1 The Hájek Central Limit Theorem for Simple Random Sampling without Replacement

Suppose we are given a sequence of finite populations, on each of which is defined an $\mathbb{R}$-valued population characteristic. Suppose on each of the finite populations, we use SRSWOR (of fixed sample size) to select a sample, observe the values of the corresponding population characteristics on the selected elements, and use the Horvitz-Thompson estimator for the population mean. We seek to determine a necessary and sufficient condition for the (associated sequence of) "standardized deviations from the mean" of the Horvitz-Thompson estimator for population mean to converge in distribution to the standard Gaussian distribution.

Theorem 1.1 (The Hájek Central Limit Theorem for SRSWOR)
Suppose we have the following:

- *Let* $\{U_\nu\}_{\nu\in\mathbb{N}}$ *be a sequence of finite populations, and* $N_\nu = |U_\nu| \geq 2$ *be the population size of* U_ν. *Let the elements of* U_ν *be indexed by* $1, 2, 3, \ldots, N_\nu$.
- *For each* $\nu \in \mathbb{N}$, *let* $y^{(\nu)} : U_\nu \longrightarrow \mathbb{R}$ *be a non-constant* $\mathbb{R}$*-valued population characteristic. For each* $i \in U_\nu$, *let* $y_i^{(\nu)}$ *denote* $y^{(\nu)}(i)$, *the value of* $y^{(\nu)}$ *evaluated at the* i^{th} *element of* U_ν.
- *For each* $\nu \in \mathbb{N}$, *let* $n_\nu \in \{1, 2, 3, \ldots, N_\nu - 1\}$ *be given, and let* $\mathcal{S}_\nu$ *be the set of all* n_ν*-element subsets of* U_ν. *Let* $\mathcal{S}_\nu$ *be endowed with the uniform probability function, namely*

$$P(s) \;=\; \frac{1}{\left(\begin{array}{c} N_\nu \\ n_\nu \end{array}\right)}, \;\; \text{for each } s \in \mathcal{S}_\nu.$$

- *For each* $\nu \in \mathbb{N}$, *let* $\widehat{\overline{Y}}_\nu : \mathcal{S}_\nu \longrightarrow \mathbb{R}$ *be the random variable defined as follows:*

$$\widehat{\overline{Y}}_\nu(s) \;\;:=\;\; \frac{1}{n_\nu}\sum_{i\in s} y_i^{(\nu)}, \;\; \text{for each } s \in \mathcal{S}_\nu$$

Let

$$\mu_\nu \;:=\; E\left[\,\widehat{\overline{Y}}_\nu\,\right] \;=\; \frac{1}{N_\nu}\sum_{i\in U_\nu} y_i^{(\nu)} \quad\text{and}\quad \sigma_\nu^2 \;:=\; \operatorname{Var}\left[\,\widehat{\overline{Y}}_\nu\,\right] \;=\; \left(1 - \frac{n_\nu}{N_\nu}\right)\frac{S_\nu^2}{n_\nu}\,,$$

where

$$S_\nu^2 \;:=\; \frac{1}{N_\nu - 1}\sum_{i\in U_\nu}\left(y_i^{(\nu)} - \mu_\nu\right)^2 \;\;>\;\; 0 \quad \left(\text{since } y^{(\nu)} : U^{(\nu)} \longrightarrow \mathbb{R} \text{ is non-constant}\right).$$

- *For each* $\nu \in \mathbb{N}$ *and each* $\delta > 0$ *define:*

$$U_\nu(\delta) \;:=\; \left\{\, i \in U_\nu \;\middle|\; |y_i^{(\nu)} - \mu_\nu| \geq \delta\sqrt{\sigma_\nu^2} \,\right\} \;\subset\; U_\nu.$$

Suppose $n_\nu \longrightarrow \infty$ *and* $N_\nu - n_\nu \longrightarrow \infty$. *Then,*

$$\lim_{\nu\to\infty} P\left\{\, s \in \mathcal{S}_\nu \;\middle|\; \frac{\widehat{\overline{Y}}_\nu(s) - \mu_\nu}{\sqrt{\sigma_\nu^2}} < x \,\right\} \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,\mathrm{d}t, \quad \text{for each } x \in \mathbb{R}$$

if and only if

$$\lim_{\nu\to\infty} \frac{\sum_{i\in U_\nu(\delta)} \left(y_i^{(\nu)} - \mu_\nu\right)^2}{\sum_{i\in U_\nu} \left(y_i^{(\nu)} - \mu_\nu\right)^2} \;=\; 0, \quad \text{for every } \delta > 0.$$

Outline of proof For each $\nu \in \mathbb{N}$, deploying the Hájek Sampling Design (see Definition 1.3 below) of size n_ν on each U_ν yields a pair of samples $\left(s_\nu^{(0)}, s_\nu^{(1)}\right)$, where $s_\nu^{(0)}$ is a simple random sample of U_ν of sample size n_ν, and $s_\nu^{(1)}$ is a Bernoulli sample of U_ν. □

Lemma 1.2
Bernoulli sampling from a finite population U of size N with individual selection probability n/N, where $n = 1, 2, \ldots, N$, is equivalent to the following two-step sampling scheme:

- ***Step 1:*** *Sample k from* $\text{Binomial}\,(N, n/N)$.
- ***Step 2:*** *Take an SRSWOR sample s of size k from U.*

Proof Note that the collection of possible samples for both schemes is the power set $\mathcal{P}(U)$ of U, i.e. all possible subsets of U. Let P_{B} and P_1 be the probability functions defined on $\mathcal{P}(U)$ under Bernoulli sampling and the two-step scheme, respectively. Then,

$$P_{\text{B}}(s) \;=\; \left(\frac{n}{N}\right)^k \left(1 - \frac{n}{N}\right)^{N-k}, \quad \text{for each } s \in \mathcal{P}(U), \text{ where } k = |\,s\,|.$$

On the other hand,

$$\begin{aligned} P_1(s) &= P(\; S = s \mid S \sim \text{SRSWOR}(k, N)\;) \cdot P(\; K = k \mid K \sim \text{Binomial}(N, n/N)\;) \\ &= \frac{1}{\binom{N}{k}} \cdot \binom{N}{k} \left(\frac{n}{N}\right)^k \left(1 - \frac{n}{N}\right)^{N-k} \\ &= \left(\frac{n}{N}\right)^k \left(1 - \frac{n}{N}\right)^{N-k}, \quad \text{for each } s \in \mathcal{P}(U), \text{ where } k = |\,s\,|. \end{aligned}$$

Thus, $P_{\text{B}} = P_1$ as (probability) functions on $\mathcal{P}(U)$. Hence, the two sampling schemes are equivalent. □

Definition 1.3 (The Hájek Sampling Design of size n)
*Suppose U is a finite population of size $N \in \mathbb{N}$ with $N \geq 3$. Let $n \in \{2, \ldots, N\}$ be fixed. Let $\mathcal{P}(U)$ be the power set of U. Let $\mathcal{S}(U, n)$ be the collection of all subsets of U with exactly n elements. The **Hájek Sampling Design of size n on** U, by definition, selects an ordered pair of samples $\left(s^{(0)}, s^{(1)}\right) \in \mathcal{S}(U, n) \times \mathcal{P}(U)$ as follows:*

- *First, select $k \in \{0, 1, 2, \ldots, N\}$ based on the binomial distribution* $\text{Binomial}(N, n/N)$.

  More precisely, let $K \sim$ $\text{Binomial}(N, n/N)$*, i.e. let K be a random variable following the binomial distribution with number of trials N and probability of success n/N. In other words,*

  $$P(K = k) \;=\; \binom{N}{k} \cdot \left(\frac{n}{N}\right)^k \cdot \left(1 - \frac{n}{N}\right)^{N-k}, \quad \text{for each } k = 0, 1, 2, \ldots, N.$$

  Let $k \in \{0, 1, 2, \ldots, N\}$ be a realization of the random variable $K \sim$ $\text{Binomial}(N, n/N)$.

- *If $k = n$, take an SRSWOR sample $s^{(0)} \subset U$ of size n, and let $s^{(1)} = s^{(0)}$.*
- *If $k > n$, take an SRSWOR sample $s^{(1)} \subset U$ of size k. Then, select an SRSWOR sample $s^{(0)}$ of $s^{(1)}$ of size n.*

- *If $k < n$, take an SRSWOR sample $s^{(0)} \subset U$ of size n. Then, select an SRSWOR sample $s^{(1)}$ of $s^{(0)}$ of size k.*

Remark 1.4

Note that the Hájek Sampling Design defines implicitly a probability function P_{H} on $\mathcal{S}(U,n) \times \mathcal{P}(U)$, making it a finite probability space. More explicitly, for each $\left(s^{(0)}, s^{(1)}\right) \in \mathcal{S}(U,n) \times \mathcal{P}(U)$, writing $k = \vert\, s^{(1)} \,\vert$, we have

$$P_{\mathrm{H}}\Big(s^{(0)}, s^{(1)}\Big) \;=\; \begin{cases} \left(\begin{array}{c} N \\ n \end{array}\right)\left(\dfrac{n}{N}\right)^{n}\left(1-\dfrac{n}{N}\right)^{N-n} \cdot \dfrac{1}{\left(\begin{array}{c} N \\ n \end{array}\right)}, & \text{if } s^{(0)} = s^{(1)} \\[2em] \left(\begin{array}{c} N \\ k \end{array}\right)\left(\dfrac{n}{N}\right)^{k}\left(1-\dfrac{n}{N}\right)^{N-k} \cdot \dfrac{1}{\left(\begin{array}{c} N \\ k \end{array}\right)} \cdot \dfrac{1}{\left(\begin{array}{c} k \\ n \end{array}\right)}, & \text{if } s^{(0)} \subsetneq s^{(1)} \\[2em] \left(\begin{array}{c} N \\ k \end{array}\right)\left(\dfrac{n}{N}\right)^{k}\left(1-\dfrac{n}{N}\right)^{N-k} \cdot \dfrac{1}{\left(\begin{array}{c} N \\ n \end{array}\right)} \cdot \dfrac{1}{\left(\begin{array}{c} n \\ k \end{array}\right)}, & \text{if } s^{(0)} \supsetneq s^{(1)} \\[2em] 0, & \text{otherwise} \end{cases}$$

Lemma 1.5 (Properties of the Hájek Sampling Design)

Suppose U is a finite population of size $N \in \mathbb{N}$ with $N \geq 3$. Let $n \in \{2, \ldots, N\}$ be fixed. Let $P_{\mathrm{H}} : \mathcal{S}(U,n) \times \mathcal{P}(U) \longrightarrow [0,1]$ be the Hájek Sampling Design. Then, the following statements are true:

(a) *The marginal sampling design induced on $\mathcal{S}(U,n)$ by P_{H} is $\mathrm{SRSWOR}(U,n)$.*

(b) *The marginal sampling design induced on $\mathcal{P}(U)$ by P_{H} is Bernoulli Sampling from U with unit selection probability n/N.*

(c) *For each fixed $k \in \{n+1, n+2, \ldots, N\}$, the sampling design induced on $\mathcal{S}(U, k-n)$ by pushing forward the conditional sampling design of $P_{\mathrm{H}}\vert_{\vert S^{(1)}\vert = k}$ via the following map:*

$$\left\{\; \Big(s^{(0)}, s^{(1)}\Big) \in \mathcal{S}(U,n) \times \mathcal{P}(U) \;\Big\vert\; \vert\, s^{(1)} \,\vert = k \;\right\} \longrightarrow \mathcal{S}(U, k-n) : \Big(s^{(0)}, s^{(1)}\Big) \longmapsto s^{(1)} \backslash\, s^{(0)}$$

is equivalent to $\mathrm{SRSWOR}(U, k-n)$.

(d) *For each fixed $k \in \{0, 1, 2, \ldots, n-1\}$, the sampling design induced on $\mathcal{S}(U, n-k)$ by pushing forward the pertinent restriction of P_{H} via the following map:*

$$\left\{\; \Big(s^{(0)}, s^{(1)}\Big) \in \mathcal{S}(U,n) \times \mathcal{P}(U) \;\Big\vert\; \vert\, s^{(1)} \,\vert = k \;\right\} \longrightarrow \mathcal{S}(U, n-k) : \Big(s^{(0)}, s^{(1)}\Big) \longmapsto s^{(0)} \backslash\, s^{(1)}$$

is equivalent to $\mathrm{SRSWOR}(U, n-k)$.

PROOF

(a) For each $s^{(0)} \in \mathcal{S}(U,n)$, it suffices to show that the marginal probability $P_{\mathrm{H}}\big(s^{(0)}, \,\cdot\,\big)$ is given by:

$$P_{\mathrm{H}}\Big(s^{(0)}, \,\cdot\,\Big) \;=\; \frac{1}{\left(\begin{array}{c} N \\ n \end{array}\right)}$$

To this end,

$$\begin{aligned}
P_{\mathrm{H}}\left(s^{(0)},\,\cdot\,\right) &= \sum_{s^{(1)}=s^{(0)}} P_{\mathrm{H}}\left(s^{(0)},s^{(1)}\right) + \sum_{s^{(1)}\supsetneq s^{(0)}} P_{\mathrm{H}}\left(s^{(0)},s^{(1)}\right) + \sum_{s^{(1)}\subsetneq s^{(0)}} P_{\mathrm{H}}\left(s^{(0)},s^{(1)}\right) \\
&= \binom{N}{n}\left(\frac{n}{N}\right)^{n}\left(1-\frac{n}{N}\right)^{N-n}\cdot\frac{1}{\binom{N}{n}} \\
&\quad + \sum_{k=n+1}^{N}\binom{N}{k}\left(\frac{n}{N}\right)^{k}\left(1-\frac{n}{N}\right)^{N-k}\cdot\frac{1}{\binom{N}{k}}\cdot\frac{1}{\binom{k}{n}}\cdot\binom{N-n}{k-n} \\
&\quad + \sum_{k=0}^{n-1}\binom{N}{k}\left(\frac{n}{N}\right)^{k}\left(1-\frac{n}{N}\right)^{N-k}\cdot\frac{1}{\binom{N}{n}}\cdot\frac{1}{\binom{n}{k}}\cdot\binom{n}{k}
\end{aligned}$$

We remark that, for a given $s^{(0)}\in\mathcal{S}(U,n)$ and $k>n$, the quantity $\binom{N-n}{k-n}$ is the number of elements in $\mathcal{P}(U)$ (i.e. number of subsets of U) of size k containing $s^{(0)}$ as a proper subset. Note also that, for $k>n$,

$$\frac{1}{\binom{N}{k}}\cdot\frac{1}{\binom{k}{n}}\cdot\binom{N-n}{k-n} \;=\; \frac{k!(N-k)!}{N!}\cdot\frac{n!(k-n)!}{k!}\cdot\frac{(N-n)!}{(k-n)!(N-k)!} \;=\; \frac{n!(N-n)!}{N!} \;=\; \frac{1}{\binom{N}{n}}.$$

Hence, we have

$$P_{\mathrm{H}}\left(s^{(0)},\,\cdot\,\right) \;=\; \frac{1}{\binom{N}{n}}\cdot\sum_{k=0}^{N}\binom{N}{k}\left(\frac{n}{N}\right)^{k}\left(1-\frac{n}{N}\right)^{N-k} \;=\; \frac{1}{\binom{N}{n}}\cdot 1 \;=\; \frac{1}{\binom{N}{n}}$$

(b) For each $s^{(1)}\in\mathcal{P}(U)$, it suffices to show that the marginal probability $P_{\mathrm{H}}\left(\,\cdot\,,s^{(1)}\right)$ is given by:

$$P_{\mathrm{H}}\left(\,\cdot\,,s^{(1)}\right) \;=\; \left(\frac{n}{N}\right)^{k}\cdot\left(1-\frac{n}{N}\right)^{N-k}, \quad \text{where } k=|\,s^{(1)}\,|.$$

To this end, first note that either $k=|\,s^{(1)}\,|\geq n$ holds, or $k=|\,s^{(1)}\,|<n$ holds. In the first case, i.e. $k=|\,s^{(1)}\,|\geq n$, we have

$$\begin{aligned}
P_{\mathrm{H}}\left(\,\cdot\,,s^{(1)}\right) &= P\left(S^{(1)}=s^{(1)}\;\middle|\;K=k\right)\cdot P(K=k) \\
&= \frac{1}{\binom{N}{k}}\cdot\binom{N}{k}\left(\frac{n}{N}\right)^{k}\left(1-\frac{n}{N}\right)^{N-k} \\
&= \left(\frac{n}{N}\right)^{k}\cdot\left(1-\frac{n}{N}\right)^{N-k}.
\end{aligned}$$

In the second case, i.e. $k = |\,s^{(1)}\,| < n$, we have

$$\begin{aligned}
P_{\mathrm{H}}\Big(\,\cdot\,,s^{(1)}\Big) &= \sum_{s^{(0)}\supsetneq s^{(1)}} P_{\mathrm{H}}\Big(s^{(0)},s^{(1)}\Big) \;=\; \sum_{s^{(0)}\supsetneq s^{(1)}} \binom{N}{k}\left(\frac{n}{N}\right)^k\left(1-\frac{n}{N}\right)^{N-k}\cdot\frac{1}{\binom{N}{n}}\cdot\frac{1}{\binom{n}{k}} \\
&= \binom{N-k}{n-k}\cdot\binom{N}{k}\left(\frac{n}{N}\right)^k\left(1-\frac{n}{N}\right)^{N-k}\cdot\frac{1}{\binom{N}{n}}\cdot\frac{1}{\binom{n}{k}} \\
&= \binom{N}{k}\left(\frac{n}{N}\right)^k\left(1-\frac{n}{N}\right)^{N-k}\cdot\frac{(N-k)!}{(n-k)!(N-n)!}\cdot\frac{n!(N-n)!}{N!}\cdot\frac{k!(n-k)!}{n!} \\
&= \binom{N}{k}\left(\frac{n}{N}\right)^k\left(1-\frac{n}{N}\right)^{N-k}\cdot\frac{k!(N-k)!}{N!} \;=\; \binom{N}{k}\left(\frac{n}{N}\right)^k\left(1-\frac{n}{N}\right)^{N-k}\cdot\frac{1}{\binom{N}{k}} \\
&= \left(\frac{n}{N}\right)^k\cdot\left(1-\frac{n}{N}\right)^{N-k}
\end{aligned}$$

We remark that, for a given $s^{(1)}\in\mathcal{P}(U)$ with $|\,s^{(1)}\,| = k < n$, the quantity $\binom{N-k}{n-k}$ is the number of elements in $\mathcal{S}(U,n)$ containing $s^{(1)}$ as a proper subset.

(c) Let $\widetilde{P}:\mathcal{S}(U,k-n)$ be the induced sampling design on $\mathcal{S}(U,k-n)$. Then, for each $s^{(2)}\in\mathcal{S}(U,k-n)$, we have

$$\begin{aligned}
\widetilde{P}\Big(s^{(2)}\Big) &= \sum_{s^{(1)}\backslash\, s^{(0)}=s^{(2)}} P_{\mathrm{H}}\Big(s^{(0)},s^{(1)}\,\Big|\,K=k\Big) \;=\; \sum_{s^{(1)}\backslash\, s^{(0)}=s^{(2)}} \frac{1}{\binom{N}{k}}\cdot\frac{1}{\binom{k}{n}} \\
&= \binom{N-k+n}{n}\cdot\frac{1}{\binom{N}{k}}\cdot\frac{1}{\binom{k}{n}} \;=\; \frac{(N-k+n)!}{n!(N-k)!}\cdot\frac{k!(N-k)!}{N!}\cdot\frac{n!(k-n)!}{k!} \\
&= \frac{(k-n)!(N-k+n)!}{N!} \;=\; 1\Big/\binom{N}{k-n}
\end{aligned}$$

This proves that $\widetilde{P}$ is indeed equivalent to $\mathrm{SRSWOR}(U,k-n)$.

(d) Let $P':\mathcal{S}(U,n-k)$ be the induced sampling design on $\mathcal{S}(U,n-k)$. Then, for each $s^{(2)}\in\mathcal{S}(U,n-k)$, we have

$$\begin{aligned}
P'\Big(s^{(2)}\Big) &= \sum_{s^{(0)}\backslash\, s^{(1)}=s^{(2)}} P_{\mathrm{H}}\Big(s^{(0)},s^{(1)}\,\Big|\,K=k\Big) \;=\; \sum_{s^{(0)}\backslash\, s^{(1)}=s^{(2)}} \frac{1}{\binom{N}{n}}\cdot\frac{1}{\binom{n}{k}} \\
&= \binom{N-n+k}{k}\cdot\frac{1}{\binom{N}{n}}\cdot\frac{1}{\binom{n}{k}} \;=\; \frac{(N-n+k)!}{k!(N-n)!}\cdot\frac{n!(N-n)!}{N!}\cdot\frac{k!(n-k)!}{n!} \\
&= \frac{(n-k)!(N-n+k)!}{N!} \;=\; 1\Big/\binom{N}{n-k}
\end{aligned}$$

This proves that P' is indeed equivalent to $\mathrm{SRSWOR}(U,n-k)$.

The proof of this Lemma is complete. $\square$

Theorem 1.6 (The Hájek Fundamental Lemma)

Suppose U is a finite population of size $N\in\mathbb{N}$ with $N\geq 3$, and $y:U\longrightarrow\mathbb{R}$ is a population characteristic. Let $n\in\{2,\ldots,N\}$ be fixed. Let $\overline{y}_U := \frac{1}{N}\sum_{i\in U} y_i$. Let $\mathcal{S}(U,n)\times\mathcal{P}(U)$ be endowed with the probability function P_{H} defined *by the Hájek Sampling Design. Define the $\mathbb{R}^2$-valued random variable $Y = \left(Y^{(0)}, Y^{(1)}\right) : \mathcal{S}(U,n) \times \mathcal{P}(U) \longrightarrow \mathbb{R}^2$ as follows: For any $\left(s^{(0)}, s^{(1)}\right) \in \mathcal{S}(U,n) \times \mathcal{P}(U)$,*

$$Y^{(0)}\left(s^{(0)}\right) \;:=\; \frac{1}{n}\sum_{i\in s^{(0)}} \left(y_i - \overline{y}_U\right), \quad \text{and} \quad Y^{(1)}\left(s^{(1)}\right) \;:=\; \frac{1}{n}\sum_{i\in s^{(1)}} \left(y_i - \overline{y}_U\right).$$

Then,

$$E\left[\left(\frac{Y^{(0)}}{\sqrt{\mathrm{Var}\left[\,Y^{(1)}\,\right]}} - \frac{Y^{(1)}}{\sqrt{\mathrm{Var}\left[\,Y^{(1)}\,\right]}}\right)^2\right] \;=\; \frac{E\left[\left(Y^{(0)} - Y^{(1)}\right)^2\right]}{\mathrm{Var}\left[\,Y^{(1)}\,\right]} \;\leq\; \sqrt{\frac{1}{n} + \frac{1}{N-n}}$$

PROOF We write $k := \left|\,s^{(1)}\,\right|$. First, observed that

$$Y^{(0)} - Y^{(1)} \;=\; \begin{cases} 0, & \text{if } k = n \\ \dfrac{|\,k-n\,|}{n}\cdot\dfrac{1}{|\,k-n\,|}\cdot\displaystyle\sum_{i\in s^{(0)}\setminus s^{(1)}} \left(y_i - \overline{y}_U\right), & \text{if } k < n \\ \dfrac{|\,k-n\,|}{n}\cdot\dfrac{1}{|\,k-n\,|}\cdot\displaystyle\sum_{i\in s^{(1)}\setminus s^{(0)}} \left(y_i - \overline{y}_U\right), & \text{if } k > n \end{cases}$$

By Lemma 1.5(c,d), for $k := \left|\,s^{(1)}\,\right|$ fixed, we may regard $s^0 \setminus s^{(1)}$ and $s^1 \setminus s^{(0)}$ as realizations from SRSWOR$(U, |\,k-n\,|)$. Hence,

$$E\left[\,\left(Y^{(0)} - Y^{(1)}\right)\,\middle|\,\left|\,s^{(1)}\,\right| = k\,\right] \;=\; \frac{|\,k-n\,|}{n}\cdot E\left[\,\widehat{\overline{T}}^{\,\mathrm{HT}}_{\mathrm{SRSWOR}}\,\right] \;=\; 0$$

Hence,

$$\begin{aligned}
E\left[\,\left(Y^{(0)} - Y^{(1)}\right)^2\,\middle|\,\left|\,s^{(1)}\,\right| = k\,\right] \;&=\; \mathrm{Var}\left[\,Y^{(0)} - Y^{(1)}\,\middle|\,\left|\,s^{(1)}\,\right| = k\,\right] \\
&=\; \frac{|\,k-n\,|^2}{n^2}\left(1 - \frac{|\,k-n\,|}{N}\right)\frac{1}{|\,k-n\,|}\frac{\sum_{i\in U}\left(y_i - \overline{y}_U\right)^2}{N-1} \\
&=\; \frac{|\,k-n\,|}{n^2}\left(\frac{N - |\,k-n\,|}{N-1}\right)\frac{\sum_{i\in U}\left(y_i - \overline{y}_U\right)^2}{N} \\
&\leq\; \frac{|\,k-n\,|}{n^2}\frac{\sum_{i\in U}\left(y_i - \overline{y}_U\right)^2}{N}
\end{aligned}$$

Consequently,

$$\begin{aligned}
E\left\{\,\left(Y^{(0)} - Y^{(1)}\right)^2\,\right\} \;&=\; E\left\{\,E\left[\,\left(Y^{(0)} - Y^{(1)}\right)^2\,\middle|\,\left|\,s^{(1)}\,\right| = k\,\right]\,\right\} \\
&\leq\; E\left\{\,E\left[\,\frac{|\,k-n\,|}{n^2}\frac{\sum_{i\in U}\left(y_i - \overline{y}_U\right)^2}{N}\,\middle|\,\left|\,s^{(1)}\,\right| = k\,\right]\,\right\} \\
&=\; \frac{1}{n^2}\frac{\sum_{i\in U}\left(y_i - \overline{y}_U\right)^2}{N}\cdot E\{\,|\,k-n\,|\,\} \;\leq\; \frac{1}{n^2}\frac{\sum_{i\in U}\left(y_i - \overline{y}_U\right)^2}{N}\cdot\sqrt{E\left\{\,|\,k-n\,|^2\,\right\}} \\
&\leq\; \frac{1}{n^2}\frac{\sum_{i\in U}\left(y_i - \overline{y}_U\right)^2}{N}\cdot\sqrt{n\left(1 - \frac{n}{N}\right)},
\end{aligned}$$

where we used the Cauchy-Schwarz inequality (Theorem 9.3, [2]) for the second last inequality. Next, we compute $\mathrm{Var}\big[\,Y^{(1)}\,\big]$. To this end, note that

$$Y^{(1)} \;=\; \sum_{i\in U} Z_i,$$

where, for each $i \in U$,

$$Z_i : \mathcal{S}(U,n)\times\mathcal{P}(U) \longrightarrow \mathbb{R} : \left(s^{(0)},s^{(1)}\right) \longmapsto \begin{cases} \dfrac{1}{n}\,(y_i \,-\, \overline{y}_U)\,, & \text{if } i\in s^{(1)} \\[2ex] \quad 0, & \text{if } i\notin s^{(1)} \end{cases}$$

Note that, since Z_i depends only on $s^{(1)}$, which can be regarded as a Bernoulli sample from U, by Lemma 1.5, we see that the $Z_i, i\in U$, are independent, and

$$P\left(\,Z_i = \frac{1}{n}\,(y_i \,-\, \overline{y}_U)\,\right) \;=\; \frac{n}{N}, \quad\text{and}\quad P(\,Z_i = 0\,) \;=\; 1-\frac{n}{N}.$$

Thus,

$$\mathrm{Var}[\,Z_i\,] \;=\; \left(\frac{y_i-\overline{y}_U}{n}\right)^2\cdot\frac{n}{N}\left(1-\frac{n}{N}\right),$$

which in turn imples

$$\mathrm{Var}\Big[\,Y^{(1)}\,\Big] \;=\; \sum_{i\in U}\mathrm{Var}[\,Z_i\,] \;=\; \sum_{i\in U}\left(\frac{y_i-\overline{y}_U}{n}\right)^2\cdot\frac{n}{N}\left(1-\frac{n}{N}\right) \;=\;\cdots\;=\; \frac{1}{n^2}\frac{\sum_{i\in U}(y_i-\overline{y}_U)^2}{N}\cdot n\left(1-\frac{n}{N}\right)$$

Thus, we see that

$$\frac{E\Big[\,\big(Y^{(0)}-Y^{(1)}\big)^2\,\Big]}{\mathrm{Var}\big[\,Y^{(1)}\,\big]} \;\leq\; \frac{\dfrac{1}{n^2}\dfrac{\sum_{i\in U}(y_i-\overline{y}_U)^2}{N}\cdot\sqrt{n\left(1-\dfrac{n}{N}\right)}}{\dfrac{1}{n^2}\dfrac{\sum_{i\in U}(y_i-\overline{y}_U)^2}{N}\cdot n\left(1-\dfrac{n}{N}\right)} \;=\; \frac{1}{\sqrt{n\left(1-\dfrac{n}{N}\right)}} \;=\;\cdots\;=\; \sqrt{\frac{1}{n}+\frac{1}{N-n}}.$$

This completes the proof of Hájek's Fundamental Lemma. □

Remark 1.7 *The importance of Hájek's Fundamental Lemma is the following Corollary, which allows one to "replace" the sequence of SRSWOR samples in Theorem 1.1 with a sister sequence of Bernoulli samples (in the sense of Hájek's sampling design), in the following sense: prove that the sequence of estimators associated to the Bernoulli samples is asymptotically normal, and then "transfer" that asymptotic behaviour to the sequence of SRSWOR samples of interest, by appealing to Hájek's Fundamental Lemma.*

Corollary 1.8

Suppose we have the following:

- *Let* $\{\,U_\nu\,\}_{\nu\in\mathbb{N}}$ *be a sequence of finite populations, and* $N_\nu = |\,U_\nu\,| \geq 2$ *be the population size of* U_ν*. Let the elements of* U_ν *be indexed by* $1,2,3,\ldots,N_\nu$*.*
- *For each* $\nu\in\mathbb{N}$*, let* $y^{(\nu)} : U_\nu \longrightarrow \mathbb{R}$ *be a non-constant* $\mathbb{R}$*-valued population characteristic. For each* $i\in U_\nu$*, let* $y_i^{(\nu)}$ *denote* $y^{(\nu)}(i)$*, the value of* $y^{(\nu)}$ *evaluated at the* i^{th} *element of* U_ν*. Let* $\overline{y}_{U_\nu} := \dfrac{1}{N_\nu}\cdot\sum_{i\in U_\nu} y_i^{(\nu)}$*.*
- *For each* $\nu\in\mathbb{N}$*, let* $n_\nu \in \{1,2,3,\ldots,N_\nu-1\}$ *be given, and let* $p_\nu : \mathcal{S}(U_\nu,n_\nu)\times\mathcal{P}(U_\nu) \longrightarrow [0,1]$ *be the Hájek Sampling Design of size* n_ν *on* U_ν*.*

- *For each $\nu \in \mathbb{N}$, let $Y_\nu = \left(Y_\nu^{(0)}, Y_\nu^{(1)}\right) : \mathcal{S}(U_\nu, n_\nu) \times \mathcal{P}(U_\nu) \longrightarrow \mathbb{R}^2$ be the $\mathbb{R}^2$-valued random variable defined as follows: For any $\left(s_\nu^{(0)}, s_\nu^{(1)}\right) \in \mathcal{S}(U_\nu, n_\nu) \times \mathcal{P}(U_\nu)$,*

$$Y_\nu^{(0)}\left(s_\nu^{(0)}\right) \;:=\; \frac{1}{n_\nu}\sum_{i\in s_\nu^{(0)}}\left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right), \quad \text{and} \quad Y_\nu^{(1)}\left(s_\nu^{(1)}\right) \;:=\; \frac{1}{n_\nu}\sum_{i\in s_\nu^{(1)}}\left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right).$$

Then, the following implication holds:

$$\left.\begin{array}{ccc} n_\nu & \longrightarrow & \infty \\ N_\nu - n_\nu & \longrightarrow & \infty \\ \dfrac{Y_\nu^{(1)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}} & \overset{\mathcal{L}}{\longrightarrow} & N(0,1) \end{array}\right\} \quad\Longrightarrow\quad \frac{Y_\nu^{(0)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(0)}\,\right]}} \;\overset{\mathcal{L}}{\longrightarrow}\; N(0,1).$$

PROOF By the Hájek Fundamental Lemma (Theorem 1.6), we have for each $\nu \in \mathbb{N}$,

$$E\left[\left(\frac{Y_\nu^{(0)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}} - \frac{Y_\nu^{(1)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}}\right)^2\right] \;\leq\; \sqrt{\frac{1}{n_\nu} + \frac{1}{N_\nu - n_\nu}}.$$

Thus, the hypotheses $n_\nu \longrightarrow \infty$ and $N_\nu - n_\nu \longrightarrow \infty$ together imply that

$$\frac{Y_\nu^{(0)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}} - \frac{Y_\nu^{(1)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}}$$

converges to 0 in the second mean (Definition 3, p.4, [1]), hence also in probability (by Theorem 1(b), p.4, [1]). This convergence to 0 in probability and the hypothesis $Y_\nu^{(1)} \Big/ \sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]} \overset{\mathcal{L}}{\longrightarrow} N(0,1)$ then together imply, by Slutsky's Theorem (Theorem 6(b), p.39, [1]),

$$\frac{Y_\nu^{(0)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}} \;\overset{\mathcal{L}}{\longrightarrow}\; N(0,1).$$

Next, recall from the proof of the Hájek Fundamental Lemma (Theorem 1.6) that

$$\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right] \;=\; \cdots \;=\; \frac{1}{n_\nu}\left(1 - \frac{n_\nu}{N_\nu}\right)\frac{\sum_{i\in U_\nu}\left(y_i^{(\nu)} \;-\; \overline{y}_{U_\nu}\right)^2}{N_\nu}.$$

On the other hand,

$$\mathrm{Var}\left[\,Y_\nu^{(0)}\,\right] \;=\; \mathrm{Var}\left[\frac{1}{n_\nu}\sum_{i\in s_\nu^{(0)}}\left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)\right] \;=\; \mathrm{Var}\left[\frac{1}{n_\nu}\sum_{i\in s_\nu^{(0)}} y_i^{(\nu)}\right] \;=\; \frac{1}{n_\nu}\left(1 - \frac{n_\nu}{N_\nu}\right)\frac{\sum_{i\in U_\nu}\left(y_i^{(\nu)} \;-\; \overline{y}_{U_\nu}\right)^2}{N_\nu - 1}$$

Hence,

$$\frac{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(0)}\,\right]}} \;=\; \left(\frac{\dfrac{1}{n_\nu}\cdot\left(1-\dfrac{n_\nu}{N_\nu}\right)\cdot\sum_{i\in U_\nu}\left(y_i^{(\nu)}-\overline{y}_{U_\nu}\right)^2\Big/\,N_\nu}{\dfrac{1}{n_\nu}\cdot\left(1-\dfrac{n_\nu}{N_\nu}\right)\cdot\sum_{i\in U_\nu}\left(y_i^{(\nu)}-\overline{y}_{U_\nu}\right)^2\Big/\,(N_\nu-1)}\right)^{1/2} \;=\; \sqrt{\frac{N_\nu-1}{N_\nu}} \;\longrightarrow\; 1,\;\text{ as }\nu\longrightarrow\infty.$$

Note that $\left\{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}\Big/\sqrt{\mathrm{Var}\left[\,Y_\nu^{(0)}\,\right]}\right\}_{\nu\in\mathbb{N}}$ is a sequence of real numbers; we may regard it as a sequence of (constant) $\mathbb{R}$-valued random variables (defined on $\mathcal{S}(U_\nu)\times\mathcal{P}(U_\nu)$). Its convergence (as a sequence of real numbers) to 1, as we have established above, implies that it converges (as constant $\mathbb{R}$-valued random variables) almost surely to 1, hence also in probability as well as in distribution (see Theorem 1, p.4, [1]). By a corollary of Slutkey's Theorem (Corollary and Example 6, p.40, [1]), we therefore have

$$\frac{Y_\nu^{(0)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(0)}\,\right]}} \;=\; \frac{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(0)}\,\right]}}\cdot\frac{Y_\nu^{(0)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}} \;\overset{\mathcal{L}}{\longrightarrow}\; 1\cdot N(0,1) \;=\; N(0,1).$$

This completes the proof of the Corollary. □

PROOF OF **Theorem 1.1 (Hájek's Central Limit Theorem for SRSWOR).**

Let $\{\,U_\nu\,\}_{\nu\in\mathbb{N}}$, $\{\,n_\nu\,\}_{\nu\in\mathbb{N}}$, and $\left\{\,y^{(\nu)}:U_\nu\longrightarrow\mathbb{R}\,\right\}_{\nu\in\mathbb{N}}$ be the sequence of finite populations, SRSWOR sample sizes, and population characteristics as given in the statement of Theorem 1.1. We seek to prove that

$$\frac{\widehat{\overline{Y}}_\nu-\mu_\nu}{\sqrt{\sigma_\nu^2}} \;\overset{\mathcal{L}}{\longrightarrow}\; N(0,1),\quad\text{as }\nu\longrightarrow\infty.$$

To this end, for each $\nu\in\mathbb{N}$, first let $P_{\mathrm{H}}^{(\nu)}:\mathcal{S}(U_\nu,n_\nu)\times\mathcal{P}(U_\nu)\longrightarrow\mathbb{R}$ be the probability function induced on $\mathcal{S}(U_\nu,n_\nu)\times\mathcal{P}(U_\nu)$ by the Hájek Sampling Design of size n_ν on U_ν. Let $\left(Y_\nu^{(0)},Y_\nu^{(1)}\right):\mathcal{S}(U_\nu,n_\nu)\times\mathcal{P}(U_\nu)\longrightarrow\mathbb{R}^2$ be the $\mathbb{R}^2$-valued random variable defined on $\mathcal{S}(U_\nu,n_\nu)\times\mathcal{P}(U_\nu)$ as in the Hájek Fundamental Lemma (Theorem 1.6). Then, note that, for each $\nu\in\mathbb{N}$, the random variables $\left(\widehat{\overline{Y}}_\nu-\mu_\nu\right):\mathcal{S}(U_\nu,n_\nu)\longrightarrow\mathbb{R}$ and $Y_\nu^{(0)}:\mathcal{S}(U_\nu,n_\nu)\longrightarrow\mathbb{R}$ are in fact equal. By Corollary 1.8, it therefore suffices to show the following:

$$\frac{Y_\nu^{(1)}}{\sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}} \;\overset{\mathcal{L}}{\longrightarrow}\; N(0,1).$$

Now, recall that for each $\nu\in\mathbb{N}$, since $Y_\nu^{(1)}:\mathcal{P}(U_\nu)\longrightarrow\mathbb{R}$ is a random variable defined on the collection of Bernoulli samples of U_ν, each $Y_\nu^{(1)}$ can be regarded as a sum of independent random variables (but not identically distributed, unless the population characteristic $y^{(\nu)}:U_\nu\longrightarrow\mathbb{R}$ is constant). More precisely, recall from the proof of Hájek's Fundamental Lemma (Theorem 1.6) that, for each $\nu\in\mathbb{N}$,

$$Y_\nu^{(1)} \;=\; \sum_{i\in U_\nu} Z_i^{(\nu)},$$

where, for each $i\in U_\nu$,

$$Z_i^{(\nu)}:\mathcal{S}(U_\nu,n_\nu)\times\mathcal{P}(U_\nu)\longrightarrow\mathbb{R}:\left(s^{(0)},s^{(1)}\right)\longmapsto\begin{cases}\dfrac{1}{n_\nu}\left(y_i^{(\nu)}-\overline{y}_{U_\nu}\right), & \text{if } i\in s^{(1)}\\[2ex] 0, & \text{if } i\notin s^{(1)}\end{cases}$$

We emphasize again that for each $\nu \in \mathbb{N}$, the random variables $Z_1^{(\nu)}, \ldots, Z_{N_\nu}^{(\nu)}$ are independent (though not necessarily identically distributed). Therefore, the classical Lindeberg Central Limit Theorem (see Chapter 5, p.27, [1]) provides a sufficient condition for the asymptotic normality of $Y_\nu^{(1)} \Big/ \sqrt{\mathrm{Var}\left[\,Y_\nu^{(1)}\,\right]}$. Thus, it remains only to show that the hypotheses in Theorem 1.1 implies the validity of Lindeberg's condition (as in Lindeberg's Central Limit Theorem) on the $Y_\nu^{(1)}$'s. Now, Lindeberg's condition on $Y_\nu^{(1)}$ can be expressed as:

$$\lim_{\nu\to\infty} \frac{1}{\tau_\nu^2} \sum_{i=1}^{N_\nu} \int_{\{|x|\,\geq\,\varepsilon\tau_\nu\}} |\,x\,|^2 \,\mathrm{d}\,P^{Z_i^{(\nu)}-E\left(Z_i^{(\nu)}\right)}(x) \;\;=\;\; 0, \quad \text{for each } \varepsilon > 0,$$

where (recalling from the proof of the Hájek Fundamental Lemma (Theorem 1.6))

$$\begin{aligned}
\tau_\nu^2 \;\;:=\;\; \mathrm{Var}\left[\,Y_\nu^{(1)}\,\right] \;\;&=\;\; \frac{1}{n_\nu}\left(1-\frac{n_\nu}{N_\nu}\right)\frac{\sum_{i\in U_\nu}\left(y_i^{(\nu)} \,-\, \overline{y}_{U_\nu}\right)^2}{N_\nu} \;\;=\;\; \frac{1}{n_\nu}\left(1-\frac{n_\nu}{N_\nu}\right)\frac{\sum_{i\in U_\nu}\left(y_i^{(\nu)} \,-\, \overline{y}_{U_\nu}\right)^2}{N_\nu-1}\cdot\frac{N_\nu-1}{N_\nu} \\
&=\;\; \sigma_\nu^2\cdot\frac{N_\nu-1}{N_\nu},
\end{aligned}$$

and where

$$\sigma_\nu^2 \;\;:=\;\; \mathrm{Var}\left[\,Y_\nu^{(0)}\,\right] \;\;=\;\; \frac{1}{n_\nu}\left(1-\frac{n_\nu}{N_\nu}\right)\frac{\sum_{i\in U_\nu}\left(y_i^{(\nu)} \,-\, \overline{y}_{U_\nu}\right)^2}{N_\nu-1}.$$

Each random variable $Z_i^{(\nu)} - E\left(Z_i^{(\nu)}\right)$ has a Bernoulli distribution, with

$$\left\{\begin{array}{l}
P\left[\,Z_i^{(\nu)} - E\left(Z_i^{(\nu)}\right) \;=\; (y_i^{(\nu)} - \overline{y}_{U_\nu})\cdot\left(\dfrac{1}{n_\nu}-\dfrac{1}{N_\nu}\right)\,\right] \;=\; \dfrac{n_\nu}{N_\nu} \\[2ex]
P\left[\,Z_i^{(\nu)} - E\left(Z_i^{(\nu)}\right) \;=\; -\,(y_i^{(\nu)} - \overline{y}_{U_\nu})\cdot\dfrac{1}{N_\nu}\,\right] \;=\; 1-\dfrac{n_\nu}{N_\nu}
\end{array}\right.$$

Hence,

$$\begin{aligned}
\int_{\{|x|\,\geq\,\varepsilon\tau_\nu\}} |\,x\,|^2 \,\mathrm{d}\,P^{Z_i^{(\nu)}-E\left(Z_i^{(\nu)}\right)}(x) \;\;=\;\; &\left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2\left(\frac{1}{n_\nu}-\frac{1}{N_\nu}\right)^2\cdot I\left\{\left|\,y_i^{(\nu)}-\overline{y}_{U_\nu}\right|\cdot\left(\frac{1}{n_\nu}-\frac{1}{N_\nu}\right) \,\geq\, \varepsilon\,\tau_\nu\right\}\cdot\frac{n_\nu}{N_\nu} \\
&+\;\; \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2\left(\frac{1}{N_\nu}\right)^2\cdot I\left\{\left|\,y_i^{(\nu)}-\overline{y}_{U_\nu}\right|\cdot\frac{1}{N_\nu} \,\geq\, \varepsilon\,\tau_\nu\right\}\cdot\left(1-\frac{n_\nu}{N_\nu}\right),
\end{aligned}$$

which implies

$$\begin{aligned}
&\sum_{i=1}^{N_\nu} \int_{\{|x|\,\geq\,\varepsilon\tau_\nu\}} |\,x\,|^2 \,\mathrm{d}\,P^{Z_i^{(\nu)}-E\left(Z_i^{(\nu)}\right)}(x) \\
&=\;\; \frac{n_\nu}{N_\nu}\left(\frac{1}{n_\nu}-\frac{1}{N_\nu}\right)^2\cdot\sum_{i\in C_\nu(\varepsilon)}\left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2 \;\;+\;\; \left(1-\frac{n_\nu}{N_\nu}\right)\left(\frac{1}{N_\nu}\right)^2\cdot\sum_{i\in B_\nu(\varepsilon)}\left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2,
\end{aligned}$$

where

$$C_\nu(\varepsilon) \;:=\; \left\{\, i\in U_\nu \;:\; \left|\,y_i^{(\nu)}-\overline{y}_{U_\nu}\right|\cdot\left(\frac{1}{n_\nu}-\frac{1}{N_\nu}\right) \,\geq\, \varepsilon\,\tau_\nu\right\} \quad\text{and}\quad B_\nu(\varepsilon) \;:=\; \left\{\, i\in U_\nu \;:\; \left|\,y_i^{(\nu)}-\overline{y}_{U_\nu}\right|\cdot\frac{1}{N_\nu} \,\geq\, \varepsilon\,\tau_\nu\right\}$$

We now make the following:

Claim: $C_\nu(\varepsilon) \;\subseteq\; U_\nu(\varepsilon)$, and $B_\nu(\varepsilon) \;\subseteq\; U_\nu(\varepsilon)$.

Proof of Claim: Note that

$$\begin{aligned}
\left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \cdot \left(\frac{1}{n_\nu} - \frac{1}{N_\nu} \right) \;\geq\; \varepsilon\,\tau_\nu
&\iff \left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \cdot \left(\frac{N_\nu - n_\nu}{n_\nu N_\nu} \right) \;\geq\; \varepsilon\,\sigma_\nu \cdot \sqrt{\frac{N_\nu - 1}{N_\nu}} \\
&\iff \left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \;\geq\; \sigma_\nu \cdot \varepsilon \cdot \frac{n_\nu\, N_\nu}{N_\nu - n_\nu} \cdot \sqrt{\frac{N_\nu - 1}{N_\nu}} \\
&\iff \left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \;\geq\; \sigma_\nu \cdot \varepsilon \cdot n_\nu \cdot \sqrt{\frac{N_\nu}{N_\nu - n_\nu} \cdot \frac{N_\nu - 1}{N_\nu - n_\nu}} \\
&\Longrightarrow \left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \;\geq\; \sigma_\nu \cdot \varepsilon
\end{aligned}$$

This proves the first inclusion. Secondly, note that

$$\begin{aligned}
\left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \cdot \frac{1}{N_\nu} \;\geq\; \varepsilon\,\tau_\nu
&\iff \left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \;\geq\; N_\nu \cdot \varepsilon \cdot \sigma_\nu \cdot \sqrt{\frac{N_\nu - 1}{N_\nu}} \\
&\iff \left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \;\geq\; \sigma_\nu \cdot \varepsilon \cdot \sqrt{N_\nu(N_\nu - 1)} \\
&\Longrightarrow \left| y_i^{(\nu)} - \overline{y}_{U_\nu} \right| \;\geq\; \sigma_\nu \cdot \varepsilon
\end{aligned}$$

This proves the second inclusion, and concludes the proof of the Claim.

The Claim implies:

$$\begin{aligned}
& \sum_{i=1}^{N_\nu} \int_{\{|x|\,\geq\,\varepsilon\tau_\nu\}} |\,x\,|^2 \,\mathrm{d}\,P^{Z_i^{(\nu)} - E\left(Z_i^{(\nu)}\right)}(x) \\
=\;& \frac{n_\nu}{N_\nu}\left(\frac{1}{n_\nu} - \frac{1}{N_\nu}\right)^2 \cdot \sum_{i\in C_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2 \;+\; \left(1 - \frac{n_\nu}{N_\nu}\right)\left(\frac{1}{N_\nu}\right)^2 \cdot \sum_{i\in B_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2 \\
\leq\;& \frac{n_\nu}{N_\nu}\left(\frac{1}{n_\nu} - \frac{1}{N_\nu}\right)^2 \cdot \sum_{i\in U_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2 \;+\; \left(1 - \frac{n_\nu}{N_\nu}\right)\left(\frac{1}{N_\nu}\right)^2 \cdot \sum_{i\in U_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2 \\
=\;& \frac{1}{N_\nu}\left(1 - \frac{n_\nu}{N_\nu}\right) \cdot \left[\left(\frac{1}{n_\nu} - \frac{1}{N_\nu}\right) + \frac{1}{N_\nu} \right] \cdot \sum_{i\in U_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2 \\
=\;& \frac{1}{n_\nu\, N_\nu}\left(1 - \frac{n_\nu}{N_\nu}\right) \cdot \sum_{i\in U_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2,
\end{aligned}$$

which in turn implies

$$\begin{aligned}
\frac{1}{\tau_\nu^2} \sum_{i=1}^{N_\nu} \int_{\{|x|\,\geq\,\varepsilon\tau_\nu\}} |\,x\,|^2 \,\mathrm{d}\,P^{Z_i^{(\nu)} - E\left(Z_i^{(\nu)}\right)}(x)
\;&\leq\; \frac{\dfrac{1}{n_\nu\, N_\nu}\left(1 - \dfrac{n_\nu}{N_\nu}\right) \cdot \displaystyle\sum_{i\in U_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2}{\dfrac{1}{n_\nu\, N_\nu}\left(1 - \dfrac{n_\nu}{N_\nu}\right) \cdot \displaystyle\sum_{i\in U_\nu} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2} \\
&=\; \frac{\displaystyle\sum_{i\in U_\nu(\varepsilon)} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2}{\displaystyle\sum_{i\in U_\nu} \left(y_i^{(\nu)} - \overline{y}_{U_\nu}\right)^2} \;\longrightarrow 0, \quad \text{for each } \varepsilon > 0,
\end{aligned}$$

where the last convergence follows by hypothesis of Theorem 1.1. This shows that $Y_\nu^{(1)}$ indeed satisfy Lindeberg's condition, and completes the proof of Theorem 1.1. $\square$

References

[1] Ferguson, T. S. *A Course in Large Sample Theory*, first ed. Texts in Statistical Science. CRC Press, 1996.

[2] Jacod, J., and Protter, P. *Probability Essentials*. Springer-Verlag, New York, 2004. Universitext.