AN ABSTRACT OF THE THESIS OF

SPENCER T. CHESTER JR. for the degree of DOCTOR OF PHILOSOPHY

in STATISTICS presented on November 27, 1979

Title: THE USE OF AN AUXILIARY VARIABLE IN OPTIMUM STRATIFICATION

Abstract approved: Redacted for Privacy

G. David Faulkenberry

When stratification is employed to reduce the variance of an estimate of the population mean, and the population does not possess an a priori partitioning, the degree of reduction provided is affected by the method of allocation, by the number of strata, and by the choice of the stratum boundaries. However, the most effective use of either the method of allocation or the choice of the number of strata is attained when one employs the points of optimum stratification. It was shown by Dalenius (1950), when the stratification and study variables are identical, and by Dalenius and Gurney (1951), when the stratification is on an auxiliary variable related to the study variable by a known regression function, that the optimum stratification boundaries were theoretically obtainable as the solution of a system of non-linear equations.

In practice, when nothing is known about the relationship between the study and auxiliary variables, Dalenius's original equations are used with the auxiliary variable as an approximation to the points of

optimum stratification. Over the range of variable populations and regression functions commonly found in practice, an empirical study is made of the effect this approximation has on the reduction in variance compared with the use of simple random sampling. The effect these approximate boundary points have on the choice of the method of allocation is also considered.

When something is known about the character of the regression function between the two variables. the cumbersome nature of the Dalenius-Gurney equations in practice suggests determining the stratum boundaries by employing one of a series of approximations proposed in the literature. An empirical study is made of the performance of these approximations, comparing the approximate boundaries with the optimum stratification boundaries.

Finally, once the method of allocation and the mechanism for computing the stratum boundaries are chosen, the number of strata to use must be determined. An algorithm is introduced that quantifies the effect the increase in the number of strata has on the reduction in variance. With this algorithm, an empirical study is made of the approximations that have appeared in the literature designed to model the relationship between the number of strata and the precision of the estimate of the population mean.

# The Use of an Auxiliary Variable in Optimum Stratification

by

Spencer T. Chester Jr.

A THESIS

submitted to

Oregon State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

June 1980

APPROVED:

Redacted for Privacy

Associate Professor of Statistics
in charge of major

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented November 27, 1979

Typed by Karen Bland for Spencer T. Chester Jr.

This dissertation is dedicated to the memory of

Barbara T. Chester, May S. Chester and Mary K. Richardson

The author wishes to thank Dr. G. David Faulkenberry for having enough faith in him to agree to act as his major professor no less than three times. The author also wishes to acknowledge the help of Dr. David S. Birkes, whose door was always open, and whose patient assistance has proved invaluable throughout the author's academic career.

The author acknowledges the financial support provided in part by a N.I.H. Biometry Traineeship, as well as an unsponsored research grant from the O.S.U. Computer Center.

## TABLE OF CONTENTS

TABLE OF CONTENTS (cont)

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (cont)

LIST OF TABLES (cont)

LIST OF TABLES (cont)

LIST OF TABLES (cont)

# The Use of an Auxiliary Variable in Optimum Stratification

## I. REVIEW OF LITERATURE AND FORMULATION OF PROBLEMS

Suppose that we have a finite population of elements and wish to estimate the population mean. It is well known that, when an auxiliary variable is available, proper use of stratified random sampling reduces the variance of the estimate of the population mean compared to the variance of the estimate obtained from simple random sampling. For example, if one has available a census of last year's crop yields for the farms in a region, this information could be used as the auxiliary variable when one wished to estimate the average crop yield for the current year. The choice of the stratification variable(s), the size of the sample, the method of allocating the sample across the strata, the number of strata, and, when the strata are not suggested *a priori*, the selection of the stratum boundaries, all influence this reduction.

The particular situation of interest in this thesis is that in which only one stratification variable is used. Additionally, we will assume that the cost of sampling the individual units does not vary between strata, nor with the number of strata employed, making it a function of only the sample size. Thus, with the sample size considered fixed, the cost does not enter into the optimum stratification process. Thus, the problem considered here will be, for a fixed sample size, to choose the number of strata, the boundaries of the strata, and the method of allocation to minimize the variance of the stratified random sampling estimate of the population mean.

Of these three variable factors, when the stratum boundaries are not suggested a priori, the most important of them is the choice of the boundary points. Though depending upon the other two factors, the boundary selection can affect both the choice of the method of allocation as well as the efficient number of strata to use. For instance, when the boundaries are carefully chosen, use of Neyman allocation provides both an improvement over the variance of the sample mean under simple random sampling, and a better reduction than that offered by proportional allocation. Yet, when the boundaries are not carefully chosen, Neyman allocation can do worse than both proportional allocation (with carefully chosen boundaries) and simple random sampling. When determining the efficient number of strata to employ, it generally takes only a few strata to exhaust the practically achievable reduction in variation with the boundaries well chosen; when the boundaries are ill chosen, even doubling that number of strata does not guarantee the attainment of the same degree of reduction.

The theory associated with the optimal choice of the stratum boundaries was first developed assuming that the stratification and the study variables were identical. Subsequent theory considered the more practical situation of stratifying on an auxiliary variable related to the study variable by a known regression function. However, virtually all practical work reported in the literature has assumed either the equivalence of the auxiliary and study variables, or that the auxiliary variable was highly correlated with the study variable. The theory provides well defined boundaries in both cases. However, one does not always have the convenience of a highly correlated auxiliary variable.

The primary direction of this thesis will be to consider the practical use of the auxiliary variable in the selection of the stratum boundaries as the magnitude of the relationship between the study and auxiliary variables ranges from small to large. Additionally, we will examine how the relationship between the variables affects the choice of the other two factors.

In the first section, we introduce the particular questions to be considered. In the second and third sections, previous studies in the literature dealing with these problems are discussed, and the model framework we will employ in this thesis is introduced.

## 1.1 The Problems

### 1.1.1. Notation and Stratified Random Sampling

Suppose that the population under study is divided into L strata and a stratified random sample of size n is drawn, i.e., a simple random sample of size $n_h$ is drawn from the h-th stratum, h=1,2,...,L, such that $\Sigma_{h=1}^{L} n_h = n$. Writing Y as the study variable, let $\mu_{hy}$ and $\sigma_{hy}^2$ denote its mean and variance within the h-th stratum. Further, suppose that $W_h$ is defined as the h-th stratum weight, the proportion of the population falling in the h-th stratum. Then the population mean is defined by $\mu = \mu_y = \Sigma_{h=1}^{L} W_h \mu_{hy}$. If $\bar{y}_h$ denotes the sample mean from the h-th stratum, then an unbiased estimate of $\mu$ is given by

$$\hat{\mu} = \Sigma_{h=1}^{L} W_h \bar{y}_h .$$

If one ignores the finite population correction factor for each stratum, the variance of $\hat{\mu}$ is

$$\mathrm{var}[\hat{\mu}] = \Sigma_{h=1}^{L} \frac{W_h^2}{n_h} \sigma_{hy}^2 .$$

When the stratum boundaries are suggested a priori, that is, the population possesses naturally occurring partitionings (e.g. geographic), $\mathrm{var}[\hat{\mu}]$ for a given number of strata is a function of the method of allocation, whose choice determines the values of $n_h$. When the population does not possess a natural partitioning, $\mathrm{var}[\hat{\mu}]$ is also a function of the location of the location of the stratum demarcation points, whose selection specifies the values of $W_h$ and $\sigma_{hy}^2$.

The theory underlying the choice of the method of allocation is well known and can be found in most sampling texts (e.g. Cochran (1977) or Kish (1965)). Which method of allocation to use is generally determined by factors outside the scope of the theory. Perhaps the three most commonly used methods of allocation employed in practice are: proportional allocation ($n_h \propto W_h$), because of the resulting self-weighting nature of the estimate $\hat{\mu}$; Neyman or optimal allocation ($n_h \propto W_h \sigma_{hy}$), because it affords the greatest decrease in the variance of the estimate for fixed cost and sample size; and equal allocation ($n_h \propto 1/L$), both because of its potential administrative convenience as well as its approximation to Neyman allocation (see Cochran (1961), §6). Under these methods of allocation, $\mathrm{var}[\hat{\mu}]$ can be reduced to

$$\text{var}[\hat{\mu}|\text{prop}] = \frac{1}{n}\Sigma_{h=1}^{L} W_h \sigma_{hy}^2$$

(1.1) $$\text{var}[\hat{\mu}|\text{Ney}] = \frac{1}{n}(\Sigma_{h=1}^{L} W_h \sigma_{hy})^2$$

$$\text{var}[\hat{\mu}|\text{eq}] = \frac{L}{n}\Sigma_{h=1}^{L} W_h^2 \sigma_{hy}^2,$$

for proportional, Neyman and equal allocation, respectively.

1.1.2. Optimum Stratification

The term "optimum stratification" refers to choosing those stratum boundaries that minimize the variance of the stratified random sampling estimate of the population mean when the number of strata and the method of allocation are fixed. The problem was first examined theoretically by Dalenius (1950); we shall briefly consider his development. The theories of sample surveys are based on sampling from a finite population. Since using this approach would result in heavy algebra, suppose that the study population's distribution can be represented by the probability distribution function, $f_Y(y)$, $a' \leq y \leq b'$. If $a' < y_1 < y_2 < \ldots < y_{L-1} < b'$ are the points of demarcation between the L strata, then the moments of the study variable within the h-th stratum can be defined by

$$W_h = \int_{y_{h-1}}^{y_h} f_Y(t)dt$$

$$\mu_{hy} = \frac{1}{W_h}\int_{y_{h-1}}^{y_h} t f_Y(t)dt$$

and $$\sigma_{hy}^2 = \frac{1}{W_h}\int_{y_{h-1}}^{y_h} t^2 f_Y(t)dt - \mu_{hy}^2$$

for $h=1,2,\ldots,L$. With $f_Y(y)$ known and the number of strata considered fixed, $\text{var}[\hat{\mu}|\text{prop}]$, $\text{var}[\hat{\mu}|\text{Ney}]$ and $\text{var}[\hat{\mu}|\text{eq}]$ are functions of only the stratum boundaries. Thus, those boundaries that provide the greatest reduction in $\text{var}[\hat{\mu}|\cdot]$ are solutions to the system of equations

$$\frac{\partial \text{var}[\hat{\mu}|\cdot]}{\partial y_h} = 0 \quad h=1,2,\ldots,L-1.$$

In particular, after some manipulation, the resulting systems of equations for the three methods of allocation listed above can be shown to be

$$(1.2) \qquad y_h - \frac{\mu_{hy}+\mu_{ky}}{2} = 0 \quad k=h+1,\ h=1,2,\ldots,L-1$$

for proportional allocation,

$$(1.3) \qquad \frac{(y_h-\mu_{hy})^2+\sigma_{hy}^2}{\sigma_{hy}} - \frac{(y_h-\mu_{ky})^2+\sigma_{ky}^2}{\sigma_{ky}} = 0 \quad k=h+1,\ h=1,2,\ldots,L-1$$

for Neyman allocation, and

$$(1.4) \qquad W_h((y_h-\mu_{hy})^2+\sigma_{hy}^2)-W_k((y_h-\mu_{ky})^2+\sigma_{ky}^2) = 0 \quad k=h+1,\ h=1,2,\ldots,L-1$$

for equal allocation. Since the moments of the strata are themselves functions of the stratum boundaries, these systems of non-linear equations do not in general lend themselves to closed form solution, but instead must be solved iteratively.

In his initial work, because of its mathematical tractability, Dalenius assumed that the auxiliary (stratification) and study variables were identical. Recognizing that this equivalence is not generally true in practice, Dalenius and Gurney (1951) considered the problem of optimum stratification when stratification was carried out on a specific auxiliary variable. The authors assumed that the two variables were related by the relationship

$$y = g(x)+e$$

where X is the auxiliary variable with probability distribution function $f_X(x)$, $a \le x \le b$, and where e is the random deviation of Y from the regression line $y=g(x)$, X and e are uncorrelated, and $E[e]=0$. Singh and Sukhatme (1969) suggested a more sensitive residual variation form with $E[e|x]=0$ and $var[e|x]=\phi(x)>0$. Through both of these models, the moments of Y within the strata can be expressed as a function of the stratum moments of g(X) and X, thus making $var[\hat{\mu}|\cdot]$ a function of the stratum boundaries with respect to the auxiliary variable, $a < x_1 < x_2 < \ldots < x_{L-1} < b$. In particular, using Singh and Sukhatme's model, we can define

$$W_h = \int_{x_{h-1}}^{x_h} f_X(t)dt$$

$$E[g(X)^s|h] = \frac{1}{W_h}\int_{x_{h-1}}^{x_h} g(t)^s f_X(t)dt, \quad s=1,2$$

$$\sigma_{hg}^2 = E[g(X)^2|h]-E[g(X)|h]^2$$

$$\text{and} \quad \sigma_{he}^2 = \frac{1}{W_h}\int_{x_{h-1}}^{x_h} \phi(t)f_X(t)dt = E[\phi(X)|h]$$

with $h=1,2,\ldots,L$. Note $\sigma_{hy}^2 = \sigma_{hg}^2+\sigma_{he}^2$. Substituting these functions into equation (1.1), the points of optimum stratification of $\text{var}[\hat{\mu}|\cdot]$ are the solutions of

$$\frac{\partial \text{var}[\hat{\mu}|\cdot]}{\partial x_h} = 0 \quad h=1,2,\ldots,L-1 \ .$$

The resulting non-linear equations have the form

$$(1.5) \qquad g(x_h) - \frac{E[g(X)|h]+E[g(X)|k]}{2} = 0 \quad k=h+1,\ h=1,2,\ldots,L-1$$

for proportional allocation,

$$(1.6) \qquad \frac{(g(x_h)-E[g(X)|h])^2+\sigma^2_{hg}+\sigma^2_{he}+\phi(x_h)}{\sqrt{\sigma^2_{hg}+\sigma^2_{he}}} - \frac{(g(x_h)-E[g(X)|k])^2+\sigma^2_{kg}+\sigma^2_{ke}+\phi(x_h)}{\sqrt{\sigma^2_{kg}+\sigma^2_{ke}}} = 0 \quad k=h+1,\ h=1,2,\ldots,L-1$$

for Neyman allocation, and

$$(1.7) \qquad W_h((g(x_h)-E[g(X)|h])^2+\sigma^2_{hg}+\sigma^2_{he}+\phi(x_h)) - W_k((g(x_h)-E[g(X)|k])^2+\sigma^2_{kg}+\sigma^2_{ke}+\phi(x_h)) = 0 \quad k=h+1,\ h=1,2,\ldots,L-1$$

for equal allocation. As was true with Dalenius's original equations, equations (1.2)-(1.4), the above Dalenius-Gurney equations* do not

---

*Though equations (1.5)-(1.7) were derived from Singh and Sukhatme's model instead of the less sensitive one suggested by Dalenius and Gurney, we shall refer to these equations in the sequel as the Dalenius-Gurney equations.

generally lend themselves to closed form solutions, but must be solved iteratively.

Let us consider the situation when $g(x)=\alpha+\beta x$ and $\phi(x)=1, \forall x$, i.e. we have a simple linear regression model with homoscedastic residual variance. Under Neyman allocation, the Dalenius-Gurney equations can be written, after some manipulation, as

$$\frac{(x_h-\mu_{hx})^2+\sigma_{hx}^2+\frac{2}{\beta^2}}{\sqrt{\sigma_{hx}^2+\frac{1}{\beta^2}}}-\frac{(x_h-\mu_{kx})^2+\sigma_{kx}^2+\frac{2}{\beta^2}}{\sqrt{\sigma_{kx}^2+\frac{1}{\beta^2}}}=0 \quad k=h+1,\ h=1,2,\ldots,L-1 .$$

Recall that $\beta=\frac{\rho}{\sqrt{1-\rho^2}}\frac{\sigma_e}{\sigma_x}$, where $\sigma_e^2=\text{var}[e]$, $\sigma_x^2=\text{var}[X]$, and $\rho$ is the correlation between X and Y. Then as $\rho\to 1$, these equations approach the form of Dalenius's original equations (1.3) when stratifying on the auxiliary variable. This result also holds under equal allocation, with equations (1.7) approaching the form of equations (1.4) as $\rho\to 1$. For proportional allocation, the Dalenius-Gurney equations for the simple linear regression model are equivalent to Dalenius's original equations (1.2) when stratifying on the auxiliary variable. Thus, the Dalenius-Gurney equations are approximated by using Dalenius's original equations with the auxiliary variable when the study and auxiliary variables are related by a simple linear regression model (or nearly so), and the correlation between them is high. Dalenius in his 1950 paper mentioned that, though his work was primarily designed to throw some light on the effect of the way in which the population is divided into strata, his results might be applicable, e.g., when trying to estimate the acreage under wheat for the current year by means of a sample of farms selected from the schedules from a previous complete census.

In fact, using Dalenius's original equations with the auxiliary variable is one of the standard techniques employed in practice for the determination of stratum boundaries, though it is not always possible to insure that the two variables can be considered "highly" correlated. However, while the purpose of optimum stratification is to provide the best reduction in the variance of the stratified random sampling estimate of the population mean, stratification is still considered successful if it offers a reasonable improvement over the variance of the estimate of the population mean under simple random sampling. Thus, if employing this approximate technique generated strata providing such reasonable reductions in variance when the two variables were only moderately correlated, the technique could be justified in its use beyond the scope of the theory.

### 1.1.3. Approximating the Dalenius-Gurney Equations

A disadvantage with approximating the Dalenius-Gurney equations by applying Dalenius's original equations on the auxiliary variable is that it ignores the regression relationship between the two variables, generating stratum boundaries independent of the study variable. When this information is available, the variance generated by the approximate technique, though an improvement over simple random sampling, could turn out to be significantly greater than the reduction offered by the use of the Dalenius-Gurney equations. The difficulty with the use of the Dalenius-Gurney equations, however, is that they can be far more cumbersome to solve than Dalenius's original equations, making them almost prohibitively unwieldy (see Appendix C for a discussion of the solution of both series of equations). Fortunately, a variety of much simpler approximate systems of equations have appeared in the literature.

The first approximations which appeared were in a paper by Singh and Sukhatme (1969). In this article, the authors expressed the stratum moments of g(X) and X as asymptotic expansions about the stratum boundaries. Substituting these expressions into the Dalenius-Gurney equations, they generated a series of approximate systems of equations for both proportional and Neyman allocations ranging in simplicity from the system

$$x_h = a + h\,\frac{b-a}{L} \qquad h=1,2,...,L-1,$$

to the fairly intricate

$$K_h^2\int_{x_{h-1}}^{x_h} G(t)\,f_X(t)\,dt = \text{constant}, \quad h=1,2,...,L-1$$

where $K_h=x_h-x_{h-1}$, and G(x) is a function of the regression model. However, the most practical, in terms of balancing ease of solution with including the most information about the functional relationship between the study and auxiliary variables, seems to be the cum H(x) rules, i.e. those equations of the form

$$(1.8) \qquad \int_{x_{h-1}}^{x_h} H(t)\,dt = \text{constant} = \frac{\int_a^b H(t)\,dt}{L} \qquad h=1,2,...,L-1$$

where the boundaries that solve these equations equalize the cumulatives of the function H(x). It is interesting to note that two of the more successful approximations to Dalenius's original equations have this cum H(x) rule form, namely Dalenius and Hodges's (1957, 1959) $\text{cum}\sqrt{f(x)}$ rule for Neyman allocation, and Thomsen's (1976) $\text{cum}\sqrt[3]{f(x)}$ rule for proportional allocation.

Two other cum H(x) rules have appeared in the literature. Singh and Parkash (1975) used the same asymptotic expansions about the stratum boundaries employed by Singh and Sukhatme to generate a cum H(x) rule for equal allocation. And, Singh (1975), after approximating the stratum variances of Y and using them to derive a system of equations under Neyman allocation which were analogous to the Dalenius-Gurney equation (1.6), employed Singh and Sukhatme's asymptotic expansions of the stratum moments about the stratum boundaries to develop another cum H(x) rule.

However, in spite of the advantages that these approximations offer when using the Dalenius-Gurney equations in practice, there has yet to appear in the literature a systematic study of the various cum H(x) rules.

### 1.1.4. Efficient Number of Strata

The problem of determining the number of strata to use is a complex question depending upon a number of factors. Additional strata generally produce further reductions in the variance of $\hat{\mu}$. In practice, it is often the case that the choice of the number of strata to use is a compromise between the increased costs of additional strata and the decrease of precision with the use of fewer strata. When it is assumed that costs do not enter into the problem, as we have in this thesis, there are two limits to having an indefinitely large number of strata. The first limit, a natural one, is the sample size. The second limit is set by the fact that the rate of the gain in efficiency of further stratification is a decreasing function of L. And while the gain is substantial for small numbers of strata, it becomes marginal beyond a certain stage. When the population does not possess a natural partitioning, the choice of the boundaries can influence the degree of reduction obtained by additional strata. In particular, if the boundaries are chosen judiciously (e.g. via optimum stratification theory), then using just a few strata will exhaust virtually all the decrease in variance that is practically attainable. Once this point is reached, any marginal increments in efficiency from further stratification can be potentially outweighed by the increase in the clerical work entailed. Thus, when one stratification variable is used, large numbers of strata are in general not employed in practice; Cochran (1977) has indicated that six or less strata effectively achieve this point, while Kish (1965) suggests using no more than 10 strata.

When one knows the study population, this second limit can be determined by computing the best reduction in variance for a series of stratum numbers. The efficient number of strata is thus defined as that point beyond which only marginal reduction in variance is afforded by any further increase in the number of strata. In the absence of the knowledge of the study variable that this procedure requires, approximations to $\mathrm{var}[\hat{\mu}]$ are employed. Let $V_L=\mathrm{var}[\hat{\mu}]$, with L strata being used.

Dalenius and Gurney (1951) suggested approximating $V_L$ by use of hypothetical distributions in the place of the study variable. When the rectangular distribution is used as a hypothetical distribution, the ratio $V_L/V_{L-1}$ is approximated by $(L-1)^2/L^2$. Dalenius (1953) demonstrated that this approximation compared well to the actual ratios when the study variable had a $\chi_2^2$ or $\chi_4^2$ distribution. Empirical studies by Cochran (1961) and Hess, Sethi and Balakrishnan (1966), working with skewed distributions, have also shown that this $(L-1)^2/L^2$ approximation worked well in practice.

A more comprehensive approach consists of using an approximation to the design effect, $V_L/V_1$ (see Kish (1965)), involving L as well as the correlation between X and Y, as parameters. (Note that $V_1$ is the variance of the sample mean under simple random sampling.) Suggesting that the density of the auxiliary variable be approximated by the rectangular distribution, and the regression relationship between the two variables be modeled by a simple linear regression model with homoscedastic error variance, Cochran (1977) derived the result $V_L/V_1 \geq \rho^2/L^2+(1-\rho^2)$, where $\rho^2$ is the correlation between X and Y. Serfling (1968) expanded this rough approximation and, for the case of Neyman allocation, derived the approximate relationship $V_L/V_1=K_1\rho^2/L^2+K_2(1-\rho^2)$, where $K_1$ and $K_2$ both depend upon the underlying distribution of the auxiliary variable. This approximation can also be shown to result when equal allocation is employed. The case when proportional allocation was used was considered by Thomsen (1976), who derived $V_L/V_1=K_3\rho^2/L^2+(1-\rho^2)$, where $K_3$ again depends upon the distribution of the auxiliary variable.

In yet another approach, Sethi suggested expressing the design effect as a quadratic function of L, i.e. $V_L/V_1=(A+BL+CL^2)^{-1}$, where A, B and C are to be computed by considering the values of the design effect when L=1,2 and 3. This function can then be used to determine the efficient number of strata for $L\geq 4$ (see Murthy (1967)).

In spite of the interest developing approximations to $V_L$ and $V_L/V_1$ has received in the literature, in only a few studies has attention been given to the performances of these approximations in practice.

### 1.1.5. Organization of Thesis

When little is known about the functional relationship between the auxiliary and study variables, one approximation to the Dalenius-Gurney equations is to employ Dalenius's original equations with the auxiliary variable. In Chapter II, we examine the improvement over the use of simple random sampling provided by this method of stratification.

When the functional relationship between the auxiliary and study variables is known, the use of the cum H(x) rule approximations to the Dalenius-Gurney equations is suggested by the equations's prohibitively unwieldy nature. How well the boundaries generated by the approximations that have appeared in the literature compare with those which solve the Dalenius-Gurney equations is examined in Chapter III.

And, finally, in Chapter IV, we examine the approximations to $V_L$ and $V_L/V_1$ which are employed in practice to determine the number of strata to use.

## 1.2. Previous Solutions

The three questions introduced in the last section have received some attention in the literature. In this section, we will review those papers.

### 1.2.1. Dalenius's Original Equations and the Auxiliary Variable

Due to the non-linear nature of Dalenius's original equations, and the computational problems they presented (particularly for Neyman allocation with large numbers of strata), a number of subsequent papers appeared devoted to developing simpler procedures providing nearly optimum stratum boundaries as approximations to the more cumbersome exact equations. Such approximations are generally out of the scope of this thesis since we will use the exact equations for our boundary computations. However, along with these approximations have also appeared a number of investigations studying how these approximations fared in practical situations (see Kpedekpo (1973) for an excellent review of this whole body of literature). While most of these studies have only dealt with the practical use of the proposed approximations, one study has partially considered the question of interest to us.

In a study designed to systematically investigate two of the proposed approximations to Dalenius's original equations, Dalenius and Hodges's cum$\sqrt{f(x)}$ rule and Thomsen's cum$\sqrt[3]{f(x)}$ rule, Anderson, Kish and Cornell (1975) chose to model their universe by a bivariate normal distribution. By treating one variable as the stratification variable, they were able to examine how the choice of the stratum boundaries on this variable affected the variance of the estimate $\hat{\mu}$ with respect to the other variable. The authors considered both Neyman and proportional allocations, as well as let the number of strata and the correlation between the two variables vary over a range of values. While the primary emphasis of this study dealt with choosing the stratum boundaries by the approximate methods, the authors also considered the situation when the boundaries were chosen by using Dalenius's original equations. For a range of correlation values and numbers of strata, the authors computed the percentage decrease in the variance of $\hat{\mu}$ compared with the variance of the sample mean under simple random sampling:

> Given that the estimation variable Y and the stratification variable X are approximately normal and a reasonable guess is available for $\rho$, an investigator can...find estimates of the gains in precision which can be achieved.... With a knowledge of the cost of additional strata and the worth of increased precision, the investigator can choose a value of L...which is a compromise between the conflicting desires of minimizing cost and maximizing precision. (1975)

While the work of Anderson, Kish and Cornell does address the question of interest to us, they have restricted the auxiliary variable's distribution to the normal and the functional relationship between the study and auxiliary variables to the simple linear regression model. Sufficient cases occur in practice in which one or both of these assumptions are not satisfied to justify an expansion of their model of the universe to a wider range of situations.

### 1.2.2. Appxoximating the Dalenius-Gurney Equations

With the appearance of their approximations, Singh and Sukhatme

(1969), Singh (1975) and Singh and Parkash (1975) also included numerical studies of their efficiencies. However, in each study, only the simple linear regression model was used for the functional relationship between variables, though they allowed the residual variation to range beyond homoscedasticity. The distributions of the auxiliary variable considered were the rectangular, the right-triangular and the exponential distribution.

Since these approximations represent a practical approach to the use of the Dalenius-Gurney equations, it would be of interest to examine their effectiveness over more than the range of functional relationships parametrized by the simple linear regression model.

### 1.2.3. Approximating the Efficient Number of Strata

An examination of the performances of the proposed approximations designed to model the effect of increasing the number of strata has been addressed in just two studies. Cochran (1961), looking at 8 data sets (all of them highly skewed), compared the $V_L/V_{L-1}$ ratios for those distributions with Dalenius's proposed $(L-1)^2/L^2$ approximation.

Hess, Sethi and Balakrishnan (1966), using the trivariate universe of United States Non-Federal short-term general hospitals*, compared the $V_L/V_{L-1}$ ratios of all three variables with $(L-1)^2/L^2$. The authors also compared the design effect for admission and inpatient days with

*The authors used bed capacity, number of admissions, and inpatient days as their variables, with beds as the auxiliary variable. The distribution of this auxiliary variable was highly skewed.

beds used as an auxiliary variable to the approximation $\rho^2/L^2+(1-\rho^2)$.

Since more potentially sensitive estimates for the ratio $V_L/V_1$ (and thus for $V_L/V_{L-1}$) are available than the two which have been examined, it would be of interest to study their performances in practice as well.

## 1.3. The Model

To examine the proposed questions, it is necessary to specify a model with the required parametric latitude. The basic components of our model will be the form of the distribution of the auxiliary variable, and the nature of the functional relationship associating the two variables.

### 1.3.1. Distribution of the Auxiliary Variable

For the form of the distribution of the auxiliary variable, we have the choice of using actual data sets (Cochran (1961) and Hess, Sethi and Balakrishnan (1966) employed this approach in their empirical studies), or of using a theoretical distribution (as did Anderson, Kish and Cornell (1975) in their investigation). We will adopt the latter approach for two reasons. Firstly, use of the theoretical distribution will afford more mathematical tractability in handling the functional relationship between the two variables. And, secondly, theoretical distributions will allow us to bypass the problems associated with the discreteness of finite populations, a factor which, while certainly relevant in any study dealing with sampling from a finite universe, nonetheless eludes quantification (e.g. two different finite populations from the same superpopulation could suffer from different discreteness problems); we will delay its consideration until a later time.

As for the distributions themselves, the populations employed in the literature for empirical comparisons were either highly skewed, modeling distributions roughly representative of those encountered in the sampling of institutions in which there were many small institutions and few large ones, or they were normally distributed. Consequently, the densities we will use include the $\chi^2_1$ and $\chi^2_2$, modeling the skewed distributions, and the normal.

### 1.3.2. Regression Function

To parametrize the second component of our model, the functional relationship between the two variables, we chose (following Anderson, Kish and Cornell) to use the regression model originally proposed by Dalenius and Gurney (1951) and later more fully parametrized by Singh and Sukhatme (1969):

$$y = g(x)+e \tag{1.9}$$

with $x \sim f_X(x)$, $E[e|x] = 0$ and $var[e|x] = \phi(x)>0$, where $f_X(x)$ is the probability distribution function of X, g(x) is the regression function of Y on X, and $\phi(x)$ the residual variation of that regression.

For the range of the regression function, g(x), we will consider, the previous studies in the literature have only employed the linear model. However, Calvin (1978) has indicated that the most common functional relationships between X and Y found in practice are the linear, quadratic and the exponential functions. Thus, we will model the regression model with the following:

$$\text{linear:} \quad g(x) = \alpha+\beta X$$

$$\text{quadratic:} \quad g(x) = \alpha+\beta X+\gamma X^2, \ \gamma\geq 0$$

$$\text{and exponential:} \quad g(x) = \alpha+\beta \exp(\gamma X), \ \gamma\geq 0.$$

By varying the values of the two parameters, $\beta$ and $\gamma$ (the linear model is a special case of the quadratic when $\gamma = 0$), a wide range of practical situations can be covered. However, the actual values of $\beta$ and $\gamma$ are themselves not informative. Consequently, we will use two other parameters that do have interpretations to us and that can be used to specify $\beta$ and $\gamma$. For this purpose, we chose to use a measure of the "linear" fit of the regression model, the correlation between X and Y, and a measure of the "non-linear" fit, the ratio of the expected squared error due to linear lack-of-fit to the pure residual error (see Appendix A for the derivation of these parameters). The last part of the functional relationship model, the form of the residual variation, $\phi(x)$, will follow Singh and Sukhatme (1969) with $\phi(x) = x^r$.

### 1.3.3. Range of Model Parameters

Let us consider the range of the models we will consider. The variance of $\hat{\mu}$ is determined by specification of (i) the distribution of X, (ii) the method of allocation, (iii) the number of strata, (iv) the form of the regression function, (v) the values of $\beta$ and $\gamma$ and (vi) the value of the exponent of the residual variation function. We have already discussed the components of (i), (ii) and (iv); it remains to specify the range of the rest of the factors.

Cochran (1977) indicated that one would not expect much further reduction in the variance of $\hat{\mu}$ beyond 6 strata; Kish (1965) has listed the practical range of the number of strata to be between 3 and 10. We will let L=2,3,...,8.

To specify the value of $\beta$ and $\gamma$ of the regression function, we introduced two parameters, $\rho$ and C (see Appendix A). Together with the exponent of the residual variation, r, specification of this triad completely determines the regression relationship between X and Y. For the measure of the "linear" fit, we will let $\rho$ = .00, .30, .60, .90 and .99; for the measure of the non-linear" fit, we will let C = 0, 1 and 2 (modeling the amount of pure error units of linear lack-of-fit in the full regression relationship). Note that when C=0 the quadratic model reduces to the linear model.

The range of the exponent of the residual variation will depend upon the particular regression function under consideration. For the linear model, r will vary over a full range of values, r = 0, 1 and 2, modeling homoscedasticity, residual variation proportional to the stratum means, and residual variation proportional to the stratum means squared, respectively. For the quadratic model, since we would not expect a quadratic regression model in practice to have homoscedastic error variation, we will restrict r to the values of 1 and 2. And, for the exponential model, since our modeling of exponential regression with an additive error was for mathematical tractability only (we would expect the model $y = \beta\exp(\gamma x+e)$ to be more realistic), the residual variation form $\phi(x) = x^2$ allows us to more closely model a multiplicative error structure. We will restrict our attention to r=2.

Finally, in Chapter II, we are considering the practical question of optimum stratification on the auxiliary variable independent of the functional relationship between the study and auxiliary variables. To allow us to examine this question over the "universe" of models we would expect to find in practice, we will consider all three regression functions. However, for Chapter III, to evaluate the performances of the available approximations to the Dalenius-Gurney equations in the literature, we will consider only the quadratic and linear models; we can get a reasonable approximation to the exponential model with judicious use of the parameters of the quadratic model.

# II. DALENIUS'S ORIGINAL EQUATIONS AND THE AUXILIARY VARIABLE

When little is known about the functional relationship between the study and auxiliary variables, one approach to determining the stratum boundaries is to employ Dalenius's original equations on the auxiliary variable. In this chapter, our interest is primarily concerned with whether this use of Dalenius's original equations provides stratification offering reasonable improvements over simple random sampling. Since our interest is in the reduction afforded by the use of stratification, we will use the design effect, the ratio of $\text{var}[\hat{\mu}]$ to the variance of the sample mean under simple random sampling, to examine this question.

In the next section, we introduce notation and background. We will then consider the question from an empirical point of view, examining the reduction afforded by stratification compared to simple random sampling over the range of models introduced in the last chapter.

In the third section, we will briefly consider the situation when the population has a natural partitioning with the number of strata considered fixed, leaving only the method of allocation variable. When little is known about the functional relationship between the two variables, Neyman allocation, theoretically offering the largest reductions in variance, must be approximated by replacing the stratum variances of the study variable with the stratum variances of the auxiliary variable. How reasonable this approximation is in practice will be addressed from an algebraic point of view.

## 2.1. Notation and Background

Recall that if $\bar{y}_h$ denotes the sample mean from the h-th stratum, h=1,2,...,L, then $\hat{\mu}=\Sigma_{h=1}^{L}W_h\bar{y}_h$ would be an unbiased estimate of the population mean, with $\text{var}[\hat{\mu}]=\Sigma_{h=1}^{L}\frac{W_h^2}{n_h}\sigma_{hy}^2$ with use of the theoretical distributions for our populations removing the finite population correction from the variance.

It is the nature of proportional and equal allocation formulae that they remain the same whether they are defined with respect to the study variable, or with respect to the auxiliary variable. However, this is not the case with Neyman allocation. In particular, we have two separate forms of this allocation formula, depending upon whether it is defined practically, with respect to the auxiliary variable (which we will refer to as X-optimal allocation), or whether it is defined theoretically, with respect to the study variable (which we will continue to refer to as Neyman allocation). The forms of these two methods of allocation can thus be written

$$\frac{n_h}{n} = \frac{W_h\sigma_{hx}}{\Sigma_{k=1}^{L} W_k\sigma_{kx}} \qquad \text{for X-optimal allocation}$$

and

$$\frac{n_h}{n} = \frac{W_h\sigma_{hy}}{\Sigma_{k=1}^{L} W_k\sigma_{ky}} \qquad \text{for Neyman allocation.}$$

The X-optimal allocation form can be substituted into $\text{var}[\hat{\mu}]$, resulting in the variance

$$(2.1) \qquad \text{var}[\hat{\mu}|X\text{-opt}] = \frac{1}{n}\left(\sum_{h=1}^{L} W_h \frac{\sigma^2_{hy}}{\sigma_{hx}}\right)\left(\sum_{h=1}^{L} W_h \sigma^2_{hx}\right) \quad .$$

In practice, when nothing is known about the functional relationship between X and Y, we would use X-optimal allocation instead of Neyman allocation. Thus, the three methods of allocation we will consider in the second section are proportional, equal and X-optimal allocations. The forms of var$[\hat{\mu}]$ for proportional and equal allocations are in (1.1).

### 2.1.1 Algebraic Forms of $\sigma^2_{hy}$

Recall that when we model the functional relationship between X and Y by (1.9), we may write $\sigma^2_{hy}=\sigma^2_{hg}+\sigma^2_{he}$. Let us examine the form of $\sigma^2_{hy}$ over our range of g(x)'s.

Linear: $g(x) = \alpha+\beta X$. Since $\beta^2=\frac{\rho^2}{1-\rho^2}\frac{\sigma^2_e}{\sigma^2_x}$, $\sigma^2_{hg}$ can be written as

$$\sigma^2_{hg} = \beta^2\sigma^2_{hx} = \rho^2\sigma^2_y \frac{\sigma^2_{hx}}{\sigma^2_x}$$

where $\rho$ is the correlation between X and Y. After some manipulation, we may write $\sigma^2_{hy}$ as

$$(2.2) \qquad \sigma^2_{hy} = \sigma^2_y\left(\rho^2\frac{\sigma^2_{hx}}{\sigma^2_x} + (1-\rho^2)\frac{\sigma^2_{he}}{\sigma^2_e}\right) .$$

Quadratic: $g(x) = \alpha+\beta x+\gamma x^2$, $\gamma\geq 0$. Let us define the following

$$C_x = \text{cov}[X,X^2] \qquad C_h = \text{cov}[X,X^2|h]$$

$$V_x = \text{var}[X^2] \qquad V_h = \text{var}[X^2|h]$$

$$D_x = V_x - C^2_x/\sigma^2_x \qquad B_x = \rho - \frac{C_x}{\sigma_x\sqrt{D_x}}\sqrt{\pi^2-\rho^2}$$

where, for instance,

$$C_h = \frac{1}{W_h}\int_{x_{h-1}}^{x_h} t\, f_X(t)dt - \left(\frac{1}{W_h}\int_{x_{h-1}}^{x_h} t f_X(t)dt\right)\left(\frac{1}{W_h}\int_{x_{h-1}}^{x_h} t\, f_X(t)dt\right)$$

and where $\pi$, introduced in Appendix A, is a measure of model fit analogous to $\rho$, i.e., for the linear model, $\sigma_y^2$ and $\sigma_e^2$ are related by $\sigma_e^2=(1-\rho^2)\sigma_y^2$, while for the non-linear model, $\sigma_y^2$ and $\sigma_e^2$ are related by $\sigma_e^2=(1-\pi^2)\sigma_y^2$. Note that for our quadratic model, with all other factors constant, $(\beta,\gamma)$ are one-to-one with $(\rho,\pi)$. The range of $\pi$ is $\rho\leq\pi\leq 1$, so that as $\pi\rightarrow\rho$, the quadratic model approaches the linear model.

Using these parameters, after some manipulation, $\sigma_{hg}^2$ can be written as

$$\sigma_{hg}^2 = \sigma_y^2\left(B_x^2\frac{\sigma_{hx}^2}{\sigma_x^2} + \frac{\pi^2-\rho^2}{D_x}V_h + \frac{2B_x}{\sigma_x\sqrt{D_x}}\sqrt{\pi^2-\rho^2}\ C_h\right)$$

$$= \sigma_y^2 \text{ var}\left[\frac{B_x}{\sigma_x}X + \frac{\sqrt{\pi^2-\rho^2}}{\sqrt{D_x}}X^2 \,\middle|\, h\right].$$

Then, since $\text{var}\left[\frac{B_x}{\sigma_x}X + \frac{\sqrt{\pi^2-\rho^2}}{\sqrt{D_x}}X^2\right]=\pi^2$, we may write $\sigma_{hy}^2$ as

$$\sigma_{hy}^2 = \sigma_y^2\left(\pi^2\frac{\text{var}[G(X)|h]}{\text{var}[G(X)]} + (1-\pi^2)\frac{\sigma_{he}^2}{\sigma_e^2}\right)$$

(2.3)

$$\text{with } G(x) = \frac{B_x}{\sigma_x}x + \frac{\sqrt{\pi^2-\rho^2}}{\sqrt{D_x}}\ x^2$$

and $B_x$ and $D_x$ are defined above. Since $B_x=\rho-\frac{C_x}{\sigma_x\sqrt{D_x}}\sqrt{\pi^2-\rho^2}$, we can write

$$G(x) = \frac{\rho}{\sigma_x} x + \frac{\sqrt{\pi^2-\rho^2}}{\sqrt{D_x}} (x^2 - \frac{C_x}{\sigma_x^2} x).$$

Thus, $\frac{\rho}{\sigma_x} x$ is the linear component of G(x) when $\pi=\rho$, while $\frac{\sqrt{\pi^2-\rho^2}}{\sqrt{D_x}} (x^2 - \frac{C_x}{\sigma_x^2} x)$ is the non-linear component which is added to $\frac{\rho}{\sigma_x} x$ when $\rho<\pi$.

Exponential: $g(x)=\alpha+\beta\exp(\gamma x)$, $\gamma\geq 0$. Let us define the random variable $U=\exp(\gamma X)$. Then we may write $\sigma^2_{hg}$ as

$$\sigma^2_{hg} = \beta^2 \mathrm{var}[\exp(\gamma X)|h]$$

$$= \beta^2\sigma^2_{hu}, \text{ say.}$$

Then, after some manipulation, $\sigma^2_{hy}$ may be written as

$$\sigma^2_{hy} = \sigma^2_y \left( \pi^2 \frac{\sigma^2_{hu}}{\sigma^2_u} + (1-\pi^2) \frac{\sigma^2_{he}}{\sigma^2_e} \right) \tag{2.4}$$

where $\sigma^2_u = E[\exp(2\gamma X)] - E[\exp(\gamma X)]^2$. Note that the major impact $\rho$ has in this exponential model is in the computation of the parameter $\gamma$, which must be determined iteratively (see Appendix A).

If we define $\omega$ as the correlation between X and U, then, since $\omega=\rho/\pi$, as $\rho\to\pi$, $\omega\to 1$. Thus, since a correlation between X and U of 1 implies we can express U as a linear function of X, it can be shown that as $\rho\to\pi$, $\frac{\sigma^2_{hu}}{\sigma^2_u} \to \frac{\sigma^2_{hx}}{\sigma^2_x}$.

However, unlike the quadratic model which does not particularly simplify as $\rho\to 0$, (2.4) can be reduced to a simpler form. To see this, let us consider the situation when X is normally distributed (the case when X has a $\chi^2_\nu$ distribution is analogous). We can write

$$E[U^r|h] = E[e^{r\gamma X}|h] = \frac{1}{W_h}T_h(r\gamma\sigma_x)\exp(r\gamma E[X] + \frac{r^2\gamma^2\sigma_x^2}{2}), \quad r=1,2$$

where $T_h(a) = \int_{x_{h-1}-a}^{x_h-a} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}dt$. Then,

$$\frac{\sigma^2_{hu}}{\sigma^2_u} = \frac{\frac{T_h(2\gamma\sigma_x)}{W_h} - e^{-\gamma^2\sigma_x^2}\left(\frac{T_h(\gamma\sigma_x)}{W_h}\right)^2}{1 - e^{-\gamma^2\sigma_x^2}}$$

As $\rho\to 0$, $\gamma\to\infty$, and since $a<x_h<b, \forall\, h$, with a and b finite, then

$T_h(r\gamma\sigma_x)\to 0$, and thus $\frac{\sigma^2_{hu}}{\sigma^2_u}\to 0$ as $\rho\to 0$. So, (2.4) takes the form

$$\sigma^2_{hy} = \sigma^2_y(1-\pi^2)\frac{\sigma^2_{he}}{\sigma^2_e}.$$

## 2.1.2. Algebraic Forms of $\text{var}[\hat{\mu}|\cdot]$

Let us now examine the algebraic form of the three variances $\text{var}[\hat{\mu}|\cdot]$ for the three regression functions.

Linear: Combining (1.1) and (2.2), we may write

$$(2.5) \qquad \text{var}[\hat{\mu}|\text{prop}] = \frac{\sigma^2_y}{n}\left(\rho^2\frac{\sum_{h=1}^{L} W_h\sigma^2_{hx}}{\sigma^2_x} + 1-\rho^2\right)$$

since $\sigma^2_e=\text{var}[e]= E[\text{var}[e|X]]+\text{var}[E[e|X]]=\sum_{h=1}^{L} W_h\sigma^2_{he}+0.$

If we imagine a stratified random sample taken from the auxiliary variable's population, with $\bar{x}_h$ the h-th stratum sample mean, then when the auxiliary variable's population mean is estimated by

$\hat{\mu}_x=\sum_{h=1}^{L} W_h\bar{x}_h$, we note $\text{var}[\hat{\mu}_x|\text{prop}]=\frac{1}{n}\sum_{h=1}^{L} W_h\sigma_{hx}^2$. Further, if $\bar{y}$ and $\bar{x}$ denote the sample means of simple random samples of size n from the study population and auxiliary population, respectively, then $\text{var}[\bar{y}]=\frac{1}{n}\sigma_y^2$ and $\text{var}[\bar{x}]=\frac{1}{n}\sigma_x^2$. Using these results, we were able to alter (2.5) to

$$\text{(2.6)} \qquad \frac{\text{var}[\hat{\mu}|\text{prop}]}{\text{var}[\bar{y}]} = \rho^2 \frac{\text{var}[\hat{\mu}_x|\text{prop}]}{\text{var}[\bar{x}]} + 1-\rho^2.$$

Combining (2.1) with (2.2), and (1.1) with (2.2), we are able to derive the following forms analogous to (2.5),

$$\text{var}[\hat{\mu}|\text{X-opt}] = \frac{\sigma_y^2}{n}\left(\rho^2\frac{(\sum_{h=1}^{L} W_h\sigma_{hx})^2}{\sigma_x^2} + (1-\rho^2)\frac{(\sum_{h=1}^{L} W_h\frac{\sigma_{he}^2}{\sigma_{hx}})(\sum_{h=1}^{L} W_h\sigma_{hx})}{\sigma_e^2}\right)$$

$$\text{(2.7)} \qquad \text{var}[\hat{\mu}|\text{eq}] = \frac{\sigma_y^2}{n}\left(\rho^2\frac{L\sum_{h=1}^{L} W_h^2\sigma_{hx}^2}{\sigma_x^2} + (1-\rho^2)\frac{L\sum_{h=1}^{L} W_h^2\sigma_{he}^2}{\sigma_e^2}\right)$$

for X-optimal and equal allocations, respectively. And, imagining a stratified random sample from the population of e's, with $\bar{e}_h$ denoting the h-th stratum sample mean, let us define $\hat{\mu}_e=\sum_{h=1}^{L}W_h\bar{e}_h$. Thus, noting, for example, that $\text{var}[\hat{\mu}_x|\text{X-opt}]=\frac{1}{n}(\sum_{h=1}^{L}W_h\sigma_{hx})^2$ and $\text{var}[\hat{\mu}_e|\text{eq}]=\frac{L}{n}\sum_{h=1}^{L}W_h^2\sigma_{he}^2$, we can alter (2.7) in an analogous manner to (2.6) to derive

$$\text{(2.8)} \quad \frac{\text{var}[\hat{\mu}|\cdot]}{\text{var}[\bar{y}]} = \rho^2 \frac{\text{var}[\hat{\mu}_x|\cdot]}{\text{var}[\bar{x}]} + (1-\rho^2) \frac{\text{var}[\hat{\mu}_e|\cdot]}{\text{var}[\bar{e}]}$$

for X-optimal and equal allocations. Since $\frac{1}{n}\sigma_e^2 = \text{var}[\hat{\mu}_e|\text{prop}]$, (2.8) is also true for proportional allocation.

Thus, for the linear model, when the correlation between X and Y is high, the design effect of $\hat{\mu}$ is approximately the design effect of $\hat{\mu}_x$ for all three allocations, while for moderate and small correlation values, it becomes a convex combination of the design effect of $\hat{\mu}_x$ and a function of the residual variation.

Quadratic: Considering G(x) defined in (2.3) as our variable of interest, and defining $\hat{\mu}_G$ as the stratified random sampling estimate of the population mean, when combining (2.3) with (1.1) and (2.1) and employing manipulation analogous to that for the linear model, we derive

$$\text{(2.9)} \quad \frac{\text{var}[\hat{\mu}|\cdot]}{\text{var}[\bar{y}]} = \pi^2 \frac{\text{var}[\hat{\mu}_G|\cdot]}{\text{var}[\overline{G(X)}]} + (1-\pi^2) \frac{\text{var}[\hat{\mu}_e|\cdot]}{\text{var}[\bar{e}]}$$

for all three methods of allocation, where $\text{var}[\overline{G(X)}] = \frac{1}{n}\text{var}[G(X)]$. Thus, the design effect of $\hat{\mu}$ under the quadratic model is a convex combination of the "design effect" of $\hat{\mu}_G$ and a function of the residual variation. Note that if $\rho<\pi$, then $\pi \to 1$ does not necessarily simplify (2.9) into familiar terms; it is necessary for $\rho \to \pi$ before (2.9) approaches the form of (2.8).

Exponential: For $U=\exp(\gamma X)$, following the quadratic model and defining $\hat{\mu}_u$ as the stratified random sampling estimate of U's population mean, we can combine (2.4) with the $\text{var}[\hat{\mu}|\cdot]$ forms to generate

$$(2.10) \qquad \frac{\mathrm{var}[\hat{\mu}|\cdot]}{\mathrm{var}[\bar{y}]} \qquad \pi^2\,\frac{\mathrm{var}[\hat{\mu}_u|\cdot]}{\mathrm{var}[\bar{u}]} + (1-\pi^2)\,\frac{\mathrm{var}[\hat{\mu}_e|\cdot]}{\mathrm{var}[\bar{e}]}$$

for all three methods of allocation. As was true with the quadratic model, the design effect of $\hat{\mu}$ is a convex combination of the "design effect" of $\hat{\mu}_u$ and a function of the residual variation, which approaches the form (2.8) as $\rho \to \pi$.

## 2.2. Empirical Comparisons

When the auxiliary and study variables are identical, optimum stratification under Neyman allocation turns out to be relatively "robust". To see this, let us define the following two functions of the stratum boundaries,

$$V_{opt}(\underset{\sim}{y}) = (\sum_{h=1}^{L} W_h \sigma_{hy})^2$$

$$V_{prop}(\underset{\sim}{y}) = \sum_{h=1}^{L} W_h \sigma_{hy}^2,$$

where $\underset{\sim}{y}$ is the vector of stratum boundaries on the study variable. Thus, $V_{opt}(\underset{\sim}{y})$ is proportional to the variance of the stratified random sampling estimate under Neyman allocation, and $V_{prop}(\underset{\sim}{y})$ is proportional to the variance of $\hat{\mu}$ under proportional allocation. If we define $\underset{\sim}{y}_o$ and $\underset{\sim}{y}_p$ by the two relationships

$$\min_{\underset{\sim}{y}} V_{opt}(\underset{\sim}{y}) = V_{opt}(\underset{\sim}{y}_o)$$

$$\min_{\underset{\sim}{y}} V_{prop}(\underset{\sim}{y}) = V_{prop}(\underset{\sim}{y}_p),$$

then we have the following inequalities:

$$V_{opt}(\underset{\sim}{y}_o) \le V_{opt}(\underset{\sim}{y}_p) \le V_{prop}(\underset{\sim}{y}_p) \le V_{prop}(\underset{\sim}{y}_o).$$

Thus, this relationship indicates that when optimum stratification is employed on the study variable, minor deviations from the optimum boundaries for Neyman allocation should still provide a better reduction in variance than the optimal use of proportional allocation.

The question of interest here, however, is when the boundaries of the strata are defined on the auxiliary variable. Since the stratum weights, $W_h$, are the same for both auxiliary and study variables, for any set of stratum boundaries, the algebraic nature of proportional allocation assures us that its use offers some reduction in variance when compared with simple random sampling (see Cochran (1977), page 99); neither Neyman allocation, in the form of X-optimal allocation, nor equal allocation can offer us this comfort.

In this section, we will examine the degree of variance reduction afforded by stratifying on the auxiliary variable using the three methods of allocation, proportional, X-optimal and equal, and will consider whether an improvement over simple random sampling is offered.

Let us initially examine the form of the variance of $\hat{\mu}$ under proportional allocation,

$$\text{var}[\hat{\mu}|\text{prop}] = \frac{1}{n}\sum_{h=1}^{L} W_h\sigma^2_{hg} + \frac{1}{n}\sigma^2_e \quad .$$

Referring to Appendix A, we note that the values of the parameters $\beta$ and $\gamma$ for the quadratic model (the linear model being a special case of it), as well as the value of $\beta$ for the exponential model ($\gamma$ being independent of the residual variation in the exponential model), depend upon the value of r only through $\sigma^2_e$. Thus, we are able to factor out $\sigma^2_e$ from the $\sigma^2_{hg}$ terms in both models. And, since $\sigma^2_y = \frac{\sigma^2_e}{1-\pi^2}$ with $\text{var}[\bar{y}] = \frac{1}{n}\sigma^2_y$, in case of proportional allocation, the design effect is independent of the residual variation parameter r.

The tables of the design effects are at the end of this section, starting on page 40: Tables 2.2-2.8 pertain to the quadratic model, with proportional allocation in Table 2.2, X-optimal allocation in Tables 2.3-2.5, and equal allocation in Tables 2.6-2.8; Tables 2.9-2.11 pertain to the exponential model, with Tables 2.9, 2.10 and 2.11 dealing with proportional, X-optimal and equal allocations, respectively.

Examination of the tables reveals the following:

1. For X-optimal and equal allocation, it is only for the linear model (C=0) when $\rho$ is small to moderate and the residual variation is not too heterogeneous ($r \leq 1$) that the use of simple random sampling offers smaller variance than the use of stratification. Proportional allocation, by its nature, can never so worse than simple random sampling. This discrepancy--that non-linear models provide better design effects than the linear models--can be explained by recalling the relationship between $\pi$ and the two model parameters, $\rho$ and C: $\pi = \sqrt{\frac{C+\rho^2}{C+1}}$. For the range

of $\rho$ and C employed in this thesis, we have computed the corresponding values of $\pi$ in Table 2.1. Recalling that $\pi^2$ is the percentage of the total variation of Y explained by the full regression line y=g(x), we note that it isn't until $\rho>.60$ that the full fit of the linear model is commensurate with the non-linear models. Thus, it is not that the non-linear models are necessarily "better" than the linear model, rather it is that the non-linear model under our parametrization explains more variation than the linear model when $\rho$ is small to moderate. We can thus conclude that when the full regression explains more than 40 percent of the variation of Y, the use of this approximate procedure to determine the optimum stratum boundaries does provide an improvement in variance over the use of simple random sampling.

Table 2.1. Values of $\pi$.

| | C | | |
|---|---|---|---|
| $\rho$ | 0 | 1 | 2 |
| .00 | .00 | .7071 | .8165 |
| .30 | .30 | .7382 | .8347 |
| .60 | .60 | .8246 | .8869 |
| .90 | .90 | .9513 | .9678 |
| .99 | .99 | .9950 | .9967 |

2. For the three parameters, $\rho$, C and r, an increase in any of them produces a decrease in the design effect, and thus an improvement in the use of stratification. This can be explained for $\rho$ and C by recalling $\pi=\sqrt{\frac{C+\rho^2}{C+1}}$ ; an increase in either parameter increases the strength of the relationship between X and Y. As for the effect an increase in r has, recall that we have modeled the heterogeneous residual

variation as a function of the first two stratum moments of the auxiliary variable. Thus, as r increases from 0 to 1 to 2, the stratum variance of Y becomes more of a function of the stratum moments of X.

3. When the choice of the method of allocation is primarily to generate the best reduction in variance, it would be of interest to compare the theoretically suggested X-optimal allocation with the more practical proportional allocation. We note, looking at the tables, that except for small L, the ordering between the design effects of the two methods of allocation remains fairly constant as L varies. Thus, to demonstrate the relationship between the allocations, we chose to plot both design effects on the same graph, using $\rho$ as the ordinate. Figures 2.1-2.4 show these relationships when L=6. It is interesting that for the normal distribution under the quadratic model, X-optimal allocation actually almost uniformly does worse than proportional allocation, except for the very high correlation ranges. For the skewed distributions under the linear model, when the residual variation is homoscedastic or moderately heterogeneous, X-optimal allocation does considerably worse than proportional allocation. When the model is non-linear, X-optimal allocation offers a definite advantage in variance reduction over proportional allocation. For the exponential model, X-optimal allocation offers a greater reduction in variance than proportional allocation for all three distributions.

4. Since one of the reasons we considered the use of equal allocation was as an approximation to Neyman allocation, we are also interested in comparing the design effects of equal and X-optimal allocations. Such a comparison indicates that equal allocation actually provides a better

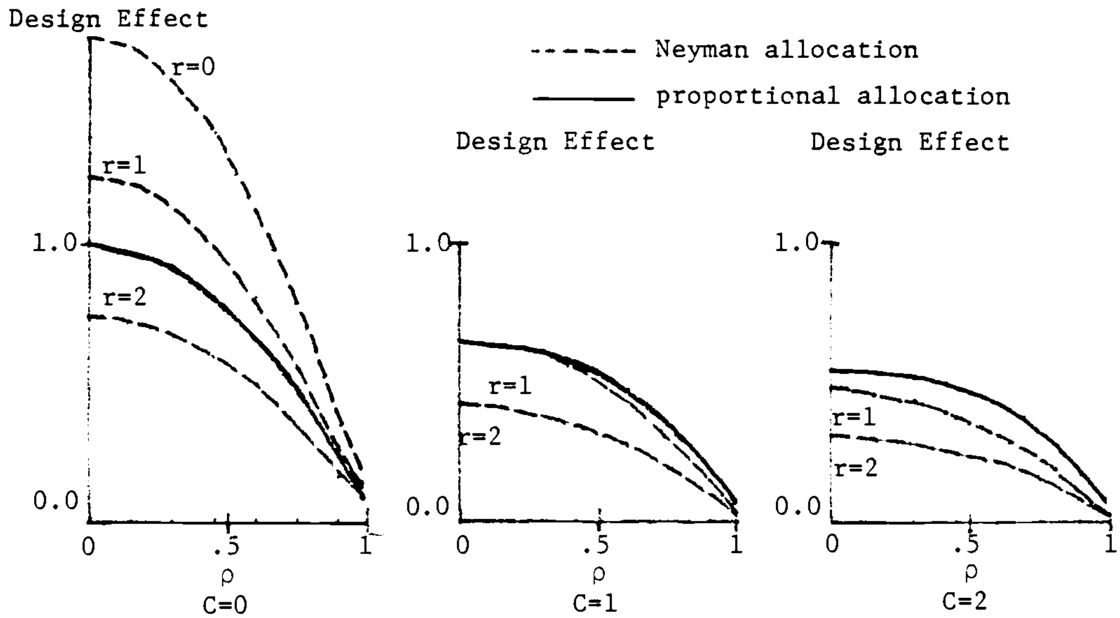



Figure 2.1. Comparison of the design effects using Dalenius's original equations for proportional and X-optimal allocations under the $\chi_1^2$ distribution for the quadratic model when L=6.

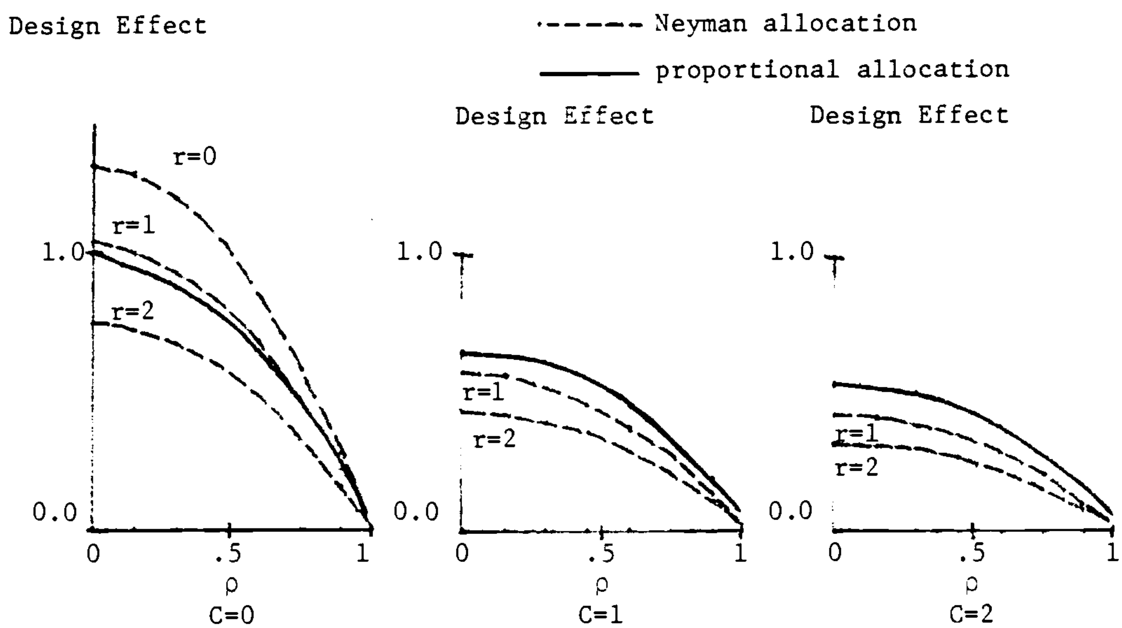



Figure 2.2. Comparison of the design effects using Dalenius's original equations for proportional and X-optimal allocations under the $\chi_2^2$ distribution for the quadratic model when L=6.

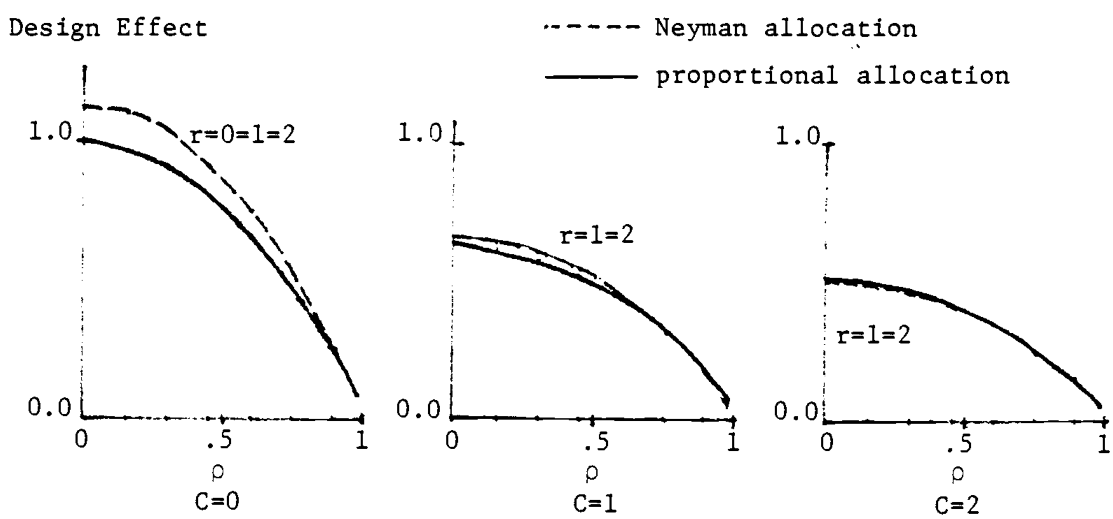



Figure 2.3. Comparison of the design effects using Dalenius's original equations for proportional and X-optimal allocations under the normal distribution for the quadratic model when L=6.

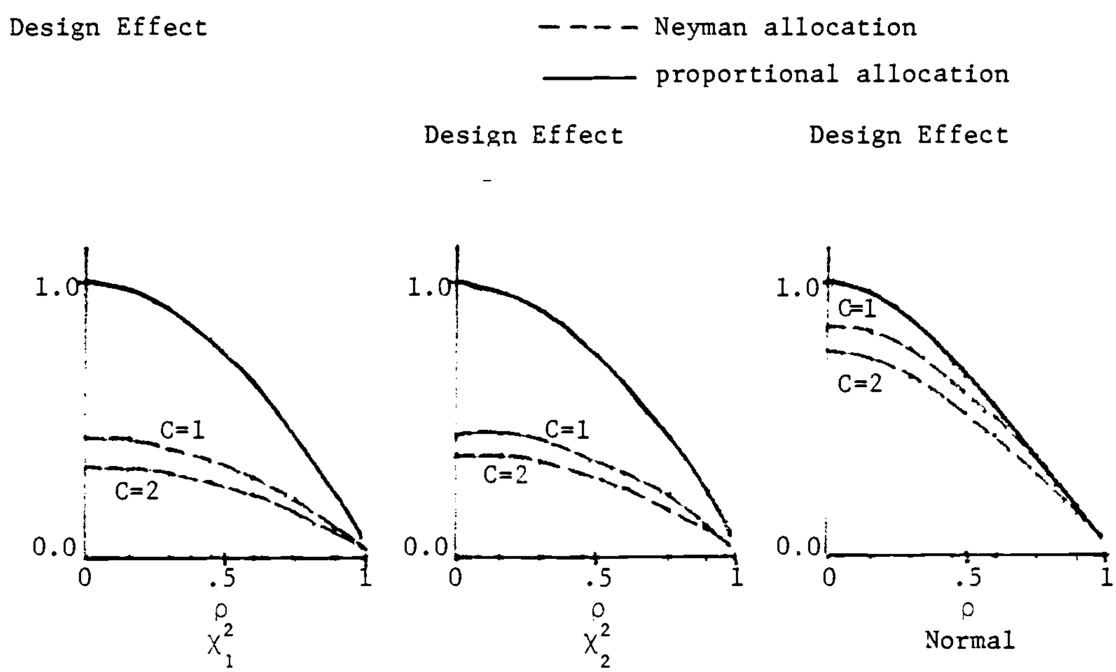



Figure 2.4. Comparison of the design effects using Dalenius's original equations for proportional and X-optimal allocations under the $\chi_1^2$, $\chi_2^2$ and normal distributions for the exponential model when L=6.

reduction in variance than does X-optimal allocation in all but the highest correlation ($\rho$=.99) and in the smallest stratum numbers ($L \leq 3$), though the difference is generally marginal. Since Dalenius's original equations for equal allocation are about as cumbersome to solve as are the equations for Neyman (X-optimal) allocation, the use of equal allocation does not provide a significant advantage in the ease of computing the stratum boundaries.

5. As was indicated by Anderson, Kish and Cornell (1975), if one has an approximate idea of the functional relationship between the two variables, Tables 2.2-2.11 can be used when the cost is not invariant to the number of strata to decide upon the number of strata to use, balancing the precision one desires with the cost of additional strata.

Table 2.2. Design effects from Dalenius's original equations for the quadratic model under proportional allocation for the $\chi_1^2$, $\chi_2^2$ and the normal distributions.

| | | $\chi_1^2$ | | | $\chi_2^2$ | | | Normal | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ρ | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 |
| | .00 | 1.000 | .982 | .977 | 1.000 | .987 | .982 | 1.000 | 1.000 | 1.000 |
| | .30 | .941 | .987 | .991 | .942 | .983 | .987 | .943 | .943 | .943 |
| L=2 | .60 | .765 | .847 | .869 | .767 | .848 | .859 | .771 | .771 | .771 |
| | .90 | .472 | .552 | .564 | .475 | .546 | .556 | .484 | .484 | .484 |
| | .99 | .361 | .390 | .395 | .365 | .391 | .395 | .376 | .376 | .376 |
| | .00 | 1.000 | .857 | .809 | 1.000 | .852 | .802 | 1.000 | .835 | .780 |
| | .30 | .926 | .841 | .804 | .926 | .831 | .792 | .927 | .777 | .727 |
| L=3 | .60 | .703 | .687 | .668 | .705 | .676 | .654 | .708 | .603 | .568 |
| | .90 | .332 | .367 | .367 | .336 | .361 | .360 | .344 | .313 | .302 |
| | .99 | .192 | .211 | .214 | .196 | .212 | .214 | .206 | .203 | .202 |
| | .00 | 1.000 | .764 | .686 | 1.000 | .757 | .676 | 1.000 | .734 | .645 |
| | .30 | .920 | .737 | .671 | .920 | .727 | .657 | .921 | .678 | .598 |
| L=4 | .60 | .678 | .581 | .539 | .679 | .571 | .526 | .682 | .512 | .455 |
| | .90 | .276 | .275 | .267 | .278 | .271 | .261 | .285 | .235 | .218 |
| | .99 | .124 | .135 | .136 | .127 | .136 | .136 | .135 | .130 | .128 |
| | .00 | 1.000 | .702 | .603 | 1.000 | .694 | .592 | 1.000 | .671 | .562 |
| | .30 | .916 | .669 | .582 | .917 | .659 | .570 | .917 | .618 | .518 |
| L=5 | .60 | .666 | .515 | .457 | .666 | .505 | .446 | .669 | .458 | .388 |
| | .90 | .247 | .223 | .210 | .249 | .220 | .205 | .255 | .192 | .171 |
| | .99 | .089 | .095 | .095 | .092 | .096 | .095 | .098 | .092 | .089 |
| | .00 | 1.000 | .659 | .545 | 1.000 | .651 | .535 | 1.000 | .631 | .508 |
| | .30 | .915 | .623 | .522 | .915 | .613 | .510 | .915 | .579 | .467 |
| L=6 | .60 | .658 | .471 | .403 | .659 | .462 | .392 | .661 | .425 | .346 |
| | .90 | .231 | .191 | .174 | .232 | .188 | .169 | .237 | .167 | .143 |
| | .99 | .070 | .072 | .071 | .071 | .072 | .071 | .077 | .069 | .067 |
| | .00 | 1.000 | .629 | .505 | 1.000 | .621 | .495 | 1.000 | .603 | .471 |
| | .30 | .913 | .590 | .480 | .914 | .581 | .468 | .914 | .553 | .432 |
| L=7 | .60 | .654 | .441 | .365 | .654 | .433 | .355 | .656 | .402 | .317 |
| | .90 | .221 | .170 | .150 | .222 | .167 | .146 | .226 | .150 | .125 |
| | .99 | .058 | .057 | .056 | .059 | .057 | .056 | .063 | .055 | .052 |
| | .00 | 1.000 | .606 | .475 | 1.000 | .599 | .466 | 1.000 | .583 | .444 |
| | .30 | .913 | .566 | .449 | .913 | .558 | .438 | .913 | .534 | .407 |
| L=8 | .60 | .651 | .419 | .338 | .651 | .411 | .328 | .652 | .386 | .297 |
| | .90 | .214 | .156 | .133 | .215 | .152 | .129 | .218 | .139 | .112 |
| | .99 | .049 | .047 | .046 | .050 | .047 | .045 | .054 | .045 | .043 |

Table 2.3. Design effects from Dalenius's original equations for the quadratic model under X-optimal allocation for the $\chi^2_1$ distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .00 | 1.437 | 1.127 | .771 | .823 | .645 | .721 | .603 |
| | .30 | 1.329 | 1.047 | .723 | .770 | .608 | .678 | .570 |
| | .60 | 1.006 | .808 | .580 | .613 | .499 | .548 | .472 |
| | .90 | .467 | .408 | .341 | .350 | .317 | .331 | .309 |
| | .99 | .264 | .257 | .250 | .251 | .248 | .249 | .247 |
| L=3 | .00 | 1.572 | 1.173 | .741 | .730 | .514 | .582 | .438 |
| | .30 | 1.440 | 1.077 | .684 | .674 | .477 | .539 | .408 |
| | .60 | 1.044 | .789 | .512 | .505 | .367 | .411 | .318 |
| | .90 | .385 | .309 | .227 | .225 | .183 | .196 | .169 |
| | .99 | .135 | .127 | .119 | .118 | .114 | .116 | .113 |
| L=4 | .00 | 1.646 | 1.201 | .735 | .689 | .457 | .519 | .364 |
| | .30 | 1.503 | 1.098 | .675 | .633 | .421 | .478 | .337 |
| | .60 | 1.075 | .790 | .492 | .463 | .314 | .354 | .254 |
| | .90 | .361 | .276 | .188 | .179 | .135 | .147 | .117 |
| | .99 | .091 | .083 | .073 | .072 | .068 | .069 | .066 |
| L=5 | .00 | 1.697 | 1.221 | .736 | .671 | .428 | .488 | .326 |
| | .30 | 1.548 | 1.115 | .673 | .614 | .393 | .447 | .300 |
| | .60 | 1.100 | .795 | .485 | .443 | .288 | .326 | .222 |
| | .90 | .353 | .263 | .171 | .158 | .112 | .124 | .093 |
| | .99 | .071 | .062 | .052 | .051 | .046 | .047 | .044 |
| L=6 | .00 | 1.739 | 1.239 | .738 | .663 | .413 | .471 | .304 |
| | .30 | 1.585 | 1.130 | .674 | .606 | .378 | .431 | .279 |
| | .60 | 1.123 | .802 | .482 | .434 | .274 | .311 | .204 |
| | .90 | .352 | .257 | .162 | .147 | .100 | .111 | .079 |
| | .99 | .061 | .051 | .041 | .039 | .034 | .035 | .032 |
| L=7 | .00 | 1.776 | 1.255 | .742 | .660 | .404 | .462 | .291 |
| | .30 | 1.618 | 1.144 | .677 | .603 | .369 | .422 | .267 |
| | .60 | 1.144 | .810 | .482 | .430 | .265 | .303 | .193 |
| | .90 | .353 | .254 | .157 | .141 | .094 | .104 | .071 |
| | .99 | .054 | .044 | .034 | .032 | .027 | .028 | .025 |
| L=8 | .00 | 1.812 | 1.271 | .746 | .661 | .399 | .458 | .283 |
| | .30 | 1.650 | 1.158 | .680 | .603 | .364 | .418 | .259 |
| | .60 | 1.165 | .819 | .483 | .428 | .260 | .298 | .186 |
| | .90 | .356 | .253 | .154 | .138 | .088 | .099 | .066 |
| | .99 | .051 | .040 | .029 | .028 | .022 | .024 | .020 |

Table 2.4. Design effects from Dalenius's original equations for the quadratic model under X-optimal allocation for the $\chi^2_2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | 1.244 | 1.009 | .780 | .838 | .724 | .781 | .705 |
| | .30 | 1.158 | .943 | .736 | .788 | .684 | .736 | .667 |
| L=2 | .60 | .899 | .748 | .602 | .639 | .566 | .603 | .554 |
| | .90 | .468 | .423 | .380 | .390 | .369 | .380 | .365 |
| | .00 | .305 | .300 | .295 | .296 | .294 | .295 | .294 |
| | .00 | 1.203 | 1.021 | .751 | .700 | .565 | .593 | .503 |
| | .30 | 1.188 | .941 | .695 | .649 | .526 | .552 | .470 |
| L=3 | .60 | .875 | .702 | .528 | .496 | .410 | .428 | .370 |
| | .90 | .354 | .302 | .251 | .241 | .215 | .221 | .203 |
| | .99 | .156 | .151 | .146 | .145 | .142 | .142 | .141 |
| | .00 | 1.310 | 1.027 | .743 | .633 | .492 | .502 | .408 |
| | .30 | 1.199 | .942 | .683 | .583 | .454 | .464 | .378 |
| L=4 | .60 | .866 | .685 | .504 | .433 | .342 | .349 | .289 |
| | .90 | .311 | .257 | .204 | .183 | .156 | .158 | .140 |
| | .99 | .101 | .096 | .090 | .088 | .085 | .085 | .083 |
| | .00 | 1.318 | 1.030 | .741 | .597 | .453 | .453 | .357 |
| | .30 | 1.204 | .942 | .679 | .548 | .416 | .417 | .329 |
| L=5 | .60 | .862 | .677 | .492 | .400 | .308 | .308 | .246 |
| | .90 | .291 | .236 | .181 | .154 | .126 | .127 | .108 |
| | .99 | .075 | .069 | .064 | .061 | .058 | .058 | .056 |
| | .00 | 1.322 | 1.032 | .740 | .576 | .430 | .424 | .326 |
| | .30 | 1.207 | .942 | .677 | .527 | .394 | .389 | .300 |
| L=6 | .60 | .859 | .673 | .486 | .381 | .288 | .284 | .221 |
| | .90 | .280 | .225 | .169 | .138 | .110 | .109 | .090 |
| | .99 | .061 | .055 | .049 | .046 | .043 | .043 | .041 |
| | .00 | 1.325 | 1.033 | .740 | .562 | .415 | .405 | .307 |
| | .30 | 1.208 | .943 | .676 | .514 | .380 | .371 | .282 |
| L=7 | .60 | .857 | .671 | .483 | .369 | .275 | .268 | .206 |
| | .90 | .273 | .217 | .162 | .128 | .100 | .098 | .079 |
| | .99 | .052 | .046 | .040 | .037 | .034 | .033 | .032 |
| | .00 | 1.327 | 1.034 | .740 | .552 | .405 | .392 | .294 |
| | .30 | 1.209 | .943 | .675 | .504 | .371 | .358 | .269 |
| L=8 | .60 | .856 | .669 | .481 | .361 | .267 | .258 | .195 |
| | .90 | .268 | .213 | .157 | .121 | .093 | .091 | .072 |
| | .99 | .046 | .040 | .034 | .031 | .028 | .027 | .025 |

Table 2.5. Design effects from Dalenius's original equations for the quadratic model under X-optimal allocation for the normal distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .943 | .943 | .943 | .943 | .943 | .943 | .943 |
| L=2 | .60 | .771 | .771 | .771 | .771 | .771 | .771 | .771 |
| | .90 | .484 | .484 | .484 | .484 | .484 | .484 | .484 |
| | .99 | .376 | .376 | .376 | .376 | .376 | .376 | .376 |
| | .00 | 1.061 | 1.061 | 1.057 | .827 | .825 | .749 | .748 |
| | .30 | .982 | .982 | .978 | .769 | .767 | .698 | .697 |
| L=3 | .60 | .745 | .745 | .742 | .595 | .594 | .545 | .544 |
| | .90 | .349 | .349 | .349 | .305 | .305 | .290 | .290 |
| | .99 | .200 | .200 | .200 | .195 | .195 | .194 | .194 |
| | .00 | 1.094 | 1.094 | 1.088 | .737 | .734 | .618 | .616 |
| | .30 | 1.005 | 1.005 | 1.000 | .680 | .678 | .572 | .570 |
| L=4 | .60 | .739 | .739 | .736 | .511 | .509 | .435 | .434 |
| | .90 | .296 | .296 | .295 | .228 | .228 | .206 | .205 |
| | .99 | .129 | .129 | .129 | .122 | .122 | .119 | .119 |
| | .00 | 1.112 | 1.112 | 1.106 | .687 | .684 | .545 | .543 |
| | .30 | 1.019 | 1.019 | 1.013 | .632 | .629 | .503 | .501 |
| L=5 | .60 | .738 | .738 | .734 | .466 | .464 | .375 | .374 |
| | .90 | .270 | .270 | .269 | .189 | .189 | .162 | .162 |
| | .99 | .093 | .093 | .093 | .085 | .085 | .082 | .082 |
| | .00 | 1.124 | 1.124 | 1.117 | .657 | .654 | .502 | .499 |
| | .30 | 1.027 | 1.027 | 1.021 | .603 | .600 | .461 | .459 |
| L=6 | .60 | .738 | .738 | .733 | .439 | .437 | .340 | .338 |
| | .90 | .255 | .255 | .254 | .167 | .166 | .137 | .137 |
| | .99 | .073 | .073 | .073 | .064 | .064 | .061 | .060 |
| | .00 | 1.131 | 1.131 | 1.124 | .638 | .634 | .473 | .471 |
| | .30 | 1.033 | 1.033 | 1.026 | .584 | .581 | .434 | .432 |
| L=7 | .60 | .738 | .738 | .733 | .422 | .420 | .317 | .315 |
| | .90 | .246 | .246 | .245 | .152 | .152 | .121 | .121 |
| | .99 | .060 | .060 | .060 | .050 | .050 | .047 | .047 |
| | .00 | 1.136 | 1.136 | 1.129 | .625 | .621 | .454 | .452 |
| | .30 | 1.037 | 1.037 | 1.030 | .571 | .568 | .416 | .414 |
| L=8 | .60 | .738 | .738 | .733 | .411 | .408 | .302 | .300 |
| | .90 | .240 | .240 | .239 | .143 | .142 | .111 | .110 |
| | .99 | .052 | .052 | .052 | .042 | .042 | .038 | .038 |

Table 2.6. Design effects from Dalenius's original equations for the quadratic model under equal allocation for the $\chi_1^2$ distribution.

| | ρ | C=0 r=0 | C=0 r=1 | C=0 r=2 | C=1 r=1 | C=1 r=2 | C=2 r=1 | C=2 r=2 |
|---|---|---|---|---|---|---|---|---|
| | .00 | 1.330 | 1.067 | .756 | .800 | .644 | .711 | .607 |
| | .30 | 1.232 | .993 | .710 | .750 | .608 | .668 | .574 |
| L=2 | .60 | .938 | .770 | .571 | .599 | .499 | .542 | .475 |
| | .90 | .448 | .398 | .339 | .347 | .317 | .330 | .310 |
| | .99 | .262 | .257 | .251 | .252 | .249 | .250 | .248 |
| | .00 | 1.456 | 1.109 | .723 | .700 | .507 | .564 | .435 |
| | .30 | 1.334 | 1.019 | .668 | .647 | .471 | .523 | .406 |
| L=3 | .60 | .970 | .748 | .501 | .487 | .363 | .399 | .317 |
| | .90 | .363 | .297 | .224 | .219 | .183 | .193 | .169 |
| | .99 | .133 | .126 | .119 | .118 | .114 | .116 | .113 |
| | .00 | 1.528 | 1.137 | .717 | .659 | .449 | .500 | .359 |
| | .30 | 1.396 | 1.040 | .658 | .605 | .414 | .460 | .332 |
| L=4 | .60 | 1.000 | .749 | .480 | .443 | .309 | .341 | .252 |
| | .90 | .339 | .264 | .185 | .174 | .134 | .143 | .117 |
| | .99 | .089 | .081 | .073 | .072 | .068 | .069 | .066 |
| | .00 | 1.578 | 1.157 | .716 | .640 | .419 | .467 | .320 |
| | .30 | 1.439 | 1.057 | .655 | .586 | .385 | .429 | .295 |
| L=5 | .60 | 1.024 | .755 | .472 | .423 | .282 | .313 | .219 |
| | .90 | .331 | .251 | .167 | .153 | .111 | .120 | .092 |
| | .99 | .069 | .060 | .052 | .050 | .046 | .047 | .044 |
| | .00 | 1.615 | 1.173 | .718 | .631 | .403 | .450 | .298 |
| | .30 | 1.472 | 1.070 | .656 | .576 | .369 | .412 | .274 |
| L=6 | .60 | 1.043 | .761 | .469 | .413 | .268 | .298 | .200 |
| | .90 | .328 | .244 | .158 | .141 | .098 | .107 | .078 |
| | .99 | .058 | .049 | .040 | .039 | .034 | .035 | .032 |
| | .00 | 1.644 | 1.186 | .720 | .626 | .393 | .440 | .284 |
| | .30 | 1.498 | 1.081 | .657 | .572 | .360 | .402 | .260 |
| L=7 | .60 | 1.059 | .766 | .468 | .408 | .259 | .288 | .189 |
| | .90 | .328 | .241 | .153 | .135 | .090 | .099 | .070 |
| | .99 | .052 | .043 | .033 | .032 | .027 | .028 | .025 |
| | .00 | 1.668 | 1.197 | .722 | .625 | .387 | .434 | .275 |
| | .30 | 1.520 | 1.091 | .659 | .570 | .354 | .396 | .252 |
| L=8 | .60 | 1.073 | .772 | .468 | .405 | .253 | .283 | .182 |
| | .90 | .329 | .240 | .149 | .131 | .086 | .094 | .064 |
| | .99 | .048 | .038 | .029 | .027 | .022 | .023 | .020 |

Table 2.7. Design effects from Dalenius's original equations for the quadratic model under equal allocation for the $\chi^2_2$ distribution.

| | $\rho$ | C=0 r=0 | C=0 r=1 | C=0 r=2 | C=1 r=1 | C=1 r=2 | C=2 r=1 | C=2 r=2 |
|---|---|---|---|---|---|---|---|---|
| | .00 | 1.207 | .992 | .780 | .834 | .728 | .781 | .710 |
| | .30 | 1.124 | .929 | .736 | .784 | .688 | .736 | .672 |
| L=2 | .60 | .875 | .738 | .602 | .636 | .569 | .603 | .557 |
| | .90 | .461 | .420 | .380 | .390 | .370 | .380 | .366 |
| | .99 | .304 | .300 | .296 | .297 | .295 | .296 | .294 |
| | .00 | 1.260 | 1.003 | .747 | .693 | .565 | .590 | .504 |
| | .30 | 1.158 | .925 | .692 | .643 | .526 | .549 | .471 |
| L=3 | .60 | .854 | .690 | .526 | .492 | .410 | .425 | .371 |
| | .90 | .347 | .299 | .250 | .240 | .215 | .220 | .204 |
| | .99 | .156 | .151 | .146 | .145 | .142 | .142 | .141 |
| | .00 | 1.282 | 1.010 | .738 | .626 | .490 | .498 | .407 |
| | .30 | 1.174 | .926 | .678 | .577 | .453 | .460 | .377 |
| L=4 | .60 | .848 | .674 | .500 | .428 | .341 | .346 | .288 |
| | .90 | .306 | .254 | .203 | .181 | .155 | .157 | .140 |
| | .99 | .101 | .096 | .090 | .088 | .085 | .085 | .084 |
| | .00 | 1.295 | 1.015 | .735 | .590 | .450 | .449 | .355 |
| | .30 | 1.183 | .928 | .673 | .542 | .414 | .413 | .328 |
| L=5 | .60 | .847 | .668 | .488 | .396 | .306 | .305 | .245 |
| | .90 | .287 | .233 | .180 | .153 | .126 | .126 | .108 |
| | .99 | .075 | .069 | .064 | .061 | .058 | .058 | .056 |
| | .00 | 1.302 | 1.019 | .734 | .569 | .427 | .419 | .325 |
| | .30 | 1.188 | .930 | .671 | .521 | .392 | .385 | .299 |
| L=6 | .60 | .846 | .665 | .483 | .377 | .286 | .281 | .220 |
| | .90 | .276 | .222 | .168 | .137 | .110 | .108 | .090 |
| | .99 | .060 | .055 | .049 | .046 | .043 | .043 | .041 |
| | .00 | 1.308 | 1.021 | .734 | .556 | .412 | .401 | .305 |
| | .30 | 1.192 | .032 | .670 | .508 | .378 | .367 | .280 |
| L=7 | .60 | .846 | .663 | .479 | .365 | .273 | .266 | .205 |
| | .90 | .269 | .215 | .161 | .127 | .099 | .097 | .079 |
| | .99 | .051 | .046 | .040 | .037 | .034 | .033 | .032 |
| | .00 | 1.311 | 1.023 | .734 | .547 | .403 | .388 | .292 |
| | .30 | 1.195 | .932 | .670 | .500 | .368 | .355 | .268 |
| L=8 | .60 | .846 | .662 | .477 | .357 | .265 | .256 | .194 |
| | .90 | .265 | .211 | .156 | .120 | .093 | .090 | .072 |
| | .99 | .046 | .040 | .034 | .030 | .028 | .027 | .025 |

Table 2.8. Design effects from Dalenius's original equations for the quadratic model under equal allocation for the normal distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .943 | .943 | .943 | .943 | .943 | .943 | .943 |
| L=2 | .60 | .771 | .771 | .771 | .771 | .771 | .771 | .771 |
| | .90 | .484 | .484 | .484 | .484 | .484 | .484 | .484 |
| | .99 | .376 | .376 | .376 | .376 | .376 | .376 | .376 |
| | .00 | 1.043 | 1.043 | 1.039 | .819 | .817 | .744 | .743 |
| | .30 | .965 | .965 | .962 | .762 | .760 | .694 | .693 |
| L=3 | .60 | .733 | .733 | .731 | .590 | .589 | .542 | .542 |
| | .90 | .346 | .346 | .345 | .304 | .303 | .289 | .289 |
| | .99 | .200 | .200 | .200 | .195 | .195 | .194 | .194 |
| | .00 | 1.070 | 1.070 | 1.065 | .726 | .723 | .611 | .609 |
| | .30 | .984 | .984 | .979 | .670 | .668 | .566 | .564 |
| L=4 | .60 | .724 | .724 | .721 | .504 | .502 | .430 | .429 |
| | .90 | .292 | .292 | .291 | .226 | .226 | .205 | .204 |
| | .99 | .128 | .128 | .128 | .122 | .122 | .119 | .119 |
| | .00 | 1.088 | 1.088 | 1.082 | .675 | .672 | .538 | .536 |
| | .30 | .996 | .996 | .991 | .621 | .618 | .496 | .494 |
| L=5 | .60 | .722 | .722 | .719 | .458 | .456 | .370 | .369 |
| | .90 | .265 | .265 | .264 | .187 | .187 | .161 | .161 |
| | .99 | .093 | .093 | .093 | .085 | .084 | .082 | .082 |
| | .00 | 1.100 | 1.100 | 1.094 | .645 | .642 | .494 | .492 |
| | .30 | 1.005 | 1.005 | 1.000 | .592 | .589 | .454 | .452 |
| L=6 | .60 | .722 | .722 | .719 | .432 | .430 | .335 | .333 |
| | .90 | .251 | .251 | .250 | .164 | .164 | .136 | .135 |
| | .99 | .073 | .073 | .072 | .063 | .063 | .060 | .060 |
| | .00 | 1.108 | 1.108 | 1.102 | .627 | .623 | .466 | .464 |
| | .30 | 1.012 | 1.012 | 1.006 | .574 | .571 | .428 | .426 |
| L=7 | .60 | .723 | .723 | .719 | .415 | .413 | .312 | .311 |
| | .90 | .242 | .242 | .241 | .150 | .150 | .120 | .119 |
| | .99 | .060 | .060 | .060 | .050 | .050 | .047 | .047 |
| | .00 | 1.115 | 1.115 | 1.108 | .614 | .611 | .447 | .445 |
| | .30 | 1.017 | 1.017 | 1.011 | .562 | .559 | .410 | .408 |
| L=8 | .60 | .724 | .724 | .720 | .404 | .402 | .297 | .296 |
| | .90 | .236 | .236 | .235 | .141 | .140 | .109 | .109 |
| | .99 | .052 | .052 | .051 | .042 | .042 | .038 | .038 |

Table 2.9. Design effects from Dalenius's original equations for the exponential model under proportional allocation for the $\chi_1^2$, $\chi_2^2$ and normal distributions.

| | ρ | $\chi_1^2$ C=1 | $\chi_1^2$ C=2 | $\chi_2^2$ C=1 | $\chi_2^2$ C=2 | Normal C=1 | Normal C=2 |
|---|---|---|---|---|---|---|---|
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .962 | .963 | .964 | .965 | .974 | .976 |
| L=2 | .60 | .835 | .839 | .838 | .842 | .840 | .850 |
| | .90 | .553 | .564 | .552 | .563 | .520 | .530 |
| | .99 | .391 | .396 | .393 | .397 | .380 | .381 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .941 | .941 | .942 | .942 | .945 | .947 |
| L=3 | .60 | .752 | .755 | .752 | .755 | .727 | .734 |
| | .90 | .387 | .394 | .385 | .391 | .347 | .348 |
| | .99 | .213 | .217 | .215 | .218 | .206 | .206 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .926 | .926 | .926 | .926 | .916 | .918 |
| L=4 | .60 | .700 | .701 | .697 | .698 | .650 | .649 |
| | .90 | .303 | .305 | .300 | .302 | .263 | .257 |
| | .99 | .137 | .139 | .138 | .139 | .132 | .131 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .915 | .915 | .913 | .913 | .890 | .891 |
| L=5 | .60 | .663 | .663 | .658 | .657 | .595 | .586 |
| | .90 | .253 | .252 | .250 | .247 | .216 | .204 |
| | .99 | .097 | .098 | .098 | .098 | .094 | .092 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .907 | .907 | .903 | .903 | .867 | .865 |
| L=6 | .60 | .636 | .634 | .628 | .625 | .554 | .537 |
| | .90 | .222 | .217 | .217 | .212 | .186 | .171 |
| | .99 | .074 | .074 | .074 | .073 | .071 | .069 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .900 | .899 | .895 | .894 | .846 | .842 |
| L=7 | .60 | .615 | .611 | .605 | .600 | .523 | .498 |
| | .90 | .199 | .192 | .195 | .186 | .167 | .149 |
| | .99 | .059 | .058 | .059 | .058 | .057 | .054 |
| | .00 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | .30 | .894 | .893 | .888 | .887 | .826 | .820 |
| L=8 | .60 | .598 | .592 | .586 | .578 | .498 | .467 |
| | .90 | .183 | .174 | .179 | .168 | .153 | .133 |
| | .99 | .049 | .048 | .048 | .047 | .047 | .044 |

Table 2.10. Design effects from Dalenius's original equations for the exponential model under X-optimal allocation for the $\chi_1^2$, $\chi_2^2$ and normal distributions.

| | | $\chi_1^2$ | | $\chi_2^2$ | | Normal | |
|---|---|---|---|---|---|---|---|
| | ρ | C=1 | C=2 | C=1 | C=2 | C=1 | C=2 |
| L=2 | .00 | .590 | .530 | .657 | .616 | 1.000 | 1.000 |
| | .30 | .564 | .510 | .632 | .595 | .974 | .976 |
| | .60 | .480 | .442 | .547 | .522 | .840 | .850 |
| | .90 | .320 | .311 | .375 | .371 | .520 | .530 |
| | .99 | .249 | .248 | .296 | .296 | .380 | .381 |
| L=3 | .00 | .499 | .421 | .558 | .494 | .947 | .910 |
| | .30 | .470 | .397 | .525 | .466 | .894 | .863 |
| | .60 | .369 | .319 | .419 | .379 | .691 | .674 |
| | .90 | .189 | .175 | .224 | .214 | .332 | .327 |
| | .99 | .115 | .114 | .143 | .142 | .198 | .198 |
| L=4 | .00 | .458 | .370 | .511 | .433 | .900 | .839 |
| | .30 | .430 | .348 | .474 | .403 | .831 | .776 |
| | .60 | .324 | .267 | .361 | .312 | .602 | .563 |
| | .90 | .140 | .124 | .164 | .150 | .248 | .233 |
| | .99 | .068 | .067 | .086 | .085 | .123 | .121 |
| L=5 | .00 | .437 | .342 | .482 | .396 | .863 | .783 |
| | .30 | .409 | .320 | .445 | .366 | .784 | .711 |
| | .60 | .301 | .239 | .330 | .275 | .546 | .491 |
| | .90 | .117 | .099 | .133 | .117 | .204 | .183 |
| | .99 | .046 | .045 | .059 | .057 | .086 | .084 |
| L=6 | .00 | .425 | .325 | .464 | .372 | .834 | .740 |
| | .30 | .396 | .303 | .426 | .342 | .747 | .661 |
| | .60 | .287 | .222 | .310 | .251 | .508 | .442 |
| | .90 | .104 | .085 | .116 | .098 | .178 | .153 |
| | .99 | .035 | .033 | .043 | .042 | .064 | .062 |
| L=7 | .00 | .424 | .318 | .448 | .352 | .808 | .705 |
| | .30 | .388 | .292 | .412 | .324 | .718 | .623 |
| | .60 | .279 | .211 | .296 | .234 | .482 | .407 |
| | .90 | .096 | .076 | .105 | .086 | .161 | .134 |
| | .99 | .028 | .025 | .034 | .032 | .051 | .048 |
| L=8 | .00 | .412 | .306 | .440 | .341 | .790 | .678 |
| | .30 | .383 | .284 | .402 | .311 | .695 | .591 |
| | .60 | .273 | .203 | .287 | .222 | .462 | .380 |
| | .90 | .091 | .070 | .098 | .078 | .150 | .121 |
| | .99 | .023 | .021 | .028 | .026 | .042 | .039 |

Table 2.11. Design effects from Dalenius's original equations for the exponential model under equal allocation for the $\chi_1^2$, $\chi_2^2$ and normal distributions.

| | ρ | $\chi_1^2$ C=1 | $\chi_1^2$ C=2 | $\chi_2^2$ C=1 | $\chi_2^2$ C=2 | Normal C=1 | Normal C=2 |
|---|---|---|---|---|---|---|---|
| L=2 | .00 | .591 | .535 | .663 | .623 | 1.000 | 1.000 |
| | .30 | .564 | .514 | .637 | .601 | .974 | .976 |
| | .60 | .481 | .447 | .551 | .528 | .840 | .850 |
| | .90 | .321 | .313 | .377 | .373 | .520 | .530 |
| | .99 | .250 | .249 | .297 | .297 | .380 | .381 |
| L=3 | .00 | .496 | .422 | .561 | .499 | .947 | .916 |
| | .30 | .466 | .398 | .527 | .471 | .894 | .868 |
| | .60 | .367 | .320 | .421 | .383 | .690 | .676 |
| | .90 | .188 | .176 | .225 | .215 | .331 | .326 |
| | .99 | .115 | .114 | .143 | .143 | .198 | .198 |
| L=4 | .00 | .453 | .369 | .512 | .437 | .899 | .843 |
| | .30 | .425 | .347 | .475 | .407 | .829 | .781 |
| | .60 | .321 | .266 | .362 | .315 | .599 | .564 |
| | .90 | .139 | .124 | .164 | .151 | .246 | .232 |
| | .99 | .068 | .067 | .086 | .085 | .123 | .121 |
| L=5 | .00 | .431 | .340 | .483 | .399 | .860 | .788 |
| | .30 | .403 | .319 | .445 | .369 | .781 | .715 |
| | .60 | .296 | .237 | .330 | .276 | .542 | .491 |
| | .90 | .116 | .098 | .133 | .118 | .202 | .182 |
| | .99 | .046 | .045 | .059 | .057 | .086 | .084 |
| L=6 | .00 | .418 | .323 | .464 | .374 | .831 | .744 |
| | .30 | .389 | .301 | .426 | .344 | .744 | .665 |
| | .60 | .282 | .220 | .309 | .252 | .503 | .441 |
| | .90 | .102 | .084 | .116 | .099 | .176 | .152 |
| | .99 | .034 | .033 | .043 | .042 | .064 | .062 |
| L=7 | .00 | .416 | .315 | .448 | .354 | .804 | .707 |
| | .30 | .380 | .288 | .412 | .326 | .715 | .625 |
| | .60 | .273 | .208 | .296 | .235 | .477 | .405 |
| | .90 | .094 | .075 | .105 | .086 | .159 | .133 |
| | .99 | .027 | .025 | .034 | .032 | .051 | .048 |
| L=8 | .00 | .403 | .301 | .440 | .342 | .786 | .680 |
| | .30 | .374 | .279 | .402 | .313 | .692 | .593 |
| | .60 | .267 | .200 | .286 | .223 | .457 | .379 |
| | .90 | .089 | .069 | .097 | .078 | .148 | .120 |
| | .99 | .023 | .021 | .028 | .026 | .042 | .039 |

## 2.3. Optimum Allocation and the Auxiliary Variable

When the stratum boundaries are suggested a priori, the only variable factors are the method of allocation and the number of strata. And, when estimating the population mean, it is known that Neyman allocation provides the smallest variance with the same size and number of strata fixed. When little is known about the functional relationships between the study and auxiliary variables, Neyman allocation is approximated by X-optimal allocation. When (1.9) is the regression model between the two variables, Neyman allocation can be written as

$$(2.11) \qquad \frac{n_h}{n} = \frac{W_h \sigma_{hy}}{\sum_{k=1}^{L} W_k \sigma_{ky}} = \frac{W_h \sqrt{\sigma^2_{hg} + \sigma^2_{he}}}{\sum_{k=1}^{L} W_k \sqrt{\sigma^2_{kg} + \sigma^2_{ke}}} \ .$$

Let us examine the form of (2.11) over the range of regression and residual variation functions we are using to model our universe, examining both when X-optimal allocation offers a good estimate to Neyman allocation, as well as what improvements to X-optimal allocation might be possible when marginal information is available.

Linear: Combining (2.11) and (2.2) and simplifying, Neymal allocation can be written as

$$(2.12) \qquad \frac{n_h}{n} = \frac{W_h \sqrt{\rho^2 \frac{\sigma^2_{hx}}{\sigma^2_x} + (1-\rho^2) \frac{\sigma^2_{he}}{\sigma^2_e}}}{\sum_{k=1}^{L} W_k \sqrt{\rho^2 \frac{\sigma^2_{kx}}{\sigma^2_x} + (1-\rho^2) \frac{\sigma^2_{ke}}{\sigma^2_e}}} \ .$$

Thus, as $\rho \to 1$, $\frac{n_h}{n} \to W_h\sigma_{hx}/\Sigma_{k=1}^{L} W_k\sigma_{kx}$, X-optimal allocation. Recalling that we defined $\phi(x)=x^r$, r=0,1 and 2, we also note that as $\rho \to 0$,

$$(2.13) \qquad \frac{n_h}{n} \to \frac{W_h\sigma_{he}}{\sum_{k=1}^{L} W_k\sigma_{ke}} \quad \propto \quad \begin{cases} W_h & r=0 \\ W_h\sqrt{\mu_{hx}} & r=1 \\ W_h\sqrt{\mu_{hx}^2+\sigma_{hx}^2} & r=2 \end{cases}$$

which we might refer to as "E-optimal" allocation. Thus, for the linear model, Neyman allocation is a combination of X- and E-optimal allocations. It is interesting that Neyman allocation becomes a combination of X-optimal allocation and proportional allocation only when the residual variation is homogeneous across the strata.

Suppose we have the marginal information of the approximate magnitude of $\rho$. It is one of the mathematical tractabilities of the linear model that knowing this and the distribution of X leaves just the form of the residual variation unknown. So, in practice, determining the sample allocations by (2.12) for each residual variation form would allow one to see the effect the choice of $\phi(x)$ has on the allocation, and perhaps suggest an improvement over the use of X-optimal allocation.

Quadratic: Combining (2.11) and (2.3) and simplifying, Neyman allocation for this model can be written as

$$\frac{n_h}{n} = \frac{W_h\sqrt{\pi^2\frac{\text{var}[G(X)|h]}{\text{var}[G(X)]} + (1-\pi^2)\frac{\sigma^2_{he}}{\sigma^2_e}}}{\sum_{k=1}^{L} W_k\sqrt{\pi^2\frac{\text{var}[G(X)|k]}{\text{var}[G(X)]} + (1-\pi^2)\frac{\sigma^2_{ke}}{\sigma^2_e}}} \tag{2.14}$$

where G(x) is defined in (2.3). Unlike the linear model, knowledge of the approximate magnitude of $\rho$ does not leave just the form of the residual unknown--the non-linearity parameter, C, is also unknown. However, if $\rho$'s magnitude is high enough, we might conjecture that $\rho\to\pi$, in which case $C\to 0$, and (2.14) can be approximated by (2.12). But, when $\rho$ is only small to moderate, it is difficult to be able to decide upon the magnitude of C, since the higher the value of $\pi$, the greater the non-linearity. Recalling that $C=\frac{\pi^2-\rho^2}{1-\pi^2}$ with $\rho\le\pi\le 1$, in Table 2.12 we list the values of C for a range of $\rho$ and $\pi$. Thus, knowing only the magnitude of $\rho$ tells us little about the size of C when $\rho$ is small to moderate. For instance, when $\rho=0$ and $C=2$, say, placing a lot of weight on E-optimal allocation, as the magnitude of $\rho$ would suggest in (2.12), compares unfavorably with the actual weight, $1-\pi^2=1/3$.

Table 2.12. Values of C.

| | $\pi$ | | | | |
|---|---|---|---|---|---|
| $\rho$ | .00 | .30 | .60 | .90 | .99 |
| .00 | 0 | .10 | .56 | 4.3 | 49 |
| .30 | | 0 | .42 | 3.8 | 45 |
| .60 | | | 0 | 2.4 | 31 |

Exponential: Combining (2.11) and (2.4) and simplifying, we can write Neyman allocation as

$$(2.15) \qquad \frac{n_h}{n} = \frac{W_h\sqrt{\pi^2\frac{\sigma^2_{hu}}{\sigma^2_u} + (1-\pi^2)\frac{\sigma^2_{he}}{\sigma^2_e}}}{\sum_{k=1}^{L} W_k\sqrt{\pi^2\frac{\sigma^2_{ku}}{\sigma^2_u} + (1-\pi^2)\frac{\sigma^2_{ke}}{\sigma^2_e}}}$$

where $U=\exp(\gamma X)$. Like the quadratic model, when $\rho$ is moderately large, we can conjecture $\rho\to\pi$, and thus approximate (2.15) with (2.12), since $\frac{\sigma^2_{hu}}{\sigma^2_u} \to \frac{\sigma^2_{hx}}{\sigma^2_x}$ as $\rho\to\pi$. However, unlike the quadratic model which does not in general reduce to something useful when $\rho\to 0$, since $\frac{\sigma^2_{hu}}{\sigma^2_u} \to 0$ in this situation for the exponential model, (2.15) reduces to the form

$$(2.16) \qquad n_h \propto W_h\sqrt{\mu^2_{hx} + \sigma^2_{hx}} \quad .$$

Thus, knowledge that g(x) is exponential allows us, for small enough $\rho$, to approximate (2.15) with (2.16).

## 2.4. Summary

One approximation to the Dalenius-Gurney equations when little is known about the functional relationship between the study and auxiliary variables is to use Dalenius's original equations with the auxiliary variable. With this technique, the variance of the stratified random sampling estimate of the population mean offers an improvement over the use of simple random sampling when the functional relationship between the study and auxiliary variables explains more than 40 percent of the variation of Y. Thus, use of this approximate procedure can be considered justified as an improvement over simple random sampling when the correlation between X and Y is only moderate. However, when the functional relationship explains less than 40 percent of the variation of Y, the use of stratification with X-optimal and equal allocations can actually provide variances of $\hat{\mu}$ which are greater than the variance under simple random sampling.

When the auxiliary variable is normally distributed, and the two variables are related by a linear or quadratic model, proportional allocation affords almost as much reduction in variance as does X-optimal allocation, and in a number of cases, actually offers more. This suggests that the simpler equations under proportional allocation can be used with at worst marginal loss of reduction instead of the more cumbersome X-optimal allocation equations. Additionally, for the two skewed distributions, except for the most heterogeneous residual variation, proportional allocation under the linear model in general offered better reductions in variance than X-optimal allocation. In all other cases, X-optimal allocation provided better improvements over simple random sampling than did proportional allocation.

A comparison between equal and X-optimal allocations indicated that except for the highest correlation, equal allocation consistently offered a greater reduction in variance than use of X-optimal allocation, though the advantage at times was only marginal. This would support Cochran's findings that equal allocation offers an excellent approximation to Neyman allocation in practical situations.

When the boundaries are fixed, X-optimal allocation offers a close approximation to Neyman allocation only when the correlation between the two variables is high. When the correlation is small to moderate, Neyman allocation is a function of the residual variation as well. Knowledge of the magnitude of $\rho$ can potentially improve X-optimal allocation's approximation of Neyman allocation when the underlying model relating the two variables is known to be linear, or when the magnitude of the correlation is high enough so that the linear model offers a reasonable approximation to the actual non-linear model. Finally, if $\rho$ is small enough and the regression model can be assumed to be exponential, then the effect of X-optimal allocation disappears, leaving Neyman allocation a function of just the residual variation.

# III. APPROXIMATIONS TO THE DALENIUS-GURNEY EQUATIONS

When information is available about the functional relationship between the study and auxiliary variables, use of the Dalenius-Gurney equations to determine the stratum boundaries offers the greatest reduction in the variance of $\hat{\mu}$ for specified method of allocation and number of strata. However, because of the almost prohibitively unwieldy nature of these equations, their solutions can be quite difficult to obtain. Hence, in this chapter we will examine several approximate systems of equations whose roots theoretically offer almost the degree of reduction as the exact boundaries, but whose solution is much simpler.

To consider this question, we must decide upon a method of comparing the optimum boundaries computed from the Dalenius-Gurney equations with those boundaries computed from the various approximations. Two methods of comparison have been used in the literature. The first is a boundary-versus-boundary approach, comparing how close the approximate boundaries are to the optimum. The second is a variance-versus-variance approach, considering the percentage increase in the variance generated by the approximate boundaries compared to the variance provided by the optimum boundaries. The first approach is clearly inadequate for our purposes, since our interest is less in the boundaries than in the reduction in variance the use of these boundaries provides. The second approach is also not entirely to our liking, since it is in percentage units of the smallest variance, and for it to be entirely useful, it is necessary to know the value of the smallest variance.

However, we would like to use the variances of $\hat{\mu}$ from the two sets of stratum boundaries in our measure. In the previous chapter, we used the design effect to judge how that approach performed against the variance under simple random sampling. The units of the design effect, the percentage reduction compared with simple random sampling, represents a practical scale for any stratification technique; we would again like to employ the design effect. Since the ratio of two design effects eliminates the units of this measure, in this chapter we will use as our mechanism of comparison the additive difference between the design effects of the variances resulting from the exact and approximate boundaries.

In the first three sections of this chapter, we introduce the three approximation techniques we will consider. In the fourth section, we will examine the results of comparing the design effects of these approximations with the optimum design effects.

## 3.1. The Cum H(x) Rules

The first techniques of interest to us are the cum H(x) rules, (1.8), that have appeared in the literature. Singh and Sukhatme (1969) proposed such a rule for both Neyman and proportional allocations, with

$$(3.1) \qquad H(x) = \begin{cases} \sqrt[3]{\dfrac{\phi'(x)^2 + 4\phi(x)g'(x)^2}{\phi(x)^{3/2}}\, f_X(x)} & \text{for Neyman allocation} \\[2ex] \sqrt[3]{g'(x)^2\, f_X(x)} & \text{for proportional allocation.} \end{cases}$$

Singh and Parkash (1975), using the same method of derivation, proposed a cum H(x) rule for equal allocation, with

$$(3.2) \qquad H(x) = \sqrt{\phi(x)}\, f_X(x)$$

It might be noted that the form of g(x) does not affect Singh-Parkash's cum H(x) rule at all.

Finally, Singh (1975), employing a second method of derivation, generated another cum H(x) rule approximation to the Dalenius-Gurney equations under Neyman allocation, with

$$(3.3) \qquad H(x) = \sqrt{f_X(x)\sqrt{g'(x)^2 + \Theta\phi(x)}}\ , \qquad \Theta = \frac{12L^2}{(b-a)^2}$$

Singh's purpose in deriving this additional approximation was to generate a cum H(x) rule under Neyman allocation which would reduce to the cum$\sqrt{f(x)}$ rule when the functional relationship between X and Y was a simple linear regression model; Singh-Sukhatme's cum H(x) rule under Neyman allocation ((3.1)) reduces to a cum $\sqrt[3]{f(x)}$ rule.

Under proportional allocation, the analog of Neyman allocation's cum$\sqrt{f(x)}$ rule is the cum $\sqrt[3]{f(x)}$ rule. Presumably, the reason Singh did not generate a cum H(x) rule analogous to (3.3) under proportional allocation that would reduce to this cum $\sqrt[3]{f(x)}$ rule under the simple linear regression model was that Singh-Sukhatme's cum H(x) rule under proportional allocation already possessed this property. However, this rule is a function of the form of the residual variation only through the values of the parameters in g(x). Consequently, to complete this series of cum H(x) rules, one additional approximation is introduced.

Following the development of Singh in his 1975 paper, an analogous cum H(x) rule under proportional allocation was derived, with

$$(3.4) \qquad H(x) = \sqrt[3]{f_X(x)\ (g'(x)^2 + \Theta\phi(x))} \quad , \quad \Theta = \frac{12L^2}{(b-a)^2} \quad ,$$

which also reduces to the cum $\sqrt[3]{f(x)}$ rule when the functional relationship between X and Y is a simple linear regression model. However, it differs from Singh-Sukhatme's cum H(x) rule for proportional allocation by the addition of the residual variation term, thus potentially offering more sensitivity to the form of the underlying regression model. The derivation of this result is in Appendix B.

## 3.2. Linearization of Regression Model

The second technique we will consider uses the Dalenius-Gurney equations, but approximates the regression function relating the two variables by a simple linear regression model. While the cum H(x) rules could be used when we have good information about the true model, but are seeking a more practical way of generating the stratum boundaries, this linearization of the regression model method of approximation might arise when we did not have a complete picture of the g(x) function. Consequently, we could assume the model was a simple linear regression model with homoscedastic residual variation, and use what information is available to estimate the slope parameter. So, while the first method is designed to bypass the difficulties associated with seeking a global minimum on the troublesome surfaces generated by the true regression model (see Appendix C), the second method instead reduces the degree of the regression model, making the exact equations easier to solve.

Let us derive this approximation technique. Recall that we have restricted ourselves in this chapter to the quadratic model. So, starting with the underlying relationship between the two variables

$$y = \alpha + \beta x + \gamma x^2 + e$$

$x \sim f_X(x)$, and $E[e|x]=0$ and $var[e|x]=\phi(x)>0$, our interest is in approximating this model by

$$y = A + Bx + z$$

with $x \sim f_X(x)$, $E[Z|x]=0$ and $var[Z|x]=1, \forall x$. Under this approximate model, the Dalenius-Gurney equations can be written as

$$(3.5) \qquad x_h - \frac{\mu_{hx} + \mu_{kx}}{2} = 0 \qquad k=h+1,\ h=1,2,\ldots,L-1,$$

for proportional allocation,

$$(3.6) \qquad \frac{B^2(x_h-\mu_{hx})^2 + B^2\sigma^2_{hx} + 2}{\sqrt{B^2\sigma^2_{hx} + 1}} - \frac{B^2(x_h - \mu_{kx})^2 + B^2\sigma^2_{kx} + 2}{\sqrt{B^2\sigma^2_{kx} + 1}} = 0$$

$k=h+1$, $h=1,2,\ldots,L-1$ for Neyman allocation, and

$$(3.7) \qquad W_h(B^2(x_h - \mu_{hx})^2 + B^2\sigma^2_{hx} + 2) - W_k(B^2(x_h - \mu_{kx})^2 + B^2\sigma^2_{kx} + 2) = 0$$

$k=h+1$, $h=1,2,\ldots,L-1$, for equal allocation. These equations, though not as simple as Dalenius's original equations (with the exception of proportional allocation), are much simpler to solve than the full Dalenius-Gurney equations; the surfaces of (3.5)-(3.7) are globally convex in nature, and thus have one minimum value.

It remains to determine the value of the slope parameter. We chose to estimate A and B of the approximation by minimizing the expected squared error, as we did in Appendix A. Thus, $\hat{A}$ and $\hat{B}$ are defined by

$$\min_{A,B} E[Z^2] = \min_{A,B} E[(Y-A-BX)^2] = E[(Y-\hat{A}-\hat{B}X)^2] = E[(\alpha+\beta X+\gamma X^2+e-\hat{A}-\hat{B}X)^2] \ .$$

The resulting estimate of B is

$$\hat{B} = \beta + \gamma \frac{\mathrm{cov}[X,X^2]}{\sigma_x^2} \ .$$

It is interesting to note that if we expand $g(x)=\alpha+\beta x+\gamma x^2$ in a Taylor series about $m=E[X]$ and use $\tilde{B}=g'(m)=\beta+2\gamma m$ as an estimate of B, then

$$\tilde{B} = \hat{B} - \gamma \frac{E[(X-m)^3]}{\sigma_x^2} \ .$$

So, the minimum expected squared error estimate is identical to the truncated Taylor series estimate when the underlying distribution of X is symmetric. The two estimates become more and more dissimilar as the degree of skewness in the population increases.

Thus, our proposed approximation consists of using the simpler forms of the Dalenius-Gurney equations, (3.5)-(3.7), with the information about the functional relationship between the two variables contained in the estimate of $\hat{B}$.

## 3.3. Linearized Cum H(x) Rule

The final technique we will consider consists of combining the two previous approaches. In particular, we chose to approximate the functional relationship between the two variables with a simple linear regression model, and then apply one of the cum H(x) rules to approximate the resulting Dalenius-Gurney equations which we derived in the last section. The resulting cum H(x) rules are

$$H(x) = \begin{cases} \sqrt[3]{f(x)} & \text{under proportional allocation ((3.1) and (3.4)) and Neyman allocation ((3.1))} \\ \sqrt{f(x)} & \text{under Neyman allocation ((3.3))} \\ f(x) & \text{under equal allocation ((3.2))} \end{cases}$$

It is interesting that this linearized cum H(x) rule approach depends only upon the form of the distribution of the auxiliary variable, and is not affected by the regression model relating the two variables. Thus, we have the most insensitive approach of all of the approximations we have considered. But, at the same time, we have perhaps the simplest systems of equations to solve.

Also note that when we apply this linearized cum H(x) rule approach, the resulting systems of equations reduce to those proposed by Dalenius and Hodges (1957) for Neyman allocation, and by Thomsen (1976) for proportional allocation. This offers us the opportunity to see how well these rules fare beyond the scope for which they were designed.

## 3.4. Empirical Comparisons

Let us now consider the comparisons themselves. These are listed in Tables 3.2-3.28, that are at the end of this section, starting on page 71. Because of the number of these tables, we offer an index to them in Table 3.1.

Table 3.1. Index to Tables of Comparisons of Proposed Approximations to the Dalenius-Gurney Equations.

| Tables | Allocation | Proposed Approximation |
|---|---|---|
| 3.2-3.4 | proportional | Singh's cum H(x): $\sqrt[3]{f_X(x)(g'(x)^2+\Theta\phi(x))}$ |
| 3.5 | proportional | Singh-Sukhatme's cum H(x): $\sqrt[3]{f_X(x)g'(x)^2}$ |
| 3.6 | proportional | Linearization: see (3.5) |
| 3.7 | proportional | Linearized cum H(x): $\sqrt[3]{f_X(x)}$ |
| 3.8-3.10 | Neyman | Singh's cum H(x): $\sqrt{f_X(x)\sqrt{g'(x)^2+\Theta\phi(x)}}$ |
| 3.11-3.13 | Neyman | Singh-Sukhatme's cum H(x): see (3.1) |
| 3.14-3.16 | Neyman | Linearization: see (3.6) |
| 3.17-3.19 | Neyman | Linearized Singh's cum H(x): $\sqrt{f_X(x)}$ |
| 3.20-3.22 | Neyman | Linearized Singh-Sukhatme's cum H(x): $\sqrt[3]{f_X(x)}$ |
| 3.23-3.25 | equal | Singh-Parkash's cum H(x): $f_X(x)\sqrt{\phi(x)}$ |
| 3.26-3.28 | equal | Linearization: see (3.7) |

Note that the boundaries under proportional allocation for the linearization approach (3.5) and the cum $\sqrt[3]{f_X(x)}$ rule are independent of the value of the residual parameter, r. Additionally, since $g'(x)=\beta+2\gamma x$, Singh-Sukhatme's cum H(x) rule depends upon r only through the value of $\sigma_e^2$ (see page 33), making the boundaries determined by this approximation also independent of r, since $\sigma_e^2$ can be factored out of (3.1) altogether. Thus, for proportional allocation, it is only Singh's cum H(x) rule which generates approximate boundaries that are not independent of the residual parameter.

For our investigation, the range of parameters, $\rho$, C and r, we will consider are subject to two restrictions. First, when $\rho$=C=r=0, there is no functional relationship between X and Y, and thus we will exclude this case throughout the comparisons. Second, recall that for the linearization approach, the slope parameter B is estimated by

$$\hat{B} = \beta + \gamma\sqrt{\frac{C}{D_x}} = \frac{\sigma_e}{\sigma_x}\frac{\rho}{\sqrt{1-\rho^2}}\sqrt{C+1} \ .$$

Thus, when $\rho$=0, $\hat{B}$=0, regardless of the values of the parameters C or r. Since $\rho$ is the only model parameter under the linearization approach, when $\rho$=0, no model is assumed to be present between X and Y. We will exclude these cases from the comparisons of the linearization approach.

In judging how well an approximation performs, it is difficult to develop an objective gauge, since the decision of when an approach offers a reasonable approximation is highly subjective. With knowledge of the purpose of the survey as well as the nature of the functional relationship between X and Y, a relevant criterion can be constructed. However, in the absence of such information, let us decide to use the poorest fit of an approximation over a specified range of parameters as our criterion. Since the full comparisons are listed in Tables 3.2-3.28, this is designed to provide a rough overview of the performances of the approximations only. We will consider the "poorest fit" over two ranges of stratum numbers: for $L \geq 4$, representing the range of stratum numbers we would expect to find in practice; and $L \geq 6$, representing the range of "large" stratum numbers.

In the following three sections, we will examine the proposed approximations under proportional, Neyman and equal allocations, respectively.

3.4.1. Proportional Allocation

For Singh's cum H(x) rule (Tables 3.2-3.4), we note that for the linear model (C=0), the approximation is quite good; for $L \geq 4$, the largest table value is .009, while for $L \geq 6$, the largest difference is .003. Note that while there is an improvement in the approximation as $\rho \to 0$, it is the nature of proportional allocation to algebraically approach 1 as the correlation between X and Y approaches 0, for both the optimum and the approximate design effects under the linear model. For the non-linear model when C=1, the fit is fair; for $L \geq 4$, the largest difference is .059, while the largest value becomes .024 for $L \geq 6$. And, for the non-linear model when C=2, the largest table value is .081 for $L \geq 4$, and it is .033 for $L \geq 6$. Note that after a large enough L, when $C \neq 0$, the fit improves as $\rho$ increases.

Singh-Sukhatme's cum H(x) rule (Table 3.5) offers about the same fit as did Singh's rule, though marginally worse. For the linear model, since both cum H(x) rules reduce to the cum $\sqrt[3]{f(x)}$ rule, the fits are identical. For C=1, the largest difference is .064 for $L \geq 4$, while it is .027 for $L \geq 6$. When C=2, the largest differences are .085 and .036 for $L \geq 4$ and $L \geq 6$, respectively. Again, note that after a large enough L, the fit improves as $\rho$ increases.

For the linearization technique (Table 3.6), when C=0, the two systems of equations are identical. For $C \neq 0$, the fit is only fair at best, being poor for small to moderate correlation and somewhat improved as $\rho$ becomes high. For C=1, the largest difference when $L\geq4$ is .169, while it is .109 for $L\geq6$. For C=2, the two values are .213 for $L\geq4$, and .142 for $L\geq6$.

Finally, let us look at the cum $\sqrt[3]{f(x)}$ rule, which is the linearized cum H(x) rule for both Singh's and Singh-Sukhatme's approximations (Table 3.7). When C=0, the fit is described in the above discussion of Singh's cum H(x) rule. Interestingly enough, for $C\neq0$, this approximation is actually superior to the linearization approach. For C=1, when $L\geq4$, the largest difference in the table is .100, while for $L\geq6$, the largest difference is .065. When C=2, for $L\geq4$, the largest difference becomes .132, and .085 is the largest value when $L\geq6$. This approximation also improves in fit as the value of the correlation increases.

### 3.4.2. Neyman Allocation

For Singh's cum H(x) rule (Tables 3.8-3.10), when the study and auxiliary variables are related by a linear model, the approximation fits very well. The worst fit is .010 when $L\geq4$, and it drops to .005 for $L\geq6$. Note that this worst fit occurs when $\rho=0$. If we consider just $\rho>0$, then the worst fit becomes .003 for both $L\geq4$ and $L\geq6$. When the model is non-linear, with C=1, the fit is still quite good, with the worst difference for $L\geq4$ being .043, and the worst for $L\geq6$ being .023. These values occurred with the normal distribution; the skewed distributions both showed much better fits. For C=2, the fit gets a little worse, though it is still reasonable. The worst value (again with respect to the normal distribution) for $L\geq4$ is .053, while it is .026

for $L \geq 6$. Again note these worst fits occurred when $\rho=0$; the fit improves for $L \geq 4$ as $\rho$ increases, making this approximation very good for even moderate correlation values.

For Singh-Sukhatme's cum H(x) rule (Tables 3.11-3.13), when C=0, we note that the fit is not as good as that of Singh's cum H(x) rule. In particular, when $L \geq 4$, the worst value is .025, while it is .010 for $L \geq 6$, though the worst fits in this approximation are at the highest correlation value. For the non-linear model, when C=1, the fit is slightly worse than Singh's rule, though still good, with the worst fit when $L \geq 4$ being .057, and the worst fit being .023 for $L \geq 6$. When C=2, the fit is still a little more ragged than Singh's, with the worst fit being .080 for $L \geq 4$, and .031 for $L \geq 6$ (these worst fits again occur with the normal distribution). So, while this approximation does offer a good fit, it does not offer as good a fit as does Singh's approximation.

For the linearization approach (Tables 3.14-3.16), when C=0, the fit is quite good, the worst fit being .003 when $L \geq 4$, and only .002 when $L \geq 6$. However, for $C \neq 0$, the fit becomes worse. When C=1, the worst fit for $L \geq 4$ is .107, while it is .063 for $L \geq 6$. When C=2, the worst fit is .132 for $L \geq 4$, and .076 when $L \geq 6$. However, as was true for the previous two approximations, the fit is better for the skewed distributions than it is for the normal (the worst fits for the skewed are .054 and .025 for C=1, and .065 and .030 for C=2, when $L \geq 4$ and $L \geq 6$, respectively). Also, note the worst fits occur at the lower correlations, with the fits improving as $\rho$ increases.

Looking at the comparisons of the cum$\sqrt{f(x)}$ rule (Singh's linearized cum H(x) rule, Tables 3.17-3.19), we note that for the linear model, the rule works very well, with the worst fit being .003 for both $L \geq 4$ and $L \geq 6$. However, for $C \neq 0$, the fit is not as good. For C=1, the worst fit for $L \geq 4$ is .094, while it is .058 for $L \geq 6$. For C=2, the worst fit for $L \geq 4$ is .115, and it is .068 for $L \geq 6$. While the approximation worked better with the skewed distributions than it did with the normal, the fit is still only fair.

Finally, examining the cum $\sqrt[3]{f(x)}$ rule (the linearized cum H(x) rule of Singh-Sukhatme, Tables 3.20-3.22), the fit for the linear model is not as good as the cum $\sqrt{f(x)}$ rule. For $L \geq 4$, the worst fit is .030, while it is .012 for $L \geq 6$. For the non-linear model, however, the fit is better than the cum $\sqrt{f(x)}$ rule. In particular, when C=1, the worst fits for $L \geq 4$ and $L \geq 6$ are .063 and .032, respectively, while for C=2, the worst fits for $L \geq 4$ and $L \geq 6$ are .075 and .038, respectively. So, while the cum $\sqrt{f(x)}$ rule is much superior to the cum $\sqrt[3]{f(x)}$ rule when the model is linear, it would appear that the cum $\sqrt[3]{f(x)}$ rule offers a better fit when the functional relationship between X and Y is non-linear.

### 3.4.3. Equal Allocation

Looking at how well Singh-Parkash's cum H(x) rule approximates the Dalenius-Gurney equations for equal allocation, we note that there are actually two separate cases, depending upon the skewness of the underlying distribution. For the symmetric normal distribution (Table 3.25), when C=0, the fit is reasonable, with the worst value for $L \geq 4$ being .039, and the worst value for $L \geq 6$ being .036, both of these occurring when $\rho=.99$. When the model is non-linear, the worst fits occur when $\rho=0$; when C=1, the worst fit is .107 for $L \geq 4$, while the worst fit is .101 for

$L \geq 6$. When C=2, the worst fit is .194 for $L \geq 4$, with .174 being the worst fit for $L \geq 6$. Thus, while the approximation offers a reasonable fit for the normal distribution when the model is linear, the approximation is not good for the non-linear model.

However, for the two skewed distributions (Tables 3.23-3.24), the fits of Singh-Parkash's cum H(x) rule are worse. For the linear model, when homoscedastic error variation is modeled (r=0), the worst fit for $L \geq 4$ is .271, which becomes .202 for $L \geq 6$. When the error variation is heterogeneous, the worst value for $L \geq 4$ is .097, and the worst value for $L \geq 6$ is .065. For the non-linear models, the worst fits for C=1 are .237 for $L \geq 4$, and .199 for $L \geq 6$; for C=2, the worst fits for $L \geq 4$ and $L \geq 6$ are .320 and .268, respectively. Thus, Singh-Parkash's approximation applied to the skewed distribution offers fits that are quite poor; they are somewhat better, but still poor, when the distribution is normal.

For the linearization approach, we note that we can again group the success of the approximation by the skewness of the auxiliary variable's distribution. For the normal distribution (Table 3.28), when C=0, we are confronted with the curious phenomenon of having the approximation become worse as L increases, though the worst it becomes over the range of L we are considering is only .104. For $C \neq 0$, increases in L produce improvements in the approximation. When C=1, for $L \geq 4$, the worst value is .051, while the worst fit is .027 for $L \geq 6$. When C=2, the worst fit is .058 for $L \geq 4$, while it is .032 for $L \geq 6$. Thus, the linearization approximation offers a reasonable fit for the normal distribution, a fit that is actually better when the model is non-linear than when it is linear.

For the skewed distributions under the linearization approach (Tables 3.26-3.27), again, when C=0, we note that the fit gets worse as the number of strata increases, though this time the worst fit for the distributions is .142. When $C \neq 0$, there is a marked difference between the fits as the correlation varies: for $\rho \geq .60$, when C=1 and $L \geq 4$, the worst fit is .101, while the worst fit is .085 for $L \geq 6$; for $\rho = .30$, however, when $L \geq 4$, the worst fit is .250, with that value also being the worst fit when $L \geq 6$. For C=2, the same difference in fits occurs: for $\rho \geq .60$, the worst difference is .116 when $L \geq 4$, and the largest difference for $L \geq 6$ is .092; when $\rho = .30$, the worst differences are .305 and .284 for $L \geq 4$ and $L \geq 6$, respectively.

Finally, the comparison of the last approximation, the linearized Singh-Parkash's cum H(x) rule, demonstrated a fit so bad that we chose not to include the table with the rest of our comparisons. In fact, differences in the design effects of the magnitude of .300 were not the largest values--they were near the average.

Table 3.2. Difference between design effects using the Dalenius-Gurney equations and Singh's cum H(x) rule under proportional allocation for the $\chi_1^2$ distribution.

| | ρ | C=0 | | | C-1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .00 | -- | .000 | .000 | .024 | .024 | .032 | .032 |
| | .30 | .004 | .004 | .005 | .000 | .000 | .001 | .001 |
| | .60 | .015 | .015 | .015 | .025 | .025 | .020 | .020 |
| | .90 | .034 | .034 | .034 | .048 | .048 | .050 | .050 |
| | .99 | .041 | .041 | .041 | .053 | .053 | .054 | .054 |
| L=3 | .00 | -- | .000 | .000 | .134 | .134 | .178 | .178 |
| | .30 | .002 | .002 | .003 | .035 | .035 | .053 | .053 |
| | .60 | .007 | .007 | .007 | .013 | .013 | .014 | .014 |
| | .90 | .015 | .015 | .015 | .026 | .026 | .026 | .026 |
| | .99 | .018 | .018 | .018 | .025 | .025 | .026 | .026 |
| L=4 | .00 | -- | .000 | .000 | .042 | .043 | .056 | .056 |
| | .30 | .001 | .001 | .002 | .042 | .042 | .061 | .061 |
| | .60 | .003 | .003 | .004 | .008 | .008 | .013 | .013 |
| | .90 | .007 | .007 | .008 | .014 | .014 | .014 | .014 |
| | .99 | .009 | .009 | .009 | .013 | .013 | .013 | .013 |
| L=5 | .00 | -- | .000 | .000 | .018 | .019 | .024 | .024 |
| | .30 | .000 | .001 | .001 | .032 | .032 | .038 | .038 |
| | .60 | .002 | .002 | .003 | .007 | .007 | .012 | .012 |
| | .90 | .004 | .004 | .004 | .008 | .008 | .008 | .008 |
| | .99 | .005 | .005 | .005 | .007 | .007 | .008 | .008 |
| L=6 | .00 | -- | .000 | .000 | .022 | .022 | .029 | .029 |
| | .30 | .000 | .000 | .001 | .014 | .014 | .017 | .017 |
| | .60 | .001 | .001 | .002 | .006 | .006 | .012 | .012 |
| | .90 | .003 | .003 | .003 | .005 | .005 | .005 | .005 |
| | .99 | .003 | .003 | .003 | .005 | .005 | .005 | .005 |
| L=7 | .00 | -- | .000 | .000 | .013 | .013 | .017 | .017 |
| | .30 | .000 | .000 | .001 | .008 | .008 | .011 | .011 |
| | .60 | .001 | .001 | .001 | .005 | .005 | .011 | .011 |
| | .90 | .002 | .002 | .002 | .003 | .003 | .004 | .004 |
| | .99 | .002 | .002 | .002 | .003 | .003 | .003 | .003 |
| L=8 | .00 | -- | .000 | .000 | .007 | .007 | .008 | .009 |
| | .30 | .000 | .000 | .001 | .007 | .007 | .011 | .011 |
| | .60 | .001 | .001 | .001 | .005 | .005 | .010 | .009 |
| | .90 | .001 | .001 | .001 | .002 | .002 | .002 | .002 |
| | .99 | .001 | .001 | .001 | .002 | .002 | .002 | .002 |

Table 3.3. Difference between design effects using the Dalenius-Gurney equations and Singh's cum H(x) rule under proportional allocation for the $\chi_2^2$ distribution.

| | ρ | C=0 r=0 | C=0 r=1 | C=0 r=2 | C=1 r=1 | C=1 r=2 | C=2 r=1 | C=2 r=2 |
|---|---|---|---|---|---|---|---|---|
| | .00 | -- | .000 | .000 | .037 | .037 | .049 | .049 |
| | .30 | .003 | .003 | .004 | .002 | .002 | .005 | .005 |
| L=2 | .60 | .011 | .011 | .011 | .018 | .018 | .012 | .012 |
| | .90 | .024 | .024 | .024 | .043 | .042 | .046 | .046 |
| | .99 | .029 | .029 | .029 | .040 | .041 | .042 | .042 |
| | .00 | -- | .000 | .000 | .127 | .127 | .160 | .169 |
| | .30 | .001 | .002 | .002 | .047 | .047 | .070 | .070 |
| L=3 | .60 | .005 | .005 | .006 | .012 | .012 | .016 | .016 |
| | .90 | .012 | .012 | .012 | .022 | .022 | .024 | .024 |
| | .99 | .014 | .014 | .014 | .020 | .020 | .020 | .020 |
| | .00 | -- | .000 | .000 | .030 | .030 | .039 | .039 |
| | .30 | .001 | .001 | .002 | .049 | .049 | .075 | .075 |
| L=4 | .60 | .003 | .003 | .004 | .010 | .010 | .018 | .018 |
| | .90 | .006 | .006 | .006 | .012 | .012 | .013 | .013 |
| | .99 | .007 | .007 | .007 | .011 | .011 | .011 | .011 |
| | .00 | -- | .000 | .000 | .020 | .020 | .026 | .026 |
| | .30 | .000 | .001 | .001 | .024 | .024 | .027 | .027 |
| L=5 | .60 | .002 | .002 | .002 | .010 | .010 | .019 | .018 |
| | .90 | .004 | .004 | .004 | .007 | .007 | .007 | .007 |
| | .99 | .004 | .004 | .004 | .006 | .006 | .006 | .006 |
| | .00 | -- | .000 | .000 | .024 | .024 | .032 | .032 |
| | .30 | .000 | .000 | .001 | .011 | .011 | .014 | .014 |
| L=6 | .60 | .001 | .001 | .002 | .010 | .009 | .017 | .017 |
| | .90 | .002 | .002 | .002 | .004 | .004 | .005 | .005 |
| | .99 | .003 | .003 | .003 | .004 | .004 | .004 | .004 |
| | .00 | -- | .000 | .000 | .010 | .010 | .012 | .012 |
| | .30 | .000 | .000 | .001 | .009 | .009 | .014 | .014 |
| L=7 | .60 | .001 | .001 | .001 | .009 | .009 | .015 | .015 |
| | .90 | .001 | .002 | .002 | .003 | .003 | .003 | .003 |
| | .99 | .002 | .002 | .002 | .003 | .003 | .003 | .003 |
| | .00 | -- | .000 | .000 | .008 | .008 | .010 | .010 |
| | .30 | .000 | .000 | .001 | .010 | .010 | .013 | .013 |
| L=8 | .60 | .000 | .001 | .001 | .008 | .008 | .009 | .009 |
| | .90 | .001 | .001 | .001 | .002 | .002 | .002 | .002 |
| | .99 | .001 | .001 | .001 | .002 | .002 | .002 | .002 |

Table 3.4. Difference between design effects using the Dalenius-Gurney equations and Singh's cum H(x) rule under proportional allocation for the normal distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .00 | -- | .000 | .000 | .152 | .152 | .203 | .203 |
| | .30 | .000 | .000 | .000 | .097 | .097 | .139 | .139 |
| | .60 | .000 | .000 | .000 | .015 | .015 | .031 | .031 |
| | .90 | .000 | .000 | .000 | .010 | .010 | .008 | .008 |
| | .99 | .000 | .000 | .000 | .002 | .002 | .003 | .003 |
| L=3 | .00 | -- | .000 | .000 | .009 | .009 | .011 | .011 |
| | .30 | .001 | .001 | .001 | .039 | .040 | .043 | .043 |
| | .60 | .003 | .003 | .003 | .070 | .070 | .105 | .105 |
| | .90 | .007 | .007 | .007 | .008 | .008 | .009 | .009 |
| | .99 | .009 | .009 | .009 | .009 | .009 | .008 | .008 |
| L=4 | .00 | -- | .000 | .000 | .059 | .059 | .081 | .081 |
| | .30 | .001 | .001 | .001 | .022 | .022 | .033 | .032 |
| | .60 | .002 | .002 | .002 | .047 | .047 | .038 | .038 |
| | .90 | .006 | .006 | .006 | .006 | .006 | .008 | .008 |
| | .99 | .007 | .007 | .007 | .006 | .006 | .006 | .006 |
| L=5 | .00 | -- | .000 | .000 | .018 | .018 | .022 | .022 |
| | .30 | .000 | .000 | .001 | .038 | .038 | .046 | .046 |
| | .60 | .002 | .002 | .002 | .010 | .010 | .014 | .014 |
| | .90 | .004 | .004 | .004 | .005 | .005 | .007 | .006 |
| | .99 | .005 | .005 | .005 | .004 | .004 | .004 | .004 |
| L=6 | .00 | -- | .000 | .000 | .023 | .023 | .033 | .033 |
| | .30 | .000 | .000 | .000 | .014 | .014 | .016 | .016 |
| | .60 | .001 | .001 | .001 | .013 | .013 | .022 | .022 |
| | .90 | .003 | .003 | .003 | .004 | .004 | .006 | .006 |
| | .99 | .003 | .003 | .003 | .003 | .003 | .003 | .003 |
| L=7 | .00 | -- | .000 | .000 | .013 | .013 | .015 | .015 |
| | .30 | .000 | .000 | .000 | .015 | .015 | .022 | .021 |
| | .60 | .001 | .001 | .001 | .014 | .014 | .016 | .016 |
| | .90 | .002 | .002 | .002 | .003 | .003 | .006 | .005 |
| | .99 | .002 | .002 | .002 | .002 | .002 | .002 | .002 |
| L=8 | .00 | -- | .000 | .000 | .012 | .012 | .016 | .016 |
| | .30 | .000 | .000 | .000 | .010 | .010 | .012 | .012 |
| | .60 | .001 | .001 | .001 | .006 | .006 | .006 | .006 |
| | .90 | .001 | .001 | .001 | .003 | .003 | .005 | .005 |
| | .99 | .002 | .002 | .002 | .002 | .002 | .002 | .002 |

Table 3.5. Difference between design effects using the Dalenius-Gurney equations and Singh-Sukhatme's cum H(x) rule under proportional allocation for the $\chi^2_1$, $\chi^2_2$ and normal distributions.

| | | $\chi^2_1$ | | | $\chi^2_2$ | | | Normal | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ρ | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 |
| | .00 | .000 | .024 | .032 | .000 | .036 | .048 | .000 | .152 | .203 |
| | .30 | .004 | .000 | .001 | .003 | .002 | .005 | .000 | .092 | .135 |
| L=2 | .60 | .015 | .025 | .020 | .011 | .018 | .013 | .000 | .014 | .029 |
| | .90 | .034 | .048 | .050 | .024 | .043 | .046 | .000 | .010 | .008 |
| | .99 | .041 | .053 | .054 | .029 | .040 | .042 | .000 | .002 | .003 |
| | .00 | .000 | .134 | .178 | .000 | .126 | .168 | .000 | .007 | .009 |
| | .30 | .002 | .034 | .052 | .001 | .045 | .069 | .001 | .039 | .042 |
| L=3 | .60 | .007 | .013 | .014 | .005 | .012 | .016 | .003 | .065 | .101 |
| | .90 | .015 | .026 | .026 | .012 | .022 | .024 | .007 | .008 | .009 |
| | .99 | .018 | .025 | .026 | .014 | .020 | .021 | .009 | .009 | .009 |
| | .00 | .000 | .042 | .056 | .000 | .029 | .038 | .000 | .063 | .085 |
| | .30 | .001 | .041 | .060 | .001 | .048 | .075 | .001 | .018 | .029 |
| L=4 | .60 | .003 | .008 | .013 | .003 | .010 | .017 | .002 | .048 | .037 |
| | .90 | .007 | .014 | .014 | .006 | .012 | .013 | .006 | .006 | .007 |
| | .99 | .009 | .013 | .013 | .007 | .011 | .011 | .007 | .006 | .006 |
| | .00 | .000 | .017 | .023 | .000 | .019 | .025 | .000 | .013 | .018 |
| | .30 | .000 | .032 | .038 | .000 | .023 | .027 | .000 | .041 | .046 |
| L=5 | .60 | .002 | .006 | .012 | .002 | .009 | .018 | .002 | .008 | .011 |
| | .90 | .004 | .008 | .008 | .004 | .007 | .007 | .004 | .004 | .006 |
| | .99 | .005 | .007 | .008 | .004 | .006 | .006 | .005 | .004 | .004 |
| | .00 | .000 | .021 | .028 | .000 | .024 | .032 | .000 | .027 | .036 |
| | .30 | .000 | .014 | .016 | .000 | .010 | .013 | .000 | .010 | .013 |
| L=6 | .60 | .001 | .005 | .011 | .001 | .008 | .017 | .001 | .009 | .019 |
| | .90 | .003 | .005 | .005 | .002 | .004 | .005 | .003 | .003 | .005 |
| | .99 | .003 | .005 | .005 | .003 | .004 | .004 | .003 | .003 | .003 |
| | .00 | .000 | .012 | .016 | .000 | .008 | .011 | .000 | .008 | .011 |
| | .30 | .000 | .008 | .010 | .000 | .008 | .013 | .000 | .013 | .021 |
| L=7 | .60 | .001 | .005 | .010 | .001 | .008 | .014 | .001 | .014 | .016 |
| | .90 | .002 | .003 | .004 | .001 | .003 | .003 | .002 | .003 | .005 |
| | .99 | .002 | .003 | .003 | .002 | .003 | .003 | .002 | .002 | .002 |
| | .00 | .000 | .006 | .008 | .000 | .007 | .009 | .000 | .013 | .018 |
| | .30 | .000 | .006 | .010 | .000 | .009 | .014 | .000 | .008 | .009 |
| L=8 | .60 | .001 | .004 | .009 | .000 | .007 | .009 | .001 | .005 | .005 |
| | .90 | .001 | .002 | .002 | .001 | .002 | .002 | .001 | .002 | .004 |
| | .99 | .001 | .002 | .002 | .001 | .002 | .002 | .002 | .002 | .002 |

Table 3.6. Differences between design effects using the Dalenius-Gurney equations and the linearization approach under proportional allocation for the $\chi^2_1$, $\chi^2_2$ and normal distributions.

| | | $\chi^2_1$ | | | $\chi^2_2$ | | | Normal | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\rho$ | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 |
| L=2 | .30 | | .283 | .361 | | .280 | .357 | | .227 | .290 |
| | .60 | | .254 | .328 | | .254 | .327 | | .210 | .268 |
| | .90 | | .077 | .104 | | .080 | .108 | | .077 | .101 |
| | .99 | | .007 | .009 | | .007 | .010 | | .009 | .012 |
| L=3 | .30 | | .231 | .291 | | .219 | .274 | | .135 | .184 |
| | .60 | | .214 | .277 | | .210 | .269 | | .149 | .185 |
| | .90 | | .064 | .087 | | .066 | .089 | | .060 | .078 |
| | .99 | | .006 | .008 | | .006 | .008 | | .007 | .009 |
| L=4 | .30 | | .169 | .211 | | .153 | .201 | | .110 | .145 |
| | .60 | | .165 | .213 | | .160 | .204 | | .100 | .119 |
| | .90 | | .050 | .067 | | .051 | .068 | | .045 | .058 |
| | .99 | | .005 | .006 | | .005 | .007 | | .005 | .007 |
| L=5 | .30 | Identical | .133 | .174 | Identical | .126 | .165 | Identical | .078 | .101 |
| | .60 | | .128 | .164 | | .123 | .155 | | .065 | .086 |
| | .90 | | .039 | .052 | | .039 | .053 | | .034 | .044 |
| | .99 | | .004 | .005 | | .004 | .005 | | .004 | .005 |
| L=6 | .30 | | .109 | .142 | | .102 | .132 | | .060 | .080 |
| | .60 | | .102 | .129 | | .096 | .121 | | .052 | .067 |
| | .90 | | .031 | .042 | | .031 | .042 | | .026 | .034 |
| | .99 | | .003 | .004 | | .003 | .004 | | .003 | .004 |
| L=7 | .30 | | .090 | .116 | | .082 | .106 | | .048 | .063 |
| | .60 | | .083 | .104 | | .077 | .095 | | .041 | .052 |
| | .90 | | .025 | .034 | | .025 | .034 | | .021 | .027 |
| | .99 | | .003 | .003 | | .002 | .003 | | .002 | .003 |
| L=8 | .30 | | .074 | .096 | | .067 | .086 | | .037 | .050 |
| | .60 | | .068 | .085 | | .063 | .076 | | .030 | .040 |
| | .90 | | .021 | .028 | | .021 | .028 | | .017 | .022 |
| | .99 | | .002 | .003 | | .002 | .003 | | .002 | .003 |

Table 3.7. Differences between design effects using the Dalenius-Gurney equations and the linearized cum H(x) rule under proportional allocation for the $\chi^2_1$, $\chi^2_2$ and normal distributions.

| | | $\chi^2_1$ | | | $\chi^2_2$ | | | Normal | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ρ | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 |
| | .00 | .000 | .196 | .261 | .000 | .191 | .255 | .000 | .152 | .203 |
| | .30 | .004 | .227 | .298 | .003 | .232 | .303 | .000 | .227 | .290 |
| L=2 | .60 | .015 | .156 | .214 | .011 | .167 | .227 | .000 | .210 | .268 |
| | .90 | .034 | .013 | .026 | .024 | .020 | .035 | .000 | .077 | .101 |
| | .99 | .041 | .012 | .009 | .029 | .006 | .004 | .000 | .009 | .012 |
| | .00 | .000 | .128 | .170 | .000 | .129 | .172 | .000 | .121 | .162 |
| | .30 | .002 | .151 | .191 | .001 | .146 | .183 | .001 | .100 | .137 |
| L=3 | .60 | .007 | .126 | .168 | .005 | .131 | .172 | .003 | .127 | .154 |
| | .90 | .015 | .017 | .028 | .012 | .023 | .036 | .007 | .060 | .075 |
| | .99 | .018 | .004 | .003 | .014 | .003 | .002 | .009 | .015 | .017 |
| | .00 | .000 | .099 | .132 | .000 | .096 | .129 | .000 | .066 | .088 |
| | .30 | .001 | .099 | .121 | .001 | .091 | .121 | .001 | .078 | .102 |
| L=4 | .60 | .003 | .098 | .128 | .003 | .100 | .129 | .002 | .079 | .091 |
| | .90 | .007 | .017 | .026 | .006 | .021 | .031 | .006 | .043 | .054 |
| | .99 | .009 | .002 | .001 | .007 | .002 | .002 | .007 | .011 | .012 |
| | .00 | .000 | .072 | .097 | .000 | .068 | .091 | .000 | .055 | .074 |
| | .30 | .000 | .078 | .103 | .000 | .077 | .101 | .000 | .053 | .067 |
| L=5 | .60 | .002 | .077 | .100 | .002 | .078 | .098 | .002 | .049 | .063 |
| | .90 | .004 | .015 | .023 | .004 | .018 | .026 | .004 | .032 | .040 |
| | .99 | .005 | .001 | .001 | .004 | .001 | .001 | .005 | .008 | .009 |
| | .00 | .000 | .056 | .074 | .000 | .054 | .072 | .000 | .038 | .051 |
| | .30 | .000 | .065 | .085 | .000 | .063 | .082 | .000 | .040 | .054 |
| L=6 | .60 | .001 | .062 | .079 | .001 | .062 | .077 | .001 | .039 | .049 |
| | .90 | .003 | .013 | .019 | .002 | .015 | .022 | .003 | .025 | .031 |
| | .99 | .003 | .001 | .001 | .003 | .001 | .001 | .003 | .006 | .007 |
| | .00 | .000 | .047 | .063 | .000 | .045 | .060 | .000 | .056 | .075 |
| | .30 | .000 | .054 | .069 | .000 | .051 | .066 | .000 | .055 | .071 |
| L=7 | .60 | .001 | .051 | .064 | .001 | .050 | .061 | .001 | .051 | .063 |
| | .90 | .002 | .012 | .017 | .001 | .013 | .019 | .002 | .036 | .043 |
| | .99 | .002 | .001 | .001 | .002 | .001 | .001 | .002 | .020 | .020 |
| | .00 | .000 | .039 | .052 | .000 | .037 | .049 | .000 | .042 | .055 |
| | .30 | .000 | .045 | .057 | .000 | .042 | .053 | .000 | .041 | .054 |
| L=8 | .60 | .001 | .043 | .053 | .000 | .041 | .049 | .001 | .036 | .047 |
| | .90 | .001 | .010 | .014 | .001 | .011 | .016 | .001 | .027 | .032 |
| | .99 | .001 | .001 | .001 | .001 | .001 | .001 | .002 | .014 | .014 |

Table 3.8. Differences between design effects using the Dalenius-Gurney equations and Singh's cum H(x) rule under Neyman allocation for the $\chi_1^2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .007 | .031 | .058 | .014 | .084 | .023 |
| | .30 | .001 | .001 | .001 | .097 | .041 | .136 | .087 |
| L=2 | .60 | .002 | .000 | .002 | .009 | .000 | .018 | .006 |
| | .90 | .000 | .001 | .004 | .000 | .002 | .001 | .002 |
| | .99 | .004 | .005 | .006 | .005 | .006 | .005 | .006 |
| | .00 | -- | .003 | .015 | .039 | .017 | .045 | .028 |
| | .30 | .001 | .000 | .000 | .010 | .003 | .014 | .009 |
| L=3 | .60 | .003 | .000 | .000 | .013 | .001 | .024 | .009 |
| | .90 | .001 | .000 | .001 | .000 | .000 | .000 | .000 |
| | .99 | .001 | .001 | .002 | .001 | .002 | .001 | .002 |
| | .00 | -- | .002 | .009 | .010 | .003 | .010 | .006 |
| | .30 | .001 | .000 | .000 | .018 | .005 | .024 | .011 |
| L=4 | .60 | .003 | .000 | .000 | .009 | .001 | .011 | .001 |
| | .90 | .002 | .000 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .001 | .000 | .001 | .000 | .001 |
| | .00 | -- | .001 | .005 | .017 | .005 | .020 | .010 |
| | .30 | .001 | .000 | .000 | .010 | .001 | .010 | .002 |
| L=5 | .60 | .003 | .000 | .000 | .004 | .000 | .004 | .001 |
| | .90 | .002 | .001 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .00 | -- | .001 | .003 | .009 | .002 | .009 | .002 |
| | .30 | .001 | .000 | .000 | .007 | .002 | .010 | .004 |
| L=6 | .60 | .003 | .001 | .000 | .003 | .001 | .005 | .001 |
| | .90 | .002 | .001 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .00 | -- | .000 | .002 | .008 | .002 | .009 | .003 |
| | .30 | .001 | .000 | .000 | .007 | .001 | .007 | .002 |
| L=7 | .60 | .002 | .000 | .000 | .003 | .001 | .005 | .001 |
| | .90 | .002 | .001 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .00 | -- | .000 | .002 | .007 | .002 | .007 | .002 |
| | .30 | .001 | .000 | .000 | .006 | .001 | .006 | .002 |
| L=8 | .60 | .002 | .000 | .000 | .003 | .001 | .004 | .001 |
| | .90 | .002 | .001 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .001 | .000 | .000 | .000 | .000 | .000 | .000 |

Table 3.9. Differences between design effects using the Dalenius-Gurney equations and Singh's cum H(x) rule under Neyman allocation for the $\chi^2_2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .007 | .027 | .084 | .021 | .124 | .077 |
| | .30 | .001 | .001 | .001 | .116 | .065 | .165 | .122 |
| L=2 | .60 | .001 | .000 | .001 | .016 | .003 | .029 | .014 |
| | .90 | .000 | .000 | .002 | .000 | .001 | .007 | .001 |
| | .99 | .001 | .002 | .002 | .002 | .002 | .002 | .002 |
| | .00 | -- | .004 | .016 | .044 | .022 | .052 | .035 |
| | .30 | .001 | .000 | .001 | .015 | .007 | .021 | .015 |
| L=3 | .60 | .002 | .000 | .001 | .020 | .005 | .034 | .018 |
| | .90 | .001 | .000 | .001 | .001 | .000 | .000 | .000 |
| | .99 | .000 | .001 | .001 | .001 | .001 | .001 | .001 |
| | .00 | -- | .002 | .010 | .013 | .006 | .016 | .012 |
| | .30 | .001 | .000 | .000 | .021 | .007 | .029 | .015 |
| L=4 | .60 | .002 | .000 | .000 | .013 | .001 | .014 | .002 |
| | .90 | .002 | .000 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .00 | -- | .002 | .007 | .019 | .006 | .023 | .012 |
| | .30 | .001 | .000 | .000 | .012 | .001 | .012 | .004 |
| L=5 | .60 | .002 | .000 | .000 | .005 | .001 | .006 | .002 |
| | .90 | .002 | .000 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .00 | -- | .001 | .005 | .009 | .003 | .010 | .005 |
| | .30 | .001 | .000 | .000 | .009 | .002 | .012 | .005 |
| L=6 | .60 | .002 | .000 | .000 | .004 | .001 | .007 | .002 |
| | .90 | .002 | .001 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .00 | -- | .001 | .004 | .010 | .003 | .011 | .005 |
| | .30 | .001 | .000 | .000 | .008 | .001 | .009 | .002 |
| L=7 | .60 | .002 | .000 | .000 | .004 | .001 | .006 | .002 |
| | .90 | .002 | .001 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .00 | -- | .001 | .003 | .007 | .002 | .008 | .001 |
| | .30 | .001 | .000 | .000 | .006 | .001 | .007 | .003 |
| L=8 | .60 | .002 | .000 | .000 | .004 | .001 | .005 | .001 |
| | .90 | .002 | .001 | .000 | .001 | .000 | .001 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |

Table 3.10. Differences between design effects using the Dalenius-Gurney equations and Singh's cum H(x) rule under Neyman allocation for the normal distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .00 | -- | .000 | .000 | .152 | .150 | .206 | .204 |
| | .30 | .000 | .000 | .000 | .131 | .127 | .180 | .176 |
| | .60 | .000 | .000 | .000 | .040 | .036 | .059 | .056 |
| | .90 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| L=3 | .00 | -- | .000 | .000 | .056 | .055 | .062 | .061 |
| | .30 | .000 | .000 | .000 | .060 | .061 | .072 | .073 |
| | .60 | .000 | .000 | .000 | .052 | .049 | .072 | .070 |
| | .90 | .000 | .000 | .000 | .001 | .000 | .001 | .001 |
| | .99 | .001 | .001 | .001 | .001 | .001 | .001 | .001 |
| L=4 | .00 | -- | .000 | .000 | .043 | .043 | .052 | .053 |
| | .30 | .000 | .000 | .000 | .036 | .034 | .040 | .039 |
| | .60 | .000 | .000 | .000 | .036 | .035 | .031 | .031 |
| | .90 | .000 | .000 | .000 | .001 | .001 | .002 | .001 |
| | .99 | .001 | .001 | .001 | .001 | .001 | .001 | .001 |
| L=5 | .00 | -- | .000 | .000 | .030 | .029 | .032 | .031 |
| | .30 | .000 | .000 | .000 | .028 | .026 | .031 | .029 |
| | .60 | .000 | .000 | .000 | .018 | .017 | .020 | .019 |
| | .90 | .000 | .000 | .000 | .001 | .001 | .002 | .002 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| L=6 | .00 | -- | .000 | .000 | .023 | .023 | .025 | .025 |
| | .30 | .000 | .000 | .000 | .021 | .020 | .022 | .022 |
| | .60 | .000 | .000 | .001 | .015 | .014 | .017 | .016 |
| | .90 | .000 | .000 | .000 | .001 | .001 | .002 | .002 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| L=7 | .00 | -- | .000 | .000 | .019 | .019 | .020 | .020 |
| | .30 | .000 | .000 | .000 | .017 | .016 | .015 | .015 |
| | .60 | .000 | .000 | .001 | .013 | .012 | .014 | .013 |
| | .90 | .000 | .000 | .001 | .002 | .001 | .003 | .002 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| L=8 | .00 | -- | .000 | .000 | .016 | .016 | .017 | .017 |
| | .30 | .000 | .000 | .000 | .014 | .014 | .015 | .015 |
| | .60 | .000 | .000 | .001 | .010 | .009 | .011 | .010 |
| | .90 | .000 | .000 | .001 | .002 | .001 | .003 | .002 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |

Table 3.11. Differences between design effects using the Dalenius-Gurney and Singh-Sukhatme's cum H(x) rule under Neyman allocation for the $\chi_1^2$ distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .002 | .001 | .023 | .053 | .025 | .050 |
| | .30 | .004 | .005 | .002 | .000 | .000 | .001 | .003 |
| L=2 | .60 | .018 | .013 | .009 | .045 | .038 | .053 | .032 |
| | .90 | .062 | .035 | .018 | .084 | .050 | .099 | .059 |
| | .99 | .102 | .055 | .023 | .076 | .034 | .079 | .036 |
| | .00 | -- | .001 | .001 | .074 | .022 | .107 | .034 |
| | .30 | .002 | .002 | .001 | .045 | .045 | .074 | .065 |
| L=3 | .60 | .008 | .005 | .003 | .011 | .011 | .011 | .007 |
| | .90 | .027 | .014 | .007 | .034 | .020 | .039 | .023 |
| | .99 | .047 | .024 | .010 | .035 | .016 | .037 | .016 |
| | .00 | -- | .000 | .001 | .014 | .017 | .018 | .024 |
| | .30 | .001 | .001 | .000 | .032 | .010 | .043 | .012 |
| L=4 | .60 | .004 | .002 | .001 | .008 | .008 | .010 | .009 |
| | .90 | .014 | .006 | .003 | .017 | .009 | .020 | .011 |
| | .99 | .025 | .012 | .005 | .019 | .008 | .020 | .009 |
| | .00 | -- | .000 | .001 | .025 | .006 | .035 | .010 |
| | .30 | .000 | .000 | .000 | .010 | .005 | .013 | .008 |
| L=5 | .60 | .002 | .001 | .001 | .006 | .007 | .010 | .010 |
| | .90 | .008 | .004 | .002 | .010 | .005 | .011 | .006 |
| | .99 | .015 | .007 | .003 | .011 | .005 | .012 | .006 |
| | .00 | -- | .000 | .001 | .008 | .005 | .012 | .006 |
| | .30 | .000 | .000 | .000 | .007 | .006 | .012 | .008 |
| L=6 | .60 | .001 | .001 | .000 | .006 | .005 | .011 | .005 |
| | .90 | .005 | .002 | .001 | .006 | .003 | .007 | .004 |
| | .99 | .010 | .005 | .002 | .007 | .003 | .008 | .004 |
| | .00 | -- | .000 | .001 | .007 | .003 | .010 | .005 |
| | .30 | .000 | .000 | .000 | .008 | .002 | .012 | .002 |
| L=7 | .60 | .001 | .001 | .000 | .005 | .003 | .007 | .003 |
| | .90 | .003 | .001 | .001 | .004 | .002 | .005 | .003 |
| | .99 | .007 | .003 | .001 | .005 | .002 | .006 | .003 |
| | .00 | -- | .000 | .000 | .006 | .002 | .009 | .003 |
| | .30 | .000 | .000 | .000 | .004 | .002 | .005 | .003 |
| L=8 | .60 | .001 | .000 | .000 | .003 | .002 | .004 | .002 |
| | .90 | .002 | .001 | .000 | .003 | .002 | .004 | .002 |
| | .99 | .005 | .002 | .001 | .004 | .002 | .004 | .002 |

Table 3.12. Differences between design effects using the Dalenius-Gurney equations and Singh-Sukhatme's cum H(x) rule under Neyman allocation for the $\chi_2^2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .001 | .004 | .046 | .046 | .056 | .095 |
| | .30 | .003 | .002 | .000 | .002 | .006 | .002 | .018 |
| L=2 | .60 | .013 | .007 | .003 | .027 | .021 | .031 | .014 |
| | .90 | .043 | .020 | .006 | .066 | .035 | .082 | .046 |
| | .99 | .069 | .029 | .007 | .047 | .014 | .050 | .015 |
| | .00 | -- | .000 | .003 | .057 | .011 | .084 | .018 |
| | .30 | .001 | .001 | .001 | .060 | .042 | .099 | .061 |
| L=3 | .60 | .006 | .004 | .002 | .113 | .008 | .013 | .011 |
| | .90 | .020 | .010 | .004 | .029 | .016 | .034 | .019 |
| | .99 | .036 | .016 | .006 | .026 | .010 | .027 | .011 |
| | .00 | -- | .000 | .002 | .014 | .026 | .019 | .039 |
| | .30 | .001 | .001 | .001 | .025 | .006 | .033 | .008 |
| L=4 | .60 | .003 | .002 | .001 | .011 | .006 | .018 | .014 |
| | .90 | .011 | .005 | .003 | .015 | .008 | .018 | .010 |
| | .99 | .020 | .010 | .004 | .015 | .007 | .016 | .007 |
| | .00 | -- | .000 | .001 | .025 | .005 | .037 | .008 |
| | .30 | .000 | .000 | .000 | .009 | .008 | .011 | .015 |
| L=5 | .60 | .002 | .001 | .001 | .009 | .007 | .017 | .010 |
| | .90 | .006 | .003 | .002 | .009 | .005 | .010 | .006 |
| | .99 | .012 | .006 | .003 | .010 | .004 | .010 | .005 |
| | .00 | -- | .000 | .001 | .007 | .009 | .009 | .013 |
| | .30 | .000 | .000 | .000 | .010 | .006 | .018 | .008 |
| L=6 | .60 | .001 | .001 | .000 | .007 | .004 | .011 | .004 |
| | .90 | .004 | .002 | .001 | .006 | .003 | .007 | .004 |
| | .99 | .008 | .004 | .002 | .007 | .003 | .007 | .003 |
| | .00 | -- | .000 | .001 | .011 | .002 | .015 | .004 |
| | .30 | .000 | .000 | .000 | .007 | .002 | .010 | .003 |
| L=7 | .60 | .001 | .000 | .000 | .005 | .003 | .006 | .002 |
| | .90 | .003 | .001 | .001 | .004 | .002 | .005 | .002 |
| | .99 | .006 | .003 | .001 | .005 | .002 | .005 | .002 |
| | .00 | -- | .000 | .001 | .004 | .004 | .007 | .003 |
| | .30 | .000 | .000 | .000 | .003 | .003 | .005 | .005 |
| L=8 | .60 | .001 | .000 | .000 | .003 | .002 | .004 | .002 |
| | .90 | .002 | .001 | .000 | .003 | .002 | .003 | .002 |
| | .99 | .004 | .002 | .001 | .003 | .002 | .004 | .002 |

Table 3.13. Differences between design effects using the Dalenius-Gurney equations and Singh-Sukhatme's cum H(x) rule under Neyman allocation for the normal distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .000 | .000 | .132 | .104 | .176 | .138 |
| | .30 | .000 | .000 | .000 | .113 | .146 | .164 | .215 |
| L=2 | .60 | .000 | .000 | .001 | .012 | .021 | .024 | .042 |
| | .90 | .000 | .001 | .003 | .017 | .011 | .017 | .011 |
| | .99 | .000 | .001 | .004 | .001 | .000 | .002 | .000 |
| | .00 | -- | .000 | .000 | .006 | .015 | .006 | .021 |
| | .30 | .001 | .001 | .001 | .026 | .012 | .025 | .009 |
| L=3 | .60 | .003 | .004 | .004 | .078 | .093 | .122 | .141 |
| | .90 | .009 | .010 | .011 | .013 | .011 | .014 | .013 |
| | .99 | .014 | .015 | .018 | .014 | .014 | .014 | .014 |
| | .00 | -- | .000 | .000 | .057 | .043 | .080 | .059 |
| | .30 | .001 | .001 | .001 | .027 | .041 | .045 | .069 |
| L=4 | .60 | .003 | .003 | .003 | .036 | .027 | .019 | .012 |
| | .90 | .007 | .008 | .008 | .009 | .008 | .011 | .011 |
| | .99 | .012 | .013 | .014 | .012 | .012 | .012 | .012 |
| | .00 | -- | .000 | .000 | .016 | .024 | .022 | .034 |
| | .30 | .000 | .000 | .001 | .034 | .023 | .035 | .022 |
| L=5 | .60 | .002 | .002 | .002 | .008 | .010 | .016 | .024 |
| | .90 | .005 | .005 | .006 | .006 | .006 | .009 | .009 |
| | .99 | .009 | .009 | .010 | .009 | .009 | .009 | .009 |
| | .00 | -- | .000 | .000 | .023 | .016 | .031 | .022 |
| | .30 | .000 | .000 | .000 | .011 | .014 | .016 | .025 |
| L=6 | .60 | .001 | .001 | .002 | .014 | .020 | .029 | .030 |
| | .90 | .003 | .004 | .004 | .005 | .005 | .007 | .008 |
| | .99 | .006 | .007 | .007 | .006 | .007 | .006 | .006 |
| | .00 | -- | .000 | .000 | .011 | .016 | .015 | .023 |
| | .30 | .000 | .000 | .000 | .019 | .016 | .022 | .016 |
| L=7 | .60 | .001 | .001 | .001 | .015 | .011 | .012 | .009 |
| | .90 | .002 | .003 | .003 | .004 | .004 | .006 | .007 |
| | .99 | .005 | .005 | .005 | .005 | .005 | .005 | .005 |
| | .00 | -- | .000 | .000 | .011 | .008 | .015 | .011 |
| | .30 | .000 | .000 | .000 | .007 | .007 | .009 | .012 |
| L=8 | .60 | .001 | .001 | .001 | .006 | .005 | .007 | .010 |
| | .90 | .002 | .002 | .002 | .003 | .003 | .005 | .006 |
| | .99 | .003 | .004 | .004 | .004 | .004 | .004 | .004 |

Table 3.14. Differences between design effects using the Dalenius-Gurney equations and the linearization approach under Neyman allocation for the $\chi_1^2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .30 | Identical | .004 | .006 | .140 | .073 | .175 | .122 |
| | .60 | | .002 | .003 | .129 | .095 | .166 | .137 |
| | .90 | | .000 | .000 | .036 | .030 | .047 | .041 |
| | .99 | | .000 | .000 | .002 | .002 | .003 | .003 |
| L=3 | .30 | | .002 | .004 | .061 | .043 | .078 | .066 |
| | .60 | | .001 | .003 | .068 | .045 | .082 | .061 |
| | .90 | | .000 | .000 | .021 | .018 | .027 | .024 |
| | .99 | | .000 | .000 | .001 | .001 | .002 | .002 |
| L=4 | .30 | | .002 | .003 | .043 | .026 | .051 | .038 |
| | .60 | | .001 | .002 | .041 | .027 | .047 | .034 |
| | .90 | | .000 | .001 | .014 | .012 | .017 | .015 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=5 | .30 | | .001 | .002 | .028 | .016 | .032 | .023 |
| | .60 | | .001 | .002 | .026 | .019 | .029 | .025 |
| | .90 | | .000 | .001 | .010 | .009 | .012 | .011 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=6 | .30 | | .001 | .002 | .019 | .013 | .023 | .018 |
| | .60 | | .001 | .001 | .019 | .015 | .023 | .019 |
| | .90 | | .000 | .001 | .007 | .007 | .009 | .008 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=7 | .30 | | .001 | .001 | .015 | .009 | .018 | .012 |
| | .60 | | .000 | .001 | .015 | .011 | .018 | .014 |
| | .90 | | .000 | .001 | .006 | .005 | .007 | .007 |
| | .99 | | .000 | .000 | .000 | .000 | .001 | .000 |
| L=8 | .30 | | .001 | .001 | .012 | .007 | .014 | .010 |
| | .60 | | .000 | .001 | .012 | .009 | .014 | .011 |
| | .90 | | .000 | .001 | .005 | .004 | .006 | .005 |
| | .99 | | .000 | .000 | .000 | .000 | .000 | .000 |

Table 3.15. Differences between design effects using the Dalenius-Gurney equations and the linearization approach under Neyman allocation for the $\chi_2^2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .30 | Identical | .003 | .003 | .169 | .115 | .218 | .179 |
| | .60 | | .001 | .001 | .154 | .126 | .199 | .176 |
| | .90 | | .000 | .000 | .043 | .036 | .055 | .048 |
| | .99 | | .000 | .000 | .003 | .003 | .004 | .004 |
| L=3 | .30 | | .003 | .003 | .077 | .059 | .100 | .090 |
| | .60 | | .001 | .002 | .085 | .065 | .103 | .087 |
| | .90 | | .000 | .000 | .027 | .024 | .034 | .030 |
| | .99 | | .000 | .000 | .002 | .002 | .003 | .002 |
| L=4 | .30 | | .002 | .003 | .054 | .037 | .065 | .054 |
| | .60 | | .001 | .002 | .051 | .037 | .059 | .047 |
| | .90 | | .000 | .001 | .019 | .016 | .023 | .021 |
| | .99 | | .000 | .000 | .001 | .001 | .002 | .002 |
| L=5 | .30 | | .001 | .002 | .035 | .024 | .040 | .035 |
| | .60 | | .001 | .002 | .033 | .027 | .038 | .035 |
| | .90 | | .000 | .001 | .013 | .012 | .017 | .015 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=6 | .30 | | .001 | .002 | .025 | .018 | .030 | .026 |
| | .60 | | .001 | .002 | .024 | .020 | .029 | .026 |
| | .90 | | .000 | .001 | .010 | .009 | .013 | .012 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=7 | .30 | | .001 | .001 | .019 | .013 | .023 | .018 |
| | .60 | | .001 | .001 | .019 | .016 | .023 | .020 |
| | .90 | | .000 | .001 | .008 | .008 | .010 | .009 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=8 | .30 | | .001 | .001 | .015 | .011 | .017 | .015 |
| | .60 | | .001 | .001 | .015 | .012 | .018 | .015 |
| | .90 | | .000 | .001 | .006 | .006 | .008 | .008 |
| | .99 | | .000 | .000 | .001 | .000 | .001 | .001 |

Table 3.16. Differences between design effects using the Dalenius-Gurney equations and the linearization approach under Neyman allocation for the normal distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .30 | Identical | .000 | .000 | .220 | .217 | .285 | .283 |
| | .60 | | .000 | .000 | .177 | .171 | .228 | .224 |
| | .90 | | .000 | .000 | .042 | .040 | .051 | .049 |
| | .99 | | .000 | .000 | .004 | .004 | .005 | .005 |
| L=3 | .30 | | .000 | .000 | .128 | .135 | .169 | .148 |
| | .60 | | .000 | .001 | .122 | .120 | .144 | .043 |
| | .90 | | .001 | .001 | .034 | .032 | .042 | .040 |
| | .99 | | .000 | .000 | .003 | .003 | .004 | .003 |
| L=4 | .30 | | .000 | .000 | .097 | .107 | .120 | .132 |
| | .60 | | .000 | .001 | .080 | .080 | .086 | .087 |
| | .90 | | .001 | .001 | .027 | .025 | .032 | .031 |
| | .99 | | .000 | .000 | .002 | .002 | .002 | .002 |
| L=5 | .30 | | .000 | .000 | .066 | .076 | .078 | .091 |
| | .60 | | .000 | .001 | .056 | .058 | .070 | .072 |
| | .90 | | .001 | .001 | .021 | .020 | .026 | .029 |
| | .99 | | .000 | .000 | .001 | .001 | .002 | .002 |
| L=6 | .30 | | .000 | .000 | .051 | .063 | .062 | .076 |
| | .60 | | .000 | .001 | .045 | .048 | .054 | .057 |
| | .90 | | .001 | .001 | .017 | .016 | .021 | .020 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=7 | .30 | | .000 | .000 | .040 | .051 | .044 | .058 |
| | .60 | | .000 | .001 | .036 | .039 | .042 | .045 |
| | .90 | | .001 | .001 | .015 | .014 | .017 | .017 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=8 | .30 | | .000 | .000 | .031 | .042 | .038 | .051 |
| | .60 | | .000 | .001 | .029 | .031 | .034 | .038 |
| | .90 | | .001 | .001 | .012 | .012 | .015 | .014 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |

Table 3.17. Differences between design effects using Dalenius-Gurney equations and the linearized Singh's cum H(x) rule under Neyman allocation for the $\chi_1^2$ distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .00 | -- | .002 | .001 | .063 | .019 | .090 | .030 |
| | .30 | .001 | .001 | .001 | .155 | .081 | .190 | .128 |
| | .60 | .002 | .000 | .002 | .140 | .089 | .171 | .124 |
| | .90 | .000 | .001 | .004 | .023 | .013 | .029 | .019 |
| | .99 | .004 | .005 | .006 | .001 | .001 | .000 | .001 |
| L=3 | .00 | -- | .001 | .000 | .092 | .056 | .112 | .082 |
| | .30 | .001 | .000 | .000 | .087 | .055 | .103 | .075 |
| | .60 | .003 | .000 | .000 | .087 | .050 | .097 | .062 |
| | .90 | .001 | .000 | .001 | .019 | .012 | .023 | .016 |
| | .99 | .001 | .001 | .002 | .000 | .000 | .000 | .000 |
| L=4 | .00 | -- | .000 | .000 | .057 | .034 | .064 | .046 |
| | .30 | .001 | .000 | .000 | .068 | .039 | .075 | .049 |
| | .60 | .003 | .000 | .000 | .060 | .034 | .064 | .039 |
| | .90 | .002 | .000 | .000 | .016 | .010 | .018 | .012 |
| | .99 | .000 | .000 | .001 | .000 | .000 | .000 | .000 |
| L=5 | .00 | -- | .000 | .000 | .048 | .027 | .053 | .035 |
| | .30 | .001 | .000 | .000 | .052 | .029 | .054 | .034 |
| | .60 | .003 | .001 | .000 | .045 | .028 | .046 | .032 |
| | .90 | .002 | .001 | .000 | .013 | .009 | .015 | .010 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| L=6 | .00 | -- | .000 | .000 | .037 | .021 | .040 | .026 |
| | .30 | .001 | .000 | .000 | .041 | .024 | .044 | .029 |
| | .60 | .003 | .001 | .000 | .036 | .023 | .038 | .026 |
| | .90 | .002 | .001 | .000 | .012 | .007 | .012 | .009 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| L=7 | .00 | -- | .000 | .000 | .032 | .018 | .034 | .022 |
| | .30 | .001 | .000 | .000 | .035 | .020 | .037 | .023 |
| | .60 | .002 | .001 | .000 | .031 | .019 | .033 | .021 |
| | .90 | .002 | .001 | .000 | .010 | .007 | .011 | .007 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| L=8 | .00 | -- | .000 | .000 | .027 | .015 | .029 | .018 |
| | .30 | .001 | .000 | .000 | .030 | .018 | .031 | .020 |
| | .60 | .002 | .001 | .000 | .027 | .017 | .028 | .018 |
| | .90 | .002 | .001 | .000 | .009 | .006 | .009 | .006 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |

Table 3.18. Differences between design effects using the Dalenius-Gurney equations and the linearized Singh's cum H(x) rule under Neyman allocation for the $\chi_2^2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .002 | .001 | .090 | .027 | .133 | .085 |
| | .30 | .001 | .001 | .001 | .184 | .121 | .234 | .183 |
| L=2 | .60 | .001 | .000 | .001 | .164 | .117 | .204 | .162 |
| | .90 | .000 | .000 | .002 | .033 | .023 | .041 | .032 |
| | .99 | .001 | .002 | .002 | .000 | .000 | .000 | .000 |
| | .00 | -- | .001 | .000 | .095 | .073 | .136 | .107 |
| | .30 | .001 | .000 | .000 | .099 | .067 | .123 | .095 |
| L=3 | .60 | .002 | .000 | .000 | .102 | .064 | .115 | .081 |
| | .90 | .001 | .000 | .001 | .025 | .017 | .030 | .022 |
| | .99 | .000 | .001 | .001 | .000 | .000 | .001 | .000 |
| | .00 | -- | .001 | .000 | .065 | .044 | .078 | .062 |
| | .30 | .001 | .000 | .000 | .078 | .047 | .089 | .062 |
| L=4 | .60 | .002 | .000 | .000 | .069 | .040 | .074 | .047 |
| | .90 | .002 | .000 | .000 | .020 | .013 | .023 | .016 |
| | .99 | .000 | .000 | .000 | .000 | .000 | .001 | .000 |
| | .00 | -- | .000 | .000 | .055 | .034 | .063 | .046 |
| | .30 | .001 | .000 | .000 | .058 | .034 | .063 | .043 |
| L=5 | .60 | .002 | .000 | .000 | .051 | .032 | .054 | .038 |
| | .90 | .002 | .000 | .000 | .016 | .011 | .018 | .013 |
| | .99 | .000 | .001 | .000 | .000 | .000 | .001 | .000 |
| | .00 | -- | .000 | .000 | .042 | .027 | .047 | .035 |
| | .30 | .001 | .000 | .000 | .046 | .029 | .051 | .035 |
| L=6 | .60 | .002 | .000 | .000 | .041 | .026 | .044 | .030 |
| | .90 | .002 | .001 | .000 | .014 | .009 | .015 | .011 |
| | .99 | .000 | .000 | .000 | .001 | .000 | .001 | .000 |
| | .00 | -- | .000 | .000 | .036 | .022 | .040 | .028 |
| | .30 | .001 | .000 | .000 | .039 | .023 | .043 | .027 |
| L=7 | .60 | .002 | .000 | .000 | .035 | .022 | .037 | .025 |
| | .90 | .002 | .001 | .000 | .012 | .008 | .013 | .009 |
| | .99 | .000 | .000 | .000 | .001 | .000 | .001 | .000 |
| | .00 | -- | .000 | .000 | .031 | .019 | .034 | .021 |
| | .30 | .001 | .000 | .000 | .033 | .020 | .035 | .024 |
| L=8 | .60 | .002 | .000 | .000 | .030 | .019 | .032 | .021 |
| | .90 | .002 | .001 | .000 | .010 | .007 | .011 | .008 |
| | .99 | .000 | .000 | .000 | .001 | .000 | .011 | .000 |

Table 3.19. Differences between design effects using the Dalenius-Gurney equations and the linearized Singh's cum H(x) rule under Neyman allocation for the normal distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .000 | .000 | .153 | .150 | .206 | .204 |
| | .30 | .000 | .000 | .000 | .220 | .217 | .285 | .283 |
| L=2 | .60 | .000 | .000 | .000 | .177 | .171 | .228 | .224 |
| | .90 | .000 | .000 | .000 | .042 | .040 | .051 | .049 |
| | .99 | .000 | .000 | .000 | .004 | .004 | .005 | .005 |
| | .00 | -- | .000 | .000 | .148 | .146 | .186 | .186 |
| | .30 | .000 | .000 | .000 | .121 | .119 | .157 | .157 |
| L=3 | .60 | .000 | .000 | .000 | .115 | .111 | .133 | .131 |
| | .90 | .000 | .000 | .000 | .033 | .031 | .042 | .040 |
| | .99 | .001 | .001 | .001 | .004 | .004 | .005 | .005 |
| | .00 | -- | .000 | .000 | .092 | .093 | .111 | .112 |
| | .30 | .000 | .000 | .000 | .094 | .092 | .114 | .113 |
| L=4 | .60 | .000 | .000 | .000 | .074 | .070 | .077 | .075 |
| | .90 | .000 | .000 | .000 | .025 | .024 | .032 | .030 |
| | .99 | .001 | .001 | .001 | .003 | .002 | .003 | .003 |
| | .00 | -- | .000 | .000 | .076 | .075 | .091 | .090 |
| | .30 | .000 | .000 | .000 | .067 | .064 | .075 | .075 |
| L=5 | .60 | .000 | .000 | .000 | .052 | .050 | .063 | .061 |
| | .90 | .000 | .000 | .000 | .020 | .019 | .025 | .023 |
| | .99 | .000 | .000 | .000 | .002 | .002 | .002 | .002 |
| | .00 | -- | .000 | .000 | .058 | .058 | .068 | .068 |
| | .30 | .000 | .000 | .000 | .055 | .053 | .065 | .064 |
| L=6 | .60 | .000 | .000 | .000 | .043 | .041 | .050 | .048 |
| | .90 | .000 | .000 | .000 | .016 | .015 | .020 | .019 |
| | .99 | .000 | .000 | .000 | .001 | .001 | .002 | .002 |
| | .00 | -- | .000 | .000 | .049 | .049 | .058 | .057 |
| | .30 | .000 | .000 | .000 | .045 | .043 | .049 | .048 |
| L=7 | .60 | .000 | .000 | .000 | .035 | .033 | .039 | .037 |
| | .90 | .000 | .000 | .000 | .014 | .013 | .016 | .015 |
| | .99 | .000 | .000 | .000 | .001 | .001 | .001 | .001 |
| | .00 | -- | .000 | .000 | .041 | .041 | .048 | .048 |
| | .30 | .000 | .000 | .000 | .037 | .037 | .044 | .043 |
| L=8 | .60 | .000 | .000 | .000 | .029 | .027 | .033 | .032 |
| | .90 | .000 | .000 | .001 | .012 | .009 | .014 | .013 |
| | .99 | .000 | .000 | .000 | .001 | .001 | .001 | .001 |

Table 3.20. Differences between design effects using the Dalenius-Gurney equations and the linearized Singh-Sukhatme's cum H(x) rule under Neyman allocation for the $\chi_1^2$ distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .016 | .037 | .078 | .053 | .107 | .063 |
| | .30 | .004 | .020 | .044 | .083 | .033 | .107 | .066 |
| L=3 | .60 | .018 | .038 | .064 | .034 | .011 | .046 | .024 |
| | .90 | .062 | .078 | .097 | .008 | .014 | .005 | .009 |
| | .99 | .102 | .105 | .109 | .073 | .074 | .069 | .070 |
| | .00 | -- | .008 | .018 | .073 | .067 | .093 | .093 |
| | .30 | .002 | .010 | .021 | .031 | .026 | .047 | .045 |
| L=3 | .60 | .008 | .017 | .030 | .023 | .005 | .030 | .012 |
| | .90 | .027 | .034 | .046 | .003 | .005 | .002 | .003 |
| | .99 | .047 | .050 | .053 | .034 | .035 | .032 | .033 |
| | .00 | -- | .005 | .010 | .034 | .035 | .044 | .049 |
| | .30 | .001 | .005 | .011 | .025 | .016 | .034 | .027 |
| L=4 | .60 | .004 | .009 | .016 | .014 | .003 | .018 | .005 |
| | .90 | .014 | .018 | .025 | .002 | .002 | .002 | .001 |
| | .99 | .025 | .027 | .030 | .018 | .019 | .017 | .018 |
| | .00 | -- | .003 | .006 | .025 | .023 | .033 | .032 |
| | .30 | .000 | .003 | .007 | .017 | .009 | .021 | .015 |
| L=5 | .60 | .002 | .005 | .010 | .009 | .003 | .011 | .006 |
| | .90 | .008 | .010 | .015 | .001 | .001 | .001 | .001 |
| | .99 | .015 | .016 | .018 | .011 | .012 | .010 | .011 |
| | .00 | -- | .002 | .004 | .017 | .015 | .021 | .021 |
| | .30 | .000 | .002 | .005 | .011 | .007 | .015 | .012 |
| L=6 | .60 | .001 | .003 | .007 | .007 | .003 | .010 | .005 |
| | .90 | .005 | .007 | .010 | .001 | .001 | .001 | .001 |
| | .99 | .010 | .011 | .012 | .007 | .008 | .007 | .008 |
| | .00 | -- | .001 | .003 | .013 | .011 | .017 | .016 |
| | .30 | .000 | .002 | .003 | .010 | .005 | .012 | .008 |
| L=7 | .60 | .001 | .002 | .005 | .007 | .002 | .008 | .004 |
| | .90 | .003 | .004 | .007 | .001 | .000 | .001 | .001 |
| | .99 | .007 | .007 | .009 | .005 | .006 | .005 | .005 |
| | .00 | -- | .001 | .002 | .010 | .008 | .013 | .012 |
| | .30 | .000 | .001 | .003 | .008 | .004 | .009 | .007 |
| L=8 | .60 | .001 | .002 | .003 | .006 | .002 | .007 | .003 |
| | .90 | .002 | .003 | .005 | .001 | .000 | .001 | .000 |
| | .99 | .005 | .005 | .007 | .003 | .004 | .003 | .004 |

Table 3.21. Differences between design effects using the Dalenius-Gurney equations and the linearized Singh-Sukhatme's cum H(x) rule under Neyman allocation for the $\chi_2^2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .014 | .026 | .103 | .047 | .143 | .101 |
| | .30 | .003 | .017 | .032 | .112 | .062 | .147 | .107 |
| L=2 | .60 | .013 | .029 | .047 | .055 | .027 | .074 | .048 |
| | .90 | .043 | .056 | .068 | .001 | .004 | .000 | .001 |
| | .99 | .069 | .071 | .073 | .043 | .044 | .040 | .040 |
| | .00 | -- | .008 | .014 | .083 | .102 | .105 | .101 |
| | .30 | .001 | .009 | .017 | .040 | .061 | .059 | .054 |
| L=3 | .60 | .006 | .011 | .025 | .034 | .014 | .044 | .026 |
| | .90 | .020 | .027 | .037 | .003 | .003 | .004 | .002 |
| | .99 | .036 | .038 | .040 | .023 | .024 | .021 | .022 |
| | .00 | -- | .005 | .008 | .038 | .040 | .051 | .056 |
| | .30 | .001 | .005 | .010 | .032 | .020 | .042 | .033 |
| L=4 | .60 | .003 | .008 | .015 | .021 | .005 | .025 | .009 |
| | .90 | .011 | .015 | .022 | .003 | .002 | .004 | .002 |
| | .99 | .020 | .022 | .023 | .013 | .014 | .012 | .013 |
| | .00 | -- | .003 | .005 | .031 | .027 | .040 | .038 |
| | .30 | .000 | .003 | .006 | .021 | .012 | .026 | .021 |
| L=5 | .60 | .002 | .005 | .009 | .013 | .005 | .016 | .010 |
| | .90 | .006 | .009 | .014 | .002 | .001 | .003 | .002 |
| | .99 | .012 | .014 | .015 | .008 | .009 | .008 | .008 |
| | .00 | -- | .002 | .004 | .020 | .019 | .026 | .027 |
| | .30 | .000 | .002 | .004 | .015 | .010 | .020 | .016 |
| L=6 | .60 | .001 | .003 | .007 | .010 | .005 | .014 | .008 |
| | .90 | .004 | .006 | .010 | .002 | .001 | .003 | .001 |
| | .99 | .008 | .009 | .011 | .006 | .006 | .005 | .006 |
| | .00 | -- | .002 | .003 | .016 | .014 | .021 | .020 |
| | .30 | .000 | .002 | .003 | .012 | .007 | .016 | .011 |
| L=7 | .60 | .001 | .002 | .005 | .009 | .004 | .011 | .006 |
| | .90 | .003 | .004 | .007 | .002 | .001 | .002 | .001 |
| | .99 | .006 | .007 | .008 | .004 | .005 | .004 | .004 |
| | .00 | -- | .001 | .002 | .012 | .011 | .016 | .013 |
| | .30 | .000 | .001 | .002 | .010 | .006 | .012 | .010 |
| L=8 | .60 | .001 | .002 | .004 | .008 | .003 | .009 | .004 |
| | .90 | .002 | .003 | .006 | .002 | .001 | .002 | .001 |
| | .99 | .004 | .005 | .006 | .003 | .003 | .003 | .003 |

Table 3.22. Differences between design effects using the Dalenius-Gurney equations and the linearized Singh-Sukhatme's cum H(x) rule under Neyman allocation for the normal distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .00 | -- | .000 | .000 | .153 | .150 | .206 | .204 |
| | .30 | .000 | .000 | .000 | .220 | .217 | .285 | .283 |
| | .60 | .000 | .000 | .000 | .177 | .171 | .344 | .224 |
| | .90 | .000 | .000 | .000 | .042 | .040 | .051 | .049 |
| | .99 | .000 | .000 | .000 | .004 | .004 | .005 | .005 |
| L=3 | .00 | -- | .000 | .000 | .106 | .104 | .130 | .129 |
| | .30 | .001 | .001 | .001 | .086 | .085 | .110 | .110 |
| | .60 | .003 | .003 | .004 | .100 | .097 | .112 | .110 |
| | .90 | .009 | .009 | .009 | .042 | .041 | .051 | .050 |
| | .99 | .014 | .014 | .014 | .017 | .017 | .018 | .018 |
| L=4 | .00 | -- | .000 | .000 | .055 | .056 | .066 | .067 |
| | .30 | .001 | .001 | .001 | .063 | .060 | .075 | .073 |
| | .60 | .003 | .003 | .003 | .058 | .054 | .055 | .053 |
| | .90 | .007 | .007 | .007 | .032 | .031 | .038 | .037 |
| | .99 | .012 | .012 | .012 | .014 | .014 | .015 | .015 |
| L=5 | .00 | -- | .000 | .000 | .045 | .044 | .054 | .054 |
| | .30 | .000 | .000 | .001 | .040 | .037 | .046 | .043 |
| | .60 | .002 | .002 | .002 | .037 | .035 | .045 | .043 |
| | .90 | .005 | .005 | .005 | .024 | .023 | .029 | .028 |
| | .99 | .009 | .009 | .009 | .011 | .010 | .011 | .011 |
| L=6 | .00 | -- | .000 | .000 | .032 | .032 | .038 | .038 |
| | .30 | .000 | .000 | .000 | .032 | .030 | .038 | .037 |
| | .60 | .001 | .001 | .001 | .030 | .028 | .034 | .032 |
| | .90 | .003 | .003 | .004 | .019 | .018 | .022 | .021 |
| | .99 | .006 | .006 | .006 | .008 | .008 | .008 | .008 |
| L=7 | .00 | -- | .000 | .000 | .027 | .026 | .032 | .032 |
| | .30 | .000 | .000 | .000 | .025 | .023 | .026 | .026 |
| | .60 | .001 | .001 | .001 | .023 | .021 | .026 | .024 |
| | .90 | .002 | .002 | .003 | .015 | .014 | .017 | .017 |
| | .99 | .005 | .005 | .005 | .006 | .006 | .006 | .006 |
| L=8 | .00 | -- | .000 | .000 | .021 | .021 | .025 | .026 |
| | .30 | .000 | .000 | .000 | .020 | .019 | .024 | .023 |
| | .60 | .001 | .001 | .001 | .018 | .016 | .020 | .019 |
| | .90 | .002 | .002 | .002 | .012 | .011 | .014 | .013 |
| | .99 | .003 | .003 | .004 | .004 | .004 | .005 | .005 |

Table 3.23. Differences between design effects using the Dalenius-Gurney equations and Singh-Parkash's cum H(x) rule under equal allocation for the $\chi_1^2$ distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .001 | .003 | .028 | .005 | .046 | .006 |
| | .30 | .002 | .003 | .004 | .097 | .064 | .156 | .091 |
| L=2 | .60 | .044 | .027 | .008 | .262 | .169 | .382 | .240 |
| | .90 | .248 | .111 | .016 | .298 | .125 | .350 | .158 |
| | .99 | .358 | .155 | .020 | .207 | .045 | .216 | .050 |
| | .00 | -- | .000 | .003 | .074 | .049 | .163 | 091 |
| | .30 | .003 | .002 | .004 | .181 | .102 | .296 | .159 |
| L=3 | .60 | .042 | .017 | .007 | .262 | .134 | .358 | .182 |
| | .90 | .213 | .081 | .016 | .252 | .102 | .298 | .127 |
| | .99 | .319 | .123 | .021 | .173 | .044 | .182 | .048 |
| | .00 | -- | .000 | .002 | .112 | .062 | .207 | .104 |
| | .30 | .002 | .001 | .003 | .191 | .095 | .290 | .141 |
| L=4 | .60 | .037 | .012 | .006 | .237 | .112 | .320 | .153 |
| | .90 | .178 | .060 | .014 | .211 | .083 | .250 | .102 |
| | .99 | .271 | .097 | .019 | .141 | .038 | .150 | .042 |
| | .00 | -- | .000 | .002 | .124 | .062 | .209 | .099 |
| | .30 | .002 | .001 | .003 | .185 | .087 | .272 | .127 |
| L=5 | .60 | .033 | .009 | .005 | .216 | .096 | .291 | .132 |
| | .90 | .152 | .046 | .011 | .181 | .069 | .216 | .086 |
| | .99 | .232 | .078 | .017 | .118 | .033 | .126 | .036 |
| | .00 | -- | .000 | .002 | .126 | .059 | .202 | .093 |
| | .30 | .002 | .001 | .002 | .176 | .079 | .256 | .116 |
| L=6 | .60 | .029 | .007 | .004 | .199 | .085 | .268 | .117 |
| | .90 | .133 | .038 | .010 | .159 | .060 | .190 | .074 |
| | .99 | .202 | .065 | .015 | .101 | .029 | .108 | .032 |
| | .00 | -- | .000 | .002 | .124 | .057 | .194 | .088 |
| | .30 | .002 | .001 | .002 | .168 | .074 | .242 | .107 |
| L=7 | .60 | .026 | .006 | .004 | .184 | .077 | .248 | .105 |
| | .90 | .117 | .031 | .008 | .142 | .052 | .170 | .065 |
| | .99 | .178 | .055 | .013 | .088 | .026 | .094 | .028 |
| | .00 | -- | .000 | .002 | .121 | .054 | .187 | .083 |
| | .30 | .002 | .000 | .002 | .159 | .069 | .229 | .099 |
| L=8 | .60 | .023 | .005 | .003 | .171 | .070 | .231 | .096 |
| | .90 | .105 | .027 | .007 | .128 | .047 | .155 | .059 |
| | .99 | .160 | .047 | .012 | .077 | .023 | .083 | .025 |

Table 3.24. Differences between design effects using the Dalenius-Gurney equations and Singh-Parkash's cum H(x) rule under equal allocation for the $\chi_2^2$ distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .00 | -- | .000 | .000 | .005 | .000 | .009 | .001 |
| | .30 | .002 | .001 | .000 | .079 | .047 | .136 | .081 |
| | .60 | .029 | .008 | .000 | .240 | .134 | .353 | .204 |
| | .90 | .155 | .039 | .000 | .215 | .068 | .265 | .095 |
| | .99 | .225 | .057 | .000 | .102 | .005 | .111 | .008 |
| L=3 | .00 | -- | .000 | .001 | .069 | .050 | .153 | .101 |
| | .30 | .002 | .001 | .001 | .162 | .081 | .265 | .136 |
| | .60 | .026 | .007 | .002 | .224 | .103 | .309 | .144 |
| | .90 | .139 | .038 | .007 | .193 | .066 | .236 | .088 |
| | .99 | .208 | .058 | .010 | .101 | .019 | .110 | .021 |
| L=4 | .00 | -- | .000 | .000 | .102 | .054 | .186 | .098 |
| | .30 | .001 | .001 | .001 | .166 | .074 | .253 | .116 |
| | .60 | .022 | .006 | .002 | .202 | .085 | .277 | .121 |
| | .90 | .116 | .032 | .007 | .165 | .057 | .201 | .074 |
| | .99 | .179 | .051 | .011 | .089 | .020 | .096 | .022 |
| L=5 | .00 | -- | .000 | .000 | .109 | .050 | .183 | .086 |
| | .30 | .001 | .000 | .001 | .159 | .065 | .236 | .102 |
| | .60 | .019 | .004 | .001 | .182 | .072 | .250 | .102 |
| | .90 | .099 | .026 | .006 | .142 | .048 | .174 | .062 |
| | .99 | .154 | .044 | .010 | .077 | .019 | .084 | .021 |
| L=6 | .00 | -- | .000 | .000 | .109 | .046 | .176 | .077 |
| | .30 | .001 | .000 | .000 | .150 | .058 | .221 | .090 |
| | .60 | .016 | .004 | .001 | .166 | .062 | .227 | .088 |
| | .90 | .085 | .022 | .005 | .125 | .041 | .153 | .053 |
| | .99 | .134 | .038 | .009 | .068 | .017 | .074 | .019 |
| L=7 | .00 | -- | .000 | .000 | .106 | .041 | .168 | .069 |
| | .30 | .001 | .000 | .000 | .141 | .052 | .208 | .081 |
| | .60 | .014 | .003 | .001 | .152 | .054 | .209 | .078 |
| | .90 | .075 | .019 | .004 | .111 | .036 | .137 | .046 |
| | .99 | .119 | .033 | .008 | .060 | .015 | .065 | .017 |
| L=8 | .00 | -- | .000 | .000 | .099 | .038 | .161 | .063 |
| | .30 | .001 | .000 | .000 | .133 | .047 | .195 | .073 |
| | .60 | .013 | .002 | .001 | .141 | .049 | .193 | .070 |
| | .90 | .067 | .016 | .003 | .100 | .032 | .124 | .041 |
| | .99 | .106 | .029 | .007 | .054 | .014 | .059 | .015 |

Table 3.25. Differences between design effects using the Dalenius-Gurney equations and Singh-Parkash's cum H(x) rule under equal allocation for the normal distribution.

| | ρ | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| | .00 | -- | .000 | .000 | .000 | .000 | .000 | .000 |
| | .30 | .000 | .000 | .000 | .077 | .067 | .125 | .109 |
| L=2 | .60 | .000 | .000 | .002 | .152 | .125 | .217 | .181 |
| | .90 | .000 | .002 | .008 | .070 | .046 | .098 | .055 |
| | .99 | .000 | .003 | .012 | .003 | .000 | .005 | .000 |
| | .00 | -- | .000 | .000 | .090 | .095 | .191 | .200 |
| | .30 | .000 | .000 | .000 | .077 | .063 | .145 | .127 |
| L=3 | .60 | .002 | .002 | .004 | .104 | .084 | .142 | .115 |
| | .90 | .014 | .016 | .022 | .072 | .051 | .095 | .070 |
| | .99 | .023 | .026 | .035 | .025 | .023 | .027 | .023 |
| | .00 | -- | .000 | .000 | .098 | .107 | .181 | .194 |
| | .30 | .000 | .000 | .000 | .083 | .074 | .155 | .147 |
| L=4 | .60 | .002 | .002 | .003 | .082 | .063 | .118 | .093 |
| | .90 | .016 | .018 | .023 | .062 | .046 | .082 | .061 |
| | .99 | .028 | .031 | .039 | .030 | .029 | .032 | .029 |
| | .00 | -- | .000 | .000 | .096 | .106 | .171 | .186 |
| | .30 | .000 | .000 | .000 | .080 | .075 | .145 | .141 |
| L=5 | .60 | .002 | .002 | .003 | .070 | .054 | .108 | .089 |
| | .90 | .015 | .017 | .021 | .053 | .039 | .069 | .052 |
| | .99 | .028 | .031 | .038 | .030 | .029 | .031 | .029 |
| | .00 | -- | .000 | .000 | .090 | .101 | .158 | .174 |
| | .30 | .000 | .000 | .000 | .075 | .072 | .134 | .133 |
| L=6 | .60 | .001 | .002 | .003 | .062 | .048 | .098 | .083 |
| | .90 | .013 | .015 | .019 | .046 | .034 | .060 | .045 |
| | .99 | .027 | .029 | .036 | .028 | .027 | .029 | .027 |
| | .00 | -- | .000 | .000 | .084 | .096 | .146 | .163 |
| | .30 | .000 | .000 | .000 | .070 | .069 | .124 | .124 |
| L=7 | .60 | .001 | .002 | .002 | .055 | .044 | .089 | .077 |
| | .90 | .012 | .013 | .017 | .040 | .030 | .052 | .039 |
| | .99 | .024 | .027 | .033 | .026 | .025 | .027 | .025 |
| | .00 | -- | .000 | .000 | .079 | .090 | .136 | .152 |
| | .30 | .000 | .000 | .000 | .065 | .065 | .115 | .116 |
| L=8 | .60 | .001 | .001 | .002 | .050 | .040 | .082 | .071 |
| | .90 | .010 | .012 | .015 | .035 | .026 | .046 | .034 |
| | .99 | .022 | .025 | .030 | .024 | .023 | .025 | .023 |

Table 3.26. Differences between design effects using the Dalenius-Gurney equations and the linearization approach under equal allocation for the $\chi_1^2$ distribution.

| | ρ | C=0 r=0 | C=0 r=1 | C=0 r=2 | C=1 r=1 | C=1 r=2 | C=2 r=1 | C=2 r=2 |
|---|---|---|---|---|---|---|---|---|
| L=2 | .30 | | .032 | .087 | .122 | .120 | .148 | .114 |
| | .60 | | .001 | .003 | .119 | .121 | .185 | .175 |
| | .90 | | .000 | .000 | .055 | .058 | .079 | .081 |
| | .99 | | .000 | .000 | .006 | .006 | .008 | .008 |
| L=3 | .30 | | .036 | .103 | .227 | .196 | .301 | .212 |
| | .60 | | .007 | .015 | .115 | .083 | .144 | .103 |
| | .90 | | .000 | .000 | .030 | .031 | .042 | .044 |
| | .99 | | .000 | .000 | .004 | .004 | .005 | .005 |
| L=4 | .30 | | .039 | .113 | .250 | .202 | .305 | .204 |
| | .60 | | .010 | .024 | .101 | .069 | .116 | .078 |
| | .90 | | .001 | .001 | .019 | .019 | .026 | .026 |
| | .99 | | .000 | .000 | .002 | .002 | .003 | .003 |
| L=5 | .30 | Identical | .041 | .122 | .253 | .200 | .204 | .196 |
| | .60 | | .013 | .032 | .091 | .061 | .102 | .065 |
| | .90 | | .001 | .002 | .014 | .013 | .018 | .017 |
| | .99 | | .000 | .000 | .001 | .002 | .002 | .002 |
| L=6 | .30 | | .043 | .129 | .250 | .197 | .284 | .189 |
| | .60 | | .015 | .038 | .085 | .056 | .092 | .058 |
| | .90 | | .002 | .003 | .011 | .009 | .014 | .012 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .002 |
| L=7 | .30 | | .045 | .136 | .246 | .194 | .275 | .182 |
| | .60 | | .017 | .044 | .080 | .055 | .085 | .053 |
| | .90 | | .002 | .004 | .009 | .007 | .011 | .009 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |
| L=8 | .30 | | .046 | .142 | .241 | .192 | .267 | .177 |
| | .60 | | .019 | .050 | .077 | .054 | .080 | .050 |
| | .90 | | .002 | .004 | .008 | .006 | .010 | .008 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |

Table 3.27. Differences between design effects using the Dalenius-Gurney equations and the linearization approach under equal allocation for the $\chi_2^2$ distribution.

| | ρ | C=0 r=0 | C=0 r=1 | C=0 r=2 | C=1 r=1 | C=1 r=2 | C=2 r=1 | C=2 r=2 |
|---|---|---|---|---|---|---|---|---|
| L=2 | .30 | Identical | .026 | .042 | .107 | .103 | .140 | .125 |
| | .60 | | .000 | .003 | .148 | .159 | .226 | .227 |
| | .90 | | .000 | .000 | .072 | .077 | .102 | .106 |
| | .99 | | .000 | .000 | .008 | .008 | .011 | .011 |
| L=3 | .30 | | .035 | .066 | .206 | .134 | .269 | .157 |
| | .60 | | .007 | .011 | .107 | .088 | .138 | .113 |
| | .90 | | .001 | .001 | .040 | .045 | .057 | .062 |
| | .99 | | .000 | .000 | .005 | .005 | .007 | .007 |
| L=4 | .30 | | .041 | .084 | .221 | .134 | .263 | .137 |
| | .60 | | .012 | .020 | .088 | .061 | .105 | .075 |
| | .90 | | .001 | .001 | .024 | .027 | .035 | .038 |
| | .99 | | .000 | .000 | .003 | .003 | .004 | .004 |
| L=5 | .30 | | .045 | .098 | .221 | .134 | .253 | .129 |
| | .60 | | .016 | .028 | .077 | .049 | .088 | .056 |
| | .90 | | .002 | .002 | .016 | .018 | .023 | .025 |
| | .99 | | .000 | .000 | .002 | .002 | .003 | .003 |
| L=6 | .30 | | .048 | .110 | .217 | .134 | .243 | .124 |
| | .60 | | .019 | .035 | .071 | .043 | .078 | .045 |
| | .90 | | .003 | .003 | .012 | .013 | .016 | .018 |
| | .99 | | .000 | .000 | .001 | .002 | .002 | .002 |
| L=7 | .30 | | .050 | .120 | .212 | .134 | .234 | .120 |
| | .60 | | .021 | .042 | .067 | .040 | .071 | .039 |
| | .90 | | .003 | .004 | .009 | .010 | .012 | .008 |
| | .99 | | .000 | .000 | .001 | .001 | .002 | .002 |
| L=8 | .30 | | .052 | .128 | .206 | .134 | .225 | .117 |
| | .60 | | .023 | .048 | .064 | .039 | .066 | .036 |
| | .90 | | .004 | .005 | .008 | .008 | .010 | .010 |
| | .99 | | .000 | .000 | .001 | .001 | .001 | .001 |

Table 3.28. Differences between design effects using the Dalenius-Gurney equations and the linearization approach under equal allocation for the normal distribution.

| | | C=0 | | | C=1 | | C=2 | |
|---|---|---|---|---|---|---|---|---|
| | ρ | r=0 | r=1 | r=2 | r=1 | r=2 | r=1 | r=2 |
| L=2 | .30 | Identical | .003 | .010 | .104 | .122 | .155 | .160 |
| | .60 | | .001 | .005 | .195 | .208 | .266 | .276 |
| | .90 | | .000 | .000 | .103 | .106 | .136 | .138 |
| | .99 | | .000 | .000 | .013 | .013 | .017 | .017 |
| L=3 | .30 | | .019 | .045 | .040 | .037 | .078 | .069 |
| | .60 | | .018 | .025 | .088 | .099 | .107 | .115 |
| | .90 | | .002 | .003 | .067 | .070 | .088 | .090 |
| | .99 | | .000 | .000 | .008 | .008 | .011 | .011 |
| L=4 | .30 | | .025 | .069 | .025 | .020 | .055 | .044 |
| | .60 | | .030 | .040 | .044 | .051 | .051 | .053 |
| | .90 | | .004 | .005 | .045 | .048 | .058 | .060 |
| | .99 | | .000 | .000 | .005 | .006 | .007 | .007 |
| L=5 | .30 | | .026 | .083 | .017 | .013 | .037 | .024 |
| | .60 | | .036 | .051 | .027 | .033 | .032 | .034 |
| | .90 | | .006 | .007 | .032 | .035 | .040 | .042 |
| | .99 | | .000 | .000 | .004 | .004 | .005 | .005 |
| L=6 | .30 | | .025 | .093 | .013 | .012 | .027 | .015 |
| | .60 | | .040 | .059 | .019 | .025 | .021 | .023 |
| | .90 | | .008 | .009 | .025 | .027 | .030 | .032 |
| | .99 | | .000 | .000 | .003 | .003 | .004 | .004 |
| L=7 | .30 | | .024 | .099 | .011 | .012 | .021 | .010 |
| | .60 | | .041 | .065 | .016 | .022 | .015 | .018 |
| | .90 | | .009 | .011 | .020 | .023 | .023 | .025 |
| | .99 | | .000 | .000 | .002 | .002 | .003 | .003 |
| L=8 | .30 | | .022 | .104 | .009 | .014 | .017 | .008 |
| | .60 | | .041 | .070 | .014 | .020 | .012 | .014 |
| | .90 | | .011 | .012 | .017 | .019 | .018 | .020 |
| | .99 | | .000 | .000 | .002 | .002 | .002 | .002 |

## 3.5. Summary

Consider Table 3.29, in which the approximations we have examined in this chapter have been ranked by their poorest fit over the range of parameters when $L \geq 4$.

Table 3.29. Ranking the proposed approximations by the worst fits with the Dalenius-Gurney equations for $L \geq 4$.

| | C=0 | | C≠0 | |
|---|---|---|---|---|
| Allocation | Approximation | Poorest Fit | Approximation | Poorest Fit |
| Proportional | Linearization | .003[a] | Singh's | .081 |
| | Singh's | .009[b] | Singh-Sukhatme's | .085 |
| | Singh-Sukhatme's | | cum $\sqrt[3]{f(x)}$ | .132 |
| | cum $\sqrt[3]{f(x)}$ | | Linearization | .213 |
| Neyman | Linearization | .003 | Singh's | .053 |
| | cum $\sqrt{f(x)}$ | .003 | cum $\sqrt[3]{f(x)}$ | .075 |
| | Singh's | .010 | Singh-Sukhatme's | .080 |
| | Singh-Sukhatme's | .025 | cum $\sqrt{f(x)}$ | .114 |
| | cum $\sqrt[3]{f(x)}$ | .030 | Linearization | .132 |
| Equal: Symmetric Distribution | Singh-Parkash's | .039 | Linearization | .060 |
| | Linearization | .104 | Singh-Parkash's | .194 |
| Equal: Skewed Distribution | Linearization | .142 | Linearization | .305 |
| | Singh-Parkash's | .271 | Singh-Parkash's | .320 |

[a]The approximation generates the optimum boundaries.

[b]These three approximations are equivalent.

Thus, for the linear model under proportional and Neyman allocations, approximate procedures exist which offer excellent fits. For the non-linear model, approximations are available which, while not as good as those for the linear model, still offer reasonable fits. It is interesting that for the non-linear model under Neyman allocation, the cum $\sqrt[3]{f(x)}$ rule, though possessing no information about the functional relationship between X and Y, nonetheless offers an approximation whose worst fit is actually better than that of Singh-Sukhatme's cum H(x) rule. It is also interesting that if we were to use the above table to decide upon one appxorimation to use for both linear and non-linear models, Singh's cum H(x) rule would be chosen for both proportional and Neyman allocations. Finally, we note that with judicious use of the cum $\sqrt{f(x)}$ and cum $\sqrt[3]{f(x)}$ rules, we are able to employ cum H(x) rules which do not depend upon the functional relationship between X and Y, and yet provide nearly as good a fit as the best approximation in all cases except the non-linear under proportional.

However, for equal allocation, it is only when the underlying distribution is symmetric that a reasonable approximation exists; for the skewed distributions, no approximation is available whose worst fit for the non-linear model is less than .300.

# IV. EFFICIENT NUMBER OF STRATA

When the method of allocation and the mechanism for computing the stratum boundaries have been chosen, it remains to decide upon the number of strata. When cost does not enter into this decision, the limiting factor is that point, the efficient number of strata, beyond which any additional strata produce only marginal reductions in the variance of $\hat{\mu}$. Let us define $V_L=\mathrm{var}[\hat{\mu}]$ when the L-1 stratum boundaries are determined by some optimum stratification procedure. Then, most criteria suggested in the literature for judging when this point is reached involve the ratio $V_L/V_{L-1}$, a function of the design effect, $D(L)=V_L/V_1$.

Since $V_L$ is not known, it is necessary to approximate it. These approximations then replace the actual variances to decide upon the number of strata to employ. However, the few studies in the literature that have examined the suggested approximations against the actual distributions chose merely to compare the ratios $D(L)/D(L-1)$ as L varied to judge their performance. It is felt this comparison is potentially too vague. Consequently, we introduce an algorithm that allows us to explicitly compute the efficient number of strata. With this algorithm, we are able to judge the effectiveness of an approximation by comparing its efficient number of strata with the efficient number generated by the actual distributions.

To employ this algorithm, it is necessary to be able to express the design effect as a continuous function of L. While the approximations suggested in the literature satisfy this condition, the design effects from the actual distributions are only defined for integer values of L. Hence, we also introduce a continuous function of L which is used to approximate these design effects.

In the first section, we introduce the algorithm and continuous function of L, while in the second section, we will discuss the approximations to the design effects suggested in the literature. In the third section, we will use the procedure derived in the first section to compare these approximations with the actual design effects from use of Dalenius's original equations in Chapter II, and from use of the Dalenius-Gurney equations in Chapter III.

## 4.1. Quantifying the Effect of Increasing the Number of Strata

We will first introduce an algorithm designed to replace the ratio of succeeding design effects as a measure of the effectiveness of the addition of another strata. If we assume that the design effect is a continuous function of L, $L>0$, then the derivative of this function evaluated at a specific integer value of L would provide an estimate of the amount of reduction experienced by the design effect as the number of strata increases from L to L+1. For instance, if we assume the design effect can be represented by $D(L)=1/L^2$, then Table 4.1 lists the derivatives of this function for several integer values of L.

The design effect as a continuous function of L is generally convex and non-increasing, with its derivatives non-decreasing. Hence, by the Mean Value Theorem, the decrease in the design effect from L to L+1 is

Table 4.1 Derivatives of $D(L)=1/L^2$.

| Number of Strata | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| $\frac{dD(L)}{dL} = -\frac{2}{L^3}$ | -.250 | -.074 | -.031 | -.016 | -.009 | -.006 | -.004 |

bounded from above* by the derivative evaluated at L, and from below by the derivative evaluated at L+1. For instance, if $D(L)=1/L^2$, then reference to Table 4.1 indicates that we would expect the percentage decrease in the design effect from 4 to 5 strata to be between 3.1 percent and 1.6 percent; the actual percentage decrease is 2.25 percent.

Thus, a conservative estimate of the expected reduction in the design effect by increasing the number of strata from L to L+1 is the value of the derivative of D(L) at L. We employ this measure in the following algorithm to decide when the addition of another stratum no longer produces a significant reduction in var$[\hat{\mu}]$: after specifying the smallest percentage decrease in the design effect which would justify the addition of another stratum, the values of the derivative of D(L) for various integer L's are examined; the "efficient" number of strata is then defined as the smallest L such that the derivative at that point is below the specified cutoff percentage. Since the derivative at L is

*Note that the negative sign of the derivative, indicating the decreasing nature of the design effect, could cause some confusion. Since our sole interest in the derivatives is measuring the magnitude of reduction of the design effect, we introduce the convention of treating the derivative in our discussion as an absolute value, thereby attempting to avoid as much misunderstanding as possible. For instance, using Table 4.1 and comparing the derivatives evaluated at L=3 and L=4, we note that the effect of incrementing the stratum number from 3 to 4 offers us more reduction than incrementing the stratum number from 4 to 5 (7.4% versus 3.1%). Thus, it makes sense to order -.074 as "greater" than -.031.

an upper bound of the actual percentage decrease incurred by incrementing the number of strata to L+1, if the derivative is above the specified cutoff percentage, there is a chance that the addition of another stratum may offer that much reduction. However, as soon as the derivative falls below that point, the non-decreasing nature of the derivatives implies that any future increments would only offer reductions below that specified.

As an example of this algorithm when $D(L)=1/L^2$, using Table 4.1 and choosing the arbitrary cutoff percentage of 5 percent, we note that when L=4, the derivative of D(L) falls "below" -.05 for the first time. Hence, 4 would be our efficient number of strata for a 5 percent cutoff point. Table 4.2 lists the efficient number of strata for several cutoff percentages when $D(L)=1/L^2$.

Table 4.2. Efficient number of strata when $D(L)=1/L^2$.

| Cutoff Percentage | 10% | 5% | 2.5% | 1% |
|---|---|---|---|---|
| Efficient Number of Strata | 3 | 4 | 5 | 6 |

Note that we choose to employ the derivative of the continuous design effect evaluated at L (the conservative upper bound of the actual reduction) instead of the equally available actual reduction to compare with our arbitrary cutoff percentage. This choice is, of course, entirely arbitrary. The reason for it is that since it is not possible to express the actual design effects as a continuous function of L, we must resort to approximations. And, since we have no guarantee that these approximations will be accurate, we choose the more conservative route to better insure that our approach will allow us to attain the degree of reduction specified by the cutoff percentage.

To employ this procedure on the design effects from the actual distributions, we used the design effects from Dalenius's original equations in Chapter II, and the design effects from the Dalenius-Gurney equations in Chapter III, and considered five functions of L as approximations to these design effects. The functions we considered were:

$$D_1(L) = A + B/L + e$$

$$D_2(L) = A + B/L + C/L^2 + e$$

$$D_3(L) = (A + BL + e)^{-1}$$

$$D_4(L) = (A + BL + CL^2 + e)^{-1}$$

$$\text{and } \ln D_5(L) = A + BL + e$$

(Note that the fourth model was the one proposed and investigated by Sethi; see page 14.)

Fitting these functions to the design effects by minimum expected squared error technique (a continuous analog to least squares procedure), it was found that $D_2$ offered the most consistently superior fit. In fact, over the range of models considered in our empirical work, in only a few cases did the $R^2$ value (percentage of total variation explained by the model) fall below .999; the smallest value of $R^2$ among all the fits for $D_2$ was .951. On the other hand, only $D_4$ (Sethi's model) offered almost as good a fit, though the only times that $D_4$ offered an $R^2$ value which was greater than $D_2$'s was when they were both greater than .999. Furthermore, in Chapter III, when dealing with the design effects from the optimum stratum boundaries, the smallest $R^2$ of $D_2$ was .997; in Chapter II, the worst $R^2$ was .951. For $D_4$, the respective smallest $R^2$ values for the design effects from Chapter III and Chapter II were .954 and .728.

## 4.2. Design Effects and the Efficient Number of Strata

There have been basically two types of approximations that have been proposed in the literature to approximate the design effect to determine the effect the addition of further strata has upon the reduction in the variance of $\hat{\mu}$. The first one, proposed initially by Dalenius and Gurney (1951), is to employ a hypothetical distribution to approximate the study variable, using its design effects to make the decisions. Empirical studies dealing with Neyman allocation found that the rectangular distribution worked well as an approximation when the study variable's distribution was skewed and when the ratio of succeeding design effects, $V_L/V_{L-1}$, was employed to judge the effect of further stratification. Assuming that the study variable's distribution is rectangular, when Dalenius's original equations are used to determine the optimum stratum boundaries, all three methods of allocation we have considered reduce to the same system of equations, $y_h - y_{h-1} = \text{constant}, \forall h$, and the var$[\hat{\mu}|\cdot]$'s all reduce to $V_L = \frac{1}{nL^2}\sigma_y^2$. Thus, Table 4.2 offers one approximation to the efficient number of strata.

We may employ this method of approximating the design effect in a more sensitive way. In particular, let us use the distribution of the auxiliary variable as our hypothetical distribution, using it to derive an approximation to the efficient number of strata. Assuming the study and auxiliary variables are identical, the best design effects for a particular method of allocation and for a reasonable range of L are determined on the auxiliary variable. A continuous function of L is then computed by fitting the function $V_L/V_1=A+B/L+C/L^2+e$ by minimum expected squared error. We then use the derivatives of this function with the algorithm outlined in the previous section to generate a more sensitive series of approximations to the efficient number of strata. These approximations are listed in Table 4.3.

Table 4.3. Efficient number of strata when the auxiliary variable is used as the hypothetical distribution.

| | Proportional Allocation | | | Neyman/Equal Allocation | | |
|---|---|---|---|---|---|---|
| Cutoff | $\chi^2_1$ | $\chi^2_2$ | Normal | $\chi^2_1$ | $\chi^2_2$ | Normal |
| 10% | 3 | 3 | 3 | 3 | 3 | 4 |
| 5% | 4 | 4 | 4 | 4 | 4 | 4 |
| 2.5% | 6 | 6 | 6 | 5 | 5 | 6 |
| 1% | 8 | 8 | 8 | 6 | 7 | 8 |

Note that comparing Table 4.2 with Table 4.3 under the $\chi^2_1$ distribution for Neyman allocation agrees with the previous results in the literature that the rectangular distribution offers a good approximation to skewed distributions for determining the effect of increasing the number of strata. However, as the degree of skewness decreases, use of the rectangular distribution as an approximation results in underestimating the actual effect of increasing the number of strata.

This approach of employing a hypothetical distribution assumed that the stratification was done on the study variable. However, since this is not true in practice, the second type of approximation which has appeared in the literature introduces a simple linear regression model relating the study variable to the auxiliary variable, and then approximates the design effect as a function of the number of strata, the correlation between the two variables, and the distribution of the auxiliary variable. The general form of these approximations is

$$(4.1) \qquad D(L) = K_1\rho^2/L^2 + K_2(1-\rho^2)$$

with $K_1$ and $K_2$ functions of both the distribution of the auxiliary variable and the method of allocation. In particular, for Neyman allocation, Serfling (1968) derived the above approximation with $K_1 = \frac{H_1^4}{12\sigma_x^2}$ and $K_2 = H_1 H_2$, where $H_1 = \int_a^b \sqrt[3]{f(t)}dt$ and $H_2 = \int_a^b f(t)^{3/2}dt$. For proportional allocation, Thomsen (1976) derived (4.1) with $K_1 = \frac{H^3}{12\sigma_x^2}$ and $K_2 = 1$, with $H = \int_a^b \sqrt[3]{f(t)}dt$. Cochran (1977) originally proposed this use of the auxiliary variable and the simple linear ergression model for the case of the rectangular distribution; in that case, $K_1 = K_2 = 1$.

One of the complaints about the use of this second approximation in Hess, Sethi and Balakrishnan's study was that the linear regression model seemed too simple--they suggested that a linear model with heteroscedastic error variation might offer a better approximation. Following the approach of Serfling and Thomsen, a more sensitive $K_2$ term was developed for (4.1), reflecting the nature of the residual form beyond just homoscedasticity. The resulting term for Neyman (and equal) allocation was $K_2 = H_1 H_r$, where $H_1 = \int_a^b \sqrt{f(t)}dt$ and $H_r = \int_a^b t^r f(t)^{3/2}dt$, when the residual variation has the form $var[e|x] = x^r$.

To employ the approach we proposed in the last section, our interest is in the derivative of (4.1),

$$\frac{dD(L)}{dL} = -\frac{2\rho^2}{L^3} K_1 \quad . \tag{4.2}$$

Note that the derivatives do not depend upon the $K_2$ term. The values of $K_1$ for the different methods of allocation and for the distributions we are considering are listed in Table 4.4. Combining Table 4.4 and (4.2), our proposed criterion generates the approximate efficient stratum numbers in Table 4.5.

Table 4.4. Values of $K_1$.

| Allocation | $\chi_1^2$ | $\chi_2^2$ | Normal |
|---|---|---|---|
| Proportional | 2.1082 | 2.2500 | 2.7207 |
| Neyman/Equal | .9570 | 1.3333 | 2.0944 |

Table 4.5. Efficient number of strata for $D(L)=K_1\rho^2/L^2+K_2(1-\rho^2)$.

| | | Proportional Allocation | | | Neyman/Equal Allocation | | |
|---|---|---|---|---|---|---|---|
| Cutoff | $\rho$ | $\chi_1^2$ | $\chi_2^2$ | Normal | $\chi_1^2$ | $\chi_2^2$ | Normal |
| | .60 | 3 | 3 | 3 | 2 | 3 | 3 |
| 10% | .90 | 4 | 4 | 4 | 3 | 3 | 4 |
| | .99 | 4 | 4 | 4 | 3 | 3 | 4 |
| | .60 | 4 | 4 | 4 | 3 | 3 | 4 |
| 5% | .90 | 5 | 5 | 5 | 4 | 4 | 5 |
| | .99 | 5 | 5 | 5 | 4 | 4 | 5 |
| | .60 | 4 | 5 | 5 | 4 | 4 | 4 |
| 2.5% | .90 | 6 | 6 | 6 | 4 | 5 | 6 |
| | .99 | 6 | 6 | 6 | 5 | 5 | 6 |
| | .60 | 6 | 6 | 6 | 5 | 5 | 6 |
| 1% | .90 | 7 | 8 | 8 | 6 | 7 | 7 |
| | .99 | 8 | 8 | 9 | 6 | 7 | 8 |

To examine these approximations, it remains to use the algorithm developed in the last section along with the approximation $D_2(L)=A+B/L+C/L^2+e$, to determine the efficient number of strata with respect to actual design effects. In particular, we determined the efficient number of strata from the design effects in Chapter II that were derived from Dalenius's original equations, since these represented the closest to the situations we would expect to find in practice. These efficient numbers of strata are in Tables 4.6-4.9, with Tables 4.6, 4.7 and 4.8 being the numbers computed from the design effects under the quadratic model for proportional, X-optimal and equal allocations, respectively, while Table 4.9 is the efficient number of strata from the design effects under the exponential distribution.

We also determined the efficient number of strata from the design effects in Chapter III that were derived from the Dalenius-Gurney equations. Since these design effects reflected the optimum reduction in variance, they represented a more theoretical situation. These efficient numbers of strata are in Tables 4.10, 4.11 and 4.12, determined from the optimum design effects under proportional, Neyman and equal allocation, respectively.

Table 4.6. Efficient number of strata for the design effects using Dalenius's original equations for the quadratic model under proportional allocation.

| | | $\chi^2_1$ | | | $\chi^2_2$ | | | Normal | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cutoff | | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 |
| | .60 | 3 | 4 | 4 | 3 | 4 | 4 | 3 | 4 | 4 |
| 10% | .90 | 3 | 4 | 4 | 3 | 4 | 4 | 3 | 4 | 4 |
| | .99 | 3 | 4 | 4 | 3 | 4 | 4 | 3 | 4 | 4 |
| | .60 | 3 | 6 | 6 | 3 | 6 | 6 | 3 | 5 | 6 |
| 5% | .90 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 |
| | .99 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 4 | 4 |
| | .60 | 4 | 8 | 9 | 4 | 8 | 9 | 4 | 7 | 8 |
| 2.5% | .90 | 5 | 7 | 7 | 5 | 7 | 7 | 5 | 6 | 6 |
| | .99 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| | .60 | 5 | 12 | 14 | 5 | 12 | 14 | 5 | 11 | 12 |
| 1% | .90 | 7 | 10 | 11 | 7 | 10 | 11 | 7 | 9 | 9 |
| | .99 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

Table 4.7. Efficient number of strata for the design effects using Dalenius's original equations for the quadratic model under X-optimal allocation.

| | | $\chi^2_1$ | | | | | | | $\chi^2_2$ | | | | | | | Normal | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | |
| | | r | | | r | | r | | r | | | r | | r | | r | | | r | | r | |
| Cutoff | ρ | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 |
| | .60 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | 4 | 4 | 2 | 2 | 2 | 4 | 4 | 4 | 4 |
| 10% | .90 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 |
| | .99 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | 2 | 3 | 3 | 4 | 4 | 4 | 4 | 2 | 3 | 3 | 4 | 4 | 5 | 5 | 3 | 3 | 3 | 5 | 5 | 5 | 5 |
| 5% | .90 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 |
| | .99 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | 2 | 3 | 4 | 4 | 5 | 5 | 5 | 3 | 4 | 4 | 5 | 6 | 6 | 6 | 3 | 3 | 3 | 6 | 6 | 7 | 7 |
| 2.5% | .90 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 7 |
| | .99 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| | .60 | 2 | 3 | 5 | 6 | 7 | 7 | 7 | 4 | 5 | 5 | 8 | 8 | 9 | 9 | 4 | 4 | 4 | 9 | 9 | 10 | 10 |
| 1% | .90 | 5 | 5 | 6 | 6 | 6 | 6 | 7 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 9 | 9 |
| | .99 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

Table 4.8. Efficient number of strata for the design effects using Dalenius's original equations for the quadratic model under equal allocation.

| | | $\chi^2_1$ | | | | | | | $\chi^2_2$ | | | | | | | Normal | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | |
| | | r | | | r | | r | | r | | | r | | r | | r | | | r | | r | |
| Cutoff | ρ | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 |
| | .60 | -* | 2 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | 4 | 4 | 2 | 2 | 2 | 4 | 4 | 4 | 4 |
| 10% | .90 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 |
| | .99 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | - | 3 | 3 | 4 | 4 | 4 | 4 | 2 | 3 | 3 | 4 | 4 | 5 | 5 | 3 | 3 | 3 | 5 | 5 | 5 | 5 |
| 5% | .90 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 |
| | .99 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | - | 3 | 4 | 5 | 5 | 5 | 5 | 3 | 3 | 4 | 5 | 6 | 6 | 6 | 3 | 3 | 3 | 6 | 6 | 7 | 7 |
| 2.5% | .90 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 |
| | .99 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| | .60 | - | 3 | 5 | 6 | 7 | 7 | 8 | 3 | 4 | 5 | 8 | 8 | 9 | 9 | 4 | 4 | 4 | 9 | 9 | 10 | 10 |
| 1% | .90 | 5 | 5 | 6 | 6 | 6 | 6 | 7 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 9 | 9 |
| | .99 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

*The design effects were an increasing function of L for this model.

Table 4.9. Efficient number of strata for the design effects using Dalenius's original equations for the exponential model.

| | | Proportional | | | | | | Neyman | | | | | | Equal | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\chi^2_1$ | | $\chi^2_2$ | | Normal | | $\chi^2_1$ | | $\chi^2_2$ | | Normal | | $\chi^2_1$ | | $\chi^2_2$ | | Normal | |
| | | C | | C | | C | | C | | C | | C | | C | | C | | C | |
| Cutoff | ρ | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| | .60 | 2 | 2 | 2 | 2 | 3 | 4 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 3 | 3 | 4 | 4 |
| 10% | .90 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 3 | 3 | 4 | 4 |
| | .99 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 3 | 3 | 4 | 4 |
| | .60 | 4 | 4 | 4 | 5 | 5 | 6 | 4 | 4 | 4 | 4 | 5 | 6 | 4 | 4 | 4 | 4 | 5 | 6 |
| 5% | .90 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 5 | 5 | 4 | 4 | 4 | 4 | 5 | 5 |
| | .99 | 5 | 5 | 5 | 5 | 4 | 5 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | 6 | 6 | 7 | 7 | 8 | 9 | 5 | 5 | 5 | 6 | 7 | 8 | 5 | 5 | 5 | 6 | 7 | 9 |
| 2.5% | .90 | 7 | 7 | 7 | 7 | 6 | 7 | 5 | 5 | 5 | 5 | 6 | 6 | 5 | 5 | 5 | 5 | 6 | 6 |
| | .99 | 6 | 6 | 6 | 6 | 6 | 6 | 5 | 5 | 5 | 5 | 6 | 6 | 5 | 5 | 5 | 5 | 6 | 6 |
| | .60 | 10 | 11 | 11 | 11 | 13 | 14 | 7 | 7 | 8 | 9 | 12 | 13 | 7 | 7 | 8 | 9 | 12 | 13 |
| 1% | .90 | 10 | 11 | 10 | 11 | 10 | 10 | 7 | 7 | 7 | 8 | 9 | 9 | 7 | 7 | 7 | 8 | 9 | 9 |
| | .99 | 8 | 9 | 8 | 9 | 8 | 8 | 6 | 6 | 7 | 7 | 8 | 8 | 6 | 6 | 7 | 7 | 8 | 8 |

Table 4.10. Efficient number of strata for the design effects using the Dalenius-Gurney equations under proportional allocation.

| | | $\chi_1^2$ | | | $\chi_2^2$ | | | Normal | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cutoff | ρ | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 | C=0 | C=1 | C=2 |
| | .60 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 10% | .90 | 3 | 4 | 4 | 3 | 4 | 4 | 3 | 3 | 3 |
| | .99 | 3 | 4 | 4 | 3 | 4 | 4 | 3 | 4 | 4 |
| | .60 | 3 | 4 | 4 | 3 | 4 | 4 | 3 | 4 | 4 |
| 5% | .90 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 4 | 4 |
| | .99 | 4 | 4 | 5 | 4 | 5 | 5 | 4 | 4 | 4 |
| | .60 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 |
| 2.5% | .90 | 5 | 6 | 6 | 5 | 6 | 6 | 5 | 5 | 5 |
| | .99 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| | .60 | 5 | 7 | 8 | 5 | 7 | 8 | 5 | 7 | 8 |
| 1% | .90 | 7 | 9 | 9 | 7 | 8 | 9 | 7 | 7 | 8 |
| | .99 | 8 | 8 | 8 | 8 | 8 | 9 | 8 | 8 | 8 |

Table 4.11. Efficient number of strata for the design effects using the Dalenius-Gurney equations under Neyman allocation.

| | | $\chi^2_1$ | | | | | | | $\chi^2_2$ | | | | | | | Normal | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | |
| | | r | | | r | | r | | r | | | r | | r | | r | | | r | | r | |
| Cutoff | ρ | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 |
| | .60 | 2 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 10% | .90 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| | .99 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 |
| 5% | .90 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .99 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 |
| 2.5% | .90 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| | .99 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| | .60 | 5 | 5 | 5 | 5 | 6 | 5 | 6 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 8 | 8 |
| 1% | .90 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| | .99 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

Table 4.12. Efficient number of strata for the design effects using the Dalenius-Gurney equations under equal allocation.

| | | $\chi_1^2$ | | | | | | | $\chi_2^2$ | | | | | | | Normal | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | | C=0 | | | C=1 | | C=2 | |
| | | r | | | r | | r | | r | | | r | | r | | r | | | r | | r | |
| Cutoff | ρ | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 2 |
| | .60 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 2 |
| 10% | .90 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| | .99 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 |
| 5% | .90 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .99 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| | .60 | 4 | 4 | 4 | 5 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 5 | 5 | 6 | 6 |
| 2.5% | .90 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| | .99 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| | .60 | 6 | 5 | 5 | 7 | 6 | 6 | 6 | 6 | 5 | 6 | 7 | 7 | 7 | 7 | 5 | 5 | 6 | 8 | 8 | 10 | 10 |
| 1% | .90 | 7 | 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| | .99 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

## 4.3. Empirical Comparisons

The largest number of strata in practice beyond which further stratification would not produce any significant reduction of the variance of $\hat{\mu}$, were listed by Cochran and Kish as 6 and 10, respectively. Examining the tables containing the efficient number of strata derived from the theoretical design effects of Chapter III (Tables 4.10-4.12), we note that if we employ a cutoff percentage of no less than 2.5 percent in our proposed algorithm, then the largest number of strata our algorithm would require is 6; if we include the more sensitive 1 percent cutoff point, the largest number is 10. Thus, using these theoretical design effects, Cochran's and Kish's suggested limits would result if we used 2.5 percent and 1 percent, respectively, as our lowest cutoff points.

However, turning to Tables 4.6-4.9, which list the efficient number of strata derived from the practical situation examined in Chapter II, the most number of strata required to attain at least a 2.5 percent and a 1 percent reduction for the quadratic model under proportional allocation, are 9 and 14, respectively, while the largest numbers of strata are 7 and 10, respectively, for Neyman and equal allocation. For the exponential model, the most number of strata required are 9 for a 2.5 percent cutoff point, and 14 for a 1 percent cutoff.

The reason for the discrepancy between the largest numbers in these tables can be explained by the fact that the stratum boundaries in Chapter II were determined independently of the functional relationship between X and Y, and hence for the more moderate values of $\rho$, they did not generate the best stratification of the study variable possible. It thus took more strata to eliminate the same amount of variation in the population than would have been required by fewer strata when they were more carefully constructed. As $\rho$ increases to the higher values, the approximation employed in Chapter II more closely approximates the actual functional relationship between X and Y, resulting in better stratum construction, and thus closer agreement to the optimum number of strata in Tables 4.10-4.12. The largest numbers occurred in Tables 4.6-4.9 when $\rho$=.60.

Thus, using the auxiliary variable to decide upon the number of strata to use could lead one to underestimate the number of strata needed to attain a specified cutoff. As an example of this, consider the situation under Neyman allocation when $\rho$=.60 and C=2 for the $\chi_2^2$ distribution. The number of strata Table 4.11 indicates is necessary for a 2.5 percent cutoff point is 4; in Table 4.7, the use of 4 strata for this model for the approximate situation examined in Chapter II corresponds to the efficient number of strata when a 10 percent cutoff point is used. Thus, the theoretical case corresponds to a reduction criterion that is 4 times coarser in the practical situation.

Let us now consider how well the approximations we've discussed in the second section fare against the values from Chapters II and III. Let us initially note a pattern that exists in Tables 4.6, 4.7, 4.8 and 4.9. Unlike the other tables, for C$\neq$0, as the correlation between X and Y increases, all other parameters considered fixed, the efficient number of strata in these tables decreases. This trend is most pronounced for the quadratic model under proportional allocation, and when the 1 percent cutoff point is used under X-optimal and equal allocations; it is most pronounced for the exponential model for the lower cutoff percentage points and for the symmetric normal distribution. This pattern can again be explained by the non-optimal computation of the stratum boundaries requiring more strata to achieve the same degree of reduction attained by fewer optimally constructed strata.

We will first examine how the approximations fare against the results from Chapter III, representing the theoretical design effects. Comparing the approximation using the hypothetical distribution in place of the study variable's distribution (Table 4.3), we note that for proportional allocation, except for the lowest correlation in the linear model (when only 36 percent of the variation of Y is explained by the functional relationship with the auxiliary variable), the approximation works very well. For Neyman and equal allocations, the approximation also does well, though suffering a little in the lowest correlation at the 1 percent cutoff point for equal allocation. When we compare the second, more sensitive, approximation that was generated from (4.1) (Table 4.5), we note that except for the lowest correlation at 1 percent cutoff, the fits for all three allocations are also quite good.

Let us now consider the estimation of the efficient number of strata derived from the practical design effects of Chapter II. For proportional allocation, because of the pattern of the stratum numbers for $C \neq 0$ we discussed above, neither approximation offers a good fit for $\rho < .99$, though both are excellent for $\rho = .99$. On the other hand, for the linear model, (4.1) offers a good fit over the correlation values covered, while the less sensitive estimate in Table 4.3 did poorly for the smaller correlation values.

For X-optimal and equal allocations, we have the interesting result that, except for the smallest correlation under the C=r=0 model under the $\chi_1^2$ distribution, the more sensitive estimate (4.1) fits well for the linear model, while the less sensitive estimate in Table 4.3 fits fairly well over all but the 1 percent cutoff point for the non-linear model.

Finally, for the exponential model in Table 4.9, except for the 1 percent cutoff point, the use of the hypothetical distribution fits reasonably well, while the pattern we discussed above makes (4.1) fit this model poorly.

## 4.4. Summary

To enable us to quantify how well the approximations suggested in the literature fared in the theoretical and practical cases we considered in Chapters III and II, respectively, we introduced an algorithm which determined the efficient number of strata to use to insure strata were added until a specified reduction could no longer be attained. Using this criterion, it was found that both proposed approximations performed reasonably well in estimating the efficient number of strata from those design effects derived from the optimum variances of $\hat{\Omega}$ in Chapter III.

For the more interesting problem of estimating the efficient number of strata from the design effects for the practical situation examined in Chapter II, it was found that (4.1) offered an approximation that fit the linear model quite well for all but the lowest correlation considered (when only 36 percent of the variation of Y is explained by the functional relationship with the auxiliary variable) under all three methods of allocation. Use of the hypothetical distribution to replace the study variable's distribution offered a reasonable fit for the non-linear quadratic model for both X-optimal and equal allocations for all but the most sensitive cutoff criterion; this approximation also offered the best approximation to the exponential model. On the other hand, neither approach fared well for the non-linear quadratic model under proportional allocation.

# V. CONCLUSIONS

When one is interested in estimating the population mean of the study variable, stratified random sampling can provide an estimate whose variance is less than that of the variance of the sample mean under simple random sampling. When the stratum boundaries are not suggested a priori, the use of optimum stratification allows us to choose those boundaries that provide the best reduction in the variance of this estimate.

In this thesis, we considered two questions dealing with the choice of the stratum boundaries. In Chapter II, when nothing is known about the functional relationship between the study and auxiliary variables, we investigated whether the approximate boundaries from Dalenius's original equations on the auxiliary variable provided a reduction in the variance compared with the use of simple random sampling. This was found to be true when the functional relationship between the two variables explained more than 40 percent of the variation of the study variable.

In Chapter III, when information was available about the functional relationship between the study and auxiliary variables, we investigated how good an approximation to the Dalenius-Gurney equations was available from a series of proposed approximations found in the literature. None of the approximations investigated performed superbly in all situations, though reasonable approximations were found for almost every case except when equal allocation is employed. Since each situation encountered will offer its own idiosyncrasies, the tables of comparisons in Chapter III can be used to decide which rule offers the best approximation in each case. When the performance of an approximation was judged by its largest table value over a specified range of parameters, it was found that Singh's cum H(x) rules for both proportional and Neyman allocations offered the best overall approximation.

In Chapter II, we also considered how the choice of the method of allocation is affected when the stratification is performed on the auxiliary variable independently of the study variable. It was discovered that for the symmetric normal distribution, proportional allocation offers design effects which are comparable to those of X-optimal allocation, suggesting the easier boundary equations under proportional allocation can be used instead of the more cumbersome boundary equations for X-optimal allocation without a significant loss of variance reduction. For the skewed distributions, except when the relationship between X and Y was meager, X-optimal allocation generally provided significantly smaller design effects than did proportional allocation.

Finally, once the method of allocation and the mechanism for determining the stratum boundaries is chosen, it remains to decide what number of strata to use. This decision is generally made by examining the design effects as the number of strata increases. An algorithm was introduced to allow us to quantify the effect the increase in the number of strata has on the reduction in variance. Employing this algorithm, we were able to investigate the approximations to the design effects proposed in the literature. It was discovered that while the approximations worked reasonably well when compared with the design effects resulting from the optimum stratification boundaries of Chapter III, the performance of the approximations was not as good when compared with the design effects resulting from the approximate stratum boundaries of Chapter II. In particular, it was discovered that the approximation modeling the design effect as a function of L, $\rho$ and the distribution of X ((4.1)), performed reasonably well only when the functional relationship between the two variables was linear. On the other hand, the approximation employing the auxiliary variable's distribution as a hypothetical distribution for the study variable, performed reasonably well only when the functional relationship between the two variables was non-linear.

As a final consideration, let us briefly address the question of how good an approximation Dalenius's original equations with the auxiliary variable is to the Dalenius-Gurney equations. This question was never specifically considered in the thesis, but we can still offer some insight for the cases of proportional and X-optimal allocations. When proportional allocation is employed, the linearization approach considered in Chapter III generates Dalenius's original equations on the auxiliary variable. However, when X-optimal allocation is employed, not only do none of the proposed approximations in Chapter III have the form of Dalenius's original equation for X-optimal allocation, but the two variance forms, $var[\hat{\mu}|X\text{-opt}]$ and $var[\hat{\mu}|Ney]$, are different. But, results in the literature have indicated that Dalenius and Hodges's $cum\sqrt{f(x)}$ rule offers a very reasonable approximation to Dalenius's original equations (see, for example, Kpedekpo (1973)). Thus, if we assume that the two variance forms, $var[\hat{\mu}|X\text{-opt}]$ and $var[\hat{\mu}|Ney]$, have roughly the same sensitivity to approximate boundaries (that is, approximate boundaries the same distance from the optimum boundaries will result in about the same difference in design effects for both variance forms), we can consider the performance of the cum$\sqrt{f(x)}$ rule under Neyman allocation as an approximation to the performance of Dalenius's original equations for X-optimal allocation.

If we decide to judge the performance of the proposed approximations by the largest table value over a specified range of parameter values, we note that for both methods of allocation, Dalenius's original equations offers an excellent approximation when the functional relationship between the variables is linear, but a poor approximation when the functional relationship is non-linear.

BIBLIOGRAPHY

Anderson, Dallas, W., Kish, Leslie and Cornell, Richard G. (1975), "Quantifying Gains from Stratification for Optimum and Approximately Optimum Strata Using a Bivariate Normal Model," Biostatistics Technical Report Series No. 4, Department of Biostatistics, University of Michigan.

Calvin, Lyle D. (1978), Personal communication.

Cochran, William G. (1977), Sampling Techniques, 3rd edition, New York: John Wiley & Sons.

Cochran, William G. (1961), "Comparison of Methods for Determining Stratum Boundaries," Bulletin of the International Statistical Institute, 38, 345-358.

Dalenius, Tore (1950), "The Problem of Optimum Stratification," Skandinavisk Aktuarietidskrift, 33, 203-213.

Dalenius, Tore (1953), "The Economics of One-Stage Stratified Sampling," Sankhyā, 12, 351-356.

Dalenius, Tore and Gurney, Margaret (1951), "The Problem of Optimum Stratification. II.," Skandinavisk Aktuarietidskrift, 34, 133-148.

Dalenius, Tore and Hodges, Joseph L., Jr. (1957), "The Choice of Stratification Points," Skandinavisk Aktuarietidskrift, 40, 198-203.

Dalenius, Tore and Hodges, Joseph L., Jr. (1959), "Minimum Variance Stratification," Journal of the American Statistical Association, 54, 88-101.

Hess, Irene, Sethi, V. K. and Balakrishnan, T. R. (1966), "Stratification: A Practical Investigation," Journal of the American Statistical Association, 61, 74-90.

Kish, Leslie (1965), Survey Sampling, New York: John Wiley & Sons.

Kpedekpo, G. M. K. (1973), "Recent Advances on Some Aspects of Stratified Sample Design," Metrika, 20, 54-64.

Murthy, N. N. (1967), Sampling Theory and Methods, Calcutta: Statistical Publishing Society.

Serfling, R. J. (1968), "Approximately Optimal Stratification," Journal of the American Statistical Association, 63, 1298-1309.

Sethi, V. K. (1963), "A Note on Optimum Stratification of Populations for Estimating the Population Means," Australian Journal of Statistics, 5, 20-33.

Singh, Ravindra (1975), "An Alternative Method of Stratification on the Auxiliary Variable," Sankhyā, Series C, 37, 100-108.

Singh, Ravindra and Sukhatme, B. V. (1969), "Optimum Stratification," Annals of the Institute of Statistical Mathematics, 21, 515-528.

Singh, Ravindra and Parkash, Dev (1975), "Optimum Stratification for Equal Allocation," Annals of the Institute of Statistical Mathematics, 27, 273-280.

Thomsen, Ib (1976), "A Comparison of Approximately Optimal Stratification given Proportional Allocation with Other Methods of Stratification and Allocation," Metrika, 23, 15-25.

# APPENDICES

# APPENDIX A

The three models we are considering are

$$\text{linear:}\quad g(x) = \alpha + \beta X$$

$$\text{quadratic:}\quad g(x) = \alpha + \beta X + \gamma X^2, \quad \gamma\geq 0$$

$$\text{and exponential:}\quad g(x) = \alpha + \beta \exp(\gamma x), \ \gamma\geq 0$$

To specify the values of the two relevant parameters, $\beta$ and $\alpha$ (because our interest is in the variance, the location parameter, $\alpha$, does not enter into the problem), we introduce a measure of the "linear" fit of the regression line $y=g(x)$, and a measure of the "non-linear" fit of the line.

Suppose the true regression model between X and Y is given in (1.9). Then, let us consider approximating it by the linear model

$$y = A + Bx + z$$

where $cov[X,Z]=0$, $E[Z]=0$ and $var[Z]=\sigma_z^2$. To estimate the parameters A and B, let us minimize the expected squared error, i.e. $\hat{A}$ and $\hat{B}$ are defined by

$$\min_{A,B} E[Z^2] = \min_{A,B} E[(Y-A-BX)^2] = E[(Y-\hat{A}-\hat{B}X)^2] \quad .$$

Using the model (1.9), we can write

$$\text{(A.1)}\quad \begin{aligned} &\text{total expected squared error} = \text{EST} = E[(Y-E[g(X)])^2] \\ &\text{linear expected squared error} = \text{ESL} = E[(\hat{A}+\hat{B}X-E[g(X)])^2] \\ &\text{non-linear expected squared error} = \text{ESNL} = E[(\hat{A}+\hat{B}X-g(X))^2] \\ &\text{pure expected squared error} = \text{ESPE} = E[(Y-g(X))^2]=\sigma_e^2 \end{aligned}$$

with EST = ESL + ESNL + ESPE.

The two measures we've chosen to specify the relevant parameters of $g(x)$ are the correlation between X and Y, and the ratio of the non-linear expected squared error (the "lack-of-linear-fit") to the pure expected squared error. Using the above notation, we may define the correlation, $0 \leq \rho \leq 1$, by

$$\text{(A.2)}\qquad \rho^2 = \frac{\text{ESL}}{\text{EST}} ,$$

and define the measure of the non-linear fit, $C \geq 0$, by

$$\text{(A.3)}\qquad C = \frac{\text{ESNL}}{\text{ESPE}} = \frac{\text{ESNL}}{\sigma_e^2} .$$

Thus, we can write (A.1) as

$$\text{ESL} = \frac{\rho^2(C+1)}{1-\rho^2}\,\sigma_e^2$$

$$\text{ESNL} = C\sigma_e^2$$

$$\text{EST} = \frac{C+1}{1-\rho^2}\,\sigma_e^2$$

However, while $\rho$ and C are well defined, they are not necessarily comparable; the correlation, $\rho$, is scale free, while C is measured in units of residual variance. Consequently, let us replace C with another measure of non-linear fit that is comparable with $\rho$. In particular, let us introduce the parameter $\pi$, $\rho \leq \pi \leq 1$, defined by

$$\text{(A.4)} \qquad \pi^2 = \frac{\text{ESL} + \text{ESNL}}{\text{EST}} = \frac{C+\rho^2}{C+1} \quad .$$

Note that this parameter is a scale free analog of $\rho$, except that it measures the amount of "full" fit in the regression line $y=g(x)$. We, thus, can write (A.1) as

$$\text{ESL} = \frac{\rho^2}{1-\pi^2}\,\sigma_e^2$$

$$\text{ESNL} = \frac{\pi^2-\rho^2}{1-\pi^2}\,\sigma_e^2$$

$$\text{EST} = \frac{1}{1-\pi^2}\,\sigma_e^2 \quad .$$

Perhaps the best feeling for the parameter $\pi$ can be gained by noting that when $g(x)$ is linear, $\rho$ relates $\sigma_e^2$ and $\sigma_y^2$ by $\sigma_e^2=(1-\rho^2)\sigma_y^2$: when $g(x)$ is non-linear, $\sigma_e^2$ and $\sigma_y^2$ are related by $\sigma_e^2=(1-\pi^2)\sigma_y^2$.

Note that as $C\to 0$, or analogously, $\pi^2\to\rho^2$, the non-linear portion of the model disappears. We shall use both sets of measures, $(\rho,C)$ and $(\rho,\pi)$ in this thesis, as each offers its own particular advantages in the consideration of the forms of the stratum variances of Y; the pair $(\rho,C)$ is superior in the empirical studies in parametrizing the range of our functional relationships.

So, the specification of either pair of these measures determines the values of the parameters $\beta$ and $\gamma$. In particular, for the quadratic model, the parameters β and γ are defined by

$$\beta = \frac{\sigma_e}{\sigma_x}\left(\frac{\rho}{\sqrt{1-\rho^2}}\sqrt{C+1} - \frac{C_x}{\sigma_x}\sqrt{\frac{C}{D_x}}\right)$$

$$\gamma = \sigma_e\sqrt{\frac{C}{D_x}}$$

where $\sigma_x^2$=var[X], $\sigma_e^2$=var[e]=E[$\phi$(X)], and $C_x$=cov[X,$X^2$] and $D_x$=var[$X^2$]-$C_x^2/\sigma_x^2$. Note that the linear model is a special case of the quadratic model when C=0, in which case $\beta=\frac{\rho}{\sqrt{1-\rho^2}}\frac{\sigma_e}{\sigma_x}$ and γ=0.

For the exponential model, with U=exp(γX), we can write

$$\beta = \frac{\sigma_e}{\sigma_u}\sqrt{\frac{C+\rho^2}{1-\rho^2}}$$

where $\sigma_u^2$=var[U]=E[exp(2γX)]-E[exp(γX)]$^2$, and thus depends upon the parameter γ, which must be determined iteratively. Let ω be defined as the correlation between X and U. Then, if $M_x(t)$ is the moment generating function of the auxiliary variable X, we may write

$$\omega^2 = \rho^2\frac{C+1}{C+\rho^2} = \frac{(M_x'(\gamma)-E[X]M_x(\gamma))^2}{\sigma_x^2(M_x(2\gamma)-M_x(\gamma)^2)} \quad .$$

So, knowing (ρ,C), we can use the above equation to determine the value of γ.

APPENDIX B

Our interest is in deriving the proportional allocation analog to Singh's cum H(x) rule for Neyman allocation. We will use both Singh (1975) as well as Singh and Sukhatme (1969).

Under proportional allocation, the variance of $\hat{\mu}$ has the form

$$\mathrm{var}[\hat{\mu}|\mathrm{prop}] = \frac{1}{n}\sum_{h=1}^{L} W_h(\sigma_{hg}^2+\sigma_{he}^2)$$

where $\sigma_{he}^2=E[e^2|h]=E[\phi(X)|h]=\mu_{h\phi}$. Singh chose to approximate $\mu_{h\phi}$ by

$$\mu_{h\phi} = \frac{12\sigma_{h\Psi}^2}{K_h^2}(1+O(K_h^2))$$

where $K_h=x_h-x_{h-1}$, and $\Psi(x)$ is such that $\frac{d\Psi(x)}{dx} = \phi(x)$. So

$$\sigma_{hy}^2 = \sigma_{hg}^2 + \frac{12\sigma_{h\Psi}^2}{K_h^2}(1+O(K_h^2)) \quad .$$

Singh argues that when the number of strata is large, the $O(K_h^2)$ terms can be neglected in comparison to 1, so $\sigma_{hy}^2 \doteq \sigma_{hg}^2+\Theta\sigma_{h\Psi}^2$, where

$\Theta=\frac{12L^2}{(b-a)^2}$ . The form of the variance of $\hat{\mu}$ thus becomes

$$\mathrm{var}[\hat{\mu}|\mathrm{prop}] = \frac{1}{n}\sum_{h=1}^{L} W_h(\sigma_{hg}^2+\Theta\sigma_{h\Psi}^2) \quad .$$

Since this variance is a function of the stratum boundaries, equations analogous to the Dalenius-Gurney equations can be generated by

$$\frac{\partial \text{var}[\hat{\mu}|\text{prop}]}{\partial x_h} = 0, \qquad h=1,2,\ldots,L-1,$$

which can be shown to reduce to

$$(B.1)\qquad \begin{aligned}&(\mu_{hg}-g(x_h))^2 + \Theta(\mu_{h\Psi}-\Psi(x_h))^2 = \\ &(\mu_{kg}-g(x_h))^2 + \Theta(\mu x_\Psi-\Psi(x_h))^2\end{aligned} \qquad \begin{aligned}&k=h+1,\\ &h=1,2,\ldots,L-1\end{aligned}$$

The stratum moments, $\mu_{hg} = \dfrac{\int_{x_{h-1}}^{x_h} g(t)f_X(t)dt}{\int_{x_{h-1}}^{x_h} f_X(t)dt}$, can be expanded about $x_h$ (see Singh and Sukhatme):

$$\mu_{hg} = g(1 - \frac{g'}{2g}K_h + \frac{g'f'+2fg''}{12fg}K_h^2 - \frac{ff''g'+ff'g''+f^2g'''-f'^2g'}{24f^2g}K_h^3$$

$$+ O(K_h^4))$$

where the functions g, f and their derivatives are evaluated at $x=x_h$. Using this, we can write

$$(\mu_{hg}-g(x_h))^2 = \frac{K_h^2}{4f}(fg'^2 - \frac{g'^2f'+2fg'g''}{3}K_h + O(K_h^2)) \quad .$$

Analogously, we can write

$$(\mu_{h\Psi}-\Psi(x_h))^2 = \frac{K_h^2}{4f}(f\Psi'^2 - \frac{\Psi'^2f'+2f\Psi'\Psi''}{3}K_h + O(K_h^2))$$

where $\Psi$ and its derivatives are also evaluated at $x=x_h$. Since $\Psi'=\sqrt{\phi}$

by assumption, we can write the left-hand side of (B.1) as

$$\text{LHS} = \frac{K_h^2}{4f}\left(f(g'^2+\Theta\phi) - \frac{2fg'g''+f'g'^2+\Theta(f'\phi+f\phi')}{3}K_h + O(K_h^2)\right)$$

or

$$\text{(B.2)} \qquad = \frac{K_h^2}{4f}\left(S - \frac{S'}{3}K_h + O(K_h^2)\right)$$

with $S=f(g'^2+\Theta\phi)$.

Because of the nature of our resulting approximate equations, we depart from Singh's derivation and instead follow Singh and Sukhatme's derivation of their cum H(x) rules, In particular, raising (B.2) to the 3/2 power, we can write

$$\text{LHS}^{3/2} = \frac{\sqrt{S'}}{8f^{3/2}}\left(K_h^3 S\left(1 - \frac{S'}{2S}K_h + O(K_h^2)\right)\right) \quad .$$

Now, expanding $\sqrt[3]{S(t)}$ in a Taylor series about $x=x_h$, we can write

$$\sqrt[3]{S(t)} = \sqrt[3]{S(x_h)} + \frac{1}{3}(t-x_h)S'(x_h)S(x_h)^{-\frac{2}{3}} + O(t-x_h)^2$$

$$\Longrightarrow \quad \int_{x_{h-1}}^{x_h}\sqrt[3]{S(t)}\,dt = K_h\sqrt[3]{S(x_h)} - \frac{S'(x_h)}{6S(x_h)^{2/3}}K_h^2 + O(K_h^3)$$

$$= K_h\sqrt[3]{S}\left(1 - \frac{S'}{6S}K_h + O(K_h^2)\right) \quad ,$$

where $S$ and $S'$ are evaluated at $x=x_h$. Thus

$$\left(\int_{x_{h-1}}^{x_h}\sqrt[3]{S(t)}\,dt\right)^3 = K_h^3 S\left(1 - \frac{S'}{2S}K_h + O(K_h^2)\right)$$

and we may write

$$LHS^{3/2} = \frac{\sqrt{S}}{8f^{3/2}}\left(\left(\int_{x_{h-1}}^{x_h}\sqrt[3]{S(t)}dt\right)^3(1+O(K_h^2))\right) \quad ,$$

where S and f are evaluated at $x=x_h$.

By identical travels, the right-hand side of (B.1) can be approximated, raised to the 3/2 power, resulting in

$$RHS^{3/2} = \frac{\sqrt{S}}{8f^{3/2}}\left(\left(\int_{x_h}^{x_{h+1}}\sqrt[3]{S(t)}dt\right)(1+O(K_i^2))\right)$$

where $K_i=x_{h+1}-x_h$, and where S and f are also evaluated at $x=x_h$. Thus, we can approximate (B.1) with

$$\text{(B.3)} \qquad \int_{x_{h-1}}^{x_h}\sqrt[3]{S(t)}dt(1+O(K_h^2)) = \int_{x_h}^{x_{h+1}}\sqrt[3]{S(t)}dt(1+O(K_i^2))$$

with $S=f(g'^2+\Theta\phi)$, $\Theta=\frac{12L}{(b-a)^2}$ . And, if we have a large number of strata so that the $K_h$ are small and their higher powers in the expansion can be neglected, (B.1) can be approximated by the system

$$\int_{x_{h-1}}^{x_h}\sqrt[3]{S(t)}dt = \text{constant} = \frac{\int_a^b\sqrt[3]{S(t)}dt}{L} \quad , h=1,2,\ldots,L-1$$

where terms of order $O(m^2)$, $m=\sup_h K_h$, have been neglected on both sides of (B.3) since $\int_{x_{h-1}}^{x_h}\sqrt[3]{S(t)}dt = O(m)$, as long as it is assumed that

$\sqrt[3]{f_X(x)(g'(x)^2+\Theta\phi(x))} > 0$, $a\leq x\leq b$.

# APPENDIX C

In determining the boundaries which offer optimum (or approximately optimum) stratification, it is necessary to solve some system of non-linear equations. In this appendix, we discuss briefly the techniques that were used in this thesis.

## Dalenius's Original Equations

The solutions to equations (1.2)-(1.4) were determined by a technique suggested by Sethi (1963). First, rearrange the equations so $x_h$ is expressed as a function of the moments of the adjacent strata. Then, e.g., under proportional allocation, the following iterative method is used:

(i) start with an arbitrary set of boundary points

$$a < x_1' < x_2' < \dots < x_{L-1}' < b$$

(ii) compute the stratum moments based on these boundaries

(iii) calculate

$$x_h'' = \frac{\mu_{hx} + \mu_{kx}}{2} \qquad k=h+1,\ h=1,2,\dots,L-1$$

(iv) repeat (ii) and (iii) until two consecutive sets of boundaries are within a specified distance of one another.

For Neyman allocation, step (iii) would be replaced by:

(iii) calculate

$$x_h'' = \frac{B+\sqrt{B^2-AC}}{A}$$

where

$$A = \frac{1}{\sigma_{hx}} - \frac{1}{\sigma_{kx}}$$

$$B = \frac{\mu_{hx}}{\sigma_{hx}} - \frac{\mu_{kx}}{\sigma_{kx}}$$

$$\text{and } C = \frac{\sigma_{hx}^2 + \mu_{hx}^2}{\sigma_{hx}} - \frac{\sigma_{kx}^2 + \mu_{kx}^2}{\sigma_{kx}} \qquad k=h+1,\ h=1,2,\ldots,L-1$$

and where the "+" sign before the radical results from the assumption $\mu_{hx} \leq x_h'' \leq \mu_{kx}$, k=h+1, h=1,2,...,L-1.

or equal allocation, step (iii) would be replaced by

(iii) calculate

$$x_h'' = \frac{T_h - T_k + \sqrt{(T_h - T_k)^2 - (S_h - S_k)(W_h - W_k)}}{W_h - W_k} \qquad k=h+1,\ h=1,2,\ldots,L-1$$

where

$$T_u = \int_{x_{u-1}'}^{x_u'} t f_X(t)dt, \quad S_u = \int_{x_{u-1}'}^{x_u'} t^2 f_X(t)dt, \quad u=h,h+1,$$

and where the "+" sign results from the assumption

$$\mu_{hx} \leq x_h'' \leq \mu_{kx}, \qquad k=h+1,\ h=1,2,\ldots,L-1.$$

## Cum H(x) Rules

The approximations to the Dalenius-Gurney equations we have considered in this study have the form of (1.8). If we assume $H(x)>0$, $a \leq x \leq b$, the $F(x)=\int_a^x H(t)dt$ is a non-decreasing function of x.

If $C=\frac{\int_a^b H(t)dt}{L}$, the equations have the form

$$F(x_k)-F(x_h)-C = 0, \quad k=h+1,\ h=1,2,\ldots,L-1.$$

We employed a first-derivative Newton procedure to consecutively compute the solutions to these equations:

(i) calculate $F(x_o^*)$, $x_o^*=a$, and set h=1

(ii) specify an arbitrary $x_h'>x_{h-1}^*$

(iii) calculate $F(x_h')$, $H(x_h')$

(iv) calculate

$$x_h'' = x_h' - \frac{F(x_h')-F(x_{h-1}^*)-C}{H(x_h')}$$

(v) repeat (iii) and (iv) until two consecutive points are within a specified distance of one another.

This computes $x_1^*$. Repeat (ii)-(v) with h=2 to compute $x_2^*$, and then h=3 to compute $x_3^*$, et cetera, until all L-1 boundaries have been computed.

The Dalenius-Gurney Equations

To determine how well the approximate boundaries computed by the cum H(x) rules fare against the best boundaries, it is necessary to solve the Dalenius-Gurney equations, (1.5)-(1.7). Because of the complexity of these equations, we were unable to use the same procedure we used to compute the solutions to Dalenius's original equations. Consequently, it was decided to employ a form of second-derivative Newton procedure:

Let $F(\underset{\sim}{x})=\text{var}[\hat{\mu}|\cdot]$, with allocation and number of strata fixed, where $\underset{\sim}{x}$ is the L-1 vector of stratum boundaries. Then, let the gradient vector and symmetric Hessian matrix of F evaluated at $\underset{\sim}{x}$ be denoted by $\nabla F(\underset{\sim}{x})$ and $H(\underset{\sim}{x})$, respectively. Then the iterative algorithm is:

(i) specify a starting point, $\underset{\sim}{x}_1$

(ii) compute $\nabla F(\underset{\sim}{x}_1)$ and $H(\underset{\sim}{x}_1)$

(iii) if $H_{ii}$ is the i-th principle minor of $H(\underset{\sim}{x}_1)$, define

$$A(\underset{\sim}{x}_1) = \begin{cases} H(\underset{\sim}{x}_1)^{-1} & \text{when } |H_{ii}|>0,\ i=1,2,\ldots,L-1 \\ (H(\underset{\sim}{x}_1)+h)^{-1} & \text{when at least one } |H_{ii}|<0, \\ & \text{but } |H_{ii}|>=.0001,\ i=1,2,\ldots,L-1 \\ & \text{and all diagonal elements of} \\ & H(\underset{\sim}{x}_1) \text{ are positive, with h being} \\ & \text{the smallest diagonal element} \\ I & \text{when } |H_{ii}|<=.0001 \text{ for at least} \\ & \text{one i} \end{cases}$$

(iv) calculate the L-1 vector $\underset{\sim}{d}=A(\underset{\sim}{x}_1)\nabla F(\underset{\sim}{x}_1)$

(v) compute by direct line search $\lambda^* \varepsilon R^+$, which is the first minimum of $F(\underset{\sim}{x})$ along the ray $\underset{\sim}{x}=\underset{\sim}{x}_1-\lambda\underset{\sim}{d}$, $\lambda>0$

(vi) calculate $\underset{\sim}{x}_2=\underset{\sim}{x}_1-\lambda^*\underset{\sim}{d}$

(vii) repeat (ii)-(vi) until two consecutive solutions vectors are within a specified distance of one another.

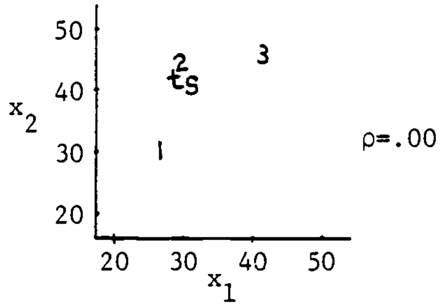
While the surface of $\text{var}[\hat{\mu}|\cdot]$ (i.e. the locus of values of $\text{var}[\hat{\mu}|\cdot]$ for specified method of allocation and fixed number of strata as the boundaries range over $\{x:a<x_1<x_2<\ldots<x_{L-1}<b\}$) for Dalenius's original equations is globally convex (with one minimum), the surface for the Dalenius-Gurney equations when g(x) is non-linear and $\rho$ small to

moderate, is not globally convex, possessing several local minima.

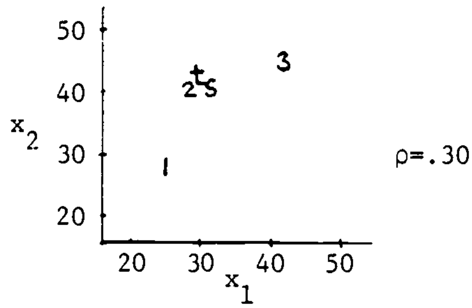
Along with the presence of several local minima, the solutions to the Dalenius-Gurney equations have another interesting facet. For fixed L, C and r, as $\rho$ varies over small to moderate values, the local minima roughly stay in the same location on the surface. However, the surface tends to "undulate", that is, the global minimum shifts from one location to another as $\rho$ increases. To see an example of this, we consider the case when X is normally distributed, and the regression function is quadratic with parameters C=1 and r=1. Employing Neyman allocation and fixing L=3 (so only two boundary points are used), the local minima are computed for the Dalenius-Gurney equations for a range of $\rho$ values. We also computed the unique minima for Singh's and Singh-Sukhatme's cum H(x) rules. The relative positions of these optimum boundaries, as well as the var$[\hat{\mu}|\text{Ney}]$ values they provide, are listed in Figure C.1. Note that the global minimum, which is located at point 2 for $\rho \leq .30$, shifts to point 3 when $\rho = .60$. Also note that the optimum boundaries for the two cum H(x) rules migrate from point 2 to point 3 as $\rho$ increases.

The only practical way to compute all local minima is to generate a grid of starting points that cover our surface. From each starting point, the above iterative procedure is employed to determine the nearest local minimum. Unfortunately, unless a very exhaustive grid is employed, there is no guarantee that the true global minimum was not the one local minimum a more modest grid missed. And such an exhaustive grid can potentially consume a great deal of computer time.

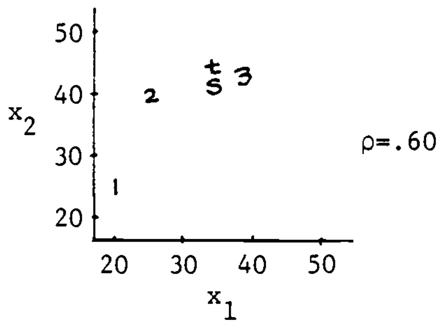
However, our interest in this study is comparing the approximate



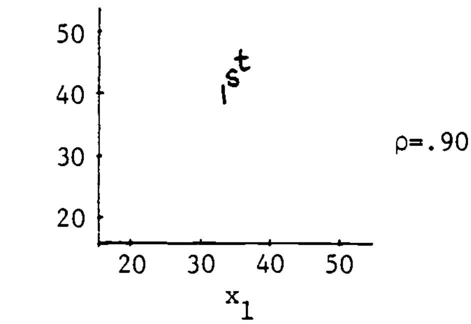

v(1)=58.30 v(s)=51.36
v(2)=47.34 v(t)=47.75
v(3)=58.39




v(1)=73.29 v(s)=53.72
v(2)=48.99 v(t)=51.04
v(3)=50.82




v(1)=111.23 v(s)=55.25
v(2)= 57.96 v(t)=58.15
v(3)= 49.42




v(1)=88.82* v(s)=89.05
v(t)=93.67

*Dalenius-Gurney equations have one solution.

1: First solution to Dalenius-Gurney equation
2: Second solution to Dalenius-Gurney equation
3: Third solution to Dalenius-Gurney equation
s: Solution to Singh's cum H(x) rule
t: Solution to Singh-Sukhatme's cum H(x) rule

v(•) = variance of $\hat{\mu}$ under specified boundary points

Figure C.1. Relative location of the solutions to the Dalenius-Gurney equations and Singh's Singh-Sukhatme's cum H(x) rules with the resulting var[$\hat{\mu}$] under Neyman allocation when C=1, r=1, L=3 and X is normally distributed.

boundaries with the best ones we can find. If the approximation does poorly when only compared to a local minimum, it will do worse when compared to the global minimum. If our approximations compare quite favorably to a local minimum that is the smallest of a number of local minima, while it may do worse when compared to the global minimum, using a simple iterative procedure, it has still compared quite well to a solution combining a fairly comprehensive search procedure with a grid of starting points. Thus, a great deal of effort will not be spent on attempting to identify the global minimum, and the results in Chapter III should be viewed as a practical comparison.