

On The Effect of Stratification When Two Independent Stratifying Variables Are Used

Dr. Al-Kassab, MMT*

Dr. Al-Elemat, A.R. **

Abstract

In this paper, we study the effect of using two independent stratifying variables on the approximate value of the variance of the stratified sample mean obtained by Al-Kassab and Al-Elemat [1]. Ten special cases are obtained and a comparison between these special cases is given using four probability distributions.

تأثير الطبقية عند استخدام متغيرين مستقلين لهذه الطبقية

الملخص

تم في هذا البحث دراسة تأثير متغيرين مستقلين للطبقية على تباين الوسط الحسابي الطبقي المقدم من قبل كل من القصاب و العليمات تم ايجاد عشرة حالات خاصة لهذا التباين و مقارنة هذه الحالات اعتمادا على اربعة توزيعات احتمالية

1. Introduction

Most of the literature on stratified survey sampling deals with a single stratifying variable, but the work when two stratifying variables has recently been studied. The pioneering work in this field was done by Thomsen (1977) [4], where he found an approximation for the variance of the stratified sample mean, called $cum f^{1/2}$.

* Prof / Dept. Statistics and Informatics / Computer Sciences and Mathematics College / Mosul University

** Lecturer / Dept. of Mathematics / Al Al-bayt University

Received Date 22/9/2013 ————— Accept Date 21/11/2013

In 1999 Al-Kassab and Al-Hasso [3] suggested a new approximation method for bivariate stratification, called $\text{cum } f^{1/3}$, where they found an approximation for the variance of the stratified sample mean. In 2000 Al-Kassab [2] suggested two new approximation methods for bivariate stratification, called $\text{cum } f^{3/5}$ and $\text{cum } f^{16/25}$, where he found an approximation for the variance of the stratified sample mean, and concluded that the $\text{cum } f^{16/25}$ is better than the $\text{cum } f^{3/5}$. They presented approximations for the variance of the study variable under the assumption of a linear regression on the two stratifying variables. The results indicate that, in many practical situations, the gain from using two stratifying variables over one is non-trivial. In 2011 Al-kassab and Al-Elamat [1] suggested a new approximate method for bivariate stratification, namely $\text{cum } f^{a^m/(a^m+1)}$, which is a generalization of the methods suggested by Thomsen [4], Al-kassab and Al-Hasso [3], and Al-kassab [2]. This approximation of the variance depends only on n (sample size), r, s (number of strata), the correlation coefficients between the study variable Y and the stratifying variables X and Z , and the simultaneous density of X and Z .

In this paper, we studied the formula obtained by Al-Kassab and Al-Elamat [1] to obtain the variance of the stratified sample mean when two stratifying variables are used. The study concentrated on ten cases with special distributions for the stratifying variables. In section two, we give the problem of the study variable Y on the two stratifying variables X and Z , and the variance of the stratified sample mean obtained by Al-Kassab and Al-Elamat [1]. In section three, we obtained the variance of the stratified sample mean for ten cases. In section four, the theoretical distributions are proposed for the stratifying variables X and Z . Section five contains the conclusions.

2. The Problem

We assume that the population values of the two stratifying variables, X and Z , are generated from a background bivariate distribution with a joint probability density function f and marginal probability density functions f_1 and f_2 , respectively. The population values of the study variable Y are also assumed to be realization of a stochastic background variable. The regression of this variable (Y) on the stratification variables is assumed to be bivariate linear, defined by

$$Y_{ijh} = c + \alpha X_{ijh} + \beta Z_{ijh} + e_{ijh},$$

where $E(e_{ijh}) = 0$, $\text{var}(e_{ijh}) = \sigma^2$, $i = 1, 2, \dots, r$ (r number of strata on X), $j = 1, 2, \dots, s$ (s number of strata on Z), and $h = 1, 2, \dots, N_{ij}$ (N_{ij} number of units in stratum (i, j)). The population is stratified into rs strata, where r strata are constructed along X , and s strata are constructed along Z . An element in the population belongs to stratum (i, j) if its X -value belongs to the i -th stratum along X , and its Z -value belongs to the j -th stratum along Z . The sample size is allocated with $n^* = n/rs$ observations in each stratum. It is assumed that there are N_{ij} units with Y -values Y_{ijh} ($h = 1, 2, \dots, N_{ij}$) in the stratum (i, j) . [1]

The population mean of the (i, j) th stratum is $\mu_{ij} = \frac{1}{N_{ij}} \sum_{h=1}^{N_{ij}} Y_{ijh}$,

and the population variance of the (i, j) th stratum is

$$\sigma_{ij}^2(Y) = \frac{1}{N_{ij}} \sum_{h=1}^{N_{ij}} (Y_{ijh} - \mu_{ij})^2.$$

We assume that $N_{ij} \geq n^*$ for all values i and j . The overall population mean is $\mu = \sum_{i=1}^r \sum_{j=1}^s w_{ij} \mu_{ij}$, while the overall population variance is

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^r \sum_{j=1}^s \sum_{h=1}^{N_{ij}} (Y_{ijh} - \mu)^2, \text{ where } N = \sum_{i=1}^r \sum_{j=1}^s N_{ij}.$$

Let y_{ijh} denote the h -th value of Y at the stratum (i, j) , then the sample mean in the stratum (i, j) is $\bar{y}_{ij} = \frac{1}{n^*} \sum_{h=1}^{N_{ij}} y_{ijh}$, and $\bar{y}_{st} = \sum_{i=1}^r \sum_{j=1}^s \frac{N_{ij}}{N} \bar{y}_{ij}$ is an unbiased estimator of the overall population mean μ , with

$$V(\bar{y}_{st}) = \sum_{i=1}^r \sum_{j=1}^s w_{ij}^2 V(\bar{y}_{ij}).$$

And the problem is to minimize $V(\bar{y}_{st})$.

Corollary (see[1]) The minimum variance of the stratified sample mean, depending on the $\text{cum} f \frac{2^m}{2^m+1}$, is given by

$$V(\bar{y}_{st}) = \left\{ V(\bar{y}) \left[M(X, Z) \frac{(r_{xy} - r_{yz} r_{xz})^2}{(1 - r_{xz}^2)^2 r^2} + M(Z, X) \frac{(r_{zy} - r_{xz} r_{xy})^2}{(1 - r_{xz}^2)^2 s^2} + N(X, Z) (1 - R_{y \cdot xz}^2) \right] \right\} \quad (1)$$

$$\left[V(\bar{y}) [N(X, Z) (1 - R_{y \cdot xz}^2)] \right], \text{ for large } r \text{ and } s$$

And if X and Z are independent,

On The Effect of Stratification When Two Independent

$$V(\bar{y}_{st}) \propto \left\{ \begin{array}{l} V(\bar{y}) \left[k_z^* k_x \frac{r_{xy}^2}{r^2} + k_x^* k_z \frac{r_{zy}^2}{s^2} + k_x^* k_z^* (1 - R_{y,xz}^2) \right] \\ V(\bar{y}) [k_x^* k_z^* (1 - R_{y,xz}^2)] , \text{ for large } r \text{ and } s \end{array} \right\} \quad (2)$$

Where,

$$K(X) = \int_{-\infty}^{\infty} f_1^{\frac{a^m}{a^m+1}}(x) dx, \quad K(Z) = \int_{-\infty}^{\infty} f_2^{\frac{a^m}{a^m+1}}(z) dz, \quad a \in \mathbb{R}, \quad m \in \mathbb{R},$$

$$M(X, Z) = \frac{K^3(X)K(Z)}{12\sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2(x, z) f_1^{\frac{-3a^m}{a^m+1}}(x) f_2^{\frac{-a^m}{a^m+1}}(z) dx dz,$$

$$N(X, Z) = N(Z, X) = K(X)K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2(x, z) f_1^{\frac{-a^m}{a^m+1}}(x) f_2^{\frac{-a^m}{a^m+1}}(z) dx dz,$$

$$K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{\frac{a^m+2}{a^m+1}} dx, \quad k^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{\frac{-a^m+2}{a^m+1}} dx,$$

$$k_x = \frac{K^3(X)k^*(X)}{12\sigma^2(X)}, \text{ and } k_x^* = K(X)K^*(X).$$

3. Special Cases

In this section we give different values for a , and m , to calculate $\text{cum } f^{\frac{a^m}{a^m+1}}$ obtained by Al-Kassab and Al-Elamat [1]. Ten cases where obtained as a special cases of the above formula by assuming different values for a and m . Four of these special cases are obtained by other authors (Thomsen [4], Al-Kassab and Al-Hasso [3], and Al-Kassab [2]).

3.1 Assuming $a=1/3$ and $m=1$

The method becomes $\text{cum } f^{1/4}$. Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{1/4} dx, K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{1/4} dz, K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{7/4} dx, \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{5/4} dx, k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, k_x^* = K(X) K^*(X), \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-3/4} [f_2(z)]^{-1/4} dx dz, \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-1/4} [f_2(z)]^{-1/4} dx dz. \end{aligned} \right\} (3)$$

3.2 Assuming a=1/2 and m=1

The method becomes *cum f^{1/3}* (Al-kassab and Al-Hasso method [3]). Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{1/3} dx, K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{1/3} dz, K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{5/3} dx, \\ k^*(X) &= \int_{-\infty}^{\infty} f_1(x) dx, k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, k_x^* = K(X) K^*(X), \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-1} [f_2(z)]^{-1/3} dx dz, \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-1/3} [f_2(z)]^{-1/3} dx dz. \end{aligned} \right\} (4)$$

3.3 Assuming a=2/3 and m=1

The method becomes *cum f^{2/5}*. Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{2/5} dx, K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{2/5} dz, K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{8/5} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{4/5} dx, k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, k_x^* = K(X) K^*(X) \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-6/5} [f_2(z)]^{-2/5} dx dz \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-2/5} [f_2(z)]^{-2/5} dx dz \end{aligned} \right\} (5)$$

On The Effect of Stratification When Two Independent

3.4 Assuming $a=12/13$ and $m=1$

The method becomes $\text{cum } f^{12/25}$. Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{12/25} dx, \quad K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{12/25} dz, \quad K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{38/25} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{14/25} dx, \quad k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, \quad k_x^* = K(X) K^*(X) \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-36/25} [f_2(z)]^{-12/25} dx dz \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-12/25} [f_2(z)]^{-12/25} dx dz \end{aligned} \right\} (6)$$

3.5 Assuming $a=1$ and $m=1$

The method becomes $\text{cum } f^{1/2}$ (Thomsen method [4]). Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{1/2} dx, \quad K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{1/2} dz, \quad K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{3/2} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{1/2} dx, \quad k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, \quad k_x^* = K(X) K^*(X) \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-3/2} [f_2(z)]^{-1/2} dx dz \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-1/2} [f_2(z)]^{-1/2} dx dz \end{aligned} \right\} (7)$$

3.6 Assuming $a=13/12$ and $m=1$

The method becomes $\text{cum } f^{13/25}$. Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{13/25} dx, \quad K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{13/25} dz, \quad K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{37/25} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{11/25} dx, \quad k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, \quad k_x^* = K(X) K^*(X) \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-39/25} [f_2(z)]^{-13/25} dx dz \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-13/25} [f_2(z)]^{-13/25} dx dz \end{aligned} \right\} (8)$$

3.7 Assuming a=4/3 and m=1

The method becomes *cum f^{4/7}*. Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{4/7} dx, \quad K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{4/7} dz, \quad K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{10/7} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{2/7} dx, \quad k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, \quad k_x^* = K(X) K^*(X) \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-12/7} [f_2(z)]^{-4/7} dx dz \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-4/7} [f_2(z)]^{-4/7} dx dz \end{aligned} \right\} (9)$$

3.8 Assuming a=3/2 and m=1

The method becomes *cum f^{3/5}* (Al-kassab method [2]). Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{3/5} dx, \quad K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{3/5} dz, \quad K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{7/5} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{1/5} dx, \quad k_x = \frac{K^3(X) k^*(X)}{12 \sigma^2(X)}, \quad k_x^* = K(X) K^*(X) \\ M(X, Z) &= \frac{K^3(X) K(Z)}{12 \sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-9/5} [f_2(z)]^{-3/5} dx dz \\ N(X, Z) &= K(X) K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-3/5} [f_2(z)]^{-3/5} dx dz \end{aligned} \right\} (10)$$

3.9 Assuming $a=4/3$ and $m=2$

The method becomes $\text{cum } f^{16/25}$ (Al-kassab method [2]). Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{16/25} dx, \quad K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{16/25} dz, \quad K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{34/25} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{2/25} dx, \quad k_x = \frac{K^3(X)k^*(X)}{12\sigma^2(X)}, \quad k_x^* = K(X)K^*(X) \\ M(X, Z) &= \frac{K^3(X)K(Z)}{12\sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-48/25} [f_2(z)]^{-16/25} dx dz \\ N(X, Z) &= K(X)K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-16/25} [f_2(z)]^{-16/25} dx dz \end{aligned} \right\} (11)$$

3.10 Assuming $a=13/7$ and $m=1$

The method becomes $\text{cum } f^{13/20}$. Depending on this method the approximate value for the variance of the stratified sample mean is given in equations (1) and (2), where

$$\left. \begin{aligned} K(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{13/20} dx, \quad K(Z) = \int_{-\infty}^{\infty} [f_2(z)]^{13/20} dz, \quad K^*(X) = \int_{-\infty}^{\infty} [f_1(x)]^{27/20} dx \\ k^*(X) &= \int_{-\infty}^{\infty} [f_1(x)]^{1/20} dx, \quad k_x = \frac{K^3(X)k^*(X)}{12\sigma^2(X)}, \quad k_x^* = K(X)K^*(X) \\ M(X, Z) &= \frac{K^3(X)K(Z)}{12\sigma^2(X)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-39/20} [f_2(z)]^{-13/20} dx dz \\ N(X, Z) &= K(X)K(Z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, z)]^2 [f_1(x)]^{-13/20} [f_2(z)]^{-13/20} dx dz \end{aligned} \right\} (12)$$

4. Some probability distributions

To illustrate the suggested methods, we present a comparison between these methods using equation (3), the comparison is done using four probability distributions, namely Normal $[N(0,1)]$, Exponential $[Exp(1)]$, Gamma $[\Gamma(2,1)]$, and Uniform distribution $[U(0,1)]$, we assume that the stratifying variables X and Z belong to these distributions, and that the

correlation coefficients between X and Y , and Z and Y , are $r_{xy} = 0.85$, $r_{zy} = 0.50$ respectively.

4.1 X and Z have Normal distribution

We apply the above ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 1 below.

Table 1: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, and the distributions of X and Z are both Normal.					
$r \backslash s$	2	3	4	5	∞
2	0.6191	0.5358	0.5066	0.4932	0.4692
3	0.3783	0.2950	0.2659	0.2524	0.2284
4	0.2941	0.2107	0.1816	0.1681	0.1441
5	0.2551	0.1717	0.1426	0.1291	0.1051
∞	0.1857	0.1024	0.0732	0.0597	0.0357

Table 1 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent and Normally distributed, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $\text{cum } f^{13/25}$, case 6). In fact it is the best case among the ten cases.

4.2 X and Z have Exponential distribution

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 2 below.

Table 2: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, and the distributions of X and Z are both Exponential .

$r \backslash s$	2	3	4	5	∞
2	0.472	0.4112	0.3899	0.3801	0.3626
3	0.2963	0.2356	0.2143	0.2044	0.1869
4	0.2349	0.1741	0.1528	0.143	0.1255
5	0.2064	0.1456	0.1244	0.1145	0.097
∞	0.1558	0.095	0.0738	0.0639	0.0464

Table 2 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent and Exponentially distributed, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $\text{cum } f^{13/25}$, case 6) . In fact it is the best case among the ten cases.

4.3 X and Z have Gamma distribution

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 3 below.

Table 3: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, and the distributions of X and Z are both Gamma.

$r \backslash s$	2	3	4	5	∞
2	0.5177	0.4493	0.4253	0.4143	0.3946
3	0.3200	0.2516	0.2277	0.2166	0.1969
4	0.2509	0.1825	0.1585	0.1475	0.1278
5	0.2188	0.1505	0.1265	0.1154	0.0958
∞	0.1619	0.0935	0.0696	0.0585	0.0388

Table 3 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent and Gamma distributed, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $\text{cum } f^{13/25}$, case 6) . In fact it is the best case among the ten cases.

4.4 X and Z have Uniform distribution

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 4 below.

Table 4: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, and the distributions of X and Z are both Uniform

$r \backslash s$	2	3	4	5	∞
2	0.2706	0.2359	0.2238	0.2181	0.2081
3	0.1703	0.1356	0.1234	0.1178	0.1078
4	0.1352	0.1004	0.0883	0.0827	0.0727
5	0.1189	0.0842	0.0720	0.0664	0.0564
∞	0.0900	0.0553	0.0431	0.0375	0.0275

On The Effect of Stratification When Two Independent

Table 4 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent and Uniform distributed, with $r_{xy} = 0.85$, $r_{zy} = 0.50$. We conclude that all the ten cases give the same results.

4.5 The distribution of X is Normal, and the distribution of Z is Exponential

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 5 below.

Table 5: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, the distribution of X is Normal, and the distribution of Z is Exponential.					
$r \backslash s$	2	3	4	5	∞
2	0.6308	0.5774	0.5588	0.5501	0.5348
3	0.3563	0.3030	0.2843	0.2757	0.2603
4	0.2602	0.2069	0.1882	0.1796	0.1642
5	0.2157	0.1624	0.1438	0.1351	0.1198
∞	0.1367	0.0834	0.0647	0.0561	0.0407

Table 5 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent, the distribution of X is Normal, and the distribution of Z is Exponential, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $cum f^{13/25}$, case 6). In fact it is the best case among the ten cases.

4.6 The distribution of X is Normal, and the distribution of Z is gamma

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 6 below.

Table 6: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, the distribution of X is Normal, and the distribution of Z is Gamma.					
$r \backslash s$	2	3	4	5	∞
2	0.5348	0.3453	0.2789	0.2482	0.1936
3	0.448	0.2584	0.192	0.1613	0.1067
4	0.4176	0.228	0.1616	0.1309	0.0763
5	0.4035	0.2139	0.1476	0.1169	0.0623
∞	0.3785	0.1889	0.1226	0.0918	0.0373

Table 6 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent, the distribution of X is Normal, and the distribution of Z is Gamma, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $\text{cum } f^{13/25}$, case 6). In fact it is the best case among the ten cases.

4.7 The distribution of X is Normal, and the distribution of Z is Uniform

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 7 below.

On The Effect of Stratification When Two Independent

Table 7: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, the distribution of X is Normal, and the distribution of Z is Uniform.

$r \backslash s$	2	3	4	5	∞
2	0.4828	0.4432	0.4294	0.4230	0.4116
3	0.2716	0.2320	0.2181	0.2117	0.2003
4	0.1976	0.1581	0.1442	0.1378	0.1264
5	0.1634	0.1238	0.1100	0.1036	0.0922
∞	0.1026	0.0630	0.0492	0.0427	0.0313

Table 7 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent, the distribution of X is Normal, and the distribution of Z is Uniform, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $\text{cum } f^{13/25}$, case 6) . In fact it is the best case among the ten cases.

4.8 The distribution of X is Exponential, and the distribution of Z is gamma

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 8 below.

Table 8: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, the distribution of X is Exponential, and the distribution of Z is Gamma.

$r \backslash s$	2	3	4	5	∞
2	0.5315	0.3154	0.2397	0.2047	0.1425
3	0.4759	0.2598	0.1842	0.1492	0.0869
4	0.4564	0.2404	0.1647	0.1297	0.0675
5	0.4474	0.2313	0.1557	0.1207	0.0585
∞	0.4314	0.2153	0.1397	0.1047	0.0425

Table 8 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent, the distribution of X is Exponential, and the distribution of Z is Gamma, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $\text{cum } f^{13/25}$, case 6) . In fact it is the best case among the ten cases.

4.9 The distribution of X is Gamma, and the distribution of Z is Uniform

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 9 below.

Table 9: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, the distribution of X Gamma , and the distribution of Z is Uniform .					
$r \backslash s$	2	3	4	5	∞
2	0.4063	0.3650	0.3506	0.3439	0.3320
3	0.2400	0.1987	0.1843	0.1776	0.1657
4	0.1818	0.1405	0.1261	0.1194	0.1075
5	0.1548	0.1136	0.0991	0.0925	0.0806
∞	0.1070	0.0657	0.0512	0.0446	0.0327

Table 9 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent, the distribution of X is Gamma, and the distribution of Z is Uniform, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $\text{cum } f^{13/25}$, case 6) . In fact it is the best case among the ten cases.

4.10 The distribution of X is Exponential, and the distribution of Z is Uniform

We apply the ten cases for obtaining the variance ratio between the variance of the stratified sample mean and that of the simple random sample mean defined in equation (3). Amongst the ten cases, the best one appeared in table 10 below.

Table 10: Ratio of variances, when $r_{xy} = 0.85$, $r_{zy} = 0.50$, the distribution of X is Exponential, and the distribution of Z is Uniform.					
$r \backslash s$	2	3	4	5	∞
2	0.3602	0.3151	0.2993	0.2920	0.2790
3	0.2251	0.1800	0.1642	0.1569	0.1439
4	0.1778	0.1327	0.1169	0.1096	0.0966
5	0.1559	0.1108	0.0950	0.0877	0.0747
∞	0.1169	0.0718	0.0560	0.0487	0.0357

Table 10 above gives the least values of the ratio between the variance of the stratified sample mean and that of the simple random sample mean, when the stratifying variables are independent, the distribution of X is Exponential, and the distribution of Z is Uniform, with $r_{xy} = 0.85$, $r_{zy} = 0.50$, which calculated by equations (8) (i.e., $cum f^{13/25}$, case 6). In fact it is the best case among the ten cases.

5 Conclusions

1. We conclude that all the ten cases give the same results when the stratifying variables have uniform distribution.
2. We conclude that the suggested method $cum f^{13/25}$ (case 6) is the best method amongst all methods.
3. Bivariate stratification is almost always more efficient than univariate stratification.

References

- [1] Al-Kassab, MMT., and Al-Eleamat, A.R., **On The Effect Of Stratification When Two Stratifying Variables Are Used (A New Method)**, accepted in "Third Conference on Mathematical Sciences(CMS'2011), Zarqa University, April 27-29, 2011, Jordan".
- [2] Al-Kassab, MMT., (2000), **Two New Suggested Approximate Methods Useful in Bivariate Stratification**, Journal of Tanmiat Al-Rafidain, Vol. 181, P. 155-167.
- [3] Al-Kassab, MMT., and Al-Hasso, A., (1999), **Use of Bivariate Regression Model in Stratified Sampling**, Journal of Tanmiat Al-Rafidain, Vol. 57, P. 143-151.
- [4] Thomsen, I. B., (1977), **On the effect of stratification when two stratifying variables are used**, Journal of the American Statistical Association, 72(357), 149-153.