

**OPTIMIZACIÓN MULTIOBJETIVO EN LA
ASIGNACIÓN ÓPTIMA EN MUESTREO
ESTRATIFICADO MULTIVARIADO**

LILIANA ULLOA CORTEZ

T E S I S

**Presentada como requisito parcial
para obtener el grado de
Maestro en Ciencias en
Estadística Experimental**

**Universidad Autónoma Agraria
Antonio Narro
Subdirección de Postgrado
Buenavista, Saltillo, Coahuila, México
Marzo de 2006**

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO

Subdirección Postgrado

TÍTULO

OPTIMIZACIÓN MULTI OBJETIVO EN LA ASIGNACIÓN ÓPTIMA

EN

MUESTREO ESTRATIFICADO MULTIVARIADO

Por:

LILIANA ULLOA CORTEZ

Elaborada bajo la supervisión del comité particular de asesoría y aprobada como requisito parcial, para optar al grado de

MAESTRO EN CIENCIAS EN

ESTADÍSTICA EXPERIMENTAL

Comité Particular

Asesor principal: _____

Dr. José Antonio Díaz García

Asesor: _____

Dr. Fernando Esquivel Bocanegra

Asesor: _____

M.C. Félix de Jesús Sánchez Pérez

Dr. Jerónimo Landeros Flores

Subdirector de Postgrado

Buenavista, Saltillo, Coahuila. Marzo de 2006

AGRADECIMIENTOS

Quiero expresar mi agradecimiento:

- A mi familia entera por el apoyo que me brindaron estos meses. En especial a mis tíos Salvador y Lucy.
- A mis amigos: Isaac, Rosy, Paola, Lucy, Sarai, Paulina, Manuel, Romo, Lupita, Karla y Tony, gracias por permanecer, gracias por quererme como soy y siempre creer en mí, son parte importante en mi vida.
- Al Dr. José Antonio Díaz García, por guiarme en la conclusión de esta etapa, por la confianza y la paciencia.
- Al Dr. Fernando Esquivel Bocanegra y al M.C. Félix de Jesús Sánchez Pérez por su colaboración en este trabajo.
- A la Universidad Autónoma Agraria Antonio Narro, por haberme brindado la oportunidad de realizar el postgrado.
- A CONACYT por hacer posibles mis metas.

DEDICATORIA

Con todo mi amor y cariño para mis hermanos, Paco, Betty y Dany, y en especial a mi Mamá Rebeca Cortez Villarreal.

Gracias por ser mis amigos y estar conmigo siempre que los necesité. Los quiero.

En memoria de mi Papá, Francisco Ulloa Valdez.

Te extraño, no con la desesperanza de extrañar a un muerto.

Te extraño sin el agudo dolor del desamor.

Te extraño sin dolor, sin llanto, sin mal sabor.

Te extraño con alegría, con ilusión, con entusiasmo.

Te extraño con optimismo, con fuerza, con fe.

COMPENDIO

OPTIMIZACIÓN MULTIOBJETIVO EN LA ASIGNACIÓN ÓPTIMA EN MUESTREO ESTRATIFICADO MULTIVARIADO

POR

LILIANA ULLOA CORTEZ

MAESTRÍA

ESTADÍSTICA EXPERIMENTAL

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA. Marzo de 2006

Dr. José Antonio Díaz García -Asesor-

Palabras claves: Muestreo estratificado, muestreo estratificado multivariado, asignación óptima, optimización multiobjetivo, método función de valor, método lexicográfico, método de las ε -restricciones, método basado en distancias.

El presente trabajo plantea el problema de la asignación óptima en muestreo estratificado multivariado como un problema de optimización multiobjetivo de enteros en tres diferentes escenarios, información completa, parcial o nula. Se concluye con un ejemplo para cada uno de los métodos en diferentes técnicas.

ABSTRACT

MULTIOBJECT PROGRAMMING IN OPTIMUM ALLOCATION IN MULTIVARIATE STRATIFIED SURVEY:

BY

LILIANA ULLOA CORTEZ

MASTER

EXPERIMENTAL STATISTICS

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA. March, 2006

Ph. D. José Antonio Díaz García -Advisor-

Key Words: Stratified survey, multivariate stratified survey, optimum allocation, multiobject programming, value function method, lexicographical method, method of the ε -restrictions, method based on distances.

In this work we consider the allocation problem for multivariate stratified surveys as a problem of integer multiobject programming in three different scenes, complete information, partisan or null. It concludes with an example for each one of the methods in different techniques.

ÍNDICE

ÍNDICE DE CUADROS	ix
I. INTRODUCCIÓN	1
II. MUESTREO ESTRATIFICADO	6
MUESTREO ESTRATIFICADO	
UNIVARIADO	6
INTRODUCCIÓN	6
NOTACIÓN	8
ESTIMADORES Y SUS PROPIEDADES	8
INSESGABILIDAD	9
VARIANZA DE LOS ESTIMADORES	10
ESTIMACIONES DE LA VARIANZA PARA MUESTRAS	
ESTRATIFICADAS	11
INTERVALOS DE CONFIANZA PARA MUESTRAS	
ESTRATIFICADAS	11
ASIGNACIÓN ÓPTIMA	12
MUESTREO ESTRATIFICADO	
MULTIVARIADO	15
INTRODUCCIÓN	15
ASIGNACIÓN COMPROMISO	15
ASIGNACIÓN COMPROMISO MINIMIZANDO	
LA PÉRDIDA TOTAL RELATIVA	17
ASIGNACIÓN COMPROMISO TOMANDO LA MEDIA	
DE LOS VALORES ÓPTIMOS	19
ASIGNACIÓN COMPROMISO MINIMIZANDO	
LA TRAZA	20
EJEMPLO	21
ASIGNACIÓN FACTIBLE	22
ASIGNACIÓN MINIMIZANDO COSTOS	23

EJEMPLO (CONTINUACIÓN)	25
III. OPTIMIZACIÓN MULTIOBJETIVO	27
INTRODUCCIÓN	27
FUNCIÓN DE VALOR	29
EJEMPLO	30
MÉTODO LEXICOGRÁFICO	30
EJEMPLO	32
MÉTODO DE LAS ε -RESTRICCIONES	33
EJEMPLO (CONTINUACIÓN)	34
MÉTODO BASADO EN DISTANCIAS	34
EJEMPLO	37
IV. ASIGNACIÓN ÓPTIMA A TRAVÉS DE OPTIMIZACIÓN	
MULTIOBJETIVO	39
INTRODUCCIÓN	39
FUNCIÓN DE VALOR	42
MÉTODO LEXICOGRÁFICO	44
MÉTODO ε -RESTRICCIONES	47
MÉTODO BASADO EN DISTANCIAS	48
TÉCNICA DE DALENIUS	53
EJEMPLO	55
FUNCIÓN DE VALOR	55
MÉTODO LEXICOGRÁFICO	56
MÉTODO ε -RESTRICCIONES	57
MÉTODO DE DISTANCIAS	58
TÉCNICA DE DALENIUS	60
CONCLUSIONES	62
LITERATURA CITADA	65
ÁPENDICE A	67

ÍNDICE DE CUADROS

CUADRO 2.1.....	21
CUADRO 2.2.....	22
CUADRO 3.1.....	38
CUADRO 4.1.....	61
CUADRO 4.2.....	61

I. INTRODUCCIÓN

Una de las áreas de la estadística más utilizada en cualquier campo de la investigación científica, es el muestreo. Hablar de obtener buenos resultados en investigaciones médicas, sociales, etc., es hablar de un proceso de muestreo exitoso y bien aplicado. Es por esto que esta rama de la estadística provoca sin duda mucho interés por parte de los investigadores. El empleo de técnicas de muestreo efectivas en una población se traduce en la obtención de información útil para el logro del conocimiento de aspectos importantes de la misma. Pretender acceder a esta información por el análisis del total de la población sería prácticamente imposible, es por esto que el muestreo ofrece múltiples ventajas, entre los beneficios que ofrece se tiene que:

1. El muestreo de una población permite obtener información confiable con costos mucho menores que los de un censo.
2. Con una muestra es más ágil la recolección de información, así como el proceso de análisis de los datos recolectados.
3. Con el muestreo se puede obtener estimaciones más precisas que las basadas en un censo, esto por la facilidad de manejar una menor cantidad de datos.

Es por estos hechos que resulta fundamental el estudio y desarrollo de las técnicas de muestreo, Cochran (1977).

Dentro de los esquemas de muestreo más utilizados se tiene el muestreo estratificado. Este tipo de muestreo es muy recurrido ya que toma en cuenta que muy difícilmente se encontrará una población homogénea. Por ejemplo, hablando de un estudio de los habitantes en un país, generalmente existen grupos (o clases sociales) con actitudes que difieren significativamente del resto de los habitantes, y es de esperarse que no estén repartidos al azar entre la población, sino que tiendan a agruparse en ciertas regiones.

Se querrá entonces minimizar la posibilidad de que la muestra pueda fallar o no representar cualquiera de estos grupos (estratos) en una muestra realizada con muestreo aleatorio simple.

Comúnmente, cierto tipo de fenómenos tienden a presentar divisiones naturales en su población (hombres y mujeres, entidades federativas, etc.). Sea este el caso, o sea el caso en que se complique un poco más estratificar a la población, el hecho de poder considerar a cada una de estas divisiones (estratos) como una población independiente y entonces poder trabajar con cualquier otro esquema de muestreo que mejor convenga al investigador, es una de las mayores ventajas del muestreo estratificado.

Para determinar el tamaño de muestra en muestreo estratificado, se tiene la asignación óptima, que como su nombre lo señala, se establece a través de un problema de optimización donde la función objetivo es la varianza, sujeta a una res-

tricción de costos, o viceversa. Tradicionalmente, este problema se resuelve usando la desigualdad de Cauchy-Schwarz (Stuart 1954), Cochran (1977) o el método de Multiplicadores de Lagrange, Sukhatme *et al.* (1984). Ambas maneras de resolver el problema ofrecen un resultado en el terreno de los números reales, lo cual, desde un punto de vista práctico y aplicado no es lo más sencillo de interpretar. Es por esto que el presente trabajo pretende integrar una restricción más al problema, y es que la solución se limite al espacio de los números naturales, es decir, que los tamaños de muestra obtenidos para cada estrato sean enteros positivos.

La manera usual en que se han utilizado las técnicas de muestreo y en particular el muestreo estratificado, es de forma univariada, es decir, cuando el tamaño de la muestra y su asignación en los estratos es propuesta tomando en cuenta una sola variable de decisión o característica, Cochran (1977).

En el contexto del muestreo estratificado, algunos esfuerzos se han realizado con el objetivo de establecer el tamaño de muestra y su asignación en los estratos tomando en cuenta varias características, Sukhatme *et al.* (1984).

Cuando la asignación óptima se hace proponiendo como función objetivo una función de costos sujeta a las restricciones de las varianzas en las diferentes características, el problema se reduce a un problema de programación matemática clásica y es tratado por Arthanari y Dodge (1981).

Alternativamente, cuando el interés es “minimizar” las varianzas sujetas a una función de costos, o a un tamaño de muestra establecido, tal problema se ha resuelto de diferentes maneras, Sukhatme *et al.* (1984). Pero como se establecerá,

todos estos métodos son casos particulares de una de las técnicas de optimización multiobjetivo, la cual tiene como primicia, que la población bajo estudio es conocida totalmente, es decir, se conoce una función escalar que establece una relación entre las varianzas de cada característica de forma única, en el contexto del muestreo. Tal hipótesis, en general no es aplicable.

En base a lo anterior, en el presente trabajo se establece propiamente, el problema de asignación óptima en el muestreo estratificado, como un problema de optimización multiobjetivo de enteros, estudiando diferentes técnicas para su solución. Soluciones establecidas en función del conocimiento previo de la población el cual se clasificará como información completa, parcial o nula.

Para poder presentar de manera organizada todo lo anterior, el trabajo se divide en cuatro capítulos y un apéndice. El Capítulo II ofrece los conceptos del muestreo estratificado univariado y multivariado. Para el caso univariado se expone un ejemplo de cuando se puede emplear el muestreo estratificado, se da la notación a utilizar en el resto del trabajo, se revisa la teoría, y la asignación óptima. En el caso multivariado, se hace una revisión a los tipos de asignación que existen en la literatura y se da un ejemplo.

Dado que para poder abordar el problema de forma multivariada se tendrá que hacer uso de las técnicas de optimización multiobjetivo, el Capítulo III muestra una revisión de dichas técnicas y ofrece ejemplos algebraicos de las mismas.

Para finalizar, el Capítulo IV plantea el problema de asignación óptima a través de optimización multiobjetivo. Se expone el problema inicial al que se

enfrenta, abordado como una función de varianzas sujeta a una función de costos, y como una función de costos sujeta a una función de varianzas. Se enfatizan cuales son las condiciones de la muestra, y también del uso de la varianza muestral en lugar de la varianza poblacional por ser generalmente desconocida. Se utilizan los métodos de optimización multiobjetivo para la solución del problema inicial y se plantea una forma alternativa de abordar el problema. Finalmente todas las técnicas estudiadas en el Capítulo II son aplicadas a un problema de la literatura, cuyas soluciones son obtenidas a través del paquete computacional LINGO.

El Apéndice A hace una revisión a la técnica de los multiplicadores de Lagrange.

II. MUESTREO ESTRATIFICADO

Muestreo Estratificado Univariado

Introducción

En múltiples ocasiones cuando pretendemos realizar una investigación o experimento, se tiene información que ayuda a decidir como se podría elaborar una muestra, en otras palabras, la naturaleza del problema y los datos a estudiar en sí, aportan esta información con la cual se puede lograr una muestra representativa.

Por ejemplo, en los alumnos aspirantes a una universidad, se sabe que los aspirantes a la carrera de derecho son muchos más que los alumnos aspirantes a la carrera de filosofía. Este es el tipo de información previa que nos ayuda a decidir el esquema de muestreo, ya que elegir una muestra directamente de ambos grupos de aspirantes podría repercutir en una mala representación de los datos. En casos como este resulta conveniente fraccionar la población original de N unidades en L subpoblaciones de tamaño N_h tales que $N_1 + N_2 + \dots + N_L = N$; de tal forma que cada unidad de la población pertenece sólo a una subdivisión y la unión de todas ellas conforma la población original. Si se extrae una muestra independiente de cada subdivisión y posteriormente se reúne la información para obtener las estimaciones globales de la población, tal esquema de muestreo se le conoce como *muestreo estratificado* y a cada subdivisión considerada de manera independiente se denomina

estrato.

Saber concretamente cuando se debe realizar el muestreo estratificado podría representar un problema, este se puede evitar si se consideran las siguientes razones:

1. Se pretende proteger contra la posibilidad de obtener una mala muestra, es decir, hablando nuevamente de la naturaleza de la población, que ésta sea demasiado heterogénea y proporcione información que sea poco representativa.
2. Si los datos deseados deben tener una precisión conocida en algunas subdivisiones de la población.
3. Una muestra estratificada podría administrarse, de manera más conveniente, a un menor costo. Esto puede verse claramente en la forma de recolección de datos, en un estrato podría aplicarse una encuesta telefónica y en otro estrato una encuesta personal.
4. Si se hace correctamente, dará estimaciones más precisas (con menor varianza).

Si la naturaleza de nuestro experimento presenta por lo menos una de estas necesidades es conveniente utilizar el muestreo estratificado, Cochran (1977). Además, una ventaja del muestreo estratificado es que, como cada estrato funciona independientemente, se puede aplicar dentro de cada uno, el esquema de muestreo que más convenga para elegir los elementos concretos que formarán parte de la muestra.

Notación

El subíndice $h = 1, 2, \dots, L$ denota el estrato, e $i = 1, 2, \dots, N_h$ la unidad dentro del estrato h . Además

N_h	número total de unidades en el estrato h
n_h	número de unidades de la muestra en el estrato h
y_{hi}	valor obtenido para la i -ésima unidad en el estrato h
$\mathbf{n} = (n_1, n_2, \dots, n_L)'$	vector del número de unidades de la muestra
$W_h = \frac{N_h}{N}$	tamaño relativo del estrato h
$f_h = \frac{n_h}{N_h}$	fracción de muestreo en el estrato h
$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h}$	media poblacional del estrato h
$\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$	media muestral en el estrato h
$S_h^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2}{N_h - 1}$	varianza poblacional en el estrato h
c_h	costo por unidad de muestreo en el estrato h

Estimadores y sus propiedades

Un estimador es una función de la muestra que se utiliza para estimar un parámetro desconocido de la población. El estimador de la media poblacional usado

en un muestreo aleatorio estratificado es:

$$\bar{y}_{st} = \frac{\sum_{h=1}^L N_h \bar{y}_h}{N} = \sum_{h=1}^L W_h \bar{y}_h.$$

El estimador \bar{y}_{st} , en general, no es el mismo que el estimador de la media muestral \bar{y} que es,

$$\bar{y} = \frac{\sum_{h=1}^L n_h \bar{y}_h}{n}.$$

La diferencia entre ambos estimadores radica en que \bar{y}_{st} , las estimaciones a partir de estratos individuales, reciben sus ponderaciones correctas $\frac{N_h}{N}$. Es evidente que \bar{y} coincide con \bar{y}_{st} cuando en cada estrato

$$\frac{n_h}{n} = \frac{N_h}{N} \quad \text{o} \quad \frac{n_h}{N_h} = \frac{n}{N} \quad \text{o} \quad f_h = f.$$

Esto es, cuando la fracción de muestreo es la misma en todos los estratos, y se describe como estratificación con asignación *proporcional* de los números n_h y da lugar a una muestra autoponderada.

Las principales propiedades del estimador \bar{y}_{st} se enuncian enseguida y son una consecuencia directa de las propiedades de los estimadores en una muestra aleatoria simple, Cochran (1977).

Insesgabilidad

Ésta es una de las propiedades deseables para un estimador. El término *insegado* se refiere al hecho de que la esperanza del estimador de la media poblacional usada en muestreo estratificado, es decir, \bar{y}_{st} , coincide con la media de la población \bar{Y} . El siguiente teorema demuestra esta propiedad.

Teorema 1. Si en cada estrato la estimación muestral \bar{y}_h es insesgada, entonces \bar{y}_{st} es una estimación insesgada de la media de población \bar{Y} .

$$\begin{aligned} E(\bar{y}_{st}) &= E\left(\sum_{h=1}^L W_h \bar{y}_h\right) \\ &= \sum_{h=1}^L W_h \bar{Y}_h \\ &= \bar{Y}. \end{aligned}$$

Varianza de los estimadores

Esta propiedad es importante ya que un estimador con menor varianza es más probable que dé lugar a estimaciones que estén más próximas al verdadero valor del parámetro.

Teorema 2. Si las muestras se extraen independientemente en los diferentes estratos, entonces

$$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 V(\bar{y}_h),$$

donde $V(\bar{y}_h)$ es la varianza de \bar{y}_h sobre muestras repetidas del estrato de prueba h .

Teorema 3. Para el muestreo aleatorio estratificado, la varianza de la estimación \bar{y}_{st} es

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h(N_h - n_h) \frac{S_h^2}{n_h} = \sum_{h=1}^L W_h^2 \frac{S_h^2}{n_h} (1 - f_h). \quad (1)$$

En los siguientes corolarios se dan algunos casos particulares de esta fórmula.

Corolario 1. Si las fracciones de muestreo $\frac{n_h}{N_h}$ son despreciables en todos los estratos, entonces

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L \frac{N_h^2 S_h^2}{n_h} = \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h}.$$

Corolario 2. Con asignación proporcional, se sustituye

$$n_h = \frac{nN_h}{N},$$

en (1). La varianza se reduce a

$$V(\bar{y}_{st}) = \sum_{h=1}^L \frac{N_h S_h^2}{N} \frac{1}{n} \left(\frac{N - n}{N}\right) = \frac{1 - f}{n} \sum_{h=1}^L W_h S_h^2.$$

Corolario 3. Si el muestreo es proporcional y las varianzas en todos los estratos tienen el mismo valor, S_w^2 , se obtiene el siguiente resultado

$$V(\bar{y}_{st}) = \frac{S_w^2}{n} \left(\frac{N-n}{N} \right).$$

Teorema 4. Si $\hat{Y}_{st} = N\bar{y}_{st}$ es la estimación del total de la población Y , entonces

$$V(\hat{Y}_{st}) = \sum_{h=1}^L N_h(N_h - n_h) \frac{S_h^2}{n_h}.$$

Estimaciones de la varianza para muestras estratificadas

Si se toma una muestra aleatoria simple dentro de cada estrato, una estimación insesgada de S_h^2 es

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2.$$

Por lo tanto, se obtiene el siguiente resultado.

Teorema 5. Con muestreo aleatorio estratificado, una estimación insesgada de la varianza de \bar{y}_{st} es

$$\hat{V}(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h(N_h - n_h) \frac{s_h^2}{n_h}.$$

Una forma alternativa para propósitos de cálculo es

$$\hat{V}(\bar{y}_{st}) = \sum_{h=1}^L \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^L \frac{W_h s_h^2}{N}.$$

Para calcular este estimador, debe haber cuando menos dos unidades provenientes de todos y cada uno de los estratos.

Intervalos de confianza para muestras estratificadas

En el contexto de estimar un parámetro poblacional, un intervalo de confianza es un rango de valores (calculado en una muestra) en el cual se encuentra posiblemente el verdadero valor del parámetro, con una probabilidad determinada.

Si se considera una muestra grande dentro de cada estrato, o el diseño de muestreo tiene una gran cantidad de estratos, un intervalo de confianza aproximado del $100(1 - \alpha)\%$ para la media, vea Lohr (2000), esta dado por;

$$\text{Media de la población: } \bar{y}_{st} \pm z_{\alpha/2} s(\bar{y}_{st}).$$

$$\text{Total de la población: } N\bar{y}_{st} \pm z_{\alpha/2} N s(\bar{y}_{st})$$

Estas fórmulas suponen que \bar{y}_{st} está distribuida normalmente y que $s(\bar{y}_{st})$ está bien determinada, de modo que el percentil $z_{\alpha/2}$ pueda encontrarse en las tablas de distribución normal estándar.

Asignación óptima

Cuando las varianzas S_h^2 en cada estrato varían mucho, se busca reducir los costos de muestreo y aumentar la precisión, para esto se hace uso de la asignación óptima. La asignación óptima funciona mejor para las unidades de muestreo que varían mucho en magnitud, así como también cuando los costos de muestreo son diferentes entre los estratos

Entonces para determinar el tamaño de muestra n_h en los diferentes estratos se debe considerar una de las 2 siguientes opciones:

1. Minimizar la $V(\bar{y}_{st})$ para una función de costos específica.
2. Minimizar la función de costos para un valor específico de la $V(\bar{y}_{st})$.

Una función de costo lineal sencilla es:

$$C = c_0 + \sum_{h=1}^L c_h n_h, \tag{2}$$

donde C es el costo total, c_0 representa un costo adicional fijo, c_h el costo de una observación en el estrato h .

Dentro de cualquier estrato el costo es directamente proporcional al tamaño de muestra, aunque el costo por unidad c_h puede variar en los diferentes estratos.

Teorema 6. *En el muestreo aleatorio estratificado, con una función de costo lineal de la forma (2), la varianza de la media estimada \bar{y}_{st} es un mínimo para el costo específico C y el costo es un mínimo para una varianza específica $V(\bar{y}_{st})$ donde n_h es proporcional a $\frac{W_h S_h}{\sqrt{c_h}}$.*

En términos del tamaño de muestra n_h en un estrato, se tiene

$$\frac{n_h}{n} = \frac{\frac{W_h S_h}{\sqrt{c_h}}}{\sum_{h=1}^L \frac{W_h S_h}{\sqrt{c_h}}} = \frac{\frac{N_h S_h}{\sqrt{c_h}}}{\sum_{h=1}^L \frac{N_h S_h}{\sqrt{c_h}}}. \quad (3)$$

Este teorema da lugar a las siguientes reglas de conducta. En un estrato dado, se toma una muestra más grande si

- El estrato es más grande.
- El estrato es más variable internamente.
- El muestreo es más barato en el estrato.

Para completar la asignación se necesita algo más, la ecuación (3) da el n_h en términos de n , pero aun no se conoce el valor de n . La solución depende de:

- Si la muestra se escoge para satisfacer un costo total especificado C .

Si este es el caso, se sustituyen los valores óptimos de n_h en la función de costo

(2) y se resuelve para n , lo que da

$$n = \frac{(C - c_0) \sum_{h=1}^L \left(\frac{N_h S_h}{\sqrt{c_h}} \right)}{\sum_{h=1}^L (N_h S_h \sqrt{c_h})}.$$

- Si la muestra se escoge para dar una varianza especificada V para \bar{y}_{st} .

En este caso, se sustituye la n_h óptima en (1) de donde finalmente

$$n = \frac{\left(\sum_{h=1}^L W_h S_h \sqrt{c_h} \right) \sum_{h=1}^L \frac{W_h S_h}{\sqrt{c_h}}}{V + \left(\frac{1}{N} \right) \sum_{h=1}^L W_h S_h^2},$$

donde $W_h = \frac{N_h}{N}$.

Un caso especial de asignación óptima ocurre cuando $c_h = c$, es decir, si el costo por unidad es el mismo en todos los estratos. Este caso es llamado *asignación Neyman*, y el costo se convierte en $C = c_0 + cn$ y la asignación óptima para un costo fijo se reduce a la asignación óptima para el tamaño de muestra fijo. El resultado para este caso particular es el siguiente.

Teorema 7. *En el muestreo aleatorio estratificado la $V(\bar{y}_{st})$ se minimiza para un tamaño de muestra total fijo n si*

$$n_h = n \frac{W_h S_h}{\sum_{h=1}^L W_h S_h} = n \frac{N_h S_h}{\sum_{h=1}^L N_h S_h}. \quad (4)$$

Una fórmula para la varianza mínima con n fija se obtiene al sustituir el valor n en (4) en (1). El resultado es

$$V_{\min}(\bar{y}_{st}) = \frac{\left(\sum_{h=1}^L W_h S_h \right)^2}{n} - \frac{\sum_{h=1}^L W_h S_h^2}{N}.$$

Hasta aquí se ha abordado sólo el caso del muestreo estratificado univariado, a continuación se revisarán los conceptos para ampliar el tema al caso multivariado.

Muestreo Estratificado Multivariado

Introducción

Cuando se obtiene una muestra, regularmente se presenta el problema de estimar varias características de la población. Esto usualmente se torna complicado dado que las diferentes características podrán tener diferentes varianzas provocando que los tamaños de muestra para cada característica puedan ser diferentes. En la presente sección se muestran las diferentes propuestas que se han desarrollado para resolver el problema de asignación óptima en muestreo estratificado multivariado.

De acuerdo con Neyman (1934) la solución del problema se basa en las varianzas, es decir que, varianzas de diferentes características son posiblemente correlacionadas si las características por si mismas son posiblemente correlacionadas. Si este es el caso, la asignación óptima para cualquier característica será también apropiada para otras características. Sin embargo, si las diferentes características no son posiblemente correlacionadas, el tamaño total de la muestra puede ser distribuido entre diferentes estratos en proporción a sus tamaños.

Así el problema de optimización multivariado consistirá en evaluar las varianzas de las j características, $j = 1, 2, \dots, G$.

Asignación compromiso

La *asignación compromiso* fue propuesta por Peters y Bucher (sin fecha) y se basa en la maximización de su eficiencia relativa promedio con respecto a la

asignación óptima, es decir, se desea

$$\begin{aligned} \max_{\mathbf{n}} \quad & \frac{1}{G} \sum_{j=1}^G \left[\frac{V(\bar{y}_{st}^j)_N}{V(\bar{y}_{st}^j)_C} \right] \\ \text{sujeto a} \quad & \\ & \sum_{h=1}^L n_h = n, \end{aligned} \tag{5}$$

donde \bar{y}_{st}^j denota la media poblacional estimada de la característica j , $j = 1, 2, \dots, G$,

$$V(\bar{y}_{st}^j)_C = \sum_{h=1}^L W_h^2 S_{hj}^2 \left(\frac{1}{n_h} - \frac{1}{N_h} \right), \tag{6}$$

es la varianza compromiso y

$$V(\bar{y}_{st}^j)_N = \frac{\left(\sum_{h=1}^L W_h S_{hj} \right)^2}{n} - \frac{1}{N} \sum_{h=1}^L W_h S_{hj}^2,$$

es la varianza de Neyman, para la j -ésima característica, y S_{hj}^2 es la varianza poblacional de la característica j en el estrato h .

Así, aplicando el método de multiplicadores de Lagrange (vea Apéndice A), se tiene

$$L(\mathbf{n}, \lambda) = \frac{1}{G} \sum_{j=1}^G \left[\frac{V(\bar{y}_{st}^j)_N}{V(\bar{y}_{st}^j)_C} \right] + \lambda \left[\sum_{h=1}^L n_h - n \right], \tag{7}$$

donde λ denota el multiplicador de Lagrange. Luego el punto crítico que maximiza el programa restringido (5) se obtiene a partir del programa sin restricciones (7) como sigue.

Derivando respecto a n_h y λ se tiene

$$\begin{aligned} \frac{\partial L(\mathbf{n}, \lambda)}{\partial n_h} &= \frac{1}{G} \sum_{j=1}^G \frac{V(\bar{y}_{st}^j)_N}{[V(\bar{y}_{st}^j)_C]^2} \cdot \frac{W_h^2 S_{hj}^2}{n_h^2} + \lambda, \quad h = 1, 2, \dots, L \\ \frac{\partial L(\mathbf{n}, \lambda)}{\partial \lambda} &= \sum_{h=1}^L n_h - n. \end{aligned}$$

Ahora igualando a cero y resolviendo el sistema de ecuaciones no lineales resultante, se obtiene el punto crítico deseado.

Asignación compromiso minimizando la pérdida total relativa

Una forma alternativa de abordar el problema de asignación óptima en el muestreo estratificado multivariado se basa en observar que, el uso de la asignación compromiso para estimar la media poblacional de la j -ésima característica tendrá cierta pérdida en la precisión relativa con respecto a la asignación óptima. Una medida de esta pérdida de precisión relativa, es

$$P_j = \frac{V(\bar{y}_{st}^j)_C - V(\bar{y}_{st}^j)_{Opt}}{V(\bar{y}_{st}^j)_{Opt}}, \quad (8)$$

donde $V(\bar{y}_{st}^j)_C$ es la varianza compromiso definida en (6) y

$$V(\bar{y}_{st}^j)_{Opt} = \frac{\left(\sum_{h=1}^L W_h S_{hj} \sqrt{c_h} \right)^2}{C} - \frac{1}{N} \sum_{h=1}^L W_h S_{hj}^2, \quad (9)$$

es la varianza óptima. Donde C es definido como en (2).

Esta aproximación propuesta por Geary (1949) consiste en minimizar esta pérdida relativa total, es decir $\sum_{j=1}^G P_j$, sujeto a la condición $\sum_{h=1}^L n_h = n$ esto es, el problema a resolver se puede plantear como el siguiente programa no lineal restringido:

$$\min_{\mathbf{n}} \sum_{j=1}^G \left[\frac{V(\bar{y}_{st}^j)_C - V(\bar{y}_{st}^j)_{Opt}}{V(\bar{y}_{st}^j)_{Opt}} \right]$$

sujeto a

$$\sum_{h=1}^L n_h = n.$$

Más generalmente, Dalenius (1957) consideró una función lineal ponderada de la forma $\sum_{j=1}^G \alpha_j P_j$, siendo α_j las ponderaciones dadas a las diferentes carac-

terísticas en proporción a su importancia. Se determinará n_h tal que $\sum_{j=1}^G \alpha_j P_j$ sea un mínimo sujeto a la condición que $\sum_{h=1}^L c_h n_h \leq C$. Es decir,

$$\min_{\mathbf{n}} \sum_{j=1}^G \left[\alpha_j \left(\frac{V(\bar{y}_{st}^j)_C - V(\bar{y}_{st}^j)_{Opt}}{V(\bar{y}_{st}^j)_{Opt}} \right) \right]$$

sujeto a

$$\sum_{h=1}^L c_h n_h \leq C. \quad (10)$$

Procediendo como anteriormente se hizo, es decir, utilizando el método de multiplicadores de Lagrange se tiene

$$L(\mathbf{n}, \lambda) = \sum_{j=1}^G \alpha_j P_j + \lambda \left[\sum_{h=1}^L c_h n_h - C \right], \quad (11)$$

donde P_j es definido como en (8) y $V(\bar{y}_{st}^j)_{Opt}$ en (9). Donde, como antes, λ denota el multiplicador de Lagrange.

Derivando con respecto a n_h como se hizo anteriormente se tiene

$$\frac{\partial L(\mathbf{n}, \lambda)}{\partial n_h} = W_h^2 \sum_{j=1}^G [(\alpha_j S_{hj}^2) \setminus V(\bar{y}_{st}^j)_{Opt}] - \lambda c_h n_h^2, \quad h = 1, 2, \dots, L.$$

Ahora igualando a cero y resolviendo el sistema de ecuaciones no lineales resultante, se obtienen los valores óptimos para n_h .

Usando el hecho que $\sum_{h=1}^L c_h n_h = C$ (cuando $c_0 = 0$), los valores óptimos de n_h son dados por el conjunto de ecuaciones

$$n_h = \frac{CW_h \sqrt{\frac{1}{c_h} \sum_{j=1}^G [(\alpha_j S_{hj}^2) \setminus V(\bar{y}_{st}^j)_{Opt}]}}{\sum_{h=1}^L W_h \sqrt{c_h \sum_{j=1}^G [(\alpha_j S_{hj}^2) \setminus V(\bar{y}_{st}^j)_{Opt}]}}.$$

La asignación anterior se reduce a la siguiente ecuación cuando todas las características son igualmente importantes y el costo por unidad de la enumeración es el mismo para todos los estratos.

$$n_h = nW_h \frac{\sqrt{\sum_{j=1}^G S_{hj}^2 / \left(\sum_{h=1}^L W_h S_{hj}\right)^2}}{\sum_{h=1}^L W_h \sqrt{\sum_{j=1}^G S_{hj}^2 / \left(\sum_{h=1}^L W_h S_{hj}\right)^2}},$$

donde $C = nc$.

Para este tipo de asignación, la varianza óptima de la media poblacional estimada, para la j -ésima característica es

$$V(\bar{y}_{st}^j) = \frac{1}{n} \sum_{h=1}^L \frac{W_h S_{hj}^2}{\sqrt{\sum_{j=1}^G S_{hj}^2 / \left(\sum_{h=1}^L W_h S_{hj}\right)^2}} \cdot \sum_{h=1}^L W_h \sqrt{\sum_{j=1}^G S_{hj}^2 / \left(\sum_{h=1}^L W_h S_{hj}\right)^2} - \frac{1}{N} \sum_{h=1}^L W_h S_{hj}^2,$$

vea Sukhatme *et al.* (1984).

Asignación compromiso tomando la media de los valores óptimos

Otra asignación compromiso sugerida por Sukhatme *et al.* (1984) es la asignación de la media dada por

$$n_h = \frac{1}{G} \sum_{j=1}^G n_{hj},$$

donde n_{hj} es el valor óptimo del tamaño de muestra que se toma del h -ésimo estrato para estimar la media poblacional para la j -ésima característica, y está dado por

$$n_{hj} = C \frac{W_h S_{hj}}{\sqrt{c_h}} \bigg/ \sum_{h=1}^L W_h S_{hj} \sqrt{c_h}.$$

Asignación compromiso minimizando la traza

Una aproximación diferente basada en la minimización de la varianza generalizada (esto es el determinante de la matriz de varianza-covarianza de la media \bar{y}_{st}^j , $j = 1, 2, \dots, G$) fue sugerida por Dalenius (1957) y explorada después por Ghosh (1958) y Aoyama (1963). Sin embargo, esta aproximación no es totalmente satisfactoria. La solución explícita es complicada y no puede ser obtenida excepto a través de métodos iterativos. Basada en consideraciones similares, una aproximación que se sugiere es la minimización de la traza de la matriz de varianzas y covarianzas sujeta a la condición $\sum_{h=1}^L c_h n_h = C$ Sukhatme *et al.* (1984).

Así el programa restringido será

$$\begin{aligned} \min_{\mathbf{n}} \sum_{j=1}^G V(\bar{y}_{st}^j)_C \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h = C, \end{aligned} \tag{12}$$

donde $\sum_{j=1}^G V(\bar{y}_{st}^j)_C$ es la traza de la matriz de varianzas y covarianzas.

Aplicando el método de multiplicadores de Lagrange como se tiene

$$L(\mathbf{n}, \lambda) = \sum_{j=1}^G V(\bar{y}_{st}^j)_C + \lambda \left[\sum_{h=1}^L c_h n_h - C \right]. \tag{13}$$

Nuevamente, derivando con respecto a n_h e igualando a cero se obtienen que los valores óptimos para n_h están dados por

$$n_h = \frac{CW_h}{\sqrt{c_h}} \sqrt{\sum_{j=1}^G S_{hj}^2} / \sum_{h=1}^L W_h \sqrt{\sum_{j=1}^G S_{hj}^2}.$$

Para este tipo de asignación, la varianza óptima de la media poblacional

estimada para la j -ésima característica es

$$V(\bar{y}_{st}^j) = \frac{1}{C} \sum_{h=1}^L \frac{W_h \sqrt{c_h S_{hj}^2}}{\sqrt{\sum_{j=1}^G S_{hj}^2}} \cdot \sum_{h=1}^L W_h \sqrt{c_h} \sqrt{\sum_{j=1}^G S_{hj}^2} - \frac{1}{N} \sum_{h=1}^L W_h S_{hj}^2.$$

Ejemplo

Para tener una idea acerca de la eficiencia relativa de la asignación compromiso sobre la asignación proporcional se considerarán los datos de una encuesta realizada en el distrito de Banaras en la India, Sukhatme *et al.* (1984). Serán evaluadas 2 características relevantes (1) el área cultivada de arroz y (2) el área cultivada total, estos datos son mostrados en el Cuadro 2.1. El número total de pueblos en el distrito fue de 4190. Se asumirá que estas dos características son igualmente importantes y el tamaño total de muestra fija es de 382, es decir la restricción de este problema es $\sum_{h=1}^L n_h = n$, donde $n = 382$.

CUADRO 2.1: Datos del ejemplo.

Estrato No.	No. de pueblos N_h	Área de arroz s_{h1}^2	Área cultivada s_{h2}^2
1	1,419	4,817.72	130,121.15
2	619	6,251.26	7,613.52
3	1,253	3,066.16	1,456.40
4	899	56,207.25	66,977.72

Se usará el estimador s_{hj}^2 como valor verdadero de S_{hj}^2 , los valores de n_h para las diferentes asignaciones y las eficiencias relativas de las asignaciones compromiso con respecto a la asignación proporcional son mostrados en el Cuadro 2.2.

CUADRO 2.2: Eficiencia Relativa de las diferentes asignaciones compromiso con respecto a la asignación proporcional.

Asignación	n_1	n_2	n_3	n_4	Traza	Eficiencia relativa
Proporcional	129	57	114	82	180	
Compromiso: min. pérdida total rel.	163	33	43	143	135.18	1.33
media de valores ópt.	159	34	42	147	136.45	1.32
min. la traza	201	28	32	121	129.33	1.39

Se puede observar que la asignación compromiso basada en la minimización de la traza de la matriz de varianzas- covarianzas es la más eficiente. Esto se calcula de la siguiente manera

$$eficiencia = \frac{\text{Traza asignación proporcional}}{\text{Traza asignación compromiso}} = \frac{V(\bar{y}_{st}^1)_P + V(\bar{y}_{st}^2)_P}{V(\bar{y}_{st}^1)_C + V(\bar{y}_{st}^2)_C},$$

donde $V(\bar{y}_{st}^1)_P$ y $V(\bar{y}_{st}^2)_P$ son las varianzas calculadas con los tamaños de muestra obtenidos por asignación proporcional, y $V(\bar{y}_{st}^1)_C$ y $V(\bar{y}_{st}^2)_C$ son las varianzas calculadas con los tamaños de muestra obtenidos por asignación compromiso. La eficiencia relativa de la *Traza asignación compromiso minimizando la traza* es mayor con respecto a la *traza asignación proporcional*, es por esto que es la más eficiente.

Asignación factible

Otro punto de vista del problema de asignación es el sugerido por Folks y Antle (1965). Menciona que cualquier vector $\mathbf{n} = (n_1, n_2, \dots, n_L)'$ tal que $n_h \geq 0$ y $\sum_{h=1}^L n_h = n$ será llamado *asignación factible*. El problema radica en elegir una asignación que sea mejor que otras. Esto puede ser aclarado por las siguientes definiciones.

Definición 1. Una asignación \mathbf{n} es mejor que \mathbf{n}^\diamond si $V(\bar{y}_{st}^j | \mathbf{n}) \leq V(\bar{y}_{st}^j | \mathbf{n}^\diamond)$ para todo j con estricta desigualdad para al menos una j .

Definición 2. Una asignación factible \mathbf{n} es llamada eficiente si no hay otra asignación la cual sea mejor que \mathbf{n} .

Definición 3. Un conjunto de asignaciones factibles eficientes es llamado completo si dada cualquier asignación \mathbf{n} que no este en el conjunto, existe una asignación \mathbf{n}^* en el conjunto la cual es mejor que \mathbf{n} .

Folks y Antle demostraron que un conjunto completo de asignaciones factibles eficientes es el conjunto $\mathbf{n}^* = (n_1, n_2, \dots, n_L)$ donde

$$n_h = \frac{nW_h \sqrt{\sum_{j=1}^G \alpha_j S_{hj}^2}}{\sum_{h=1}^L W_h \sqrt{\sum_{j=1}^G \alpha_j S_{hj}^2}}, \quad \sum_{j=1}^G \alpha_j \quad \text{y} \quad \alpha_j \geq 0.$$

Asignación minimizando costos

Las diferentes aproximaciones al caso multivariado propuestas hasta ahora, son extensiones naturales de la aproximación a el caso univariado cuando el costo total disponible para la encuesta es fijado por adelantado.

Si c_h es el costo por unidad, de la colección de datos en el h -ésimo estrato, el costo total de la encuesta para la colección de datos es $\sum_{h=1}^L c_h n_h$. Yates (1960) propuso una asignación que minimiza el costo total sujeto a la condición de que $V(\bar{y}_{st}^j)_c = V_o^j$, donde los valores V_o^j ($j = 1, 2, \dots, G$) son especificados por adelantado.

Es decir,

$$\min_{\mathbf{n}} \sum_{h=1}^L c_h n_h$$

sujeto a

$$V(\bar{y}_{st}^j)_c = V_o^j, \quad j = 1, 2, \dots, G$$

donde $V(\bar{y}_{st}^j)_c$ es la varianza para la j -ésima característica, $j = 1, 2, \dots, G$.

Aplicando los multiplicadores de Lagrange se tiene la siguiente función

$$\phi = \sum_{h=1}^L c_h n_h + \sum_{j=1}^G \lambda_j [V(\bar{y}_{st}^j)_c - V_o^j],$$

donde λ_j , $j = 1, 2, \dots, G$ son los multiplicadores de Lagrange. Derivando con respecto a \mathbf{n} se puede ver que los valores óptimos para n_h son

$$n_h = \frac{W_h}{\sqrt{c_h}} \sqrt{\sum_{j=1}^G \lambda_j S_{hj}^2},$$

donde las constantes λ_j , $j = 1, 2, \dots, G$ son determinadas por las G restricciones,

$$V(\bar{y}_{st}^j)_c = V_o^j, \quad j = 1, 2, \dots, G.$$

En la práctica, puede no ser sencillo satisfacer exactamente las restricciones sobre las varianzas de los estimadores \bar{y}_{st}^j . Una aproximación más razonable, es minimizar el costo total sujeto a la condición de que $V(\bar{y}_{st}^j)_c \leq V_o^j$, $j = 1, 2, \dots, G$.

Es decir,

$$\min_{\mathbf{n}} \sum_{h=1}^L c_h n_h$$

sujeto a

$$V(\bar{y}_{st}^j)_c \leq V_o^j,$$

Sea

$$\frac{1}{n_h} - \frac{1}{N_h} = \xi_h.$$

Entonces nuestro problema es elegir n_h , $h = 1, 2, \dots, L$ tal que

$$\begin{aligned} & \min_{\mathbf{n}} \sum_{h=1}^L \frac{c_h N_h}{N_h \xi_h + 1} \\ & \text{sujeto a} \\ & \sum_{h=1}^L W_h^2 S_{hj}^2 \xi_h \leq V_o^j \quad j = 1, 2, \dots, G \\ & \text{y} \quad 0 \leq \xi_h \leq 1 - \frac{1}{N_h} \quad h = 1, 2, \dots, L \end{aligned} \tag{14}$$

o equivalentemente, elegir n_h tal que

$$\begin{aligned} & \max_{\mathbf{n}} \left[- \sum_{h=1}^L \frac{c_h N_h}{N_h \xi_h + 1} \right] \\ & \text{sujeto a} \\ & \sum_{h=1}^L W_h^2 S_{hj}^2 \xi_h \leq V_o^j \quad j = 1, 2, \dots, G \\ & \text{y} \quad 0 \leq \xi_h \leq 1 - \frac{1}{N_h} \quad h = 1, 2, \dots, L. \end{aligned} \tag{15}$$

La función objetivo (15) es no lineal y convexa, y las restricciones son lineales en los ξ 's. Dado que la función objetivo es convexa, si una solución factible $(\xi_1, \xi_2, \dots, \xi_L)$ existe, una solución óptima también existirá, vea Arthanari y Dodge (1981) y Sukhatme *et al.* (1984).

Ejemplo (continuación)

Para ilustrar lo anterior se volverán a tomar los datos del Cuadro 2.1. El área cultivada estimada de arroz y el área total cultivada estimada son 409,999.5 y 1,119,149 respectivamente. Tomando los valores estimados como valores verdaderos y asumiendo que el costo por unidad de enumeración es el mismo para todos los

estratos, el problema de determinar n_h se reduce a

$$\max_{\mathbf{n}} \left[- \sum_{h=1}^4 n_h \right] = - \sum_{h=1}^4 \frac{N_h}{N_h \xi_h + 1}$$

sujeto a

$$\sum_{h=1}^4 W_h^2 S_{hj}^2 \xi_h \leq (0.05 \bar{Y}^1)^2,$$

$$\sum_{h=1}^4 W_h^2 S_{hj}^2 \xi_h \leq (0.05 \bar{Y}^2)^2$$

$$\text{y} \quad 0 \leq \xi_h \leq 1 - \frac{1}{N_h}.$$

Realizando las operaciones puede verse que los valores óptimos para n_h son $n_1 = 88$, $n_2 = 43$, $n_3 = 61$, $n_4 = 190$. Se puede observar que cuando se decide aplicar una asignación óptima por minimización de costos, el tamaño de la muestra en cada estrato cambia, aunque para este ejemplo, el tamaño de muestra en los estratos 1 y 4 sigue siendo mayor, como en el caso de las demás asignaciones.

III. OPTIMIZACIÓN MULTI OBJETIVO

Introducción

En los problemas de la vida común continuamente nos topamos con que estos problemas tienen cierta complejidad debido a que aparecen varios objetivos comúnmente múltiples, por ejemplo una familia que planea adquirir un coche tendrá en cuenta, entre una enorme gama de alternativas que ofrece el mercado, sus objetivos, su precio, velocidad máxima, seguridad, confort, capacidad, gasto por km, reparaciones futuras, vida útil, color, etc. esto aunado a que probablemente el padre de familia no sea el decisor único y tenga que consultar la opinión de varios integrantes de su familia. Considerando esta cantidad de objetivos se plantean las técnicas de optimización multiobjetivo, Ríos, Ríos Insua y Ríos Insua (1989).

Nuestro problema consistirá en minimizar la función vectorial $\mathbb{Z}(\mathbf{x})$, es decir

$$\begin{aligned} \min_{\mathbf{x}} \mathbb{Z}(\mathbf{x}) \\ \text{sujeto a} \\ \mathbb{G}(\mathbf{x}) \leq 0 \\ \mathbf{x} \geq 0, \end{aligned} \tag{16}$$

donde $Z : \mathbb{R}^n \rightarrow \mathbb{R}^r$, es decir, $Z(\mathbf{x}) = \begin{bmatrix} Z_1(\mathbf{x}) \\ \vdots \\ Z_r(\mathbf{x}) \end{bmatrix}$ y $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$, tal que $G(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{bmatrix}$.

Dependiendo de la cantidad de información que pueda suministrar el decisor se clasificará en tres el problema:

1. Información nula.
2. Información parcial.
3. Información completa.



Bajo la clasificación de *información completa* el planteamiento es a través de una función de valor $v : \mathcal{X} \rightarrow \mathbb{R}$ esto es

$$\min_{\mathbf{x} \in \mathcal{X}} v(Z(\mathbf{x})).$$

Donde si \mathcal{X} es continuo, el problema es un problema de programación matemática clásico que entra dentro de los “Métodos de optimización multiobjetivo” y si \mathcal{X} es discreto el problema es llamado “Análisis de decisión multiatributo”.

En seguida se proponen cuatro métodos para solucionar estos problemas. Los primeros dos métodos, es decir, el método de la función de valor y el método le-

xicográfico son propuestos bajo la clasificación de información completa. El método que se propone bajo información parcial es el método de las ε -restricciones y el método basado en distancias es el que se lleva a cabo cuando la información es nula.

Función de valor

En este caso el decisor es capaz de construir una función de valor $v : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$\mathbb{Z} \succ \mathbb{Z}' \Leftrightarrow v(\mathbb{Z}) < v(\mathbb{Z}').$$

Esto es que \mathbb{Z} es preferible de \mathbb{Z}' si y sólo si $v(\mathbb{Z})$ es menor que $v(\mathbb{Z}')$.

Con esto el problema de optimización es reducido a un problema de programación clásico:

$$\min_{\mathbf{x} \in \mathcal{X}} v(\mathbb{Z}).$$

Uno de los métodos de la función de valor más frecuentemente empleado es el “método de las ponderaciones”, el cual se define como

$$\min_{\mathbf{x} \in \mathcal{X}} g(\mathbf{x})$$

sujeto a

$$\sum_{j=1}^r \lambda_j = 1, \quad \lambda_j \geq 0 \quad \forall \quad j = 1, 2, \dots, r,$$

donde $g(\mathbf{x}) = \sum_{j=1}^r \lambda_j Z_j(\mathbf{x})$.

Ejemplo

Se desea encontrar

$$\min_{\mathbf{x} \in \mathcal{X}} \begin{bmatrix} Z_1(\mathbf{x}) \\ Z_2(\mathbf{x}) \end{bmatrix}, \quad \mathbf{x} \in \mathbb{R}^2,$$

donde $Z_1(x_1, x_2) = -2x_1 + x_2$, $Z_2(x_1, x_2) = x_1 - 3x_2$ y

$$\mathcal{X} = \left\{ \begin{array}{l} \left(\begin{array}{l} x_1 \\ x_2 \end{array} \right) \left| \begin{array}{l} x_1 + 2x_2 \leq 11 \\ -x_1 + 3x_2 \leq 9 \\ x_1 \leq 5 \\ x_1, x_2 \geq 0 \end{array} \right. \end{array} \right\}.$$

Bajo este criterio la solución está dada al minimizar el problema

$$\min_{\mathbf{x}} \lambda_1(-2x_1 + x_2) + \lambda_2(x_1 - 3x_2)$$

sujeto a

$$x_1 + 2x_2 \leq 11$$

$$-x_1 + 3x_2 \leq 9$$

$$x_1 \leq 5$$

$$x_1, x_2 \geq 0.$$

Método lexicográfico.

Método de optimización secuencial es otro nombre con el que se conoce a este método. Aquí se requiere que el decisor ordene por importancia los objetivos que habrán de considerarse en el problema. La solución bajo este criterio, se define como aquella que maximiza el mayor número de objetivos posibles, comenzando por el de más importancia y descendiendo en el orden jerárquico establecido.

Suponga que dado el problema de optimización multiobjetivo

$$\min_{\mathbf{x} \in \mathcal{X}} \mathbb{Z}(\mathbf{x}),$$

donde $\mathbb{Z}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^r$, es decir, $\mathbb{Z}(\mathbf{x}) = \begin{bmatrix} Z_1(\mathbf{x}) \\ \vdots \\ Z_r(\mathbf{x}) \end{bmatrix}$, y \mathcal{X} pertenece a los continuos.

El orden de importancia establecido entre las funciones objetivos viene dado por la ordenación Z_{i_1}, \dots, Z_{i_r} . El método empieza por resolver el problema

$$\min_{\mathbf{x} \in \mathcal{X}} Z_{i_1}(\mathbf{x}),$$

si a_1 es un óptimo de Z_{i_1} , en la segunda etapa se resuelve el problema

$$\min_{\mathbf{x} \in \mathcal{X}} Z_{i_2}(\mathbf{x})$$

sujeto a

$$Z_{i_1}(\mathbf{x}) = a_1,$$

si el óptimo del problema anterior es a_2 , en la tercer etapa el óptimo a resolver es

$$\min_{\mathbf{x} \in \mathcal{X}} Z_{i_3}(\mathbf{x})$$

sujeto a

$$Z_{i_j}(\mathbf{x}) = a_j, \quad j = 1, 2.$$

De tal forma que en la etapa k se resuelve el problema

$$\min_{\mathbf{x} \in \mathcal{X}} Z_{i_k}(\mathbf{x})$$

sujeto a

(17)

$$Z_{i_j}(x) = a_j, \quad j = 1, 2, \dots, k - 1.$$

Entonces

1. Si (17) tiene solución única o bien $k = r$ ésta será la solución al problema multiobjetivo.
2. Si (17) no está acotada, no hay solución al problema multiobjetivo.
3. En otro caso, ir al paso $k + 1$.

Ejemplo

$$\max \begin{bmatrix} x_1 \\ -x_1 + 3x_2 \end{bmatrix}$$

sujeto a.

$$x_1 + 2x_2 \leq 11$$

$$-x_1 + 3x_2 \leq 9$$

$$x_1 \leq 5$$

$$x_1, x_2 \geq 0.$$

Solución. Suponga que la jerarquización es $Z_1(x)$, $Z_2(x)$. Así recordando que $\max f(x) = \min (-f(x))$, se tiene

Paso (1)

$$\min -x_1$$

sujeto a

$$x_1 + 2x_2 \leq 11$$

$$-x_1 + 3x_2 \leq 9$$

$$x_1 \leq 5$$

$$x_1, x_2 \geq 0.$$

Esto da el siguiente resultado

$$\mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \quad a_1 = 5.$$

Paso (2)

$$\min x_1 - 3x_2$$

sujeto a

$$x_1 + 2x_2 \leq 11$$

$$-x_1 + 3x_2 \leq 9$$

$$x_1 = 5$$

$$x_1, x_2 \geq 0.$$

Esto da el siguiente resultado

$$\mathbf{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \quad a_2 = 4.$$

Por lo tanto la solución máxima es

$$\mathbf{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \text{ con } \mathbb{Z}(\mathbf{x}) = \begin{bmatrix} 5 \\ 4 \end{bmatrix}.$$

Método de las ε -restricciones.

Dado el problema

$$\min_{\mathbf{x} \in \mathcal{X}} \mathbb{Z}(\mathbf{x}),$$

donde $\mathbb{Z}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^r$, es decir, $\mathbb{Z}(\mathbf{x}) = \begin{bmatrix} Z_1(\mathbf{x}) \\ \vdots \\ Z_r(\mathbf{x}) \end{bmatrix}$, y \mathcal{X} pertenece a los continuos,

este método supone que hay un objetivo Z_k , más preferido que los demás, y que

se pueden dar $r - 1$ números ε_j , $j \neq k$, $j = 1, 2, \dots, r$ que corresponden a cotas superiores de estos $r - 1$ objetivos restantes. De esta forma, se plantea el problema uniobjetivo

$$\begin{aligned} & \min_{\mathbf{x} \in \mathcal{X}} Z_k(\mathbf{x}) \\ & \text{sujeto a} \\ & Z_j(\mathbf{x}) \leq \varepsilon_j, \quad j = 1, 2, \dots, r \quad j \neq k. \end{aligned}$$

Ejemplo. (Continuación)

Para el ejemplo que se trató en la sección “Función de Valor”, suponga que Z_1 es el objetivo de mayor importancia y $\varepsilon_2 = 5$. El problema a resolver es:

$$\begin{aligned} & \min_{\mathbf{x}} -2x_1 + x_2 \\ & \text{sujeto a} \\ & -x_1 + 3x_2 \geq 5 \\ & x_1 + 2x_2 \leq 11 \\ & -x_1 + 3x_2 \leq 9 \\ & x_1 \leq 5 \\ & x_1, x_2 \geq 0. \end{aligned}$$

Método basado en distancias.

En múltiples situaciones prácticas un decisor, en base al nivel de logro o alcance de las desviaciones de los objetivos en comparación con valores fijados con anticipación, define el concepto de mejor o preferido.

Una forma en que esto se mide es por medio de una distancia que, de alguna manera refleja que la solución elegida es aquella que menos se aleje de las

metas deseadas.

Para resolver este problema se especifica, para cada objetivo Z_i , un valor \widehat{Z}_1 que se considera como idealmente bueno, de tal forma que el vector

$$\widehat{\mathbf{Z}} = \begin{bmatrix} \widehat{Z}_1 \\ \vdots \\ \widehat{Z}_r \end{bmatrix},$$

es el conjunto de metas ideal que deben alcanzar los objetivos simultáneamente. Si $\widehat{\mathbf{Z}}$ es factible, el problema en tal caso consiste en resolver el sistema de ecuaciones

$$Z_1(\mathbf{x}) = \widehat{Z}_i, \quad i = 1, 2, \dots, r.$$

Pero, si $\widehat{\mathbf{Z}}$ no es factible, debemos determinar un punto que sea lo más “próximo” posible al vector de metas, $\widehat{\mathbf{Z}}$. Este enfoque se puede plantear como un problema de minimizar una distancia d entre los puntos de un espacio r -dimensional.

El problema será

$$\min_{\mathbf{x} \in \mathcal{X}} d(\mathbf{Z}(\mathbf{x}), \widehat{\mathbf{Z}}).$$

Ya que evidentemente la distancia no es única, la solución del problema anterior dependerá de la distancia que se tome, así como del punto $\widehat{\mathbf{Z}}$ fijado por el decisor.

Definición 4. Una pseudodistancia d en \mathbb{Z} , es una función

$$d : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{R},$$

tal que para todo $Z, Z', Z'' \in \mathbb{Z}$, cumple las siguientes propiedades

1. $d(Z, Z') \geq 0 \quad (Z \neq Z')$.
2. $d(Z, Z) = 0$.
3. $d(Z, Z'') \leq d(Z, Z') + d(Z', Z'')$.
4. $d(Z, Z') = d(Z', Z)$.

Si además cumple

$$d(Z, Z') = 0 \Rightarrow Z = Z'.$$

La función d es llamada distancia.

Fijado el vector de metas \widehat{Z} y dado $Z \in \mathbb{Z}$, se llama " pseudodistancia radial" del par ordenado (Z, \widehat{Z}) a la pseudodistancia $d(Z, \widehat{Z})$ en $Z \times \{\widehat{Z}\}$.

Se pueden hacer diferentes elecciones de d , por ejemplo

1. L_q -ponderada

$$d_q^*(\mathbb{Z}(\mathbf{x}), \widehat{Z}) = \left[\sum_{j=1}^r \lambda_j |Z_j(\mathbf{x}) - \widehat{Z}_j|^q \right]^{\frac{1}{q}}, \quad 1 \leq q \leq \infty, \quad \lambda_j \geq 0,$$

- 2.

$$d_q(\mathbb{Z}(\mathbf{x}), \widehat{Z}) = \left[\sum_{j=1}^r \lambda_j (Z_j(\mathbf{x}) - \widehat{Z}_j)^q \right]^{\frac{1}{q}}, \quad 1 \leq q \leq \infty, \quad \lambda_j \geq 0,$$

ó equivalentemente

$$d'_q(\mathbb{Z}(\mathbf{x}), \widehat{Z}) = \left[\sum_{j=1}^r \lambda_j (Z_j(\mathbf{x}) - \widehat{Z}_j)^q \right].$$

Estas distancias contienen varios casos particulares. Por ejemplo si \widehat{Z} es el punto ideal Z^* donde Z_j^* es el óptimo del problema

$$\min_{\mathbf{x} \in \mathcal{X}} Z_j(\mathbf{x}).$$

De esta forma se dice que \mathbf{x}_q^* es una solución compromiso respecto de q .

Tres distancias de inters son d_1 , d_2 y d_∞

Con d_1 , la solución compromiso \mathbf{x}_1^* se obtiene de

$$\min_{\mathbf{x} \in \mathcal{X}} \sum_{j=1}^r (|Z_j(\mathbf{x}) - Z_j^*|),$$

Con d_∞ sólo se tiene en cuenta la máxima desviación, luego la solución compromiso \mathbf{x}_∞^* se obtiene de

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{j=1,2,\dots,r} (Z_j(\mathbf{x}) - Z_j^*).$$

Finalmente, para d_2 , la solución compromiso \mathbf{x}_2^* se obtendrá de

$$\min_{\mathbf{x} \in \mathcal{X}} \sum_{j=1}^r (Z_j(\mathbf{x}) - Z_j^*)^2,$$

que corresponde con la distancia euclídea entre \mathbb{Z}^* y $\mathbb{Z}(\mathbf{x}_2^*)$.

Ejemplo

$$\max_{\mathbf{x}} \mathbb{Z}(\mathbf{x}) = \begin{bmatrix} 5x_1 - 3x_2 \\ -x_1 + 5x_2 \end{bmatrix}$$

sujeto a.

$$x_1 + 3x_2 \leq 15$$

$$2x_1 + 3x_2 \leq 18$$

$$x_1 - x_2 \leq 4$$

$$x_1, x_2 \geq 0.$$

Suponga, además que $\lambda_1 = \lambda_2 = 1$, determine el punto compromiso.

Recordando nuevamente que $\max f(x) = \min(-f(x))$, se resuelven los siguientes problemas

$$\begin{array}{ll} \min_{\mathbf{x}} -5x_1 + 3x_2 & \min_{\mathbf{x}} x_1 - 5x_2 \\ \text{sujeto a} & \text{sujeto a} \\ x_1 + 3x_2 \leq 15 & x_1 + 3x_2 \leq 15 \\ 2x_1 + 3x_2 \leq 18 & 2x_1 + 3x_2 \leq 18 \\ x_1 - x_2 \leq 4 & x_1 - x_2 \leq 4 \\ x_1, x_2 \geq 0. & x_1, x_2 \geq 0. \end{array}$$

Se tiene el punto ideal $Z^* = \begin{bmatrix} Z_1^* \\ Z_2^* \end{bmatrix} = \begin{bmatrix} 24 \\ 25 \end{bmatrix}$, de donde los resultados para las tres distancias de interés se presentan en el Cuadro 3.1.

CUADRO 3.1: Resultado de las distancias

q	x_q^*	$Z(x_q^*)$
∞	$\begin{pmatrix} 3.19 \\ 3.87 \end{pmatrix}$	$\begin{pmatrix} 11.03 \\ 12.03 \end{pmatrix}$
1	$\begin{pmatrix} 6 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 24 \\ 4 \end{pmatrix}$
2	$\begin{pmatrix} 4.66 \\ 2.89 \end{pmatrix}$	$\begin{pmatrix} 14.61 \\ 9.81 \end{pmatrix}$

donde x_q^* es el punto crítico y $Z(x_q^*)$ es el óptimo de las funciones.

IV. ASIGNACIÓN ÓPTIMA A TRAVÉS DE OPTIMIZACIÓN MULTIOBJETIVO

Introducción

En la experimentación científica es muy común determinar el tamaño de muestra para un problema con un costo fijo, ya que en muchos trabajos el presupuesto es asignado anticipadamente. Así, el problema se soluciona minimizando la varianza $V(\bar{y}_{st})$, sujeta a ese costo fijo C , es decir, el problema se convierte en un problema de optimización donde la función objetivo es precisamente la varianza y la restricción es la función de costo.

$$\begin{aligned} \min_{\mathbf{n}} V(\bar{y}_{st}) \\ \text{sujeto a} \\ \mathbf{c}' \mathbf{n} + c_0 = C, \end{aligned}$$

o alternativamente, cuando lo importante es obtener cierta precisión V_0 y ésta es asignada a la varianza $V(\bar{y}_{st})$, el problema ahora consiste en minimizar la función de costos. Este caso, de igual manera, es un problema de optimización.

$$\begin{aligned} \min_{\mathbf{n}} \mathbf{c}' \mathbf{n} + c_0 \\ \text{sujeto a} \\ V(\bar{y}_{st}) \leq V_0. \end{aligned}$$

Ahora nuestro problema se extiende al campo multivariado, entonces para el caso donde la varianza es fija, se tendrá el siguiente problema

$$\min_{\mathbf{n}} \mathbf{c}' \mathbf{n} + c_0$$

sujeto a

$$V(\bar{y}_{st}^j) \leq V_0^j, \quad j = 1, 2, \dots, G,$$

cuyo problema fue abordado en la sección “Asignación minimizando costos” del Capítulo II.

Se debe notar que:

1. n_h pertenece a los naturales y no a los reales es decir, $n_h \in \mathbb{N}$ y $n_h \notin \mathbb{R}$, lo que convierte el problema en un problema de programación no lineal multiobjetivo de enteros.
2. N_h debe ser mayor que n_h , además n_h debe ser mayor o igual 2; es decir, la muestra no debe rebasar la población y la muestra debe de ser al menos igual a 2 para así poder estimar las varianzas.
3. $V(\bar{y}_{st}^j)$ es definida por (6) donde se emplean las varianzas poblacionales S_h^2 , $h = 1, 2, \dots, L$ que generalmente no se conocen, por lo tanto se sustituirán por las varianzas muestrales s_h^2 , $h = 1, 2, \dots, L$ y entonces $V(\bar{y}_{st}^j)$ se sustituirá por la varianza estimada $\widehat{V}(\bar{y}_{st}^j)$, la cual está dada por

$$\widehat{V}(\bar{y}_{st}^j) = \sum_{h=1}^L \frac{W_h^2 s_{hj}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{hj}^2}{N}.$$

Así el problema a resolver es

$$\begin{aligned} & \min_{\mathbf{n}} \mathbf{c}' \mathbf{n} + c_0 \\ & \text{sujeto a} \\ & \widehat{V}(\bar{y}_{st}^j) \leq V_0^j, \quad j = 1, 2, \dots, G \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}. \end{aligned}$$

Cuya solución se puede obtener a partir de cualquiera de los métodos de optimización clásica de enteros, vea la sección “Asignación minimizando costos” del Capítulo II.

Similarmente, para el caso donde los costos son fijos

$$\begin{aligned} \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}) = \min_{\mathbf{n}} & \begin{pmatrix} \widehat{V}(\bar{y}_{st}^1) \\ \vdots \\ \widehat{V}(\bar{y}_{st}^G) \end{pmatrix} \\ & \text{sujeto a} \\ & c_1 n_1 + c_2 n_2 + \dots + c_L n_L + c_0 = C \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}. \end{aligned} \tag{18}$$

Así (18) define el programa de optimización multiobjetivo no lineal de enteros.

En la siguiente sección se obtiene la solución del problema (18) a través de los cuatro métodos de optimización multiobjetivo expuestos en el Capítulo anterior.

1. Función de valor.

2. Método lexicográfico.
3. Método de las ε -restricciones.
4. Método basado en distancias.

Como se verá a continuación los métodos propuestos en la literatura para resolver el programa de optimización (18) descritos en la sección *Muestreo Estratificado Multivariado* son casos particulares del método de la *función de valor*. Estos métodos son:

1. **Asignación compromiso minimizando la pérdida total relativa.** Donde la función objetivo es la siguiente

$$\min_{\mathbf{n}} \sum_{j=1}^G \left[\alpha_j \left(\frac{V(\bar{y}_{st}^j)_C - V(\bar{y}_{st}^j)_{Opt}}{V(\bar{y}_{st}^j)_{Opt}} \right) \right]$$

sujeto a

$$\sum_{h=1}^L c_h n_h \leq C.$$

2. **Asignación compromiso minimizando la traza.** Con función objetivo siguiente

$$\min_{\mathbf{n}} \sum_{j=1}^G V(\bar{y}_{st}^j)_C$$

sujeto a

$$\sum_{h=1}^L c_h n_h + c_0 = C.$$

Función de valor

Este método tiene su aplicación en experimentos donde la información de las características evaluadas es completa, es decir, cuando se conoce totalmente la

importancia que cada una tiene. Por ejemplo, para la función de valor aplicando el método de las ponderaciones, se debe conocer tan perfectamente su jerarquía que el decisor es capaz de poder otorgar una ponderación a cada una ellas.

Se parte del problema inicial (18).

Bajo la técnica de función de valor se tiene que el problema a resolver es

$$\begin{aligned}
 & \min_{\mathbf{n}} \mathbf{v}(\widehat{V}(\bar{\mathbf{y}}_{st})), \\
 & j = 1, 2, \dots, G \\
 & \text{sujeto a} \\
 & \sum_{h=1}^L c_h n_h + c_0 = C \\
 & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\
 & n_h \in \mathbb{N}.
 \end{aligned} \tag{19}$$

En particular aplicando el método de las ponderaciones, (19) toma la forma

$$\begin{aligned}
 & \min_{\mathbf{n}} \sum_{j=1}^G \lambda_j \widehat{V}(\bar{\mathbf{y}}_{st}^j), \\
 & \text{sujeto a} \\
 & \sum_{h=1}^L c_h n_h + c_0 = C \\
 & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\
 & n_h \in \mathbb{N}.
 \end{aligned}$$

tal que $\sum_{j=1}^G \lambda_j = 1, \quad \lambda_j \geq 0 \quad \forall \quad j = 1, 2, \dots, G.$

Este es sin duda el método más explorado por los investigadores. Esto puede apreciarse en la literatura citada en el presente trabajo, Sukhatme *et al.* (1984). Todos los procedimientos de asignación descritos en la sección *Muestreo*

Estratificado Multivariado del Capítulo II son realizados por medio de una función de valor. El hecho de ser un método recurrido se debe a que la *función de valor* no es única, puede elegirse la función que más convenga a los investigadores. Este método además se aplica en experimentos donde se tienen antecedentes que nos ayudan a ponderar las características evaluadas, en otras palabras, el método función de valor se utiliza para investigaciones frecuentes donde con el paso de tiempo se han obtenido resultados que ayudan a una mejor inferencia en experimentos futuros por medio de la adecuada ponderación. Cuando se tienen estos antecedentes o información completa para jerarquizar, este procedimiento es el más adecuado. Es por esta razón que en diversos campos de la ciencia donde una investigación lleva tiempo realizándose por la importancia que representa (por ejemplo, investigaciones médicas de laboratorios farmacéuticos, encuestas de opinión, etc) el presente método es el más apropiado.

Método lexicográfico.

Este método al igual que el anterior, requiere de información completa del fenómeno que ayude a jerarquizar por importancia a las características evaluadas, en este caso medidas por medio de sus varianzas. A diferencia del método de la función de valor, no es necesario conocer que ponderación tiene cada una, sólo con saber el orden de importancia que representan para la obtención de la muestra, es suficiente. Esto en la práctica es muy útil, ya que en diversas ocasiones no se contará con el valor del peso que tiene cada característica evaluada, probablemente los decisores solamente conozcan el orden en que cada característica influye en su investigación.

En este caso el problema a optimizar es nuevamente (18).

El decisor debe ordenar la varianzas comenzando por aquella varianza de la característica más importante y descendiendo en orden de importancia, obteniéndose $\widehat{V}(\bar{y}_{st}^{i_1}), \widehat{V}(\bar{y}_{st}^{i_2}), \dots, \widehat{V}(\bar{y}_{st}^{i_G})$, donde i_1, \dots, i_G es una permutación con el orden descendiente deseado del conjunto de súper índices $1, 2, \dots, G$. Luego el primer problema que debe resolverse es:

$$\begin{aligned} & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^{i_1}) \\ & \text{sujeto a} \\ & \sum_{h=1}^L c_h n_h + c_0 = C \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}. \end{aligned} \tag{20}$$

Si el mínimo del problema (20) es v_1 , en la siguiente etapa se deberá resolver el problema

$$\begin{aligned} & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^{i_2}) \\ & \text{sujeto a} \\ & \sum_{h=1}^L \frac{W_h^2 s_{h1}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{h1}^2}{N} = v_1 \\ & \sum_{h=1}^L c_h n_h + c_0 = C \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}. \end{aligned} \tag{21}$$

Ahora sea v_2 el mínimo del problema (21), en la tercera etapa se resolverá

el problema

$$\begin{aligned}
& \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^{i3}) \\
& \text{sujeto a} \\
& \sum_{h=1}^L \frac{W_h^2 s_{h1}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{h1}^2}{N} = v_1 \\
& \sum_{h=1}^L \frac{W_h^2 s_{h2}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{h2}^2}{N} = v_2 \\
& \sum_{h=1}^L c_h n_h + c_0 = C \\
& 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\
& n_h \in \mathbb{N}.
\end{aligned}$$

De tal forma que se llegará a la etapa G , donde se deberá resolver el siguiente problema.

$$\begin{aligned}
& \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^{iG}) \\
& \text{sujeto a} \\
& \sum_{h=1}^L \frac{W_h^2 s_{h1}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{h1}^2}{N} = v_1 \\
& \sum_{h=1}^L \frac{W_h^2 s_{h2}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{h2}^2}{N} = v_2 \\
& \sum_{h=1}^L \frac{W_h^2 s_{h3}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{h3}^2}{N} = v_3 \\
& \vdots \\
& \sum_{h=1}^L \frac{W_h^2 s_{hG-1}^2}{n_h} - \sum_{h=1}^L \frac{W_h s_{h2G-1}^2}{N} = v_{G-1} \\
& \sum_{h=1}^L c_h n_h + c_0 = C \\
& 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\
& n_h \in \mathbb{N}.
\end{aligned}$$

Así el vector obtenido en esta etapa será la solución óptima para el problema.

Método de las ε -restricciones.

Este es un método para resolver el problema cuando se tiene información parcial. Para poder realizar este método se necesita solamente identificar la característica más importante. Este método es sumamente útil en investigaciones donde los precedentes han podido reconocer cual es la característica que más influye para obtención de la muestra y los límites permitidos para el resto de ellas.

Nuevamente partiendo del problema (18) y asumiendo que la característica más importante en el estudio es la k -ésima, $k \in \{1, 2, \dots, G\}$, el problema se puede replantear bajo esta técnica como:

$$\begin{aligned}
 & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^k) \\
 & \text{sujeto a} \\
 & \widehat{V}(\bar{y}_{st}^r) \leq v_r, \quad r \neq k, \quad r = 1, 2, \dots, G \\
 & \sum_{h=1}^L c_h n_h + c_0 = C \\
 & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\
 & n_h \in \mathbb{N},
 \end{aligned} \tag{22}$$

donde en este caso, v_r es una cota preestablecida para cada una de las $G-1$ varianzas restantes que aparecen como restricciones. Para fines prácticos, estos valores v_r se pueden tomar como el límite superior del intervalo de confianza para cada varianza

o alternativamente, se pueden definir como los mínimos de los problemas siguientes:

$$\begin{aligned} & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^r) \\ & \text{sujeto a} \\ & \sum_{h=1}^L c_h n_h + c_0 = C \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}, r = 1, 2, \dots, G, \quad r \neq k. \end{aligned}$$

Se hace hincapié en que la elección de la k característica y las cotas inferiores v_r , representan preferencias subjetivas del decisor, de forma que si no existiera solución para el problema (22), vendría a significar que las cotas inferiores se han tomado demasiado altas y al menos una debería modificarse, Ríos, Ríos Insua y Ríos Insua (1989).

Método basado en distancias.

En múltiples ocasiones, al realizar alguna investigación se tendrá el problema de no tener antecedentes que ayuden a resolverlo, o el investigador puede enfrentarse a la dificultad de decidir cuál de las características que se evalúan es la más importante, en casos como este el método presentado en esta sección es de mucha ayuda. No será necesario ningún antecedente, ya que para resolver el problema por este método sólo se deberá contar con un vector de metas “ideal” determinado con la nula información que el problema expone.

Luego partiendo del problema (18), este método logra obtener los valores óptimos minimizando la distancia entre el óptimo y el vector de metas si-

multáneamente. Sea v_j el punto o meta ideal para el objetivo $\widehat{V}(\bar{y}_{st}^j)$, $j = 1, \dots, G$, i.e. el vector de metas \mathbb{V} esta dado como

$$\mathbb{V} = \begin{pmatrix} v_1 \\ \vdots \\ v_G \end{pmatrix}.$$

Una gran ventaja de este método radica en que dicho vector de metas \mathbb{V} puede ser calculado sin información adicional. Esto puede ser llevado a cabo, minimizando por separado cada objetivo $\widehat{V}(\bar{y}_{st}^j)$, $j = 1, \dots, G$ de tal forma que el vector \mathbb{V} quedará definido como el vector de sus mínimos individuales, obtenido al resolver los siguientes G programas de minimización:

$$\begin{array}{ccc} \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^1), & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^2), \quad \dots, & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^G) \\ \text{sujeto a} & \text{sujeto a} & \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C, & \sum_{h=1}^L c_h n_h + c_0 = C, \quad \dots, & \sum_{h=1}^L c_h n_h + c_0 = C, \\ 2 \leq n_h \leq N_h, & 2 \leq n_h \leq N_h, & 2 \leq n_h \leq N_h, \\ h = 1, 2, \dots, L & h = 1, 2, \dots, L & h = 1, 2, \dots, L \\ n_h \in \mathbb{N}. & n_h \in \mathbb{N}. & n_h \in \mathbb{N}. \end{array}$$

Así se obtendrá el vector de mínimos

$$\mathbb{V} = \begin{pmatrix} v_1 \\ \vdots \\ v_G \end{pmatrix}.$$

Una vez calculados los valores de este vector se procede a plantear el pro-

blema a optimizar con la nueva función objetivo. La cual será

$$\begin{aligned} \min_{\mathbf{n}} d(V(\bar{y}_{st}^j), \mathbb{V}) \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ n_h \in \mathbb{N}. \end{aligned}$$

Lo cual como se vió en el Capitulo III, la función d corresponde a una norma ponderada, y se va a considerar más generalmente la norma L_q -ponderada, así el problema a optimizar tomará la forma

$$\begin{aligned} \min_{\mathbf{n}} \left[\sum_{j=1}^G \lambda_j |V(\bar{y}_{st}^j) - v_j|^q \right]^{\frac{1}{q}} \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ n_h \in \mathbb{N}, \end{aligned}$$

con $1 \leq q \leq \infty$ y $\lambda_j \geq 0$ que es el peso o prioridad que se le da a cada objetivo j .

Para ejemplificar esta técnica se tomará $\lambda_j = 1$, es decir, se dará la misma prioridad a cada característica, y se usarán los casos particulares $q = 1$, $q = 2$ y $q = \infty$.

Así con $q = 1$, se tendrá el problema siguiente

$$\begin{aligned} \min_{\mathbf{n}} \left[\sum_{j=1}^G |V(\bar{y}_{st}^j) - v_j| \right] \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ n_h \in \mathbb{N}, \end{aligned}$$

como v_j son constantes para todo $j = 1, \dots, G$, el problema se reduce a

$$\begin{aligned} \min_{\mathbf{n}} \sum_{j=1}^G V(\bar{y}_{st}^j) \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ n_h \in \mathbb{N}. \end{aligned}$$

Con $q = \infty$ debe tenerse en cuenta sólo la máxima desviación, así el problema a optimizar será

$$\begin{aligned} \min_{\mathbf{n}} \max_{j=1,2,\dots,G} [V(\bar{y}_{st}^j) - v_j] \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ n_h \in \mathbb{N}. \end{aligned}$$

y para $q = 2$ el problema será

$$\begin{aligned} \min_{\mathbf{n}} \left[\sum_{j=1}^G [V(\bar{y}_{st}^j) - v_j]^2 \right]^{1/2} \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ n_h \in \mathbb{N}. \end{aligned}$$

Alternativamente, otra distancia propuesta en la literatura por Khuri y Cornell (1987), es

$$\begin{aligned} \min_{\mathbf{n}} \sum_{j=1}^G \left[\frac{(V(\bar{y}_{st}^j) - v_j)^2}{v_j^2} \right] \\ \text{sujeto a} \\ \sum_{h=1}^L c_h n_h + c_0 = C \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ n_h \in \mathbb{N}. \end{aligned}$$

Nota. Se puede observar que en todos estos métodos de optimización se ha utilizado la restricción sobre costos, $\sum_{h=1}^L c_h n_h + c_0 = C$. Pero en diferentes ocasiones, las restricciones no son dadas sobre los costos, sino por la disponibilidad de horas hombre para realizar la encuesta o simplemente por el tiempo total para llevar a cabo la misma. Estas limitantes pueden ser establecidas a través de la siguiente expresión, ver Arthanari y Dodge (1981):

$$\sum_{h=1}^L n_h = n.$$

Técnica de Dalenius

Esta técnica, como se mencionó en la sección “Asignación compromiso minimizando la traza”, del Capítulo II, fue propuesta por Dalenius (1957), la cual se basa en una aproximación por medio del determinante de la matriz de varianza-covarianza de la media \bar{y}_{st}^j , $j = 1, 2, \dots, G$, Sukhatme *et al.* (1984). Observe que en el problema (18) se ha asumido que las covarianzas entre las diferentes características son cero, lo cual si se asume normalidad sería equivalente a establecer que las características son estocásticamente independientes, cuyo hecho no necesariamente se presenta en las aplicaciones. Esta idea motiva a plantear el problema de asignación óptima en muestreo estratificado multivariado como sigue

$$\begin{aligned} & \min_{\mathbf{n}} \Theta \\ & \text{sujeto a} \\ & c_1 n_1 + c_2 n_2 + \dots + c_L n_L + c_0 = C \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}. \end{aligned} \tag{23}$$

donde $\Theta = \text{Cov}(\bar{\mathbf{y}}_{st})$ es la matriz de varianzas-covarianzas del vector $\bar{\mathbf{y}}_{st} = \begin{bmatrix} \bar{y}_{st}^1 \\ \vdots \\ \bar{y}_{st}^G \end{bmatrix}$.

Obviamente la dificultad de plantear de esta forma el problema es como definir que significa el mínimo de una matriz. Sin embargo, si dicho problema es tratado a través del método de la función de valor, fácilmente surgen interpretaciones del mínimo de una matriz. De esta forma, el programa empleando la técnica de la

función de valor está dado por

$$\begin{aligned} & \min_{\mathbf{n}} \mathbf{v}(\Theta) \\ & \text{sujeto a} \\ & \sum_{h=1}^L c_h n_h + c_0 = C \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}. \end{aligned}$$

Note que, si en particular $\mathbf{v}(\Theta) = \text{tr}(\Theta)$, y se da a todas las características la misma ponderación, tal que $\sum_{j=1}^G \lambda_j = 1$, $\lambda_j \geq 0 \quad \forall \quad j = 1, 2, \dots, G$ entonces se obtiene la solución particular tratada en la sección función de valor. La función de valor empleada para el problema (23) tratada por Dalenius, donde define $\mathbf{v}(\Theta) = \det(\Theta) = |\Theta|$, es

$$\begin{aligned} & \min_{\mathbf{n}} |\Theta| \\ & \text{sujeto a} \\ & \sum_{h=1}^L c_h n_h + c_0 = C \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, L \\ & n_h \in \mathbb{N}. \end{aligned}$$

Observe que la traza y el determinante no son las únicas funciones de valor que se pueden emplear. Definiciones alternativas de la función de valor empleadas en otros contextos de la estadística pueden ser

1. *La suma de todos los elementos de la matriz.* $\mathbf{v}(\Theta) = \sum_{k,l=1}^G \theta_{kl}$.
2. $\mathbf{v}(\Theta) = \lambda_1(\Theta)$, donde λ_1 es el máximo eigenvalor de de la matriz de covarianzas Θ .

3. $\mathbf{v}(\Theta) = \lambda_G(\Theta)$, donde λ_G es el mínimo eigenvalor de la matriz de covarianzas Θ .

A continuación se utilizarán los datos del Cuadro 2.1 del ejemplo descrito en el Capítulo II para calcular los tamaños de muestra por medio de los cinco métodos descritos anteriormente. Dado que la restricción para estos métodos puede estar sujeta a costos fijos o sujeta a un tamaño de muestra fijo, se tomará la restricción inicial del problema $\sum_{h=1}^4 n_h = 382$. Los resultados obtenidos por medio de programa de cómputo LINGO, para todas las técnicas, serán mostrados en el Cuadro 4.2.

Ejemplo

Función de Valor

El problema a optimizar, donde se evalúan dos características, es

$$\begin{aligned} \min_{\mathbf{n}} \quad & \begin{pmatrix} \widehat{V}(\bar{y}_{st}^1) \\ \widehat{V}(\bar{y}_{st}^2) \end{pmatrix} \\ \text{sujeto a} \quad & \\ & \sum_{h=1}^4 n_h = 382 \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ & n_h \in \mathbb{N}. \end{aligned} \tag{24}$$

Como debe cumplirse que $\sum_{j=1}^2 \lambda_j = 1$, $\lambda_j \geq 0 \quad \forall \quad j = 1, 2$ se otorgará la misma ponderación a las características evaluadas, $\lambda_1 = 0.5$ y $\lambda_2 = 0.5$. De acuerdo

con lo anterior y utilizando la teoría ya mencionada, el problema de optimización es

$$\begin{aligned} \min_{\mathbf{n}} \quad & \left(0.5 \left(\widehat{V}(\bar{y}_{st}^1) \right) + 0.5 \left(\widehat{V}(\bar{y}_{st}^2) \right) \right) \\ \text{sujeto a} \quad & \\ & \sum_{h=1}^4 n_h = 382 \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ & n_h \in \mathbb{N}. \end{aligned}$$

Método lexicográfico.

Se parte del problema (24). Lo primero que se hace para resolver el problema por medio del método lexicográfico es ordenar las características por orden de importancia, siendo para este caso $\widehat{V}(\bar{y}_{st}^1)$ la más importante, de esta manera la primera etapa de este método es resolver el siguiente programa:

$$\begin{aligned} \min_{\mathbf{n}} \quad & \widehat{V}(\bar{y}_{st}^1) \\ \text{sujeto a} \quad & \\ & \sum_{h=1}^4 n_h = 382 \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ & n_h \in \mathbb{N}. \end{aligned}$$

cuyo mínimo obtenido a través del programa de cómputo LINGO es $v_1 = 23.86$.

En la segunda etapa, y para este caso la etapa final, se tiene el siguiente

programa de optimización:

$$\begin{aligned}
 & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^2) \\
 & \text{sujeto a} \\
 & \widehat{V}(\bar{y}_{st}^1) \leq 23.86 \\
 & \sum_{h=1}^4 n_h = 382 \\
 & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\
 & n_h \in \mathbb{N}.
 \end{aligned}$$

Método de las ε -restricciones.

Bajo esta técnica el problema (24) se resuelve, eligiendo la característica 1 como la prioridad, de la siguiente forma

$$\begin{aligned}
 & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^1) \\
 & \text{sujeto a} \\
 & \widehat{V}(\bar{y}_{st}^2) \leq 92.50 \\
 & \sum_{h=1}^4 n_h = 382 \\
 & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\
 & n_h \in \mathbb{N},
 \end{aligned}$$

donde $v_2 = 92.50$, es la cota superior para la varianza de la característica dos y

es calculada al minimizar el problema siguiente:

$$\begin{aligned} & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^2) \\ & \text{sujeto a} \\ & \sum_{h=1}^4 n_h = 382 \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ & n_h \in \mathbb{N}, \end{aligned}$$

Método basado en distancias.

Partiendo del problema (24), se aplicará el procedimiento para la distancia $q = 1$ (norma del valor absoluto), $q = 2$ (norma euclídea, propuesta por Salukvadze en 1971 para $\lambda_j = 1$) y la distancia propuesta por Khuri y Cornell (1987) que se propondrá al final de esta sección. El vector de comparación se obtiene evaluando por separado las dos características

$$\begin{array}{ll} \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^1) & \min_{\mathbf{n}} \widehat{V}(\bar{y}_{st}^2) \\ \text{sujeto a} & \text{sujeto a} \\ \sum_{h=1}^4 n_h = 382 & \sum_{h=1}^4 n_h = 382 \\ 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ n_h \in \mathbb{N}. & n_h \in \mathbb{N}, \end{array}$$

así se obtiene

$$\mathbb{V} = \begin{pmatrix} 23.86 \\ 92.50 \end{pmatrix}.$$

Con $q = 1$ se tiene el problema

$$\begin{aligned} \min_{\mathbf{n}} \quad & \left(|\widehat{V}(\bar{y}_{st}^1) + \widehat{V}(\bar{y}_{st}^2)| \right) \\ \text{sujeto a} \quad & \\ & \sum_{h=1}^4 n_h = 382 \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ & n_h \in \mathbb{N}. \end{aligned}$$

Con $q = 2$ se tiene

$$\begin{aligned} \min_{\mathbf{n}} \quad & \left[\left(\widehat{V}(\bar{y}_{st}^1) - 23.86 \right)^2 + \left(\widehat{V}(\bar{y}_{st}^2) - 92.50 \right)^2 \right] \\ \text{sujeto a} \quad & \\ & \sum_{h=1}^4 n_h = 382 \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ & n_h \in \mathbb{N}. \end{aligned}$$

Con la distancia propuesta por Khuri y Cornell (1987) el problema a optimizar es el siguiente

$$\begin{aligned} \min_{\mathbf{n}} \quad & \left[\frac{\left(\widehat{V}(\bar{y}_{st}^1) - 23.86 \right)^2}{23.86^2} + \frac{\left(\widehat{V}(\bar{y}_{st}^2) - 92.50 \right)^2}{92.50^2} \right] \\ \text{sujeto a} \quad & \\ & \sum_{h=1}^4 n_h = 382 \\ & 2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4 \\ & n_h \in \mathbb{N}. \end{aligned}$$

Técnica Dalenius

El problema a resolver, dado que son dos características que se evalúan, es

$$\min_{\mathbf{n}} \left| \begin{array}{cc} \widehat{V}(\bar{y}_{st}^1) & \widehat{\text{Cov}}(\bar{y}_{st}^1, \bar{y}_{st}^2) \\ \widehat{\text{Cov}}(\bar{y}_{st}^2, \bar{y}_{st}^1) & \widehat{V}(\bar{y}_{st}^2) \end{array} \right|$$

sujeto a

$$\sum_{h=1}^4 n_h = 382$$

$$2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4$$

$$n_h \in \mathbb{N}.$$

Desarrollando el determinante se tiene

$$\min_{\mathbf{n}} \left[\widehat{V}(\bar{y}_{st}^1) \cdot \widehat{V}(\bar{y}_{st}^2) - \left(\widehat{\text{Cov}}(\bar{y}_{st}^1, \bar{y}_{st}^2) \right)^2 \right]$$

sujeto a

$$\sum_{h=1}^4 n_h = 382 \tag{25}$$

$$2 \leq n_h \leq N_h, \quad h = 1, 2, \dots, 4$$

$$n_h \in \mathbb{N},$$

ya que $\widehat{\text{Cov}}(\bar{y}_{st}^1, \bar{y}_{st}^2) = \widehat{\text{Cov}}(\bar{y}_{st}^2, \bar{y}_{st}^1)$. Donde

$$\widehat{V}(\bar{y}_{st}^1) = \sum_{h=1}^4 \frac{W_h^2 s_{h1}^2}{n_h} - \sum_{h=1}^4 \frac{W_h s_{h1}^2}{N},$$

$$\widehat{V}(\bar{y}_{st}^2) = \sum_{h=1}^4 \frac{W_h^2 s_{h2}^2}{n_h} - \sum_{h=1}^4 \frac{W_h s_{h2}^2}{N},$$

$$\widehat{\text{Cov}}(\bar{y}_{st}^1, \bar{y}_{st}^2) = \sum_{h=1}^4 \frac{W_h^2 s_{h12}^2}{n_h} - \sum_{h=1}^4 \frac{W_h s_{h12}^2}{N}.$$

Como el problema no menciona cuales son los valores para s_{h12}^2 , se simuló una muestra con el programa de cómputo S-Plus y se obtuvo el Cuadro 4.1.

CUADRO 4.1: Covarianzas calculadas con la simulación de muestra.

Estrato No.	Covarianza s_{h12}^2
1	-23,432,794
2	8,613,673
3	174,845.7
4	773,133,590

CUADRO 4.2: Tamaños de muestra de las diferentes asignaciones calculadas.

Asignación	n_1	n_2	n_3	n_4
Función de valor	201	28	32	121
Lexicográfico	207	27	31	117
ε -restricciones	120	39	55	168
Distancias q=1	201	28	32	121
Distancias q=2	190	29	36	127
Distancia Khuri y Cornell	191	29	35	127
Método Dalenius	356	2	22	2
Método Costos s.a. Varianzas	88	43	61	190

Se puede observar que los estratos con mayor número de unidades asignadas son el estrato 1 y el estrato 4, independientemente de la técnica de optimización multiobjetivo aplicada. Esto se debe a que las varianzas de las características evaluadas son mayores en estos estratos, ver Cuadro 2.1. Este comportamiento se mantiene al aplicar el método donde se minimizan los costos, pero la asignación cambia drásticamente sobre todo en el estrato 1, según se puede observar en el Cuadro 4.2. Las asignaciones óptima más alejada del resto de las asignaciones es obtenida a través del método de Dalenius la cual asigna casi el 93 por ciento de la muestra al estrato 1.

Cabe destacar que se ilustraron todos los métodos con el mismo ejemplo, pero esto en la vida real no se presenta, ya que cuando se planteó el método a

utilizar se deberá contemplar si se dispone de información completa, parcial o nula, y en consecuencia aplicar el método más adecuado para cada situación, según la experiencia del investigador.

También es importante destacar lo mucho que difieren los tamaños de muestra cuando por naturaleza del experimento, o por interés del investigador, se debe decidir entre minimizar costos o varianzas.

Conclusiones.

El problema de asignación óptima en muestreo estratificado multivariado ha sido tratado en la literatura estadística anteriormente, pero las soluciones que se han propuesto son casos particulares de una *técnica de optimización multiobjetivo*. Más aún, todas las soluciones propuestas no son más que un caso particular de la técnica conocida como *función de valor*, en el contexto de optimización multiobjetivo. En la solución de un programa de optimización multiobjetivo se tienen tres escenarios posibles, tener: información nula, información parcial o información completa. Desafortunadamente la técnica de la función de valor es propuesta bajo el contexto de información completa. Llevar esta idea al contexto del muestreo, implica que el investigador debe de conocer perfectamente a la población en estudio, al grado de ser capaz de proponer una función de valor que contenga la importancia de todas y cada una de las varianzas de las características estudiadas, aspecto que en la práctica rara vez se presenta. Tomando en cuenta esta situación, en el presente trabajo se expusieron técnicas alternativas en los contextos de información

parcial (donde puede ser suficiente conocer a la característica más importante) y en el contexto de nula información (no es necesario conocer información adicional a los estimadores de los parámetro de estudio) para una solución más adecuada.

Una vez planteado el problema de asignación óptima en el muestreo estratificado multivariado como un programa de optimización multiobjetivo, las técnicas propuestas para resolver este problema son ilustradas a través de un ejemplo de la literatura, ver Sukhatme *et al.* (1984). Al abordar el problema de asignación óptima en muestreo multivariado, primeramente debe analizarse en cual de los tres contextos (información total, información parcial e información nula) se encuentra dicho problema, una vez hecho esto, se procede a decidir la técnica que por consecuencia de la información disponible debe aplicarse. Es importante destacar que la solución para un problema de asignación se realiza exclusivamente por un sólo método. Es por esto que los resultados obtenidos para el ejemplo son comparables únicamente dentro del contexto en el que se localicen.

Así por ejemplo observe los tamaños de muestra asignados por los métodos *función de valor* y *lexicográfico* (ambos en el contexto información completa) tienen poca diferencia. De igual forma se puede comparar los tamaños de muestra asignados por el *método de distancias* (información nula) en las tres diferentes distancias empleadas, *distancia $q=1$* , *distancia $q=2$* y *distancia Khuri y Cornell*, se aprecia que los tamaños de muestra para los diferentes estratos son muy similares, particularmente para la distancia $q=2$ y la distancia propuesta por Khuri y Cornell, el tamaño de los estratos sólo difiere por una unidad en el estrato uno y el estrato tres.

Es primordial recalcar que la filosofía propuesta por Dalenius, se acerca un poco más a las situaciones reales, pues considera la posible correlación entre las características en estudio, cuyo supuesto no ha sido considerado en ninguno de los contextos aquí propuestos. Así el desarrollo de estudios en esta dirección queda abierto, ya que como se puede observar en el ejemplo las asignaciones obtenidas por este método son muy diferentes a las obtenidas cuando no se considera correlación entre las características.

Resumiendo, el presente trabajo realiza las siguientes aportaciones:

1. Extiende los métodos de optimización multiobjetivo al campo de la estadística, específicamente al área del muestreo estratificado multivariado.
2. Establece la solución al problema de asignación óptima para el caso multivariado, en los casos en que no se cuenta con información total.
3. Expone una forma de abordar el problema de asignación cuando se considera correlación entre las características estudiadas. Motivando a la investigación posterior de este problema.

LITERATURA CITADA

Arthanari, T.S. and Dodge Yadolah (1981). *Mathematical Programming in Statistics*.

A Wiley-Interscience Publication John Wiley and Sons, Inc.

Cochran, W. G. 1977. *Técnicas de muestreo*. C.E.C.S.A. México.

Khuri, A.I. and Cornell, J.A. (1987). *Response Surface: Designs Analysis*. Marcel

Dekker, Inc. New York.

Lohr, Sh. L. 2000. *Muestreo: Diseño y Análisis*. Internacional Thomson Editores.

Prawda Witenberg, J. (1981). *Métodos y modelos de investigación de operaciones*.

Editorial Limusa.

Prékopa A. (1978). "*The use of stochastic programing for the solution of the some*

problems in statistics and probability". Technical Summary report 1983. University of Wisconsin-Madison.

Rao, S.S. 1978-79. *Optimization theory and applications*. Wiley Eastern Limited.

Ríos, S., Ríos Insua, S., y Ríos Insua, M.J. (1989). *Procesos de decisión Multicriterio*.

EUDEMAUNIVERSIDAD Manuales.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S., and Asok, C. (1984). *Sampling theory of surveys with application*. State University Press.

APÉNDICE A

Multiplicadores de Lagrange.

El problema a resolver es:

$$\text{Optmizar } f(X_1, X_2, \dots, X_n)$$

sujeto a

$$g_1(X_1, X_2, \dots, X_n) = 0,$$

$$g_2(X_1, X_2, \dots, X_n) = 0,$$

\vdots

$$g_m(X_1, X_2, \dots, X_n) = 0,$$

donde $f(X_1, X_2, \dots, X_n)$, $g_1(X_1, X_2, \dots, X_n)$, \dots , $g_m(X_1, X_2, \dots, X_n)$ son funciones de n variables independientes, y la palabra optimizar puede significar maximización o minimización. La función $f(\cdot)$, se le llama *función objetivo* y las funciones $g_1(\cdot)$, $g_2(\cdot)$, \dots , $g_m(\cdot)$, *restricciones*.

El método clásico para resolver este problema, es utilizando los *multiplicadores de Lagrange*, Prawda (1981). Esto se realiza definiendo de la siguiente manera una nueva función objetivo, llamada el *Lagrangeano*:

$$\begin{aligned} F(X_1, X_2, \dots, X_n, \lambda_1, \lambda_2, \dots, \lambda_m) = & f(X_1, X_2, \dots, X_n) \\ & - \lambda_1 g_1(X_1, X_2, \dots, X_n) - \lambda_2 g_2(X_1, X_2, \dots, X_n) - \\ & \dots - \lambda_m g_m(X_1, X_2, \dots, X_n). \end{aligned}$$

El *Lagrangeano* $F(\cdot)$ con $m + n$ variables independientes, es equivalente a la función objetivo original, $f(\cdot)$ con n variables independientes. La gran ventaja de tener el *Lagrangeano*, es que ahora se tiene un problema de optimización *sin restricciones*, aunque con más variables. A continuación habrá que obtener todas las siguientes derivadas parciales, igualarlas a cero y resolver el sistema de ecuaciones resultantes.

$$\begin{aligned} \frac{\partial F}{\partial X_1} &= 0 \\ \frac{\partial F}{\partial X_2} &= 0 \\ &\vdots \\ \frac{\partial F}{\partial X_n} &= 0 \\ \frac{\partial F}{\partial \lambda_1} &= 0 \\ \frac{\partial F}{\partial \lambda_2} &= 0 \\ &\vdots \\ \frac{\partial F}{\partial \lambda_m} &= 0 \end{aligned}$$

La solución del sistema de ecuaciones genera un punto crítico $(X_1^0, X_2^0, \dots, X_n^0, \lambda_1^0, \lambda_2^0, \dots, \lambda_m^0)$, el cual deberá probarse si da un máximo, mínimo o punto de silla. Esto, para el caso de una función de dos variables se hace calculando $\frac{\partial^2 F}{\partial X_1^2}$, $\frac{\partial^2 F}{\partial X_2^2}$ y $\frac{\partial^2 F}{\partial X_1 \partial X_2}$ y evaluándolas en los puntos críticos obtenidos (X_1^0, X_2^0) , denotando el resultado por

$$\frac{\partial^2 F}{\partial X_1^2} \Big|_{X_1^0, X_2^0}, \frac{\partial^2 F}{\partial X_2^2} \Big|_{X_1^0, X_2^0}, \frac{\partial^2 F}{\partial X_1 \partial X_2} \Big|_{X_1^0, X_2^0}.$$

Ahora se calcula

$$\left(\frac{\partial^2 F}{\partial X_1^2} \Big|_{X_1^0, X_2^0} \right) \left(\frac{\partial^2 F}{\partial X_2^2} \Big|_{X_1^0, X_2^0} \right) - \left(\frac{\partial^2 F}{\partial X_1 \partial X_2} \Big|_{X_1^0, X_2^0} \right)^2 = \Delta$$

Si $\Delta > 0$, se tiene un máximo en $(X_1^0, X_2^0, F(X_1^0, X_2^0))$ si

$$\frac{\partial^2 F}{\partial X_1^2} \Big|_{X_1^0, X_2^0} < 0,$$

o un mínimo en $(X_1^0, X_2^0, F(X_1^0, X_2^0))$ si

$$\frac{\partial^2 F}{\partial X_1^2} \Big|_{X_1^0, X_2^0} > 0,$$

si $\Delta \leq 0$, la regla falla y es necesario investigar la función en el entorno al punto (X_1^0, X_2^0) . En este caso, si no hay ni un máximo o mínimo en $(X_1^0, X_2^0, F(X_1^0, X_2^0))$, este punto recibe el nombre de punto de silla.

Ejemplo. Encuentre el máximo o mínimo de

$$f(X, Y) = 5X^2 + 6Y^2 - XY$$

sujeto a

$$g(X, Y) = X + 2Y - 24 = 0$$

El *Lagrangeano* es:

$$\begin{aligned} F(X, Y, \lambda) &= f(X, Y) - \lambda g(X, Y) \\ &= 5X^2 + 6Y^2 - XY - \lambda(X + 2Y - 24). \end{aligned}$$

Entonces

$$\begin{aligned} \frac{\partial F}{\partial X} &= 10X - Y - \lambda = 0 \\ \frac{\partial F}{\partial Y} &= 12Y - X - 2\lambda = 0 \\ \frac{\partial F}{\partial \lambda} &= -X - 2Y + 24 = 0 \end{aligned}$$

La solución de este problema lineal de 3 ecuaciones con tres incógnitas, genera el punto crítico $X = 6$, $Y = 9$, $\lambda = 51$

$$\begin{aligned}\frac{\partial^2 F}{\partial X^2} \Big|_{(6,9,51)} &= 10 \\ \frac{\partial^2 F}{\partial Y^2} \Big|_{(6,9,51)} &= 12 \\ \frac{\partial^2 F}{\partial X \partial Y} \Big|_{(6,9,51)} &= -1 \\ \frac{\partial^2 F}{\partial Y \partial X} \Big|_{(6,9,51)} &= -1\end{aligned}$$

$$\left(\frac{\partial^2 F}{\partial X^2} \Big|_{X_0, Y_0} \right) \left(\frac{\partial^2 F}{\partial Y^2} \Big|_{X_0, Y_0} \right) - \left(\frac{\partial^2 F}{\partial X \partial Y} \Big|_{X_0, Y_0} \right)^2 = \Delta$$

$$\Delta = (10)(12) - (-1)^2 = 119 > 0.$$

Como $\Delta > 0$ y las segundas parciales en el punto crítico son positivas, hay un mínimo en $X = 6$ y $Y = 9$, siendo el valor mínimo de la función, el siguiente:

$$f(6, 9) = 5(6)^2 + 6(9)^2 - (6)(9) = 612.$$