1 Poisson sampling

1.1 Poisson sampling, Horvitz-Thompson estimator

We make the following assumptions on finite population and sampling design:

Assumption 1.1.1. *The finite population* $U_N = \{y_1, \cdots, y_N\}$ *is a realization of a sequence of iid random variables from Uniform* $[0,1]$ *distribution.*

Assumption 1.1.2. *A Poisson sample* $A_n = \{y_1, \cdots, y_n\}$ *is selected from* U_N*, with inclusion probabilities* π_1 *between* 0 *and* 1*. Assume* $\kappa_N = \frac{1}{N}\sum_U \frac{1-\pi_i}{\pi_i}$*, and* $\kappa = \lim_{N\to\infty} \frac{1}{N}\sum_U \frac{1-\pi_i}{\pi_i}$ *is a well-defined finite quantity.*

Assumption 1.1.3. *Let* $\boldsymbol{d}_N = (\mathrm{I}_1, \cdots, \mathrm{I}_N)$ *be a random vector with each component supported on* $\{0,1\}$ *and* $\boldsymbol{d}_N$ *is independent of* $(y_1, \cdots, y_N)$*.*

We use the following estimator to estimate the distribution function,

$$\tilde{F}_n(y) = \frac{1}{N}\sum_A \frac{1}{\pi_i}\mathrm{I}_{(\,y_i \le y)} \tag{1}$$

To obtain asymptotic normality for finite dimensions, we denote,

$$\boldsymbol{z}_i = \begin{bmatrix} \mathrm{I}_{(y_i \le t)} \\ \mathrm{I}_{(y_i \le s)} \\ 1 \end{bmatrix}, \boldsymbol{z} = \begin{bmatrix} t \\ s \\ 1 \end{bmatrix},$$

for $0 \le s, t \le 1$.

We define

$$\bar{\boldsymbol{z}}_{HT} = \frac{1}{N}\sum_A \frac{1}{\pi_i}\boldsymbol{z}_i = \frac{1}{N}\sum_U \frac{1}{\pi_i}\boldsymbol{z}_i\mathrm{I}_i,$$

where I_i denotes sampling indicator.

Now, let us work out the expectation and variance of $\bar{\boldsymbol{z}}_{HT}$.

$$\begin{aligned} E(\bar{\boldsymbol{z}}_{HT}) &= E_\xi E_p \frac{1}{N}\sum_U \frac{1}{\pi_i}\boldsymbol{z}_i \mathrm{I}_i \\ &= E_\xi \left[E_p \left(\frac{1}{N}\sum_U \frac{1}{\pi_i}\boldsymbol{z}_i \mathrm{I}_i \middle| \mathcal{F}_N \right) \right] \\ &\overset{\text{A1.1.3}}{=} E_\xi \frac{1}{N}\sum_U \boldsymbol{z}_i = \boldsymbol{z}. \end{aligned}$$

$$\begin{aligned} Var(\bar{\boldsymbol{z}}_{HT} - \boldsymbol{z}) &= E_\xi \left[Var_p \left(\bar{\boldsymbol{z}}_{HT} | \mathcal{F}_N \right) \right] + Var_\xi \left[E_p(\bar{\boldsymbol{z}}_{HT} | \mathcal{F}_N) \right] \\ &\overset{A1.1.3}{=} \frac{1}{N^2} E_\xi \sum_U \frac{1-\pi_i}{\pi_i} \begin{bmatrix} \mathrm{I}_{(y_i \le t)} \\ \mathrm{I}_{(y_i \le s)} \\ 1 \end{bmatrix} \begin{bmatrix} \mathrm{I}_{(y_i \le t)} \\ \mathrm{I}_{(y_i \le s)} \\ 1 \end{bmatrix}^T + Var_\xi \left[\frac{1}{N}\sum_U \boldsymbol{z}_i \right] \\ &= \frac{N-1}{N^2} \begin{bmatrix} t & t\wedge s & t \\ t \wedge s & s & s \\ t & s & 1 \end{bmatrix} \kappa_N + \frac{1}{N} \begin{bmatrix} t(1-t) & t\wedge s - ts & 0 \\ t\wedge s - ts & s(1-s) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

where $t \wedge s$ denotes the smaller one between t and s.

Assume the limit of $Var(\bar{\boldsymbol{z}}_{HT} - \boldsymbol{z})$ to be $V(\bar{\boldsymbol{z}}_{HT} - \boldsymbol{z})$, which is of order $\frac{\kappa_N+1}{N}$. To establish asymptotic normality of $\bar{\boldsymbol{z}}_{HT}$, we define

$$\boldsymbol{x}_i = \frac{1}{N} Var^{-1/2}(\bar{\boldsymbol{z}}_{HT} - \boldsymbol{z}) \left\{ \frac{1}{\pi_i}\boldsymbol{z}_i \mathrm{I}_i - \boldsymbol{z} \right\},$$

then it is obvious that,

$$E_\xi E_p \boldsymbol{x}_i = \boldsymbol{0}, \text{and } \sum_U E_\xi E_p \boldsymbol{x}_i^T \boldsymbol{x}_i = \mathbb{I}_k,$$

where $\mathbb{I}_k$ denotes the identity matrix of order k. Assumption 1.1.2 suggests that $(\kappa_N + 1)\sqrt{N} \to \infty$, then $\|\boldsymbol{x}_i\| \to 0$ as $N \to \infty$. The following Lindeberg condition is then satisfied,

$$\lim_{N\to\infty}\sum_U E_\xi E_p\left[\|\boldsymbol{x}_i\|^2 \mathrm{I}_{(\|\boldsymbol{x}_i\|\geq\epsilon)}\right]=0.$$

Then we can use the multivariate Lindeberg CLT to conclude that

$$Var^{-1/2}(\bar{\boldsymbol{z}}_{HT}-\boldsymbol{z})(\bar{\boldsymbol{z}}_{HT}-\boldsymbol{z})\xrightarrow{d}N(\boldsymbol{0},\mathbb{I}_k), \tag{2}$$

where we replaced the covariance matrix $Var(\bar{\boldsymbol{z}}_{HT}-\boldsymbol{z})$ by its limit.

Now let us derive the finite dimensional distribution for the random function $\tilde{F}_n(y)$. For simplicity purpose, we only consider the joint distribution of $\tilde{F}_n(t)-t$ and $\tilde{F}_n(s)-s$. By Taylor linearization,

$$\begin{bmatrix}\tilde{F}_n(t)-t\\ \tilde{F}_n(s)-s\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}(\bar{\boldsymbol{z}}_{HT}-\boldsymbol{z})+\text{smaller order terms},$$

then we can use Delta method to conclude that

$$\sqrt{\frac{N^2}{\kappa_N(N-1)+N}}\begin{bmatrix}\tilde{F}_n(t)-t\\ \tilde{F}_n(s)-s\end{bmatrix}\xrightarrow{d}N\left(\boldsymbol{0},\begin{bmatrix}t(1-t)&t\wedge s-ts\\ t\wedge s-ts&s(1-s)\end{bmatrix}+\frac{\kappa_N}{\kappa_N+(1+\frac{1}{N-1})}\begin{bmatrix}t^2&ts\\ ts&s^2\end{bmatrix}\right),$$

which has the same finite dimensional distribution as the following stochastic process,

$$B(y)=W^0(y)+yX_0\sqrt{\frac{\kappa_N}{\kappa_N+(1+\frac{1}{N-1})}}, \tag{3}$$

after multiplying by a constant on both sizes. Here, $W^0(y)$ is a Brownian bridge with $W(0)=0, W(1)=0$, and $X_0\sim N(0,1)$ independent of W^0. A Brownian bridge $W^0(t)\overset{d}{=}W(t)-tW(1)$, where $W(t)$ is a standard Wiener process with mean 0 and variance t.

Now, let us show that

$$E\left[|\tilde{F}_n(t)-\tilde{F}_n(t_1)|^2|\tilde{F}_n(t_2)-\tilde{F}_n(t)|^2\right]\leq [H(t_2)-H(t_1)]^{2\alpha},$$

for some continuous, nondecreasing function $H(\cdot)$ on $[0,1]$, and $0\leq t_1\leq t\leq t_2\leq 1$. We need to make the assumption of $\frac{1}{N^2}\sum_i \frac{1}{\pi_i}<\infty$ in the limit.

$$\begin{aligned}
& E\left[|\tilde{F}_n(t)-\tilde{F}_n(t_1)|^2|\tilde{F}_n(t_2)-\tilde{F}_n(t)|^2\right] \\
= \;& c_1\frac{1}{N^4}E_\xi E_p \sum_{i_1,i_2,j_1,j_2}\frac{1}{\pi_{i_1}\pi_{i_2}\pi_{j_1}\pi_{j_2}}\mathrm{I}_{(t_1\leq y_{i_1}\leq t)}\mathrm{I}_{(t_1\leq y_{i_2}\leq t)}\mathrm{I}_{(t\leq y_{j_1}\leq t_2)}\mathrm{I}_{(t\leq y_{j_2}\leq t_2)}\mathrm{I}_{i_1}\mathrm{I}_{i_2}\mathrm{I}_{j_1}\mathrm{I}_{j_2} \\
\leq \;& \underbrace{c_1\frac{1}{N^4}\frac{N!}{(N-4)!}(t-t_1)^2(t_2-t)^2}_{\text{None of } i_1,i_2,j_1,j_2 \text{ are equal}} \\
& +\underbrace{c_1\frac{1}{N^4}\sum_i\frac{1}{\pi_i}(t-t_1)(t_2-t)^2N(N-1)}_{i_1=i_2,j_1\neq j_2} \\
& +\underbrace{c_1\frac{1}{N^4}\sum_j\frac{1}{\pi_j}(t-t_1)^2(t_2-t)N(N-1)}_{i_1\neq i_2,j_1=j_2} \\
& +\underbrace{c_1\frac{1}{N^4}\sum_{i\neq j}\frac{1}{\pi_i\pi_j}(t-t_1)(t_2-t)N(N-1)}_{i_1=i_2,j_1=j_2} \\
\leq \;& c_2(t-t_1)(t_2-t) \\
\leq \;& c_2(t_2-t_1)^2.
\end{aligned}$$

Now, by Theorem 15.6 of Billingsley (1968), we can conclude that,

$$\sqrt{\frac{N^2}{\kappa_N(N-1)+N}}\left(\tilde{F}_n(y)-F(y)\right)\xrightarrow{d}B(y), \tag{4}$$

and then,

$$\sqrt{\frac{N^2}{\kappa_N(N-1)+N}}\sup_y\left|\left(\tilde{F}_n(y)-F(y)\right)\right|\xrightarrow{d}\sup_y|B(y)|. \tag{5}$$

1.2 Poisson sampling, Hajek estimator

Let $U_N = \{y_1, \cdots, y_N\}$ be a collection of *iid* random variables from Uniform$[0,1]$ distribution, and $\pi_1, \cdots, \pi_N$ be a vector of constants between 0 and 1 associated with each random variable. A Poisson sample $A_n = \{y_1, \cdots, y_n\}$ is drawn from U_N and the following estimator can be used to estimate the distribution function,

$$\hat{F}_n(y) = \frac{\sum_A \frac{1}{\pi_i} \mathrm{I}_{(y_i \le y)}}{\sum_A \frac{1}{\pi_i}} \tag{6}$$

Now let us derive the finite dimensional distribution for the random function $\hat{F}_n(y)$. For simplicity purpose, we only consider the joint distribution of $\hat{F}_n(t) - t$ and $\hat{F}_n(s) - s$. By Taylor linearization,

$$\begin{bmatrix} \hat{F}_n(t) - t \\ \hat{F}_n(s) - s \end{bmatrix} = \begin{bmatrix} 1 & 0 & -t \\ 0 & 1 & -s \end{bmatrix} (\bar{\boldsymbol{z}}_{HT} - \boldsymbol{z}) + \text{smaller order terms},$$

then we can use Delta method to conclude that

$$\begin{bmatrix} \hat{F}_n(t) - t \\ \hat{F}_n(s) - s \end{bmatrix} \xrightarrow{d} N\left(\mathbf{0}, \frac{(N-1)\kappa_N + N}{N^2} \begin{bmatrix} t(1-t) & t \wedge s - ts \\ t \wedge s - ts & s(1-s) \end{bmatrix}\right),$$

given (2). The limiting distribution is the same as finite dimensional distribution of Brownian bridge $W^0(\cdot)$ after multiplying by $\sqrt{\frac{N^2}{(N-1)\kappa_N + N}}$ on both sides. A Brownian bridge $W^0(t) \stackrel{d}{=} W(t) - tW(1)$, where $W(t)$ is a standard Wiener process with mean 0 and variance t.

Now, let us show that

$$E\left[|\hat{F}_n(t) - \hat{F}_n(t_1)|^2 |\hat{F}_n(t_2) - \hat{F}_n(t)|^2\right] \le [H(t_2) - H(t_1)]^{2\alpha},$$

for some continuous, nondecreasing function $H(\cdot)$ on $[0,1]$, and $0 \le t_1 \le t \le t_2 \le 1$.

We need to make the assumption of $\frac{1}{N^2}\sum_i \frac{1}{\pi_i} < \infty$ in the limit.

$$
\begin{aligned}
& E\left[|\hat{F}_n(t) - \hat{F}_n(t_1)|^2 |\hat{F}_n(t_2) - \hat{F}_n(t)|^2\right] \\
= \; & E\frac{\left(\sum_A \frac{1}{\pi_i}\mathrm{I}_{(t_1 \le y_i \le t)}\right)^2 \left(\sum_A \frac{1}{\pi_i}\mathrm{I}_{(t \le y_i \le t_2)}\right)^2}{\left(\sum_A \frac{1}{\pi_i}\right)^4} \\
\le \; & c_1 \frac{1}{N^4} E\left(\sum\nolimits_A \frac{1}{\pi_i}\mathrm{I}_{(t_1 \le y_i \le t)}\right)^2 \left(\sum\nolimits_A \frac{1}{\pi_i}\mathrm{I}_{(t \le y_i \le t_2)}\right)^2 \\
\le \; & c_1 \frac{1}{N^4} E \sum_{i_1, i_2, j_1, j_2} \frac{1}{\pi_{i_1}\pi_{\pi_2}\pi_{j_1}\pi_{j_2}} \mathrm{I}_{(t_1 \le y_{i_1} \le t)}\mathrm{I}_{(t_1 \le y_{i_2} \le t)}\mathrm{I}_{(t \le y_{j_1} \le t_2)}\mathrm{I}_{(t \le y_{j_2} \le t_2)}\mathrm{I}_{i_1}\mathrm{I}_{i_2}\mathrm{I}_{j_1}\mathrm{I}_{j_2} \\
\le \; & \underbrace{c_1 \frac{1}{N^4} \frac{N!}{(N-4)!}(t - t_1)^2 (t_2 - t)^2}_{\text{None of } i_1, i_2, j_1, j_2 \text{ are equal}} \\
& + \underbrace{c_1 \frac{1}{N^4} \sum_i \frac{1}{\pi_i}(t - t_1)(t_2 - t)^2 N(N-1)}_{i_1 = i_2, j_1 \ne j_2} \\
& + \underbrace{c_1 \frac{1}{N^4} \sum_j \frac{1}{\pi_j}(t - t_1)^2 (t_2 - t) N(N-1)}_{i_1 \ne i_2, j_1 = j_2} \\
& + \underbrace{c_1 \frac{1}{N^4} \sum_{i \ne j} \frac{1}{\pi_i \pi_j}(t - t_1)(t_2 - t) N(N-1)}_{i_1 = i_2, j_1 = j_2} \\
\le \; & c_2 (t - t_1)(t_2 - t) \\
\le \; & c_2 (t_2 - t_1)^2.
\end{aligned}
$$

Now, by Theorem 15.6 of Billingsley (1968), we can conclude that,

$$
\sqrt{\frac{N^2}{\kappa_N(N-1) + N}} \left(\hat{F}_n(y) - F(y)\right) \xrightarrow{d} W^0(y), \tag{7}
$$

and then,

$$
\sqrt{\frac{N^2}{\kappa_N(N-1) + N}} \sup_y \left|\left(\hat{F}_n(y) - F(y)\right)\right| \xrightarrow{d} \sup_y \left|W^0(y)\right|. \tag{8}
$$

The standard theory on Kolmogorov-Smirnov test would give

$$P(\sup_{y} \left|W^0(y)\right| \leq x) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2x^2}, 0 < x < \infty,$$

which provides the theoretical distribution of our test statistic.

1.3 General sampling design, Hajek estimator

Assumption 1.3.1. *The finite population $U_N = \{y_1, \cdots, y_N\}$ is a realization of a sequence of iid random variables from Uniform $[0,1]$ distribution.*

We consider a general sampling design, but assume the design normality for any nonrandom vector $\boldsymbol{z}$ with suitable moment conditions.

Assumption 1.3.2. *Assume that for any $\boldsymbol{z}$ with positive variance-covariance matrix and finite $2+\delta$ moment*

$$n^{*1/2}(\bar{\boldsymbol{z}}_\pi - \bar{\boldsymbol{z}}_N)|\mathcal{F}_N \xrightarrow{d} N(\mathbf{0}, \Sigma_{\boldsymbol{z}}), \quad a.s. \tag{9}$$

where $\bar{\boldsymbol{z}}_N = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{z}_i$ and $\bar{\boldsymbol{z}}_\pi = \frac{1}{N}\sum_{A} \frac{\boldsymbol{z}_i}{\pi_i}$.

Assumption 1.3.3. *We have the following assumptions on inclusion and joint inclusion probabilities,*

1. *$K_L \le \pi_i \frac{N}{n^*} \le K_U$ for all i where K_L and K_U are positive constants.*
2. *$\sum\sum\sum \frac{\pi_{ijk}}{\pi_j\pi_j\pi_k} = O(N^3), a.s.$*
3. *$\frac{\pi_{ijkl}}{\pi_j\pi_j\pi_k\pi_l}$ are uniformly bounded by a positive constant c_3, for all unequal i,j,k,l.*

Let $\boldsymbol{z}_i = (\mathrm{I}_{(y_i \le t)}, \mathrm{I}_{(y_i \le s)}, 1)'$, Assumption 1.3.2 implies that

$$\frac{n^{*1/2}}{N}\sum_U \left\{ \frac{1}{\pi_i}\begin{bmatrix} \mathrm{I}_{(y_i\le t)} \\ \mathrm{I}_{(y_i\le s)} \\ 1 \end{bmatrix} - \begin{bmatrix} F_N(t) \\ F_N(s) \\ 1 \end{bmatrix} \right\}\Bigg| \mathcal{F}_N \xrightarrow{d} N(0, \Sigma_{\boldsymbol{z}}), \quad a.s.,$$

where $\Sigma_{\boldsymbol{z}} = \frac{n^*}{N}\sum_i\sum_j \Delta_{ij}\frac{\boldsymbol{z}_i\boldsymbol{z}_i'}{\pi_i\pi_j}$.

Now we can use Taylor linearization to find the finite dimensional distribution of

$\hat{F}_n(y)$ as defined in (6).

$$\begin{bmatrix} \hat{F}_n(t) - F_N(t) \\ \hat{F}_n(s) - F_N(s) \end{bmatrix} = \begin{bmatrix} 1 & 0 & -t \\ 0 & 1 & -s \end{bmatrix} (\bar{\boldsymbol{z}}_{HT} - \bar{\boldsymbol{z}}_N) + \text{smaller order terms},$$

and obtain the following asymptotic normality,

$$n^{*1/2} \begin{bmatrix} \hat{F}_n(t) - F_N(t) \\ \hat{F}_n(s) - F_N(s) \end{bmatrix} \Bigg| \mathcal{F}_N \xrightarrow{d} N\left(0, \Sigma_{HA}\right), \ a.s., \tag{10}$$

where

$$\Sigma_{HA} = \frac{n^*}{N} \sum_i \sum_j \frac{\Delta_{ij}}{\pi_i \pi_j} \begin{bmatrix} (\mathrm{I}_{(y_i \le t)} - F_N(t))^2 & (\mathrm{I}_{(y_i \le t)} - F_N(t))(\mathrm{I}_{(y_i \le s)} - F_N(s)) \\ (\mathrm{I}_{(y_i \le t)} - F_N(t))(\mathrm{I}_{(y_i \le s)} - F_N(s)) & (\mathrm{I}_{(y_i \le s)} - F_N(s))^2 \end{bmatrix},$$

where $\Delta_{ij} = \pi_{ij} - \pi_i \pi_j$.

It follows from the standard Central Limit Theorem that

$$n^{*1/2} \begin{bmatrix} F_N(t) - t \\ F_N(s) - s \end{bmatrix} \xrightarrow{d} N\left(0, \frac{n^*}{N} \begin{bmatrix} t - t^2 & (1 - t \vee s)(t \wedge s) \\ (1 - t \vee s)(t \wedge s) & s - s^2 \end{bmatrix}\right). \tag{11}$$

Assume there exists a function $C(t,s)$ defined on $[0,1] \times [0,1]$, such that

$$C(t,s) = \lim_{N \to \infty} \frac{n^*}{N} \sum \sum_U \Delta_{ij} \frac{\left(\mathrm{I}_{(y_i \le t)} - F_N(t)\right)\left(\mathrm{I}_{(y_i \le s)} - F_N(s)\right)}{\pi_i \pi_j} + \lim_{N \to \infty} \frac{n^*}{N} (1 - t \vee s)(t \wedge s), \ a.s., \tag{12}$$

where $t \vee s = \max\{t, s\}$, $t \wedge s = \min\{t, s\}$.

We can verify that $C(t,s)$ is a non-negative definite kernel function, and by Lemma

1.3.1 of Fuller (2006),

$$n^{*1/2}\begin{bmatrix}\hat{F}_n(t)-t\\ \hat{F}_n(s)-s\end{bmatrix}\xrightarrow{d} N\left(0,\begin{bmatrix}C(t,t) & C(t,s)\\ C(t,s) & C(s,s)\end{bmatrix}\right), \tag{13}$$

for all $t,s\in[0,1]$. The asymptotic normality in (13) can be generalized to all finite dimensions fairly easily.

Now, let us show that

$$E\left[|\hat{F}_n(t)-\hat{F}_n(t_1)|^2|\hat{F}_n(t_2)-\hat{F}_n(t)|^2\right]\le[H(t_2)-H(t_1)]^{2\alpha},$$

for some continuous, nondecreasing function $H(\cdot)$ on $[0,1]$, and $0\le t_1\le t\le t_2\le 1$.

$$\begin{aligned}
&E\left[|\hat{F}_n(t)-\hat{F}_n(t_1)|^2|\hat{F}_n(t_2)-\hat{F}_n(t)|^2\right]\\
=\;& E\frac{\left(\sum_A\frac{1}{\pi_i}\mathrm{I}_{(t_1\le y_i\le t)}\right)^2\left(\sum_A\frac{1}{\pi_i}\mathrm{I}_{(t\le y_i\le t_2)}\right)^2}{\left(\sum_A\frac{1}{\pi_i}\right)^4}\\
\le\;& c_1\frac{1}{N^4}E\left(\sum\nolimits_A\frac{1}{\pi_i}\mathrm{I}_{(t_1\le y_i\le t)}\right)^2\left(\sum\nolimits_A\frac{1}{\pi_i}\mathrm{I}_{(t\le y_i\le t_2)}\right)^2\\
=\;& c_1\frac{1}{N^4}E\sum_{i_1,i_2,j_1,j_2}\frac{1}{\pi_{i_1}\pi_{\pi_2}\pi_{j_1}\pi_{j_2}}\mathrm{I}_{(t_1\le y_{i_1}\le t)}\mathrm{I}_{(t_1\le y_{i_2}\le t)}\mathrm{I}_{(t\le y_{j_1}\le t_2)}\mathrm{I}_{(t\le y_{j_2}\le t_2)}\mathrm{I}_{i_1}\mathrm{I}_{i_2}\mathrm{I}_{j_1}\mathrm{I}_{j_2}\\
\le\;& \underbrace{c_1c_3\frac{1}{N^4}\frac{N!}{(N-4)!}(t-t_1)^2(t_2-t)^2}_{\text{None of } i_1,i_2,j_1,j_2 \text{ are equal}}+\underbrace{\frac{1}{N^4}\sum_{i,j,k}\frac{\pi_{i,j,k}}{\pi_i\pi_j\pi_k}}_{\text{o(1), a.s., by Assumption 1.3.3}}\\
\le\;& c_2(t-t_1)(t_2-t)\\
\le\;& c_2(t_2-t_1)^2.
\end{aligned}$$

Now the conditions are satisfied to apply Theorem 15.6 of Billingsley (1968), and we can conclude that $n^{*1/2}(\hat{F}_n(y)-y)$ converges to a Gaussian process with zero mean and covariance function $C(t,s)$.

Remark: In stratified sampling, given Assumption 1.1.3, we can simplify the expression of $C(t,s)$ as,

$$C(t,s) = \lim_{N\to\infty} n \sum_{h=1}^{H} W_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{1}{n_h}(1 - t \vee s)(t \wedge s) + \lim_{N\to\infty} \frac{n}{N}(1 - t \vee s)(t \wedge s), a.s.,$$

where n_h and N_h denote the stratum sample size and strutum size, respectively, $W_h = \frac{n_h}{N_h}$, and H is the number of strata.

Remark: We can not simplify the covariance expression (12) for general sampling designs, as the right hand side usually involves dependent summands, which makes it very difficult to show almost sure convergence even when the sampling is non-informative.

References

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sons.

Fuller, W. (2006). *Sampling Statistics.*