frac{\sum_{j \in w} h_{C'|T}(j|x)}{\sum_{k \in w} \tilde{f}_{C'|T}(k|x)}, \quad (x, c) \in S_F.
\]

Thus \( 0 = 0 \left( \sum_{j \in w \notin \{c\}} \frac{h_{C'|T}(c|x)}{0} + 0 \right) \). It implies \( h_{C'|T}(c|x) = 0 \) as \( \frac{0 \text{def}}{0} = 1 \). Since \( \tilde{f}_{C'|T}(c|x) = h_{C'|T}(c|x) = 0 \), it reduces to case (2) by replacing \( J - 1 \) for \( J \).

Now consider case (2). Then (8.7) is equivalent to

\[
1 = \sum_{w \in J, y \in w} \frac{q(w|x)}{S_R(x^-)} \frac{\sum_{j \in w} h_{C'|T}(j|x)}{\sum_{k \in w} \tilde{f}_{C'|T}(k|x)}, \quad y \in \{1, \ldots, J\}. \tag{8.9}
\]

Treating \( \tilde{f}_{C'|T} \) as given constant, \( \mathbf{h}' = (h_{C'|T}(1|x), \ldots, h_{C'|T}(J|x)) \) as an undetermined vector, and \( 1 \) as the vector with coordinates 1, Eq. (8.9) is actually a system of linear equations in \( h_{C'|T} \), say \( 1 = \mathbf{A} \mathbf{h} \), where \( \mathbf{A} = K'_{\phi} \mathbf{K}_q \mathbf{K}_\phi \), \( \mathbf{K}_q \) is the \( k_t \times k_t \) diagonal matrix with diagonal elements \( \frac{q(W_i|x)}{S_R(x^-)\sum_{k \in w_j} \tilde{f}_{C'|T}(k|x)} \). \( W_i \), \( \ldots \), \( W_{k_t} \) are all the elements of \( J \) such that \( q(W_i|x) > 0 \), and \( \mathbf{K}_\phi \) is the \( k_t \times J \) matrix with the \( j^{th} \) row \( \mathbf{K}_\phi(W_j)' \) (see A1, which is implied by A2). Verify that \( \mathbf{K}_\phi \) is of full rank \( J \) and \( \mathbf{K}_q \) is a diagonal matrix of full rank \( k_t \), thus \( \mathbf{K}'_{\phi} \mathbf{K}_q \mathbf{K}_\phi \) is of full rank \( J \). Consequently, the solution of the equation \( 1 = \mathbf{A} \mathbf{h} \) is unique and \( h_{C'|T} = \tilde{f}_{C'|T} \), as the latter is a solution to (8.9). This solves the proof of the lemma. \( \square \)

**Proof of Lemma 7.7.** By assumption \( ||S_F|| \) is finite, let \( (t_1, c_1), \ldots, (t_m, c_m) \) be all the distinct points in \( S_F \cap (D_V \times \mathbb{C}_r) \). WLOG, we can assume that \( \tau < \infty \) and \( (\infty, J) \in S_F \). Otherwise, the proof is simpler as \( S_F \subset (D_V \times \mathbb{C}_r) \). Let \( (t_{m+1}, c_{m+1}) = (\infty, J) \). Let \( g_i(x) = 1_{(x \geq (t_i, c_i))} \) and \( H_i = R_{S_F}(g_i), i = 1, \ldots, m + 1 \).

By Lemma 7.4, \( S_{R_F} \subset S_F \). We shall show that \( \forall i \in \{1, \ldots, m + 1\}, \)

\[
H_i(-\infty, -\infty) = 0, \quad H_i(\infty, \infty) = 1 \quad \text{and} \quad \mu_{H_i} \left( \{(t_k, c_k)\} \right) \geq 0, \quad \text{for all possible (i, k).} \tag{8.10}
\]

Now \( \forall i, H_i(-\infty, -\infty) = 0 \) follows from Lemma 7.4 and (1) in (2.2). Moreover,

\[
H_i(\infty, \infty) = \int_{v < \infty} \frac{\tilde{F}(\infty, \infty) - \tilde{F}(v, \infty)}{1 - \tilde{F}(v)} (1 - g_i(v, J)) dF_R(v)
\]

\[
+ \sum_{w \in J} \int_{v \leq \infty} \int_{c \in w} \sum_{h \leq \infty, h \in w} \tilde{f}_{C'|T}(h|v) \frac{\sum_{j \in w} \tilde{f}_{C'|T}(j|v)}{\sum_{k \in w} \tilde{f}_{C'|T}(k|v)} q(w|v)dg_i(v, c)
= \int_{v < \infty} (1 - g_i(v, J))dF_R(v) + \sum_{w \in J} \int_{v \leq \infty} \int_{c \in w} q(w|v)dg_i(v, c)
= \int_{v < t_i} dF_R(v) + \sum_{w \in J, c_i \in w} q(w|t_i)
= \int_{v < t_i} dF_R(v) + \sum_{w \in J, c_i \in w} \sum_{h \in P_h} f(h)S_{R|\Delta}(t_i - |h|
= P(R < t_i) + S_R(t_i-) = 1.
\]
Thus the first 2 equations in (8.10) hold. Notice that

\[ H_i(t_k, c_k) = \int_{v < t_k} \frac{\hat{F}(t_k, c_k) - \hat{F}(v, c_k)}{1 - \hat{F}(v)} (1 - g_i(v, J)) dF(v) \]

\[ + \sum_{w \in J} \int_{v < t_k} \int_{c \in w} \frac{\sum_{h \leq c_k, h \in w} \hat{f}_{C \mid T}(h|v)}{\hat{f}_{C \mid T}(j|v)} q(w|v) dK(v) \]

\[ = \sum_{h=1}^{J} 1_{(h \leq c_k)} \int_{v < t_k} (\hat{F}_h^a(t_k) - \hat{F}_h^a(t_k)) \frac{1 - g_i(v, J)}{1 - \hat{F}(v)} dF(v) \]

\[ + \sum_{w \in J} \int_{v < t_k} \int_{c \in w} \frac{1_{(h \in w)} \hat{f}_{C \mid T}(h|v)}{\hat{f}_{C \mid T}(j|v)} q(w|v) dK(v) \]

Thus (8.10) holds. (8.10) implies that \( H_i \in \Theta_0 \).

Notice that \( (g_1, ..., g_{m+1}) \) is a base of the linear space \( \mathcal{D}_F \) with \( g_i(\infty, J) = 1 \) and \( g_i \in \Theta_0 \), thus each \( H \) in \( \Theta_0 \cap \mathcal{D}_F \) satisfies \( H = \sum_{i=1}^{m+1} a_i g_i \), where \( \sum_i a_i = 1 \) and \( a_i \geq 0 \). It follows that \( R_F(H) \in \Theta_0 \).

**Proof of Lemma 7.8.** Since \( ||S_F|| \) is finite, by Lemma 7.4, \( R_F \) is a linear map from \( \mathcal{D}_F \) to \( \mathcal{D}_F \). WLOG, we can assume that \( (\infty, J) \) belongs to \( S_F \) (= \{ (t_1, c_1), ..., (t_{m+1}, c_{m+1}) \}), where \( (t_{m+1}, c_{m+1}) = (\infty, J) \). In order to show that \( R_F^{-1} \) exists on \( \mathcal{D}_F \cap \mathcal{D}_1 \), it suffices to show that (1) \( R_F \) is 1-1 on \( \mathcal{D}_F \cap \mathcal{D}_1 \) and (2) \( R_F(H) (\infty, J) = i \) if \( H \in \mathcal{D}_F \cap \mathcal{D}_1 \).

Now suppose that \( R_F(H) = 0 \) and \( H \in \mathcal{D}_F \cap \mathcal{D}_0 \). Since \( R_F(\hat{F}) = \hat{F} \) by Lemma 7.2 \( \hat{F} + 0 = R_F(\hat{F}) + R_F(H) = R_F(\hat{F} + H) \). Write \( H_1 = \hat{F} + H \), then \( H_1(\infty, J) = 1 \), \( H_1 \in \mathcal{D}_F \) and \( R_F(H_1) = \hat{F} \). Thus by Lemma 7.6, \( H_1 = \hat{F} \), that is, \( \hat{F} + H = \hat{F} \). Hence \( H = 0 \), i.e., \( R_F \) is 1-1 on \( \mathcal{D}_F \cap \mathcal{D}_1 \). Thus (1) holds for \( i = 0 \).

If \( H \in \mathcal{D}_F \cap \mathcal{D}_0 \), then \( H(x, y) = \sum_{i=1}^{m+1} d_i 1_{(x,y) \geq (t_i, c_i)} \), where \( \sum_{i=1}^{m+1} d_i = 0 \) and \( m+1 = ||S_F|| \).

It follows that \( H = H_+ - H_- \), where \( H_+(x, y) = \sum_{i=1}^{m+1} d_i 1_{((x,y) \geq (t_i, c_i))} \) and \( H_-(x, y) = -\sum_{i=1}^{m+1} d_i 1_{((x,y) \geq (t_i, c_i), d_i < 0)} \). Let \( d = \sum_{i=1}^{m+1} d_i 1_{d_i \geq 0} \). Then \( H_+/d \) and \( H_-/d \) both belong to \( \Theta_0 \). By Lemma 7.7, \( R_F(H_+/d) \in \Theta_0 \) and \( R_F(H_-/d) \in \Theta_0 \). Thus \( R_F(H) \in \mathcal{D}_0 \). That is, (2) holds for \( i = 0 \). It follows that \( R_F^{-1} \) exists on \( \mathcal{D}_F \cap \mathcal{D}_0 \).

If \( H \) and \( G \) are in \( \mathcal{D}_F \cap \mathcal{D}_1 \) and \( R_F(H) = R_F(G) \), then \( 0 = R_F(H - G) \). Notice that \( H - G \in \mathcal{D}_0 \cap \mathcal{D}_F \), thus \( H - G = R_F^{-1}(0) = 0 \), that is \( H = G \). As a consequence, \( R_F \) is 1-1 on \( \mathcal{D}_1 \cap \mathcal{D}_F \). By Lemma 7.7, \( R_F(H) \in \Theta_0 \) if \( H \in \Theta_0 \), it follows that \( R_F^{-1} \) exists on \( \mathcal{D}_F \cap \mathcal{D}_1 \).

It is easy to show that \( R_F^{-1} \) is also linear. \( \Box \)
Proof of Lemma 7.9. By (7.5), it is obvious that statement (1) holds, as

\[ R_F(U_o)(x, y) = \int_{v < x} \frac{\tilde{F}(x, y) - \tilde{F}(v, y)}{1 - \tilde{F}_T(v)} (0 - 0) dF_R(v) \]

\[ + \sum_{w \in \mathcal{J}} \int_{v \leq x} \int_{c \in \mathcal{W}} \int_{d \in \mathcal{S}_F} \frac{\tilde{f}_C(T(i)v)}{\tilde{f}_C(T(j)v)} q(w|v) d\tilde{F}_T(v, w) \]

Moreover, if \( x < \tilde{t}_o \), then \( R_F(U_F)(x, y) = U_F(x, y) = 0 \). If \( x \geq \tilde{t}_o \), then \( R_F(U_F)(x, y) \)

\[ = \int_{v < x} \frac{\tilde{F}(x, y) - \tilde{F}(v, y)}{1 - \tilde{F}_T(v)} (U_F(\infty, J) - U_F(v, J)) dF_R(v) \]

\[ + \sum_{w \in \mathcal{J}} \int_{v \leq x} \int_{c \in \mathcal{W}} \int_{d \in \mathcal{S}_F} \frac{\tilde{f}_C(T(i)v)}{\tilde{f}_C(T(j)v)} q(w|v) d\tilde{F}_T(v, w) \]

= \int_{\inf \mathcal{S}_F \leq v \leq \tilde{t}_o} \tilde{F}(x, y) dF_R(v) + \sum_{w \in \mathcal{J}} \sum_{i \leq y} \tilde{f}_C(T(i)v) \sum_{j \in \mathcal{J}} \tilde{f}_C(T(j)v) q(w|v)|_{v=\tilde{t}_o} \]

(8.11)

In particular, if \( \tilde{F} = F \), then \( R_F(U_F)(x, y) = \sum_{i \leq y} f_C(T(i)v) S_R(v-)|_{v=\tilde{t}_o} 1(\geq \tilde{t}_o) = U_F(x, y) \).

Moreover, (8.11) yields statement (3) of the lemma by A3.

Proof of Lemma 7.10. For \( \tilde{F} \in \Theta_o \) with \( t_o = t_o \), let \( C_k \) be the collection of all the distinct points among \( c_k,j \) s, where \( c_k,j = \inf \{ x : \tilde{F}_T(x) \geq j/2^k \}, j = 0, ..., 2^k, k \geq 1 \). Let \( F_k \) be a function in \( \Theta_o \) such that \( F_k(t, c) = \tilde{F}(t, c) \forall (t, c) \in C_k \times C_r \) and its \( F_k^* \) is a step function with its discontinuity points in \( C_k \). Denote \( W_k \) the subclass of \( \mathcal{W} \) such that each member \( H \) satisfies that \( H^*_k \) is a step function and the collection of discontinuity points is a subset of \( S_{F_k^*} \).

Set \( i = 1 \) first. For each \( g \in \mathcal{D}_F \cap \mathcal{D}_i \) and \( k \geq 1 \), let \( g_k \in W_k \cap \mathcal{D}_i \) be such that \( g_k(x, c) = g(x, c) \) if \( (x, c) \in C_k \times C_r \). Then \( \| g_k - g \| \to 0 \), since \( S_{g_k} \in \mathcal{S}_F, S_g \in \mathcal{S}_F \) and the set \( \cup_k C_k \) is dense in \( S_F \). By Lemma 7.8, \( R_F^{-1} \) exists, so \( \exists \ H_k \in W_k \cap \mathcal{D}_i \) such that \( g_k = R_{F_k}(H_k), \forall K > k \) and \( H \in \mathcal{D}_F \cap \mathcal{D}_i, W_k \subset W_K, ||F_k - F_K|| \leq 1/2^k \) and \( ||R_{F_k}(H) - R_{F_K}(H)|| \to 0 \) as \( k \to \infty \) by the BCT. Lemma 7.9 says that \( U_o \) and \( U_F \) play the role of 0 and 1 in some sense, respectively, and if \( k \) is large, one can write

\[ (R_{F_k} - R_{F_K})(H) = o(1)U_F \] and \( (R_{F_k} - R_{F_K})(H) = o(1)R_{F_K}(U_F), \]

(8.12)

as \( ||R_{F_k}(U_F) - U_F|| \to 0 \) by Lemma 7.9. Let \( G_k \in \mathcal{D}_{F_k} \cap \mathcal{D}_i \) such that \( G_k = R_{F_k}^{-1}(g_k) \).

\[ R_{F_k}(G_k - H_k) = R_{F_k}(G_k) - R_{F_k}(H_k) + R_{F_k}(H_k) - R_{F_k}(H_k) \]

\[ = g_k - g_k + R_{F_k}(H_k) - R_{F_k}(H_k) \]

\[ = (R_{F_k} - R_{F_K})(H_k) \]

(8.13)

(8.14)

as \( H_k = R_{F_k}^{-1}(g_k) \).
By (7.6) and (7.7), the solution of Lemma 7.3, and the second statement of Lemma 7.3 holds. Then the solution \( R_0 = G \Theta \int \). In other words, the solution satisfies \( H \in J \). The solution \( H \in S \), \( F \) follows almost line by line as the proof in the first paragraph making use of the hypotheses (7.9) as \( \| g_k - g \| \to 0 \) and thus \( \| g_k - g \| = o(1)U_{F_k} \) \( \| R \| \to 0 \) as \( k \to \infty \) (by (8.13)).

Thus \( \{ H_k \}_{k \geq 1} \) is a Cauchy sequence. Since \( D_F \) is a Banach space, \( \exists H_0 \in D_F \) such that \( \| H_k - H_0 \| \to 0 \). By the BCT, \( g = \lim_{k \to -\infty} R_{F_k}(H_k) = R_F(H_0) \). Define \( H_0 = R^{-1}_F(g) \).

Verify that the foregoing arguments hold also for \( i = 0 \).

**Proof of Lemma 7.13.** Since \( S_H = S_{\hat{F}}, D_H = D = D_F \) by (7.2). Given \( H \in \Theta_o \), let \( G \in \Theta_o \) be an arbitrary solution to \( H = R_H(G) \). Since \( H = R_H(H) \) by (3) of Lemma 7.2, \( 0 = H - H = R_H(G) - R_H(H) = R_H(G - H) \) by Lemma 7.4. That is, \( 0 = R_H(G - H) \). Since \( R^{-1}_H \) exists on \( D_H \cap D_0 \) by Lemma 7.10 and \( G - H \in D_H \cap D_0 \), \( 0 = R^{-1}_H(0) = G - H \). That is, \( G = H \). In other words, the solution \( G \in \Theta_o \) to the equation \( H = R_H(G) \) is unique. Hence the second statement of Lemma 7.3 holds. Then the solution \( H \) to \( R_H(F) = H \) is unique in \( \Theta_o \) by Lemma 7.3, and \( H = F \) by Lemma 7.2. Since \( B_H(Q) = R_H(F) \) for each \( H \in \Theta_o \) by Lemma 7.2, the solution \( H \in \Theta_o \) to \( H = B_H(Q) \) is unique and is \( H = F \).

Now consider \( H = B_{H,F_0}(Q) \). Notice that \( H \in \Theta_o \) and thus \( h_{C|T,w} \) are constant in \( w \in J \).

By (7.6) and (7.7),

\[
B_{H,F_0}(Q)(x,y) = \int_{v \leq x, v \in S_{H_T}} \frac{\sum_{i \leq y, i \in w} f_{C|T}(i|v)}{\sum_{j \in w} f_{C|T}(j|v)} dQ(v,1,w) + \sum_{v \leq x, v \in D_T} \frac{\sum_{i \leq y, i \in w} f_{C|T}(i|v)}{\sum_{j \in w} h_{C|T}(j|v)} F(v,c) + \int_{v \leq x} \frac{H(x,y) - H(v,y)}{1 - H_T(v)} dQ(v,0,C_r)
\]

Now verify that replacing \((B_o, R_H)\) for \((B_H, R_H)\) in Lemmas 7.2 through 7.9, except for Lemma 7.5, one can show that those modified lemmas hold too. Then the proof of the lemma regarding \( H = B_{H,F_0}(Q) \) follows almost line by line as the proof in the first paragraph making use of the modified lemmas. In particular, Eq. (8.5) and Eq. (8.6) become

\[
\hat{f}_T(x) = \hat{f}_T(x) \int_{v \leq x} \frac{1 - H_T(v)}{1 - T(v)} dF_R(v) + S_R(x-)[1_{(x \in D_T^c)]} \hat{f}_T(x) + 1_{(x \in D_T)} h_T(x),
\]
\[
\hat{f}_{C|T}(y|x)S_R(x) = \sum_{w \in J, y \in w} q(w|x) \left( \frac{\sum_{j \in w} h_{C|T}(j|x)}{\sum_{k \in w} \hat{f}_{C|T}(k|x)} \mathbf{1}_{(x \in D_T)} + \frac{\sum_{j \in w} \hat{f}_{C|T}(j|x)}{\sum_{k \in w} \hat{f}_{C|T}(k|x)} \mathbf{1}_{(x \in D_T')} \right),
\]

\(\forall (x, y) \in S_R\) and \(H \in D\). This completes the proof of the lemma. \(\square\)

**Proof of Lemma 7.15.** **Step 1** (preliminary). By definition, \(\hat{F}^*\) is the GMLE of the modified data \((V_{mi}, \delta_i, M_i)s\), thus it maximizes the modified likelihood function

\[
\Lambda_2 = \prod_{i=1}^{n} \mu_{\hat{F}}(I_{mi}), \text{ where } I_{mi} = \begin{cases} (V_{i}, \infty) \times M_i & \text{if } \delta_i = 0 \\ \{V_{mi}\} \times M_i & \text{otherwise,} \end{cases}
\]

and \(\hat{F}\) is a cdf.

It can be shown as in Lemma 7.1 that \(H = \hat{F}^*\) is a solution to the equation \(H = B_H(\hat{Q}_m)\), where \(\hat{Q}_m\) is the empirical cdf of \(Q\), based on \((V_{mi}, \delta_i, M_i)s\), and \(B_H\) is defined in (7.4). Let \(\Omega_6\) be the event that \(\hat{Q} \rightarrow Q\). By the SLLN \(P(\Omega_6) = 1\). Let \(\Omega_7\) be the event that

\[
\sup_{t \in [-\infty, \infty], u \in \{0,1\}, w \in J} |\hat{Q}_m(t, u, w) - Q(t, u, w)| \rightarrow 0.
\]  

(8.14)

**Step 2** (to show \(\Omega_6 \subset \Omega_7\)). Fix \(\omega \in \Omega_6\). Since \(V_{mi} = V_i\) if \(\delta_i = 0\), for each \(w \in J\),

\[
\sup_{t,u} |\hat{Q}_m(t, u, w) - Q(t, u, w)| = \sup_t |\hat{Q}_m(t, 1, w) - \hat{Q}(t, 1, w)| \overset{\text{def}}{=} d_{nt}.
\]

Then \(d_{nt} = \frac{1}{n} \sum_{i=1}^{n} d_i\), where \(d_i = 1_{(V_{mi} \leq t, \delta_i = 1, M_i = w)} - 1_{(V_i \leq t, \delta_i = 1, M_i = w)}\). Verify that \(d_{nt} = 0\) if \(t = 0\) (by the definition of \(V_{mi}s\)). For \(t > 0\), \(\exists j\) such that \(t \in (s_{n,j-1}, s_{n,j}]\). Then due to the definition of \(s_{n,j}\),

\[
d_{nt} = \frac{1}{n} \sum_{i=1}^{n} 1_{(V_i \in (s_{n,j-1}, t], \ s_{n,j-1} < t < s_{n,j})} \leq \mu_{\hat{F}}((s_{n,j-1}, t])1_{(t \in s_{n,j-1}, s_{n,j})} \leq 1/\sqrt{n} \rightarrow 0.
\]

That is, \(\sup_{t,u} |\hat{Q}_m(t, u, w) - Q(t, u, w)| \rightarrow 0\). Since \(\omega \in \Omega_6\), \(\sup_{t,u} |\hat{Q}(t, u, w) - Q(t, u, w)| \rightarrow 0\) by assumption on \(\Omega_6\). Thus it yields that for \((u, w) \in \{0,1\} \times J\),

\[
\sup_t |\hat{Q}_m(t, u, w) - Q(t, u, w)| \leq \sup_t |\hat{Q}_m(t, u, w) - \hat{Q}(t, u, w)| + |\hat{Q}(t, u, w) - Q(t, u, w)| \rightarrow 0.
\]

Hence (8.14) holds. That is, \(\omega \in \Omega_7\). Thus \(\Omega_6 \subset \Omega_7\).

**Step 3** (to derive \(\lim \hat{f}_{C|T}\)). For each \(j \in C_r\), \(\hat{F}^{*^j}\) is monotone and bounded by 0 and 1, each given subsequence of \(\hat{F}^{*^j}\) has a further convergent subsequence. Thus each subsequence of \((\hat{F}^{*^j}, \ldots, \hat{F}^{*^j})\), has a further convergent subsequence, as \(J < \infty\), and so is \(\hat{F}^*(\cdot, c)\), \(c \in C_r\). Denote the limiting point by \(H^*\) together with \(H^{*^j}\), \(c \in C_r\). For the given \(\omega \in \Omega_6\) and for each \(n\), one can define \(\hat{f}_{C|T}(c|t)\) for each \((c, t) \in C_r \times \{V_i : i = 1, \ldots, n\}\). Notice also that \(\hat{F}^{*^j}(x) = \int_{t \leq x} \hat{f}_{C|T}(c|t)d\hat{F}^{*^j}_T(t) = \int_{t \leq x} \hat{f}_{C|T}(c|t)d\hat{F}^{*^j}_T(t)\) (see step 2.4). Since \(\hat{f}_{C|T}\) is bounded and \(C_r \times \{V_i : i = 1, \ldots, n\}\), is countable, by Helly’s selection theorem there is a convergent subsequence of \(\hat{f}_{C|T}\) with limiting point \(h_{C|T}\) such that \(\hat{f}_{C|T}(c|t) \rightarrow h_{C|T}(c|t) \forall (c, t) \in C_r \times \{V_i : i = 1, \ldots, n\}\).

By taking further subsequence, without loss of generality, one can assume that \(\{\hat{F}^{*^j}_{n_h}, \hat{f}_{C|T,n_h}\}_{h \geq 1}\) converges. Now by the BCT, \(H^{*^j}(x) = \int_{t \leq x} h_{C|T}(c|t)dH^{*^j}_T(t)\) \(\forall x\). Thus for each \(c \in C_r\), \(h_{C|T}(c|\cdot) = h^{*^j}_{C|T}(c|\cdot)\) (induced by \(H^*\)) (except for a set of zero \(\mu_{H^*_T}\) measure) is a limiting point of \(\hat{f}_{C|T}\).

**Step 4** (conclusion). Verify that \(\sup_{x \in D_V} |\hat{F}^*_n(x) - F_T(x)| \rightarrow 0\). Under A1, for each \(t\), if \(n\) is large enough, there are \(M_{k_j}\), \(t \in (s_{n,j-1}, s_{n,j})\) and \(\delta_{k_j} = 1\) such that \(\forall c \in C_r\).
constants $b_{c_j}$ satisfying $\phi(\{c\}) = \sum_j b_{c_j}\phi(M_{k_j})$. It follows that $\hat{f}_{C|T}(\cdot|t)$ does not based on only one $w \in J_t$, at least for large enough $n$. In fact, if $n$ is large, each $w$ in $\{W \in J_t : t \in (s_{n,j-1}, s_{n,j})\}$ would be observed and thus $\hat{f}_{C|T}(\cdot|t)$ is based on all these $w$\'s. Then for the convergent subsequence $\{\tilde{F}_{nh}\}_{h \geq 1}$, the sequences $\{f_{C|T,n_h}\}_{h \geq 1}$ induced by $\{\tilde{F}_{nh}\}_{h \geq 1}$ converges to the limiting functions $h_{C|T}$ on the set $\{V_i : i = 1, ..., n\}$, which contains the support of $F_T$. Notice that each of the $||J||+1$ integrands of $B_{P^*}$ (see (7.4)), namely, $\frac{\tilde{F}_{nh}(x,y) - F_{nh}^*(v,y)}{1 - F_{T,n_h}(v)}$ and $\frac{\sum_{i \leq j < n} w f_{C|T,n_h}(i|v)}{\sum_{j \leq n} w f_{C|T,n_h}(i|v)}$ where $w \in J_t$, are bounded by 1 and converge as $h \to \infty$, for each $x \in D_V$, $y \in C_r$ and $v \in \{V_i : i = 1, ..., n\}$. Thus by the BCT, the limiting equation of $\tilde{F}^*(x,y) = B_{P^*}(Q)(x,y)$ is $H^*(x,y) = B_{H^*}(Q)(x,y)$ for all $(x,y) \in D_V \times C_r$. It can be shown that $S_{H^*} = S_F$ and $H^* = H$. By Lemma 7.13, the solution to $B_H(Q) = H$, $H = F$, is unique in the sense specified there. It follows that each limiting point $h_{C|T}$ of $f_{C|T}$ satisfies that $h_{C|T}(\cdot|t) = f_{C|T}(\cdot|t)$ for each $(t,c) \in \{(V_i : i = 1, ..., n) \times C_r\}$, and $\tilde{F}^* \to F$. Since $\tilde{F}_{c*}(x) = \int_{t \leq x} \hat{f}_{C|T}(t) dF_T(t)$, by the BCT, $\tilde{F}_{c*}(t) \to F_{c*}(t)$ on $D_V$ and $F \to F$ for the given $\omega$. Since $\omega$ is arbitrary in $\Omega_0$ and $P(\Omega_0) = 1$, the lemma is proved. 


§9.1. Remark 9.1. Dinse (1982, p.426) provides a data set with $J = 2$. Dinse comments that the GML of $f_{C|T}$ is “extremely erratic”, and partitions the observations into several equal-sized groups and derives a new smoothed estimator of $f_{C|T}$ based on the grouped data. The new estimator assigns positive weights to both $(V_i, 1)$ and $(V_i, 2)$, even if the MI is not $(V_i, 1) \times \{2\}$. Thus the new estimator is a GML based on the original data. Moreover, most discrete GMLEs of continuous density functions are not consistent, thus most of them are erratic. However, the cdfs of based on the inconsistent GMLE of the densities are often consistent, just like the case we are studying.

§9.2. Proofs in Example 2.1. Under given assumptions, the log likelihood function is

$$L = n_1 \ln(p_1 + p_2) + n_2 \ln(1 - p_1 - p_2) + n_3 \ln(p_1 + p_3) + n_4 \ln(1 - p_1 - p_3),$$

where $n_i = \sum_{j=1}^{n} 1_{\{M_i = W_{i}\}}$, $p_i = f(1, i)$ and $n = n_1 + \cdots + n_4$. The normal equations are

$$\frac{n_1}{p_1 + p_2} - \frac{n_2}{1 - p_1 - p_2} + \frac{n_3}{p_1 + p_3} - \frac{n_4}{1 - p_1 - p_3} = 0,$$

$$\frac{n_1}{p_1 + p_2} - \frac{n_2}{1 - p_1 - p_2} + \frac{n_3}{p_1 + p_3} - \frac{n_4}{1 - p_1 - p_3} = 0.$$

Reducing these two equations leads to the GMLE in Example 2.1: $p_2 = r_1 - p_1$, $p_3 = r_2 - p_1$, and $p_4 = 1 - p_1 - p_2 - p_3$, where $p_1$ is arbitrary in $[\max(0, r_1 + r_2 - 1), \min\{r_1, r_2\}]$, $r_1 = \frac{n_1}{n_3 + n_4}$ and $r_2 = \frac{n_3 + n_4}{n_3 + n_4}$. $

§9.3. Proofs in Example 2.2 (existence of both inconsistent GMLE and consistent GMLE).

Suppose that $J = 2$; partitions $P_0 = \{\{1\}, \{2\}\}$ and $P_1 = \{C_r\}$; $F_{c*}'$ satisfy $\frac{\partial}{\partial t} F_{c*}(t) = p_1 1_{\{t(1,2)\}}$ and $\frac{\partial}{\partial t} F_{c*}'(t) = p_2 1_{\{t(0,3)\}}$, where $p_1 + p_2 = 1$ and $p_i \geq 0$; $f_{\Delta}(h) = 1/2$; there is no censoring; $(T, C) \perp (\Delta, R)$ and A1 holds. Verify that

(a) all $n(V_i, \delta_i, M_i)$s are of the forms $(1, V_i, \{j\})$, $j \in \{1, 2\}$, or $(2, V_i, C_r)$;
(b) they are all distinct and thus each of them is an MI induced by these $n$ observations.

Thus the GMLE $\hat{F}_1$ assigns weight $\frac{1}{n}$ to $\{V_i \times \{j\}\}$ if $M_i = \{j\}$, and assigns weight $\frac{1}{2n}$ to $\{V_i \times \{j\}\}$ if $M_i = C_r$ and $j \in C_r$. Now $\hat{F}_1(t, 1) = \frac{1}{2n} \sum_{i=1}^{n} 1_{\{V_i \leq t, M_i = C_r\}} + \frac{1}{n} \sum_{i=1}^{n} 1_{\{V_i \leq t, M_i = \{1\}\}} \to \frac{1}{2} \frac{p_2 p_2}{f_{\Delta}(2)} = p_2/12 ≠ F(t, 1)$. Thus $\hat{F}_1$ is not consistent on $\{0, 1\}$. In the aforementioned example, since the data are of the form either (1) or (2), there is a consistent GMLE with a closed form solution: $\hat{F}_3(t, 1) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{V_i \leq t\}} \hat{f}_{C|T}(1|V_i)$, where $\hat{f}_{C|T}(1|t) = \begin{cases} \frac{n_1}{n_t} & \text{if } n_t \neq 0 \\ \frac{1}{2} & \text{otherwise} \end{cases}$.

$n_t = \sum_{i=1}^{n} 1_{\{V_i - t|t| \leq \frac{1}{2}, \delta_i = 1, M_i = \{1\} \} or \{2\}}$ and $n_{1t} = \sum_{i=1}^{n} 1_{\{V_i - t|t| \leq \frac{1}{2}, \delta_i = 1, M_i = \{1\}\}}$. $

§9.4. Proofs in Example 6.1. Suppose $T \equiv 1$. $J = 3$, $P_2 = \{\{1, 2\}, \{3\}\}$, $P_3 = \{\{1, 3\}, \{2\}\}$, $f_{\Delta}(2) = f_{\Delta}(3) = 1/2$. $R \equiv 2$. The possible observations are of the forms $(1, 1, W_i)$, $i = 1, \ldots, 4$, where $W_1 = \{1, 2\}$, $W_2 = \{2\}$, $W_3 = \{3\}$ and $W_4 = \{1, 3\}$. Let $N_j = \sum_{i=1}^{n} 1_{M_i = W_j}$,
$j = 1, \ldots, 4$. Then the GMLEs are $\hat{f}_{C|T}(2|1) = \frac{N_2}{N_2+N_4}$ and $\hat{f}_{C|T}(3|1) = \frac{N_3}{N_1+N_3}$, provided that $\hat{\theta} \leq 1$, where $\hat{\theta} = \frac{N_2}{N_2+N_4} + \frac{N_3}{N_1+N_3}$. However, if $f_C(2) = f_C(3) = 1/2$, then $\sqrt{n}(\hat{\theta} - 1)$ converges in distribution to $N(0,\sigma^2)$, where $\sigma > 0$. Thus $P\{\hat{\theta} > 1\} \rightarrow 1/2$. It follows that the GMLE $\hat{F}(1,1) = \hat{f}_{C|T}(1|1) = \begin{cases} 1 - \hat{\theta} & \text{if } \hat{\theta} < 1 \\ 0 & \text{otherwise} \end{cases}$ Thus it is not asymptotically normally distributed.