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 Hàjek Central Limit Theorem for SRSWOR)
Suppose we have the following:

- *Let* $\{\,U_\nu\,\}_{\nu\in\mathbb{N}}$ *be a sequence of finite populations, and* $N_\nu = |\,U_\nu\,|$ *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 an* $\mathbb{R}$*-valued population characteristic. For each* $i\in U_\nu$*, let* $y^{(\nu)}_i$ *denote* $y^{(\nu)}(i)$*, the value of* $y^{(\nu)}$ *evaluated at the* $i^{\textrm{th}}$ *element of* U_ν*.*
- *For each* $\nu\in\mathbb{N}$*, let* $n_\nu\in\{1,2,3,\ldots,N_\nu\}$ *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*

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

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

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

Let

$$\mu_\nu \;:=\; E\!\left[\,\widehat{\overline{Y}}_\nu\,\right] \;=\; \frac{1}{N_\nu}\sum_{i\in U_\nu} y^{(\nu)}_i \quad\textrm{and}\quad \sigma^2_\nu \;:=\; \textrm{Var}\!\left[\,\widehat{\overline{Y}}_\nu\,\right] \;=\; \left(1-\frac{n_\nu}{N_\nu}\right)\frac{S^2_\nu}{n_\nu}\,,$$

where

$$S^2_\nu \;:=\; \frac{1}{N_\nu-1}\sum_{i\in U_\nu}\left(y^{(\nu)}_i-\mu_\nu\right)^2$$

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

$$U_\nu(\delta) \;:=\; \left\{\, i\in U_\nu \;\middle\vert\; |y^{(\nu)}_i-\mu_\nu| > \delta\,\sqrt{\sigma^2_\nu}\,\right\} \;\subset\; U_\nu.$$

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

$$\lim_{\nu\to\infty} P\left\{\, s\in\mathcal{S}_\nu \;\middle\vert\; \frac{\widehat{\overline{Y}}_\nu(s)-\mu_\nu}{\sqrt{\sigma^2_\nu}} < x \,\right\} \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,\textrm{d}t$$

if and only if

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

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,

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

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

Definition 1.3 (The Hàjek 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 $\mathcal{P}(U)$ be the power set of U. Let $\mathcal{S}(U, n)$ be the collection of all subsets of U with exactly n elements. The Hàjek Sampling Design, 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,

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

  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 Hàjek 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

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

Lemma 1.5 (Properties of the Hàjek 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 Hàjek 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:

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

To this end,

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

We remark that, for a given $s^{(0)}\in\mathcal{S}(U,n)$ and $k>n$, the quantity $\left(\begin{array}{c} N-n \\ k-n \end{array}\right)$ 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 $k>n$,

$$\frac{1}{\left(\begin{array}{c} N \\ k \end{array}\right)}\cdot\frac{1}{\left(\begin{array}{c} k \\ n \end{array}\right)}\cdot\left(\begin{array}{c} N-n \\ k-n \end{array}\right) \;=\; \frac{k!(N-k)!}{N!}\cdot\frac{n!(k-n)!}{k!}\cdot\frac{(N-n)!}{(k-n)!(N-k)!} \;=\; \frac{n!(N-n)!}{N!} \;=\; \frac{1}{\left(\begin{array}{c} N \\ n \end{array}\right)}.$$

Hence, we have

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

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

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

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

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

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 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 SRSWOR$(U,n-k)$.

The proof of this Lemma is complete. □

Theorem 1.6 (Hàjek's 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 Hàjek Sampling Design. Define the* $\mathbb{R}^2$*-valued random variable* $X = \left(X^{(0)}, X^{(1)}\right) : \mathcal{S}(U,n) \times \mathcal{P}(U) \longrightarrow \mathbb{R}^2$ *as follows: For any* $\left(s^{(0)}, s^{(1)}\right) \in \mathcal{S}(U,n) \times \mathcal{P}(U)$*,*

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

Then,

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

Proof First, observed that

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

We argue conditionally on $k := \left|\,s^{(1)}\,\right|$. 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$(U, \left|\,k-n\,\right|)$. Hence,

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

□

References