Allocation of sample sizes
to (n_i's) → two diff. strata
There are generally three types of allocation:

(i) Proportional Allocation →
P.A is given by Bowley.

(ii) Neyman Allocation = Neyman
(iii) Optimum Allocation

* (i) Proportional Allocation →
Allocation of n_i's to various strata is called proport^n; if
$n_i \propto N_i$; $i = 1, 2 \ldots k$
$n_i = \ell N_i$
where ℓ is the constant of

$n_i = \frac{10}{100} \times \_\_ = ?$

proportionality (for the value of l)

$$\Rightarrow \sum_{i=1}^{k} n_i = l \sum_{i=1}^{k} N_i$$

$n = lN$

$$l = \frac{n}{N}$$

$$\therefore n_i = \frac{n}{N} N_i = n p_i \quad \because p_i = \frac{N_i}{N}$$

where $p_i = \frac{N_i}{N}$, $i = 1, 2, 3 \ldots k)$

$$\therefore \boxed{n_{i,p} = n p_i}$$

In Proportional Allocat^n the variance is given by

$$V(\bar{y}_{st}) = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i} \right) p_i^2 S_i^2$$

$$V(\bar{y}_{st})_p = \sum_{i=1}^{k} \left(\frac{1}{n_{i,p}} - \frac{1}{N_i} \right) p_i^2 S_i^2$$

$$= \sum_{i=1}^{k} \left(\frac{1}{n p_i} - \frac{1}{N_i} \right) p_i^2 S_i^2$$

$$= \sum_{i=1}^{k} \left(\frac{p_i}{n p_i} - \frac{p_i}{N_i} \right) p_i S_i^2$$

$$= \sum_{i=1}^{k} \left(\frac{1}{n} - \frac{1}{N} \right) p_i S_i^2$$

$$\boxed{V(\bar{y}_{st})_p = \left(\frac{1}{n} - \frac{1}{N} \right) \sum_{i=1}^{k} p_i S_i^2}$$

(2) Neyman Allocation :→

$n_i \propto N_i S_i$

$\Rightarrow n_i = l N_i S_i \quad i = 1, 2, \ldots k$

where l = proportionality const.

or $\sum_{i=1}^{k} n_i = l \sum_{i=1}^{k} N_i S_i$

$$l = \frac{n}{\sum_{i=1}^{k} N_i S_i}$$

or $n_{i,N} = \frac{n}{\sum_{i=1}^{k} N_i S_i} N_i S_i$

$\frac{N_i}{N} = p_i$

$$\boxed{n_{i,N} = \frac{n}{\sum p_i S_i} p_i S_i}$$

Under Neyman allocation the var. is given by

$$Var(\bar{y}_{st}) = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i} \right) p_i^2 S_i^2$$

$$Var(\bar{y}_{st})_N = \sum_{i=1}^{k} \left(\frac{1}{n_{i,N}} - \frac{1}{N_i} \right) p_i^2 S_i^2$$

$$= \sum_{i=1}^{k} \left[\frac{1}{\frac{n p_i S_i}{\sum_{i=1}^{k} p_i S_i}} - \frac{1}{N_i} \right] p_i^2 S_i^2$$

$$= \sum_{i=1}^{k}\left[\frac{\sum_{i=1}^{k} p_i s_i}{n p_i s_i} - \frac{1}{N_i}\right] p_i^2 s_i^2$$

$$= \frac{\sum_{i=1}^{k} p_i s_i \sum_{i=1}^{k} p_i s_i}{n} - \frac{1}{N}\sum_{i=1}^{k} p_i s_i^2$$

$$= \frac{1}{n}\left(\sum_{i=1}^{k} p_i s_i\right)^2 - \frac{1}{N}\sum_{i=1}^{k} p_i s_i^2$$

cost देखते होते हैं n_i find करते हैं & var. min होता है। C_0 = given cost

(3) Optimum Allocation:→

In this principle of allocation allocation of sample size n_i's; $i=1,2,..k$ to different strata is done in such a way so as to maximize the precision for a given cost or to minimize cost for a given degree of precision [Precision $\propto \frac{1}{\text{Var.}}$]

खर्च

Case-(i) When cost of the (होना) survey is given. incurred

Let c_i be the cost per unit of the sample from the i^{th} strata.

Therefore total cost of the sample from all strata will be $\sum_{i=1}^{k} c_i n_i$

Let a be the overhead cost, therefore total cost of

Survey will be

$$C = a + \sum_{i=1}^{k} c_i n_i \le C_0 \text{ (given)} \quad \text{—(1)}$$

where C_0 is the given cost

Our aim is to max precision or to min. var. subject to

restruct^n (कम करना) restriction of the cost C_0

We have

$$V(\bar{y}_{st}) = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i} \right) p_i^2 s_i^2$$

Consider a funct^n ϕ

V = Var.

$$\phi = V + \lambda (C - C_0)$$

λ is the langrages constant

C = cost

C_0 = given cost

$$\phi = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i} \right) p_i^2 s_i^2 + \lambda \left(a + \sum_{i=1}^{k} c_i n_i - C_0 \right) \quad \text{—(2)}$$

Differentiating (2) w.r.t n_i and equating to zero, we get

$$\frac{\partial \phi}{\partial n_i} = -\frac{1}{n_i^2} p_i^2 s_i^2 + \lambda c_i = 0$$

$$\Rightarrow \frac{p_i^2 s_i^2}{n_i^2} = \lambda c_i$$

$$n_i = \sqrt{\frac{p_i^2 s_i^2}{\lambda c_i}} = \frac{p_i s_i}{\sqrt{\lambda c_i}} \quad \text{—(3)}$$

$$i = 1, 2, \dots k$$

Diffⁿ (2) w.r.t λ and equating to zero, we get;

$$\frac{\partial \phi}{\partial \lambda} = a + \sum_{i=1}^{k} c_i n_i - c_0 = 0$$

$$\Rightarrow c_0 = a + \sum_{i=1}^{k} c_i n_i \quad \text{——} \quad (4)$$

Sub. the value of n_i from eqⁿ (3) in eqⁿ (4) we get;

$$c_0 = a + \sum_{i=1}^{k} c_i \frac{p_i s_i}{\sqrt{\lambda c_i}}$$

$$c_0 = a + \frac{1}{\sqrt{\lambda}} \sum_{i=1}^{k} p_i s_i \sqrt{c_i}$$

$$\Rightarrow \sqrt{\lambda} = \frac{\sum_{i=1}^{k} p_i s_i \sqrt{c_i}}{c_0 - a} \quad \text{——} \quad (5)$$

Sub. the value of λ in eq (3) we get

$$n_i = \frac{p_i s_i}{\sqrt{c_i} \; \frac{\sum_{i=1}^{k} p_i s_i \sqrt{c_i}}{c_0 - a}}$$

$$\boxed{n_{i,opt.} = \frac{p_i s_i}{\sqrt{c_i}} \cdot \frac{c_0 - a}{\sum_{i=1}^{k} p_i s_i \sqrt{c_i}}} \; ; \; i = 1, 2 \ldots k$$

or

$$n_{i,opt.} = \frac{N_i s_i}{\sqrt{c_i}} \cdot \frac{c_0 - a}{\sum_{i=1}^{k} N_i s_i \sqrt{c_i}}$$

Therefore the optimum var under

Optimum Allocatⁿ will be.

$$Var.(\bar{y}_{st})opt. = \sum_{i=1}^{k}\left(\frac{1}{n_{i\,opt}} - \frac{1}{N_i}\right) p_i^2 S_i^2$$

$$= \sum_{i=1}^{k}\left(\frac{1}{\frac{p_i S_i}{\sqrt{c_i}} \cdot \frac{C_0 - a}{\sum_{i=1}^{k} p_i S_i \sqrt{c_i}}} - \frac{1}{N_i}\right) p_i^2 S_i^2$$

$\because p_i = \frac{N_i}{N}$

$$= \sum_{i=1}^{k} p_i S_i \sqrt{c_i} \cdot \frac{\sum_{i=1}^{k} p_i S_i \sqrt{c_i}}{C_0 - a} - \frac{1}{N}\sum_{i=1}^{k} p_i S_i^2$$

$$\therefore Var.(\bar{y}_{st})opt. = \frac{\left(\sum_{i=1}^{k} p_i S_i \sqrt{c_i}\right)^2}{C_0 - a} - \frac{1}{N}\sum_{i=1}^{k} p_i S_i^2$$

Case: (II) When var. (precision) is given we have to determine the sample size from every stratum so as to min. the cost of the survey

Consider a funcⁿ ϕ as

$$\phi = C + \lambda(V - V_0) \quad \text{--- (1)}$$

V_0 is the given Var. and stands for variance

$$\phi = a + \sum_{i=1}^{k} c_i n_i + \lambda\left[\sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_i^2 - V_0\right] \text{--- (2)}$$

Diff. (2) w.r.t. n_i and equating to zero we get;

$$\frac{\partial \phi}{\partial n_i} = c_i + \lambda\left(-\frac{1}{n_i^2}\right) p_i^2 s_i^2 = 0$$

$$\Rightarrow \frac{\lambda p_i^2 s_i^2}{n_i^2} = c_i \Rightarrow n_i = \sqrt{\frac{\lambda p_i^2 s_i^2}{c_i}}$$

$$\boxed{n_i = \sqrt{\frac{\lambda}{c_i}}\, p_i s_i} \quad \text{—— (3)}$$

Diff. (2) w.r.t. λ and equating to zero we get;

$$\frac{\partial \phi}{\partial \lambda} = \sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 s_i^2 - V_0 = 0$$

$$V_0 = \sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 s_i^2 \quad \text{—— (4)}$$

Sub. the value of n_i from (3) in eq.^{n} (4) we get;

$$V_0 = \sum_{i=1}^{k}\left(\frac{1}{\sqrt{\frac{\lambda}{c_i}}\, p_i s_i} - \frac{1}{N_i}\right) p_i^2 s_i^2$$

$$V_0 = \sum_{i=1}^{k}\left(\frac{\sqrt{c_i}}{\sqrt{\lambda}\, p_i s_i} - \frac{1}{N_i}\right) p_i^2 s_i^2$$

$$V_0 = \frac{1}{\sqrt{\lambda}} \sum_{i=1}^{k} p_i s_i \sqrt{c_i} - \frac{1}{N} \sum_{i=1}^{k} p_i s_i^2$$

$$\Rightarrow \sqrt{\lambda} = \frac{\sum_{i=1}^{k} p_i s_i \sqrt{c_i}}{V_0 + \frac{1}{N} \sum_{i=1}^{k} p_i s_i^2} \quad \text{—— (5)}$$

Sub. the value of $\sqrt{\lambda}$ in eq. ③ we get;

$$n_{i,opt} = \frac{p_i S_i}{\sqrt{c_i}} \left[\frac{\sum_{i=1}^{k} p_i S_i \sqrt{c_i}}{V_0 + \frac{1}{N} \sum_{i=1}^{k} p_i S_i^2} \right]$$

Now, the min cost will be

$$C = a + \sum_{i=1}^{k} c_i n_{i,opt}$$

$$= a + \sum_{i=1}^{k} c_i \frac{\sum_{i=1}^{k} p_i S_i \sqrt{c_i}}{V_0 + \frac{1}{N} \sum_{i=1}^{k} p_i S_i^2} \times \frac{p_i S_i}{\sqrt{c_i}}$$

$$C = a + \frac{\sum_{i=1}^{k} p_i S_i \sqrt{c_i} \sum_{i=1}^{k} p_i S_i \sqrt{c_i}}{V_0 + \frac{1}{N} \sum_{i=1}^{k} p_i S_i^2}$$

$$\boxed{C = a + \frac{\left(\sum_{i=1}^{k} p_i S_i \sqrt{c_i}\right)^2}{V_0 + \frac{1}{N} \sum_{i=1}^{k} p_i S_i^2}}$$

$$\sum_{i=1}^{k} p_i = \sum_{i=1}^{k} \frac{N_i}{N} = \frac{N}{N} = 1$$

10/010

$$C_0 = a + \sum_{i=1}^{k} c_i n_i$$

Remarks. ① We have $n_{i,opt} = \frac{p_i S_i}{\sqrt{c_i}}$

$$= \frac{p_i S_i}{\sqrt{c_i}} \cdot \frac{C_0 - a}{\sum_{i=1}^{k} p_i S_i \sqrt{c_i}} \quad ; i = 1, 2 \ldots k$$

→ when the cost is given

If the cost inf.ⁿ is not given we can assume

$$c_1 = c_2 = \ldots c_k = c'$$

Now

$$C_0 - a = \sum_{i=1}^{k} c_i n_i = c' \sum_{i=1}^{k} n_i = n c'$$

$$\therefore n_{i,opt} = \frac{p_i S_i}{\sqrt{c'}} \cdot \frac{n c'}{\sum_{i=1}^{k} p_i S_i \sqrt{c'}} = n \frac{p_i S_i}{\sum_{i=1}^{k} p_i S_i}$$

$$= n_{i,N}$$

If the cost is equal and the cost ~~inf.ⁿ~~ inf.ⁿ isn't given then the optimum Allocatⁿ tends to Neyman Allocatⁿ

→ When infⁿ abt strata variability is not given, then we can assume

$$S_1 = S_2 = \ldots = S_k = S$$

$$n_{i,N} = n \frac{p_i S_i}{\sum_{i=1}^{k} p_i S_i} = \frac{n p_i S}{\sum_{i=1}^{k} p_i S}$$

as $\sum_{i=1}^{k} p_i = 1$

$$= n p_i = n_{i,p}$$

$\overline{Y}_{Ni}$ = stratum mean

therefore them strata variabilities are not given then Neyman Allocation tends to proportional Allocation

Comparison of efficiency of different Allocation of stratified sampling with SRS.

08/10/00

(1) We have

$$V(\overline{y}_n) = \left(\frac{1}{n} - \frac{1}{N}\right) S^2 \text{ in SRSWOR} \quad \text{---} (1)$$

where $S^2 = \frac{1}{N-1} \sum_{i=1}^{k} \sum_{j=1}^{N_i} (Y_{ij} - \overline{Y}_N)^2$ is

the pop^n mean square

$$V(\overline{y}_{st}) = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_i^2 \quad \text{---} (2)$$

$$V(\overline{y}_{st})_{prop} = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{i=1}^{k} p_i S_i^2 \quad \text{---} (3)$$

We have

$$S^2 = \frac{1}{N-1} \sum_{i=1}^{k} \sum_{j=1}^{N_i} (Y_{ij} - \overline{Y}_N)^2$$

$$= \frac{1}{N-1} \sum_{i=1}^{k} \sum_{j=1}^{N_i} (Y_{ij} - \overline{Y}_{Ni} + \overline{Y}_{Ni} - \overline{Y}_N)^2$$

$$= \frac{1}{N-1} \Big[\sum_i \sum_j (Y_{ij} - \overline{Y}_{Ni})^2 + \sum\sum (\overline{Y}_{Ni} - \overline{Y}_N$$

$$+ 2 \sum_i \sum_j (Y_{ij} - \overline{Y}_{Ni})(\overline{Y}_{Ni} - \overline{Y}_N)$$

$N =$ pop^n size
$N_i =$ strata size

The product term is zero. Hence the algebric sum of the deviat^n from mean is always zero.

[as $\sum_j (y_{ij} - \bar{Y}_{N_i}) = 0$] $\quad S_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (y_{ij} - \bar{Y}_{N_i})^2$

$$S^2 = \frac{1}{N-1} \left[\sum_i (N_i - 1) S_i^2 + \sum_i N_i (\bar{Y}_{N_i} - \bar{Y}_N)^2 \right]$$

$$\boxed{\because \sum_j^{N_i} 1 = N_i}$$

If strata sizes are large then $N_i - 1 \approx N_i \ \forall i$ and $N - 1 \approx N$

$$\therefore S^2 = \frac{1}{N} \left[\sum_i N_i S_i^2 + \sum_i N_i (\bar{Y}_{N_i} - \bar{Y}_N)^2 \right]$$

$$= \left[\sum_i k_i S_i^2 + \sum_i k_i (\bar{Y}_{N_i} - \bar{Y}_N)^2 \right] \quad \text{--- (4)}$$

$$V(\bar{y}_n)_{SRSWOR} = \left(\frac{1}{n} - \frac{1}{N} \right) \left[\sum_i k_i S_i^2 + \sum_i k_i (\bar{Y}_{N_i} - \bar{Y}_N)^2 \right] \quad \text{--- (5)}$$

Now $V(\bar{y}_n)_{SRSWOR} - V(\bar{y}_{st})_{prop}$

$$= \left(\frac{1}{n} - \frac{1}{N} \right) \sum_{i=1}^{k} k_i S_i^2 + \left(\frac{1}{n} - \frac{1}{N} \right) \sum_{i=1}^{k} k_i (\bar{Y}_{N_i} - \bar{Y}_N)^2$$

$$- \left(\frac{1}{n} - \frac{1}{N} \right) \sum_{i=1}^{k} k_i S_i^2$$

$$= \left(\frac{1}{n} - \frac{1}{N} \right) \sum_{i=1}^{k} N_i (\bar{Y}_{N_i} - \bar{Y}_N)^2 \geq 0$$ (a non -ve quantity)

Hence $V(\bar{y}_n)_{SRSWOR} \geq V(\bar{y}_{st})_{prop}$ --- (A)

Hence stratified sampling under proportional allocatⁿ is more efficient than SRS.

② We have

$$V(\bar{y}_{st})_p = \left(\frac{1}{n}-\frac{1}{N}\right)\sum_{i=1}^{k} p_i s_i^2 \quad \text{—— (}$$

$$V(\bar{y}_{st})_N = \frac{1}{n}\left(\sum_{i=1}^{k} p_i s_i\right)^2 - \frac{1}{N}\sum_{i=1}^{k} p_i s_i^2$$

Now

$$V(\bar{y}_{st})_p - V(\bar{y}_{st})_N = \left(\frac{1}{n}-\frac{1}{N}\right)\sum_{i=1}^{k} p_i s_i^2 - \frac{1}{n}\left(\sum_{i=1}^{k} p_i s_i\right)^2 + \frac{1}{N}\sum_{i=1}^{k} p_i s_i^2$$

$$= \frac{1}{n}\sum_{i=1}^{k} p_i s_i^2 - \frac{1}{N}\sum_{i=1}^{k} p_i s_i^2$$

$\frac{1}{N}\sum S_i^2$

$$- \frac{1}{n}\left(\sum_{i=1}^{k} p_i s_i\right)^2 + \frac{1}{N}\sum_{i=1}^{k} p_i s_i^2$$

N_i

9 10

$$= \frac{1}{n}\left[\sum_{i=1}^{k} p_i s_i^2 - \left(\sum_{i=1}^{k} p_i s_i\right)^2\right]$$

$= \frac{1}{N}\sum N_i$

$\frac{1}{N}\sum N_i S_i = \bar{S}_w$

$$= \frac{1}{n}\left[\sum_{i=1}^{k} p_i s_i^2 - \bar{S}_w^2\right]$$

[where $\bar{S}_w = \sum_{i=1}^{k} p_i s_i$ = weighted mean square or $\bar{S}_w = \frac{1}{N}\sum N_i S_i$

the weights being equal to the stratum size

$$= \frac{1}{n}\left[\sum_{i=1}^{k} p_i (s_i - \bar{S}_w)^2\right]$$

≥ 0 (non-ve term)

$V(\bar{y}_{st})_p \geq V(\bar{y}_{st})_N$

if var. decr.

var. कम होता है तो efficient होता है

from A and B we get

$$\boxed{V(\bar{y}_{st})_N \le V(\bar{y}_{st})_P \le V(\bar{y}_n)_{SRSWOR}}$$

Estimate of percentage gain in efficiency due to stratified sampling (stratification) over SRSWOR :→ The estimate of percentage gain in efficiency due to stratification over SRS is define as;

$\hat{V}(\bar{y}_{st})$ stratified var. का estim

$$= \frac{\hat{V}(\bar{y}_n)_{SRS} - \hat{V}(\bar{y}_{st})}{\hat{V}(\bar{y}_{st})} \quad \text{---} (1)$$

we have

$$V(\bar{y}_{st}) = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_i^2$$

$$\hat{V}(\bar{y}_{st}) = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 \hat{S}_i^2$$

$$\therefore \hat{V}(\bar{y}_{st}) = \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 \hat{S}_i^2$$

[as $E(s_i^2) = S_i^2$ by the prop of SRSWO"

where s_i^2 = sample mean square of the i^{th} stratum

$$\hat{S}_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{n_i})^2$$

$\bar{Y}_N = \sum k_i \bar{Y}_{Ni}$ sample Mean pop^n Mean are unbiased estimate of 1

we also have

$$V(\bar{y}_n)_{SRS} \approx \left(\frac{1}{n}-\frac{1}{N}\right)\left[\sum_{i=1}^{k} k_i S_i^2 + \sum_{i=1}^{k} k_i (\bar{Y}_{Ni}\right.$$

or $$V(\bar{y}_n)_{SRS} \approx \left(\frac{1}{n}-\frac{1}{N}\right)\left[\sum_{i=1}^{k} k_i S_i^2 + \sum_{i=1}^{k} k_i (\bar{Y}_{Ni}^2 - \bar{Y}_N\right.$$

$$\hat{V}(\bar{y}_n)_{SRS} \approx \left(\frac{1}{n}-\frac{1}{N}\right)\left[\sum_{i=1}^{k} k_i \hat{S}_i^2 + \sum_{i=1}^{k} k_i (\hat{\bar{Y}}_{Ni}^2 - \hat{\bar{Y}}_N^2\right.$$

Consider

$$V(\bar{y}_{ni}) = E(\bar{y}_{ni}^2) - [E(\bar{y}_{ni})]^2$$

sample mean of i^{th} stratum

$$= E(\bar{y}_{ni}^2) - \bar{Y}_{Ni}^2$$

$$\Rightarrow \bar{Y}_{Ni}^2 = E(\bar{y}_{ni}^2) - V(\bar{y}_{ni})$$

$$= E(\bar{y}_{ni}^2) - \left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_i^2$$

$$\Rightarrow \hat{\bar{Y}}_{Ni}^2 = \bar{y}_{ni}^2 - \left(\frac{1}{n_i}-\frac{1}{N_i}\right) s_i^2 \quad (3)$$

again Consider

$$V(\bar{y}_{st}) = E(\bar{y}_{st}^2) - [E(\bar{y}_{st})]^2$$

$$\sum_{i=1}^{k}\left(\frac{1}{n_i}-\frac{1}{N_i}\right) k_i^2 S_i^2 = E(\bar{y}_{st})^2 - \bar{Y}_N^2$$

$$\Rightarrow \bar{Y}_N^2 = E(\bar{y}_{st}^2) - \sum_{i=1}^{k}\left(\frac{1}{n_i}-\frac{1}{N_i}\right) k_i^2 S_i^2$$

$$\Rightarrow \hat{\bar{Y}}_N^2 = \bar{y}_{st}^2 - \sum_{i=1}^{k}\left(\frac{1}{n_i}-\frac{1}{N_i}\right) k_i^2 s_i^2$$

$E(s_i^2) = S_i^2 \quad S_i^2 \stackrel{a}{=} s_i^2$

$\hat{S}_i^2 = s_i^2$

Sub the value from eqn (3) & (4) in eqn (2) we get;

$$\hat{V}(\bar{y}_n)_{SRS} \simeq \left(\frac{1}{n} - \frac{1}{N}\right)\left[\sum_{i=1}^{k} p_i s_i^2 + \sum_{i=1}^{k} p_i \left\{ \bar{y}_{n_i}^2 - \left(\frac{1}{n_i} - \frac{1}{N_i}\right) s_i^2 \right\} - \bar{y}_{st}^2 + \sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 s_i^2 \right] \quad \text{--- (5)}$$

by putting the values $\hat{V}(\bar{y}_{st})$ and $\hat{V}(\bar{y}_n)_{SRS}$ in eqn. (1)

we can get, the estimate of percentage gain in efficiency due to stratification over SRS.

~~Cluster Sampling~~ :→

$v(x) = E(x^2) - E(x)^2$