ASIGNACIÓN ÓPTIMA EN MUESTREO ESTRATIFICADO: OPTIMIZACIÓN ESTOCÁSTICA

MARÍA MAGDALENA GARAY TAPIA

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 Postgrado

TÍTULO

ASIGNACIÓN ÓPTIMA EN MUESTREO ESTRATIFICADO: OPTIMIZACIÓN ESTOCÁSTICA

Por:

MARÍA MAGDALENA GARAY TAPIA

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. Octubre de 2005

AGRADECIMIENTOS

Quiero agradecer a todas las personas que han contribuido para poder finalizar esta etapa de mi vida:

- A mi familia, que con su amor, me han brindado la tranquilidad y ánimo para llevar a cabo este proyecto de vida.

- Al Dr. José Antonio Díaz García a quien agradezco toda su dedicación, para guiar mi formación académica y la idea de este trabajo, pero más aun agradezco sus consejos y amistad, tanto de él como de su familia, quienes me han apoyado valiosamente en diversos aspectos de mi vida.

- Al Dr. Fernando Esquivel Bocanegra y al M.C. Félix de Jesús Sánchez Pérez por su tiempo y aportaciones a este trabajo.

- A la Universidad Autónoma Agraria Antonio Narro, por la oportunidad y el apoyo que se me ha brindado.

- A CONACYT por hacer viable mi evolución académica y profesional.

- También, agradezco el apoyo parcial a este trabajo, mediante el proyecto de investigación 45974-F del CONACYT-México.

DEDICATORIA

Con todo mi amor, para Álvaro

Gracias por aparecer en mi vida, y convertirte en mi cómplice desde el primer día que decidí tomar este camino. Lo que hoy vivo, es un sueño que tuve hace cuatro años.

Gracias.

COMPENDIO

ASIGNACIÓN ÓPTIMA EN MUESTREO ESTRATIFICADO: OPTIMIZACIÓN ESTOCÁSTICA

POR

MARÍA MAGDALENA GARAY TAPIA

MAESTRÍA

ESTADÍSTICA EXPERIMENTAL

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA. Octubre de 2005

Dr. José Antonio Díaz García -Asesor-

Palabras claves: Muestreo estratificado, E-modelo, optimización estocástica, asignación óptima, programación de enteros, restricciones aleatorias.

En el presente trabajo se plantea el problema de la asignación óptima en muestreo estratificado como un problema de optimización estocástica de enteros. Se concluye con un ejemplo en donde se proponen las soluciones determinísticas clásicas, de enteros y las soluciones a través del E-modelo modificado y la técnica de restricciones aleatorias.

ABSTRACT

OPTIMUM ALLOCATION IN STRATIFIED SURVEY: STOCHASTIC PROGRAMMING

BY

MARÍA MAGDALENA GARAY TAPIA

MASTER

EXPERIMENTAL STATISTICS

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA. October, 2005

Ph. D. José Antonio Díaz García -Advisor-

Key Words: Stratified sampling, E-model, stochastic programming, optimum allocation, integer programming, chance constraints.

In this work we consider the allocation problem for stratified surveys as a problem of integer non-linear stochastic programming. An example is solved by mean of the classical Lagrangian multiplier, modified E-model and chance constraints techniques.

ÍNDICE

ÍNDICE DE CUADROS

I. INTRODUCCIÓN

El muestreo es una de las áreas de la estadística que ha ganado reconocimiento y confianza en las últimas décadas, lo cual se ve reflejado en la gran cantidad de información a la que se puede acceder hoy en día, y que es generada a partir del estudio de una muestra.

El creciente interés radica principalmente en la cantidad de ventajas que otorga el trabajar con una muestra en lugar de una población completa, algunas de estas razones son, las siguientes:

- Se puede trabajar con mayor rapidez tanto, en la recolección de información, como en el análisis de datos.
- Es mucho más económico y práctico (en muchos sentidos) obtener y analizar una menor cantidad de datos.
- Se tiene mayor precisión, en el sentido de que se puede tener un mejor control en la metodología de recolección de información.
- Cuando se obtiene una "buena muestra", ésta representa las características globales de la población.

Las técnicas de muestreo han evolucionado en su capacidad para modelar diseños de muestreo ante situaciones muy complejas. Además, en la aplicación ha contribuido el desarrollo de procesadores y programas para analizar una mayor cantidad de información.

Además del muestreo por conglomerados en una o varias etapas, una de las técnicas de muestreo más utilizada, es el muestreo estratificado; esto se debe a la capacidad que tiene está técnica para dividir una población en subpoblaciones (estratos) con caracterśticas similares. A diferencia de otras técnicas, el muestreo estratificado considera las características de la población desde la etapa inicial del muestreo, donde estas características se convierten en criterios de decisión para formar los estratos.

Una de las razones más importantes para trabajar con muestreo estratificado, es que puede transformar una población muy heterogńea en un conjunto de subpoblaciones (estratos) homogéneos, lo cual estadísticamente es importante, ya que incrementa las estimaciones. Cada estrato es tratado como una población independiente, lo cual permite trabajar con diferentes técnicas de muestreo para cada estrato.

En el muestreo estratificado existen diferentes métodos para asignar el número de unidades muestrales en los estratos, dichos métodos se basan en algunas caracterśticas de los estratos como la varianza, el tamaño, y el costo de obtener una observación. Posiblemente uno de los métodos de asignación más eficiente, es el de asignación óptima, el cual permite obtener un tamaño de muestra y la asignación del número de unidades a los estratos, para un presupuesto fijo o para una varianza deseada. El planteamiento para este método es mediante un problema de optimización, donde se minimiza la varianza de un estimador, sujeto a un costo fijo o bien, se minimiza la función de costo para una precisión deseada de la varianza.

Tradicionalmente, la solución a este problema se ha dado usando la desigualdad de Cauchy-Schwars, (Stuart 1954), o mediante métodos de optimización como Multiplicadores de Lagrange. Sin embargo, la función para la varianza de la media estimada, por ejemplo, involucra a las varianzas poblacionales para cada estrato, las cuales en la práctica rara vez son conocidas; entonces, se requiere utilizar las varianzas muestrales, las cuales son variables aleatorias.

Bajo estas condiciones, observe que el problema de optimización en la asignación óptima no es determinístico, ya que está en términos de las varianzas muestrales que son variables aleatorias. Entonces, el problema de asignación óptima se define como un problema de optimización estocástica.

En este trabajo, se considera la idea de asumir el problema de asignación óptima donde las varianzas muestrales y/o los costos son variables aleatorias, además se desea obtener una solución en términos de enteros para los tamaños de muestra en cada estrato, este problema se define como un problema de optimización estocástica de enteros.

El objetivo de este documento es mostrar formalmente el problema de la asignación óptima a travéz de la optimización estocástica, así como el planteamiento para las soluciones en los distintos casos que pueden presentarse.

Este trabajo se encuentra compuesto por cuatro capítulos y dos apéndices. En el Capítulo II, se abordan los conceptos mas importantes del muestreo estratificado. Se da una idea general de esta técnica de muestreo, los casos y las razones por lo que es conveniente utilizar la estratificación, la teoría del muestreo estratificado, y los métodos de asignación más utilizados, cabe mencionar que la notación que se emplea en el desarrollo de este tema esta dada en forma vectorial. El capítulo finaliza con un ejemplo de asignación óptima, donde se aborda tanto el problema de minimizar la varianza, como el de minimizar la función de costo.

La optimización estocástica, es un tema relativamente de reciente desarrollo, las primeras contribuciones y trabajos aparecen alrededor de 1950, y es hasta hace poco tiempo que es posible resolver problemas a gran escala, gracias al desarrollo tecnológico. Existe una gran diversidad de métodos y técnicas de optimización estocástica, algunas de estas técnicas basan sus soluciones en extensiones de técnicas de optimización determinística. Otro enfoque, son los métodos basados en la obtención de soluciones a través de proponer un problema determinístico equivalente, estos están basados en ideas estadísticas y probabilísticas, algunas de estas técnicas son el E-Modelo, V-Modelo y el P-Modelo. En el Capítulo III, se revisan algunos de estos métodos de optimización estocástica, principalmente los que se basan en ideas estadísticas, y se estudian sus soluciones.

Por último, en el Capítulo IV se trata el problema de la asignación óptima a través de la optimización estocástica. Se establecen las condiciones bajo las cuales se hace el planteamiento formal del problema de asignación óptima con variables estocásticas, y se abordan diversos casos que pueden ser de interés en este problema, así como la solución a cada uno de estos casos. El capítulo termina con la exposición de un ejemplo, en el cual se utilizan técnicas de optimización estocástica de enteros y se comparan las soluciones con las obtenidas en el Capítulo II.

Las últimas dos secciones están compuestas por el Apéndice A, en el cual se exponen algunos temas que son requeridos para la comprensión de este trabajo, y el Apéndice B, en donde se muestran los datos que fueron utilizados en los ejemplos de los Capítulo II y IV.

II. MUESTREO ESTRATIFICADO

Introducción

Al realizar un estudio o investigación donde se tenga que aplicar alguna de las diversas técnicas para extraer muestras de una población, lo que naturalmente se desearía es obtener la mejor alternativa en cuanto a la representatividad de las muestras. Un buen diseño de muestreo muchas veces radica en pensar, tanto en la naturaleza de los datos, como en el contexto del problema o estudio, lo cual da idea de algunas variables auxiliares que pueden contribuir a diseñar la muestra de interés.

Un ejemplo de esto es el peso de las personas, se sabe incluso antes de realizar el estudio que las mujeres suelen pesar menos que los hombres, si estudiamos una población donde por cada hombre hay dos mujeres, podría representar un error el tomar una muestra directamente de la población de mujeres y hombres para estudiar el peso promedio, en este caso, sería mas preciso considerar dos subpoblaciones una de mujeres y otra de hombres, lo cual concedería homogeneidad en cada estrato.

En el muestreo estratificado se trabaja precisamente con esta idea, si se conoce información adicional que contribuya a diseñar una muestra, se puede dividir a la población en subpoblaciones, sujeto a cierta variable de interés (el genero en el ejemplo mencionado), de tal forma que se obtengan muestras que permitan estimaciones mas precisas de las cantidades poblacionales. A este tipo de muestras se les conoce como **muestras estra-**

tificadas.

En el muestreo estratificado la población es dividida en H subpoblaciones llamadas estratos. Es decir una población de N unidades se divide en subpoblaciones de $N_1, N_2, ..., N_H$ unidades donde las subpoblaciones no se traslapan y el conjunto comprende a toda la población de modo que cada unidad de la muestra pertenece sólo a un estrato y se cumple que

$$N_1 + N_2 + \cdots + N_H = N.$$

Es necesario conocer los valores de cada N_h, ya que las estimaciones se basan en estas cantidades.

Una vez seleccionada la variable de estratificación y construidos los estratos, se extrae una muestra independiente de cada uno de ellos. Los tamaños de muestra dentro de cada estrato se les denota por $n_1, n_2, \ldots, n_H$ respectivamente, y

$$n_1 + n_2 + \cdots + n_H = n.$$

Algunas razones importantes, para utilizar un muestreo estratificado son:

1. Si se quiere reducir la posibilidad de obtener una mala muestra, es decir, que no sea representativa, sobre todo en aquellos casos donde la población es demasiado heterogénea, el formar estratos permite tener subpoblaciones homogéneas, y las estimaciones ganan precisión.

2. Si se desea cierta precisión en los datos de alguna de las subpoblaciones, es recomendable tratar la subpoblación como "población" por si misma.

3. Las muestras estratificadas pueden ser administradas de manera más conveniente, y a un menor costo: Por ejemplo, en cuanto a los métodos de recolección de datos, se puede realizar una encuesta por correo a grandes empresas y hacer visitas personalizadas a las más pequeñas.

4. En muestreo estratificado se pueden utilizar distintos esquemas de muestreo en cada uno de los estratos. Esto depende, tanto de las características de los estratos, como de las especificaciones o necesidades de información en el estudio o investigación que se este realizando.

5. El muestreo estratificado, al realizarse de forma correcta, dará estimaciones más precisas (con menor varianza), para toda la población. La estratificación permite reducir la varianza total, ya que la varianza en cada estrato, es menor que la varianza en toda la población.

Teoría del muestreo estratificado.

La teoría del muestreo estratificado se ocupa de estudiar las propiedades de los estimadores en una muestra estratificada, así como también, de los métodos para la elección de los tamaños de muestras n_h que otorguen una precisión máxima. En este capítulo se revisarán estas propiedades, así como algunas de las técnicas de asignación más utilizadas.

Suponemos que, la población de N unidades de muestreo es dividida en H capas o estratos, con N_h unidades de muestreo en el estrato h, los valores $N_1, N_2, ..., N_H$ se conocen y se cumple que

$$N_1 + N_2 + ... + N_H = N.$$

Antes de iniciar con la revisión de la teoría, se establece la notación que se utiliza en este capítulo.

Notación

La notación que se emplea en la exposición del presente trabajo, está dada en las siguientes definiciones. Se hace notar que la notación en "negrito" (por ejemplo, **n**) indica que se está haciendo referencia a un vector, mientras que los escalares se denotan por letras sencillas (por ejemplo, n).

Definición 1. ***(Vector de unidades poblacionales)*** *Sea $N_1, N_2, \ldots, N_H$ el número de unidades poblacionales en cada uno de los estratos, entonces se define*

$$\mathbf{N} = \begin{bmatrix} N_1 \\ \vdots \\ N_H \end{bmatrix} \in \mathbb{R}^H. \tag{1}$$

Así el número total de las unidades de la población esta dado por

$$N = \mathbf{1}_H^{'}\mathbf{N} \text{ donde } \mathbf{1} = (1, 1, \ldots, 1)' \in \mathbf{R}^H$$

Definición 2. ***(Vector de unidades muestrales)*** *El número de unidades muestrales en cada uno de los estratos esta dado por $n_1, n_2, \ldots, n_H$, esto es*

$$\mathbf{n} = \begin{bmatrix} n_1 \\ \vdots \\ n_H \end{bmatrix} \in \mathbb{R}^H. \tag{2}$$

Luego, el número total de unidades en la muestra esta dado por

$$n = \mathbf{1}_H^{'}\mathbf{n}.$$

Definición 3. ***(Vector de valores poblacionales)*** *Para cada estrato,* $\mathbf{Y}_1, \ldots, \mathbf{Y}_H$ *donde* $\mathbf{Y}_h = (Y_{h1}, Y_{h2}, \ldots, Y_{hN_h})' \in \mathbb{R}^{N_h}, \quad h = 1, 2, \ldots, H$*, son los valores poblacionales de la variable de interés en el estrato* h*, se tiene*

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_H \end{bmatrix} \in \mathbb{R}^N. \tag{3}$$

Definición 4. ***(Vector de valores muestrales)*** *Dentro de cada estrato,* $\mathbf{y}_1, \ldots, \mathbf{y}_H$ *donde* $\mathbf{y}_h = (y_{h1}, y_{h2}, \ldots, y_{hn_h})' \in \mathbb{R}^{n_h}, \quad h = 1, 2, \ldots, H$*, son los valores muestrales de la variable de estudio en el estrato* h*, se define*

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_H \end{bmatrix} \in \mathbb{R}^n. \tag{4}$$

Definición 5. ***(Vector de totales poblacionales)*** *Sea* $t_1, t_2, \ldots, t_H$ *los totales poblacionales en cada estrato, se tiene*

$$\mathbf{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_H \end{bmatrix} \in \mathbb{R}^H \tag{5}$$

Donde cada t_h es calculado mediante

$$t_h = \mathbf{1}_{N_h}^{'} \mathbf{Y}_h \text{ donde } \mathbf{1} \in \mathbb{R}^{N_H}, \quad h = 1, 2, \ldots, H$$

y el total de la población esta dado por

$$t = \mathbf{1}_H^{'} \mathbf{t}.$$

Definición 6. ***(Vector de ponderaciones)*** *Las ponderaciones o pesos que corresponden a los tamaños relativos de cada uno de los estratos, están denotados por* $W_1, W_2, \dots, W_H$*, donde*

$$\mathbf{W} = \begin{bmatrix} W_1 \\ \vdots \\ W_H \end{bmatrix} = \begin{bmatrix} \frac{N_1}{N} \\ \vdots \\ \frac{N_H}{N} \end{bmatrix} = \frac{1}{N}(\mathbf{N}) \in \mathbb{R}^H. \tag{6}$$

Definición 7. ***(Vector de fracciones de muestreo)*** *Para cada estrato, las fracciones de muestreo están dadas por* $f_1, f_2, \dots, f_H$*, se define*

$$\mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_H \end{bmatrix} = \begin{bmatrix} \frac{n_1}{N_1} \\ \vdots \\ \frac{n_H}{N_H} \end{bmatrix} = \mathbf{n} \odot \mathbf{N}^{(-)} \in \mathbb{R}^H. \tag{7}$$

donde $\odot$ es el producto Hadamar (ver Apéndice A), y

$$\mathbf{N}^{(-)} = \begin{bmatrix} \frac{1}{N_1} \\ \vdots \\ \frac{1}{N_H} \end{bmatrix} \in \mathbb{R}^H.$$

Definición 8. ***(Vector de medias poblacionales)*** *Las medias poblacionales correspondientes a cada uno de los estratos están denotadas por* $\bar{Y}_1, \dots, \bar{Y}_H$ *y se define por*

$$\bar{\mathbf{Y}} = \begin{bmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_H \end{bmatrix} = \begin{bmatrix} \frac{1}{N_1}\sum_{i=1}^{N_1} Y_{1i} \\ \vdots \\ \frac{1}{N_H}\sum_{i=1}^{N_H} Y_{Hi} \end{bmatrix} = \begin{bmatrix} \frac{1}{N_1}\mathbf{1}'_{N_1}\mathbf{Y}_1 \\ \vdots \\ \frac{1}{N_H}\mathbf{1}'_{N_H}\mathbf{Y}_H \end{bmatrix} = \mathbf{N}^{(-)} \odot \mathbf{t} \in \mathbb{R}^H. \tag{8}$$

Mientras que la media global de la población, es

$$\bar{Y} = \frac{t}{N} = \frac{\mathbf{1}'_H\mathbf{t}}{\mathbf{N}'\mathbf{1}_H}. \tag{9}$$

Definición 9. ***(Vector de medias muestrales)*** *Para las medias muestrales dentro de cada estrato, definidas por $\bar{y}_1, \ldots, \bar{y}_H$, se tiene*

$$\bar{\mathbf{y}} = \begin{bmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_H \end{bmatrix} = \begin{bmatrix} \frac{1}{n_1}\sum_{i=1}^{n_1} y_{1i} \\ \vdots \\ \frac{1}{n_H}\sum_{i=1}^{n_H} y_{Hi} \end{bmatrix} = \begin{bmatrix} \frac{1}{n_1}\mathbf{1}_{n_1}^{'}\mathbf{y}_1 \\ \vdots \\ \frac{1}{n_H}\mathbf{1}_{n_H}^{'}\mathbf{y}_H \end{bmatrix} \in \mathbb{R}^H \tag{10}$$

.

Definición 10. ***(Vector de varianzas poblacionales)*** *Sean $S_1, \ldots, S_H$ las varianzas poblacionales para cada uno de los estratos, se define*

$$\mathbf{S^2} = \begin{bmatrix} S_1^2 \\ \vdots \\ S_H^2 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{N_1 - 1}\mathbf{Y}_1^{'}\left(\mathbf{I}_{N_1} - \dfrac{1}{N_1}\mathbf{1}_{N_1}\mathbf{1}_{N_1}^{'}\right)\mathbf{Y}_1 \\ \vdots \\ \dfrac{1}{N_H - 1}\mathbf{Y}_H^{'}\left(\mathbf{I}_{N_H} - \dfrac{1}{N_H}\mathbf{1}_{N_H}\mathbf{1}_{N_H}^{'}\right)\mathbf{Y}_H \end{bmatrix} \in \mathbb{R}^H. \tag{11}$$

Definición 11. ***(Vector de varianzas muestrales)*** *Sean $s_1, \ldots, s_H$ las varianzas muestrales correspondientes a cada estrato, se tiene*

$$\mathbf{s^2} = \begin{bmatrix} s_1^2 \\ \vdots \\ s_H^2 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{n_1 - 1}\mathbf{y}_1^{'}\left(\mathbf{I}_{n_1} - \dfrac{1}{n_1}\mathbf{1}_{n_1}\mathbf{1}_{n_1}^{'}\right)\mathbf{y}_1 \\ \vdots \\ \dfrac{1}{n_H - 1}\mathbf{y}_H^{'}\left(\mathbf{I}_{n_H} - \dfrac{1}{n_H}\mathbf{1}_{n_H}\mathbf{1}_{n_H}^{'}\right)\mathbf{y}_H \end{bmatrix} \in \mathbb{R}^H. \tag{12}$$

Propiedades de las estimaciones

El **muestreo aleatorio estratificado**, es la forma mas sencilla de realizar un muestreo estratificado, y consiste en considerar una muestra aleatoria de manera independiente en cada estrato, es por ello que las propiedades de los estimadores son heredadas directamente de la teoría del muestreo aleatorio simple. En esta sección se estudian las principales propiedades de los estimadores, tanto en el muestreo aleatorio estratificado, como en el muestreo estratificado en general.

En el contexto de muestreo estratificado, la expresión que se utiliza para la estimación de la media de la población, es $\bar{y}_{st}$ y se calcula como sigue

$$\bar{y}_{st} = \frac{1}{N}(\mathbf{N}'\bar{\mathbf{y}}) = \mathbf{W}'\bar{\mathbf{y}}. \tag{13}$$

Cabe mencionar que generalmente $\bar{y}_{st}$ no tiene el mismo valor que la media muestral $\bar{y}$, la cual puede escribirse como

$$\bar{y} = \frac{1}{n}(\mathbf{n}'\bar{\mathbf{y}}). \tag{14}$$

La diferencia radica básicamente en el valor que toma la pondelación $\mathbf{W}$ en relación con $\frac{1}{n}(\mathbf{n})$. Es evidente que $\bar{y}$ coincide con $\bar{y}_{st}$ cuando en cada estrato, vea Cochran (1977)

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

lo que significa que la fracción de muestreo es la misma en todos los estratos, es entonces que $\bar{y}_{st}$ y $\bar{y}$ coinciden.

A este tipo de estratificación se le conoce como **estratificación con asignación proporcional**, donde cada unidad de la muestra tiene el mismo peso y representa el mismo número de unidades de la población, en este caso se dice que la muestra es **autoponderada**. La asignación proporcional puede ser recomendable cuando no existe mucha variabilidad entre los estratos.

La expresión para la estimación del total de la población en el muestreo aleatorio estratificado esta dado por;

$$\hat{t}_{st} = N\bar{y}_{st} = N(\mathbf{W}'\bar{\mathbf{y}}). \tag{15}$$

Mientras que el total estimado en cada estrato se calcula mediante;

$$\hat{\mathbf{t}} = \begin{bmatrix} \hat{t}_1 \\ \vdots \\ \hat{t}_H \end{bmatrix} = \begin{bmatrix} \bar{y}_1 N_1 \\ \vdots \\ \bar{y}_H N_H \end{bmatrix} \in \mathbb{R}^H. \tag{16}$$

Cabe hacer notar que $\hat{t}_{st} = \mathbf{1}_H'\hat{\mathbf{t}}$.

Los siguientes teoremas exhiben las propiedades más importantes de la estimación de $\bar{y}_{st}$, el teorema 1 y 3 son aplicables al muestreo estratificado en general, es decir que no se requiere que la muestra dentro de un estrato sea una muestra aleatoria simple.

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 poblacional $\bar{Y}$.*

Demostración.

$$\mathrm{E}\left[\bar{y}_{st}\right] = \mathrm{E}(\mathbf{W}'\bar{\mathbf{y}}) = \mathbf{W}'\mathrm{E}(\bar{\mathbf{y}}) = \mathbf{W}'\bar{\mathbf{Y}} = \begin{bmatrix} \frac{N_1}{N} & ,\dots, & \frac{N_H}{N} \end{bmatrix} \begin{bmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_H \end{bmatrix} = \bar{Y}.$$

Recuerde que $\mathrm{E}(\bar{\mathbf{y}})$ está dada por

$$\mathrm{E}(\bar{\mathbf{y}}) = \begin{bmatrix} \mathrm{E}(\bar{y}_1) \\ \vdots \\ \mathrm{E}(\bar{y}_H) \end{bmatrix} = \begin{bmatrix} \frac{1}{n_1}\mathbf{1}_{n_1}'\mathrm{E}(\mathbf{y}_1) \\ \vdots \\ \frac{1}{n_H}\mathbf{1}_{n_H}'\mathrm{E}(\mathbf{y}_H) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{n_1}\mathbf{1}_{n_1}^{'}\bar{Y}_1\mathbf{1}_{n_1} \\ \vdots \\ \frac{1}{n_H}\mathbf{1}_{n_H}^{'}\bar{Y}_H\mathbf{1}_{n_H} \end{bmatrix} = \begin{bmatrix} \frac{1}{n_1}\mathbf{1}_{n_1}^{'}\mathbf{1}_{n_1}\bar{Y}_1 \\ \vdots \\ \frac{1}{n_H}\mathbf{1}_{n_H}^{'}\mathbf{1}_{n_H}\bar{Y}_H \end{bmatrix}$$

$$= \begin{bmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_H \end{bmatrix} = \bar{\mathbf{Y}}.$$

■

Teorema 2. *Para el total estimado, si en cada estrato la estimación muestral $\bar{y}_h$ es insesgada, entonces $\hat{t}_{st}$ es un estimador insesgado de t.*

Demostración.

$$\mathrm{E}(\hat{t}_{st}) = \mathrm{E}(N\bar{y}_{st}) = N\mathrm{E}(\bar{y}_{st}) = N\bar{Y} = N\frac{t}{N} = t.$$

■

Teorema 3. *Si las muestras se extraen de forma independiente en los diferentes estratos se tiene que*

$$\mathrm{Var}\,(\bar{y}_{st}) = \mathrm{Var}(\mathbf{W}^{'}\bar{\mathbf{y}}) = \mathbf{W}^{'}\mathrm{Var}(\bar{\mathbf{y}})\mathbf{W}. \tag{17}$$

Este resultado se desprende de la distribución de vectores aleatorios, donde $\mathbf{W} \in \mathbb{R}^H$, es vector de constantes (ver Apéndice A), y $\mathrm{Var}(\bar{\mathbf{y}})$ es la matriz de varianza y covarianzas del vector de medias muestrales en cada uno de los estratos, está matriz será definida más adelante.

Teorema 4. *Específicamente, en el muestreo aleatorio estratificado, la varianza del estimador $\bar{y}_{st}$ está dada por*

$$\text{Var}(\bar{y}_{st}) = \mathbf{W}' \left[\text{diag} \left(\frac{S_h^2}{n_h}(1 - f_h) \right) \right] \mathbf{W} \tag{18}$$

donde

$$\text{diag} \left(\frac{S_h^2}{n_h}(1 - f_h) \right) = \text{diag} \left(\frac{S_1^2}{n_1} \left(1 - \frac{n_1}{N_1} \right), \ldots, \frac{S_H^2}{n_H} \left(1 - \frac{n_H}{N_H} \right) \right)$$

Demostración.

Utilizando el hecho de que $\bar{y}_h$ es un estimador insesgado de $\bar{Y}_h$, partimos de la ecuación (17), donde el problema es determinar $\text{Var}(\bar{\mathbf{y}})$.

De la teoría del muestreo aleatorio simple, (ver Apéndice A) se tiene el siguiente resultado,

$$\text{Var}(\bar{y}) = E(\bar{y} - \bar{Y})^2 = \frac{S^2}{n} \frac{(N - n)}{N}.$$

Aplicando este resultado al vector de medias en cada estrato, se observa que

$$\text{Var}(\bar{\mathbf{y}}) = \begin{pmatrix} \text{Var}(\bar{y}_h), \quad \bar{y}_h = \bar{y}_j \\ \\ \text{Cov}(\bar{y}_h, \bar{y}_j), \bar{y}_h \neq \bar{y}_j \end{pmatrix}$$

Dado que las muestras se extraen de forma independiente en los diferentes estratos, la $\text{Cov}(\bar{y}_h, \bar{y}_j) = 0$, de modo que la matriz de varianzas y covarianzas esta dada por la matriz diagonal de las $\text{Var}(\bar{y}_h)$, $\;h = 1, 2, \ldots, H.$, esto es

$$\text{Var}(\bar{\mathbf{y}}) = \text{diag} \left(\frac{S_1^2}{n_1}(1 - f_1), \ldots, \frac{S_H^2}{n_H}(1 - f_H) \right).$$

Finalmente sustituyendo el resultado de $\text{Var}(\bar{\mathbf{y}})$ en la ecuación (17), se tiene

$$\text{Var}\left(\bar{y}_{st}\right) = \mathbf{W}' \text{Var}(\bar{\mathbf{y}}) \mathbf{W} = \mathbf{W}' \left[\text{diag} \left(\frac{S_h^2}{n_h}(1 - f_h) \right) \right] \mathbf{W}, \quad h = 1, 2, \ldots, H.$$

$\blacksquare$

Corolario 1. *Si las fracciones de muestreo $f_h = \frac{n_h}{N_h}$ son insignificantes, la varianza de $\bar{y}_{st}$ queda definida como*

$$\text{Var}(\bar{y}_{st}) = \mathbf{W}' \left[\text{diag} \left(\frac{S_h^2}{n_h} \right) \right] \mathbf{W} \tag{19}$$

Lo cual es una expresión apropiada cuando las correcciones para poblaciónes finitas, se puedan ignorar, vea Cochran (1977).

Corolario 2. *Cuando se utiliza la asignación proporcional, se sustituye $n_h = n\frac{N_h}{N}$ en la ecuación (18), y la varianza Var($\bar{y}_{st}$) puede rescribirse como*

$$\begin{aligned}
\text{Var}(\bar{y}_{st}) &= \mathbf{W}'\text{diag}\left[\frac{S_h^2}{n} \frac{N}{N_h} \left(1 - \frac{nN_h}{NN_h} \right) \right] \mathbf{W} \\
&= \mathbf{W}' \frac{N}{n}(1-f) \left[\text{diag} \left(\frac{S_h^2}{N_h} \right) \right] \mathbf{W}
\end{aligned}$$

recordando que $\mathbf{W}' = \frac{1}{N}(\mathbf{N}')$, sustituyendo este resultado se tiene;

$$\begin{aligned}
\text{Var}(\bar{y}_{st}) &= \frac{1}{N}(\mathbf{N}') \left[\frac{N}{n}(1-f)\text{diag} \left(\frac{S_h^2}{N_h} \right) \right] \mathbf{W} \\
&= \frac{(1-f)}{n} \left[\mathbf{N}'\text{diag} \left(\frac{S_h^2}{N_h} \right) \right] \mathbf{W} \\
&= \frac{(1-f)}{n} \left[(N_1, \ldots, N_H) \begin{pmatrix} \frac{S_1^2}{N_1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \frac{S_H^2}{N_H} \end{pmatrix} \mathbf{W} \right] \\
&= \frac{(1-f)}{n}(S_1^2, \ldots, S_H^2)\mathbf{W} \\
&= \frac{(1-f)}{n} (\mathbf{S^2}'\mathbf{W}).
\end{aligned} \tag{20}$$

Teorema 5. *Para el estimador del total de la población $\hat{t}_{st}$, la varianza está dada por*

$$\text{Var}(\hat{t}_{st}) = \mathbf{N}' \left[\text{diag} \left(\frac{S_h^2}{n_h}(1-f_h) \right) \right] \mathbf{N}. \tag{21}$$

Demostración.

$$
\begin{aligned}
\mathrm{Var}(\hat{t}_{st}) &= \mathrm{Var}(N\bar{y_{st}}) \\
&= N^2\left[\mathbf{W}'\mathrm{diag}\left(\frac{S_h^2}{n_h}(1-f_h)\right)\mathbf{W}\right] \\
&= N^2\left[\frac{1}{N}\mathbf{N}'\mathrm{diag}\left(\frac{S_h^2}{n_h}(1-f_h)\right)\frac{1}{N}\mathbf{N}\right] \\
&= \mathbf{N}'\left[\mathrm{diag}\left(\frac{S_h^2}{n_h}(1-f_h)\right)\right]\mathbf{N}.
\end{aligned}
$$

■

Varianza estimada e intervalos de confianza para muestras estratificadas.

Dentro de la teoría del muestreo aleatorio simple, se establece que para la varianza de la población S^2, una estimación insesgada está dada por s^2, (ver Apéndice A), cuya expresión es la siguiente;

$$
s^2 = \frac{1}{n-1}\mathbf{y}'\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\mathbf{y}
$$

Si se considera una muestra aleatoria dentro de cada estrato, se puede obtener un estimador insesgado de $\mathbf{S}^2$, mediante

$$
\mathbf{s}^2 = \begin{bmatrix} s_1^2 \\ \vdots \\ s_H^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{n_1-1}\mathbf{y_1}'\left(\mathbf{I}_{n_1} - \frac{1}{n_1}\mathbf{1}_{n_1}\mathbf{1}_{n_1}'\right)\mathbf{y_1} \\ \vdots \\ \frac{1}{n_H-1}\mathbf{y_H}'\left(\mathbf{I}_{n_H} - \frac{1}{n_H}\mathbf{1}_{n_H}\mathbf{1}_{n_H}'\right)\mathbf{y_H} \end{bmatrix}.
$$

Definido $\mathbf{s}^2$, si las varianzas poblacionales $\mathbf{S}^2$, son sustituidas por las estimaciones muestrales $\mathbf{s}^2$ en la ecuación (18), se puede obtener un estimador insesgado de $\mathrm{Var}(\bar{y_{st}})$, esto es expresado en el siguiente teorema.

Teorema 6. *En el muestreo aleatorio estratificado, una estimación insesgada de la $Var(\bar{y}_{st})$ es*

$$\widehat{\mathrm{Var}}(\bar{y}_{st}) = \mathbf{W}' \left[\mathrm{diag} \left(\frac{s_h^2}{n_h}(1-f_h) \right) \right] \mathbf{W}. \tag{22}$$

Demostración.

Note que

$$\begin{aligned}
\mathrm{E}\left(\widehat{\mathrm{Var}}(\bar{y}_{st})\right) &= \mathrm{E}\left(\mathbf{W}' \left[\mathrm{diag}\left(\frac{s_h^2}{n_h}(1-f_h)\right)\mathbf{W}\right]\right) \\
&= \mathbf{W}' \left[\mathrm{E}\left(\mathrm{diag}\left(\frac{s_h^2}{n_h}(1-f_h)\right)\right)\right]\mathbf{W} \\
&= \mathbf{W}' \left[\mathrm{diag}\left(\frac{\mathrm{E}(s_h^2)}{n_h}(1-f_h)\right)\right]\mathbf{W}, \\
&= \mathbf{W}' \left[\mathrm{diag}\left(\frac{S_h^2}{n_h}(1-f_h)\right)\right]\mathbf{W} \\
&= \mathrm{Var}(\bar{y}_{st}).
\end{aligned}$$

Dado a que $\mathrm{E}(s_h^2) = S_h^2$, ver Apéndice A. ■

Una nota importante es que para estimar la varianza $\widehat{\mathrm{Var}}(\bar{y}_{st})$, es necesario una muestra de cuando menos dos observaciones por estrato. Además se observa que

$$\begin{aligned}
\widehat{\mathrm{Var}}(\bar{y}_{st}) &= \mathbf{W}' \left[\mathrm{diag}\left(\frac{s_h^2}{n_h}\right) - \mathrm{diag}\left(\frac{s_h^2}{n_h}f_h\right)\right]\mathbf{W} \\
&= \mathbf{W}' \left[\mathrm{diag}\left(\frac{s_h^2}{n_h}\right) - \mathrm{diag}\left(\frac{s_h^2}{n_h}\frac{n_h}{N_h}\right)\right]\mathbf{W} \\
&= \mathbf{W}' \left[\mathrm{diag}\left(\frac{s_h^2}{n_h}\right)\right]\mathbf{W} - \mathbf{W}' \left[\mathrm{diag}\left(\frac{s_h^2}{N_h}\right)\right]\mathbf{W} \\
&= \mathbf{W}' \left[\mathrm{diag}\left(\frac{s_h^2}{n_h}\right)\right]\mathbf{W} - \frac{1}{N}(\mathbf{s}^2)'\mathbf{W}. \qquad (23)
\end{aligned}$$

Se observa que la ecuación (23) es mas sencilla para realizar cálculos, que la dada por (22).

Utilizando nuevamente el hecho de que $\mathbf{s}^2$ es un estimador insesgado de la varianza poblacional $\mathbf{S}^2$, se tiene que para la varianza estimada del total, la cual está denotada por $\widehat{\text{Var}}(\hat{t}_{st})$, se sustituyen las varianzas poblacionales $\mathbf{S}^2$, por las estimaciones $\mathbf{s}^2$, en la ecuación (21), entonces

$$\widehat{\text{Var}}(\hat{t}_{st}) = \mathbf{N}^{'} \left[\text{diag} \left(\frac{s_h^2}{n_h}(1 - f_h) \right) \right] \mathbf{N}. \tag{24}$$

Finalmente, el error estándar de $\bar{y}_{st}$, esta dado por

$$EE(\bar{y}_{st}) = \sqrt{\widehat{\text{Var}}(\bar{y}_{st})}. \tag{25}$$

Intervalos de confianza para muestras estratificadas. 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;

$$\bar{Y} : \qquad \bar{y}_{st} \pm z_{\alpha/2} EE(\bar{y_{st}}). \tag{26}$$

Análogamente para el total t, la fórmula para el intervalo de confianza es;

$$t : \quad \hat{t}_{st} \pm z_{\alpha/2} EE(\bar{y}_{st}). \tag{27}$$

Estas fórmulas, suponen que $\bar{y}_{st}$, esta distribuido normalmente y que además, $EE(\bar{y}_{st})$ está bien definido, de modo que $z_{\alpha/2}$ pueda ser fácilmente encontrado en la tablas de la distribución normal estándar.

En el caso donde sólo se dan algunos grados de libertad en cada estrato, el procedimiento usual para tomar en cuenta el error de muestreo inherente a una cantidad como $EE(\bar{y}_{st})$ consiste en leer el valor de $z_{\alpha/2}$ en las tablas de la t de *Student* en lugar de las tablas de la distribución normal estándar. La distribución de $EE(\bar{y}_{st})$ es en general muy complicada, para permitir una aplicación estricta del procedimiento comentado. Satterthwaite (1946) propone un método aproximado para asignar un número efectivo de grados de libertad a $EE(\bar{y}_{st})$, la descripción de este método puede encontrarla en Cochran (1977).

Asignación de las observaciones en los estratos.

En está sección se analizan algunos de los métodos más utilizados para la asignación de las observaciones en los estratos, así como los criterios que se consideran para dicha asignación.

Asignación Proporcional.

Si se extrae una muestra estratificada con el propósito de que ésta refleje la población con respecto a la variable de estratificación, y si además se desea que ésta fuera una versión miniatura de la población, entonces la asignación proporcional es el diseño más adecuado a emplear.

En la asignación proporcional, simplemente se toma un número de observaciones en la muestra proporcional al tamaño del estrato, donde la probabilidad de selección es la misma en todos los estratos

$$\frac{n}{N} = \frac{n_h}{N_h}.$$

Por ejemplo, se considera una población con 750 estudiantes de los cuales 500 son estudiantes locales y 250 son estudiantes foráneos, si se considera como variable de estratificación la procedencia de los estudiantes, una distribución proporcional con una muestra del 10%, implica tomar una muestra de 50 estudiantes locales y 25 foráneos.

Al utilizar asignación proporcional, cada unidad de la muestra representa el mismo número de unidades de la población, y como se comentó anteriormente la muestra es llamada **autoponderada**. En una muestra de este tipo, $\bar{y}_{st}$ es el promedio de todas las observaciones que existen en dicha muestra ($\bar{y}$), vea (Lohr, 2000, sec.4.4).

Asignación Óptima.

La asignación proporcional, es probablemente la mejor alternativa en términos de precisión cuando las varianzas $\mathbf{S}^2$ son parecidas o iguales a través de todos los estratos. Pero si las varianzas $\mathbf{S}^2$ varían demasiado entonces la asignación óptima puede producir un menor costo y una mayor precisión.

En muchos casos aplicados, cuando se trabaja con unidades de muestreo que difieran mucho en magnitud, es decir que los valores que toman las $y_{hi},\ i = 1, 2, \ldots, n_h$ son bastante heterogéneos, es común que las unidades de mayor magnitud tengan más variabilidad que las unidades de menor magnitud, y que éstas sean extraídas en una mayor proporción. 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. No debemos olvidar que un objetivo del muestreo, es obtener más información a un menor costo.

En el muestreo estratificado con asignación óptima, el valor de los tamaños de muestra n_h en los diversos estratos, se pueden elegir para minimizar la varianza $\mathrm{Var}(\bar{y}_{st})$ dada una función de costo determinado, o de forma alternativa, para minimizar la función de costo, sujeto a un valor especifico de $\mathrm{Var}(\bar{y}_{st})$.

Una de las funciones de costo más sencilla es la siguiente;

$$C = c_0 + \mathbf{c}'\mathbf{n} \tag{28}$$

donde

$$\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_H \end{bmatrix} \in \mathbb{R}^H,$$

y se define como el vector de costos para cada uno de los estratos, los costos pueden variar de estrato a estrato. C es el costo total, c_0 es un costo adicionales y finalmente $\mathbf{n}$ es el vector de unidades muestrales en cada estrato. Como se observa en la ecuación (28), el costo será proporcional al tamaño de la muestra, es decir C es mayor, cuando en uno o varios estratos, el tamaño de muestra n_h es relativamente grande.

Algunos estudios empíricos y matemáticos sugieren, que cuando los costos de viaje entre las unidades son sustanciales, estos son mejor representados cuando se utiliza la expresión,

$$(\mathbf{c_v}'\mathbf{n}^{1/2}).$$

donde $\mathbf{c_v} \in \mathbf{R}^H$ es el costo de viaje por unidad en cada estrato, vea Cochran (1977, Pág.133). Para nuestros propósitos únicamente se considera la función lineal de costo (28).

Teorema 7. *En el muestreo aleatorio estratificado, con una función de costo lineal de la forma (28), la varianza de la media estratificada $\bar{y}_{st}$ es un mínimo para un costo especifico C, y el costo es un mínimo para una varianza especifica $Var(\bar{y}_{st})$, donde*

$$\mathbf{n} \propto (\mathbf{W} \odot (\mathbf{c}^{1/2(-)} \odot \mathbf{S})) \tag{29}$$

Demostración.

Consideremos el problema de elegir $\mathbf{n}$ para minimizar la varianza $\mathrm{Var}(\bar{y}_{st})$ sujeto a una función de costo, dado un costo fijo C, entonces el problema puede expresarse como

$$\min_{\mathbf{n}} \mathrm{Var}(\bar{y}_{st}) = \mathbf{W}' \left[\mathrm{diag} \left(\frac{S_h^2}{n_h} \right) \right] \mathbf{W} - \frac{1}{N} (\mathbf{S}^2)' \mathbf{W}. \tag{30}$$

Sujeto a la restricción $(s.a)$

$$\mathbf{c}'\mathbf{n} = C - c_0. \tag{31}$$

Utilizando el metódo de los multiplicadores de Lagrange, seleccionamos $\mathbf{n}$ y el multiplicador λ para minimizar

$$\begin{aligned} f(\mathbf{n}) &= \mathrm{Var}(\bar{y_{st}}) + \lambda \left(\mathbf{c}'\mathbf{n} - C + c_0 \right) \qquad (32) \\ &= \mathbf{W}' \left[\mathrm{diag} \left(\frac{S_h^2}{n_h} \right) \right] \mathbf{W} - \frac{1}{N} (\mathbf{S}^2)' \mathbf{W} + \lambda (\mathbf{c}'\mathbf{n} - C + c_0). \\ &= \mathbf{W}' \left[\mathrm{diag} \left(\mathbf{S}^2 \odot \mathbf{n}^{(-)} \right) \right] \mathbf{W} - \frac{1}{N} (\mathbf{S}^2)' \mathbf{W} + \lambda (\mathbf{c}'\mathbf{n} - C + c_0). \end{aligned}$$

Donde $\mathbf{n}^{(-)}$ se define como

$$\mathbf{n}^{(-)} = \begin{bmatrix} \frac{1}{n_1} \\ \vdots \\ \frac{1}{n_H} \end{bmatrix}$$

Diferenciando con respecto a $\mathbf{n}$ tenemos

$$\begin{aligned} df(\mathbf{n}) &= d \left[\mathbf{W}' \left(\mathrm{diag} \left(\mathbf{S}^2 \odot \mathbf{n}^{(-)} \right) \right) \mathbf{W} - \frac{1}{N} (\mathbf{S}^2)' \mathbf{W} + \lambda (\mathbf{c}'\mathbf{n} - C + c_0) \right] \\ &= d \left[\mathbf{W}' \left(\mathrm{diag} \left(\mathbf{S}^2 \odot \mathbf{n}^{(-)} \right) \right) \mathbf{W} \right] + d \left[\lambda (\mathbf{c}'\mathbf{n} - C + c_0) \right] \\ &= - \left(\frac{W_1^2 S_1^2}{n_1^2}, \ldots, \frac{W_H^2 S_H^2}{n_H^2} \right)' d\mathbf{n} + (\lambda \mathbf{c}') d\mathbf{n}. \end{aligned}$$

Por lo tanto

$$\begin{aligned}
\frac{\partial f(\mathbf{n})}{\partial \mathbf{n}} &= \lambda \mathbf{c} - \begin{bmatrix} \frac{W_1^2 S_1^2}{n_1^2} \\ \vdots \\ \frac{W_H^2 S_H^2}{n_H^2} \end{bmatrix} \\
&= \lambda \mathbf{c} - (\mathbf{n}^{(-)} \odot \mathbf{n}^{(-)}) \odot \begin{bmatrix} W_1^2 S_1^2 \\ \vdots \\ W_H^2 S_H^2 \end{bmatrix} = 0 \\
\lambda \mathbf{c} &= (\mathbf{n}^{(-)} \odot \mathbf{n}^{(-)}) \odot \mathbf{W}^2 \odot \mathbf{S}^2. \qquad (33)
\end{aligned}$$

Donde $\mathbf{W}^2 = (\mathbf{W} \odot \mathbf{W})$ y además $\mathbf{n}^{(-)} \odot \mathbf{n}^{(-)}$ es ;

$$\mathbf{n}^{(-)} \odot \mathbf{n}^{(-)} = \begin{bmatrix} \frac{1}{n_1} \\ \vdots \\ \frac{1}{n_H} \end{bmatrix} \odot \begin{bmatrix} \frac{1}{n_1} \\ \vdots \\ \frac{1}{n_H} \end{bmatrix} = \begin{bmatrix} \frac{1}{n_1^2} \\ \vdots \\ \frac{1}{n_H^2} \end{bmatrix}$$

Ahora para despejar a $\mathbf{n}$ se tiene que

$$\lambda \mathbf{c} = (\mathbf{n}^{(-)} \odot \mathbf{n}^{(-)}) \odot \mathbf{W}^2 \odot \mathbf{S}^2.$$

Sacando raíz cuadrada en ambos lados

$$\begin{aligned}
\sqrt{\lambda} \mathbf{c}^{1/2} &= \mathbf{n}^{(-)} \odot \mathbf{W} \odot \mathbf{S} \\
\sqrt{\lambda} \mathbf{n} &= \mathbf{c}^{1/2(-)} \odot \mathbf{W} \odot \mathbf{S} \\
&= \begin{bmatrix} \frac{W_1 s_1}{\sqrt{c_1}} \\ \vdots \\ \frac{W_H s_H}{\sqrt{c_H}} \end{bmatrix} \qquad (34)
\end{aligned}$$

Para la proporción, se consideran las cantidades totales, esto es la suma sobre todos los estratos,

$$
\begin{aligned}
\sqrt{\lambda}\mathbf{1}'(\mathbf{n}) &= \mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)}) \\
\sqrt{\lambda} n &= \mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)}),
\end{aligned}
$$

donde n es el tamaño de la muestra global.

La proporción en términos del tamaño de muestra en cada estrato está dada por

$$
\begin{aligned}
\frac{\sqrt{\lambda}\mathbf{n}}{\sqrt{\lambda} n} &= \frac{\mathbf{c}^{1/2(-)} \odot \mathbf{W} \odot \mathbf{S}}{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})} \\
\frac{1}{n}(\mathbf{n}) &= \frac{1}{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})} \begin{bmatrix} \frac{W_1 S_1}{\sqrt{c_1}} \\ \vdots \\ \frac{W_H S_H}{\sqrt{c_H}} \end{bmatrix}.
\end{aligned}
\tag{35}
$$

■

El resultado anterior implica que se extraen más unidades de un estrato, si se cumple lo siguiente;

- Si la varianza en el interior de un estrato es mayor, se extrae una cantidad más grande de unidades para compensar la heterogeneidad
- El estrato representa una gran parte de la población
- El muestreo en el estrato es menos costoso que en otros estratos

Como se puede observar en la ecuación (35), el vector de unidades muestrales $\mathbf{n}$ está expresada en términos de n, donde n, la muestra global no se conoce.

La solución depende de, si la muestra es seleccionada para satisfacer un costo especifico C, o para una precisión deseada V_0 asignada a la varianza $\text{Var}(\bar{y}_{st})$.

Para el primer caso, si el costo es fijo, sustituimos los valores óptimos de $\mathbf{n}$ desde la ecuación (35), en la función de costo (28)

$$
\begin{aligned}
C &= c_0 + \mathbf{c}' \left(\frac{n}{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})} \begin{bmatrix} \frac{W_1 S_1}{\sqrt{c_1}} \\ \vdots \\ \frac{W_H S_H}{\sqrt{c_H}} \end{bmatrix} \right) \\
C - c_0 &= \frac{n}{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})} \ \mathbf{c}' \left(\mathbf{W} \odot \mathbf{S} \odot \mathbf{c}^{1/2(-)}\right) \\
&= \frac{n}{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})} \ \mathbf{c}^{(1/2)'}(\mathbf{W} \odot \mathbf{S}),
\end{aligned}
$$

y resolviendo para n se tiene

$$
n = \frac{(C - c_0)\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})}{\mathbf{c}^{1/2'}(\mathbf{W} \odot \mathbf{S})}. \tag{36}
$$

Si la decisión es que la varianza sea fija para minimizar la función de costo, se utiliza la ecuación (30) para la varianza y se sustituye el $\mathbf{n}$ óptimo a partir de la ecuación (35)

$$
\begin{aligned}
\mathrm{Var}(\bar{y}_{st}) &= \mathbf{W}' \left[\mathrm{diag}\left(\frac{S_h^2}{n_h}\right)\right] \mathbf{W} - \frac{1}{N}(\mathbf{S}^2)'\mathbf{W} \\
\mathrm{Var}(\bar{y}_{st}) + \frac{1}{N}(\mathbf{S}^2)'\mathbf{W} &= \mathbf{W}' \left[\mathrm{diag}\left(\mathbf{S}^2 \odot \mathbf{n}^{(-)}\right)\right] \mathbf{W} \\
&= \mathbf{W}' \left[\mathrm{diag}\left(\mathbf{S}^2 \odot \left[\frac{n}{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})} \left(\mathbf{W} \odot \mathbf{S} \odot \mathbf{c}^{1/2(-)}\right)\right]^{(-)}\right)\right] \mathbf{W} \\
&= \frac{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})}{n} \mathbf{W}' \left[\mathrm{diag}\left(\mathbf{S}^2 \odot \left[\mathbf{W} \odot \mathbf{S} \odot \mathbf{c}^{1/2(-)}\right]^{(-)}\right)\right] \mathbf{W}
\end{aligned}
$$

despejando a n, se obtiene

$$
\begin{aligned}
n &= \frac{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})}{\mathrm{Var}(\bar{y}_{st}) + \frac{1}{N}(\mathbf{S}^2)'\mathbf{W}} \left(\mathbf{W}' \left[\mathrm{diag}\left(\mathbf{S}^2 \odot \left[\mathbf{W} \odot \mathbf{S} \odot \mathbf{c}^{1/2(-)}\right]^{(-)}\right)\right] \mathbf{W}\right) \\
&= \frac{\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2(-)})\mathbf{W}'(\mathbf{S} \odot \mathbf{c}^{1/2})}{\mathrm{Var}(\bar{y}_{st}) + \frac{1}{N}(\mathbf{S}^2)'\mathbf{W}}
\end{aligned} \tag{37}
$$

Un caso particular surge cuando el costo por unida es el mismo en todos los estratos, esto es que $c = c_h$, entonces la función de costo se limita a

$$c_0 + \mathbf{c}^{'}\mathbf{n} = c_0 + [c, \ldots, c] \begin{bmatrix} n_1 \\ \vdots \\ n_H \end{bmatrix} = c_0 + cn \tag{38}$$

y la asignación óptima para un costo fijo se reduce a la asignación óptima para un tamaño de muestra fijo, por lo que para este caso particular se tiene el siguiente teorema.

Teorema 8. *En el muestreo aleatorio estratificado, la varianza $Var(\bar{y}_{st})$ se minimiza para un tamaño de muestra total fijo n si*

$$\mathbf{n} = \frac{n}{\mathbf{W}'\mathbf{S}}(\mathbf{W} \odot \mathbf{S}). \tag{39}$$

Este tipo de asignación es conocida como **Asignación de Neyman**, por Neyman (1934) cuya prueba dió promidencia al resultado. Más tarde se descubrió una prueba anterior de Tschuprow (1923), vea Cochran (1977).

La expresión para la varianza mínima con n fija, es obtenida sustituyendo el valor de $\mathbf{n}$ a partir de (39), en la ecuación de la varianza $\mathrm{Var}(\bar{y}_{st})$, de donde se obtiene el siguiente resultado;

$$\begin{aligned} \mathrm{Var}(\bar{y}_{st}) &= \mathbf{W}' \left[\mathrm{diag}\left(\mathbf{S}^2 \odot \mathbf{n}^{(-)}\right)\right] \mathbf{W} - \frac{1}{N}(\mathbf{S}^2)'\mathbf{W} \\ &= \frac{\mathbf{W}'\mathbf{S}}{n}\mathbf{W}' \left[\mathrm{diag}\left(\mathbf{S}^2 \odot (\mathbf{W} \odot \mathbf{S})^{(-)}\right)\right] \mathbf{W} - \frac{1}{N}(\mathbf{S}^2)'\mathbf{W} \\ &= \frac{1}{n}(\mathbf{W}'\mathbf{S})^2 - \frac{1}{N}(\mathbf{S}^2)'\mathbf{W}. \end{aligned} \tag{40}$$

La asignación de Neyman es un caso particular de la distribución óptima, la cual se utiliza cuando los costos de los estratos son muy similares o iguales, y no cuando las varianzas lo son.

Con esta asignación, si las varianzas $\mathbf{S}^2$ son correctamente especificadas, entonces se proporciona estimaciones con menor varianza que en la asignación proporcional, vea Lohr (2000).

Finalmente observe que, realmente el problema que se tienen en la asignación óptima al minimizar la varianza sujeta a la función de costos, es

$$\min_{\mathbf{n}} \operatorname{Var}(\bar{y}_{st})$$

$$s.a$$

$$\mathbf{c}^{'}\mathbf{n} + c_0 = C$$

$$n_h \in \mathbb{N}, h = 1, 2, \ldots, H,$$

donde $\mathbb{N} = \{n | n = 0, 1, 2, \ldots\}$, el conjunto de los números naturales. De igual forma cuando se minimizan la función de costos sujeta a una varianza fija, el problema es

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

$$s.a$$

$$\mathbf{W}' \left[\operatorname{diag}\left(\frac{S_h^2}{n_h} \right) \right] \mathbf{W} - \tfrac{1}{N}(\mathbf{S}^2)'\mathbf{W} \leq V_0$$

$$n_h \in \mathbb{N}, h = 1, 2, \ldots, H,$$

Los cuales realmente son problemas de programación no lineal de enteros, para los cuales existen algoritmos adecuados, vea Rao (1978).

La solución particular de estos problemas de enteros, será presentada más adelante en el ejemplo dentro del capitulo IV.

Ejemplo de asignación óptima.

Se tiene 187 datos (ver Apéndice B), los cuales representan la medición de la longitud de una hoja en cierta "flor". La población fue dividida en 4 estratos, dado a que la "flor" se cultiva en 4 diferentes regiones.

Problema 1

El problema que se desea resolver, es el de encontrar los valores óptimos de n_h con $h = 1, 2, 3, 4$ para una función de costo con $C = 5000$, que es el presupuesto con el que se cuenta para poder realizar el muestreo. El costo por unidad para los estratos son; $c_1 = 25$, $c_2 = 30$, $c_3 = 15$, y $c_4 = 25$ pesos, el costo adicional es de $c_0 = 0$.

Este problema de optimización esta dado por

$$\min_{\mathbf{n}} \widehat{\mathrm{Var}}(\bar{y}_{st})$$

$$s.a$$

$$\mathbf{c}^{'}\mathbf{n} = 5000$$

Donde

$$\widehat{\mathrm{Var}}(\bar{y}_{st}) = \mathbf{W}^{'} \left[\mathrm{diag}\left(\frac{s_h^2}{n_h} \right) \right] \mathbf{W} - \frac{1}{N}(\mathbf{s}^2)^{'}\mathbf{W}$$

Los valores de los vectores se encuentras definidos en el Apéndice B.

Utilizando la ecuación (36), se tiene que el tamaño de la muestra es de $n = 189.6944$, sustituyendo este valor en la ecuación (35), y despejando el vector $\mathbf{n}$, se obtienen los valores óptimos para el tamaño de muestra en cada estrato, los cuales se muestran en el siguiente cuadro.

Cuadro 2.1: Soluciones al problema de asignación óptima en muestreo estratificado bajo los diferentes enfoques de optimización para el problema 1.

	Solución Lagrange	Solución Ajustada [a]
n_1	13.7976	14
n_2	81.7554	82
n_3	15.1137	15
n_4	79.0277	79
$\sum_{h=1}^{H} n_h$	189.6944	190
$f(\mathbf{n})$ [b]	0.0096	0.0096

[a] Esta solución se obtiene a partir de la solución de Lagrange, redondeando cada n_i, vea Cochran (1977). Alternativamente, en la literatura estadística también se suele ajustar al entero inmediato superior.

[b] Donde $f(\mathbf{n}) = \widehat{Var}(\bar{y}_{st}) = \sum_{h=1}^{H} W_h^2 \frac{s_h^2}{n_h} - \sum_{h=1}^{H} W_h \frac{s_h^2}{N}$

La mayor cantidad de unidades asignadas a los tamaños de muestra n_h ocurre en los estratos 2 y 4, los costos por unidad en estos estratos no son más económicos, pero son los que presentan una mayor varianza (ver Apéndice B). Los estratos 1 y 4 tiene el mismo costo por unidad, pero el estrato 1 tiene una varianza mas pequeña, en cuanto al estrato 3, aunque es más económico que los demás estratos, la varianza en este estrato es mucho menor, es por esto que se asignan pocas unidades muestrales a este estrato.

Problema 2

En el segundo problema, el interés es el de obtener los valores óptimos n_h, para una precisión $V_0 = .015$. En está situación el problema de optimización esta dado por

$$\min_{\mathbf{n}} \mathbf{c}^{'}\mathbf{n}$$

$$s.a$$

$$\mathbf{W}' \left[\text{diag} \left(\frac{s_h^2}{n_h} \right) \right] \mathbf{W} - \frac{1}{N} (\mathbf{s}^2)^{'} \mathbf{W} \leq 0.015$$

Utilizando la ecuación (37), se obtiene un tamaño de muestra $n = 122.872$, sustituyendo el valor de n en la ecuación (35), y despejando $\mathbf{n}$, se obtienen los tamaño de muestra óptimos, los cuales se resumen en el siguiente cuadro.

Cuadro 2.2: Soluciones al problema de asignación óptima en muestreo estratificado bajo los diferentes enfoques de optimización para el problema 2.

	Solución Lagrange	Solución Ajustada [a]
n_1	8.9372	9
n_2	52.9560	53
n_3	9.7879	10
n_4	51.1891	51
$n = \sum_{h=1}^{H} n_h$	122.8722	123
$f(\mathbf{n})$ [b]	3238.6820	3240

[a] Esta solución se obtiene a partir de la solución de Lagrange, redondeando cada n_i, vea Cochran (1977). Alternativamente, en la literatura estadística también se suele ajustar al entero inmediato superior.

[b] Donde $f(\mathbf{n}) = \sum_{h=1}^{H} c_h n_h$

Observe que el mayor número de unidades asignadas en los tamaños de muestra n_h, corresponden a los estratos 2 y 4, que son los que tiene una mayor varianza, en comparación de los estratos 1 y 3 que tienen varianzas muy pequeñas.

III. OPTIMIZACIÓN ESTOCÁSTICA

Introducción

La optimización estocástica es una rama de la programación matemática, ya que trata de resolver problemas de optimización donde algunos de los parámetros o variables son de carácter aleatorio.

La aleatoriedad está presente en diversos problemas y, hasta hace poco tiempo, es posible dar soluciónes a problemas de optimización en grandes sistemas considerando explícitamente el cáracter aleatorio de las variables y/o parámetros de forma.

La aleatoriedad puede deberse a diversos motivos, errores de medición, carencia de datos confiables, naturaleza de las variables (que por si mismas sean variables aleatorias), o por tratarse de parámetros o variables que representan información sobre el futuro, es decir estimaciones.

En la optimización determinística se supone que los parámetros del problema son conocidos con certeza, o por lo menos en su valor medio. Mientras que en optimización estocástica esta condición es relajada, no se conocen los valores, sólo sus distribuciones o el supuesto de éstas.

La optimización estocástica aparece en 1955, como una extensión de la programación lineal con enfásis en problemas con un gran número de variables y parámetros, Beale (1955); Dantzig (1955). Por otra parte, como una extensión de la programación lineal para grandes sistemas con estructuras especiales, aparecen las técnicas de descomposición Benders (1962), Datzing (1963) y Van Slyke and Wets (1969), también denominadas como optimización matemática a gran escala, vea Ramos y Cerisela (2005).

Aunque estas investigaciones datan de hace más de cinco décadas, es recientemente que el avance en la tecnología de las computadoras ha permitido la solución a problemas de gran tamaño y ha devuelto el interés al tema de la optimización estocástica, produciendo además un avance en la teoría matemática que la sustenta.

En estadística y probabilidad, el uso de la optimización estocástica como una herramienta para el planteamiento y solución de problemas tampoco es de reciente aparición; en la literatura sobre el tema, Prékopa (1978) ha planteado diferentes problemas, tales como, problemas de pruebas de hipótesis, obtención de regiones de confianza, determinación de regiones de tolerancia y su aplicación a procesos estocásticos.

Formalmente, considere el siguiente problema de optimización

$$\begin{aligned} &\min_{x} h(\mathbf{x}, \boldsymbol{\xi}_0) \\ &s.a \quad \text{(sujeto a)} \\ &g_1(\mathbf{x}, \boldsymbol{\xi}_1) \leq 0 \\ &g_2(\mathbf{x}, \boldsymbol{\xi}_2) \leq 0 \\ &\vdots \\ &g_m(\mathbf{x}, \boldsymbol{\xi}_m) \leq 0, \end{aligned} \tag{41}$$

donde $\mathbf{x} \in \mathbb{R}^n$ (variables de decisión), y $\boldsymbol{\xi}_j \in \mathbb{R}^{K_j}$ (parámetros de forma), $j = 0, 1, \ldots, m$. Si $\mathbf{x}$ y/o $\boldsymbol{\xi}_j$ son de carácter aleatorio, entonces este problema se define como un problema de optimización estocástica.

Esencialmente, existen dos enfoques para resolver un problema de optimización estocástica. En el primer enfoque se extienden las técnicas determinísticas al caso estocástico, ejemplos de estos procedimientos son: La técnica de direcciones factibles de Ruszczynski, el método de la proyección del subgradiente, las funciones de Lagrange extendidas y el procedimiento minimax, por mencionar algunas. Otro enfoque, son aquellos métodos basados en ideas estadísticas y probabilísticas que abordan el problema estocástico a través del planteamiento de un problema determinístico equivalente. Se dice que es equivalente en el sentido de que la solución del nuevo problema determinístico es una solución del problema estocástico original. Algunos de los métodos que trabajan con estas ideas son el E-Modelo, V-Modelo y P-Modelo, entre otros. Como se puede observar, este es un tema de gran amplitud, y resultaría muy complicado el abordarlo en su totalidad. Es por ello que este capítulo únicamente aborda la descripción de algunos de los métodos basados en ideas estadísticas y probabilísticas, donde se muestra como transformar un problema estocástico en un problema determinístico equivalente. Una vez transformado el problema estocástico en determinístico, este puede ser resuelto mediante cualquiera de los métodos de optimización determinísticos conocidos, tales como, programación lineal, geométrica, dimámica, nolineal, etc.

Optimización lineal estocástica

Es frecuente que en optimización estocástica, se pueda tener una o más variables aleatorias en un mismo problema. Esta sección tiene como propósito estudiar algunos de estos casos, la forma en como se plantean las soluciones incluso donde simultáneamente se presentan un conjunto de variables aleatorias. Para comprender lo comentado, considere el siguiente planteamiento.

El problema de optimización que se desea resolver es el siguiente:

$$\min_{\mathbf{x}} \mathbf{c}^{'}\mathbf{x} \tag{42}$$

$$s.a$$

$$\mathrm{P}\left[\mathbf{a}_i^{'}\mathbf{x} \leq b_i\right] \geq p_i, \quad i = 1, 2, \ldots, m \tag{43}$$

$$x_j \geq 0, \quad j = 1, 2, \ldots, n. \tag{44}$$

Donde $\mathbf{c} = (c_1, c_2, \ldots, c_n)^{'}$, $\mathbf{a_i} = (a_{i1}, a_{i2}, \ldots, a_{in})^{'}$ y b_i son variables aleatorias, p_i es una probabilidad especificada y x_j denota la variable de decisión, la cual se considera determinística. Note que la ecuación (43) indica la i-ésima restricción,

$$\mathbf{a_i}^{'}\mathbf{x} \leq b_i$$

la cual tiene que ser satisfecha con probabilidad de al menos p_i, donde $0 \leq p_i \leq 1$, vea Rao (1978, Pág. 597).

Para abordar la solución a este problema, primero se presenta el caso particular cuando sólo las $\mathbf{a_i}$ son aleatorias, donde se supone que las demás variables y parámetros son no aleatorios. Posteriormente se estudia los casos particulares para cuando sólo los b_i o las c_j

son variables aleatorias, y finalmente se presentara el caso donde simultáneamente a_{ij}, b_i y c_j son aleatorias.

Caso 1.

Cuando sólo las a_{ij} son aleatorias. Suponga que

$$\mathrm{E}(\mathbf{a}_i) = \bar{\mathbf{a}}_i = \begin{bmatrix} \bar{a}_{i1} \\ \vdots \\ \bar{a}_{in} \end{bmatrix}, \quad i = 1, 2, \ldots, m. \tag{45}$$

y

$$\mathrm{Cov}(\mathbf{a}_i) = \mathbf{V}_i = \begin{bmatrix} \mathrm{Var}(a_{i1}) & \mathrm{Cov}(a_{i1}, a_{i2}) & \ldots & \mathrm{Cov}(a_{i1}, a_{in}) \\ \mathrm{Cov}(a_{i2}, a_{i1}) & \mathrm{Var}(a_{i2}) & \ldots & \mathrm{Cov}(a_{i2}, a_{in}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(a_{in}, a_{i1}) & \mathrm{Cov}(a_{in}, a_{i2}) & \ldots & \mathrm{Var}(a_{in}) \end{bmatrix}. \tag{46}$$

Entonces,

$$\mathrm{E}(\mathbf{a}_i^{'}\mathbf{x}) = \mathrm{E}(\mathbf{a}_i^{'})\mathbf{x} = \bar{\mathbf{a}}_i^{'}\mathbf{x} \tag{47}$$

y

$$\mathrm{Var}(\mathbf{a}_i^{'}\mathbf{x}) = \mathrm{Var}(\mathbf{x}^{'}\mathbf{a}_i) = \mathbf{x}^{'}\mathrm{Cov}(\mathbf{a}_i)\mathbf{x} = \mathbf{x}^{'}\mathbf{V}_i\mathbf{x}. \tag{48}$$

Ahora, bajo la suposición de que $\mathbf{a}_i$ está normalmente distribuida para todo i, se tiene que

$$\mathrm{P}[\mathbf{a}_i^{'}\mathbf{x} \leq b_i] \geq p_i, \tag{49}$$

estandarizando la ecuación (49)

$$\mathrm{P}\left[\frac{\mathbf{a}_i^{'}\mathbf{x} - \bar{\mathbf{a}}_i^{'}\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{a}_i^{'}\mathbf{x})}} \leq \frac{b_i - \bar{\mathbf{a}}_i^{'}\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{a}_i^{'}\mathbf{x})}}\right] \geq p_i, \quad i = 1, 2, \ldots, m. \tag{50}$$

Recordando que, $\mathrm{P}(Y \leq y) = \mathrm{F}_y(y) = \Phi(y)$ es la función de distribución acumulativa normal, entonces la ecuación (50) puede ser escrita de la siguiente forma

$$\mathrm{P}\left[\frac{\mathbf{a}_i'\mathbf{x} - \bar{\mathbf{a}}_i'\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{a}_i'\mathbf{x})}} \leq \frac{b_i - \bar{\mathbf{a}}_i'\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{a}_i'\mathbf{x})}}\right] = \Phi\left(\frac{b_i - \bar{\mathbf{a}}_i'\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{a}_i'\mathbf{x})}}\right). \tag{51}$$

Sea e_i el valor de la variable aleatoria tal que $\Phi(e_i) = p_i$, entonces utilizando la ecuación (50) y (51) se tiene que

$$\Phi\left(\frac{b_i - \bar{\mathbf{a}}_i'\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{a}_i'\mathbf{x})}}\right) \geq \Phi(e_i), \quad i = 1, 2, \ldots, m. \tag{52}$$

Considerando que $\Phi(v) \geq \Phi(w)$ sólo si se cumple que $v \geq w$, entonces de forma equivalente, se tiene

$$\frac{b_i - \bar{\mathbf{a}}_i'\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{a}_i'\mathbf{x})}} \geq e_i,$$

o alternativamente

$$\bar{\mathbf{a}}_i'\mathbf{x} + e_i\sqrt{\mathbf{x}'\mathbf{V}_i\mathbf{x}} - b_i \leq 0, \quad i = 1, 2, \ldots, m.$$

De tal forma que, para este caso particular, la solución para el problema de optimización estocástica (42-44) puede ser obtenida resolviendo el problema determinístico equivalente, el cual está dado por

$$\begin{gathered}\min_{\mathbf{x}} f(x) = \mathbf{c}'\mathbf{x} \\ s.a \\ \bar{\mathbf{a}}_i'\mathbf{x} + e_i\sqrt{\mathbf{x}'\mathbf{V}_i\mathbf{x}} - b_i \leq 0, \quad i = 1, 2, \ldots, m \\ x_j \geq 0, \quad j = 1, 2, \ldots, n\end{gathered} \tag{53}$$

Este problema puede ser resuelto mediante algún método de optimización tradicional.

Caso 2.

Para el caso donde solo los b_i son aleatorios, el planteamiento es el siguiente. Primero, suponga que $\mathbf{b} \sim N_m(\bar{\mathbf{b}}, \mathrm{Cov}(\mathbf{b}))$, donde $\mathbf{b} = (b_1, \ldots, b_m)'$ entonces

$$\begin{aligned} \mathrm{p}[\mathbf{a}_i'\mathbf{x} \leq b_i] &= \mathrm{p}\left[\frac{\mathbf{a}_i'\mathbf{x} - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \leq \frac{b_i - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}\right] \\ &= \mathrm{p}\left[\frac{b_i - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \geq \frac{\mathbf{a}_i'\mathbf{x} - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}\right] \\ &= 1 - \mathrm{p}\left[\frac{b_i - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \leq \frac{\mathbf{a}_i'\mathbf{x} - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}\right] \\ &= 1 - \Phi\left(\frac{\mathbf{a}_i'\mathbf{x} - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}\right). \end{aligned} \tag{54}$$

Donde la restricción del problema (43) puede ser escrita de la siguiente forma

$$-\Phi\left(\frac{\mathbf{a}_i'\mathbf{x} - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}\right) \geq -1 + p_i$$

o

$$\Phi\left(\frac{\mathbf{a}_i'\mathbf{x} - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}\right) \leq 1 - p_i. \tag{55}$$

De forma similar al caso 1. Sea E_i el valor de la variable aleatoria normal estándar, tal que $\Phi(E_i) = 1 - p_i$, de forma que la desigualdad (55) es válida sólo si se cumple

$$\frac{\mathbf{a}_i'\mathbf{x} - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \leq E_i, \quad i = 1, 2, \ldots, m$$

o

$$\mathbf{a}_i'\mathbf{x} - \bar{b}_i - E_i\sqrt{\mathrm{Var}(b_i)} \leq 0, \quad i = 1, 2, \ldots, m. \tag{56}$$

Entonces para este caso en particular, el problema determinístico equivalente al problema estocástico (42-44) está dado por

$$\begin{gathered}\min_{\mathbf{x}} f(x) = \mathbf{c}'\mathbf{x} \\ s.a \\ \mathbf{a}_i'\mathbf{x} - \bar{b}_i - E_i\sqrt{\mathrm{Var}(b_i)} \leq 0, \quad i = 1, 2, \ldots, m \\ x_j \geq 0, \quad j = 1, 2, \ldots, n\end{gathered} \tag{57}$$

Caso 3.

Cuando sólo las variables c_j son aleatorias, implica que la función objetivo es estocástica y las restricciones son fijas. La solución esta dada mediante el siguiente planteamiento. Suponga que $\mathbf{c} \sim N_n(\bar{c}, \mathbf{V})$ donde $\mathbf{c} = (c_1, c_2, \ldots, c_n)'$, $\mathrm{E}(\mathbf{c}) = \bar{\mathbf{c}}$ y $\mathbf{V} = \mathrm{Cov}(\mathbf{c})$. Entonces la función objetivo $f(x)$ también tiene una distribución normal. Más aún

$$\mathrm{E}(f(x)) = \mathrm{E}(\mathbf{c}'\mathbf{x}) = \mathrm{E}(\mathbf{c}')\mathbf{x} = \bar{\mathbf{c}}'\mathbf{x}$$

y

$$\mathrm{Var}(f(x)) = \mathrm{Cov}(\mathbf{c}'\mathbf{x}) = \mathrm{Cov}(\mathbf{x}'\mathbf{c}) = \mathbf{x}'\mathrm{Cov}(\mathbf{c})\mathbf{x} = \mathbf{x}'\mathbf{V}\mathbf{x}. \tag{58}$$

Entonces, la nueva función objetivo a minímizar se establece como

$$F(x) = k_1\bar{\mathbf{c}}'\mathbf{x} + k_2\sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}}. \tag{59}$$

Donde k_1 y k_2 son constantes no negativas. Las cuales significan la ponderación o importancia relativa de $\mathrm{E}(f(x))$ y $(\mathrm{Cov}(f(x)))^{1/2}$ en el problema a minimizar. Si $k_1 = k_2 = 1$ indica que, tanto el valor esperado $\mathrm{E}(f(x))$, como la desviación estándar $(\mathrm{Cov}(f(x)))^{1/2}$ son de igual importancia, Rao (1978).

De esta forma el problema determinístico equivalente al problema estocástico (42-44), para este caso esta dado por

$$\begin{aligned} \min_{\mathbf{x}} F(x) &= k_1\bar{\mathbf{c}}'\mathbf{x} + k_2\sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \\ &s.a \\ \mathbf{a}_i'\mathbf{x} - b_i &\leq 0, \quad i = 1, 2, \ldots, m \\ x_j &\geq 0, \quad j = 1, 2, \ldots, n \end{aligned} \tag{60}$$

Caso 4.

Finalmente, se aborda el caso donde simultáneamente $\mathbf{a}_i$,b_i, y c_j son aleatorios. Las variables c_j solo aparecen en la función objetivo. Tal como se vió en el caso 3, la nueva función objetivo está dada por

$$F(x) = k_1\bar{\mathbf{c}}'\mathbf{x} + k_2\sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}}.$$

Ahora, la función para las restricciones, puede ser escrita de la siguiente forma

$$\mathrm{P}[h_i \leq 0] \geq p_i, \qquad i = 1, 2, \ldots, m. \tag{61}$$

Donde h_i son las nuevas variables aleatorias, las cuales de definen como

$$h_i = \mathbf{a}_i'\mathbf{x} - b_i = \sum_{k=1}^{n+1} q_{ik}y_k$$

con

$$\begin{aligned} q_{ik} &= a_{ik}, \quad k = 1, 2, \ldots, n \\ q_{i,n+1} &= b_i \\ y_k &= x_k, \quad k = 1, 2, \ldots, n \\ y_{n+1} &= -1. \end{aligned}$$

Las variables h_i son una combinación lineal de variables aleatorias normales q_{ik}, por lo que h_i también tiene una distribución normal, más aún

$$\mathrm{E}(h_i) = \bar{h}_i = \sum_{k=1}^{n+1} \mathrm{E}(q_{ik}) y_k = \sum_{j=1}^{n} \bar{a}_{ij} x_j - \bar{b}_i$$

y $\mathrm{Var}(h_i) = \mathbf{Y}' \mathbf{V_i} \mathbf{Y}$, específicamente

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n+1} \end{bmatrix}$$

y

$$\mathbf{V}_i = \begin{bmatrix} \mathrm{Var}(q_{i1}) & \mathrm{Cov}(q_{i1}, q_{i2}) & \dots & \mathrm{Cov}(q_{i1}, q_{in+1}) \\ \mathrm{Cov}(q_{i2}, q_{i1}) & \mathrm{Var}(q_{i2}) & \dots & \mathrm{Cov}(q_{i2}, q_{in+1}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(q_{i,n+1}, q_{i1}) & \mathrm{Cov}(q_{i,n+1}, q_{i2}) & \dots & \mathrm{Var}(q_{i,n+1}) \end{bmatrix}.$$

Además se observa que

$$\mathrm{Var}(h_i) = \sum_{j=1}^{n} \left[x_j^2 \mathrm{Var}(a_{ij}) + 2 \sum_{l=j+1}^{n} x_j x_l \mathrm{Cov}(a_{ij}, a_{il}) \right] + \mathrm{Var}(b_i) - 2 \sum_{j=1}^{n} x_j \mathrm{Cov}(a_{ij}, b_i).$$

Definido h_i, y estandarizando la ecuación (61) la expresión para las restricciones puede escribirse como

$$\mathrm{P} \left[\frac{h_i - \bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}} \leq \frac{-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}} \right] \geq p_i, \quad i = 1, 2, \dots, m.$$

De nuevo, sea e_i el valor de la variable aleatoria normal tal que $\Phi(e_i) = p_i$, entonces

$$\Phi \left(\frac{-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}} \right) \geq \Phi(e_i), \quad i = 1, 2, \dots, m.$$

Lo cual es cierto si se cumple

$$\frac{-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}} \geq e_i, \quad i = 1, 2, \ldots, m$$

o equivalentemente

$$\bar{h}_i + e_i\sqrt{\mathrm{Var}(h_i)} \leq 0, \quad i = 1, 2, \ldots, m.$$

De donde finalmente, el problema determinístico equivalente al problema de optimización estocástica (42-44) esta planteado como

$$\begin{gathered} \min_{\mathbf{x}} F(x) = k_1\bar{\mathbf{c}}'\mathbf{x} + k_2\sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \\ s.a \\ \bar{h}_i + e_i\sqrt{\mathrm{Var}(h_i)} \leq 0, \quad i = 1, 2, \ldots, m \\ x_j \geq 0, \quad j = 1, 2, \ldots, n. \end{gathered} \tag{62}$$

E-Modelo.

Anteriormente se estudió, el procedimiento para cuando las c_j son aleatorias, el cual es también conocido como el E-Modelo Modificado, ahí se vió que k_1 y k_2 son constantes que indican la importancia relativa para $\mathrm{E}(f(x))$ y $(\mathrm{Var} f(x))^{1/2}$. El E-Modelo es un caso especial del E-Modelo modificado, en el que se establece que $k_1 = 1$ y $k_2 = 0$, el planteamiento bajo el E-Modelo esta dado por

$$\begin{gathered} \min_{\mathbf{x}} \mathrm{E}(f(x)) = \mathrm{E}(\mathbf{c}'\mathbf{x}) = \bar{\mathbf{c}}'\mathbf{x} \\ s.a \\ \mathrm{P}[\mathbf{a}_i'x \leq b_i] \geq p_i, \quad i = 1, 2, \ldots, m \\ x_j \geq 0, \quad j = 1, 2, \ldots, n. \end{gathered} \tag{63}$$

Optimización con mínimo riesgo

Considere nuevamente el problema de optimización estocástica (3.2-3.4), en el que las restricciones son determinísticas o bien se han sustituido por equivalentes determinísticas, de tal forma que el problema del que se parte es

$$\begin{gathered}
\min_{\mathbf{x}} f(x) = \mathbf{c}'\mathbf{x} \\
s.a \\
\mathbf{a}_i'\mathbf{x} - b_i \leq 0, \quad i = 1, 2, \ldots, m \\
x_j \geq 0, \quad j = 1, 2, \ldots, n.
\end{gathered} \tag{64}$$

Con $\mathbf{c} \sim N_n(\bar{\mathbf{c}}, \mathbf{V})$.

V-Modelo.

(o modelo de mínima varianza). El objetivo consiste en minimizar la varianza de la función objetivo estocástica $f(x)$, obteniendo el siguiente problema determinístico equivalente

$$\begin{gathered}
\min_{\mathbf{x}} \operatorname{Var}(f(x)) = \mathbf{x}'\mathbf{V}\mathbf{x} \\
s.a \\
\mathbf{a}_i'\mathbf{x} - b_i \leq 0, \quad i = 1, 2, \ldots, m \\
x_j \geq 0, \quad j = 1, 2, \ldots, n.
\end{gathered} \tag{65}$$

Como se puede observar, el V-Modelo es un caso especial del E-Modelo modificado, donde se considera que $k_1 = 0$ y $k_2 = 1$.

P-Modelo.

(o modelo de máxima probabilidad). Dentro de los diversos modelos que existen dentro del enfoque denominado de máxima probabilidad, se estudia el modelo de mínimo riesgo k, el cual busca maximizar la probabilidad de que la función objetivo estocástica sea menor que un valor fijo k. De este modo, la función objetivo es transformada en

$$\max_{x} \mathrm{P}(f(x) = \mathbf{c}^{'}\mathbf{x} \leq k).$$

Recordando que $\mathbf{c} \sim N_n(\bar{\mathbf{c}}, \mathbf{V})$ entonces se tiene que $\mathbf{c}^{'}\mathbf{x} \sim N(\bar{\mathbf{c}}^{'}\mathbf{x}, \mathbf{x}^{'}\mathbf{V}\mathbf{x})$, estandarizando la función objetivo se tiene

$$\begin{aligned} & \max_{x} \mathrm{P}\left[\frac{\mathbf{c}^{'}\mathbf{x} - \bar{\mathbf{c}}^{'}\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{c}^{'}\mathbf{x})}} \leq \frac{k - \bar{\mathbf{c}}^{'}\mathbf{x}}{\sqrt{\mathrm{Var}(\mathbf{c}^{'}\mathbf{x})}}\right] \\ = & \max_{x} \Phi\left(\frac{k - \bar{\mathbf{c}}^{'}\mathbf{x}}{\sqrt{\mathbf{x}^{'}\mathbf{V}\mathbf{x}}}\right) \\ = & \max_{x} \mathrm{F}_{c^{'}x}(k). \end{aligned}$$

En este caso, el problema determinístico está dado por

$$\begin{aligned} & \max_{\mathbf{x}} \mathrm{F}_{c^{'}x}(k) \\ & s.a \\ & \mathbf{a}_i^{'}\mathbf{x} - b_i \leq 0, \quad i = 1, 2, \ldots, m \\ & x_j \geq 0, \quad j = 1, 2, \ldots, n. \end{aligned} \tag{66}$$

Aquí es importante notar, que maximizar una función de distribución es equivalente a maximizar el valor de su variable, de tal forma que el problema determinístico puede ser expresado como

$$\begin{aligned} & \max_{\mathbf{x}} \frac{k - \bar{\mathbf{c}}^{'}x}{\sqrt{\mathbf{x}^{'}\mathbf{V}\mathbf{x}}} \\ & s.a \\ & \mathbf{a}_i^{'}\mathbf{x} - b_i \leq 0, \quad i = 1, 2, \ldots, m \\ & x_j \geq 0, \quad j = 1, 2, \ldots, n. \end{aligned} \tag{67}$$

Optimización estocástica no lineal.

Cuando el problema de optimización estocástica es no lineal, este puede ser resuelto mediante dos enfoques:

1. El primero de ellos, es mediante la expansión de las funciones que aparecen como restricciones y como función objetivo, alrededor del vector de medias. De manera que el problema a minimizar es de la forma

$$\begin{gathered} \min_{\mathbf{x}} f(\mathbf{y}) \\ s.a \\ \mathrm{P}[g_i(\mathbf{y}) \leq 0] \geq p_i, \quad i = 1, 2, \ldots, m. \end{gathered} \tag{68}$$

   Donde el vector $\mathbf{y} = (y_1, y_2, \ldots, y_N)'$ incluye las variables de decisión $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Entonces

$$f(y) \approx f(\bar{\mathbf{y}}) + \nabla' f(\bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})$$

   y

$$g_i(y) \approx g_i(\bar{\mathbf{y}}) + \nabla' g_i(\bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})$$

   Dadas estas funciones, es posible aplicar cualquiera de los métodos descritos anteriormente.

2. En un segundo enfoque, existen problemas de optimización estocástica no lineal para los cuales los métodos estudiados como el E-Modelo, V-Modelo y P-Modelo, se pueden aplicar casi sin modificación. Se introduce este enfoque através del siguiente ejemplo;

En el contexto de la métodologia de superficies de respuesta, típicamente se está interesado en optimizar un modelo estimado de segundo orden de la forma

$$\min_{\mathbf{X}} \hat{Y}(\mathbf{X}, \hat{\beta})$$

$$s.a$$

$$\|\mathbf{X}\|^2 = X_1^2 + X_2^2 + \ldots + X_k^2 \leq c^2.$$

Donde c es una constante fija conocida, y

$$\hat{Y}(\mathbf{X}, \hat{\beta}) = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i X_i + \sum_{i=1}^{k} \hat{\beta}_{ii} X_i^2 + \sum\sum_{i \neq j} \hat{\beta}_{ij} X_i x_j \equiv \mathbf{Z}'(\mathbf{X})\hat{\beta}$$

Donde $\mathbf{Z}'(\mathbf{X})$ contiene los correspondientes términos polinomiales en las variables $X_1, X_2, \ldots, X_k$ y $\hat{\beta}$ es el estimador de mínimos cuadrados de los parámetros de regresión.

En este modelo se considera que $Y_i = \mathbf{Z}'(X_i)\beta + \varepsilon_i$, con $\varepsilon_i \sim N(0, \sigma^2)$ independientes, $i = 1, 2, \ldots, n$. Entonces $\hat{\beta} \sim N_p(\beta, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$, $p = 1 + 2k + \dfrac{k(k+1)}{2}$, y $\dfrac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p}$ donde $\hat{\sigma}^2$ es el estimador insesgado de σ^2 además $\mathbf{X} \in \mathbb{R}^{n \times p}$, es la matriz de diseño de rango p. Entonces, el problema puede plantearse como

$$\begin{gathered} \min_{\mathbf{X}} \hat{Y}(\mathbf{X}, \hat{\beta}) \\ s.a \\ \|\mathbf{X}\|^2 \leq c^2, \end{gathered} \tag{69}$$

con

$$\hat{\beta} \sim N_p(\beta, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}) \text{ y } \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p}.$$

Considerando este planteamiento, el problema determinístico equivalente al estocástico, pueden ser estudiado bajo los métodos descritos anteriormente.

Note que

$$\hat{\mathbf{Y}}(\mathbf{X},\hat{\beta}) = \mathbf{Z}^{'}(\mathbf{X})\hat{\beta} \sim N(\mathbf{Z}^{'}(\mathbf{X})\beta, \sigma^2\mathbf{Z}^{'}(\mathbf{X})(\mathbf{X}^{'}\mathbf{X})^{-\mathbf{1}}\mathbf{Z}(\mathbf{X})).$$

Entonces;

(a) Considerando el **E-Modelo modificado**, el problema determinístico equivalente esta dado por

$$\begin{gathered} \min_{\mathbf{X}} \ k_1 \ \mathrm{E}(\hat{Y}(\mathbf{X},\hat{\beta})) + k_2 \ \sqrt{\mathrm{Var}(\hat{\mathbf{Y}}(\mathbf{X},\hat{\beta}))} \\ s.a \\ \|\mathbf{X}\|^2 \leq c^2. \end{gathered} \tag{70}$$

Explícitamente,

$$\begin{gathered} \min_{\mathbf{X}} \ k_1 \ \mathbf{Z}^{'}(\mathbf{X})\beta + k_2 \ \sigma\sqrt{\mathbf{Z}^{'}(\mathbf{X})(\mathbf{X}^{'}\mathbf{X})^{-\mathbf{1}}\mathbf{Z}(\mathbf{X})} \\ s.a \\ \|\mathbf{X}\|^2 \leq c^2 \end{gathered} \tag{71}$$

Si β y σ son no conocidos, entonces se pueden sustituir por sus estimadores, de tal forma que el problema determinístico es

$$\begin{gathered} \min_{\mathbf{X}} \ k_1 \ \mathbf{Z}^{'}(\mathbf{X})\hat{\beta} + k_2 \ \hat{\sigma}\sqrt{\mathbf{Z}^{'}(\mathbf{X})(\mathbf{X}^{'}\mathbf{X})^{-\mathbf{1}}\mathbf{Z}(\mathbf{X})} \\ s.a \\ \|\mathbf{X}\|^2 \leq c^2 \end{gathered} \tag{72}$$

(b) Para el **V-Modelo**, recordemos que éste es un caso particular del E-Modelo modificado, donde $k_1 = 0$ y $k_2 = 1$, por lo que el problema determinístico equivalente está dado por

$$\begin{gathered} \min_{\mathbf{X}} \hat{\sigma}\sqrt{\mathbf{Z}^{'}(\mathbf{X})(\mathbf{X}^{'}\mathbf{X})^{-\mathbf{1}}\mathbf{Z}(\mathbf{X})} \\ s.a \\ \|\mathbf{X}\|^2 \leq c^2 \end{gathered} \tag{73}$$

(c) Finalmente, para el **P-Modelo**, sea R un valor conocido, de tal forma que el problema determinístico está definido como

$$\begin{gathered}\max_{\mathbf{X}} \mathrm{P}(\hat{\mathbf{Y}}(\mathbf{X}, \hat{\beta}) \leq R) \\ s.a \\ \|\mathbf{X}\|^2 \leq c^2,\end{gathered} \tag{74}$$

estandarizando la función objetivo

$$\begin{aligned}\mathrm{P}(\hat{\mathbf{Y}}(\mathbf{X}, \hat{\beta}) \leq R) &= \mathrm{P}\left[\frac{\mathbf{Z}'(\mathbf{X})\hat{\beta} - \mathbf{Z}'(\mathbf{X})\beta}{\sigma\sqrt{\mathbf{Z}'(\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'(\mathbf{X})}} \leq \frac{R - \mathbf{Z}'(\mathbf{X})\beta}{\sigma\sqrt{\mathbf{Z}'(\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'(\mathbf{X})}}\right] \\ &= \Phi\left(\frac{R - \mathbf{Z}'(\mathbf{X})\beta}{\sigma\sqrt{\mathbf{Z}'(\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'(\mathbf{X})}}\right) \\ &= \mathrm{F}_{\mathbf{Z}'(\mathbf{X})\hat{\beta}}(R).\end{aligned}$$

Maximizando el valor de la variable, de la función de distribución $\mathrm{F}_{\mathbf{Z}'(\mathbf{X})\hat{\beta}}(R)$, entonces el problema determinístico equivalente es

$$\begin{gathered}\max_{\mathbf{X}} \frac{R - \mathbf{Z}'(\mathbf{X})\hat{\beta}}{\hat{\sigma}\sqrt{\mathbf{Z}'(\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'(\mathbf{X})}} \\ s.a \\ \|\mathbf{X}\|^2 \leq c^2\end{gathered} \tag{75}$$

Donde aparecen los estimadores $\hat{\beta}$ y $\hat{\sigma}$ si es que los parámetros correspondientes no se conocen.

IV. ASIGNACIÓN ÓPTIMA A TRAVÉS DE OPTIMIZACIÓN ESTOCÁSTICA

Introducción

Considere el problema de minimizar la varianza $\text{Var}(\bar{y}_{st})$, para un costo fijo C, que debe de cumplir la función de costo, dado por

$$\min_{\mathbf{n}} \text{Var}(\bar{y}_{st})$$

$$s.a$$

$$\mathbf{c}^{'}\mathbf{n} + c_0 = C$$

o alternativamente, el problema de minimizar la función de costo, para una precisión deseada V_0, asignada a la varianza $\text{Var}(\bar{y}_{st})$, el cual está dado por

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

$$s.a$$

$$\text{Var}(\bar{y}_{st}) \leq V_0,$$

donde

$$\begin{aligned}\text{Var}(\bar{y}_{st}) &= \mathbf{W}^{'}\left[\text{diag}\left(\frac{S_h^2}{n_h}\right)\right]\mathbf{W} - \frac{1}{N}(\mathbf{S}^2)^{'}\mathbf{W} \\ &= \sum_{h=1}^{H}\frac{W_h^2 S_h^2}{n_h} - \sum_{h=1}^{H}\frac{W_h S_h^2}{N}\end{aligned}$$

Note que las varianzas poblacionales S_h^2, $h = 1, 2, \ldots, H$ generalmente no se conocen, además como se comentó en el Capítulo II, el problema de optimización está también sujeto a que n_h sea entero, entonces el problema a optimizar, para una varianza mínima sujeta a una función de costo con C fijo, está dado por

$$\begin{gathered}\min_{\mathbf{n}} \widehat{\text{Var}}(\bar{y}_{st}) \\ s.a \\ \mathbf{c}'\mathbf{n} + c_0 = C \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,\end{gathered} \tag{76}$$

y para minimizar una función de costo, sujeta a una precisión deseada V_0, para la varianza

$$\begin{gathered}\min_{\mathbf{n}} \mathbf{c}'\mathbf{n} + c_0 \\ s.a \\ \widehat{\text{Var}}(\bar{y}_{st}) \leq V_0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,\end{gathered} \tag{77}$$

donde ahora

$$\widehat{\text{Var}}(\bar{y}_{st}) = \sum_{h=1}^{H} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{H} \frac{W_h s_h^2}{N}.$$

Observe que en la varianza estimada $\widehat{\text{Var}}(\bar{y}_{st})$, las varianzas muestrales s_h^2, $h = 1, 2, \ldots, H$, son variables aleatorias. Entonces los problemas de asignación óptima (76) y (77), se definen como problemas de optimizacón estocástica de enteros. Adicionalmente tanto para el problema (76), como en el (77), los costos c_h también pueden ser considerados variables aleatorias (esto cuando los c_h no puedan ser determinados, en alguna situación especial), bajo estas condiciones, los problemas (76) y (77) son problemas estocásticos, aun cuando las varianzas s_h^2 no lo sean.

En este capítulo se abordan precisamente estas ideas, considerar la aleatoriedad de las varianzas muestrales s_h^2 y/o los costos c_h, en el problema de asignación óptima tradicional, donde mediante los métodos de optimización estocástica revisados en el Capítulo III, se busca proponer las soluciones a los problemas estocásticos en la asiganación óptima.

En el muestreo estratificado con asignación óptima, el valor de los tamaños de muestra n_h en los distintos estratos depende de si se requiere minimizar la varianza $\text{Var}(\bar{y}_{st})$ sujeta a un costo fijo C que debe de cumplir la función de costos (Problema 1), o si la situación es la de minimizar la función de costo, sujeta a una presición deseada V_0 asignada a la varianza $\text{Var}(\bar{y}_{st})$ (Problema 2). Estos dos problemas serán tratados por separado en el resto del capítulo.

Problema 1

Cuando el interés sea el de minimizar la varianza $\widehat{\text{Var}}(\bar{y}_{st})$ sujeta a una función de costo, el problema es expresado como

$$\min_{\mathbf{n}} f(\mathbf{n})$$

$$s.a$$

$$P\left[\sum_{h=1}^{H} c_h n_h = C - c_0\right] \geq p_0.$$

$$n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,$$

con

$$f(\mathbf{n}) = \widehat{\text{Var}}(\bar{y}_{st}) = \sum_{h=1}^{H} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{H} \frac{W_h s_h^2}{N},$$

donde $\mathbf{s}^2 = (s_1^2, s_2^2, \ldots, s_H^2)'$ y $\mathbf{c} = (c_1, c_2, \ldots, c_H)'$ son vectores aleatorios, y además $p_0,\ 0 \leq p_0 \leq 1$ es una probabilidad especifica.

Bajo estas condiciones se pueden presentar diferentes casos, el primero es donde sólo las varianzas $s_h^2,\ h = 1, 2, \ldots, H$ son variables aleatorias, el segundo donde sólo los costos $c_h,\ h = 1, 2, \ldots, H$ son aleatorios, y se considera un tercer caso, donde s_h^2 y c_h son aleatorios, estos 3 casos son estudiados en esta sección.

Caso 1.

Cuando sólo las s_h^2 son variables aleatorias, el problema a optimizar es

$$\begin{gathered} \min_{\mathbf{n}} f(\mathbf{n}) \\ s.a \\ \textstyle\sum_{h=1}^{H} c_h n_h = C - c_0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H, \end{gathered} \tag{78}$$

donde

$$f(\mathbf{n}) = \sum_{h=1}^{H} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{H} \frac{W_h s_h^2}{N}$$

Utilizando la distribución límite de las varianzas muestrales (ver Apéndice A). Considere la variable aleatoria ξ_h, esto es

$$\xi_h = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{ih} - \bar{Y}_h)^2.$$

La cual tiene una distribución asintóticamente normal con media $\mathrm{E}(\xi_h)$ y varianza $\mathrm{Var}(\xi_h)$, dadas por

$$\mathrm{E}(\xi_h) = \frac{n_h}{n_h - 1} S_h^2. \tag{79}$$

y

$$\mathrm{Var}(\xi_h) = \frac{n_h}{(n_h - 1)^2} [C_{y_h}^4 - (S_h^2)^2], \tag{80}$$

donde $C_{y_h}^4$ es el cuarto momento central, y se calcula mediante

$$C_{y_h}^4 = \frac{1}{N_h}\sum_{i=1}^{N_h}(y_{ih} - \bar{Y}_h)^4.$$

Observe que

$$s_h^2 = \xi_h - \frac{n_h}{n_h - 1}(\bar{y}_h - \bar{Y}_h)^2,$$

donde (ver Apéndice A)

$$\frac{n_h}{n_h - 1} \to 1$$

y

$$(\bar{y}_h - \bar{Y}_h)^2 \to 0 \quad \text{en probabilidad.}$$

Por consecuencia las varianzas muestrales s_h^2 tienen distribución asintóticamente normal con media $\mathrm{E}(\xi_h)$ y varianza $\mathrm{Var}(\xi_h)$.

La función objetivo es una función lineal en las variables s_h^2, por lo que también tiene distribución normal, más aún

$$\begin{aligned}
\mathrm{E}(f(\mathbf{n})) &= \mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st})) \\
&= \mathrm{E}\left(\sum_{h=1}^{H}\frac{W_h^2}{n_h}s_h^2 - \sum_{h=1}^{H}\frac{W_h}{N}s_h^2\right) \\
&= \sum_{h=1}^{H}\frac{W_h^2}{n_h}\mathrm{E}(\xi_h) - \sum_{h=1}^{H}\frac{W_h}{N}\mathrm{E}(\xi_h) \\
&= \sum_{h=1}^{H}\frac{W_h^2}{(n_h - 1)}S_h^2 - \sum_{h=1}^{H}\frac{W_h}{N}\left(\frac{n_h}{n_h - 1}\right)S_h^2
\end{aligned} \tag{81}$$

y

$$
\begin{aligned}
\operatorname{Var}(f(\mathbf{n})) &= \operatorname{Var}(\widehat{\operatorname{Var}}(\bar{y}_{st})) \\
&= \operatorname{Var}\left(\sum_{h=1}^{H}\frac{W_h^2}{n_h}s_h^2-\sum_{h=1}^{H}\frac{W_h}{N}s_h^2\right) \\
&= \sum_{h=1}^{H}\frac{W_h^4}{n_h^2}\operatorname{Var}(\xi_h)-\sum_{h=1}^{H}\frac{W_h^2}{N^2}\operatorname{Var}(\xi_h) \quad \text{(por independencia)} \\
&= \sum_{h=1}^{H}\frac{W_h^4}{n_h^2}\left[\frac{n_h}{(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)\right]-\sum_{h=1}^{H}\frac{W_h^2}{N^2}\left[\frac{n_h}{(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)\right] \\
&= \sum_{h=1}^{H}\frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)-\sum_{h=1}^{H}\frac{W_h^2}{N^2}\left[\frac{n_h}{(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)\right]. \quad (82)
\end{aligned}
$$

Entonces, la nueva función objetivo puede ser determinada tal como se hizo en el Capítulo III, por

$$
\begin{aligned}
f(\mathbf{n}) &= k_1\mathrm{E}(f(\mathbf{n}))+k_2\sqrt{\operatorname{Var}(f(\mathbf{n}))} \\
&= k_1\mathrm{E}(\widehat{\operatorname{Var}}(\bar{y}_{st}))+k_2\sqrt{\operatorname{Var}(\widehat{\operatorname{Var}}(\bar{y}_{st}))}, \qquad (83)
\end{aligned}
$$

explícitamente

$$
\begin{aligned}
f(\mathbf{n}) &= k_1\left[\sum_{h=1}^{H}\frac{W_h^2}{(n_h-1)}S_h^2-\sum_{h=1}^{H}\frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right)S_h^2\right] \\
&+k_2\left[\sum_{h=1}^{H}\frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)-\sum_{h=1}^{H}\frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)\right)\right]^{1/2}
\end{aligned}
$$

La cual es una función objetivo determinística. De tal forma que el problema determinístico equivalente al problema estocástico está expresado como,

$$
\begin{gathered}
\min_{\mathbf{n}} f(\mathbf{n}) \\
s.a \\
\sum_{h=1}^{H} c_h n_h = C - c_0 \qquad (84) \\
n_h \in \mathbb{N}, \quad h=1,2,\ldots,H,
\end{gathered}
$$

donde

$$f(\mathbf{n}) = k_1 \left[\sum_{h=1}^{H} \frac{W_h^2}{(n_h - 1)} S_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} \left(\frac{n_h}{n_h - 1} \right) S_h^2 \right]$$
$$+k_2 \left[\sum_{h=1}^{H} \frac{W_h^4}{n_h(n_h - 1)^2} (C_{y_h}^4 - (S_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2} \left(\frac{n_h}{(n_h - 1)^2} (C_{y_h}^4 - (S_h^2)^2) \right) \right]^{1/2}$$

Cuyo planteamiento, es propuesto bajo la técnica del E-Modelo modificado, donde dando valores a las ponderaciones k_1 y k_2, el modelo puede enfocarse a la E($f(\mathbf{n})$) (E-Modelo) o Var($f(n_n)$))(V-Modelo).

Observación:

Es importante hacer notar, que la función objetivo está en términos de las varianzas poblacionales S_h^2, dado que estas no son conocidas entonces se utiliza un estimador de dichas varianzas, en particular las varianzas muestrales s_h^2. En tal caso el problema de optimización determinístico equivalente está dado por

$$\begin{gathered} \min_{\mathbf{n}} \hat{f}(\mathbf{n}) \\ s.a \\ \sum_{h=1}^{H} c_h n_h = C - c_0. \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H, \end{gathered} \tag{85}$$

donde, $\hat{f}(\mathbf{n})$ está dado por

$$\hat{f}(\mathbf{n}) = k_1 \left[\sum_{h=1}^{H} \frac{W_h^2}{(n_h - 1)} s_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} \left(\frac{n_h}{n_h - 1} \right) s_h^2 \right]$$
$$+k_2 \left[\sum_{h=1}^{H} \frac{W_h^4}{n_h(n_h - 1)^2} (C_{y_h}^4 - (s_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2} \left(\frac{n_h}{(n_h - 1)^2} (C_h^4 - (s_h^2)^2) \right) \right]^{1/2}$$

Caso 2.

Cuando sólo el vector de costos $\mathbf{c}$ es aleatorio. El problema de optimización puede establecerse como

$$\begin{gathered}\min_{\mathbf{n}} \left[\sum_{h=1}^{H} \frac{W_h^2 S_h^2}{n_h} - \sum_{h=1}^{H} \frac{W_h S_h^2}{N} \right] \\ s.a \\ P\left[\mathbf{c}'\mathbf{n} = C - c_0\right] \geq p_0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H. \end{gathered} \tag{86}$$

Suponga que $\mathbf{c}$ tiene cierta distribución n-dimensional conocida, con

$$\mathrm{E}(\mathbf{c}) = \bar{\mathbf{c}} = (\bar{c_1}, \ldots, \bar{c_H})' \tag{87}$$

y

$$\mathrm{Cov}(\mathbf{c}) = \mathbf{V} = \begin{bmatrix} \mathrm{Var}(c_1) & \mathrm{Cov}(c_1, c_2) & \ldots & \mathrm{Cov}(c_1, c_H) \\ \mathrm{Cov}(c_2, c_1) & \mathrm{Var}(c_2) & \ldots & \mathrm{Cov}(c_2, c_H) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(c_H, c_1) & \mathrm{Cov}(c_H, c_2) & \ldots & \mathrm{Var}(c_H) \end{bmatrix}. \tag{88}$$

Entonces, la esperanza y la varianza para la función lineal de costos son las siguientes.

$$\mathrm{E}(\mathbf{c}'\mathbf{n}) = \mathrm{E}(\mathbf{c}')\mathbf{n} = \bar{\mathbf{c}}'\mathbf{n}, \tag{89}$$

y

$$\mathrm{Var}(\mathbf{c}'\mathbf{n}) = \mathrm{Var}(\mathbf{n}'\mathbf{c}) = \mathbf{n}'\mathrm{Cov}(\mathbf{c})\mathbf{n} = \mathbf{n}'\mathbf{V}\mathbf{n}. \tag{90}$$

A partir de (86), estandarizando la función para las restricciones, se tiene que

$$\begin{aligned} P\left[\mathbf{c}'\mathbf{n} \leq C - c_0\right] &= P\left[\frac{\mathbf{c}'\mathbf{n} - \bar{\mathbf{c}}'\mathbf{n}}{(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2}} \leq \frac{(C - c_0) - \bar{\mathbf{c}}'\mathbf{n}}{(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2}}\right] \\ &= F_{c'n}\left(\frac{(C - c_0) - \bar{\mathbf{c}}'\mathbf{n}}{(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2}}\right). \end{aligned}$$

Sea e el valor de la variable aleatoria bajo la distribución propuesta para los costos, tal que $F_{c'n}(e) = p_0$, entonces la desigualdad se puede establecer como

$$F_{c'n}\left(\frac{(C - c_0) - \bar{\mathbf{c}}'\mathbf{n}}{(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2}}\right) \geq F_{c'n}(e).$$

Lo cual es cierto sólo si se cumple que

$$\frac{(C - c_0) - \bar{\mathbf{c}}'\mathbf{n}}{(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2}} > e$$

equivalentemente

$$\bar{\mathbf{c}}'\mathbf{n} + e(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2} - (C - c_0) \leq 0. \tag{91}$$

De esta forma el problema estocástico puede ser resuelto mediante el siguiente problema determinístico equivalente

$$\begin{gathered} \min_{\mathbf{n}} \left[\sum_{h=1}^{H} \frac{W_h^2}{n_h} S_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} S_h^2\right] \\ s.a \\ \bar{\mathbf{c}}'\mathbf{n} + e(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2} - (C - c_0) \leq 0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H. \end{gathered} \tag{92}$$

Caso 3.

Considere ahora que s_h^2 y c_h, $h = 1, 2, \ldots, H$ son variables aleatorias. Para este caso el problema de optimización es expresado como

$$\begin{gathered} \min_{\mathbf{n}} \; f(\mathbf{n}) \\ s.a \\ P\left[\sum_{h=1}^{H} c_h n_h = C - c_0\right] \geq p_0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H, \end{gathered} \tag{93}$$

donde

$$f(\mathbf{n}) = \widehat{\mathrm{Var}}(\bar{y}_{st}) = \sum_{h=1}^{H} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{H} \frac{W_h s_h^2}{N}.$$

Observe que tanto la función objetivo como la restricción sobre la varianza, dependen de variables aleatorias. Más aún, asumiendo que las varianzas muestrales $s_h^2 \sim N(\mathrm{E}(\xi_h), \mathrm{Var}(\xi_h))$, $h = 1, 2, \ldots, H$, independientes, donde $\mathrm{E}(\xi_h)$ y $\mathrm{Var}(\xi_h)$ están dadas por (79) y (80) respectivamente. Además, $\mathbf{c}$ tiene cierta distribución conocida, con media $\mathrm{E}(\mathbf{c}) = \bar{\mathbf{c}}$ y varianza $\mathrm{Var}(\mathbf{c}) = \mathbf{V}$. Entonces, la solución al problema estocástico (93), está dado por una combinación de los dos resultados presentados para los casos 1 y 2. De está manera, la función objetivo estocástica en (93), es reemplazada por la función determinística dada por (83) como en el caso 1, y la función determinística para las restricciones en (93), es considerada tal como se resolvió para el caso 2 en la ecuación (91). Por lo que el problema determinístico equivalente, está dado por

$$\begin{gathered} \min_{\mathbf{n}} f(\mathbf{n}) \\ s.a \\ \bar{\mathbf{c}}'\mathbf{n} + e(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2} - (C - c_0) \leq 0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H, \end{gathered} \tag{94}$$

donde

$$f(\mathbf{n}) = k_1 \left[\sum_{h=1}^{H} \frac{W_h^2}{(n_h - 1)} S_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} \left(\frac{n_h}{n_h - 1} \right) S_h^2 \right]$$
$$+ k_2 \left[\sum_{h=1}^{H} \frac{W_h^4}{n_h (n_h - 1)^2} (C_{y_h}^4 - (S_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2} \left(\frac{n_h}{(n_h - 1)^2} (C_{y_h}^4 - (S_h^2)^2) \right) \right]^{1/2}.$$

Finalmente, sustituyendo S_h^2 por su estimador s_h^2 (ver la **observación**, en el caso 1), el problema determinístico es

$$\begin{gathered}\min_{\mathbf{n}} \hat{f}(\mathbf{n}) \\ s.a \\ \bar{\mathbf{c}}^{'}\mathbf{n} + e(\mathbf{n}^{'}\mathbf{V}\mathbf{n})^{1/2} - (C - c_0) \leq 0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,\end{gathered} \tag{95}$$

donde

$$\hat{f}(\mathbf{n}) = k_1 \left[\sum_{h=1}^{H} \frac{W_h^2}{(n_h - 1)} s_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} \left(\frac{n_h}{n_h - 1} \right) s_h^2 \right]$$
$$+ k_2 \left[\sum_{h=1}^{H} \frac{W_h^4}{n_h (n_h - 1)^2} (C_{y_h}^4 - (s_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2} \left(\frac{n_h}{(n_h - 1)^2} (C_{y_h}^4 - (s_h^2)^2) \right) \right]^{1/2}.$$

Problema 2

Cuando el interés sea el de minimizar la función de costo, para un valor especifico de $\mathrm{Var}(\bar{y}_{st})$, el problema puede ser expresado como

$$\begin{gathered}\min_{\mathbf{n}} \mathbf{c}^{'}\mathbf{n} + c_0 \\ s.a \\ P\left[\widehat{\mathrm{Var}}(\bar{y}_{st}) \leq V_0 \right] \geq p_0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,\end{gathered}$$

con

$$\widehat{\mathrm{Var}}(\bar{y}_{st}) = \sum_{h=1}^{H} \frac{W_h^2}{n_h} s_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} s_h^2,$$

donde $\mathbf{s}^2 = (s_1^2, \ldots, s_H^2)'$ y $\mathbf{c} = (c_1, \ldots, c_H)'$ son variables aleatorias, V_0 es una constante no negativa conocida, c_0 son los costos adicionales (constante) y p_0, $0 \leq p_0 \leq 1$ es una probabilidad especifica.

Como en el problema 1, se pueden abordar diferentes casos de interés, el primero es cuando sólo los costos $\mathbf{c}$ son vector aleatorio, el segundo caso ocurre cuando sólo las varianzas $\mathbf{s}^2$ son aleatorias, y un tercer caso, cuando simultáneamente $\mathbf{c}$ y $\mathbf{s}^2$ son vectores aleatorios.

Caso 1.

Cuando sólo los costos $\mathbf{c} = (c_1, \dots, c_H)'$ son vector aleatorio, el problema de optimización está dado por

$$\begin{gathered} \min_{\mathbf{n}} \,(\mathbf{c}'\mathbf{n} + c_o) \\ s.a \\ \sum_{h=1}^{H} \frac{W_h^2}{n_h} S_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} S_h^2 \leq V_0 \\ n_h \in \mathbb{N} \quad h = 1, 2, \dots, H. \end{gathered} \tag{96}$$

Note que la función objetivo es una función estocástica, debido a que depende del vector aleatorio $\mathbf{c}$. Para dar solución a este problema se supone nuevamente que $\mathbf{c}$ tiene cierta distribución conocida con $\mathrm{E}(\mathbf{c}) = \bar{\mathbf{c}}$ y $\mathrm{Var}(\mathbf{c}) = \mathbf{V}$ (ver 87 y 88), entonces para la función lineal de costo en el problema (96), se tiene que

$$\mathrm{E}(\mathbf{c}'\mathbf{n} + c_0) = \mathrm{E}(\mathbf{c}')\mathbf{n} + c_0 = \bar{\mathbf{c}}'\mathbf{n} + c_0 \tag{97}$$

y varianza

$$\mathrm{Var}(\mathbf{c}'\mathbf{n} + c_0) = \mathrm{Var}(\mathbf{n}'\mathbf{c}) = \mathbf{n}'\mathrm{Cov}(\mathbf{c})\mathbf{n} = \mathbf{n}'\mathbf{V}\mathbf{n} \tag{98}$$

Donde ahora $\bar{\mathbf{c}}'\mathbf{n} + c_0$ y $\mathbf{n}'\mathbf{V}\mathbf{n}$ no son aleatorios, de tal forma que la nueva función objetivo, está dada por

$$k_1(\bar{\mathbf{c}}'\mathbf{n} + c_0) + k_2(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2} \tag{99}$$

Donde k_1 y k_2, son constantes no negativas que ponderan la importancia de $\mathrm{E}(\mathbf{c}'\mathbf{n}+c_0)$ y $(\mathrm{Var}(\mathbf{c}'\mathbf{n} + c_0))^{1/2}$.

Entonces, bajo el E-Modelo modificado, el problema determinístico equivalente al problema estocástico (96), es

$$\begin{gathered}\min_{\mathbf{n}} k_1(\bar{\mathbf{c}}'\mathbf{n}+c_0)+k_2(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2} \\ s.a \\ \sum_{h=1}^{H}\frac{W_h^2}{n_h}S_h^2-\sum_{h=1}^{H}\frac{W_h}{N}S_h^2\leq V_0 \\ n_h\in\mathbb{N} \quad h=1,2,\ldots,H.\end{gathered} \tag{100}$$

<u>Caso 2.</u>

Cuando sólo las varianzas s_h^2 son variables aleatorias. El problema de optimización estocástica es el siguiente

$$\begin{gathered}\min_{\mathbf{n}} \mathbf{c}'\mathbf{n}+c_o \\ s.a \\ P\left[\widehat{\mathrm{Var}}(\bar{y}_{st})\leq V_0\right]\geq p_0 \\ n_h\in\mathbb{N} \quad h=1,2,\ldots,H,\end{gathered} \tag{101}$$

donde

$$\widehat{\mathrm{Var}}(\bar{y}_{st})=\sum_{h=1}^{H}\frac{W_h^2}{n_h}s_h^2-\sum_{h=1}^{H}\frac{W_h}{N}s_h^2.$$

Utilizando nuevamente el hecho de que s_h^2 tiene una distribución asintóticamente normal con media $\mathrm{E}(\xi_h)$ y varianza $\mathrm{Var}(\xi_h)$ (vea 78 y 79).

Entonces la varianza estimada $\widehat{\mathrm{Var}}(\bar{y}_{st})$ en el problema estocástico (101), también tiene distribución asintóticamente normal con

$$\mathrm{E}\left(\widehat{\mathrm{Var}}(\bar{y}_{st})\right)=\sum_{h=1}^{H}\frac{W_h^2}{(n_h-1)}S_h^2-\sum_{h=1}^{H}\frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right)S_h^2,$$

y

$$\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st})) = \sum_{h=1}^{H} \frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4 - (S_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2}\left[\frac{n_h}{(n_h-1)^2}(C_{y_h}^4 - (S_h^2)^2)\right].$$

Ahora estandarizando la función para las restricciones, en la ecuación (101), se tiene que

$$\mathrm{P}\left[\frac{\widehat{\mathrm{Var}}(\bar{y}_{st}) - \mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}{\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}} \leq \frac{V_0 - \mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}{\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}}\right] \geq p_0,$$

con

$$p_0 = \Phi\left(\frac{V_0 - \mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}{\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}}\right),$$

donde $\Phi(\cdot)$, denota la función de distribución normal estándar. Entonces sea e, el valor de la variable aleatorio normal estándar tal que $\Phi(e) = p_0$, de tal forma que la desigualdad puede establecerse como

$$\Phi\left(\frac{V_0 - \mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}{\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}}\right) \geq \Phi(e),$$

lo cual se cumple sólo si

$$\frac{V_0 - \mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}{\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}} \geq e,$$

o equivalentemente

$$\mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st})) + e\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))} - V_0 \leq 0 \tag{102}$$

De forma explícita,

$$\begin{aligned} &\sum_{h=1}^{H} \frac{W_h^2}{(n_h-1)} S_h^2 - \sum_{h=1}^{H} \frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right) S_h^2 \\ +e&\left[\sum_{h=1}^{H} \frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4 - (S_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}(C_{y_h}^4 - (S_h^2)^2)\right)\right]^{1/2} - V_0 \leq 0. \end{aligned}$$

Entonces el problema determinístico equivalente, al problema estocástico (101), está dado por

$$\begin{gathered}\min_{\mathbf{n}} \mathbf{c}^{'}\mathbf{n} + c_o\\ s.a\\ \mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st})) + e\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))} - V_0 \leq 0\\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,\end{gathered} \tag{103}$$

donde

$$\begin{aligned}&\mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st})) + e\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}\\ &= \sum_{h=1}^{H} \frac{W_h^2}{(n_h - 1)} S_h^2 - \sum_{h=1}^{H} \frac{W_h}{N}\left(\frac{n_h}{n_h - 1}\right) S_h^2\\ &+e\left[\sum_{h=1}^{H} \frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4 - (S_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}(C_{y_h}^4 - (S_h^2)^2)\right)\right]^{1/2}.\end{aligned}$$

Note nuevamente que la función para las restricciones en (103) está en términos de las varianzas poblacionales S_h^2, utilizando el estimador de las varianzas s_h^2 en lugar de las S_h^2 (ver la **observación** en el problema 1, caso 1). Así, se tiene el siguiente problema determinístico

$$\begin{gathered}\min_{\mathbf{n}} \mathbf{c}^{'}\mathbf{n} + c_o\\ s.a\\ \widehat{\mathrm{E}}(\widehat{\mathrm{Var}}(\bar{y}_{st})) + e\sqrt{\widehat{\mathrm{Var}}(\widehat{\mathrm{Var}}(\bar{y}_{st}))} - V_0 \leq 0\\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,\end{gathered} \tag{104}$$

donde

$$\begin{aligned}&\widehat{\mathrm{E}}(\widehat{\mathrm{Var}}(\bar{y}_{st})) + e\sqrt{\widehat{\mathrm{Var}}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}\\ &= \sum_{h=1}^{H} \frac{W_h^2}{(n_h - 1)} s_h^2 - \sum_{h=1}^{H} \frac{W_h}{N}\left(\frac{n_h}{n_h - 1}\right) s_h^2\\ &+e\left[\sum_{h=1}^{H} \frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4 - (s_h^2)^2) - \sum_{h=1}^{H} \frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}(C_{y_h}^4 - (s_h^2)^2)\right)\right]^{1/2}\end{aligned}$$

Caso 3.

Cuando $\mathbf{s^2}$ y $\mathbf{c}$ son vectores aleatorios, el problema de optimización es el siguiente

$$\begin{gathered}\min_{\mathbf{n}} \mathbf{c}'\mathbf{n} + c_o \\ s.a \\ P\left[\widehat{\mathrm{Var}}(\bar{y}_{st}) \leq V_0\right] \geq p_0 \\ n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H,\end{gathered} \tag{105}$$

donde

$$\widehat{\mathrm{Var}}(\bar{y}_{st}) = \sum_{h=1}^{H} \frac{W_h^2}{n_h} s_h^2 - \sum_{h=1}^{H} \frac{W_h}{N} s_h^2$$

Como se puede observar, tanto la función objetivo, como la función para las restricciones dependen de variables aleatorias, por lo que el problema (105), se define como un problema de optimización estocástico.

Asumiendo, que los costos $\mathbf{c}$ tienen cierta distribución conocida o propuesta, con $\mathrm{E}(\mathbf{c}) = \bar{\mathbf{c}}$ y $\mathrm{Var}(\mathbf{c}) = \mathbf{n}'\mathbf{V}\mathbf{n}$ (ver 87 y 88); Además considerando que las varianzas muestrales s_h^2, $h = 1, 2, \ldots, H$ se distribuyen asintóticamente $N(\mathrm{E}(\xi_h), \mathrm{Var}(\xi_h))$ (ver 78-79).

La solución al problema estocástico (105), está dada por la combinación de los casos 1 y 2, en el **problema 2**, esto es: La función objetivo estocástica en (105), es reemplazada por la función objetivo determinística en el caso 1 (96), y la función determinística para las restricciones, está dada por (102), tal como se resolvió en el caso 2. Por lo que el problema determinístico equivalente al problema estocástico (105), es el siguiente

$$
\begin{gathered}
\min_{\mathbf{n}} k_1(\bar{\mathbf{c}}'\mathbf{n}+c_0)+k_2(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2} \\
s.a \\
\mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st}))+e\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}-V_0 \leq 0 \\
n_h \in \mathbb{N}, \quad h=1,2,\ldots,H,
\end{gathered} \tag{106}
$$

donde

$$
\begin{aligned}
&\mathrm{E}(\widehat{\mathrm{Var}}(\bar{y}_{st}))+e\sqrt{\mathrm{Var}(\widehat{\mathrm{Var}}(\bar{y}_{st}))} \\
&=\sum_{h=1}^{H}\frac{W_h^2}{(n_h-1)}S_h^2-\sum_{h=1}^{H}\frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right)S_h^2 \\
&+e\left[\sum_{h=1}^{H}\frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)-\sum_{h=1}^{H}\frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}(C_{y_h}^4-(S_h^2)^2)\right)\right]^{1/2}
\end{aligned}
$$

De nuevo, se observa que la función para las restricciones está en términos de las varianzas poblacionales S_h^2, las cuales no se conocen (ver la **observación** en el **problema 1**, caso 1), sustituyendo S_h^2 por su estimador s_h^2, se tiene que el problema determinístico equivalente al problema estocástico (105), está dado por

$$
\begin{gathered}
\min_{\mathbf{n}} k_1(\bar{\mathbf{c}}'\mathbf{n}+c_0)+k_2(\mathbf{n}'\mathbf{V}\mathbf{n})^{1/2} \\
s.a \\
\widehat{\mathrm{E}}(\widehat{\mathrm{Var}}(\bar{y}_{st}))+e\sqrt{\widehat{\mathrm{Var}}(\widehat{\mathrm{Var}}(\bar{y}_{st}))}-V_0 \leq 0 \\
n_h \in \mathbb{N} \quad h=1,2,\ldots,H,
\end{gathered} \tag{107}
$$

donde

$$
\begin{aligned}
&\widehat{\mathrm{E}}(\widehat{\mathrm{Var}}(\bar{y}_{st}))+e\sqrt{\widehat{\mathrm{Var}}(\widehat{\mathrm{Var}}(\bar{y}_{st}))} \\
&=\sum_{h=1}^{H}\frac{W_h^2}{(n_h-1)}s_h^2-\sum_{h=1}^{H}\frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right)s_h^2 \\
&+e\left[\sum_{h=1}^{H}\frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4-(s_h^2)^2)-\sum_{h=1}^{H}\frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}(C_{y_h}^4-(s_h^2)^2)\right)\right]^{1/2}
\end{aligned}
$$

Ejemplo de optimización estocástica en asignación óptima

Se consideran nuevamente los datos del ejemplo en el Capítulo II (ver Apéndice B), donde se tienen 187 datos obtenidos en un estudio previo, la población cuenta con un total de 10600 unidades, la variable de estratificación es la region de cultivo, por lo que se tienen nuevamente 4 estratos con $N_1 = 2500$, $N_2 = 2300$, $N_3 = 2800$, $N_4 = 3000$, y los costos por unidad en cada estrato son $c_1 = 25$, $c_2 = 20$, $c_3 = 40$, $c_4 = 15$ y $c_0 = 0$.

Problema 1

El problema que se desea resolver, es el de minimizar la varianza $\text{Var}(\bar{y}_{st})$ sujeta a una función de costo tal que $C = 5000$, note que las varianzas poblacionales $\mathbf{S}^2$ en realidad no se conocen, y en su lugar se puede utilizar el vector de varianzas estimadas $\mathbf{s}^2$, el cual es aleatorio. Por lo que el problema a resolver es el siguiente

$$\min_{\mathbf{n}} \sum_{h=1}^{4} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{4} \frac{W_h s_h^2}{10600}$$

$$s.a$$

$$\sum_{h=1}^{H} c_h n_h = 5000$$

$$n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H.$$

Este es un problema de optimización estocástica de enteros, y la solución radica en encontrar un problema determinístico equivalente al problema estocástico. Tal como se vió en el **problema 1**, caso 2, del Capítulo IV. El problema de optimización determinístico equivalente al problema de optimización estocástica está dado por

$$\min_{\mathbf{n}} f(\mathbf{n})$$

$$s.a$$

$$\sum_{h=1}^{H} c_h n_h = 5000$$

$$n_h \in \mathbb{N}, \quad h = 1, 2, \ldots, H.$$

Donde

$$f(\mathbf{n}) = k_1 \left[\sum_{h=1}^{4} \frac{W_h^2}{(n_h - 1)} s_h^2 - \sum_{h=1}^{4} \frac{W_h}{10600} \left(\frac{n_h}{n_h - 1} \right) s_h^2 \right]$$
$$+ k_2 \left[\sum_{h=1}^{4} \frac{W_h^4}{n_h (n_h - 1)^2} (C_{y_h}^4 - (s_h^2)^2) - \sum_{h=1}^{4} \frac{W_h^2}{(10600)^2} \left(\frac{n_h}{(n_h - 1)^2} (C_{y_h}^4 - (s_h^2)^2) \right) \right]^{1/2}$$

Los valores de W_h, s_h^2, $C_{y_h}^4$ y c_h, $h = 1, 2, 3, 4$ se muestran en el Apéndice B.

El problema de optimización determinístico se resolvió utilizando el programa de cómputo LINGO, las soluciones se muestran en el Cuadro 4.1. Como se comentó en el Capítulo II y IV, el problema de asignación óptima, es realmente un **problema de programación no lineal de enteros**. En el ejemplo también se utilizó programación no lineal de enteros, los resultados se muestran en la última columna del Cuadro 4.1.

CUADRO 4.1: Soluciones al problema de asignación óptima en muestreo estratificado bajo los diferentes enfoques de optimización determinística para el problema 1, donde $k_1 = k_2 = 1$.

	Solución Estocástica	Solución Estocástica Ajustada [a]	Solución Estocástica Enteros
n_1	14.5361	15	14
n_2	81.2104	81	82
n_3	15.8253	16	16
n_4	78.5161	79	78
$\sum_{h=1}^{H} n_h$	190.0878	191	190
$g(\mathbf{n})$ [b]	.0096	.0096	.0096

[a] Esta solución se obtiene a partir de la solución de Lagrange, redondeando cada n_i, vea Cochran (1977). Alternativamente, en la literatura estadística también se suele ajustar al entero inmediato superior.

[b] Donde $g(\mathbf{n}) = V(\bar{y}_{st}) = \sum_{h=1}^{H} W_h^2 \frac{S_h^2}{n_h} - \sum_{h=1}^{H} W_h^2 \frac{S_h^2}{N_h}$

Las solución estocástica y la solución estocática de enteros son muy similares, en ambas soluciones los estratos con un mayor número de unidades muestrales asignadas son el 2 y el 4, esto se debe a que estos dos estratos contienen las varianzas mas grandes.

A continuación se muestra un cuadro comparativo, que exhibe todos los resultados obtenidos para el **problema 1**, tanto en el ejemplo del Capítulo II, como los resultados obtenidos en el ejemplo del Capítulo IV. Se agregaron las soluciones mediante programación no lineal de enteros para el ejemplo del Capítulo II (columna 4), los resultados son los siguientes.

Cuadro 4.2: Soluciones al problema de asignación óptima en muestreo estratificado bajo los diferentes enfoques de optimización para el **problema 1.**

	Solución Lagrange	Solución Ajustada[a]	Solución Enteros	Solución Estocástica	Solución Estocástica Ajustada	Solución Estocástica Enteros
n_1	13.7976	14	14	14.5361	15	14
n_2	81.7554	82	82	81.2104	81	82
n_3	15.1137	15	16	15.8253	16	16
n_4	79.0277	79	78	78.5161	79	78
$\sum_{h=1}^{H} n_h$	189.6944	190	190	190.0878	191	190
$g(\mathbf{n})$ [b]	0.0096	.0096	.0096	.0096	.0096	.0096

[a]Esta solución se obtiene a partir de la solución de Lagrange, redondeando cada n_i, vea Cochran (1977). Alternativamente, en la literatura estadística también se suele ajustar al entero inmediato superior.

[b]Donde $g(\mathbf{n}) = V(\bar{y}_{st}) = \sum_{h=1}^{H} W_h^2 \frac{s_h^2}{n_h} - \sum_{h=1}^{H} W_h \frac{s_h^2}{N}$

En las columnas 2, 3 y 4 se muestran los resultados obtenidos para el **problema 1** en el Capítulo II, Observe que el tamaño de muestra es el mismo tanto para la solución mediante Multiplicadores de Lagrange ajustado, como cuando se utiliza programación no lineal de enteros (190 unidades muestrales); sin embargo, la asignación varía por una unidad en los estratos 3 y 4, la asignación generada por programación no lineal de enteros, da 16 unidades al estrato 3 y 78 al estrato 4, lo cual es una asignación más conveniente dado que el estrato 3 es más económico que el estrato 4. Como era de esperar, la mayor cantidad de unidades muestrales se asignaron a los estratos 2 y 4, debido a que estos dos estratos son lo que tienen una varianza mayor.

Referente a los resultados obtenidos para el **problema 1** en el Capítulo IV, las soluciones estocásticas varían muy poco, en relación con las soluciones determinística, principalmente en los estratos 1 y 3 que son los de menor varianza. Aunque los tamaños de muestra en los estratos son casí los mismos para todas las soluciones (190 unidades en la muestra global), la solución estocástica asigna una unidad más a los estratos con menor varianza, lo cual se aproxima a la asignación dada mediante la solución estocástica de enteros.

Problema 2

El segundo problema consiste en minimizar la función de costo, para una precisión fija $V_0 = .015$, donde nuevamente $\mathbf{s}^2$ es vector aleatorio, bajo estas condiciones el problema de optimización está dado por

$$\min_{\mathbf{n}} \mathbf{c}^{'}\mathbf{n}$$

$$s.a$$

$$P\left[\widehat{\mathrm{Var}}(\bar{y}_{st}) \leq V_0\right] \geq 0.99,$$

$$n_h \in \mathbb{N}, \quad h = 1, 2, 3, 4,$$

donde

$$\widehat{\mathrm{Var}}(\bar{y}_{st}) = \sum_{h=1}^{4} \frac{W_h^2}{n_h}\ s_h^2 - \sum_{h=1}^{4} \frac{W_h}{10600}\ s_h^2$$

Este es un problema de optimización estocástica, y su solución depende de encontrar un problema determinístico equivalente. Como se vió en el problema 2, caso 2, en el Capítulo IV, el problema determinítico equivalente al estocástico está dado por

$$\min_{\mathbf{n}} \mathbf{c}^{'}\mathbf{n}$$

$$s.a.$$

$$\sum_{h=1}^{4}\frac{W_h^2}{(n_h-1)}s_h^2-\sum_{h=1}^{4}\frac{W_h}{10600}\left(\frac{n_h}{n_h-1}\right)s_h^2+2.3263$$
$$\times\left[\sum_{h=1}^{4}\frac{W_h^4}{n_h(n_h-1)^2}(C_{y_h}^4-(s_h^2)^2)-\sum_{h=1}^{4}\frac{W_h^2}{(10600)^2}\left(\frac{n_h}{(n_h-1)^2}(C_{y_h}^4-(s_h^2)^2)\right)\right]^{1/2}-.015\leq 0$$

Los valores de W_h, s_h^2, $C_{y_h}^4$ y c_h, $h=1,2,3,4$ se muestran en el Apéndice B.

Las soluciones se obtuvieron utilizando el programa de cómputo LINGO, los resultados se muestran en el Cuadro 4.2. Para este ejemplo también se utilizó programación no lineal de enteros, los valores resultantes se muestran en la columna 4.

Cuadro 4.3: Soluciones al problema de asignación óptima en muestreo estratificado bajo los diferentes enfoques de optimización determinística para el problema 2, donde $k_1=k_2=1$.

	Solución Estocástica	Solución Estocástica Ajustada [a]	Solución Estocástica Enteros
n_1	9.9345	9	10
n_2	53.9656	54	54
n_3	10.7847	11	11
n_4	52.1605	52	52
$\sum_{h=1}^{H} n_h$	126.8453	126	127
$g(\mathbf{n})$ [b]	3333.117	3310	3335

[a] Esta solución se obtiene a partir de la solución de Lagrange, redondeando cada n_i, vea Cochran (1977). Alternativamente, en la literatura estadística también se suele ajustar al entero inmediato superior.

[b] Donde $g(\mathbf{n})=\sum_{h=1}^{H}c_h n_h+c_0$

La mayor cantidad de unidades muestrales se asignaron a los estratos 2 y 4, aunque no son los estratos más económicos, son lo que contienen las varianzas mayores; la solución ajustada es similar a la solución de enteros; la diferencia se encuentra sólo en el estrato 1, en donde el método de enteros asignó una unidad más.

Se presenta un cuadro comparativo que muestra todas las soluciones obtenidas para el **problema 2**, tanto las del ejemplo en el Capítulo II, como las del ejemplo en el Capítulo IV. Se agrega la solución utilizando programación no lineal de enteros, para el ejemplo en el Capítulo II, estos resultados aparecen en la columna 4.

Cuadro 4.4: Soluciones al problema de asignación óptima en muestreo estratificado bajo los diferentes enfoques de optimización para el **problema 2.**

	Solución Lagrange	Solución Ajustada[a]	Solución Enteros	Solución Estocástica	Solución Estocástica Ajustada	Solución Estocástica Enteros
n_1	8.9372	9	10	9.9345	9	10
n_2	52.9560	53	52	53.9656	54	54
n_3	9.7897	10	11	10.7847	11	11
n_4	51.1891	51	51	52.1605	52	52
$\sum_{h=1}^{H} n_h$	122.872	123	124	126.8453	126	127
$g(\mathbf{n})^b$	3238.6820	3240	3250	3333.117	3310	3335

[a]Esta solución se obtiene a partir de la solución de Lagrange, redondeando cada n_i, vea Cochran (1977). Alternativamente, en la literatura estadística también se suele ajustar al entero inmediato superior.

[b]Donde $g(\mathbf{n}) = \sum_{h=1}^{H} c_h n_h + c_0$

En las columnas 2, 3 y 4 se muestran los resultados obtenidos para el **problema 2** en el Capítulo II. Observe que el tamaño de muestra aumenta sólo en una unidad cuando se utiliza programación no lineal de enteros (124 unidades muestrales), pero la asignación de unidades muestrales básicamente cambia en todos los estratos (por una unidad). Esta asignación es más conveniente dado que los estratos donde se aumenta una unidad son los más económicos. La diferencia quizá no sea muy radical para este caso, dado a que la variación entre los costo no es extrema, sin embargo si el costo de muestrear una unidad fuera muy elevado, esta diferencia representaría un importante beneficio económico.

Para este problema, las soluciones estocásticas difieren por casi 4 unidades, en la muestra global, de las dadas mediante Multiplicadores de Lagrange, y aunque el presupuesto se ve incrementado la diferencia en unidades muestrales podría representar una mayor aportación en términos de precisión.

Conclusiones.

Los resultados desarrollados a través de este documento se pueden resumir de forma sencilla en los siguientes puntos:

- En este trabajo se logro proponer formalmente el problema de asignacińń óptima a través de optimizacin estocástica, esto bajo la consideración de que s_h^2 y c_h, $h = 1, 2, \ldots, H$ son variables aleatorias, enfatizando la condición de que n_h sea entero.

- Se obtuvieron los planteamientos determinísticos equivalentes a los problemas estocásticos, en asignación óptima, para 6 casos de interés.

- Una aportación adicional en este trabajo, es que se desarrollo y uso de notación matricial, para el capiítulo II, en la teoría de muestreo estratificado.

LITRATURA CITADA

Beale, E. M. L. (1955). On Minimizing a Convex Fuction Subject to Lineal Inequalities. *Journal of the Royal Statistical Society*, **17** b, págs.173-184.

Benders, J. F. (1962). *Partitioning Produres for solving Mixed-Variables Programming Problems. Numerische Mathematik,* **4** págs.238-252.

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

Dantzig, G. B. (1955). Linear Programming Under Uncertainty. *Management Science,* **1**, No. 3-4, págs. 197-206

Dantzig, G. B. (1963). *Linear Programming and Extensions.* Princeton University Press. Princeton, USA.

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

Melaku, A. (1986). Asymptotic Normality of the Optimal Allocation in Multivariate Stratified Randon Sampling. *The Indian Journal of Statistics*, **48**, Series B, págs.224-232.

Peña, D. (2002). *Análisis de datos multivariantes.* McGraw-Hill.

Prékopa A. (1978). *"The use of stochastic programing for the solution of the some ploblems* *in statistics and probability".* Technical Summary report 1983. University of Wisconsin-Madison.

Ramos, A. y Cerisela, S. (2005). *Optimización Estocástica.* Universidad Ponticifia de Madrid.

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

Searle, Sh. R. (1987). *Matrix Algebra Useful For Statistics* John Wiley and Sons, Inc.

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

Van Slyke, R. M. and Wets, R. J. B. (1962). L-Shaped Linear Programs with Application to Optimal Control and Stochastic Programming. *SIAM Journal an Applied Mathematical*, **17**, No. 4, págs. 638-663.

APÉNDICE A

Tópicos de algebra de matrices.

Producto Hadamar.

El producto Hadamar de las matrices $\mathbf{A} = (a_{ij})$ y $\mathbf{B} = (b_{ij})$, se denota por $\mathbf{A} \odot \mathbf{B}$, y está definido solo cuando $\mathbf{A}$ y $\mathbf{B}$ tienen la misma dimensión. La matriz es formada por el producto elemento a elemento de los correspondientes elementos en $\mathbf{A}$ y $\mathbf{B}$:

$$\mathbf{A} \odot \mathbf{B} = (a_{ij}b_{ij}),$$

entonces el ij-ésimo elemento del producto Hadamar $\mathbf{A} \odot \mathbf{B}$ es el producto de los ij-ésimos elementos de $\mathbf{A}$ y $\mathbf{B}$.

Ejemplo:

$$\begin{bmatrix} 2 & 0 & -1 \\ 4 & 1 & 3 \end{bmatrix} \odot \begin{bmatrix} 1 & 3 & 7 \\ 2 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -7 \\ 8 & 0 & 18 \end{bmatrix}$$

Vea Searle (1987).

Tópicos de estadística.

Esperanzas y varianzas de vectores aleatorios.

Sea $\mathbf{x}$ un vector aleatorio de dimensión p, y se define un nuevo vector aleatorio $\mathbf{y}$ de dimensión $m, (m \leq p)$, con

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{108}$$

donde $\mathbf{A}$ es una matriz de constantes rectangular de dimensiones $m \times p$. Sean μ_x, μ_y los vectores de medias respectivos a los vectores aleatorios $\mathbf{x}$, $\mathbf{y}$, y además $\mathbf{V}_x$, $\mathbf{V}_y$ sus matrices de covarianzas, entonces

$$\mu_y = \mathbf{A}\mu_x, \tag{109}$$

y, además

$$\mathbf{V}_y = \mathbf{A}\mathbf{V}_{\mathbf{x}}\mathbf{A}^{'}, \tag{110}$$

donde $\mathbf{A}^{'}$ es la matriz transpuesta de $\mathbf{A}$. Vea Peña (2002).

Condiciones concernientes a la distribución límite de las varianzas muestrales.

En orden de establecer las condiciones bajo las cuales las varianzas muestrales s_h^2, $h = 1, 2, \ldots, H$ son asintóticamentes normales, para el caso de la matriz diagonal de covarianzas, una versión modificada del teorema de Hajek (Hajek, 1960) está dada por

Teorema 1. *La variable aleatoria ξ_h tiene una distribución asintóticamente normal con*

$$\xi_h \xrightarrow{a} N(\mathrm{E}(\xi_h), \mathrm{Var}(\xi_h)),$$

si y sólo si la condición de Noether se cumple

$$\text{lím}_{h\to\infty}\frac{\max\limits_{1\le i\le N_h}\ [(x_{hi}-\bar{X}_h)-S_h^2]^2}{N_h S_h^2}=0$$

Para $n_h \to \infty$, $N_h - n_h \to \infty$ y $N_h \to \infty$ donde

$$\xi_h = \frac{1}{n_h-1}\sum_{i=1}^{n_h}(x_{hi}-\bar{X}_h)^2 \qquad \text{y} \qquad \bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} x_{hi} \tag{111}$$

$$\mathrm{E}(\xi_h) = \frac{n_h}{n_h-1}S_h^2 \qquad \text{y} \qquad S_h^2 = \frac{1}{N_h}\sum_{i=1}^{N_h}(x_{hi}-\bar{X}_h)^2 \tag{112}$$

$$\mathrm{Var}(\xi_h) = \frac{n_h}{(n_h-1)^2}[C_{y_h}^4 - (S_h^2)^2] \qquad \text{y} \qquad C_{x_h}^4 = \frac{1}{N_h}\sum_{i=1}^{N_h}(x_{hi}-\bar{X}_h)^4, \tag{113}$$

donde $C_{x_h}^4$ es el cuarto momento central, y n_h es el tamaño de muestra, para una muestra aleatoria simple de la h-ésima subpoblación de tamaño N_h.

Observe que

$$\begin{aligned} s_h^2 &= \frac{1}{n_h-1}\sum_{i=1}^{n_h}(x_{hi}-\bar{x}_h)^2 \\ &= \xi_h - \frac{n_h}{n_h-1}(\bar{x}_h-\bar{X}_h)^2. \end{aligned} \tag{114}$$

donde

$$\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}, \qquad \frac{n_h}{n_h-1} \to 1$$

y

$$(\bar{x}_h - \bar{X}_h)^2 \to 0 \quad \text{en probabilidad.}$$

Por consecuencia, s_h^2 tiene una distribución asintóticamente normal, con media y varianza asintótica, dada por (112) y (113), vea Melaku (1986).

Tópicos de muestreo.

Varianzas de las estimaciones.

La varianza de los valores y_i en una población finita usualmente se define como

$$\sigma^2 = \frac{\sum_{1}^{N}(y_i - \bar{Y})^2}{N}.$$

Por motivos de notación, los resultados se presentan en términos de una expresión ligeramente diferente, en donde el divisor $(N-1)$ se usa en lugar de N. Entonces se tiene:

$$S^2 = \frac{\sum_{1}^{N}(y_i - \bar{Y})^2}{N-1}.$$

Esta convención la usan quienes enfocan la teoría del muestreo por medio del análisis de varianza. Su ventaja es que la mayoría de los resultados toman una forma ligeramente más simple. Siempre y cuando se use consistentemente la misma notación, todos los resultados son equivalentes en ambas notaciones.

Varianza de la media muestral.

El siguiente resultado pertenece a la teoría del muestreo aleatorio simple, vea Cochran (1977).

Teorema 2. *La varianza de la media $\bar{y}$ de una muestra aleatoria simple es*

$$Var(\bar{y}) = E(\bar{y} - \bar{Y})^2 = \frac{S^2(N-n)}{nN} = \frac{S^2}{n}(1-f) \qquad (115)$$

en donde $f = \frac{n}{N}$ es la fracción de muestreo, y $(1-f)$ es el factor de corrección por finitud.

Demostración.

Sea

$$n(\bar{y}-\bar{Y}) = (y_1-\bar{Y}) + (y_2-\bar{Y}) + \ldots + (y_n-\bar{Y}). \tag{116}$$

Observe además que

$$\mathrm{E}(y_1+y_2+\ldots+y_n) \;\text{ debe ser un múltiplo de }\; y_1+y_2+\ldots+y_N. \tag{117}$$

El multiplicador debe ser n/N, ya que la expresión de la izquierda tiene n términos y la de la derecha tiene N términos.

Por el argumento anterior, se tiene que

$$\mathrm{E}[(y_1-\bar{Y})^2+(y_2-\bar{Y})^2+\ldots+(y_n-\bar{Y})^2] = \frac{n}{N}[(y_1-\bar{Y})^2+(y_2-\bar{Y})^2+\ldots+(y_N-\bar{Y})^2] \tag{118}$$

así como también

$$\begin{aligned} &\mathrm{E}[(y_1-\bar{Y})(y_2-\bar{Y}) + (y_1-\bar{Y})(y_3-\bar{Y}) + \ldots + (y_{n-1}-\bar{Y})(y_n-\bar{Y})] \qquad (119)\\ &= \frac{n(n-1)}{N(N-1)}[(y_1-\bar{Y})(y_2-\bar{Y}) + (y_1-\bar{Y})(y_3-\bar{Y}) + \ldots + (y_{N-1}-\bar{Y})(y_N-\bar{Y})]. \end{aligned}$$

En la ecuación (119), la suma de los productos se extiende sobre todas las parejas de unidades en la muestra y en la población, respectivamente. La suma del lado izquierdo contiene $n(n-1)/2$ términos, y la suma de la derecha $N(N-1)/2$ términos.

Ahora elevando al cuadrado la ecuación (116), y promediando sobre todas las muestras aleatorias simples. Usando (118) y (119) se obtiene

$$\begin{aligned} n^2\mathrm{E}(\bar{y}-\bar{Y})^2 &= \frac{n}{N}\Big\{ (y_1-\bar{Y})^2 + \ldots + (y_N-\bar{Y})^2 \\ &+ \frac{2(n-1)}{N-1}[(y_1-\bar{Y})(y_2-\bar{Y}) + \ldots + (y_{N-1}-\bar{Y})(y_N-\bar{Y})]\Big\}. \end{aligned} \tag{120}$$

Completando el cuadrado en el término que incluye el doble producto, se tiene

$$\begin{aligned} n^2\mathrm{E}(\bar{y}-\bar{Y})^2 &= \frac{n}{N}\{\left(1-\frac{n-1}{N-1}\right)[(y_1-\bar{Y})^2+\ldots+(y_N-\bar{Y})^2] \\ &+ \frac{n-1}{N-1}[(y_1-\bar{Y})+\ldots+(y_N-\bar{Y})]^2\}. \end{aligned} \tag{121}$$

El segundo término dentro de la llave desaparece, ya que la suma de y_i es igual a $N\bar{Y}$.
La división entre n^2 da

$$\mathrm{Var}(\bar{y}) = \mathrm{E}(\bar{y}-\bar{Y})^2 = \frac{N-n}{nN(N-1)}\sum_{i=1}^{N}(y_i-\bar{Y})^2 = \frac{S^2(N-n)}{nN}. \tag{122}$$

■

Estimador insesgado s^2.

Muchas fórmulas en la teoría de muestreo involucran a S^2, la varianza de la población. Esta no será conocida en la práctica pero puede ser estimada de los datos de la muestra. Un resultado importante se establece en el siguiente teorema.

Teorema 3. *En una muestra aleatoria simple,*

$$s^2 = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}$$

es un estimador insesgado de

$$S^2 = \frac{\sum_{i=1}^{N}(y_i-\bar{Y})^2}{N-1}.$$

Demostración.

Note que s^2 se puede escribir como

$$\begin{aligned} s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}\left[(y_i-\bar{Y})^2-(\bar{y}-\bar{Y})^2\right] && (123)\\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i-\bar{Y})^2-n(\bar{y}-\bar{Y})^2\right]. && (124)\end{aligned}$$

Ahora, tomando el promedio sobre todas las muestras simples al azar de tamaño n. Por el argumento de simetría usado en el teorema 5.0.2,

$$E\left[\sum_{i=1}^{n}(y_i-\bar{Y})^2\right]=\frac{n}{N}\sum_{i=1}^{N}(y_i-\bar{Y})^2=\frac{n(N-1)}{N}S^2$$

por definición de S^2. Además, por el teorema 5.0.2,

$$E[n(\bar{y}-\bar{Y})^2]=\frac{N-n}{N}S^2.$$

Por lo tanto,

$$E(s^2)=\frac{S^2}{(n-1)N}[n(N-1)-(N-n)]=S^2.$$

■

<u>Determinación del tamaño de muestra.</u>

Los distintos métodos para distribuir las observaciones en los estratos proporcionan los tamaños relativos de muestras $\mathbf{n}/n$. Después de construir los estratos y distribuir las observaciones en estos, se puede utilizar la ecuación (21) para determinar el tamaño de muestra necesario para lograr una varianza predeterminada. Como

$$\mathrm{Var}(\hat{t}_{st})\leq\frac{1}{n}\left[\mathbf{N}^{'}(n)\mathrm{diag}\left(\frac{S_h^2}{n_h}\right)\mathbf{N}\right]=\frac{\nu}{n}. \qquad (125)$$

Un intervalo de confianza aproximado del 95% sería $\hat{t}_{st}\pm z_{\alpha/2}\sqrt{\nu/n}$, esto siempre y cuando se pueda ignorar la corrección para poblaciones finitas y además sea válida la aproximación normal. Cumpliéndose estas condiciones, se hace $n = z^2_{\alpha/2}\nu/e^2$ para obtener un intervalo de confianza con la mitad de ancho e. Este enfoque requiere de conocer los valores S^2_h, vea Lohr (2000).

APÉNDICE B

Datos utilizados en los ejemplos.

El conjunto de datos en la Tabla 7, contiene 187 observaciones seleccionadas aleatoriamente, mediante un estudio previo. El valor en los datos corresponde a la medición de la longitud de una hoja en cierta especie de "flor", que es cultivada en 4 diferentes regiones. Los 187 datos se obtuvieron en las 4 regiones, 49 en la primera, 48 en la segunda, 47 en la tercera y 43 en la cuarta. La variable que se considero para la estratificación es la región, de tal forma que la población es dividida en 4 estratos, sin traslape, uno para cada región. Los vectores con los valores utilizados en los ejemplos del Capítulo II y IV, son los siguientes.

El vector de unidades poblacionales correspondientes a cada estrato es $\mathbf{N}' = (2500,\ 2300,\ 2800,\ 3000)$. El total de la población es $N = 10600$

Vector de unidades en cada estrato obtenidas en el estudio previo.

$$\mathbf{n} = \begin{bmatrix} 49 \\ 48 \\ 47 \\ 43 \end{bmatrix}.$$

Vector de pesos o ponderaciones.

$$\mathbf{W} = \begin{bmatrix} .2358 \\ .2170 \\ .2642 \\ .2830 \end{bmatrix}.$$

Vector de costos por unidad en cada estrato.

$$\mathbf{c} = \begin{bmatrix} 25 \\ 30 \\ 15 \\ 25 \end{bmatrix}.$$

Vector de varianzas en cada estrato.

$$\mathbf{s^2} = \begin{bmatrix} .1694 \\ 8.4317 \\ .0972 \\ 3.8590 \end{bmatrix}.$$

Vector correspondiente al cuarto momento muestral, en cada estrato.

$$\mathbf{C^4} = \begin{bmatrix} .0884 \\ 330.4106 \\ .0319 \\ 34.1001 \end{bmatrix}.$$

Cuadro B.1: Tabla de datos.

Fuente: Datos simulados.

Estrato 1	Estrato 2	Estrato 3	Estrato 4	Estrato 1	Estrato 2	Estrato 3	Estrato 4
3.625	11.000	3.154	2.641	3.000	6.000	3.267	3.530
2.125	8.000	3.300	2.497	2.000	6.000	3.917	0.796
3.056	4.750	3.308	6.446	2.667	7.000	3.357	3.155
2.550	8.000	2.941	1.625	3.111	14.000	2.769	2.862
3.056	7.000	2.267	4.383	2.333	2.667	3.077	6.504
3.211	6.500	3.077	2.806	2.478	5.667	3.143	5.956
2.450	6.500	3.455	3.189	2.545	8.000	3.000	6.729
3.222	8.000	3.000	2.061	3.500	9.500	3.077	6.039
2.400	4.353	3.143	7.957	2.440	4.667	3.000	3.174
3.333	4.000	3.333	4.300	2.217	7.000	3.667	5.780
3.400	6.500	3.357	5.455	2.636	3.250	2.667	5.750
2.571	7.500	3.286	2.298	2.684	7.500	3.067	0.425
2.667	15.000	3.000	3.975	4.000	7.500	3.462	8.207
2.261	7.500	2.938	5.489	3.143	4.333	3.188	4.801
2.304	7.000	3.545	4.182	2.647	3.750	4.100	3.467
2.500	7.000	3.250	0.233	2.833	3.750	3.308	6.643
3.350	15.000	3.500	4.643	2.722	4.250	3.700	3.074
2.722	4.667	3.133	5.409	2.632	7.000	3.385	5.870
3.333	7.500	2.800	3.759	2.280	3.750	4.000	2.439
2.652	8.000	2.813	3.469	2.619	5.000	3.231	3.132
2.789	3.400	3.300	5.476	2.684	4.667	3.000	6.029
2.565	5.000	3.500	7.217	2.333	6.500	3.538	
3.045	7.500	2.786		2.667	7.000	3.231	
2.348	7.000	3.500		3.200	14.000		
2.714							