Journal of Modern Applied Statistical Methods

Volume 17 | Issue 1 Article 13

6-19-2018

# Optimum Stratification in Bivariate Auxiliary Variables under Neyman Allocation

Faizan Danish
*Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, J&K India*, secsportsfbsc@gmail.com

S.E.H. Rizvi
*Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, J&K India*, rizvisehstat@gmail.com

# Optimum Stratification in Bivariate Auxiliary Variables under Neyman Allocation

**Cover Page Footnote**

Respected Sir/Madam, Sir/Madam I have prepared the paper as suggested now and made the changes suggested by reviewers by including some Lemma's that may provide the decorum of the work.The rebuttal file has also been prepared Sir/Madam Thanking you Sir/Madam

Journal of Modern Applied Statistical Methods
May 2018, Vol. 17, No. 1, eP 2580
doi: 10.22237/jmasm/1529418671

בס"ד


ISSN 1538 − 9472

# Optimum Stratification in Bivariate Auxiliary Variables under Neyman Allocation

**Faizan Danish**
Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu
Jammu, India

**S. E. H. Rizvi**
Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu
Jammu, India

When the complete data set of the study variable is unknown it produces a possible stumbling block in attempting various stratification techniques. A technique is proposed under Neyman allocation when the stratification is done on the two auxiliary variables having one estimation variable under consideration. Because of complexities made by minimal equations, approximate optimum strata boundaries are obtained. An empirical study illustrates the proposed method when the auxiliary variables have standard Cauchy and power distributions.

*Keywords:* Stratification points, bi-variate distribution, power distribution, standard Cauchy distribution

## Introduction

A populace might be homogenous or heterogeneous. For the latter, one approach is to isolate it into sub-populaces, which are known as strata. Limiting change by stratifying is known as ideal stratification. There are different factors that are responsible for minimum variance, which include choosing the variable on the basis of which stratification would be done, total number of strata, the design by which sample size will be selected from each stratum, and, the most influential factor, the demarcation of strata.

The use of a single stratification variable may be problematic, and in any case, using more than one stratification variable increases the level of exactness. Dalenius (1950) pioneered the work of obtaining optimum strata boundaries by minimizing the variance. See also Dalenius and Gurney (1951), Mahalanobis

doi: 10.22237/jmasm/1529418671 | Accepted: September 29, 2017; Published: June 19, 2018.
Correspondence: Faizan Danish, danishstat@gmail.com

(1952), Aoyama (1954), Dalenius and Hodges (1959), Durbin (1959), Singh (1975, 1977), and Verma (2008). However, their equations were mostly taken by using an estimation variable as a stratification variable. Others used a variable highly correlated to the study variable as stratification variable, such as Taga (1967), Serfling (1968), Singh and Sukhatme (1969, 1972, 1973), Singh (1971), Singh and Parkash (1975), Schneeberger and Goller (1979), and Rizvi, Gupta, and Singh (2002). Iterative procedures were also proposed by Rivest (2002) for obtaining stratification points, and Gunning and Horgan (2004) proposed a new algorithm for the skewed population.

The motivation behind the present examination is to consider a solitary report variable and two factors exceedingly connected with it. The stratification will be conducted based on auxiliary variables. For numerical illustration of the proposed method, two different distributions will be considered for the auxiliary variables.

## Stratification Points

Let a population of size $N$ units be divided into $T \times U$ strata, and let $N_{rs}$ denote the number of units in the $(r, s)^{th}$ stratum. Suppose a sample size $n$ is to be taken from the whole population, and let $n_{rs}$ denote the sample size allocated to the $(r, s)^{th}$ stratum and $z_{rsi}$ the values of the population units in the $(r, s)^{th}$, $i = 1, 2, 3, \ldots$ Let the variable $Z$ be the study variable defined by

$$Z = \sum_{r=1}^{T} \sum_{s=1}^{U} \sum_{i=1}^{N_{rs}} z_{rsi}$$

The unbiased estimate of population mean is

$$\bar{z}_{st} = \sum_{r=1}^{T} \sum_{s=1}^{U} W_{rs} \bar{z}_{rs}$$

where $W_{rs}$ is the weight for the $(r, s)^{th}$ stratum.

For obtaining strata boundaries, assume that a finite population consists of $N$ units. The stratification points $[z_{rs}]$ for the case of optimum allocation can be obtained by the following equations:

$$\frac{\left(\sigma_{rsz}^2+\left(z_{rs}-\mu_{rsz}\right)^2\right)-2\sigma_{rsz}^2 z_{rs}\mu_{rsz}}{\sigma_{rsz}^2}=\frac{\left(\sigma_{ijz}^2+\left(z_{rs}-\mu_{ijz}\right)^2\right)-2\sigma_{ijz}^2 y_{rs}\mu_{ijz}}{\sigma_{ijz}^2}$$

where $i = h + 1, h + 2,\ldots, T - 1$, and $j = k + 1, k + 2,\ldots, U - 1$, and $\mu_{rsz}$ denotes the mean of the population for $Z$ in $(r, s)^{\text{th}}$ stratum.

The minimal equations given above were obtained by Dalenius (1950) when the stratification variable is same as the estimation variable. When the density functions of the auxiliary variables $Y$ and $X$ are known, then the distribution of $Z$ is not known due to the auxiliary variables used to obtain optimum points. Assume the regression line of $Z$ on $Y$ and $X$ is linear, given as

$$Z=\lambda\left(Y,X\right)+e \tag{1}$$

where $\lambda(Y, X)$ is a function of $Y$ and $Z$ and $e$ denotes the error term such that

$$\mathrm{E}\left(\frac{e}{y,x}\right)=0 \text{ and } \mathrm{V}\left(\frac{e}{y,x}\right)=\phi\left(y,x\right)=0$$

defined in $(a, b)$. Let $f(z, y, x)$ denote the density function of the population $(Z, Y, X)$ and let $f(y, x)$ denote the joint marginal density function of $Y$ and $X$. Also, let $f(y)$ and $f(x)$ be the marginal density functions of $Y$ and $X$, respectively. According to the model defined in (1),

$$\mu_{rsz}=\mu_{rs\lambda}=\frac{1}{W_{rs}}\int_{y_{r-1}}^{y_r}\int_{x_{s-1}}^{x_s}\lambda\left(y,x\right)\mathrm{f}\left(y,x\right)\partial y\partial x \tag{2}$$

which denotes the mean of the $(r, s)^{\text{th}}$ stratum, where

$$W_{rs}=\int_{y_{r-1}}^{y_r}\int_{x_{s-1}}^{x_s}\mathrm{f}\left(y,x\right)\partial y\partial x \tag{3}$$

denotes the weight of the $(r, s)^{\text{th}}$ stratum and the variance of the stratum is given by

$$\sigma_{rsz}^2=\sigma_{rsc}^2+\mu_{rs\phi} \tag{4}$$

$$\sigma^2_{rs\lambda} = \frac{1}{W_{rs}} \int_{y_{r-1}}^{y_r} \int_{x_{s-1}}^{x_s} \lambda^2(y,x) \mathrm{f}(y,x) \partial y \partial x - \left( \frac{1}{W_{rs}} \int_{y_{r-1}}^{y_r} \int_{x_{s-1}}^{x_s} \phi(y,x) \mathrm{f}(y,x) \partial y \partial x \right)^2 \quad (5)$$

where ($y_{r-1}$, $y_r$, $x_{s-1}$, $x_s$) denotes the boundaries and $\mu_{rs\phi}$ is the expected value of the function $\phi(y, x)$ of the $(r, s)^{\text{th}}$ stratum.

## Optimum Variance Equation

Let ($g$, $h$) and ($k$, $L$) be the defined intervals for the variables $Y$ and $X$ which are needed to estimate the stratification point ($y_r$, $x_s$) so the variance of the estimate is minimum. The stratification points so obtained would be the result of taking partial derivatives of (4) with respect to the stratification points. In order to obtain the stratification points of the $(r, s)^{\text{th}}$ stratum, find the partial derivatives. Differentiate (3) with respect to $y_r$ and $x_s$:

$$\alpha = \int_{x_{s-1}}^{x_s} \mathrm{f}(y_r, x) \partial x \quad (6)$$

and

$$\alpha = \int_{y_{r-1}}^{y_r} \mathrm{f}(y, x_s) \partial y \quad (7)$$

where $\alpha$ and $\beta$ are the first partial derivatives of (3) with respect to $y_r$ and $x_s$, respectively.

Also by differentiating

$$\mu_{rs\phi} = \frac{1}{W_{rs}} \int_{y_{r-1}}^{y_r} \int_{x_{s-1}}^{x_s} \phi(y,x) \mathrm{f}(y,x) \partial y \partial x$$

with respect to $y_r$ and $x_s$,

$$\gamma = \int_{x_{s-1}}^{x_s} A_1 \frac{\mathrm{f}\left(y_r, x\right)}{w_{rs}} \partial x \tag{8}$$

$$\delta = \int_{y_{r-1}}^{y_r} A_2 \frac{\mathrm{f}\left(y, x_s\right)}{w_{rs}} \partial y \tag{9}$$

where $A_1 = \phi(y_r, x) - \mu_{rs\phi}$, $A_2 = \phi(y, x_s) - \mu_{rs\phi}$, and $\gamma$ and $\delta$ denote the first partial derivatives of $\mu_{rs\phi}$ with respect to $y_r$ and $x_s$, respectively.

Similarly, while finding the first partial derivatives of (2) with respect to $y_r$ and $x_s$, respectively,

$$\int_{x_{s-1}}^{x_s} A_3 \partial x \tag{10}$$

$$\int_{y_{r-1}}^{y_r} A_4 \partial y \tag{11}$$

where

$$A_3 = \frac{\mathrm{f}\left(y_r, x\right)}{w_{rs}} \left[\lambda\left(y_r, x\right) - \mu_{rs\lambda}\right] \text{ and } A_4 = \frac{\mathrm{f}\left(y, x_s\right)}{w_{rs}} \left[\lambda\left(y, x_s\right) - \mu_{rs\lambda}\right]$$

Again, partially differentiate (5) with respect to $y_r$ and $x_s$:

$$\frac{1}{w_{rs}} \int_{x_{s-1}}^{x_s} A_5 \partial x \tag{12}$$

$$\frac{1}{w_{rs}} \int_{y_{r-1}}^{y_r} A_6 \partial y \tag{13}$$

where

$$A_5 = \mathrm{f}\left(y_r, x\right)\left[\lambda\left(y_r, x\right) - \mu_{rs\lambda}\right] - \sigma_{rs\lambda}^2 \text{ and } A_6 = \mathrm{f}\left(y, x_s\right)\left[\lambda\left(y, x_s\right) - \mu_{rs\lambda}\right]^2 - \sigma_{rs\lambda}^2$$

Similarly, for the $(r + 1, s + 1)^{th}$ stratum, while taking the partial derivatives with respect to $y_r$ and $x_s$, we get

$$W'_{(r+1)s} = -\int_{x_{s-1}}^{x_s} \mathrm{f}\left(y_r, x\right)\partial x \tag{14}$$

$$W'_{r(s+1)} = \int_{y_{r-1}}^{y_r} \mathrm{f}\left(y, x_s\right)\partial y \tag{15}$$

$$\mu'_{(r+1)s\phi} = -\int_{x_{s-1}}^{x_s} A_7 \partial x \tag{16}$$

$$\mu'_{r(s+1)\phi} = -\int_{y_{r-1}}^{y_r} A_8 \partial y \tag{17}$$

$$\mu'_{(r+1)s\lambda} = -\int_{x_{s-1}}^{x_s} A_9 \partial x \tag{18}$$

$$\mu'_{r(s+1)\lambda} = -\int_{y_{r-1}}^{y_r} A_{10} \partial y \tag{19}$$

$$\sigma'_{(r+1)s\lambda} = -\frac{1}{w_{(r+1)s}} \int_{x_{r-1}}^{x_r} \left\{ \mathrm{f}\left(y_r, x\right) \left[ \lambda\left(y_r, x\right) - \frac{1}{W_{(r+1)s}} \int_{y_r}^{y_{r+1}} \int_{x_{s-1}}^{x_s} \lambda\left(y, x\right) \mathrm{f}\left(y, x\right) \partial y \partial x \right]^2 - \sigma^2_{(r+1)s\lambda} \right\} \partial x \tag{20}$$

$$\sigma'_{r(s+1)\lambda} = -\frac{1}{w_{r(s+1)}} \int_{y_{s-1}}^{y_s} \left\{ \mathrm{f}\left(y, x_s\right) \left[ \lambda\left(y, x_s\right) - \frac{1}{W_{r(s+1)}} \int_{y_{r-1}}^{y_r} \int_{x_s}^{x_{s+1}} \lambda\left(y, x\right) \mathrm{f}\left(y, x\right) \partial y \partial x \right]^2 - \sigma^2_{r(s+1)\lambda} \right\} \partial y \tag{21}$$

where

$$A_7 = \frac{\mathrm{f}\left(y_r, x\right)}{w_{(r+1)s}} \left[ \phi\left(y_r, x\right) - \mu_{(r+1)s\phi} \right], \quad A_8 = \frac{\mathrm{f}\left(y, x_s\right)}{w_{r(s+1)}} \left[ \phi\left(y, x_s\right) - \mu_{r(s+1)\phi} \right],$$

$$A_9 = \frac{\mathrm{f}\left(y_r, x\right)}{w_{(r+1)s}} \left[ \lambda\left(y_r, x\right) - \mu_{(r+1)s\lambda} \right], \quad A_{10} = \frac{\mathrm{f}\left(y, x_s\right)}{w_{r(s+1)}} \left[ \lambda\left(y, x_s\right) - \mu_{r(s+1)\lambda} \right]$$

Under the Neyman allotment, the fluctuation of the example mean is

$$\mathrm{V}\left(\overline{z}_{st}\right) = \frac{\left[ \sum_{r=1}^{T} \sum_{k=1}^{U} W_{rs} \sigma_{rsz} \right]^2}{n} \tag{22}$$

However, if the finite population correction is ignored, minimizing (22) is equivalent to minimizing

$$\mathrm{V}\left(\overline{z}_{st}\right) = \sum_{r=1}^{T} \sum_{s=1}^{U} W_{rs} \sqrt{\sigma^2_{rsz}}$$

This can be rewritten, using (4), as

$$\mathrm{V}\left(\overline{z}_{st}\right) = \sum_{r=1}^{T} \sum_{s=1}^{U} W_{rs} \sqrt{\sigma^2_{rs\lambda} + \mu_{rs\lambda}} \tag{23}$$

By differentiating (23) with respect to $y_r$ and then equating to zero,

$$\begin{aligned} W_{rs}\frac{\partial}{\partial y_r}\left(\sqrt{\sigma^2_{rs\lambda}+\mu_{rs\phi}}\right)+\sqrt{\sigma^2_{rs\lambda}+\mu_{rs\phi}}\frac{\partial}{\partial y_r}W_{rs} \\ +W_{(r+1)s}\frac{\partial}{\partial y_r}\left(\sqrt{\sigma^2_{(r+1)s\lambda}+\mu_{(r+1)s\phi}}\right)+\sqrt{\sigma^2_{(r+1)s\lambda}+\mu_{(r+1)s\phi}}\frac{\partial}{\partial y_r}W_{(r+1)s} \end{aligned} \tag{24}$$

Using equations (14)-(19),

$$\frac{\partial}{\partial y_r}\left(\sigma^2_{rs\lambda}+\mu_{rs\phi}\right)=\int_{x_{s-1}}^{x_s}\frac{\mathrm{f}\left(y_r,x\right)}{W_{rs}}\left\{I_1+\phi\left(y_r,x\right)-\mu_{rs\lambda}\right\}\partial x \tag{25}$$

$$\frac{\partial}{\partial y_r}\left(\sigma^2_{(r+1)s\lambda}+\mu_{(r+1)s\phi}\right)=\int_{x_{s-1}}^{x_s}\frac{\mathrm{f}\left(y_r,x\right)}{W_{rs}}\left\{I_s+\phi\left(y_r,x\right)-\mu_{(r+1)s\lambda}\right\}\partial x \tag{26}$$

where

$$I_1=\left[\lambda\left(y_r,x\right)-\mu_{rs\lambda}\right]^2-\sigma^2_{rs\lambda} \text{ and } I_2=\left[\lambda\left(y_r,x\right)-\mu_{(r+1)s\lambda}\right]^2-\sigma^2_{(r+1)s\lambda}$$

The minimal equations can be obtained by substituting the values obtained in (25) and (26):

$$\frac{\int_{x_{s-1}}^{x_s}\left\{I_3+\mu_{rs\phi}\right\}\partial x}{\sqrt{\sigma^2_{rs\lambda}+\mu_{rs\phi}}}=\frac{\int_{x_{s-1}}^{x_s}\left\{I_4+\mu_{(r+1)s\phi}\right\}\partial x}{\sqrt{\sigma^2_{(r+1)s\lambda}+\mu_{(r+1)s\phi}}} \tag{27}$$

where

$$\begin{aligned} I_3&=\mathrm{f}\left(y_r,x\right)\left[\lambda\left(y_r,x\right)-\mu_{rs\phi}\right]^2+\sigma^2_{rs\lambda}+\phi\left(y_r,x\right) \\ I_4&=\mathrm{f}\left(y_r,x\right)\left[\lambda\left(y_r,x\right)-\mu_{(r+1)s\phi}\right]^2+\sigma^2_{(r+1)s\lambda}+\phi\left(y_r,x\right) \end{aligned}$$

By differentiating (23) with respect to $x_s$ and equating to zero, the minimal equations are

$$\frac{\int_{y_{r-1}}^{y_r}\left\{f\left(y,x_s\right)\left[\lambda\left(y,x_s\right)-\mu_{rs\phi}\right]^2+I_5\right\}\partial y}{\sqrt{\sigma_{rs\lambda}^2+\mu_{rs\phi}}} = \frac{\int_{y_{r-1}}^{y_r}\left\{f\left(y,x_s\right)\left[\lambda\left(y,x_s\right)-\mu_{r(s+1)\phi}\right]^2+I_6\right\}\partial y}{\sqrt{\sigma_{r(s+1)\lambda}^2+\mu_{r(s+1)\phi}}} \tag{28}$$

where

$$I_5=\sigma_{rs\lambda}^2+\phi\left(y,x_s\right)+\mu_{rs\phi} \text{ and } I_6=\sigma_{r(s+1)\lambda}^2+\phi\left(y,x_s\right)+\mu_{r(s+1)\phi}$$

Equations (27) and (28) result in strata boundaries ($y_r$, $x_s$) corresponding to the minimum of V($\bar{z}_{st}$) of the function

$$\psi\left(y,x\right)=f\left(x,z\right)\frac{\left[4\lambda'^2\left(y,x\right)+I_7'^2\right]}{\left[I_7\right]^{\frac{3}{2}}}\in\Delta,\forall y\in\left[g,h\right],x\in\left[k,L\right]$$

where $I_7 = \phi(y, x)$.

Assuming that the regression model defined in (1) is linear of the form $z = a + by + cx + e$,

$$\sigma_{rsz}^2=b^2\sigma_{rsy}^2+c^2\sigma_{rsx}^2+\sigma_e^2 \tag{29}$$

where the expected value and variance of the error term $e$ are 0 and $\sigma^2$, respectively, and assume that the auxiliary variables are independent of each other. Equation (23) can be written as

$$\sum_{r=1}^{T}\sum_{s=1}^{U}\sqrt{W_{rs}^2\left(h\right)} \tag{30}$$

where $h=b^2\sigma_{rsy}^2+c^2\sigma_{rsx}^2+\sigma_e^2+\mu_{rs\phi}$. From the above,

$$\sum_{r=1}^{T}\sum_{s=1}^{U}\sqrt{W_{rs}^2\left(h\right)}=\frac{\sum_{h=1}^{L}\sum_{k=1}^{M}\varphi_{rs}^2\varphi_r^{-\frac{3}{2}}\varphi_s K_1^2 K_2^2\sqrt{\varphi_r\varphi_s}}{4\times 3} \tag{31}$$

where $K_1 = [y_r - y_{r-1}]$, $K_2 = [x_s - x_{s-1}]$, and $\varphi_r$, $\varphi_s$ represent unchanged values of the marginal density functions of $Y$ and $X$ in $(r, s)^{\text{th}}$ stratum, respectively.

***Lemma 1.*** If the function $\mathrm{I}_{ij}(y, x)$ is defined as

$$\mathrm{I}_{ij}\left(y,x\right)=\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left(t_1-y_1\right)^i\left(t_2-x_1\right)^j \mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2,\quad y_1<y_2, x_1<x_2$$

where $\mathrm{f}(t_1, t_2)$ is a function of two variables, then

$$\begin{aligned}\mathrm{I}_{ij}\left(y,x\right)&=\frac{k_1^{i+1}k_2^{j+1}}{(i+1)(j+1)}\mathrm{f}+\frac{k_1^{i+2}k_2^{j+1}}{(i+2)(j+1)}\mathrm{f}_y+\frac{k_1^{i+1}k_2^{j+2}}{(i+1)(j+2)}\mathrm{f}_x\\&+\frac{1}{2!}\left[\frac{k_1^{i+3}k_2^{j+1}}{(i+3)(j+1)}\mathrm{f}_{yy}+2\frac{k_1^{i+2}k_2^{j+2}}{(i+2)(j+2)}\mathrm{f}_{yx}+\frac{k_1^{i+1}k_2^{j+3}}{(i+1)(j+3)}\mathrm{f}_{xx}\right]+\mathrm{O}\left(k^{i+j+5}\right)\end{aligned}\tag{32}$$

***Proof.*** If $(t_1, t_2)$ is near $(y_1, x_1)$ and derivatives of f are continuous, then expand $\mathrm{f}(t_1, t_2)$ with the help of Taylor's theorem. Define $t_1 = y_1 + (t_1 - y_1)$ and $t_2 = x_1 + (t_2 - x_1)$. Then $\mathrm{f}(t_1, t_2) = \mathrm{f}(t_1 = y_1 + (t_1 - y_1), t_2 = x_1 + (t_2 - x_1))$. Using the Taylor series formula for a function of two variables, the expansion of $\mathrm{f}(t_1, t_2)$ is given by

$$\begin{aligned}\mathrm{f}\left(t_1,t_2\right)=f\left(y_1,x_2\right)&+\left(t_1-y_1\right)\frac{\partial \mathrm{f}}{\partial t_1}+\left(t_2-x_1\right)\frac{\partial \mathrm{f}}{\partial t_2}\\&+\frac{\left(t_1-y_1\right)^2}{2!}\frac{\partial^2 \mathrm{f}}{\partial t_1^2}+\frac{\left(t_2-x_1\right)^2}{2!}\frac{\partial^2 \mathrm{f}}{\partial t_2^2}+\dots\end{aligned}$$

which yields

$$\begin{aligned}\mathrm{I}_{ij}\left(y,x\right)=\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left(t_1-y_1\right)^i\left(t_2-x_1\right)^j \mathrm{f}\left(y_1,x_2\right)&+\left(t_1-y_1\right)\frac{\partial \mathrm{f}}{\partial t_1}+\left(t_2-x_1\right)\frac{\partial \mathrm{f}}{\partial t_2}\\&+\frac{\left(t_1-y_1\right)^2}{2!}\frac{\partial^2 \mathrm{f}}{\partial t_1^2}+\frac{\left(t_2-x_1\right)^2}{2!}\frac{\partial^2 \mathrm{f}}{\partial t_2^2}+\dots\end{aligned}$$

$$\mathrm{I}_{ij}(y,x)=\frac{k_1^{i+1}k_2^{j+1}}{(i+1)(j+1)}\mathrm{f}(y_1,x_1)+\frac{k_1^{i+2}k_2^{j+1}}{(i+2)(j+1)}\frac{\partial \mathrm{f}}{\partial t_1}+\frac{k_1^{i+1}k_2^{j+2}}{(i+1)(j+2)}\frac{\partial \mathrm{f}}{\partial t_2}$$
$$+\frac{1}{2!}\frac{k_1^{i+3}k_2^{j+1}}{(i+3)(j+1)}\frac{\partial^2 \mathrm{f}}{\partial t_2^2}+\frac{1}{2!}\frac{k_1^{i+1}k_2^{j+3}}{(i+1)(j+3)}\frac{\partial^2 \mathrm{f}}{\partial t_1^2}+\frac{1}{2!}\frac{k_1^{i+2}k_2^{j+2}}{(i+2)(j+2)}\frac{\partial^2 \mathrm{f}}{\partial t_1\partial t_2}+\mathrm{O}\left(k^{i+j+5}\right)$$

at $t_1 = y_1$, where $k_1 = y_2 - y_1$ and $k_2 = x_2 - x_1$. Denote

$$\mathrm{f}(y_1,x_1)=\mathrm{f},\frac{\partial \mathrm{f}}{\partial t_1}=\mathrm{f}_y,\frac{\partial \mathrm{f}}{\partial t_2}=\mathrm{f}_x,\frac{\partial^2 \mathrm{f}}{\partial t_1^2}=\mathrm{f}_{yy},\frac{\partial^2 \mathrm{f}}{\partial t_2^2}=\mathrm{f}_{xx},\frac{\partial^2 \mathrm{f}}{\partial t_1\partial t_2}=\mathrm{f}_{yx}$$

Then $\mathrm{I}_{ij}(y, x)$ can be written as

$$\begin{aligned}\mathrm{I}_{ij}(y,x)&=\frac{k_1^{i+1}k_2^{j+1}}{(i+1)(j+1)}\mathrm{f}+\frac{k_1^{i+2}k_2^{j+1}}{(i+2)(j+1)}\mathrm{f}_y+\frac{k_1^{i+1}k_2^{j+2}}{(i+1)(j+2)}\mathrm{f}_x\\&+\frac{1}{2!}\left[\frac{k_1^{i+3}k_2^{j+1}}{(i+3)(j+1)}\mathrm{f}_{yy}+\frac{k_1^{i+1}k_2^{j+3}}{(i+1)(j+3)}\mathrm{f}_{xx}+\frac{2k_1^{i+2}k_2^{j+2}}{(i+2)(j+2)}\mathrm{f}_{yx}\right]+\mathrm{O}\left(k^{i+j+5}\right)\end{aligned}\tag{33}$$

where $k$ indicates $k_1$ or $k_2$.

For $i = j = 0$,

$$\begin{aligned}\mathrm{I}_{00}(y,x)&=k_1k_2\,\mathrm{f}+\frac{k_1^2k_2}{(2)}\mathrm{f}_y+\frac{k_1k_2^2}{(2)}\mathrm{f}_x\\&+\frac{1}{2!}\left[\frac{k_1^3k_2}{(3)}\mathrm{f}_{yy}+\frac{k_1k_2^3}{(3)}\mathrm{f}_{xx}+\frac{2k_1^2k_2^2}{(2)(2)}\mathrm{f}_{yx}\right]+\mathrm{O}\left(k^5\right)\end{aligned}\tag{34}$$

***Lemma 2.*** Let $\mu_\eta(y, x)$ denote the conditional expectation of the function $\eta(t_1, t_2)$ so that

$$\mu_\eta(y,x)=\frac{\int_{x_1}^{x_2}\int_{y_1}^{y_2}\eta(t_1,t_2)\,\mathrm{f}(t_1,t_2)\,\partial t_1\partial t_2}{\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}(t_1,t_2)\,\partial t_1\partial t_2}$$

Then the series expansion of $\mu_\eta(y, x)$ at point $(t_1, t_2)$ is given by

$$
\begin{aligned}
\mu_\eta(y,x) = \eta + \Bigg[ & \frac{\eta'\,\mathrm{f}}{2}\left(k_1^2 k_2 + k_1 k_2^2\right) + \frac{4k_1^3 k_2 + 3k_1^2 k_2^2}{12}\left(\eta''\,\mathrm{f} + \eta'\,\mathrm{f}_y + \eta'\,\mathrm{f}_x\right) \\
& + \frac{k_1^4 k_2}{8}\left(\eta''\,\mathrm{f}_{yy} + 2\eta''\,\mathrm{f}_y\right) + \frac{k_1^3 k_2^2}{6}\left(\eta''\,\mathrm{f}_{yx} + \eta''\,\mathrm{f}_y + \eta''\,\mathrm{f}_x + \frac{\eta'\,\mathrm{f}_{yy}}{2}\right) \\
& + \frac{k_1^2 k_2^3}{12}\left(\eta'\,\mathrm{f}_{yx} + \eta'\,\mathrm{f}_{xx} + 2\eta''\,\mathrm{f}_y + 2\eta''\,\mathrm{f}_x\right) + \frac{k_1 k_1^4}{8}\left(\eta'\,\mathrm{f}_{xx} + \eta''\,\mathrm{f}_x\right) + o\left(k^6\right) \Bigg] \\
& \div \left[ k_1 k_2\,\mathrm{f} + \frac{k_1^2 k_2}{2}\mathrm{f}_y + \frac{k_1 k_2^2}{2}\mathrm{f}_x + \frac{1}{2!}\left(\frac{k_1^3 k_2}{3}\mathrm{f}_{yy} + \frac{2k_1^2 k_2^2}{4}\mathrm{f}_{yx} + \frac{k_1 k_2^3}{3}\mathrm{f}_{xx}\right)\right]
\end{aligned}
\tag{35}
$$

***Proof.*** From the definition of

$$
\begin{aligned}
\mu_\eta(y,x)\int_{x_1}^{x_2}\int_{y_1}^{y_2} \mathrm{f}(t_1,t_2)\,\partial t_1 \partial t_2 &= \int_{x_1}^{x_2}\int_{y_1}^{y_2} \eta(t_1,t_2)\,\mathrm{f}(t_1,t_2)\,\partial t_1 \partial t_2 \\
&= \mu_\eta(y,x)\,\mathrm{I}_{00}(y,x) \\
&= \int_{x_1}^{x_2}\int_{y_1}^{y_2} \eta(t_1,t_2)\,\mathrm{f}(t_1,t_2)\,\partial t_1 \partial t_2
\end{aligned}
$$

Using the Taylor series expansion, defined for two variables, of $\eta(t_1, t_2)$ about the point $(t_1, t_2) = (y, x)$,

$$
\begin{aligned}
\eta(t_1,t_2) = \eta &+ (t_1 - y)\eta' + (t_2 - x)\eta' + \frac{(t_1-y)^2}{2!}\eta'' + \frac{(t_2-x)^2}{2!}\eta'' \\
&+ \frac{2(t_1-y)(t_2-x)}{2!}\eta'' + \frac{(t_1-y)^3}{3!}\eta''' + \frac{(t_2-x)^3}{3!}\eta''' + \frac{(t_1-y)^2(t_2-x)}{3!}\eta''' \\
&+ \frac{(t_1-y)(t_2-x)^2}{3!}\eta''' + \dots
\end{aligned}
$$

Thus,

$$\begin{aligned}\mu_\eta(y,x)\mathrm{I}_{00}(y,x) &= \int_{x_1}^{x_2}\int_{y_1}^{y_2}\left[\eta+(t_1-y)\eta'+(t_2-x)\eta'+\frac{(t_1-y)^2}{2!}\eta''+\frac{(t_2-x)^2}{2!}\eta''\right.\\ &\qquad\left.+\frac{2(t_1-y)(t_2-x)}{2!}\eta''+\ldots\right]\mathrm{f}(t_1,t_2)\,\partial t_1\partial t_2\\ &= \mathrm{I}_{00}(y,x)\eta+\mathrm{I}_{10}(y,x)\eta'+\mathrm{I}_{01}(y,x)\eta'+\mathrm{I}_{20}(y,x)\eta''+\mathrm{I}_{02}(y,x)\eta''\\ &\qquad+\mathrm{I}_{11}(y,x)\eta''+\mathrm{o}\left(k^5\right)\end{aligned}$$

Neglecting the higher-order terms, it can be written as

$$\begin{aligned}&\mu_\eta(y,x)\\ &\quad=\eta+\frac{\mathrm{I}_{10}(y,x)\eta'+\mathrm{I}_{01}(y,x)\eta'+\mathrm{I}_{20}(y,x)\eta''+\mathrm{I}_{02}(y,x)\eta''+\mathrm{I}_{11}(y,x)\eta''}{\mathrm{I}_{00}(y,x)}\end{aligned}\tag{36}$$

By substituting values of $\mathrm{I}_{00}(y, x)$, $\mathrm{I}_{10}(y, x)$, $\mathrm{I}_{01}(y, x)$, $\mathrm{I}_{20}(y, x)$, $\mathrm{I}_{02}(y, x)$, and $\mathrm{I}_{11}(y, x)$ from (33) in (36), we get

$$\begin{aligned}\mu_\eta(y,x)=\eta+&\left[\frac{\eta'\mathrm{f}}{2}\left(k_1^2k_2+k_1k_2^2\right)+\frac{4k_1^3k_2+3k_1^2k_2^2}{12}\left(\eta''\mathrm{f}+\eta'\mathrm{f}_y+\eta'\mathrm{f}_x\right)\right.\\ &+\frac{k_1^4k_2}{8}\left(\eta''\mathrm{f}_{yy}+2\eta''\mathrm{f}_y\right)+\frac{k_1^3k_2^2}{6}\left(\eta''\mathrm{f}_{yx}+\eta''\mathrm{f}_y+\eta''\mathrm{f}_x+\frac{\eta'\mathrm{f}_{yy}}{2}\right)\\ &\left.+\frac{k_1^2k_2^3}{12}\left(\eta'\mathrm{f}_{yx}+\eta'\mathrm{f}_{xx}+2\eta''\mathrm{f}_y+2\eta''\mathrm{f}_x\right)+\frac{k_1k_1^4}{8}\left(\eta'\mathrm{f}_{xx}+\eta''\mathrm{f}_x\right)\right]\\ &\div\left[k_1k_2\,\mathrm{f}+\frac{k_1^2k_2}{2}\mathrm{f}_y+\frac{k_1k_2^2}{2}\mathrm{f}_x+\frac{1}{2!}\left(\frac{k_1^3k_2}{3}\mathrm{f}_{yy}+\frac{2k_1^2k_2^2}{4}\mathrm{f}_{yx}+\frac{k_1k_2^3}{3}\mathrm{f}_{xx}\right)+\mathrm{o}\left(k^6\right)\right]\end{aligned}$$

Continuing in a similar manner and utilizing Taylor's theorem at the point $z$, the expansion for $\mu_\eta(y, x)$ is obtained as

$$
\begin{aligned}
\mu_\eta(y,x) = \eta + \Bigg[ & \frac{\eta' \mathrm{f}}{2}\left(k_1^2 k_2 - k_1 k_2^2\right) + \frac{4k_1^3 k_2 + 3k_1^2 k_2^2}{12}\left(\eta'' \mathrm{f} + \eta' \mathrm{f}_y + \eta' \mathrm{f}_x\right) \\
& - \frac{k_1^4 k_2}{8}\left(\eta'' \mathrm{f}_{yy} + 2\eta'' \mathrm{f}_y\right) - \frac{k_1^3 k_2^2}{6}\left(\eta'' \mathrm{f}_{yx} + \eta'' \mathrm{f}_y + \eta'' \mathrm{f}_x + \frac{\eta' \mathrm{f}_{yy}}{2}\right) \\
& + \frac{k_1^2 k_2^3}{12}\left(\eta' \mathrm{f}_{yx} + \eta' \mathrm{f}_{xx} + 2\eta'' \mathrm{f}_y + 2\eta'' \mathrm{f}_x\right) + \frac{k_1 k_1^4}{8}\left(\eta' \mathrm{f}_{xx} + \eta'' \mathrm{f}_x\right) + o\left(k^6\right)\Bigg] \\
& \div \left[ k_1 k_2 \mathrm{f} - \frac{k_1^2 k_2}{2}\mathrm{f}_y - \frac{k_1 k_2^2}{2}\mathrm{f}_x - \frac{1}{2!}\left(\frac{k_1^3 k_2}{3}\mathrm{f}_{yy} + \frac{2k_1^2 k_2^2}{4}\mathrm{f}_{yx} - \frac{k_1 k_2^3}{3}\mathrm{f}_{xx}\right)\right]
\end{aligned} \tag{37}
$$

***Lemma 3.*** If $\sigma_\eta^2(y,x)$ denotes the conditional variance of the function $\eta(t_1, t_2)$ in the interval $(y, x)$ such that

$$
\sigma_\eta^2(y,x) = \mu_\eta^2(y,x) - \left(\mu_\eta(y,x)\right)^2
$$

then

$$
\begin{aligned}
\sigma_\eta^2(y,x) = \Bigg\{ & \eta\eta' \mathrm{f}^2\left(k_1^3 k_2^2 + k_1^2 k_2^2\right) + \frac{4k_1^4 k_2^2 + 3k_1^3 k_2^3}{6}\Big[\eta\eta' \mathrm{f}\mathrm{f}_y + \eta\eta' \mathrm{f}\mathrm{f}_x \\
& + \mathrm{f}^2\left(\eta'^2 + \eta\eta'\right)\Big] - \frac{\eta' \mathrm{f}}{4}\left(k_1^4 k_2^2 + k_1^2 k_2^4 + k_1^3 k_2^3\right)\Bigg\} \\
& \div \left\{ k_1^2 k_2^2 \mathrm{f} + \frac{k_1^4 k_2^2}{4}\mathrm{f}_y^2 + \frac{k_1^2 k_2^4}{4}\mathrm{f}_x^2 + k_1^3 k_2^2 \mathrm{f}\mathrm{f}_y + \frac{k_1^3 k_2^3}{2}\mathrm{f}_y \mathrm{f}_x \right\}
\end{aligned} \tag{38}
$$

***Proof.*** This result can be established by using the expression forms of $\mu_\eta^2(y,x)$, replacing the function $\eta(t_1, t_2)$ by $\eta^2(t_1, t_2)$:

$$
\begin{aligned}
\mu_\eta^2(y,x) = \eta^2 + \Bigg[ & \eta\eta' \mathrm{f}\left(k_1^2 k_2 + k_1 k_2^2\right) + \frac{4k_1^3 k_2 + 3k_1^2 k_2^2}{6}\left(\eta\eta' \mathrm{f}_y + \eta\eta' \mathrm{f}_x + \mathrm{f}\left(\eta'^2 + \eta\eta'\right)\right) \\
& + \frac{k_1^4 k_2}{4}\left(\eta\eta' \mathrm{f}_{yy} + 2\,\mathrm{f}_y\left(\eta'^2 + \eta\eta''\right)\right) + \frac{k_1^3 k_2^2}{3}\Bigg(\eta\eta' \mathrm{f}_{yx} + \mathrm{f}_y\left(\eta'^2 + \eta\eta''\right) + \mathrm{f}_x\left(\eta'^2 + \eta\eta''\right) \\
& + \frac{\eta\eta' \mathrm{f}_{yy}}{2}\Bigg) + \frac{k_1^2 k_2^3}{6}\left(\eta\eta' \mathrm{f}_{yx} + \eta\eta' \mathrm{f}_{xx} + 2\,\mathrm{f}_y\left(\eta'^2 + \eta\eta''\right) + 2\,\mathrm{f}_x\left(\eta'^2 + \eta\eta'\right)\right) \\
& + \frac{k_1 k_1^4}{4}\left(\eta\eta' \mathrm{f}_{xx} + 2\,\mathrm{f}_x\left(\eta'^2 + \eta\eta'\right)\right) + \mathrm{O}\left(k^6\right)\Bigg] \\
& \div \left[ k_1 k_2\,\mathrm{f} + \frac{k_1^2 k_2}{(2)}\mathrm{f}_y + \frac{k_1 k_2^2}{(2)}\mathrm{f}_x - \frac{1}{2!}\left(\frac{k_1^3 k_2}{(3)}\mathrm{f}_{yy} + \frac{k_1 k_2^3}{(3)}\mathrm{f}_{xx} + \frac{2k_1^2 k_2^2}{(2)(2)}\mathrm{f}_{yx}\right)\right]
\end{aligned}
$$

where f and η are functions and their derivations are evaluated at the point $(t_1, t_2) = (y, x)$.

Similarly to equation (33),

$$
\mu_\eta^2(y,x) = \frac{\eta^2 + \left\{\frac{\eta'^2\,\mathrm{f}^2}{4}\left(k_1^4 k_2^2 + k_1^2 k_2^4 + 2k_1^3 k_2^3\right)\right\}}{k_1^2 k_2^2\,\mathrm{f}^2 + k_1^3 k_2^2\,\mathrm{ff}_y + k_1^2 k_2^3\,\mathrm{ff}_x + \frac{k_1^4 k_2^2}{3}\mathrm{ff}_{yy} + \frac{k_1^3 k_2^2}{2}\mathrm{f}_y \mathrm{f}_x + \frac{k_1^2 k_2^4}{3}\mathrm{ff}_{xx}} + \mathrm{O}\left(k^7\right)
$$

where f and η are functions and their derivations are evaluated at the point $(t_1, t_2) = (y, x)$.

After simplification,

$$
\begin{aligned}
\mu_\eta^2(y,x) = \Bigg[ & \eta\eta' \mathrm{f}^2\left(k_1^3 k_2^2 + k_1^2 k_2^2\right) + \frac{4k_1^4 k_2^2 + 3k_1^3 k_2^3}{6}\left[\eta\eta' \mathrm{ff}_y + \eta\eta' \mathrm{ff}_x + \mathrm{f}\left(\eta'^2 + \eta\eta''\right)\right] \\
& - \frac{\eta' \mathrm{f}}{4}\left(k_1^4 k_2^2 + k_1^2 k_2^4 + 2k_1^3 k_2^3\right)\Bigg] \\
& \div \left\{ k_1^2 k_2^2\,\mathrm{f} + \frac{k_1^4 k_2^2}{4}\mathrm{f}_y^2 + \frac{k_1^2 k_2^4}{4}\mathrm{f}_x^2 + k_1^3 k_2^2\,\mathrm{ff}_y + \frac{k_1^3 k_2^3}{2}\mathrm{f}_y \mathrm{f}_x \right\}
\end{aligned}
$$

***Lemma 4.***

$$\mu_\eta(y,x)=\frac{\sqrt{\eta(y)\eta(x)}\eta'\left(k_1^4k_2^2-k_1^2k_2^4\right)}{4k_1^2k_2^2\,\mathrm{f}-2k_1^3k_2\,\mathrm{f}_y}+\mathrm{O}\left(k^7\right) \tag{39}$$

***Proof.*** To prove the lemma, use the relation obtained in equations (35) and (36). On multiplying and taking square roots on both sides, obtain

$$\mu_\eta(y,x)=\sqrt{\eta(y)\eta(x)}\left[\frac{\frac{\eta'\,\mathrm{f}}{2}\left(k_1^2k_2-k_1k_2^2\right)}{k_1k_2\,\mathrm{f}+k_1^2k_2\,\mathrm{f}_y+\frac{k_1k_2^2\,\mathrm{f}_x}{2}}+\mathrm{O}\left(k^4\right)\right]^{\frac{1}{2}}$$
$$\times\left[\frac{\frac{\eta'\,\mathrm{f}}{2}\left(k_1^2k_2-k_1k_2^2\right)}{k_1k_2\,\mathrm{f}-k_1^2k_2\,\mathrm{f}_y-\frac{k_1k_2^2\,\mathrm{f}_x}{2}}+\mathrm{O}\left(k^4\right)\right]^{\frac{1}{2}}$$

Continuing the simplification,

$$\mu_\eta(y,x)=\frac{\sqrt{\eta(y)\eta(x)}\eta'\left(k_1^4k_2^2-k_1^2k_2^4\right)}{4k_1^2k_2^2\,\mathrm{f}-2k_1^3k_2\,\mathrm{f}_y}+\mathrm{O}\left(k^7\right)$$

Hence the lemma is proven.

***Lemma 5.***

$$\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2=\frac{1}{yx}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left(t_1t_2\right)^2\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2\left[1+\mathrm{O}\left(k^3\right)\right] \tag{40}$$

***Proof.*** Consider a function

$$\lambda(x)=\frac{1}{yx}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2-\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2$$

so that

$$\lambda(y)=\left(\frac{\partial\lambda(x)}{\partial x}\right)_{y=x}=\left(\frac{\partial^2\lambda(x)}{\partial x^2}\right)_{y=x}=0$$

Because there are the initial coefficients of $k_1$ or $k_2$ and $k_1^2$ or $k_2^2$ in the Taylor series expansion of $\lambda(x)$ about $n$, find

$$\lambda(z)=\mathrm{O}\left(k^3\right)$$

where $k$ is either $k_1$ or $k_2$.

$$\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2=\frac{1}{yx}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left(t_1t_2\right)^2\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2+\mathrm{O}\left(k^3\right)$$

$$\therefore\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2=\frac{1}{yx}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left(t_1t_2\right)^2\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2\left[1+\mathrm{O}\left(k^3\right)\right]$$

***Lemma 6.***

$$\left(k_1k_2\right)^{\lambda-1}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2=\left[\int_{x_1}^{x_2}\int_{y_1}^{y_2}\sqrt[\lambda]{\mathrm{f}\left(t_1,t_2\right)}\partial t_1\partial t_2\right]^{\lambda}\left[1+\mathrm{O}\left(k^2\right)\right] \tag{41}$$

***Proof.*** To prove the above lemma, expand the term

$$\left[\int_{x_1}^{x_2}\int_{y_1}^{y_2}\sqrt[\lambda]{\mathrm{f}\left(t_1,t_2\right)}\partial t_1\partial t_2\right]^{\lambda}$$

in powers of $k_1$ and $k_2$. Using Taylor's theorem and expanding $\sqrt[\lambda]{\mathrm{f}\left(t_1,t_2\right)}$ about the point $(t_1, t_2) = (y, x)$, obtain

$$\left[\int_{x_1}^{x_2}\int_{y_1}^{y_2}\sqrt[\lambda]{\mathrm{f}\left(t_1,t_2\right)}\partial t_1\partial t_2\right]^{\lambda}=\left[\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left(\sqrt[\lambda]{\mathrm{f}\left(t_1,t_2\right)}+\frac{\left(t_1-y_1\right)}{\lambda\,\mathrm{f}^{1-\frac{1}{2}}}\mathrm{f}_y+\frac{\left(t_2-x_1\right)}{\lambda\,\mathrm{f}^{1-\frac{1}{2}}}\mathrm{f}_x+\mathrm{O}\left(k^2\right)\right)\partial t_1\partial t_2\right]^{\lambda}$$

$$=\left[k_1k_2\sqrt[\lambda]{\mathrm{f}\left(t_1,t_2\right)}+\frac{k_1^2k_2^2}{4\lambda\,\mathrm{f}^{1-\frac{1}{\lambda}}}\mathrm{f}_y+\frac{k_1^2k_2^2}{4\lambda\,\mathrm{f}^{1-\frac{1}{\lambda}}}\mathrm{f}_x+\mathrm{O}\left(k^5\right)\right]^{\lambda}$$

$$=\left(k_1k_2\right)^{\lambda}\mathrm{f}\left(t_1,t_2\right)\left[1+\frac{k_1^2k_2^2}{4\lambda\,\mathrm{f}^{1-\frac{1}{\lambda}}}\left(\mathrm{f}_y-\mathrm{f}_x\right)+\mathrm{O}\left(k^3\right)\right]^{\lambda}$$

$$=\left(k_1k_2\right)^{\lambda-1}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2\left[1+\mathrm{O}\left(k^2\right)\right]$$

This may be rewritten as

$$\left(k_1k_2\right)^{\lambda-1}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\mathrm{f}\left(t_1,t_2\right)\partial t_1\partial t_2=\left[\int_{x_1}^{x_2}\int_{y_1}^{y_2}\sqrt[\lambda]{\mathrm{f}\left(t_1,t_2\right)}\partial t_1\partial t_2\right]^{\lambda}\left[1+\mathrm{O}\left(k^2\right)\right]$$

Using Lemmas 1-6,

$$\sum_{r=1}^{T}\sum_{s=1}^{U}W_{rs}^2\sigma_{rsy}^2=\mathrm{h}\left(y,x\right)\sigma_y^2r^{-3}s^{-1} \tag{42}$$

where

$$\mathrm{h}\left(y,x\right)=\frac{\mathrm{g}^3\left(y\right)\mathrm{g}\left(x\right)}{\sigma_x^2}G_1$$

$$G_1=\frac{1}{12}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{f}^2\left(y,x\right)\mathrm{f}^{-\frac{3}{2}}\left(y\right)\mathrm{f}^{-\frac{1}{2}}\left(x\right)\partial y\partial x$$

Similarly,

$$\sum_{r=1}^{T}\sum_{s=1}^{U}W_{rs}^2\sigma_{rsx}^2=\mathrm{h}\left(x,y\right)\sigma_x^2r^{-3}s^{-1} \tag{43}$$

Again using the approximation method discussed above,

$$\sum_{h=1}^{L}\sum_{k=1}^{M} W_{hk}^2 = \mathrm{g}(x,z)(rs)^{-1}\sum_{r=1}^{T}\sum_{s=1}^{U}\varphi_{rs}^2 (k_1 k_2)^2 \tag{44}$$

where $\mathrm{g}(y, x) = \mathrm{g}(y)\mathrm{g}(x)(rs)^{-1}G_2$,

$$G_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{f}^2(y,x)\left(\sqrt{\mathrm{f}}\right)^{-1}(y)\left(\sqrt{\mathrm{f}}\right)^{-1}(x)\,\partial y \partial x$$

From equations (42), (43), (44), and (30),

$$\sigma_{rsz}^2 = \sqrt{rs}n^{-1}\left\{\frac{1}{r^2}\left[b^2\,\mathrm{g}(y,x)\sigma_y^2 + c^2\,\mathrm{g}(x,y)\sigma_x^2\right] + \sigma^2\,\mathrm{g}(y,x)\right\}^{\frac{1}{2}} \tag{45}$$

Every pair ($y$, $x$) is made of factors that are stochastic. Thus the linear relationship between them can be obtained, from which the coefficient of regression will be a result.

## Numerical Illustration

The proposed strategy is appropriate when the likelihood thickness elements of the stratification factors are known. For example, let the auxiliary variable $Y$ follow the standard Cauchy distribution with density function

$$\mathrm{f}(y) = \frac{1}{\prod\left(1+y^2\right)}, -\infty < y < \infty \tag{46}$$

and the other auxiliary variable $X$ have the density function

$$\mathrm{f}(x) = \begin{cases} \dfrac{\delta x^{\delta-1}}{\theta^{\delta}}, & 0 \le x \le \theta \\ 0, & \text{otherwise} \end{cases} \tag{47}$$

where $\delta > 0$ and $\theta > 0$. In order to find the stratification points, find the value of (3) and (5),

$$W_{rs} = \frac{[I_1][I_2]}{\pi\theta^{\delta}} \tag{48}$$

$$\sigma_{rsy}^2 = u_s \left[ \frac{4\left(v_r - I_1\right) I_1^3 I_2^3 - \pi u_s \theta^{3\delta} \left( \frac{\log\left(1+\left(v_r + y_{r-1}\right)^2\right)}{1+y_{r-1}^2} \right)}{4\pi\theta^{\delta}\left(I_1 I_2\right)^2} \right] \tag{49}$$

and

$$\begin{aligned} \sigma_{rsx}^2 = \frac{\pi\delta}{(\delta+2)\left(I_1 I_2\right)^2} \Big[ & v_r \left( \left(u_s + x_{s-1}\right)^{\delta+2} - \left(x_{s-1}\right)^{\delta+2} \right)\left(I_1 I_2\right) \\ & - \pi\delta(\delta+2)\left[ \left(u_s + x_{s-1}\right)^{\delta+1} - \left(x_{s-1}\right)^{\delta+1} \right]^2 \Big] \end{aligned} \tag{50}$$

where $I_1 = \tan^{-1}(v_r - y_{r-1}) - \tan^{-1}(y_{r-1})$ and $I_2 = (u_s + x_{s-1})^{\delta} - (x_{s-1})^{\delta}$. By substituting values obtained in equations (48), (49), and (50) in (45), the optimum strata boundaries can be obtained.

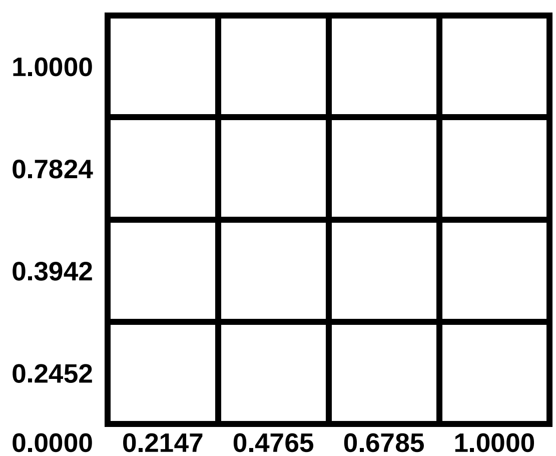


**Figure 1.** OSB for bi-variate auxiliary variables

**Table 1.** Strata boundaries and total variance

| OSB($y_r$, $x_s$) | Total Variance |
|---|---|
| (0.2147, 0.2452) | |
| (0.4765, 0.2452) | |
| (0.6785, 0.2452) | |
| (1.0000, 0.2452) | |
| (0.2147, 0.3942) | |
| (0.4765, 0.3942) | |
| (0.6785, 0.3942) | |
| (1.0000, 0.3942) | 0.05461 |
| (0.2147, 0.7824) | |
| (0.4765, 0.7824) | |
| (0.6785, 0.7824) | |
| (1.0000, 0.7824) | |
| (0.2147, 1.0000) | |
| (0.4765, 1.0000) | |
| (0.6785, 1.0000) | |
| (1.0000, 1.0000) | |

Consider a population of size 2,000, which is to be sub-isolated into 16 strata with $T = 4$ and $U = 4$, and an example of size 500 is to be taken from the population. Apply equations (48), (49), and (50) in condition (44) by using the underlying estimation of $Y = 0$ and $X = 0$ and by the maximum value of $Y = 1$ and $X = 1$ and $\theta = 1$, $\delta = 3$, respectively. Thus the total width of $Y$ and $X$ is 1 and 1, respectively.

The OSB so obtained can be displayed as above with corresponding total fluctuation in Table 1.

## Conclusion

Most of the time the complete set of data related to the study variable is unknown, which becomes a stumbling block to obtaining stratification points. However, in such situations using auxiliary variables has an increasing trend of precision. A strategy was proposed under the Neyman allocation when there is one investigation variable and two auxiliary variables based on the auxiliary variables. A numerical example demonstrated the diminishing pattern of the fluctuation when the quantity of strata is to be expanded. Along these lines, it can be concluded that this strategy for discovering stratification points can be recommended instead of existing techniques.

## References

Aoyama, H. (1954). A study of the stratified random sampling. *Annals of the Institute of Statistical Mathematics, 6*(1), 1-36. doi: 10.1007/bf02960514

Dalenius, T. (1950). The problem of optimum stratification. *Scandinavian Actuarial Journal, 33*(3-4), 203-213. doi: 10.1080/03461238.1950.10432042

Dalenius, T., & Gurney, M. (1951).The problem of optimum stratification. II. *Scandinavian Actuarial Journal, 34*(1-2), 133-148. doi: 10.1080/03461238.1951.10432134

Dalenius, T., & Hodges, J. L., Jr. (1959). Minimum variance stratification. *Journal of the American Statistical Association, 54*(285), 88-101. doi: 10.1080/01621459.1959.10501501

Durbin, J. (1959). Review of *Sampling in Sweden*. *Journal of the Royal Statistical Society. Series A (General), 122*(2), 246-248. doi: 10.2307/2342611

Gunning, P., & Horgan, J. M. (2004). A new algorithm for the construction of stratum boundaries in skewed populations. *Survey Methodology, 30*(2), 159-166. Retrieved from http://www.statcan.gc.ca/pub/12-001-x/2004002/article/7749-eng.pdf

Mahalanobis, P. C. (1952). Some aspects of the design of sample surveys. *Sankhyā: The Indian Journal of Statistics, 12*(1/2), 1-17. Available from http://www.jstor.org/stable/25048108

Rivest, L.-P. (2002). A generalization of the Lavallée and Hidiroglou algorithm for stratification in business surveys. *Survey Methodology, 28*(2), 191-198. Retrieved from http://www.statcan.gc.ca/pub/12-001-x/2002002/article/6432-eng.pdf

Rizvi, S. E. H., Gupta, J. P., & Singh, R. (2002). Approximately optimum stratification for two study variable using auxiliary information. *Journal of the Indian Society of Agricultural Statistics, 53*(3), 287-298.

Schneeberger, H., & Goller, W. (1979). On the problem of feasibility of optimal stratification points according to Dalenius. *Statistische Hefte, 20*(4), 250-256. doi: 10.1007/bf02932794

Serfling, R. J. (1968). Approximately optimal stratification. *Journal of the American Statistical Association, 63*(324), 1298-1309. doi: 10.1080/01621459.1968.10480928

Singh, R. (1971). Approximately optimum stratification on the auxiliary variable. *Journal of the American Statistical Association*, *66*(336), 829-833. doi: 10.1080/01621459.1971.10482352

Singh, R. (1975). An alternative method of stratification on the auxiliary variable. *Sankhyā: The Indian Journal of Statistics*, *Series C*, *37*(1), 100-108.

Singh, R. (1977). A note on optimum stratification for equal allocation with ratio and regression methods of estimation. *Australian & New Zealand Journal of Statistics*, *19*(2), 96-104. doi: 10.1111/j.1467-842x.1977.tb01275.x

Singh, R., & Parkash, D. (1975). Optimum stratification for equal allocation. *Annals of the Institute of Statistical Mathematics*, *27*(1), 273-280. doi: 10.1007/bf02504646

Singh, R., & Sukhatme, B. V. (1969). Optimum stratification. *Annals of the Institute of Statistical Mathematics*, *21*(3), 515-528. doi: 10.1007/bf02532275

Singh, R., & Sukhatme, B. V. (1972). Optimum stratification in sampling with varying probabilities. *Annals of the Institute of Statistical Mathematics*, *24*(1), 485-494. doi: 10.1007/bf02479777

Singh, R., & Sukhatme, B. V. (1973). Optimum stratification with ratio and regression methods of estimation. *Annals of the Institute of Statistical Mathematics*, *25*(1), 627-633. doi: 10.1007/bf02479404

Taga, Y. (1967). On optimum stratification for the objective variable based on concomitant variables using prior information. *Annals of the Institute of Statistical Mathematics*, *19*, 101-129. doi: 10.1007/bf02911670

Verma, M. R. (2008). Approximately optimum stratification for ratio and regression methods of estimation. *Applied Mathematics Letters*, *21*(2), 200-207. doi: 10.1016/j.aml.2007.03.012