



Universidad de Valladolid

FACULTAD DE CIENCIAS

DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

Tesis Doctoral:

Análisis de eficiencia de unidades productivas que prestan servicios al público: la opinión del cliente como output estocástico.

Presentada por Jesús A. Tapia García para optar al grado de Doctor por la
Universidad de Valladolid

Dirigida por:

Bonifacio Salvador González y Jesús María Rodríguez Rodríguez

Octubre 2015

*Dedicado
a mis padres, a mi hijo José
y a los que han confiado en mí.*

Agradecimientos

Por fin llega la parte más agradable de este trabajo, agradecer. Por supuesto lo primero que quiero es agradecer a mis padres, su enorme esfuerzo para que yo haya podido alcanzar las metas que, por sus circunstancias, ellos no pudieron lograr, con la fortuna de que, a pesar de sus dificultades, lo hayan podido ver y disfrutar.

La segunda parte de los agradecimientos es para los que han confiado en mí. A mi Director de tesis Jesús M. porque fue el primero en motivarme para llevar a cabo este proyecto y porque, junto a la Dirección de la Dirección General de Estadística de la Junta de Castilla y León, confiaron en mí para desarrollar un trabajo motivador de alguno de los problemas tratados en esta tesis. A Boni, que con su paciencia, trabajo y esfuerzo ha conseguido dirigir y encauzar mi trabajo y del que he aprendido a investigar y, sobre todo, a disfrutar de investigar. A José Antonio, que desde el primer día me ayudó a desarrollar mi trabajo e integrarme en el Departamento. A Cristina con quien tanto tiempo he compartido y de quien tantas veces he recibido su apoyo. No puedo olvidar a Juan, excepcional persona, compañero y profesional que me inició en el estudio de buena parte de los contenidos tratados en esta tesis.

Tampoco quiero dejar de agradecer a todos mis compañeros del Departamento por su actitud e interés en que yo lograra este proyecto y, particularmente, a Alfonso por mediar en el procedimiento de tramitación de mi estancia temporal en la administración autonómica para trabajar en la estadística pública.

Por último quiero agradecer a Teresa, la que más ha confiado en mí, por haber estado ahí para escucharme y animarme en todo momento.

Índice general

| | |
|---|-----------|
| Agradecimientos | 5 |
| Lista de Figuras | 9 |
| Lista de Tablas | 11 |
| Introducción | 13 |
| | |
| I Metodología muestral en encuestas de opinión de usuarios de servicios: el tiempo de atención como medida para tomar la muestra | 19 |
| | |
| Objetivos y motivación | 21 |
| | |
| 1. Preliminares | 23 |
| 1.1. Diseños Muestrales | 23 |
| 1.2. Estimación Bootstrap de la varianza de los estimadores | 29 |
| | |
| 2. Metodología muestral en servicios con clientes | 35 |
| 2.1. Notación | 35 |
| 2.2. Planteamiento del Problema | 36 |
| 2.3. Encuesta e información auxiliar | 38 |
| 2.4. Diseño muestral Tiempo-Espacio | 39 |
| 2.4.1. Selección de Espacios | 39 |
| 2.4.2. Selección de Servicios | 40 |
| 2.4.3. Selección de Unidades de Tiempo | 40 |
| 2.4.4. Selección de individuos | 42 |
| 2.4.4.1. Censo de Individuos | 42 |
| 2.4.4.2. Muestra de individuos proporcional a su tiempo de atención | 44 |
| 2.4.4.3. Postestratificación de unidades de tiempo | 47 |
| 2.4.5. Efecto del Diseño | 48 |
| 2.4.6. Coste del Diseño | 49 |
| 2.5. Validación de la metodología muestral | 49 |
| 2.5.1. Simulación de una población con todas las unidades de tiempo de uso saturado | 50 |
| 2.5.2. Simulación de una población con unidades de tiempo de uso saturado y no saturado | 54 |
| 2.6. Estudios de aplicación del censo de usos | 58 |
| 2.6.1. Estudio social de opinión de usuarios de bibliotecas | 58 |
| 2.6.2. Encuesta a usuarios de servicios públicos de la administración | 59 |
| 2.6.3. Sondeo en día electoral | 62 |
| 2.7. Metodología alternativa de selección de la muestra de UDTs | 63 |

| | |
|---|------------|
| 2.8. Estimación en población de usuarios inferior a la de usos | 64 |
| Conclusiones | 69 |
| II Análisis de eficiencia con outputs estimados | 71 |
| Objetivos y motivación | 73 |
| 3. Preliminares | 75 |
| 3.1. Análisis envolvente de datos | 75 |
| 3.2. Eficiencia DEA con datos imprecisos o estocásticos | 82 |
| 3.3. DEA en servicios públicos | 92 |
| 4. DEA con outputs estimados | 95 |
| 4.1. Planteamiento del problema | 95 |
| 4.2. Cálculo del tamaño muestral | 97 |
| 4.2.1. Modelo CCR-O con un input conocido y un output estocástico | 98 |
| 4.2.2. Modelos CCR-O y BCC-O con más de un input conocido y un output estocástico | 102 |
| 4.2.3. Un estudio de aplicación | 105 |
| 4.3. Intervalo de confianza de eficiencia DEA (ICE) | 107 |
| 4.3.1. Intervalo de confianza <i>Tipo1</i> | 107 |
| 4.3.2. Intervalo de confianza <i>Tipo2</i> | 108 |
| 4.3.3. Intervalo de confianza <i>Tipo3</i> | 108 |
| 4.3.4. Intervalo de confianza <i>Tipo4</i> | 109 |
| 4.3.5. Intervalo de confianza Bootstrap | 114 |
| 4.3.6. Intervalo observado | 115 |
| 4.3.7. Evaluación de los métodos de obtención de Intervalos de confianza DEA . . | 116 |
| 4.3.8. Un estudio de aplicación | 123 |
| 4.4. Alternativa de Intervalo de eficiencia DEA | 126 |
| 4.5. Región de confianza simultánea | 126 |
| 4.5.1. Modelo CCR-O con un input conocido y un output estocástico | 127 |
| 4.5.2. Modelos CCR-O y BCC-O con más de un input conocido y un output estocástico | 129 |
| 4.5.3. Un estudio de aplicación | 135 |
| Conclusiones y Líneas futuras de Investigación | 137 |
| A. Códigos de los programas MATLAB del Bloque I | 139 |
| A.1. Simulación de poblaciones | 139 |
| A.2. Diseños muestrales | 147 |
| A.3. Simulaciones de confianza de intervalos, de errores de estimación y de efecto del diseño | 148 |
| B. Códigos de los programas MATLAB del Bloque II | 149 |
| B.1. Simulación de poblaciones | 149 |
| B.2. Cálculo de tamaño de muestra | 149 |
| B.3. Índices de eficiencia DEA. Modelos BCC y CCR. | 150 |
| B.4. Intervalos de confianza de eficiencia DEA | 152 |
| B.5. Región de confianza DEA | 164 |
| Bibliografía | 173 |

Índice de figuras

| | |
|---|-----|
| 2.1. Marco de dos semanas con unidad de tiempo la hora, en un espacio con cuatro servicios | 38 |
| 2.2. Marco y muestra de unidades de tiempo en cuatro servicios ubicados en un mismo espacio | 41 |
| 2.3. Población con todas las UDTs saturadas: Error de estimación del total de usos \hat{N} | 52 |
| 2.4. Población con todas las UDTs saturadas: Error de estimación del total de usuarios \hat{N}_u con tamaño de usos, N , desconocido y conocido | 52 |
| 2.5. Población con todas las UDTs saturadas: Error de estimación de la media poblacional de usos $\hat{\mu}$ | 53 |
| 2.6. Población con todas las UDTs saturadas: Error de estimación de la media poblacional de usuarios $\hat{\mu}_u$ | 53 |
| 2.7. Población con todas las UDTs saturadas: Efecto del diseño (censo con y sin tamaño de población conocido - polietápico) en la estimación de la satisfacción media de los usos | 54 |
| 2.8. Población con UDTs saturadas y no saturadas: Error de estimación del total de usuarios, \hat{N}_u | 56 |
| 2.9. Población con UDTs saturadas y no saturadas: Error de estimación de la media poblacional de usos, $\hat{\mu}$ | 56 |
| 2.10. Población con UDTs saturadas y no saturadas: Error de estimación de la media poblacional de usuarios, $\hat{\mu}_u$ | 57 |
| 2.11. Población con UDTs saturadas y no saturadas: Efecto del diseño (censo con tamaño de población de usos conocido - censo con tamaño de población de usos desconocido - polietápico con postestratificación) en la estimación de la satisfacción media de los usos | 57 |
| 2.12. Marco y muestra de unidades de tiempo en una biblioteca del Campus de la UVa | 59 |
| 2.13. Marco y unidades de tiempo muestrales de 14 servicios localizados en cuatro espacios diferentes en la provincia de Zamora | 61 |
| 2.14. Marco y UDTs muestrales para dos colegios electorales encuestados por el mismo equipo de encuestadores | 62 |
| 4.1. Diferencias absolutas entre extremos inferiores de ICEs <i>Tipo3</i> y <i>Tipo4</i> . Modelo CCR. | 112 |
| 4.2. Diferencias absolutas entre extremos superiores de ICEs <i>Tipo3</i> y <i>Tipo4</i> . Modelo CCR. | 112 |
| 4.3. Diferencias absolutas entre extremos inferiores de ICEs <i>Tipo3</i> y <i>Tipo4</i> . Modelo BCC. | 113 |
| 4.4. Diferencias absolutas entre extremos superiores de ICEs <i>Tipo3</i> y <i>Tipo4</i> . Modelo BCC. | 113 |
| 4.5. Amplitud de intervalos ICEs por tipo y DMU. Modelo CCR-O, $\alpha = 0.1$ y $1 - \alpha' = 0.98$ | 121 |
| 4.6. Amplitud de intervalos ICEs por tipo y DMU. Modelo BCC-O, $\alpha = 0.1$ y $1 - \alpha' = 0.98$ | 122 |
| 4.7. Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.1$ y $\frac{2p}{1+p} = 0.1818$. Modelo CCR-O. | 131 |
| 4.8. Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.1$ y $\frac{2p}{1+p} = 0.1818$. Modelo BCC-O. | 131 |

| | |
|--|-----|
| 4.9. Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.05$ y $\frac{2p}{1+p} = 0.0952$. Modelo CCR-O. | 131 |
| 4.10. Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.05$ y $\frac{2p}{1+p} = 0.0952$. Modelo BCC-O. | 132 |
| 4.11. Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo CCR-O, $\delta = 0.1$ y $\alpha = 0.05$ | 133 |
| 4.12. Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo BCC-O, $\delta = 0.1$ y $\alpha = 0.05$ | 133 |
| 4.13. Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo CCR-O, $\delta = 0.1$ y $\alpha = 0.1$ | 134 |
| 4.14. Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo BCC-O, $\delta = 0.1$ y $\alpha = 0.1$ | 134 |
| 4.15. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo CCR-O, $\delta = 0.15$ y $\alpha = 0.05$ | 134 |
| 4.16. Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo BCC-O, $\delta = 0.15$ y $\alpha = 0.05$ | 135 |

Índice de tablas

| | |
|--|----|
| 2.1. Parámetros de la población simulada con todas las UDTs saturadas | 50 |
| 2.2. Población con todas las UDTs saturadas: Aproximación de confianza y error de estimación del total de usuarios con con tamaño de población de usos, N , desconocido | 51 |
| 2.3. Población con todas las UDTs saturadas: Aproximación de confianza y error de estimación del total de usuarios con con tamaño de población de usos, N , conocido . | 52 |
| 2.4. Población con todas las UDTs saturadas: Estimación de la media poblacional de usos | 52 |
| 2.5. Población con todas las UDTs saturadas: Estimación de la media poblacional de usuarios | 53 |
| 2.6. Parámetros de la población simulada con UDTs saturadas y no saturadas | 54 |
| 2.7. Población con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población de usos, N , desconocido | 55 |
| 2.8. Población con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población de usos, N , conocido | 55 |
| 2.9. Población con UDTs saturadas y no saturadas: Estimación de la media poblacional de usos | 55 |
| 2.10. Población con UDTs saturadas y no saturadas: Estimación de la media poblacional de usuarios | 55 |
| 2.11. Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Parámetros de la población simulada | 64 |
| 2.12. Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación del total de usuarios con tamaño de población de usos, N , desconocido . | 64 |
| 2.13. Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación del total de usuarios con tamaño de población de usos, N , conocido . . . | 65 |
| 2.14. Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación de la media poblacional de usos | 65 |
| 2.15. Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación de la media poblacional de usuarios | 65 |
| 2.16. Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Parámetros de la población simulada | 66 |
| 2.17. Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población N desconocido . | 66 |
| 2.18. Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población N conocido . . . | 66 |
| 2.19. Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación de la media poblacional de usos | 66 |
| 2.20. Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación de la media poblacional de usuarios | 67 |
| 3.1. Modelos DEA CCR y BCC. Orientaciones input y output | 80 |
| 3.2. E-modelo chance-constrained DEA. Orientación output. | 84 |

| | |
|---|-----|
| 3.3. E-modelo chance-constrained DEA con inputs y outputs estadísticamente independientes y normalmente distribuidos. Orientación output. | 84 |
| 3.4. Método optimista/pesimista de Entani y otros (2002) [52] de obtención del intervalo de eficiencia | 89 |
| 4.1. Modelos DEA, para la DMU _o , con rendimientos constantes (CCR) y variables de escala (BCC); orientación output. | 97 |
| 4.2. Número de médicos, enfermeras, pacientes dados de alta e ingresados en 12 centros de salud | 102 |
| 4.3. Modelo poblacional simulado | 103 |
| 4.4. Tamaño muestral obtenido para cada DMU: Previo y tomando $\delta = 0.2$ y $\delta = 0.1$ | 104 |
| 4.5. Aproximación de la probabilidad: $p\left(\varphi_i \in \left[\widehat{\Phi}_i - \delta, \min\left\{\widehat{\Phi}_i + \delta, 1\right\}\right]\right)$ | 104 |
| 4.6. Proporción de veces que los valores estimados de los índices de eficiencia toman el valor 1 | 105 |
| 4.7. Muestra aleatoria simple previa: tamaño de muestra, estimación de la satisfacción media de los usuarios y de la varianza poblacional | 106 |
| 4.8. Base de datos compuesta por 15 bibliotecas: Tamaño poblacional, préstamos de libros y puestos para usuarios (inputs), tamaño muestral y estimación de la satisfacción media de los usuarios (output) con $\delta = 0.1$, $\alpha = 0.1$, valores estimados del índice de eficiencia BCC-O, e intervalo del Teorema 2 | 106 |
| 4.9. Extremo inferior y superior de los ICEs <i>Tipo3</i> y <i>Tipo4</i> fijado $p = 0.05$ y $\alpha = 0.1$. Modelo CCR. | 114 |
| 4.10. Extremo inferior y superior de los ICEs <i>Tipo3</i> y <i>Tipo4</i> fijado $p = 0.05$ y $\alpha = 0.1$. Modelo BCC. | 114 |
| 4.11. Tamaño muestral obtenido para cada DMU fijado un error de estimación para los outputs $e=0.9$ y una confianza $1-\alpha=0.9$ | 117 |
| 4.12. Confianza simulada de los ICEs. Modelo CCR-O. | 118 |
| 4.13. Confianza simulada de los ICEs. Modelo BCC-O. | 118 |
| 4.14. Valor esperado de los extremos de los intervalos de confianza de eficiencia ICEs, $\alpha=0.1$ y $1-\alpha'=0.98$. Modelo CCR | 119 |
| 4.15. Valor esperado de los extremos de los intervalos de confianza de eficiencia ICEs, $\alpha=0.1$ y $1-\alpha'=0.98$. Modelo BCC | 120 |
| 4.16. Tamaño de Personal y de Gasto del Sistema Sanitario Español por CCAA en 2009 (inputs), Tamaño muestral y estimación de la satisfacción media con el Sistema Sanitario Español en 2009 (output). Índices de eficiencia estimados CCR-O y BCC-O. | 124 |
| 4.17. Eficiencia sanitaria estimada en CCAAs de manera puntual y por intervalo utilizando datos de la Estadística de Establecimientos Sanitarios con Régimen de Internado y Barómetro Sanitario del 2009. Modelo BCC-O. | 125 |
| 4.18. Clasificación de las CCAAs en los conjuntos E^{++} de unidades eficientes en cualquier caso, E^+ de unidades eficientes en sentido maximal y E^- de las unidades ineficientes. | 125 |
| 4.19. Simulación Teorema 5. Tamaño de muestra | 132 |
| 4.20. Teorema 5. Confianza aproximada de la región simultánea del vector de eficiencias | 132 |
| 4.21. Estudio de aplicación. Resultados de la muestra aleatoria simple previa: tamaño de muestra, estimación de la satisfacción media de los usuarios y de la varianza poblacional. | 136 |
| 4.22. Estudio de aplicación. Tamaño de muestra, estimación del índice de eficiencia y región de confianza para la eficiencia de 15 bibliotecas, $\delta = 0.1$, $\alpha = 0.1$. Modelos BCC-O y CCR-O. | 136 |

Introducción

Los estudios de opinión de los usuarios, y en particular de satisfacción, son herramientas ampliamente utilizadas para medir la percepción sobre la calidad de servicios concretos (Hill y Jim, 2003 [72]). De manera tradicional se consideran dos enfoques para definir la calidad del servicio, el cumplimiento de expectativas, entendido como la aptitud de un servicio para satisfacer las necesidades de los usuarios, y el cumplimiento de especificaciones, capacidad de un servicio para cumplir las especificaciones con las que fue diseñado (Ruiz y otros, 2006 [117]).

Tanto en el ámbito público como privado, son inmediatas las referencias que ponen de manifiesto la relevancia de los estudios de opinión-satisfacción, entendiendo por nivel de satisfacción de los usuarios de un servicio la relación entre la percepción sobre el servicio y las expectativas previas de los usuarios. Dentro de la Administración Pública, al definir el Marco General para la Mejora de la Calidad en la Administración General del Estado, el Real Decreto 951/2005 de 29 de julio, artículos 5 y 6, contempla la norma reguladora de análisis de la demanda para la detección de necesidades y expectativas de los usuarios de los servicios públicos, así como mediciones sistemáticas y permanentes del grado de satisfacción de los usuarios con los servicios públicos que reciben. Dicha norma contempla explícitamente la utilización de técnicas cualitativas y cuantitativas de análisis y la aplicación de metodologías homologadas dentro del ámbito de la investigación social.

Otra referencia al interés de los estudios de opinión-satisfacción de usuarios o clientes se halla en la Norma ISO 9001:2008, orientada al aseguramiento de la calidad del producto y a aumentar la satisfacción del cliente. Esta Norma especifica los requisitos para certificar sistemas de gestión de la calidad. En concreto, en el capítulo 8, Medición, Análisis y Mejora, se encuentra el apartado de satisfacción de los clientes; la Norma establece lo siguiente respecto a este requisito: “8.2.1 Satisfacción del cliente. Como una de las medidas del desempeño del sistema de gestión de la calidad, la organización debe realizar el seguimiento de la información relativa a la percepción del cliente respecto al cumplimiento de sus requisitos por parte de la organización. Deben determinarse los métodos para obtener y utilizar dicha información”.

La planificación completa de la realización de una encuesta de opinión de usuarios precisa definir su alcance, qué se quiere medir y preguntar, a quién se va a preguntar, cuántas encuestas se van a realizar y con qué diseño se va a tomar la muestra, qué cuestionario se va a utilizar y como se va a recoger las respuestas, en qué momento temporal se van a realizar todas las entrevistas, cómo se va a tratar la información y cómo se van a presentar los resultados. En muchos casos prácticos, cuyos resultados forman parte del contenido de numerosos artículos, proyectos y tesis, los mayores esfuerzos se centran en decidir los objetivos, el cuestionario y en sintetizar, a través de herramientas estadísticas, la información recogida con las preguntas realizadas, relajando tanto el cómo tomar la muestra, muchas veces de manera no probabilística, como el cálculo previo del número de encuestas necesario. Sin embargo, atender a la precisión de estos dos aspectos determina que los resultados finales del experimento sean o no un reflejo verdadero de la realidad de toda la población objetivo.

En el primer bloque de esta tesis se va a proponer una solución a las cuestiones de a cuántos usuarios o clientes encuestar y cómo encuestarlos, cuando la población de referencia son las personas usuarias de algún servicio de un organismo o entidad pública o privada con unas características espacio-temporales específicas, lo que equivale a una población humana móvil (Sudman 1980 [132], Kalton 1991 [79], Himelein y otros 2014 [73]), y se utiliza un cuestionario como herramienta para

recoger la opinión del cliente con respecto al servicio del que acaban de hacer uso. La literatura con respecto a diseños muestrales utilizados en encuestas de opinión de usuarios de servicios es muy prolija, siendo frecuente encontrar el uso de diseños no probabilísticos, como por ejemplo el estudio de satisfacción en el sector bancario (Mihelis y otros, 2001G. [101]) en que no se especifica diseño muestral, o el estudio en la industria griega de ferris (Grigoroudis y otro 2009 [68]) en que la forma de proceder es repartir 5000 cuestionarios entre los pasajeros recogiéndose 605 contestaciones, o el estudio de opinión de usuarios de una cadena de cafeterías (Hayes, Bob E. 2008 [71]) en que la participación en el estudio es voluntaria. Entre algunas de las referencias de encuestas de opinión a clientes de servicios en las que la muestra se toma con diseño probabilístico, se encuentra la encuesta sobre Calidad de vida en la ciudad y la satisfacción con los servicios públicos en el caso de la ciudad de Madrid, donde la ficha técnica especifica que el tipo de muestreo es polietápico, con sección censal como unidad de primera etapa y selección aleatoria en todas las etapas del muestreo (Carrillo y Tamayo, 2008 [17]). Otro ejemplo interesante es la línea de investigación de Sociología Tributaria del Instituto de Estudios Fiscales, que recoge la percepción social de la fiscalidad en España a través de una encuesta personal en que se toma una muestra de 7800 individuos con diseño muestral polietápico, estratificado por conglomerados, con selección de las unidades primarias de muestreo (municipios) y de las unidades secundarias (secciones) de forma aleatoria proporcional y de las unidades últimas, individuos, por rutas aleatorias y cuotas de sexo y edad (Instituto de estudios fiscales 2009 [53]). Estos y otros ejemplos, Barómetro Sanitario 2011 del Ministerio de Sanidad, Servicios Sociales e Igualdad, ponen de manifiesto una metodología muestral, similar entre ellos, pero diferente a la que nosotros propondremos.

Las técnicas estándar de muestreo y estimación requieren la selección de una muestra de individuos en un marco muestral con una probabilidad de selección conocida. En la situación que vamos a plantear, a priori, no hay una lista de los individuos que usarán el servicio; el espacio y el tiempo serán lo que nos permita construir un marco, identificando unidades de tiempo con miembros de la población objetivo ubicados en espacios específicos (Muhib y otros 2001 [104]). El diseño muestral tiempo-espacio (time-space sampling) ha sido usado por numerosos experimentadores en poblaciones con alto riesgo de enfermedades infecciosas, entre otros: Gould y otros (1990) [67], Stueve y otros (2001) [131], Mansergh y otros (2006) [98], MacKellar y otros (2007) [97] y Ramo y otros (2010) [113]. La literatura concerniente a la aplicación, en otros ámbitos, de los diseños muestrales probabilísticos con marco muestral tiempo-espacio no es muy extensa (Sudman 1980 [132], Kalton, 1991 [79] y 2001 [78], Ardilly and Him Blanc 2001 [4], Figuera y otros 2011 [58], Golinelli y otros 2014 [66]). Sudman (1980) [132] discute como mejorar la calidad de las encuesta llevadas a cabo en centros comerciales introduciendo un diseño muestral tiempo-espacio que requiere el conocimiento previo, aproximado o estimado, del número total de clientes, en cada periodo de tiempo, de cara a obtener la probabilidad de selección de las unidades de tiempo (UDTs) y de los individuos. Ardilly and Him Blanc (2001) [4] y Golinelli y otros (2014) [66] proponen diseños con marco muestral tiempo-espacio con objeto de conocer mejor la población de los “sin techo”; en este caso se toma una muestra probabilística primero de centros, luego de días de atención, a continuación de intervalos de tiempo de los días y por último de personas para cada uno de los intervalos de tiempo, de día y de centro; mediante un diseño muestral diferente e independiente del de las personas, se seleccionan servicios para, finalmente, estimar los parámetros con los denominados estimadores “weight sharing”. Location sampling (Kalton, 1991 [79] y 2001 [78]) es un diseño bietápico donde las unidades de primera etapa son los periodos de tiempo o unidades de tiempo (UDT) y las unidades de segunda etapa son las visitas al centro. Si el flujo de visitas es constante en el tiempo, a todas las UDTs se las asigna la misma probabilidad de selección y se toma una muestra de usos sistemática, del mismo tamaño, en cada UDT. Si el flujo de clientes no es uniforme en el tiempo, las UDTs se toman con probabilidad proporcional a su número de visitas con objeto de que el tamaño de muestra sistemática de unidades de segunda etapa sea el mismo en cada UDT muestral.

Definir la población no es tarea sencilla dado que una misma persona puede utilizar varias veces algún

servicio en distintos momentos del tiempo de referencia o varios servicios ubicados en el mismo espacio físico, por lo que es importante diferenciar entre uso del servicio y usuario del servicio. En el caso de que no se puede, o no se quiere, identificar al usuario se considera como población de referencia los usos. Kalton (2001) [78] pone de manifiesto que utilizar los usuarios como unidad de análisis es complicado, primero, por el desconocimiento previo, exacto o estimado, del tamaño de esta población para poder hacer los cálculos del tamaño muestral necesario. Segundo, por la necesidad de identificar si el individuo tomado en la muestra ya ha sido encuestado y, si ya ha sido encuestado, como afecta a la hora de continuar tomando la muestra probabilística de usuarios. Por último, las inferencias en la población de usuarios siempre van a estar condicionadas a la necesidad de la estimación del tamaño de esta población.

Las principales novedades teóricas de los procedimientos expuestos en el primer bloque de la tesis son dos: en primer lugar, y a diferencia de Sudman (1980) [132] y Kalton (1991 [79] y 2011 [78]), la metodología muestral no precisa el conocimiento previo exacto o estimado del número de usos en las UDTs del marco; la segunda novedad es utilizar el tiempo de atención de los clientes para tomar una muestra probabilística de los mismos. Otros aspectos abordados son el cálculo del tamaño de muestra de UDTs y de clientes y la estimación de los parámetros tanto para la población de usos como de usuarios. Además, la metodología propuesta atiende a la consideración de que el coste es probablemente el aspecto más importante a tener en cuenta para llevar a cabo los nuevos métodos de muestreo (Brick J.M., 2011 [15]); la puesta en práctica de la metodología que vamos a proponer en la sección 2.4 es de coste “razonable” dado que, el trabajo de campo se puede llevar a cabo con entrevista cara-a-cara (face-to-face) y con un reducido número de encuestadores, sin perder el control del error de muestreo. Las encuestas cara-a-cara aúnan el potencial de hacer el uso conjunto de las comunicaciones oral y visual, mientras que, la presencia de un encuestador, hace posible obtener respuestas completas (Dillman 2005 [48]) introduciendo más oportunidades de interacciones complejas encuestado-encuestador (West y otros 2013 [149]). El mayor problema de este tipo de encuestas es su alto coste, aspecto que, en gran medida, condiciona su aplicación (Salant and Dillman 1994 [120]). Una nueva forma de economizar este tipo de encuestas es implementarlas, cuando resulta posible, de manera conjunta a la metodología Web (Tapia y Menéndez 2015 [134]).

En conclusión, este primer bloque proporciona una alternativa “económica” de diseño muestral probabilístico en encuestas cara-a-cara en las que se recoge la opinión de individuos sobre diversos aspectos de un servicio del que acaban de hacer uso.

El contenido presentado en el segundo bloque de la tesis tiene por objeto atender al problema de determinar la eficiencia de un conjunto homogéneo de unidades de producción que prestan un servicio (el mismo en todas las unidades) a clientes, utilizando, como outputs, la opinión de estos clientes, opinión estimada a partir de una muestra probabilística, y, como inputs, los recursos que pueden influir en esta opinión. Por ejemplo, en un hospital, el tiempo de estar hospitalizado y los recursos (camas, personal, . . .) pueden explicar la satisfacción media de los pacientes al ser dados de alta. La metodología que nos proponemos utilizar es el análisis envolvente de datos (DEA), introducido por Charnes y otros (1978) [24] y expuesta en la sección 3.1. El error de estimación del output se transfiere al cálculo de la eficiencia DEA, el proceso DEA pasa a ser de naturaleza estocástica (Sengupta 1987 [124]), por la variabilidad estocástica de los outputs, y la obtención de la eficiencia relativa no es elemental (Morita and Seiford 1999 [103]). Ceyhan and Benneyan (2014) [19] discuten diversas aproximaciones (Monte Carlo, bootstrapping (Simar and Wilson, 1998 [126]), chance-constrained y optimista/pesimista DEA (Entani y otros 2002)) al problema DEA en presencia de valores de proporciones estimadas con muestras, tales como proporción de defectos, satisfacción o mortalidad. Estos autores concluyen que el uso de proporciones estimadas en los modelos DEA convencionales puede producir resultados estadísticamente sesgados e investigan el impacto del tamaño de la muestra en las medidas de eficiencia, aunque no proponen una solución al cálculo del tamaño de muestra necesario. Además obtienen intervalos concluyendo que el método MC produce mejores intervalos que los métodos bootstrap y optimista/pesimista. Sin embargo, la confianza de que los intervalos MC

contengan al correspondiente verdadero valor es muy baja, concluyendo que la aproximación MC produce resultados sesgados.

Los problemas de programación lineal sujetos a restricciones definidas en términos de probabilidad, problemas chance-constraint (Charnes and Cooper 1959 [21] y 1961 [22], Land y otros 1993 [90], Olesen et al. 1995 [107], Cooper y otros 1996 [34], Desai y otros (2005) [44], Olesen 2006 [106]), utilizan inputs y outputs que son variables aleatorias y la caracterización determinista de “eficiente” se cambia por la caracterización probabilística “probablemente eficiente”; por supuesto, en la situación que nos planteamos sería posible utilizar esta técnica pero lo que obtendríamos sería una medida de eficiencia para cada DMU. Más referencias concernientes a soluciones matemáticas de aplicaciones DEA en presencia de datos con variaciones estocásticas, han sido estudiadas por numerosos autores incluyendo: Land y otros (1993) [90], Olesen and Petersen (1995) [107], Cooper y otros (1996 [34], 1998 [35], 2002 [37], 2004 [38]), Huang and Li (2001) [75], Bruni y otros (2009) [16], Khodabakhshi y otros (2010) [87], Khodabakhshi and Asgharian (2009) [86], Khodabakhshi (2009 [85] y 2010 [84], 2011 [83]), Wu and Lee (2010) [151] y Hoseinzadeh Lotfi y otros (2012) [74].

Ante el deseo de tomar una muestra probabilística en cada DMU, surge la necesidad de decidir el tamaño de muestra necesario para lograr una precisión, la deseada por el experimentador, en la estimación, en este caso, de la eficiencia DEA. La solución que propondremos a este problema va a dar lugar a diversos resultados de inferencia estadística en el ámbito de DEA, como por ejemplo el estudio de la consistencia de los estimadores o contrastes de hipótesis. Este tipo de resultados no es muy habitual en DEA y, entre otras referencias, encontramos: Banker 1993 [9] y 1996 [8]; Kneip y otros 1998 [88] y Gijbels y otros 1999 [65].

Tomar una muestra probabilística en cada DMU da lugar a la posibilidad de estimar los outputs con intervalos de confianza planteándose el problema de como medir la eficiencia con métodos no paramétricos de programación lineal utilizando esta información. En la sección 3.2 se introducen los métodos de obtención de intervalos de eficiencia, en presencia de datos estocásticos o datos intervalo, de Entani y otros (2002) [52], Despotis and Smirlis (2002) [46] y Wang y otros (2005) [146]. Nosotros vamos a proponer cuatro metodologías de obtención de intervalos, algunas de ellas basadas en las ideas de intervalos de eficiencia, que van a dar lugar a intervalos de confianza de eficiencia DEA. Aunque el concepto de intervalo de confianza DEA no es nuevo, Simar and Wilson (1998) [126] consideran los DMUs disponibles una muestra de una población de DMUs y utilizan la técnica bootstrap para obtener intervalos de confianza, nuestros modelos de intervalos de confianza de eficiencia DEA, uno de ellos basado en la técnica bootstrap, van a ser definidos de una manera muy diferente.

En la sección 3.3 se contempla que en la evaluación de la eficiencia en el sector público, la eficiencia técnica resulta lo más apropiado (Martínez-Cabrera 2003 [99]); por esta razón, todos los resultados de esta memoria se van a referir a la eficiencia técnica. En cuanto a la elección del modelo y a la orientación, se opta por presentar los resultados para la orientación output, es decir, manteniendo los mismos recursos evaluar si la opinión es o no mejorable. Respecto a la elección del modelo, el BCC parece el más razonable puesto que no parece lógico suponer que el aumento proporcional de recursos en un servicio implique la mejora proporcional de la opinión de sus usuarios, lo que en principio descarta el modelo CCR. Sin embargo, los resultados también se van a presentar para este modelo por dos razones, ser el modelo DEA más habitual y por la facilidad de la demostración de los resultados en este modelo.

Como punto final a esta introducción vamos a poner en conexión los dos bloques de esta tesis proponiendo un protocolo de actuación a la hora de llevar a cabo un estudio social (encuesta) en un conjunto de unidades de producción (DMUs) que prestan servicio a clientes, con el fin de recoger la opinión de los usuarios, sintetizarla, en cada DMU, en uno o más índices (output(s)) y, junto a una información conocida de las unidades (inputs) evaluar la eficiencia DEA de estas unidades.

El primer, y principal problema, es decidir el diseño muestral y el tamaño de la muestra. Respecto al diseño muestral, las condiciones del experimento, unidades que prestan servicio a clientes con un

horario, permiten utilizar el diseño polietápico propuesto en el primer bloque de la tesis en alguna de sus variantes, censo o muestreo proporcional al tiempo de atención, en las unidades de tiempo (UDTs).

Respecto al cálculo del tamaño de muestra podemos considerar dos opciones en función de que previamente al trabajo de campo podamos o no tener el tamaño de la población (exacto o estimado) de usos. Si el tamaño de la población es conocido, por ejemplo en un servicio con cita previa con todas las citas de todas las UDTs completas (odontología, pediatría,...) podemos calcular el tamaño de muestra de usos utilizando el Teorema 2, expuesto en el segundo bloque, y así garantizar un error de estimación para el índice de eficiencia poblacional, en cada DMU. La capacidad de encuestación en cada UDT determinará cuantas UDTs muestrales debemos tomar. La teoría desarrollada en el primer bloque de la tesis permitirá la estimación del output(s) en cada DMU. Si el tamaño de la población de usos es desconocido, procederemos como se expone en la sección 2.4.3, para determinar el tamaño de UDTs muestrales, y como se expone en la sección 2.4.4 para tomar la muestra de clientes.

Estimada la opinión en cada DMU, decidiremos los inputs, conocidos, para evaluar la eficiencia técnica en cada DMU utilizando DEA.

La obtención de un intervalo de confianza bootstrap para la eficiencia de cada DMU, expuesto en la sección 4.3.5, nos va a permitir contrastar si las unidades de producción son eficientes y clasificar las unidades en eficientes en cualquier caso, eficientes en sentido máximo y en unidades ineficientes, clasificación recogida en 4.36.

Parte I

Metodología muestral en encuestas de opinión de usuarios de servicios: el tiempo de atención como medida para tomar la muestra

Objetivos y motivación

El objetivo de este bloque de la tesis es proponer un diseño muestral probabilístico para realizar encuestas de opinión en la población compuesta por los usuarios de un conjunto homogéneo de servicios ubicados en uno o más espacios. Con este diseño cubrimos la necesidad de saber cómo obtener la opinión de los clientes para, mediante técnicas estadísticas, poder sintetizarla en la estimación de uno o más indicadores en cada servicio, utilizables como output(s) a la hora de medir su eficiencia.

La participación en dos estudios sociológicos, por parte del autor de esta memoria, son el detonante del planteamiento del problema de buscar un diseño muestral en una población móvil finita, pero desconocida, conformada por los individuos que acaban de hacer uso de un servicio (usuarios), utilizando el tiempo de atención al cliente. El primero de estos estudios es la “*Encuesta de satisfacción de usuarios*” realizada con objeto de evaluar la Biblioteca de la Universidad de Valladolid dentro del II Plan de Calidad de Universidades (Tapia y otros 2004 [135]). La primera fase de la encuesta se lleva a cabo en cuatro de las bibliotecas de la UVa (Tapia y Sacristán 2003 [136]) y se utiliza un diseño muestral donde las aulas de las Facultades o Escuelas se toman como conglomerados de estudiantes. Esta primera fase sirve como estudio piloto previo al estudio social completo que se realiza en el año 2004. En este caso, como se recoge en la ficha técnica del informe (Tapia y otros 2004 [135]), se encuestan, durante la segunda quincena del mes de mayo, a 3.909 usuarios distribuidos entre las catorce bibliotecas de la UVa, mediante un cuestionario completado por los propios individuos. El segundo estudio social es la “*Encuesta sobre la satisfacción de los ciudadanos con los servicios prestados por la administración de la Comunidad de Castilla y León*” (DGE 2009 [47]) en la que Rodríguez J.M. y Tapia J.A. fueron los coordinadores de la ejecución de las distintas fases del estudio. Con respecto al ámbito de aplicación, la encuesta se realiza a 3.798 usuarios, mediante preguntas realizadas por un encuestador, durante cuatro semanas, siendo la población objetivo todos los individuos que en ese mes hacen uso de alguno de los 135 servicios distribuidos entre oficinas de información y atención al ciudadano, puntos de información y atención al ciudadano, oficinas departamentales y servicios territoriales de las nueve provincias de la Comunidad de Castilla y León. Estas dos experiencias ponen de manifiesto que la idea general de usar unidades de tiempo como conglomerado de usuarios funciona como diseño probabilístico, pero, también, se ponen de manifiesto diversos problemas. El principal, lograr un control de trabajo de campo que permita evitar el riesgo de tener una elevada falta de respuesta en momentos de tiempo en los que el uso del servicio es tal que se crean colas de espera.

El interés de este capítulo se centra en las novedades del diseño propuesto en esta memoria frente a otros diseños probabilísticos llevados a la práctica en poblaciones humanas móviles (Sudman (1980) [132], Kalton (1991 [79] y 2001 [78])) como son: primero, que no precisa del conocimiento previo, exacto o estimado, del número de visitas o usos en el período de realización del trabajo de campo. Segundo, la unidad de medida para tomar la muestra es el tiempo de atención de los clientes. Por último, las estimaciones se realizan tanto para la población de todos los usos como de usuarios distintos. En cuanto al interés práctico de la metodología propuesta, destaca el equilibrio que se logra entre economía y calidad puesto que el trabajo de campo se puede llevar a cabo con entrevista personal, pero con un reducido número de encuestadores, y el experimentador puede controlar el error de muestreo.

Respecto a los aspectos tratados en este bloque de la tesis, se comenzará con un capítulo de preliminares donde se hace un breve repaso de diversos diseños muestrales y a la estimación bootstrap de la varianza de los estimadores. Puesto que el objetivo es facilitar la comprensión de los resultados de los apartados de aportaciones contenidos en el tercer capítulo, no se hará una exposición exhaustiva de estos aspectos. En el capítulo 2 se describe el procedimiento teórico proponiendo una metodología muestral que utiliza el tiempo como elemento fundamental del diseño, esta es la novedad más importante de este capítulo. El tiempo va a tener dos connotaciones, el tiempo que el servicio presta a los clientes y el tiempo que el cliente dedica a realizar el servicio. Se concluye este bloque dando una relación de las principales conclusiones.

Capítulo 1

Preliminares

En este capítulo se incluyen los diseños muestrales y las técnicas de estimación necesarias durante el desarrollo del resto del bloque. Puesto que el objetivo es facilitar la comprensión de la metodología muestral que se aporta como novedad en el capítulo 2, no se realizará una exposición exhaustiva de los distintos apartados. Comenzaremos introduciendo los diseños muestrales básicos y repasaremos los estimadores de los parámetros y de las varianzas de los estimadores asociados a estos diseños. A continuación se estudia el método bootstrap de estimación de las varianzas de los estimadores en diseños muestrales que, por su complejidad, no admiten expresiones estándar y, en particular, como se aplica la técnica bootstrap en la estimación de varianzas en poblaciones finitas y en el diseño bietápico.

1.1. Diseños Muestrales

Muestreo por Conglomerados

La implementación del diseño por conglomerados sobre una población finita precisa de una población dividida en subpoblaciones, conglomerados, que no se solapen. Los conglomerados pueden ser de igual o de distinto tamaño y deben estar constituidos por unidades lo más heterogéneas posibles, sin embargo, los conglomerados deben ser homogéneos entre ellos de tal forma que la situación ideal es cuando un único conglomerado represente fielmente la población. Se selecciona una muestra probabilística de conglomerados y se mide la variable de interés en todos los individuos de los conglomerados elegidos. Una de las principales ventajas del muestreo por conglomerados es que, para su utilización, no se precisa un marco o lista de todos los individuos sino tan sólo un marco de conglomerados, es más, ni siquiera es preciso conocer el tamaño de la población. Esta propiedad va a resultar fundamental en la justificación de la elección inicial del diseño muestral probabilístico que será propuesto en el capítulo 2. Por otro lado, el muestreo por conglomerados tiene desventajas como la pérdida en la precisión de las estimaciones por la presencia de homogeneidad dentro de los conglomerados y/o de heterogeneidad entre los conglomerados. Utilizaremos la notación estándar del diseño por conglomerados (Särndal y otros 2003 [118]) para presentar la metodología conducente a la estimación de los parámetros de interés:

N_I Número de conglomerados.

$\{N_i\}_{i=1, \dots, N_I}$ Número de individuos en cada conglomerado.

$N = \sum_{i=1}^{N_I} N_i$ Tamaño de la población.

n_I Número de conglomerados en la muestra.

$n = \sum_{i=1}^{n_I} N_i$ Número de individuos en la muestra de conglomerados.

$\{y_{ij}\}$ $j = 1, \dots, n_I$ Respuesta del individuo i del conglomerado j .
 $i = 1, \dots, N_j$

$T_j = \sum_{i=1}^{N_j} y_{ij}$ Total de las respuestas de la variable de interés de los individuos del j th conglomerado.

La expresión de los estimadores de los parámetros, en muestreo por conglomerados, está condicionada al conocimiento del tamaño de la población y a si los conglomerados son o no del mismo tamaño. Nos centramos en la situación de conglomerados de distinto tamaño por ser la que surge en el planteamiento del diseño muestral considerado como novedad en este bloque. En este caso, el estimador de la media, independientemente de disponer del tamaño poblacional, equivale a un estimador de razón (Särndal y otros 2003 [118])

$$\hat{\mu}_C = \frac{\sum_{i=1}^{n_I} \left(\frac{T_i}{\pi_i} \right)}{\sum_{i=1}^{n_I} \left(\frac{N_i}{\pi_i} \right)} \quad (1.1)$$

Si se conoce el tamaño de la población, una expresión del estimador de la varianza del estimador de la media es

$$\widehat{Var}(\hat{\mu}_C) = \frac{(1 - \frac{n}{N})}{\left(\frac{N}{N_I} \right)^2 n_I} \left(\frac{\sum_{i=1}^{n_I} (T_i - \hat{\mu}_C N_i)^2}{(n_I - 1)} \right) \quad (1.2)$$

Si el tamaño de la población N es estimado con \hat{N} , un estimador sesgado de la varianza es

$$\widehat{Var}(\hat{\mu}_C) = \frac{(1 - \frac{n}{\hat{N}})}{\left(\frac{\hat{N}}{N_I} \right)^2 n_I} \left(\frac{\sum_{i=1}^{n_I} (T_i - \hat{\mu}_C N_i)^2}{(n_I - 1)} \right) \quad (1.3)$$

Muestreo por Etapas

La forma de actuar en el diseño por conglomerados es investigar a todos los individuos de los conglomerados seleccionados. Habitualmente, esta forma de proceder es menos eficiente que muestrear directamente el mismo número de individuos de la población. Una forma habitual de proceder es utilizar un diseño en varias etapas, siendo entonces un diseño multietápico. El diseño consistente en tomar en una primera etapa conglomerados (unidades primarias) y, a continuación, seleccionar de forma aleatoria, con el diseño pertinente, un número específico de unidades de cada conglomerado seleccionado (unidades secundarias) se denomina muestreo bietápico o en dos etapas (Singh and Mangat 1996 [129]).

La notación del diseño bietápico, para presentar la metodología conducente a la estimación de los parámetros de interés, coincide, parcialmente, con la del diseño por conglomerados. Como nueva notación añadimos:

n_k Tamaño de la muestra de individuos en el conglomerado k ; $k = 1, \dots, n_I$

$n = \sum_{k=1}^{n_I} n_k$ Muestra total de individuos

$\pi_{i/k}$ Probabilidad de inclusión en la muestra del individuo i th del conglomerado k ; $k = 1, \dots, n_I$; $i = 1, \dots, n_k$

La expresión de los estimadores de los parámetros, en el diseño bietápico, como en el caso anterior, dependen del conocimiento del tamaño de la población y de si los conglomerados son o no del mismo tamaño. Si se conoce el tamaño de la población y se consideran conglomerados de aproximadamente el mismo tamaño, la expresión de los estimadores de la media y de su error son (Särndal y otros 2003 [118])

$$\hat{\mu}_1 = \frac{\sum_{k=1}^{n_I} \left(\frac{\hat{T}_k}{\pi_{Ik}} \right)}{N} \quad (1.4)$$

donde $\hat{T}_k = \sum_{j=1}^{n_k} \frac{y_{jk}}{\pi_{j/k}}$ es la estimación del total de la variable de interés en el kth conglomerado. Un estimador, sesgado, de la varianza de este estimador es

$$\widehat{AVar}(\hat{\mu}_1) = \left\{ \sum_{k,l \in s_I} \check{\Delta}_{kl} \frac{\hat{T}_k}{\pi_{Ik}} \frac{\hat{T}_l}{\pi_{Il}} + \sum_{k \in s_I} \frac{\widehat{AVar}(\hat{T}_k)}{\pi_{Ik}} \right\} \quad (1.5)$$

donde $\check{\Delta}_{kl} = \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}}$. La estimación del error en la segunda etapa, $\widehat{AVar}(\hat{T}_k)$, está determinada por el diseño con que se toma la muestra de individuos en cada conglomerado.

Si el tamaño de la población se desconoce ó se consideran conglomerados de distinto tamaño la expresión de los estimadores de la media poblacional y de la varianza de este estimador son

$$\hat{\mu}_2 = \frac{\sum_{k=1}^{n_I} \left(\frac{\hat{T}_k}{\pi_{Ik}} \right)}{\sum_{k=1}^{n_I} \left(\frac{N_k}{\pi_{Ik}} \right)} \quad (1.6)$$

$$\widehat{AVar}(\hat{\mu}_2) = \left\{ \sum_{k,l \in s_I} \check{\Delta}_{kl} \frac{(\hat{T}_k - \hat{\mu}_2 N_k)}{\pi_{Ik}} \frac{(\hat{T}_l - \hat{\mu}_2 N_l)}{\pi_{Il}} + \sum_{k \in s_I} \frac{\widehat{AVar}(\hat{T}_k)}{\pi_{Ik}} \right\} \quad (1.7)$$

Sí, además, el tamaño de los conglomerados de la muestra es desconocido, el estimador de la media poblacional es

$$\hat{\mu}_3 = \frac{\sum_{k=1}^{n_I} \left(\frac{\hat{T}_k}{\pi_{Ik}} \right)}{\sum_{k=1}^{n_I} \left(\frac{\hat{N}_k}{\pi_{Ik}} \right)} \quad (1.8)$$

donde \hat{N}_k es la estimación del número de individuos del conglomerado k . La aplicación de la técnica bootstrap (Efron and Tibshirani 1994 [50]) permite estimar la varianza de este estimador; en la sección 1.2. se introduce el bootstrap en diseño bietápico para estimar la varianza.

Estimador de Razón

Una estimación de razón aparece ante el problema de estimar el cociente entre dos totales poblacionales desconocidos

$$R = \frac{T_y}{T_x} = \frac{\sum_U y_k}{\sum_U x_k} \quad (1.9)$$

Sí, por ejemplo, T_y es el total de una variable cuantitativa Y ($T_y = \sum_U y_k$) y T_x es el total de individuos de la población, R se identifica con la media poblacional de la variable Y . Ahora, si los dos totales se estiman, respectivamente, por $\hat{T}_{y\pi} = \sum_s \frac{y_k}{\pi_k}$ y $\hat{T}_{x\pi} = \sum_s \frac{x_k}{\pi_k}$ el estimador sesgado resultante (no lineal) de R es

$$\hat{R} = \frac{\hat{T}_{y\pi}}{\hat{T}_{x\pi}} \quad (1.10)$$

Para encontrar una aproximación a la varianza y a la estimación de la varianza de \widehat{R} , se utiliza la técnica de linealización de Taylor. Las expresiones de la varianza del estimador y del estimador de esta varianza que se obtienen son

$$AVar(\widehat{R}) = \frac{1}{T_x^2} \sum \sum_U \Delta_{kl} \frac{y_k - Rx_k}{\pi_k} \frac{y_l - Rx_l}{\pi_l}; \quad \Delta_{kl} = \pi_{kl} - \pi_k \pi_l \quad (1.11)$$

$$\widehat{AVar}(\widehat{R}) = \frac{1}{\widehat{T}_{x\pi}^2} \sum \sum_U \check{\Delta}_{kl} \frac{y_k - \widehat{R}x_k}{\pi_k} \frac{y_l - \widehat{R}x_l}{\pi_l} \quad (1.12)$$

Los estimadores de razón pueden ser aplicados en diseños polietápicos (pp. 180 Särndal y otros 2003 [118]). En el caso particular bietápico de muestreo aleatorio simple sin reemplazamiento (m.a.s.) de n_l individuos de N_l , en primera etapa, y m.a.s. de n_i individuos de N_i individuos en cada uno de los n_l elementos muestrales de la primera etapa, la razón (1.9) se estima con

$$\widehat{R} = \frac{\sum_{i=1}^{n_l} N_i \bar{y}_{s_i}}{\sum_{i=1}^{n_l} N_i \bar{x}_{s_i}} \quad (1.13)$$

donde $\bar{y}_{s_i} = \frac{\sum_{j=1}^{n_i} y_j}{n_i}$ y $\bar{x}_{s_i} = \frac{\sum_{j=1}^{n_i} x_j}{n_i}$
y el estimador de la varianza de este estimador es

$$\widehat{Var}(\widehat{R}) = \frac{1}{\left(\frac{\sum_{i=1}^{n_l} N_i \bar{x}_{s_i}}{n_l}\right)^2} \left[\left(\frac{1}{n_l} - \frac{1}{N_l}\right) \frac{\sum_{i=1}^{n_l} N_i \bar{e}_{s_i}^2}{n_l - 1} + \frac{1}{N_l n_l} \sum_{i=1}^{n_l} N_i^2 \left(\frac{1}{n_i} - \frac{1}{N_i}\right) \frac{\sum_{j=1}^{n_i} N_j (e_j - \bar{e}_{s_i})^2}{n_i - 1} \right] \quad (1.14)$$

$$\text{con } e_i = y_i - \widehat{R}x_i \text{ y } \bar{e}_{s_i} = \bar{y}_{s_i} - \widehat{R}\bar{x}_{s_i}$$

El método para estimar una razón conduce a uno de los estimadores indirectos del total poblacional más utilizados, el estimador de razón para el total, el cual requiere el conocimiento de la información del total de la variable auxiliar, $T_x = \sum_{k \in U} x_k$. La expresión del estimador de razón del total de la variable de interés, T_y , es

$$\widehat{T}_{yr\pi} = \widehat{R}T_x \quad (1.15)$$

donde \widehat{R} es (1.10) y, por tanto, el estimador de la varianza del estimador de razón es

$$\widehat{Var}(\widehat{T}_{yr\pi}) = \frac{T_x^2}{\widehat{T}_{x\pi}^2} \sum \sum_s \check{\Delta}_{kl} \frac{y_k - \widehat{R}x_k}{\pi_k} \frac{y_l - \widehat{R}x_l}{\pi_l} \quad (1.16)$$

La estimación de una razón también va a permitir estimar la media poblacional. Si el tamaño de la población, N , es conocido y se dispone de la información auxiliar, T_x , el estimador de razón de la media es

$$\widehat{\mu}_{yr\pi} = \frac{\widehat{T}_{yr\pi}}{N} \quad (1.17)$$

Si el tamaño de la población es desconocido, un estimador de razón de la media es

$$\widehat{\mu}_{ys} = \frac{\widehat{T}_{yr\pi}}{\widehat{N}} = \frac{\sum_s \frac{y_k}{\pi_k}}{\sum_s \frac{1}{\pi_k}} \quad (1.18)$$

y un estimador de la varianza de este estimador es

$$\widehat{Var}(\widehat{\mu}_{ys}) = \frac{1}{\widehat{N}^2} \sum \sum_s \check{\Delta}_{kl} \left(\frac{y_k - \widehat{\mu}_{ys}}{\pi_k} \right) \left(\frac{y_l - \widehat{\mu}_{ys}}{\pi_l} \right) \quad (1.19)$$

Muestreo Sistemático Proporcional al Tamaño

Como paso previo a introducir el muestreo sistemático proporcional al tamaño, es preciso realizar un breve repaso tanto del diseño sistemático como del diseño proporcional al tamaño.

Tomar una muestra con diseño sistemático tiene la ventaja de la gran rapidez y facilidad de la selección de individuos, asignando la misma probabilidad de inclusión a todos los individuos. Para lograr que el tamaño de la muestra sistemática de una población de tamaño N sea de un tamaño exacto, n , se dispone del método del intervalo fraccional: extraemos un valor ξ de forma aleatoria de una distribución uniforme en $(0, k)$, donde $k = \frac{N}{n}$, y la muestra queda conformada por los individuos

$$s = \{l : l - 1 < \xi + (j - 1)k \leq l; j = 1, \dots, n\}$$

Uno de los problemas teóricos del diseño sistemático surge a la hora de estimar la varianza del estimador, procediendo, habitualmente, a considerar equivalentes el diseño sistemático y el muestreo aleatorio simple; esta forma de proceder sólo es correcta si la disposición de los individuos de la población es aleatoria, es decir, el coeficiente de correlación intraclases (pp. 79 cap 3 Särndal y otros (2003) [118]) es nulo o negativo. La discusión completa de las distintas alternativas de estimadores para la varianza en este diseño la podemos encontrar en el capítulo 8 del texto de Wolter (2007) [150] donde se muestran ocho opciones de estimadores de la varianza.

Si disponemos de una variable auxiliar, aproximadamente proporcional a la variable de interés, observada a priori en todos los individuos de la población e identificable como una medida de tamaño, el diseño proporcional al tamaño (pps si hay reemplazamiento y πps si no hay reemplazamiento) permite tomar una muestra asignando mayor probabilidad a los individuos con mayor tamaño; en Särndal y otros (2003) [118] se expone la dificultad de implementar el diseño πps cuando se desea tomar una muestra de más de dos elementos.

El muestreo sistemático proporcional al tamaño es un diseño sistemático con probabilidades de selección de individuos desiguales, considerado tipo πps . Supongamos $\{x_k\}_{k=1, \dots, N}$ la respuesta de la variable que mide el tamaño en toda la población. Consideremos $T_0 = 0$ y $T_k = x_k + T_{(k-1)}$; $k = 1, \dots, N$. Fijamos el tamaño de muestra n y sea a la parte entera de $\frac{T_N}{n}$, entonces

$$T_N = na + c; \quad 0 \leq c \leq a$$

Si $c > 0$ la muestra puede ser de tamaño n ó $n + 1$. Supongamos que $nx_k \leq T_N - c$, que corresponde a la suposición $nx_k \leq \sum_U x_k$. El muestreo sistemático πps se define de la siguiente manera:

- Seleccionamos, con la misma probabilidad, $1/a$, un valor entero, digamos r , entre el 1 y a .
- La muestra seleccionada está conformada por

$$s = \{k : T_{k-1} < r + (j - 1)a \leq T_k \forall j = 1, \dots, n_s\}; \quad n_s = n \text{ ó } n + 1$$

A partir de esta definición se deduce que la probabilidad de inclusión de un individuo en la muestra es

$$\pi_k = \frac{nx_k}{T_N - c}$$

y, si $\{y_k\}_{k \in s}$ son las respuestas de los individuos de la muestra, el estimador de la media poblacional es

$$\hat{\mu} = \frac{\sum_{k \in s} y_k \pi_k}{N} \quad (1.20)$$

El problema de la estimación de la varianza del estimador en muestreo sistemático proporcional al tamaño admite una amplia discusión (Wolter 2007 [150]) y, al igual que en el diseño sistemático con

probabilidades iguales, se dispone de diversos estimadores alternativos más la posibilidad de aplicar la técnica bootstrap. Los estimadores de la varianza de uso más habitual son

$$v_1 = \frac{1}{N^2} \sum_{i=1}^n \sum_{i < k}^n \left(1 - \pi_i - \pi_k + \sum_{j=1}^N \frac{\pi_k^2}{n} \right) \left(\frac{y_i}{\pi_i} - \frac{y_k}{\pi_k} \right)^2 \quad (1.21)$$

$$v_2 = \frac{1}{n(n-1)} \sum_{i=1}^n \left(\frac{y_i}{p_i} - N\hat{\mu} \right)^2; p_i = \frac{x_i}{T_N} \quad (1.22)$$

$$v_3 = \frac{\frac{1}{n} \sum_{i=2}^n \left(\frac{y_i}{p_i} - \frac{y_{i-1}}{p_{i-1}} \right)^2}{2(n-1)} \quad (1.23)$$

Los estimadores v_1 y v_2 se recomiendan en poblaciones ordenadas, aproximadamente, de manera aleatoria, mientras que se propone utilizar v_3 en poblaciones ordenadas según alguna tendencia.

Postestratificación

La existencia de una subdivisión de la población U en subconjuntos disjuntos, ($U = U_1 \cup \dots \cup U_G$), de acuerdo a un principio de clasificación dado, da lugar a un modelo denominado “*the group mean model*” (Särndal y otros 2003 [118]). Los subconjuntos se denominan grupos de población, o simplemente grupos, y no son estratos, excepto en el caso especial de que sean usados para tomar una muestra estratificada. La expresión del estimador del total en este modelo es

$$\hat{T}_{yr} = \sum_{g=1}^G N_g \tilde{y}_{s_g} \quad (1.24)$$

donde

$$\tilde{y}_{s_g} = \sum_{g=1}^G \frac{\check{y}_{s_g}}{\hat{N}_g}$$

$$\hat{N}_g = \sum_{k \in s_g} \frac{1}{\pi_k}$$

$$\check{y}_{s_g} = \sum_{k \in s_g} \frac{y_k}{\pi_k}$$

N es el tamaño de la población U

N_g es el tamaño del subconjunto U_g ; $g = 1, \dots, G$, tal que $N = \sum_{g=1}^G N_g$

s es la muestra de tamaño n tomada en la población U

s_g es la muestra de tamaño n_g de U_g ; $g = 1, \dots, G$, tal que $n = \sum_{g=1}^G n_g$

π_k es la probabilidad de inclusión del individuo en la muestra s .

Una situación donde es aplicable “*the group mean model*” surge cuando la pertenencia a los posibles grupos viene dada en el marco, para todos los elementos de la población, pero nos abstenemos de usar los grupos como estratos. En este caso la información de la pertenencia de un individuo a un grupo se utiliza en la etapa de estimación y no en la etapa de diseño y representa un caso de postestratificación; nos centramos en este caso por ser el que se nos presenta en la propuesta de metodología realizada en el capítulo 2. Si la muestra s de los individuos se realiza con muestreo aleatorio simple sin reemplazamiento, la expresión del estimador del total, en el modelo postestratificado, es:

$$\hat{T}_{ypos} = \sum_{g=1}^G N_g \bar{y}_{s_g} \quad (1.25)$$

donde $\bar{y}_{s_g} = \sum_{k=1}^{n_g} \frac{y_k}{n_g}$ y n_g es el número de individuos de la muestra aleatorio simple sin reemplazamiento s de n individuos que pertenecen al grupo g ; $g = 1, \dots, G$.

El estimador de la varianza del estimador es

$$\widehat{Var}(\widehat{T}_{ypos}) = (1-f) \sum_{g=1}^G N_g^2 \frac{S_{s_g}^2}{n_g} \quad (1.26)$$

$$\text{donde } S_{s_g}^2 = \frac{1}{n_g-1} \sum_{k=1}^{n_g} (y_k - \bar{y}_{s_g})^2.$$

1.2. Estimación Bootstrap de la varianza de los estimadores

Un requisito fundamental en la correcta práctica de aplicación de diseños muestrales es disponer de una medida de precisión de los estimadores asociados al diseño. La medida de precisión más comúnmente utilizada es la varianza del estimador. En general, la varianza no se conoce pero puede ser estimada con los datos del experimento. Si los diseños son complejos, no es viable utilizar los estimadores habituales de la varianza, como el estimador de la varianza del pi-estimador, y se precisa de técnicas alternativas de estimación. Los diseños muestrales complejos aparecen cuando se combinan la estratificación, el muestreo multietápico, la selección de individuos con probabilidades desiguales, muestreo polifásico, postestratificación, utilización de estimadores de razón o regresión, desconocimiento del tamaño de la población y/o riesgo de una elevada falta de respuesta. Varios de estos aspectos se van a combinar en el diseño muestral probabilístico propuesto en el capítulo 2 y, para estimar la varianza, vamos a proponer utilizar la técnica bootstrap (Efron 1979 [49]).

Método Bootstrap

El método bootstrap ha sido objeto de multitud de investigaciones en el último cuarto de siglo. Originalmente, este método se describe considerando Y_1, \dots, Y_n una muestra de variables aleatorias i.i.d. con función de distribución F , θ un parámetro que se desea estimar por $\hat{\theta}$, un estimador, función de la muestra observada. Una muestra bootstrap, Y_1^*, \dots, Y_n^* , es una muestra, del mismo tamaño que la muestra original, tomada con reemplazamiento de la muestra original. El estimador bootstrap $\hat{\theta}^*$ es un estimador con la misma forma funcional que $\hat{\theta}$ pero aplicado a la muestra bootstrap. El estimador bootstrap ideal de la $Var(\hat{\theta})$ se define como $v_1(\hat{\theta}) = Var_*(\hat{\theta}^*)$ siendo Var_* la varianza condicional, dada la muestra original. Para estimar la varianza del estimador se utiliza la varianza de las estimaciones bootstrap obtenidas de un número B de remuestras.

En problemas simples donde $\hat{\theta}$ es lineal, los pasos del procedimiento de estimación bootstrap de la varianza del estimador son:

1. Extraer un número elevado, B , de replicas bootstrap, independientes, de la muestra original.
2. Para cada replica evaluar la correspondiente estimación bootstrap del parámetro de interés $\hat{\theta}_i^*$; $i = 1, \dots, B$
3. Calcular la varianza de las estimaciones $\hat{\theta}_i^*$; $i = 1, \dots, B$:

$$v_1(\hat{\theta}) = \frac{1}{B-1} \sum_{i=1}^B (\hat{\theta}_i^* - \widehat{\bar{\theta}}^*)^2 \quad (1.27)$$

donde $\widehat{\bar{\theta}}^* = \frac{\sum_{i=1}^B \hat{\theta}_i^*}{B}$ y se verifica

$$v_1(\hat{\theta}) \xrightarrow{B \rightarrow \infty} v_1(\hat{\theta})$$

Bootstrap en poblaciones finitas

En poblaciones finitas, el uso del método bootstrap precisa de nuevas consideraciones. Atendiendo a las necesidades de estimación de la metodología muestral propuesta en el capítulo 2, estudiamos la adaptación del bootstrap original a poblaciones finitas en el muestreo aleatorio simple sin reemplazamiento y en el muestreo proporcional al tamaño sin reemplazamiento (Wolter 2007 [150]).

BOOTSTRAP EN MUESTREO ALEATORIO SIMPLE SIN REEMPLAZAMIENTO

Supongamos que el parámetro de interés es la media poblacional μ , estimada con la media muestral \bar{Y} . Consideremos s la muestra aleatoria simple sin reemplazamiento original de tamaño n y s^* la muestra bootstrap de tamaño n^* . A la hora de obtener un estimador insesgado bootstrap de la varianza del estimador, disponemos de diversas opciones dependiendo de si la razón entre tamaño de muestra y población ($f = \frac{n}{N}$) es o no próxima a 0:

- Si $f \approx 0$ el tamaño de las remuestras bootstrap tiene que ser una unidad menos que el tamaño de la muestra original ($n^* = n - 1$) para que el estimador bootstrap de la varianza sea insesgado.
- Si $f \neq 0$ se dispone de diversas opciones para obtener estimadores bootstrap insesgados de la varianza como son (Wolter 2007 [150]):

Correction Factor: El tamaño de las remuestras bootstrap se toma una unidad menos que el tamaño de la muestra original ($n^* = n - 1$) siendo el estimador insesgado de la varianza

$$v_{1F}(\bar{Y}) = (1 - f) v_1(\bar{Y}) \quad (1.28)$$

Rescaling: Si $n^* = n - 1$, el estimador bootstrap de la varianza se define en términos de las observaciones rescaladas:

$$y_i^\# = \bar{y} + (1 - f)^{1/2} (y_i^* - \bar{y}) \quad (1.29)$$

BWR: Este método bootstrap con reemplazamiento trata de eliminar el sesgo del estimador de la varianza tomando como tamaño de la muestra bootstrap

$$n^* = \frac{(n - 1)}{(1 - f)} \quad (1.30)$$

BWO: En este método la muestra bootstrap se toma sin reemplazamiento. El procedimiento tiene los siguientes pasos:

- a) Tomamos $k = N/n$ y replicamos cada elemento de la muestra original k veces con lo que creamos una nueva pseudopoblación U_s de tamaño N . Denotamos sus valores unitarios como $\{y'_j\}_{j=1}^N$. Exactamente k de los valores y'_j son iguales a y_i para $i = 1, \dots, n$.
- b) Extraemos la muestra bootstrap s^* tomando una muestra sin reemplazamiento en U_s de tamaño n^* .
- c) Evaluamos la media bootstrap

$$\hat{y}^* = \frac{1}{n^*} \sum_{i=1}^{n^*} y_i^* \quad (1.31)$$

- d) Repetimos los apartados a)-c) un número B elevado de veces y computamos la versión de Monte Carlo del estimador teórico bootstrap de la varianza del estimador

$$v_{2BWO}(\bar{Y}) = \frac{1}{B-1} \sum_{i=1}^B \left(\bar{y}_i^* - \frac{1}{B} \sum_{j=1}^B \bar{y}_j^* \right)^2 \quad (1.32)$$

Si $n^* = n$ el estimador de la varianza se debe multiplicar por el factor $C = \frac{(N-1)n}{N(n-1)}$.

Mirror-Match: Este procedimiento también tiene cuatro pasos:

- a) Seleccionamos una submuestra (o un grupo aleatorio) de tamaño entero m ($1 \leq m \leq n$) de la muestra original s con muestreo aleatorio simple sin reemplazamiento
b) Repetimos el paso a) k veces, de manera independiente,

$$k = \frac{n(1-e)}{m(1-f)}; \text{ donde } e = 1 - \frac{m}{n} \quad (1.33)$$

La muestra bootstrap está conformada por los grupos aleatorios seleccionados y es de tamaño $n^* = mk$.

- c) Evaluamos la media bootstrap

$$\hat{y}^* = \frac{1}{n^*} \sum_{i=1}^{n^*} y_i^* \quad (1.34)$$

- d) Repetimos los apartados a)-c) un número B elevado de veces y computamos la versión de Monte Carlo del estimador teórico bootstrap de la varianza del estimador

$$v_{2MM}(\bar{Y}) = \frac{1}{B-1} \sum_{i=1}^B \left(\bar{y}_i^* - \frac{1}{B} \sum_{j=1}^B \bar{y}_j^* \right)^2 \quad (1.35)$$

BOOTSTRAP EN MUESTREO PROPORCIONAL AL TAMAÑO SIN REEMPLAZAMIENTO (πps)

Supongamos $\{x_k\}_{k=1, \dots, N}$ la respuesta de una variable auxiliar que mide el tamaño en los individuos de la población. En el diseño πps , la probabilidad de inclusión de primer orden de cada individuo, fijado un tamaño de muestra n , es

$$\pi_i = \frac{nx_i}{T_x}; \quad T_x = \sum_{i=1}^n x_i$$

y el estimador de Horwitz-Thompson para el total poblacional es

$$\hat{T}_y = \sum_{i=1}^n \frac{y_i}{\pi_i} = \sum_{i=1}^n y_i w_i = \frac{1}{n} \sum_{i=1}^n u_i \text{ donde } u_i = ny_i w_i$$

Para aplicar bootstrap se considera la muestra $s = \{u_1, \dots, u_n\}$ y se toma una muestra aleatoria simple con reemplazamiento de tamaño $n^* = n - 1$ obteniéndose la muestra bootstrap $s^* = \{u_1^*, \dots, u_{n^*}^*\}$. La estimación bootstrap del total es

$$\hat{T}_y^* = \frac{1}{n^*} \sum_{i=1}^{n^*} u_i^*$$

Si se toman B remuestras, computamos la versión de Monte Carlo del estimador teórico bootstrap de la varianza del estimador

$$v_2(\hat{T}_y) = \frac{1}{B-1} \sum_{i=1}^B \left(\hat{T}_{yi}^* - \frac{1}{B} \sum_{j=1}^B \hat{T}_{yj}^* \right)^2 \quad (1.36)$$

y se corrige el sesgo de este estimador introduciendo un factor de corrección

$$v_{2F}(\hat{T}_y) = (1 - \bar{f}) v_2(\hat{T}_y) \text{ con } \bar{f} = \frac{1}{n} \sum_{i=1}^n \pi_i \quad (1.37)$$

Estimación bootstrap de la varianza en diseño bietápico sin reemplazamiento

Rao and Wu (1988) [114] estudian la metodología de remuestreo bootstrap en diseños complejos sin reemplazamiento y, en particular, en el diseño bietápico con probabilidades iguales y sin reemplazamiento. El método es el siguiente:

- Suponemos una población compuesta por N_I conglomerados con N_i individuos en el i th conglomerado ($i = 1, \dots, N_I$); el tamaño de la población es $N = \sum_{i=1}^{N_I} N_i$.
- Tomamos una muestra sin reemplazamiento de n_I conglomerados y, en cada conglomerado muestral, tomamos una muestra sin reemplazamiento de individuos en los que se mide la variable de interés; por tanto, se dispone de la información muestral $\{y_{ij}\}_{\substack{i=1, \dots, n_I \\ j=1, \dots, n_i}}$.
- Los estimadores insesgados del total y la media poblacional son

$$\hat{T} = \frac{N_I}{n_I} \sum_{i=1}^{n_I} N_i \bar{y}_i \quad (1.38)$$

si N es conocido

$$\hat{\mu} = \frac{\hat{T}}{N} \quad (1.39)$$

si N es desconocido

$$\hat{\mu} = \frac{\hat{T}}{\hat{N}} \quad \text{con} \quad \hat{N} = \frac{N_I}{n_I} \sum_{i=1}^{n_I} N_i \quad (1.40)$$

- La obtención de las remuestras bootstrap y_{ij}^{**} a partir de y_{ij} se realiza de la siguiente manera:
 - a) Tomamos una muestra aleatoria simple con reemplazamiento de tamaño n_I de la muestra de n_I conglomerados y tomamos una muestra aleatoria simple con reemplazamiento de tamaño n_i entre los n_i elementos del conglomerado i th de la remuestra (remuestreo independiente del mismo conglomerado si es escogido más de una vez).

Usamos la siguiente notación:

y_{ij}^{**} valor de la variable de interés del elemento j th de la remuestra en el i th conglomerado de la remuestra.

n_i^* valor n_i del i th conglomerado de la remuestra (similarmente N_i^*).

$$\hat{T}_i^* = N_i^* \left(\frac{\sum_{j=1}^{n_i^*} y_{ij}^{**}}{n_i^*} \right)$$

Calculamos

$$\tilde{y}_{ij} = \hat{\mu} + \lambda_{1i} \left(\frac{\hat{T}_i^*}{\left(\frac{N}{N_I}\right)} - \hat{\mu} \right) + \lambda_{2i}^* \left(\frac{N_i^* y_{ij}^{**}}{\left(\frac{N}{N_I}\right)} - \frac{\hat{T}_i^*}{\left(\frac{N}{N_I}\right)} \right) \quad (1.41)$$

y

$$\tilde{\mu} = \frac{1}{n_I} \sum_{i=1}^{n_I} \frac{1}{n_i^*} \sum_{j=1}^{n_i^*} \tilde{y}_{ij} \quad (1.42)$$

donde $\lambda_{1i}^2 = \frac{n_I}{n_I-1} \left(1 - \frac{n_I}{N_I}\right)$ y $\lambda_{2i}^{*2} = \frac{n_I}{N_I} \left(1 - \frac{n_i}{N_i}\right) \frac{n_i^*}{n_i^*-1}$

b) Repetimos B veces el apartado a) y obtenemos $\tilde{\mu}^1, \dots, \tilde{\mu}^B$; el estimador bootstrap de la varianza del estimador $\hat{\mu}$ es

$$v(\hat{\mu}) = \frac{1}{B-1} \sum_{i=1}^B (\tilde{\mu}^i - \bar{\tilde{\mu}})^2 \quad (1.43)$$

Field y Welsh (2007) [57] introducen la denominada “estrategia 2: bootstrap dos-etapas” en la estimación de la varianza en diseño bietápico, consistente en:

- Se toma una muestra bootstrap, $s_I^* = \{u_{b_1}^*, \dots, u_{b_{n_I}}^*\}$, muestra con reemplazamiento del mismo tamaño que la muestra original de unidades de primera etapa $s_I = \{u_1, \dots, u_{n_I}\}$.
- Para cada unidad de la muestra s_I^* disponemos de la muestra original de unidades de segunda etapa $s_{b_i} \in s = \{s_1, \dots, s_{n_I}\}$ donde s_i es la muestra de individuos de la unidad muestral de primera etapa; en cada una de ellas tomamos una muestra con reemplazamiento del mismo tamaño, muestra bootstrap, y obtenemos $s^* = \{s_{b_1}^*, \dots, s_{b_{n_I}}^*\}$.

El estimador de un parámetro, en diseño bietápico, es función de las muestras s_I y s , $\hat{\theta} = f(s_I, s)$; con las muestras s^* y s_I^* estimamos $\hat{\theta}^* = f(s_I^*, s^*)$ y, repetido el proceso de remuestreo B veces, el estimador bootstrap de la varianza del estimador $\hat{\theta}$ es

$$v(\hat{\theta}) = \frac{1}{B-1} \sum_{i=1}^B (\hat{\theta}_i^* - \bar{\hat{\theta}}^*)^2 \quad (1.44)$$

donde $\bar{\hat{\theta}}^* = \frac{\sum_{i=1}^B \hat{\theta}_i^*}{B}$

Capítulo 2

Metodología muestral en unidades productivas que prestan servicios a clientes

La estructura de este capítulo es la siguiente: se introduce la notación, a continuación planteamos el problema y analizamos diversos aspectos de la encuesta y de la información auxiliar que se debe tomar paralelamente a la realización de la encuesta. Seguidamente introducimos diseños para tomar la muestra de espacios, de servicios, de unidades de tiempo y de clientes de los servicios y analizamos el coste y el efecto del diseño y comprobamos, mediante simulaciones, la calidad de los diseños muestrales propuestos. A continuación mostramos tres estudios de aplicación de la metodología censo en unidades de tiempo, dos llevados ya a cabo y una nueva propuesta. En la octava sección proponemos una alternativa a la selección de la muestra de unidades de tiempo si el tamaño de la población de usos puede ser conocido o estimado a priori. Por último analizamos el efecto en la estimación de los parámetros de usuarios en una población donde los clientes repiten, con bastante frecuencia, el uso del servicio a lo largo del tiempo.

2.1. Notación

Utilizaremos la notación estándar del diseño por conglomerados introducida en la sección 1.1, adaptándola para poder realizar inferencias tanto en la población de usos como de usuarios en un servicio:

UDT Unidad de tiempo

N_I Número total de UDTs del marco

N_i Número de usos en la i th UDT; $i = 1, \dots, N_I$

$N = \sum_{i=1}^{N_I} N_i$ Tamaño de la población de usos

n_I Número de UDTs muestrales y en las que al menos un usuario contesta

n_i Tamaño de muestra de usos en la i th UDT

$n = \sum_{i=1}^{n_I} n_i$ Tamaño de muestra de usos en las UDTs muestrales

$T_i = \sum_{j=1}^{N_i} y_{ij}$ Respuesta total a la variable de interés de los usos de la i th UDT

$\mu = \frac{\sum_{i=1}^{N_I} T_i}{N}$ Respuesta media de la población de usos a la variable de interés

$I_j = \begin{cases} 1 & \text{si individuo } j \text{ usuario nuevo} \\ 0 & \text{si individuo } j \text{ usuario ya encuestado} \end{cases}; j = 1, \dots, N$

$N_{iu} = \sum_{j=1}^{N_i} I_j$ Número de usuarios, usos no encuestados previamente, en la i th UDT; $i = 1, \dots, N_I$

$N_u = \sum_{i=1}^{N_I} N_{iu}$ Tamaño de la población de usuarios, usos encuestados sólo una vez

$N_{i\bar{u}} = N_i - N_{iu}$ Número de no usuarios, usos ya encuestados, en la i th UDT; $i = 1, \dots, N_I$

$N_{\bar{u}} = \sum_{i=1}^{N_I} N_{i\bar{u}}$ Tamaño de la población de no usuarios.

$T_{iu} = \sum_{j=1}^{N_i} y_{ij} I_j$ Respuesta total a la variable de interés de los usuarios de la i th UDT

$\mu_u = \frac{\sum_{i=1}^{N_I} T_{iu}}{N_u}$ Respuesta media de la población de usuarios a la variable de interés

n_{iu} Tamaño de muestra de usuarios en la i th UDT

$n_u = \sum_{i=1}^{N_I} n_{iu}$ Tamaño de muestra de usuarios del total de UDTs muestrales

$n_{i\bar{u}} = n_i - n_{iu}$ Tamaño de muestra de usuarios ya encuestados en la i th UDT

$n_{\bar{u}} = \sum_{i=1}^{N_I} n_{i\bar{u}}$ Tamaño de muestra de usuarios ya encuestados

π_{Ik} Probabilidad de selección de la k th UDT

π_{Ikl} Probabilidad de selección simultanea de la k th y la l th UDT

$\check{\Delta}_{Ikl} = \frac{\pi_{Ikl} - \pi_{Ik}\pi_{Il}}{\pi_{Ikl}}$

2.2. Planteamiento del Problema

Nuestro objetivo es estimar un parámetro (media, total) a partir de la opinión, recogida con una encuesta cara-a-cara, de los usuarios de un cierto conjunto de servicios ubicados en un mismo espacio. Esto exige, en primer lugar, determinar el ámbito de aplicación, la población objetivo y el marco. Definidos estos aspectos, propondremos el diseño muestral con objeto de tomar una muestra de unidades de tiempo y una muestra de individuos. Recogida la información de la muestra de individuos, estimaremos los parámetros de interés.

ÁMBITO DE APLICACIÓN

Los servicios aptos a la aplicación de la metodología propuesta en este capítulo precisan de clientes y de un horario diario perfectamente especificado. En el espacio donde se ubica el servicio, el encuestador debe tener la capacidad de poder controlar visualmente a los usuarios de cara a contabilizar tanto el tiempo de atención como el número de individuos que lo utilizan. La metodología que se va a proponer considera la posibilidad de que en un mismo espacio (edificio o centro) se ofrezca más de un servicio perfectamente identificable por ser prestado, cada servicio, en una ubicación física distinta pero dentro del mismo espacio. En cada servicio se supone una única cola de clientes y un único servidor (persona que atiende al usuario). Algunos ejemplos válidos son las gestiones, trámites o procesos de información realizados en centros de la administración pública o en una sucursal bancaria, las consultas médicas con cita previa y las cajas de pago de los centros comerciales. En concreto, en una sucursal bancaria hay distintos lugares (mostrador, mesa,..) para los distintos servicios.

POBLACIÓN

Para cada servicio se distinguen dos poblaciones, la primera, conformada por todos los usos de ese servicio en el tiempo determinado con el marco. La segunda población son los usuarios distintos de un mismo servicio, es decir, un individuo sólo está una vez aunque utilice el servicio más veces en

el marco tiempo. Como notación, a la unidad estadística, o individuo, de la primera población se la denomina uso y a la de la segunda usuario.

MARCO

El marco espacio-tiempo determina la localización geográfica y temporal de la población de referencia. Un espacio o centro es una ubicación física donde se prestan uno o más servicios y que tiene un conjunto de usos, en principio desconocido en número, desde el comienzo del trabajo de campo hasta su finalización (días, semanas, mes, meses, año, ...).

En cuanto al tiempo, el marco recoge dos aspectos, la duración del trabajo de campo y el tamaño de la unidad de tiempo muestral. En la mayoría de los casos, las condiciones impuestas, caso del presupuesto, determinan que se tenga que realizar el trabajo de campo en un período más o menos corto. Sin embargo, es frecuente detectar y prever ciertas periodicidades de comportamiento con respecto al uso de los servicios, por ejemplo, el número de usuarios de los servicios de la agencia tributaria es mucho más elevado en el período de declaración de la renta. En estos casos, el tiempo de encuestación, si es posible, debe ser lo suficientemente amplio como para evitar los sesgos originados por comportamientos periódicos de uso.

La decisión sobre la elección del tamaño de la UDT muestral está condicionada a varios factores: encuesta autoaplicada o con entrevista personal, duración de cada encuesta, como se distribuyen los clientes en el tiempo, el horario de atención diario, capacidad de encuestación, consideración de costes y periodicidades temporales tanto de uso del espacio como de tipo de usuario. Por ejemplo, períodos de UDT de encuestación de una hora durante un mes de trabajo de campo parecen razonables pero, obviamente, existen más opciones.

La Figura 2.1 muestra un ejemplo de marco de dos semanas correspondiente a un espacio con cuatro servicios y horario de apertura de lunes a viernes; en este caso, la UDT muestral es la hora.

| Marco | Servicio 1 | Servicio 2 | Servicio 3 | Servicio 4 |
|-----------|---------------|------------|------------|------------|
| Lunes | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Martes | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Miércoles | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Jueves | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Viernes | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| Lunes | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Martes | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Miércoles | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Jueves | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |
| | 16:30 - 17:30 | | | |
| | 17:30 - 18:30 | | | |
| Viernes | 9:00 - 10:00 | | | |
| | 10:00 - 11:00 | | | |
| | 11:00 - 12:00 | | | |
| | 12:00 - 13:00 | | | |
| | 13:00 - 14:00 | | | |

Figura 2.1: Marco de dos semanas con unidad de tiempo la hora, en un espacio con cuatro servicios

2.3. Encuesta e información auxiliar

A la hora de proponer un diseño muestral en el ámbito de aplicación expuesto en la sección anterior, la información disponible determina distintas alternativas de diseño muestral. En esta sección introducimos la información considerada imprescindible y aquella otra información cuyo conocimiento va a dar lugar a mejoras en la calidad de las estimaciones de los parámetros de interés.

Respecto a la encuesta, esta debe recoger la percepción del individuo sobre aspectos del servicio del que acaba de hacer uso, caso de las encuestas de calidad de los servicios tipo SERVQUAL (Parasuraman y otros 1988 [108]) que proporcionan un índice de satisfacción de cada cliente y, la media de estos índices en la población de clientes, es un indicador de satisfacción (estimador) del servicio encuestado. La síntesis de las respuestas de cada individuo en un índice, por ejemplo de satisfacción, justifica el interés en centrarnos en la estimación del parámetro respuesta media poblacional de los clientes.

Para realizar inferencias sobre la población de usos, resulta imprescindible conocer el total de usos de cada UDT muestral, $N_i; i = 1, \dots, n_I$. Para tomar esta medida, en el caso práctico de DGE (2009) [47], el encuestador dispone de un contador manual que utiliza simultáneamente a la realización de la encuesta; en Kalton (1991) [79], se propone disponer de dos encuestadores, uno que realice las encuestas y otro que cuantifique el número de usuarios y obtenga información que permita caracterizar el perfil de los mismos.

La estimación del tamaño de la población de usuarios, y de la respuesta media poblacional, precisa el conocimiento de la clasificación de toda la muestra de individuos de cada UDT muestral en usuarios, individuos aún no encuestados, y no usuarios, individuos encuestados previamente en ese servicio

$$\{n_i = n_{iu} + n_{i\bar{u}}\}_{i=1, \dots, n_I} \quad (2.1)$$

Para lograrlo, la primera pregunta a realizar a toda la muestra de usos es si ya ha sido encuestado.

El conocimiento del total de usos en todo el período de encuestación, N , finalizado el trabajo de campo, disponible en algunos de los ámbitos de aplicación en función de registros documentales, no es un dato imprescindible pero mejora, como mostrarán las simulaciones, la estimación del tamaño de la población de usuarios.

Una información auxiliar, metodológicamente importante, es la clasificación inicial del marco de UDTs en saturadas y no saturadas, entendiéndose por UDT saturada aquella en la que el servicio que se presta es continuo o intensivo, sin falta de clientes en ningún momento. Disponer de esta información, en la realidad, es bastante factible, algunas veces por la característica del servicio, citas previas, y otras por la experiencia de los que atienden el servicio.

2.4. Diseño muestral Tiempo-Espacio

Adecuar la metodología muestral al objetivo que se pretende lograr precisa, una vez fijado el marco, considerar un diseño polietápico, donde las unidades de primera etapa son los espacios o centros, las unidades de segunda etapa son los servicios, las unidades de tercera etapa son las unidades de tiempo y las unidades de última etapa son los individuos usuarios de los servicios.

En el resto de esta sección se proponen opciones para tomar la muestra de las unidades de las distintas etapas, siendo el principal objetivo, por su complejidad y lo novedoso, las propuestas de selección de las UDTs y de los clientes.

2.4.1. Selección de Espacios

La población de espacios está conformada por el conjunto de espacios o centros donde se ubican los servicios en los que se desea conocer la opinión de sus usuarios. El censo de la población de espacios es, siempre que los recursos los permitan, la mejor opción. La alternativa habitual es estratificar los espacios en grupos homogéneos y tomar una muestra aleatoria simple de espacios en cada estrato. La estimación de los parámetros de interés en cada espacio se realiza de manera independiente y las características del diseño estratificado permiten extrapolar la estimación en los espacios muestrales a la población de espacios. La extrapolación de las estimaciones de los espacios muestrales a la población de espacios es un problema clásico que no consideramos de interés abordar

en esta tesis, por tanto, nos centramos en el problema de estimación puntual y por intervalo de confianza de los parámetros en un espacio.

2.4.2. Selección de Servicios

La metodología de selección de las UDTs muestrales en cada servicio, que será propuesta en la sección 2.4.3, va a permitir poder realizar un censo de servicios en cada espacio seleccionado en primera etapa.

2.4.3. Selección de Unidades de Tiempo

La forma habitual de tomar la muestra de UDTs es con diseño proporcional a su tamaño, o a su tamaño estimado de usos (Sudman 1980 [132] y Kalton 1991 [79]). Esto implica la necesidad de conocer, a priori, el tamaño de la población de usos exacto o estimado tanto global como de cada UDT del marco. El diseño muestral sistemático de UDTs, propuesto en esta sección, no precisa del conocimiento exacto o estimado del número de usos.

Para determinar el tamaño de muestra de UDTs en todos los servicios ubicados en un mismo espacio muestral proponemos la siguiente forma de proceder:

- Fijamos un primer servicio y realizamos los cálculos como si se fuera a tomar una muestra aleatoria simple de UDTs para estimar una proporción; el número de UDTs muestrales de este servicio es

$$n_I = \frac{n_{I0}}{\frac{n_{I0}}{N_I} + 1} \quad (2.2)$$

donde

$$n_{I0} = \frac{k^2 P(1-P)}{B^2} \left(\frac{N_I}{N_I - 1} \right)$$

P es el parámetro proporción de individuos, con cierta característica, que se desea estimar

B es el error de estimación para el parámetro P

k está determinada por la confianza que se fije

- Las UDTs muestrales del primer servicio se seleccionan del marco mediante diseño sistemático fraccional con objeto de garantizar el tamaño de muestra exacto y evitar sesgos originados por las periodicidades temporales y de uso.
- Para el segundo servicio se considera la misma muestra de UDTs del primero con dos UDTs de diferencia; son dos porque una se dedica a realizar las encuestas y la otra se emplea en concluir la última encuesta y a desplazarse al siguiente servicio a encuestar. Si hay más servicios se reitera el procedimiento con la condición de que no coincidan las UDTs muestrales de distintos servicios y, cuando se llega al final del marco de UDTs en un servicio, continuaremos por el principio; si el número de servicios en un espacio es elevado será necesario más de un equipo de encuestación pudiendo utilizarse la misma muestra de UDTs para los dos (o más) equipos de encuestación en sus servicios muestrales.

En la Figura 2.2 se muestra un marco de dos semanas ($N_I = 66$) en cuatro servicios con la siguiente selección sistemática fraccional de UDTs ($n_I = 7$) del primer servicio

$$s_1 = \{8, 17, 26, 35, 45, 54, 63\}$$

por tanto, la muestra de UDTs del resto de servicios es

$$s_2 = \{10, 19, 28, 37, 47, 56, 65\}$$

$$s_3 = \{12, 21, 30, 39, 49, 58, 1\}$$

$$s_4 = \{14, 23, 32, 41, 51, 60, 3\}$$

| Marco y Muestra | | Servicio 1 | Servicio 2 | Servicio 3 | Servicio 4 |
|-----------------|---------------|------------|------------|------------|------------|
| Lunes | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Martes | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Miércoles | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Jueves | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Viernes | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| Lunes | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Martes | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Miércoles | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Jueves | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |
| | 16:30 - 17:30 | | | | |
| | 17:30 - 18:30 | | | | |
| Viernes | 9:00 - 10:00 | | | | |
| | 10:00 - 11:00 | | | | |
| | 11:00 - 12:00 | | | | |
| | 12:00 - 13:00 | | | | |
| | 13:00 - 14:00 | | | | |

Figura 2.2: Marco y muestra de unidades de tiempo en cuatro servicios ubicados en un mismo espacio

La elección del número de UDTs para la muestra sistemática debe realizarse con la intención de tomar UDTs que representen convenientemente los distintos días y unidades de tiempo del marco decidido y, además, permitan, si el número de servicios no es excesivamente elevado, ejecutar el trabajo de campo de todos los servicios ubicados en un mismo espacio con un único equipo de encuestación.

Si a priori se conoce, de manera exacta o aproximada, el total de individuos que van a hacer uso del servicio durante el período de encuestación, una alternativa de cálculo de tamaño de muestra de UDTs es obtener, en primer lugar, el tamaño de muestra de individuos, como si se fuera a tomar una muestra aleatoria simple de ellos para estimar una proporción; a partir de este cálculo, y en función de la capacidad de encuestación en cada UDT, determinamos el número de UDTs muestrales. La selección de las UDTs en el marco se realiza con muestreo sistemático fraccional. Situaciones donde se dispone de este tipo de conocimiento a priori son, por ejemplo, servicios con demanda permanente y en el que se realizan citas previas para cubrir toda la capacidad de atención de cada UDT, caso de una consulta de una especialidad en un centro sanitario.

En algunos estudios sociales, la existencia de un presupuesto será lo que determine el tamaño de la muestra de UDTs.

2.4.4. Selección de individuos

Como ya hemos citado, en Sudman (1980) [132] se toma la muestra de UDTs con probabilidad proporcional al número de usos esperado en la UDT. Respecto a la muestra de individuos, en ese mismo artículo se propone tomar una muestra de clientes en cada UDT de manera inversamente proporcional a la probabilidad de selección de la UDT. Kalton (1991) [79] introduce un procedimiento de trabajo de campo, para tomar las unidades de segunda etapa, individuos, consistente en contar las personas que entran (o salen) del servicio objetivo durante la UDT muestral, seleccionando cada k th persona mediante muestreo sistemático, y solicitando a las personas seleccionadas que completen el cuestionario. Una limitación general a este muestreo sistemático de individuos es que si el intervalo muestral tiene la amplitud que permite a los entrevistadores atender los picos de afluencia, éstos pueden estar mucho de su tiempo sin trabajar. Por otro lado, si el intervalo muestral es reducido, los entrevistadores están completamente ocupados, pero no pueden atender los picos de afluencia. Se han propuesto diversos métodos para resolver estos problemas. El primero es tomar una muestra sistemática de tiempos (por ejemplo, cada 10 minutos) y seleccionar el siguiente cliente que entra después de cada tiempo muestral. Una segunda alternativa es tomar el siguiente individuo que entra (o sale) después de la última encuesta completada. Ambos diseños son no probabilísticos.

La mejor opción de selección de individuos en las UDTs muestrales de un servicio es, indudablemente, el censo; las nuevas tecnologías de implementación de las encuestas (tablets, smartphones, ...) facilitan ponerlo en práctica. Sin embargo, en UDTs saturadas el riesgo de falta de respuesta es muy elevado y, como alternativa, si se dispone del conocimiento previo del grado de saturación de las UDTs del marco del servicio, proponemos utilizar el tiempo de atención para tomar una muestra sistemática de clientes y así lograr el control del trabajo de campo que garantice realizar las encuestas necesarias con los recursos disponibles utilizando un diseño probabilístico.

2.4.4.1. Censo de Individuos

El censo de individuos en las UDTs muestrales tiene dos ventajas, precisión y la posibilidad de aplicación en servicios en los que el usuario no tiene un tiempo de atención perfectamente definido, como por ejemplo la visita a un museo o el uso de las instalaciones de una biblioteca o, como veremos en las aplicaciones, en sondeos electorales. El inconveniente se presenta cuando el flujo de clientes en las UDTs muestrales es tan elevado que no es posible encuestar a todos los individuos con los recursos disponibles. Ante esta situación se proponen varias soluciones para mantener el censo como diseño, la primera, la más costosa, aumentar el equipo de encuestadores; la segunda, que la encuesta pase

a ser completada directamente por el encuestado (Tapia, J. y otros 2004 [135]). Una última acción es tratar la falta de respuesta originada tanto por la negativa de los individuos a responder como por la incapacidad de los encuestadores para preguntar a todos los individuos objetivo, para ello, se completa la muestra con individuos del mismo perfil que los que no han contestado. El perfil vendrá determinado por el servicio, la UDT y por aquellas características del individuo que se hayan podido recoger (tipo de usuario, edad, sexo, ...). Un ejemplo de aplicación práctica de este tratamiento de falta de respuesta se encuentra en DGE (2009) [47] donde para el reemplazamiento se toman individuos de la misma UDT del mismo día de la semana pero en otra semana o semanas posteriores dentro del marco tiempo, añadiendo una semana más para poder atender la falta de respuesta de la última semana. Obviamente, esta forma de proceder hace discutible que el diseño muestral siga siendo probabilístico.

ESTIMACIÓN DEL TAMAÑO POBLACIONAL

La estimación del número total de usos se obtiene a partir de la expresión del estimador del total en muestreo sistemático (Särndal y otros 2003 [118]):

$$\hat{N} = \sum_{i=1}^{n_I} \left(\frac{N_i}{\pi_i} \right) \quad (2.3)$$

donde $\pi_i = \frac{n_I}{N_I}$ es la probabilidad de inclusión de la UDT en la muestra.

Atendiendo a la consideración habitual de equivalencia entre diseños sistemático y aleatorio simple, una posible expresión del estimador de la varianza del estimador del tamaño de la población de usos, aunque no la más adecuada, es

$$\widehat{Var}(\hat{N}) = \frac{N_I^2 \left(1 - \frac{n_I}{N_I}\right)}{n_I} \left(\frac{\sum_{i=1}^{n_I} \left(N_i - \left(\frac{\sum_{i=1}^{n_I} N_i}{n_I} \right) \right)^2}{(n_I - 1)} \right) \quad (2.4)$$

La práctica indiscriminada de identificación de muestreo sistemático y muestreo aleatorio simple conduce a estimadores sesgados de la varianza y a inferencias incorrectas sobre el parámetro poblacional de interés. En muestreo sistemático, con probabilidad de selección de los individuos igual, disponemos de ocho estimadores de la varianza alternativos (Wolter 2007 [150]). Otra opción de estimación de la varianza es aplicar la técnica bootstrap (Efron and Tibshirani 1994 [50]) y, como nos encontramos en una situación de población finita, utilizar el bootstrap con factor de corrección propuesto en la sección 1.2; tomadas B remuestras para estimar el parámetro de interés, $\{\hat{N}_i^*\}_{i=1, \dots, B}$, la estimación bootstrap, con factor de corrección, de la varianza es

$$\widehat{Var}_F(\hat{N}) = \left(1 - \frac{n_I}{N_I}\right) \frac{1}{B-1} \sum_{i=1}^B \left(\hat{N}_i^* - \hat{N}^* \right)^2; \hat{N}^* = \frac{\sum_{i=1}^B \hat{N}_i^*}{B} \quad (2.5)$$

El grado de información disponible conduce a distintos estimadores del tamaño de la población de usuarios (N_u):

- Si se conoce el total de usos poblacional, N , y dado que se dispone de la información auxiliar considerada imprescindible (2.1) que descompone el tamaño de muestra de usos de cada UDT muestral, N_i , por ser censo, en tamaño de muestra de usuarios, N_{iu} , y tamaño de muestra de usuarios ya encuestados, $N_{i\bar{u}}$,

$$\{N_i = N_{iu} + N_{i\bar{u}}\}_{i=1, \dots, n_I} \quad (2.6)$$

estimamos la proporción de nuevos usuarios de cada UDT con métodos indirectos

$$\hat{R} = \left(\frac{\sum_{i=1}^{n_I} \frac{N_{iu}}{\pi_i}}{\sum_{i=1}^{n_I} \frac{N_i}{\pi_i}} \right) \quad (2.7)$$

y la estimación del número total de usuarios distintos es

$$\hat{N}_u = N \times \hat{R} \quad (2.8)$$

El bootstrap con factor de corrección permite estimar la varianza de este estimador.

- Si el número total de usos se desconoce, el estimador del total de usuarios es

$$\hat{N}_u = \sum_{i=1}^{n_I} \left(\frac{N_{iu}}{\pi_i} \right) \quad (2.9)$$

y, como estimador del error, consideramos

$$\widehat{Var}(\hat{N}_u) = \frac{N_I^2 \left(1 - \frac{n_I}{N_I}\right)}{n_I} \left(\frac{\sum_{i=1}^{n_I} \left(N_{iu} - \left(\frac{\sum_{i=1}^{n_I} N_{iu}}{n_I} \right) \right)^2}{(n_I - 1)} \right) \quad (2.10)$$

ESTIMACIÓN DE LA RESPUESTA MEDIA POBLACIONAL

El número de clientes, en cada UDT, no tiene por qué ser el mismo, en consecuencia, las UDTs se pueden considerar conglomerados de clientes de distinto tamaño. Para este diseño, el estimador de la media poblacional de los usos, μ , es el estimador de razón (Särndal y otros 2003 [118])

$$\hat{\mu}_C = \frac{\sum_{i=1}^{n_I} \left(\frac{T_i}{\pi_i} \right)}{\sum_{i=1}^{n_I} \left(\frac{N_i}{\pi_i} \right)} \quad (2.11)$$

Si se conoce el total de usos en el marco del tiempo, un estimador de la varianza del estimador de la media es

$$\widehat{Var}(\hat{\mu}_C) = \frac{\left(1 - \frac{n}{N}\right)}{\left(\frac{N}{N_I}\right)^2 n_I} \left(\frac{\sum_{i=1}^{n_I} (T_i - \hat{\mu}_C N_i)^2}{(n_I - 1)} \right) \quad (2.12)$$

Si se desconoce el número total de usos, N , estimamos la varianza utilizando bootstrap con factor de corrección.

Con respecto a la población de usuarios, el estimador de la media poblacional, μ_u , es

$$\hat{\mu}_{Cu} = \frac{\sum_{i=1}^{n_I} \left(\frac{T_{iu}}{\pi_i} \right)}{\sum_{i=1}^{n_I} \left(\frac{N_{iu}}{\pi_i} \right)} \quad (2.13)$$

y el error se estima utilizando bootstrap con factor de corrección.

2.4.4.2. Muestra de individuos proporcional a su tiempo de atención

Si todas las UDTs del marco son saturadas y la afluencia de clientes hace imposible encuestarles a todos con los recursos disponibles, una alternativa al censo es tomar, en cada UDT muestral, la muestra de individuos n_i con diseño sistemático proporcional a su tiempo de atención. Kalton (1991) [79] propone como forma de actuar, tomar una muestra sistemática de tiempos (por ejemplo cada 10 minutos) y seleccionar al individuo que entra después de cada tiempo que conforma la

muestra; este diseño muestral es no probabilístico.

La principal ventaja de utilizar el tiempo de atención para tomar una muestra sistemática es el control del trabajo de campo dado que se garantiza la capacidad de encuestación en cada UDT con los recursos disponibles. El inconveniente es que el tiempo de atención debería ser un buen predictor de la opinión, aspecto, que en la práctica, puede no ocurrir, sin embargo, el diseño es probabilístico y las simulaciones muestran una calidad aceptable de las estimaciones debido a su similitud con las obtenidas con censo.

Los datos imprescindibles para llevar a cabo esta metodología muestral son el tiempo de atención del cliente, sin incluir la espera, y el tiempo total de encuestación de las UDTs muestrales, definido como la medida de tiempo de la UDT más el tiempo que el último cliente sobrepasa esta cantidad si éste está en la muestra. La forma más precisa de tomar estas medidas es por el propio encuestador o un equipo de al menos dos encuestadores, uno encuesta y el otro toma el tiempo, dado que el cliente puede que no responda con exactitud a este dato.

Una vez determinado el tamaño muestral de las UDTs, el número de encuestas a realizar en cada UDT se calcula como la razón entre el tamaño temporal de la UDT y la duración aproximada de una encuesta. Por ejemplo, si la UDT es de 60 minutos y t son los minutos que dura una encuesta, el número de individuos (usos) a encuestar es

$$n_k = 60/t$$

Para tomar la muestra de usos, se selecciona, de manera aleatoria, una semilla i en el intervalo $(0, t]$ y se realiza el cuestionario a los individuos atendidos en el servicio en los tiempos $i, i+t, i+2t, \dots$. En conclusión, el diseño muestral de individuos es sistemático proporcional a su tiempo de atención, por lo que la probabilidad de inclusión en la muestra de un individuo que es atendido en la k th UDT es (Särndall y otros 2003 [118])

$$\left(\pi_{i/k} = \begin{cases} \frac{n_k x_i}{T_k - c} & \text{si } x_i \leq t \\ 1 & \text{si } x_i > t \end{cases} \right) \begin{matrix} k = 1, \dots, n_I \\ i = 1, \dots, n_k \end{matrix}$$

donde

T_k Tiempo total de encuestación de la UDT k th.

x_i Tiempo de atención del i th individuo de la muestra, sin contar el tiempo de espera.

$c = T_k - tn$

n Parte entera de T_k/t

Aunque, a priori, el número de encuestados en cada UDT es el mismo

$$\{n_k = n_t\}_{k=1,2,\dots,n_I}$$

esto no ocurre necesariamente si el tiempo de atención de uno o más individuos de la muestra supera el tiempo t que dura una encuesta. En muestreo bietápico el error se estima sumando la estimación de los errores en cada etapa. Si esto no es posible, aplicaremos bootstrap en dos etapas como se ha introducido en la sección 1.2.

ESTIMACIÓN DEL TAMAÑO POBLACIONAL

Si la unidad estadística es el uso, las expresiones (2.3) y (2.4) siguen siendo válidas para estimar el tamaño de esta población.

Para estimar el total de usuarios, el primer paso es estimar el número de usuarios en cada UDT, como el total de un dominio (Särndall y otros 2003 [118]), utilizando el conocimiento la

descomposición de los usos muestrales de cada UDT en usuarios nuevos y usuarios ya encuestados, $\{n_k = n_{ku} + n_{k\bar{u}}\}_{k=1, 2, \dots, n_I}$

$$\hat{N}_{ku} = \sum_{j=1}^{n_{ku}} \frac{1}{\pi_{j/k}}; \quad k = 1, \dots, n_I \quad (2.14)$$

El estimador del tamaño de la población de usuarios distintos es

$$\hat{N}_u = \sum_{k=1}^{n_I} \frac{\hat{N}_{ku}}{\pi_{Ik}} \quad (2.15)$$

y la estimación del error es la suma de los errores en primera y segunda etapa

$$\widehat{AVar}(\hat{N}_u) = \sum_{l=1}^{n_I} \sum_{k=1}^{n_I} \check{\Delta}_{lkl} \frac{\hat{N}_{lu} \hat{N}_{ku}}{\pi_{Il} \pi_{Ik}} + \sum_{k=1}^{n_I} \frac{\widehat{AVar}(\hat{N}_{ku})}{\pi_{Ik}} \quad (2.16)$$

El error en primera etapa se estima considerando que la muestra de unidades en primera etapa es aleatoria simple. Para estimar el error en la segunda etapa, Wolter (2007) [150] ofrece diferentes opciones de estimadores de la varianza del total en muestreo sistemático proporcional, recogidas en la sección 1.1; de entre ellas, nuestra elección es

$$\widehat{AVar}(\hat{N}_{ku}) = \frac{1}{n_I(n_I - 1)} \sum_{i=1}^{n_k} \left(\frac{1}{p_i} - \hat{N}_{ku} \right)^2 I_i; \quad p_i = \frac{x_i}{T_N} \quad (2.17)$$

Si el total de usos, N , se conoce, los estimadores de razón determinan la expresión del estimador del total de usuarios

$$\hat{N}_u = \hat{P}_u N \quad (2.18)$$

donde

$$\hat{P}_u = \left(\frac{\sum_{k=1}^{n_I} \frac{\hat{N}_{ku}}{\pi_{Ik}}}{\sum_{k=1}^{n_I} \frac{\hat{N}_k}{\pi_{Ik}}} \right) \quad (2.19)$$

y, para estimar el error, se aplica bootstrap en dos etapas.

ESTIMACIÓN DE LA MEDIA POBLACIONAL

En la población de usos, dado que el número de individuos atendidos no tiene por qué ser el mismo en todas las UDTs, conglomerados de distinto tamaño, la media se estima con un estimador de razón (Särndal y otros 2003 [118])

$$\hat{\mu} = \frac{\sum_{k=1}^{n_I} \left(\frac{\hat{T}_k}{\pi_{Ik}} \right)}{\sum_{k=1}^{n_I} \left(\frac{\hat{N}_k}{\pi_{Ik}} \right)} \quad (2.20)$$

$$\text{donde } \hat{T}_k = \sum_{j=1}^{n_k} \frac{y_{jk}}{\pi_{j/k}} \text{ y } \hat{N}_k = \sum_{j=1}^{n_k} \frac{1}{\pi_{j/k}}$$

y para estimar el error, se aplica bootstrap en dos etapas.

Si la población son los usuarios, la media poblacional también se estima con el estimador de razón

$$\hat{\mu}_u = \frac{\sum_{k=1}^{n_I} \left(\frac{\hat{T}_{ku}}{\pi_{Ik}} \right)}{\sum_{k=1}^{n_I} \left(\frac{\hat{N}_{ku}}{\pi_{Ik}} \right)} \quad (2.21)$$

$$\text{donde } \hat{T}_{ku} = \sum_{j=1}^{n_k} \frac{y_{jk}}{\pi_{j/k}} I_j$$

y, para estimar el error, se aplica bootstrap en dos etapas.

2.4.4.3. Postestratificación de unidades de tiempo

En la práctica, el censo de clientes en las UDTs muestrales conlleva, en muchos casos, falta de respuesta en horas saturadas mientras que el muestreo proporcional al tiempo de atención está condicionado a un marco del servicio con todas las UDTs saturadas. Las limitaciones prácticas que por separado tienen estas dos metodologías se resuelven si se dispone de la información auxiliar previa que identifica, en el marco, las UDTs saturadas y no saturadas. En estas circunstancias hay dos opciones, tomar una única muestra de UDTs y postestratificar o tomar dos muestras sistemáticas independientes de UDTs, una en el estrato de UDTs saturadas y otra en el estrato de UDTs no saturadas. La opción de estratificación exigiría un aumento del tamaño de muestra de unidades de primera etapa para tener estimaciones de calidad en ambos estratos, lo cual no tiene interés, y, además, sería necesario tomar dos muestras sistemáticas distintas en cada servicio, una para cada estrato, lo cual hace inviable que un sólo equipo de encuestación pueda realizar las encuestas de todos los servicios de un mismo espacio. Por tanto, las razones económicas y el interés en obtener estimaciones en toda la población de UDTs justifican tomar como opción la postestratificación y descartar la estratificación previa de UDTs.

Suponemos que la información auxiliar clasifica, a priori, todas las UDTs del marco en dos estratos, de uso parcial, N_{IC} , y de uso saturado N_{IP}

$$N_I = N_{IC} + N_{IP} \quad (2.22)$$

La muestra sistemática fraccional de n_I UDTs contendrá unidades saturadas, s_{IC} , como no saturadas, s_{IP} ,

$$s_I = s_{IC} \cup s_{IP}$$

por tanto,

$$n_I = n_{IC} + n_{IP} \quad (2.23)$$

Esta clasificación determina la postestratificación de las UDTs.

En las UDTs de uso saturado la muestra de individuos se toma proporcional al tiempo de atención y en las de uso no saturado se realiza censo.

A la hora de proponer los distintos estimadores, hay que tener en cuenta que las muestras de individuos, entre las UDTs muestrales, se toman de manera independiente.

ESTIMACIÓN DEL TAMAÑO POBLACIONAL

Si el tamaño de la población de usos se desconoce, las expresiones (2.3) y (2.4) son válidas para estimarlo ya que se dispone de la información necesaria para poder utilizarlas.

Dado que se conoce la descomposición de todo el marco de UDTs en los dos estratos, $N_I = N_{IC} + N_{IP}$, el estimador del tamaño de la población de usuarios, N_u , atendiendo a la postestratificación de las UDT muestrales, es (Sudgen and Smith 2006 [133])

$$\hat{N}_u = N_{IC} \sum_{i \in s_{IC}} \left(\frac{N_{iu}}{n_{IC}} \right) + N_{IP} \sum_{i \in s_{IP}} \left(\frac{\hat{N}_{iu}}{n_{IP}} \right) \quad (2.24)$$

Para estimar el error, aplicaremos bootstrap en dos etapas, introducido en 1.2 (Field and Welsh, 2007 [57]), remuestreando, en primer lugar, las unidades de primera etapa, UDTs; sólo en las UDTs saturadas de cada remuestra bootstrap remuestrearemos en las unidades de segunda etapa, usos o usuarios; en las UDTs no saturadas dispondremos de la información censal.

Si el total de usos, N , se conoce, estimamos el total de usuarios como un estimador de razón

$$\hat{N}_u = \hat{P}_u N \quad (2.25)$$

donde

$$\hat{P}_u = \frac{N_{IC} \sum_{i \in s_{IC}} \left(\frac{N_{iu}}{n_{ic}} \right) + N_{IP} \sum_{i \in s_{IP}} \left(\frac{\hat{N}_{iu}}{n_{IP}} \right)}{N_{IC} \sum_{i \in s_{IC}} \left(\frac{N_i}{n_{ic}} \right) + N_{IP} \sum_{i \in s_{IP}} \left(\frac{\hat{N}_i}{n_{IP}} \right)}$$

y, para estimar el error, se aplica bootstrap en dos etapas.

ESTIMACIÓN DE LA MEDIA POBLACIONAL

A la hora de estimar la media poblacional de los usos hay que tener en cuenta que las UDTs son conglomerados de clientes de distinto tamaño y el estimador es (Sudgen and Smith 2006 [133])

$$\hat{\mu} = \frac{N_{IC} \sum_{i \in s_{IC}} \left(\frac{T_i}{n_{ic}} \right) + N_{IP} \sum_{i \in s_{IP}} \left(\frac{\hat{T}_i}{n_{IP}} \right)}{N_{IC} \sum_{i \in s_{IC}} \left(\frac{N_i}{n_{ic}} \right) + N_{IP} \sum_{i \in s_{IP}} \left(\frac{\hat{N}_i}{n_{IP}} \right)} \quad (2.26)$$

La técnica bootstrap en dos etapas permite estimar la varianza de este estimador.

Como estimador de la media poblacional de los usuarios se propone

$$\hat{\mu}_u = \frac{N_{IC} \sum_{i \in s_{IC}} \left(\frac{T_{iu}}{n_{ic}} \right) + N_{IP} \sum_{i \in s_{IP}} \left(\frac{\hat{T}_{iu}}{n_{IP}} \right)}{N_{IC} \sum_{i \in s_{IC}} \left(\frac{N_{iu}}{n_{ic}} \right) + N_{IP} \sum_{i \in s_{IP}} \left(\frac{\hat{N}_{iu}}{n_{IP}} \right)} \quad (2.27)$$

y estimar la varianza mediante la técnica bootstrap en dos etapas.

2.4.5. Efecto del Diseño

La utilización de un diseño cluster o polietápico lleva asociado un posible incremento del error de estimación. En un diseño cluster, la varianza de un estimador es habitualmente mayor que la varianza obtenida con una muestra aleatoria simple sin reemplazamiento del mismo tamaño. Este incremento se cuantifica como efecto del diseño y se define como el cociente entre la varianza computada con el diseño y la varianza que se obtendría si la muestra del mismo tamaño se tomara aleatoria simple sin reemplazamiento. En la metodología definida en esta tesis como censo, teniendo en cuenta que el tamaño de la población es desconocido, el efecto del diseño para estimar la media poblacional de los usos es (chap 8, Särndal y otros 2003 [118]):

$$def f = 1 + \frac{N - N_I}{N_I - 1} \delta + \frac{Cov}{\bar{N} S_{yU}^2} \quad (2.28)$$

donde

S_{yU}^2 es la varianza poblacional de la variable de interés

$$\bar{N} = \frac{\sum_{i=1}^{N_I} N_i}{N_I}$$

δ es el coeficiente de homogeneidad dado por $\delta = 1 - \frac{S_{yW}^2}{S_{yU}^2}$

con $S_{yW}^2 = \frac{\sum_{i=1}^{N_I} (N_i - 1) S_{yU_i}^2}{\sum_{i=1}^{N_I} (N_i - 1)}$ donde $S_{yU_i}^2$ es la varianza de la variable de interés en el i th cluster

Cov es la covarianza entre N_i y $A_i = N_i \left(\frac{T_i}{N_i} \right)^2$.

Para el caso polietápico, con o sin postestratificación, resulta imposible obtener una expresión analítica del efecto del diseño. En estos casos, la imposibilidad del cálculo directo del efecto del diseño conduce a la necesidad del uso de las simulaciones para poder obtenerlo. Esta simulación se realiza en la sección 2.5 donde obtendremos una estimación del efecto del diseño.

2.4.6. Coste del Diseño

En un estudio social, el presupuesto está determinado, principalmente, por el coste de los investigadores y/o el coste del personal del trabajo de campo (Fricker y otros 2002 [59]). El coste de la metodología de encuesta cara-a-cara propuesto en este capítulo de la tesis se puede considerar “razonable” porque se minimiza el número de desplazamientos de los equipos de encuestación y la actividad de estos equipos se maximiza al ser casi permanente. En el ejemplo de la Figura 2.2, un equipo de encuestación puede realizar la encuesta en cuatro servicios cada día, sólo necesita la capacidad de desplazamiento, durante la duración de una UDT, de uno a otro servicio de la muestra. En el caso real de las bibliotecas de la Universidad de Valladolid de la sección 2.6.1, cinco entrevistadores realizan el trabajo de campo en 14 bibliotecas, localizadas en cuatro capitales de provincia, en una semana.

Una función de coste, tradicional, del trabajo de campo es (chap. 12 Särndal y otros 2003 [118])

$$C = C_o + \sum_{h=1}^H n_h C_h \quad (2.29)$$

donde

C_o : Coste inicial del experimento

H : número de estratos

n_h : número de encuestas en estrato $h = 1, \dots, H$

C_h : Coste de cada encuesta en el estrato $h = 1, \dots, H$

En el modelo propuesto en este capítulo, la función de coste pasa a ser

$$C = C_o + \sum_{i=1}^{n_e} C_t * t_i + \sum_{i=1}^{n_e} C_{nd} * nd_i \quad (2.30)$$

donde

n_e : número de encuestadores (équipo de)

C_t : coste de encuestar en cada UDT

t_i : número de UDTs que el i th encuestador (equipo de) está encuestando

C_{nd} : coste de desplazamiento de un encuestador (equipo de) entre puntos de encuestación

nd_i : número de desplazamientos del i th encuestador (equipo de)

El número de encuestas realizadas no se considera en esta función pero sí el tiempo dedicado a encuestar, que está perfectamente controlado por el diseño. El coste de desplazamientos se minimiza por el control que de los mismos puede tener el responsable(s) del estudio.

2.5. Validación de la metodología muestral

El objeto de esta sección es comprobar, mediante simulaciones, la calidad de los distintos diseños muestrales, propuestos en este capítulo, a la hora de estimar tanto el tamaño de la población de usos o usuarios como la respuesta media y compararemos los errores observados en las estimaciones realizadas con censo en todas las UDTs frente a los errores cometidos con las otras dos alternativas de selección de individuos en que se utiliza el tiempo de atención del cliente, comprobando que el control

del trabajo de campo y de la falta de respuesta, por utilizar este tiempo de atención, no conlleva una pérdida significativa en los errores de estimación. Aunque la metodología muestral se ha desarrollado considerando todas las etapas (muestra de espacios, censo de servicios, muestra de UDTs y muestra de individuos) nos centramos en la estimación de los parámetros de un servicio por ser la etapa donde reside la mayor novedad, y dificultad, de la metodología propuesta.

2.5.1. Simulación de una población con todas las unidades de tiempo de uso saturado

En esta primera población simulada consideramos todas las UDTs del marco de un servicio saturadas con objeto de comparar el censo de individuos en todas las UDTs (conglomerados) y muestra de individuos proporcional al tiempo de atención en todas las UDTs (polietápico).

Las características de la población simulada son las siguientes:

- Todas las UDTs del marco son de uso saturado.
- La encuesta se realiza en un servicio de un espacio, durante 20 días (4 semanas de 5 días con atención al público) y cada día se atiende a los clientes durante 8 horas.
- Se supone un único servidor y una única cola en cada servicio.
- Las 160 horas de posible encuestación se consideran el marco de UDTs.
- El tiempo máximo de atención de un cliente es de 20 minutos y la duración mínima de 3 minutos.
- Aunque la UDT es de 60 minutos se considera que se puede extender hasta un máximo de 79 minutos, situación que ocurre si un cliente llega en el minuto 59 y se le atiende el máximo de tiempo, 20 minutos.
- El tiempo estimado de realización de una encuesta es de 6 minutos.
- Para determinar los usos que dejan de ser usuarios (clientes no encuestados con anterioridad) se utiliza una función que determina la probabilidad de cada individuo de ser o no usuario en cada UDT del marco

$$p_t = 0.0005(t - 1)_{(t=1, \dots, 160)}$$

es decir, en la UDT 160 la probabilidad de que un individuo sea usuario es del 92,1 % y, por tanto, el 7,9 % de los clientes ya ha sido encuestado con anterioridad.

- La satisfacción del cliente es más elevada si tarda menos en ser atendido.

La Tabla 2.1 recoge los valores de los parámetros de la población simulada

Tabla 2.1: Parámetros de la población simulada con todas las UDTs saturadas

| Usos N | Usuarios N_u | Satisfacción media de los usos $\hat{\mu}$ | Satisfacción media de los usuarios $\hat{\mu}_u$ |
|-------------|-------------------|---|---|
| 2018 | 1920 | 7.20 | 7.20 |

Los pasos del procedimiento de simulación para comparar los diseños son:

- I. Se determina el tamaño de muestra de UDTs fijando el error de estimación, para estimar una proporción, menor o igual al 2% con muestreo aleatorio simple sin reemplazamiento y una confianza del 95%; se obtiene $n_I = 21$.
- II. La muestra de UDTs se toma con muestreo sistemático fraccional. En cada UDT muestral se toma el censo de individuos, para el diseño por conglomerados, o una muestra sistemática de individuos proporcional a su tiempo de atención, para el diseño polietápico; el tamaño de la muestra sistemática es de 10 individuos en cada UDT teniendo en cuenta que la duración de la encuesta es de 6 minutos.
- III. Los parámetros ($\theta = \mu, N_u$ ó μ_u) son estimados, en los dos diseños y en las distintas opciones de estimación, y se obtiene el intervalo

$$I_\theta = [\hat{\theta} - B_\theta, \hat{\theta} + B_\theta] \quad (2.31)$$

donde $B_\theta = 1.96\sqrt{\widehat{Var}(\hat{\theta})}$ es el error de estimación.

- IV. Se realizan mil iteraciones de los pasos II. y III. , obteniéndose los intervalos $\{I_\theta^{(k)}\}_{k=1, \dots, 1000}$, los errores de estimación $\{B_\theta^{(k)}\}_{k=1, \dots, 1000}$, y el efecto del diseño $def f^{(k)} = B_\theta^{(k)} / B_{\theta SI}^{(k)}$, donde $B_{\theta SI}^{(k)}$ es el error de estimación tomando una muestra aleatoria simple sin reemplazamiento del mismo tamaño.
- V. Se aproxima la confianza con

$$C_\theta = \left(\frac{1}{1000} \sum_{k=1}^{1000} I_{(\theta \in I_\theta^{(k)})} \right) \quad (2.32)$$

- VI. Se analiza el error de estimación a partir de boxplots y calculando

$$\bar{B}_\theta = \left(\frac{1}{1000} \sum_{k=1}^{1000} B_\theta^{(k)} \right) \quad (2.33)$$

RESULTADOS DE LA SIMULACIÓN

Las Tablas 2.2 y 2.3 comparan confianzas y errores de estimación de los estimadores del tamaño de la población de usuarios propuestos para el diseño por conglomerados y polietápico, cuando el tamaño poblacional de usos es desconocido y conocido, respectivamente.

Tabla 2.2: Población con todas las UDTs saturadas: Aproximación de confianza y error de estimación del total de usuarios con con tamaño de población de usos, N , desconocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|----------------------|-------------|----------------------------|------------|------------------|
| Conglomerados | (ec. 2.9) | (ec. 2.10) | 97.8% | 299.83 |
| Polietápico | (ec. 2.15) | (ec. 2.16) | 91% | 326.09 |

Tabla 2.3: Población con todas las UDTs saturadas: Aproximación de confianza y error de estimación del total de usuarios con con tamaño de población de usos, N , conocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|---------------|-------------|----------------------------|------------|------------------|
| Conglomerados | (ec. 2.8) | Correction Factor | 99 % | 51.51 |
| | | Bootstrap | | |
| Polietápico | (ec. 2.18) | Bootstrap en dos etapas | 97.1 % | 96.07 |

La Figura 2.3 muestra los errores de estimación de los estimadores del total de usos utilizando como estimador (2.3) y como estimador de la varianza (2.4) o (2.5) (estimación bootstrap con factor de corrección). La Figura 2.4 muestra los errores de estimación de los estimadores del total de usuarios propuestos para el diseño por conglomerados y polietápico con el tamaño de usos desconocido y conocido.

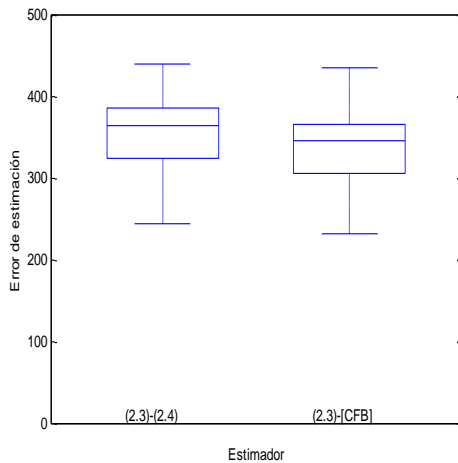


Figura 2.3: Población con todas las UDTs saturadas: Error de estimación del total de usos \hat{N}

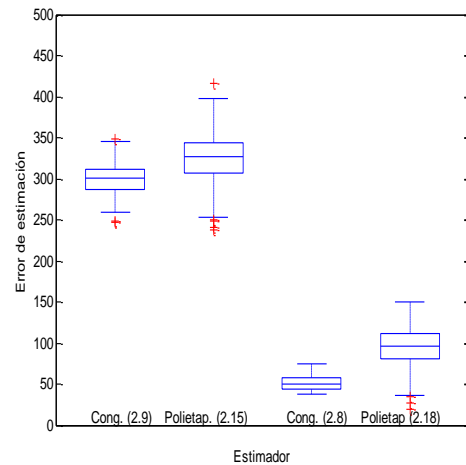


Figura 2.4: Población con todas las UDTs saturadas: Error de estimación del total de usuarios \hat{N}_u con tamaño de usos, N , desconocido y conocido

Las Tablas 2.4 y 2.5 comparan los resultados de la simulación entre los diseños conglomerados y polietápico para estimar la media poblacional de los usos y de los usuarios, respectivamente.

Tabla 2.4: Población con todas las UDTs saturadas: Estimación de la media poblacional de usos

| Diseño | $\hat{\mu}$ | $\widehat{Var}(\hat{\mu})$ | C_θ | \bar{B}_θ |
|---------------|-------------|----------------------------|------------|------------------|
| Conglomerados | (ec. 2.11) | (ec. 2.12) | 98.7 % | 0.3458 |
| | | Correction Factor | 98.1 % | 0.3568 |
| Polietápico | (ec. 2.20) | Bootstrap | 99.1 % | 0.3991 |
| | | Bootstrap en dos etapas | | |

Tabla 2.5: Población con todas las UDTs saturadas: Estimación de la media poblacional de usuarios

| Diseño | $\hat{\mu}_u$ | $\widehat{Var}(\hat{\mu}_u)$ | C_θ | \bar{B}_θ |
|----------------------|---------------|--------------------------------|------------|------------------|
| Conglomerados | (ec. 2.13) | Correction Factor Bootstrap | 98.1 % | 0.3564 |
| Poiletápico | (ec. 2.21) | Bootstrap en dos etapas | 99.7 % | 0.5065 |

Las Figuras 2.5 y 2.6 comparan, entre los diseños conglomerados y polietápico, los errores de estimación de la respuesta media poblacional de los usos y los usuarios, respectivamente.

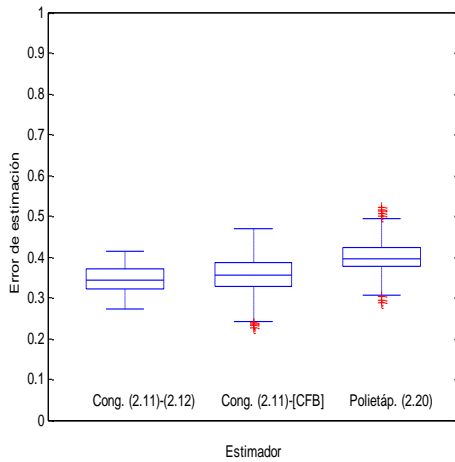


Figura 2.5: Población con todas las UDTs saturadas: Error de estimación de la media poblacional de usos $\hat{\mu}$

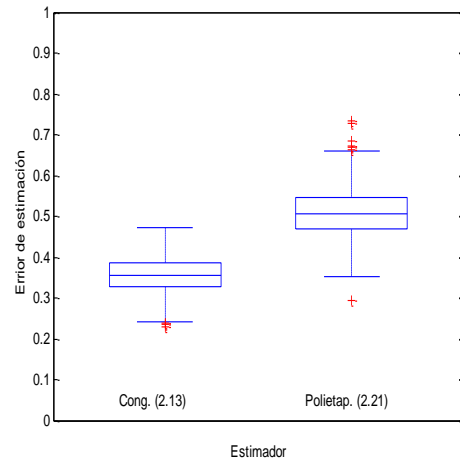


Figura 2.6: Población con todas las UDTs saturadas: Error de estimación de la media poblacional de usuarios $\hat{\mu}_u$

En la Figura 2.7 se compara el efecto del diseño (censo de usos con tamaño de población conocido - censo de usos con tamaño de población de usos desconocido - diseño polietápico) para estimar la satisfacción media de los usos. El error cometido con los tres diseños es mayor que el cometido tomando una muestra del mismo tamaño con muestreo aleatorio simple sin reemplazamiento. La comparación entre ellos pone de manifiesto un mayor error en el diseño polietápico y menor en el censo con tamaño de población de usos conocido. De todos modos, se puede considerar que, en los tres casos, hay bastante similitud en el efecto del diseño.

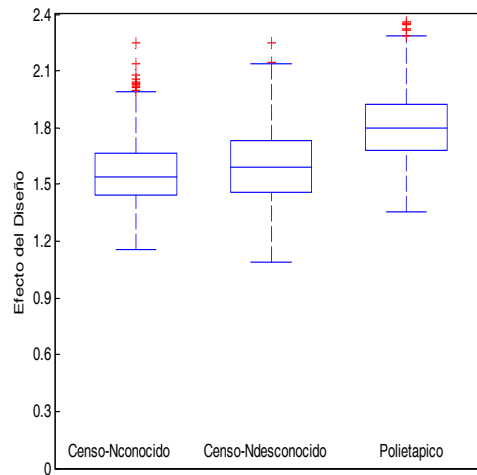


Figura 2.7: Población con todas las UDTs saturadas: Efecto del diseño (censo con y sin tamaño de población conocido - polietápico) en la estimación de la satisfacción media de los usos

CONCLUSIONES DE LA SIMULACIÓN

Si un servicio tiene todas las UDTs saturadas, las Tablas 2.2, 2.3, 2.4 y 2.5 muestran la similitud existente entre diseño polietápico y por conglomerados tanto en la confianza de los intervalos como en el error de estimación. En la Figura 2.4 se observa una notable reducción en el error de estimación del total de usuarios lograda sí, una vez concluido el trabajo de campo, se dispone del total de usos, N . Del análisis del efecto del diseño se concluye que la pérdida observada en el error cometido con censo de usos o diseño polietápico se compensa por la viabilidad en su ejecución del trabajo de campo y el ahorro en costes; en esta población humana móvil (Kalton 1991 [79]) tomar una muestra aleatoria simple es inviable.

2.5.2. Simulación de una población con unidades de tiempo de uso saturado y no saturado

En la segunda simulación se va a considerar un marco con UDTs saturadas y no saturadas con el fin de comparar el censo de individuos en todas las UDTs muestrales y el diseño polietápico con postestratificación de UDTs muestrales en saturadas y no saturadas; en las unidades saturadas se toma la muestra de individuos proporcionalmente a su tiempo de atención mientras que en las no saturadas se toma censo de individuos.

Las características de la población simulada coinciden con las de la primera población simulada excepto en la consideración de que todas las UDTs del marco son de uso saturado; en este caso las UDTs tercera, cuarta y séptima de cada día se consideran de uso saturado y el resto no. La Tabla 2.6 recoge los valores de los parámetros de la población simulada

Tabla 2.6: Parámetros de la población simulada con UDTs saturadas y no saturadas

| Usos N | Usuarios N_u | Satisfacción media de los usos $\hat{\mu}$ | Satisfacción media de los usuarios $\hat{\mu}_u$ |
|-------------|-------------------|---|---|
| 1868 | 1800 | 7.494 | 7.491 |

RESULTADOS DE LA SIMULACIÓN:

En las Tablas 2.7 y 2.8 se comparan las confianzas y los errores de estimación de los estimadores del tamaño de la población de usuarios propuestos para el diseño por conglomerados y polietápico con postestratificación, cuando el tamaño poblacional de usos es desconocido y conocido, respectivamente.

Se adopta como notación, en tablas y figuras, diseño por conglomerados al censo en las UDTs muestrales y diseño postestratificado al diseño polietápico con postestratificación de las UDTs en saturadas y no saturadas.

Tabla 2.7: Población con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población de usos, N , desconocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|--------------------------|-------------|----------------------------|------------|------------------|
| Conglomerados | (ec. 2.9) | (ec. 2.10) | 100 % | 346.1 |
| Postestratificado | (ec. 2.24) | Bootstrap en dos etapas | 97.5 % | 384.1 |

Tabla 2.8: Población con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población de usos, N , conocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|--------------------------|-------------|--------------------------------|------------|------------------|
| Conglomerados | (ec. 2.8) | Correction Factor Bootstrap | 94.6 % | 46.24 |
| Postestratificado | (ec. 2.25) | Bootstrap en dos etapas | 94 % | 66.82 |

Las Tablas 2.9 y 2.10 comparan los diseños por conglomerados y polietápico con postestratificación para estimar la media poblacional de usos y usuarios respectivamente.

Tabla 2.9: Población con UDTs saturadas y no saturadas: Estimación de la media poblacional de usos

| Diseño | $\hat{\mu}$ | $\widehat{Var}(\hat{\mu})$ | C_θ | \bar{B}_θ |
|--------------------------|-------------|--|------------------|------------------|
| Conglomerados | (ec. 2.11) | (ec. 2.12) Correction Factor Bootstrap | 89.4 % 89.3 % | 0.2175 0.2249 |
| Postestratificado | (ec. 2.26) | Bootstrap en dos etapas | 94.5 % | 0.3250 |

Tabla 2.10: Población con UDTs saturadas y no saturadas: Estimación de la media poblacional de usuarios

| Diseño | $\hat{\mu}_u$ | $\widehat{Var}(\hat{\mu}_u)$ | C_θ | \bar{B}_θ |
|--------------------------|---------------|--------------------------------|------------|------------------|
| Conglomerados | (ec. 2.13) | Correction Factor Bootstrap | 88.8 % | 0.2312 |
| Postestratificado | (ec. 2.27) | Bootstrap en dos etapas | 98.8 % | 0.4494 |

Las Figuras 2.8, 2.9 y 2.10 comparan, en los dos diseños, los errores de estimación de los estimadores del total de usuarios, de la media poblacional de usos y de la media poblacional de usuarios, respectivamente.

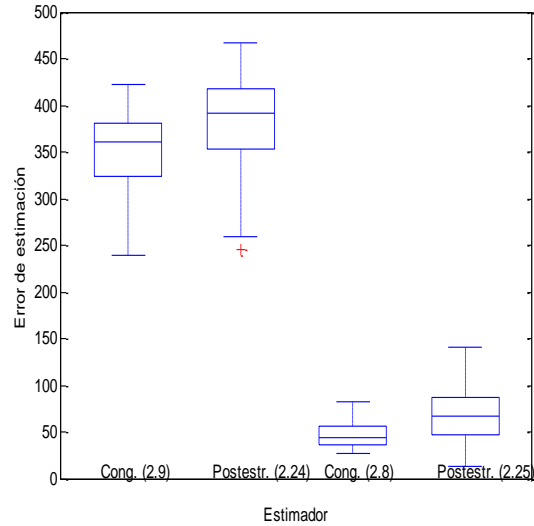


Figura 2.8: Población con UDTs saturadas y no saturadas: Error de estimación del total de usuarios, \hat{N}_u

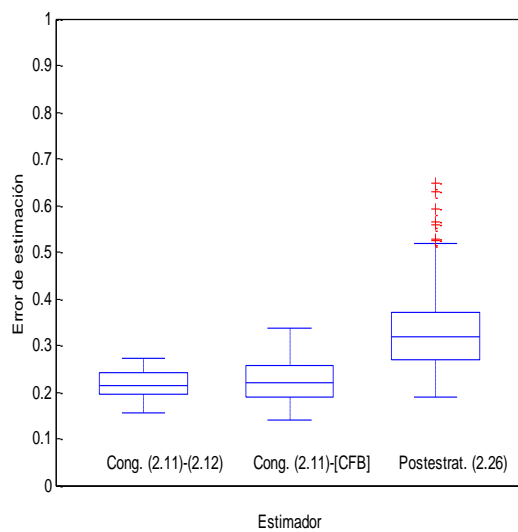


Figura 2.9: Población con UDTs saturadas y no saturadas: Error de estimación de la media poblacional de usos, $\hat{\mu}$

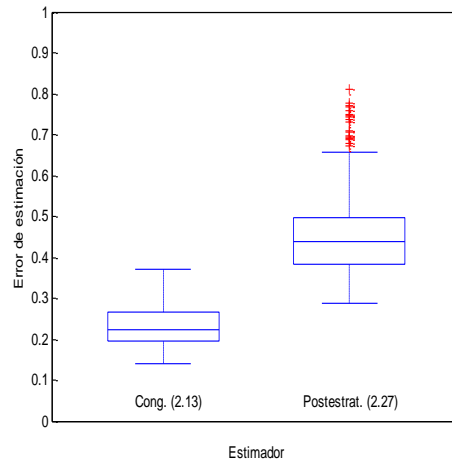


Figura 2.10: Población con UDTs saturadas y no saturadas: Error de estimación de la media poblacional de usuarios, $\hat{\mu}_u$

En la Figura 2.11 se compara el efecto del diseño (censo de usos con tamaño de población conocido - censo de usos con tamaño de población de usos desconocido - diseño polietápico con postestratificación) para estimar la satisfacción media de los usos. El error cometido con los tres diseños es mayor que el cometido tomando una muestra del mismo tamaño con muestreo aleatorio simple sin reemplazamiento. Sin embargo, hay que tener en cuenta que, dadas las características de la población móvil objetivo, en la práctica resulta inviable tomar directamente una muestra aleatoria simple de usos o de usuarios.

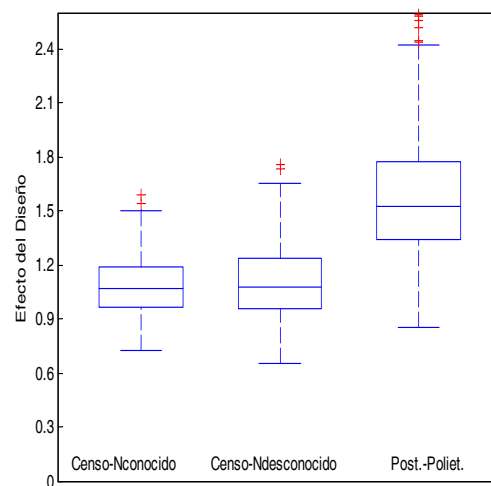


Figura 2.11: Población con UDTs saturadas y no saturadas: Efecto del diseño (censo con tamaño de población de usos conocido - censo con tamaño de población de usos desconocido - polietápico con postestratificación) en la estimación de la satisfacción media de los usos

CONCLUSIONES DE LA SIMULACIÓN

Si un servicio tiene UDTs saturadas y no saturadas, las Tablas 2.7 y 2.8 muestran la similitud existente entre diseño polietápico con postestratificación y diseño por conglomerados tanto en la confianza de los intervalos como en los errores de estimación para estimar el tamaño de la población

de usuarios. En la Figura 2.8 se observa una notable reducción en el error de estimación del tamaño de la población de usuarios si se conoce el tamaño de la población de usos, N , finalizado el trabajo de campo. En las Tablas 2.9 y 2.10 y en las Figuras 2.9 y 2.10 se observa un error de estimación de la media poblacional de usos y de usuarios algo superior en el diseño polietápico con postestratificación que en el censo, sin embargo, la aproximación a las confianzas de los intervalos propuestos en el diseño polietápico con postestratificación son más elevadas. Además, el aumento en el error de estimación utilizando diseño polietápico con postestratificación frente al censo se compensa por la disminución en la falta de respuesta que se logra por el control del trabajo de campo, ya que la postestratificación permite ajustar los recursos a las posibles encuestas realizables en una UDT.

2.6. Estudios de aplicación del censo de usos

En esta sección presentamos dos experimentos llevados a cabo bajo la supervisión total o parcial del autor de esta tesis, en diferentes ámbitos, utilizando la metodología de censo de usos en las UDTs muestrales. Además se propone el posible uso de la metodología de censo en sondeos electorales.

2.6.1. Estudio social de opinión de usuarios de bibliotecas

La “*Encuesta de satisfacción de usuarios*” realizada con objeto de evaluar la Biblioteca de la Universidad de Valladolid dentro del II Plan de Calidad de Universidades (Tapia y otros 2004 [135]) es la primera aplicación real llevada a cabo de la metodología de censo de usuarios en UDTs muestrales. La primera fase de la encuesta se realiza en cuatro de las bibliotecas de la UVa (Tapia y Sacristán 2003 [136]) y se usa un diseño muestral donde las aulas de las Facultades o Escuelas se toman como conglomerados de estudiantes. Esta primera fase se utiliza como estudio piloto previo al estudio social completo que se realiza en el año 2004. En este caso, como se recoge en la ficha técnica del informe (Tapia y otros 2004 [135]), se encuestan, durante la segunda quincena del mes de mayo, a 3.909 usuarios distribuidos entre las catorce bibliotecas de la UVa, mediante un cuestionario completado por los propios individuos. Para llevar a cabo el trabajo de campo se cuenta con dos encuestadores y la ayuda del personal de las bibliotecas. Las características del diseño muestral son:

- La población objetivo son los usuarios distintos de las 14 bibliotecas.
- Cada biblioteca se considera un estrato y todas las bibliotecas se encuestan independientemente.
- La unidad de tiempo (UDT) es la hora.
- Se toma como marco muestral el horario de apertura de la biblioteca durante dos semanas; todas las bibliotecas tienen el mismo horario.
- Se considera cada una de las horas en que la biblioteca presta servicio durante una semana de lunes a sábado un conglomerado de usuarios de distinto tamaño.
- El número de horas muestrales de cada biblioteca se determina fijando un error de estimación menor o igual que el 3%, para estimar una proporción, con muestreo aleatorio simple sin reemplazamiento, con una confianza del 95%.
- Las horas de encuesta se toman con muestreo sistemático.
- En la primera hora seleccionada se encuesta a todos los individuos que están en la biblioteca y, hasta que termina la hora, a los que entran en el recinto.
- En el resto de las horas, se solicita que rellenen el cuestionario los individuos que entran o salen y no la han realizado anteriormente.

- La falta de respuesta es despreciable.
- Terminado el estudio, el tamaño de la población objetivo, el número de usos de las bibliotecas en las semanas de encuestación, se considera desconocido.

La metodología de este ejemplo corresponde a una población objetivo de usuarios con censo de usuarios en cada UDT muestral.

La Figura 2.12 contiene el marco y la muestra de UDTs de cada biblioteca. Al ser llevado el trabajo de campo, en su mayor medida, por el propio personal de la biblioteca, se utilizó la misma muestra en todas las bibliotecas.

| Marco y Muestra | Biblioteca |
|-----------------|---------------|
| Lunes | 9:00 - 10:00 |
| | 10:00 - 11:00 |
| | 11:00 - 12:00 |
| | 12:00 - 13:00 |
| | 13:00 - 14:00 |
| | 14:00 - 15:00 |
| | 15:00 - 16:00 |
| | 16:00 - 17:00 |
| | 17:00 - 18:00 |
| | 18:00 - 19:00 |
| | 19:00 - 20:00 |
| | 20:00 - 21:00 |
| | 21:00 - 22:00 |
| | 22:00 - 23:00 |
| Martes | 9:00 - 10:00 |
| | 10:00 - 11:00 |
| | 11:00 - 12:00 |
| | 12:00 - 13:00 |
| | 13:00 - 14:00 |
| | 14:00 - 15:00 |
| | 15:00 - 16:00 |
| | 16:00 - 17:00 |
| | 17:00 - 18:00 |
| | 18:00 - 19:00 |
| | 19:00 - 20:00 |
| | 20:00 - 21:00 |
| | 21:00 - 22:00 |
| | 22:00 - 23:00 |
| Miércoles | 9:00 - 10:00 |
| | 10:00 - 11:00 |
| | 11:00 - 12:00 |
| | 12:00 - 13:00 |
| | 13:00 - 14:00 |
| | 14:00 - 15:00 |
| | 15:00 - 16:00 |
| | 16:00 - 17:00 |
| | 17:00 - 18:00 |
| | 18:00 - 19:00 |
| | 19:00 - 20:00 |
| | 20:00 - 21:00 |
| | 21:00 - 22:00 |
| | 22:00 - 23:00 |
| Jueves | 9:00 - 10:00 |
| | 10:00 - 11:00 |
| | 11:00 - 12:00 |
| | 12:00 - 13:00 |
| | 13:00 - 14:00 |
| | 14:00 - 15:00 |
| | 15:00 - 16:00 |
| | 16:00 - 17:00 |
| | 17:00 - 18:00 |
| | 18:00 - 19:00 |
| | 19:00 - 20:00 |
| | 20:00 - 21:00 |
| | 21:00 - 22:00 |
| | 22:00 - 23:00 |
| Viernes | 9:00 - 10:00 |
| | 10:00 - 11:00 |
| | 11:00 - 12:00 |
| | 12:00 - 13:00 |
| | 13:00 - 14:00 |
| | 14:00 - 15:00 |
| | 15:00 - 16:00 |
| | 16:00 - 17:00 |
| | 17:00 - 18:00 |
| | 18:00 - 19:00 |
| | 19:00 - 20:00 |
| | 20:00 - 21:00 |
| | 21:00 - 22:00 |
| | 22:00 - 23:00 |

Figura 2.12: Marco y muestra de unidades de tiempo en una biblioteca del Campus de la UVa

2.6.2. Encuesta a usuarios de servicios públicos de la administración

La segunda aplicación es la “*Encuesta sobre la satisfacción de los ciudadanos con los servicios prestados por la Administración de la Comunidad de Castilla y León*” (DGE 2009 [47]) en la que Rodríguez J.M. y Tapia J.A. fueron los responsables de la ejecución de todas las fases del estudio. Con respecto al ámbito de aplicación, la encuesta se realiza a 3.798 usuarios, mediante preguntas realizadas por un encuestador, durante cuatro semanas, siendo la población objetivo todos los individuos que en ese mes hacen uso de alguno de los 135 servicios distribuidos entre oficinas de información y atención al ciudadano, puntos de información y atención al ciudadano, oficinas departamentales y servicios territoriales de las nueve provincias de la Comunidad de Castilla y León. Las características del diseño muestral son:

- El marco espacio-tiempo está conformado por todos los espacios o centros donde se encuentran los 135 servicios objetivo y las horas de atención de estos servicios entre el 23 de Noviembre y 23 de Diciembre de 2008.
- El diseño muestral es estratificado, polietápico y con tamaño de población desconocido. La estratificación se realiza a nivel de provincia y de tipo de servicio prestado.
- Se toma censo de espacios y de servicios. Cada servicio se localiza por su mostrador de atención y persona o personas que prestan el servicio.
- La unidad de tiempo (UDT) es la hora de atención.
- La muestra en cada uno de los servicios se toma polietápica siendo la hora de atención la unidad de primera etapa y los individuos que utilizan el servicio las unidades de segunda etapa.
- El número de horas muestrales de cada servicio, 11, se determina fijando un error de estimación para estimar una proporción menor o igual al 3% con una confianza del 95%. La muestra de UDTs de cada servicio se toma mediante muestro sistemático de lunes a viernes durante un mes. Las muestras sistemáticas de los servicios ubicados en el mismo espacio se toman de tal manera que un mismo equipo de encuestación pueda realizar las entrevistas de todos los servicios.
- En cada una de las horas muestrales se mide, en primer lugar, el aforo, número de individuos que realiza en esa hora una gestión en ese servicio de la Administración autonómica. Por otro lado, se intenta realizar la encuesta a todos los individuos que conforman el aforo y, para evitar la falta de respuesta, cuando el número de encuestas no coincide con el aforo observado en esa hora muestral, se llevan a cabo nuevas encuestas a individuos que realizan alguna gestión en el mismo centro, en la misma hora y en el mismo día de la semana, pero de otra semana hasta alcanzar el aforo observado el primer día. Una vez subsanada la falta de respuesta las estimaciones se realizan suponiendo que, desde un principio, contestan todos los usuarios.

Este ejemplo corresponde a una población objetivo de usos, con censo de usos en cada hora de atención muestral.

La Figura 2.13 muestra el marco y las UDTs muestrales de 14 servicios, de los 135 servicios de la población, localizados en cuatro espacios diferentes en la provincia de Zamora.

| | | C/ Leopoldo Alas Clarín 4 C.P. 49016 Zamora | | | | | Avda. de Requejo, 4 | | | | C/ Prado Tuelto, 17 | | | | Plaza de Alemania, 1 |
|-----------|---------------|---|--|------------------------|---|---------------------------------|------------------------|--------------------------------|-----------------|---|---|-----------------------------------|---------------------------------------|--|------------------------------------|
| | | Delegación Territorial | Oficina de Información y Atención al Ciudadano | S.T. de Medio Ambiente | Servicio Territorial de Economía y Hacienda | Servicio Territorial de Fomento | Delegación Territorial | Oficina Territorial de Trabajo | S.T. de Cultura | Departamento Territorial de Familia e Igualdad de Oportunidades | Servicio Territorial de Agricultura y Ganadería | Dirección Provincial de Educación | S.T. de Industria, Comercio y Turismo | Gerencia Territorial de Servicios Sociales | S.T. de Sanidad y Bienestar Social |
| Lunes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Martes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Miércoles | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Jueves | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Viernes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Lunes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Martes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Miércoles | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Jueves | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Viernes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Lunes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Martes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Miércoles | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Jueves | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |
| Viernes | 9:00 - 10:00 | | | | | | | | | | | | | | |
| | 10:00 - 11:00 | | | | | | | | | | | | | | |
| | 11:00 - 12:00 | | | | | | | | | | | | | | |
| | 12:00 - 13:00 | | | | | | | | | | | | | | |
| | 13:00 - 14:00 | | | | | | | | | | | | | | |

Figura 2.13: Marco y unidades de tiempo muestrales de 14 servicios localizados en cuatro espacios diferentes en la provincia de Zamora

2.6.3. Sondeo en día electoral

Un sondeo en día electoral se entiende como una encuesta donde a los encuestados se les pregunta sobre el voto que han emitido y, lo que se pretende, es realizar un pronóstico de los resultados antes de que los verdaderos resultados sean conocidos. A continuación detallamos una nueva metodología, basado en lo expuesto en este capítulo, para llevar a cabo este tipo de sondeos.

DISEÑO MUESTRAL

Los colegios electorales son los espacios donde se van a realizar las encuestas. Para la selección de los colegios consideramos una estratificación de las secciones censales a la que pertenecen; la base de tal estratificación será geográfica (CCAA, provincia), ámbito rural o urbano, y utilización de variables de homogenización de las secciones (tamaño de población, tipo de población,...). La elección de secciones censales, en cada estrato, puede realizarse con muestreo proporcional a su número de votantes o con muestreo aleatorio simple sin reemplazamiento o con muestreo sistemático si tenemos las secciones en un listado; de igual modo puede tomarse la selección de colegios censales, en cada sección censal muestral.

En cada colegio electoral, la unidad de primera etapa que proponemos es ½ hora (UDT); se podría pensar en otra unidad de tiempo, por ejemplo de 20 minutos, pero se debe tener en cuenta que el encuestador en ese tiempo tiene que tener capacidad de desplazarse entre los colegios electorales muestrales y, por otro lado, la UDT debe ser suficientemente amplia como para poder recoger la opinión de un número de votantes significativo. Para determinar el tamaño muestral de UDTs, en cada colegio, se considera que la población son las UDTs y que se va a tomar una muestra aleatoria simple de ellas para estimar una proporción. En España, el tiempo de apertura de los colegios electorales es de 9:00h-20:00h, por tanto hay 22 UDTs y una muestra de 5 UDTs puede ser razonable. Las UDTs se seleccionan con muestreo sistemático, tomando una UDT de diferencia entre colegios encuestados por el mismo equipo de encuestadores; un equipo puede encuestar dos colegios. La Figura 2.14 contiene un marco y las UDTs muestrales para dos colegios electorales encuestados por el mismo equipo de encuestadores.

| Marco y Muestra | Colegio Electoral 1 | Colegio Electoral 2 |
|-----------------|---------------------|---------------------|
| Día electoral | 9:00-9:30 | |
| | 9:30-10:00 | |
| | 10:00-10:30 | |
| | 10:30-11:00 | |
| | 11:00-11:30 | |
| | 11:30-12:00 | |
| | 12:00-12:30 | |
| | 12:30-13:00 | |
| | 13:00-13:30 | |
| | 13:30-14:00 | |
| | 14:00-14:30 | |
| | 14:30-15:00 | |
| | 15:00-15:30 | |
| | 15:30-16:00 | |
| | 16:00-16:30 | |
| | 16:30-17:00 | |
| | 17:00-17:30 | |
| | 17:30-18:00 | |
| | 18:00-18:30 | |
| | 18:30-19:00 | |
| | 19:00-19:30 | |
| | 19:30-20:00 | |

Figura 2.14: Marco y UDTs muestrales para dos colegios electorales encuestados por el mismo equipo de encuestadores

Respecto a la selección de votantes, Kalton (1991) [79] propone tomarlos con muestreo sistemático cuando los individuos salen del colegio electoral, escogiendo una amplitud k de intervalo para producir un muestreo de individuos aproximadamente probabilístico en cada estrato; este procedimiento tiene diversos problemas como son el número de encuestadores necesario y la selección de la amplitud k . Frente a esta opción, nosotros proponemos la alternativa censo de individuos, es decir, encuestar a todos los individuos que han votado durante la UDT muestral. En este caso la población de usos y usuarios coinciden.

TRABAJO DE CAMPO

El entrevistador debe permanecer tan próximo a la salida del colegio como las leyes electorales lo permitan. La respuesta se recoge una vez que el votante ha emitido el voto y sale del colegio durante el tiempo de la UDT muestral. La forma de recoger la respuesta debe ser con implementación online (tablet, ordenador portátil, smartphone,...) con conexión a internet siendo, a ser posible, el mismo encuestado el que pulse la opción correspondiente. Esta implementación, por ejemplo con Google-Docs, creará automáticamente una base de datos donde se recoge la información del votante así como el número de votantes en cada UDT muestral.

ANÁLISIS ESTADÍSTICO DE LA BASE DE DATOS

La estimación del número total de votantes en un colegio electoral se obtiene utilizando el estimador (2.3) y la estimación del voto con el estimador (2.11). La estimación de los parámetros número de votantes e intención de voto se irá elevando a los distintos niveles de estratificación, sección censal, estrato de secciones, municipio, provincia, CCAA y nacional. El número total de votos, N , puede ser conocido nada más cerrar los colegios electorales y el peso obtenido como la razón entre el número de votos estimado y el verdadero número de votos, \hat{N}/N , podría ser utilizado para calibrar los resultados.

2.7. Metodología alternativa de selección de la muestra de unidades de tiempo

El conocimiento, previo al trabajo de campo, exacto o estimado de la población de clientes en el período de encuestación proporciona una opción alternativa de cálculo de tamaño de muestra. En este caso, el tamaño de la muestra de usos se calcula como si se fuera a tomar una muestra aleatoria simple de ellos y, atendiendo a la capacidad de encuestación y a la afluencia de individuos de cada UDT, determinaremos el número de UDTs muestral. La aplicación de esta técnica es, sin duda, más adecuada cuando todas las UDTs son saturadas dado que eso permite conocer cuántos clientes utilizan el servicio en cada UDT y en todo el período determinado por el marco, al menos, de manera aproximada.

Como ejemplo de aplicación de esta metodología, consideremos la población de 2018 usos, y todas las UDTs saturadas, simulada en la sección 2.5.1. Supongamos que el tamaño de la población se conoce a priori. Si fijamos un error de estimación de $B = 0.06$ y una confianza del 95% para estimar una proporción en la población de usos, la fórmula (2.2) del cálculo de tamaño de muestra en muestreo aleatorio simple sin reemplazamiento nos determina una muestra de 236 individuos. Si suponemos que la duración de cada encuesta es de 6 minutos y que en cada UDT, de 60 minutos, se pueden realizar un máximo de 10 encuestas a usos, el tamaño de muestra de UDTs necesario para cubrir la muestra de individuos es de 24 UDTs, que se tomarán del marco poblacional de 160 UDTs.

2.8. Efecto en la estimación de los parámetros en una población de usuarios significativamente inferior a la población de usos

El objeto de este apartado es observar, mediante simulaciones, el efecto en las estimaciones de los distintos diseños muestrales propuestos si, en la población objetivo, el número de usos y usuarios es relativamente diferente, es decir, los clientes que repiten el uso del servicio, en el período determinado por el marco, es elevado.

PRIMERA SIMULACIÓN

En esta simulación consideramos todas las UDTs del marco de un servicio saturadas. La población se simula de la misma manera que la población simulada en la sección 2.5.1 añadiendo la siguiente modificación:

- Para determinar los usos que dejan de ser usuarios (clientes no encuestados con anterioridad) se utiliza, como función que determina la probabilidad de cada individuo de ser o no usuario en cada UDT del marco

$$\{p_t = 0.005(t - 1)\}_{t=1, \dots, 160}$$

es decir, en la UDT 160 la probabilidad de que un individuo sea usuario es del 20.5% y, por tanto, el 79.5% de los clientes ya ha sido encuestado con anterioridad.

La Tabla 2.11 recoge los valores de los parámetros de la población simulada

Tabla 2.11: Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Parámetros de la población simulada

| Usos | Usuarios | Satisfacción media de los usuarios |
|------|----------|------------------------------------|
| N | N_u | $\hat{\mu}_u$ |
| 1871 | 1125 | 7.0315 |

Resultados de la Simulación

Las Tablas 2.12 y 2.13 comparan la aproximación de las confianzas de los intervalos y de los errores de estimación de los estimadores del tamaño de la población de usuarios propuestos para el diseño por conglomerados (2.4.4.1) y polietápico (2.4.4.2), cuando el tamaño poblacional de usos es desconocido y conocido, respectivamente.

Tabla 2.12: Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación del total de usuarios con tamaño de población de usos, N , desconocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|----------------------|-------------|----------------------------|------------|------------------|
| Conglomerados | (ec. 2.9) | (ec. 2.10) | 99.8% | 286.62 |
| Polietápico | (ec. 2.15) | (ec. 2.16) | 98.5% | 326.36 |

Tabla 2.13: Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación del total de usuarios con tamaño de población de usos, N , conocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|----------------------|-------------|--------------------------------|------------|------------------|
| Conglomerados | (ec. 2.8) | Correction Factor Bootstrap | 100 % | 221.09 |
| Polietápico | (ec. 2.18) | Bootstrap en dos etapas | 100 % | 277.57 |

Las Tablas 2.14 y la 2.15 comparan los resultados de la simulación entre los diseños por conglomerados (2.4.4.1) y polietápico (2.4.4.2) para estimar la media poblacional de los usos y de los usuarios, respectivamente.

Tabla 2.14: Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación de la media poblacional de usos

| Diseño | $\hat{\mu}$ | $\widehat{Var}(\hat{\mu})$ | C_θ | \bar{B}_θ |
|----------------------|-------------|--|--------------|------------------|
| Conglomerados | (ec. 2.11) | (ec. 2.12) Correction Factor Bootstrap | 91 % 92 % | 0.4232 0.4398 |
| Polietápico | (ec. 2.20) | Bootstrap en dos etapas | 94.4 % | 0.4776 |

Tabla 2.15: Población de usuarios inferior a la de usos y marco con todas las UDTs saturadas: Estimación de la media poblacional de usuarios

| Diseño | $\hat{\mu}_u$ | $\widehat{Var}(\hat{\mu}_u)$ | C_θ | \bar{B}_θ |
|----------------------|---------------|--------------------------------|------------|------------------|
| Conglomerados | (ec. 2.13) | Correction Factor Bootstrap | 88.8 % | 0.4401 |
| Polietápico | (ec. 2.21) | Bootstrap en dos etapas | 100 % | 1.09 |

En conclusión, la calidad obtenida en las estimaciones obtenidas con los dos diseños que se comparan son muy similares en esta población simulada con grandes diferencias entre el número de usos y usuarios y todas las UDTs saturadas, excepto en la estimación de la satisfacción media de los usuarios donde se observa un error de estimación del parámetro inferior utilizando el diseño por conglomerados aunque la aproximación de la confianza de los intervalos es bastante menor, en este diseño, que la obtenida con el diseño polietápico.

SEGUNDA SIMULACIÓN

En esta simulación consideramos que en las UDTs del marco del servicio hay tanto UDTs saturadas como no saturadas. La población se simula de la misma manera que la población simulada en la sección 2.5.2 añadiendo la siguiente modificación:

- Para determinar los usos que dejan de ser usuarios (clientes no encuestados con anterioridad) se utiliza, como función que determina la probabilidad de cada individuo de ser o no usuario en cada UDT del marco, la siguiente:

$$\{p_t = 0.005(t - 1)\}_{t=1, \dots, 160}$$

es decir, en la UDT 160 la probabilidad de que un individuo sea usuario es del 20.5 % y, por tanto, el 79.5 % de los clientes ya ha sido encuestado con anterioridad.

La Tabla 2.16 recoge los valores de los parámetros de la población simulada

Tabla 2.16: Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Parámetros de la población simulada

| Usos | Usuarios | Satisfacción media de los usuarios |
|------|----------|------------------------------------|
| N | N_u | $\hat{\mu}_u$ |
| 1835 | 1091 | 7.5353 |

Resultados de la Simulación

Las Tablas 2.17 y 2.18 comparan la aproximación de las confianzas de los intervalos y de los errores de estimación de los estimadores del tamaño de la población de usuarios propuestos para el diseño por conglomerados (2.4.4.1) y polietápico con postestratificación (2.4.4.3), cuando el tamaño poblacional de usos es desconocido y conocido, respectivamente.

Tabla 2.17: Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población N desconocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|--------------------------|-------------|----------------------------|------------|------------------|
| Conglomerados | (ec. 2.9) | (ec. 2.10) | 96.8 % | 293.07 |
| Postestratificado | (ec. 2.24) | Bootstrap en dos etapas | 98.8 % | 334.41 |

Tabla 2.18: Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación del total de usuarios con tamaño de población N conocido

| Diseño | \hat{N}_u | $\widehat{Var}(\hat{N}_u)$ | C_θ | \bar{B}_θ |
|--------------------------|-------------|--------------------------------|------------|------------------|
| Conglomerados | (ec. 2.8) | Correction Factor Bootstrap | 96.8 % | 208.14 |
| Postestratificado | (ec. 2.25) | Bootstrap en dos etapas | 97.9 % | 224.83 |

Las Tablas 2.19 y 2.20 comparan los diseños por conglomerados (2.4.4.1) y polietápico con postestratificación (2.4.4.3), para estimar la media poblacional de usos y usuarios respectivamente.

Tabla 2.19: Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación de la media poblacional de usos

| Diseño | $\hat{\mu}$ | $\widehat{Var}(\hat{\mu})$ | C_θ | \bar{B}_θ |
|--------------------------|-------------|--|------------------|------------------|
| Conglomerados | (ec. 2.11) | (ec. 2.12) Correction Factor Bootstrap | 85.7 % 90.8 % | 0.2442 0.2511 |
| Postestratificado | (ec. 2.26) | Bootstrap en dos etapas | 89.9 % | 0.3090 |

Tabla 2.20: Población de usuarios inferior a la de usos y marco con UDTs saturadas y no saturadas: Estimación de la media poblacional de usuarios

| Diseño | $\hat{\mu}_u$ | $\widehat{Var}(\hat{\mu}_u)$ | C_θ | \bar{B}_θ |
|--------------------------|---------------|--------------------------------|------------|------------------|
| Conglomerados | (ec. 2.13) | Correction Factor Bootstrap | 95.3 % | 0.3056 |
| Postestratificado | (ec. 2.27) | Bootstrap en dos etapas | 100 % | 0.7472 |

Al igual que en la primera simulación, la calidad obtenida en las estimaciones obtenidas con los dos diseños que se comparan son muy similares en esta segunda población simulada con grandes diferencias entre el número de usos y usuarios y UDTs saturadas y no saturadas, excepto en la estimación de la satisfacción media de los usuarios donde se observa un error de estimación del parámetro inferior utilizando el diseño por conglomerados aunque la aproximación de la confianza de los intervalos es algo menor, en este diseño, que la obtenida con el diseño postestratificado.

Conclusiones

Este bloque de la tesis plantea un procedimiento, con diversas alternativas, que permiten de una manera económica y fiable realizar cuestionarios cara-a-cara a individuos que tienen una opinión muy reciente respecto a un servicio del que acaban de hacer uso. Las alternativas de metodologías probabilísticas propuestas son aplicables en todos los servicios con clientes y un horario perfectamente especificado, por ejemplo, servicios donde el cliente aguarde en cola para hacer uso de dicho servicio; esta situación se presenta en los servicios prestados por la Administración, bancos, aeropuertos, centros comerciales, consultas de especialistas médicos ... En particular, el diseño de censo de individuos, introducida en la sección 2.4.4.1, también es aplicable en servicios que se prestan de manera simultánea a un conjunto de individuos, por ejemplo bibliotecas, hospitales, colegios o centros de educación universitaria. Otra aplicación del censo, tratada teóricamente en la sección 2.6.3, es en encuestas en sondeos en día electoral; aunque no está contrastada en una situación real, su ejecución práctica parece perfectamente viable.

A continuación propondremos un protocolo para llevar a la práctica, en su totalidad, la metodología detallada en el capítulo 2:

1. Detectar periodicidades temporales de uso de los espacios para fijar el período de trabajo de campo.
2. Determinar la medida de la unidad de tiempo (UDT).
3. Construir el marco tiempo-espacio para todos los servicios de un mismo espacio.
4. Calcular el tamaño de muestra de UDTs de cada servicio.
5. Las UDTs del primer servicio se seleccionan del marco mediante diseño sistemático fraccional. Para el segundo servicio, ubicado en el mismo espacio, se considera la misma muestra de UDTs del primero con dos UDTs de diferencia. Si hay más servicios, ubicados en el mismo espacio, se reitera el procedimiento con la condición de que no coincidan las UDTs muestrales de distintos servicios.
6. Decidir si la población objetivo son los usos o los usuarios.
7. Además de la opinión del individuo, se debe tomar, en cada UDT muestral:
 - Total de usos.
 - Clasificación de los individuos de la muestra en usuarios y usuarios ya encuestados.
 - Tiempo de atención de cada individuo si se decide aplicar diseño polietápico en todas las UDTs o en las saturadas.
8. Para cada servicio de la muestra, analizar la disponibilidad de la siguiente información:
 - Total de usos terminado el período de encuestación.
 - Clasificación previa de las UDTs del marco en saturadas y no saturadas.

9. Elegir la metodología de selección de individuos de cada UDT muestral atendiendo al siguiente orden de alternativas:

Opción 1: Censo de individuos. Es la metodología más precisa. Para escoger esta opción, la encuesta debe ser autoaplicada por el encuestado o el equipo de encuestación debe ser lo suficientemente amplio como para poder realizar la encuesta a todos los clientes.

Opción 2:

Si, a priori, conocemos que todas las UDTs del marco son saturadas, la muestra de individuos proporcional al tiempo de atención en cada UDT muestral proporciona el control de trabajo de campo necesario para evitar la falta de respuesta.

Si, a priori, conocemos las UDTs del marco que son saturadas y las no saturadas, realizaremos censo de individuos en las UDTs muestrales no saturadas y muestra proporcional al tiempo de atención en las UDTs muestrales saturadas.

Con respecto a mejorar la precisión de las estimaciones en la población de usos, se puede plantear utilizar calibración si se dispone de una información auxiliar poblacional más amplia, como por ejemplo el número de gestores, de particulares y de otros tipos de clientes que hacen uso del servicio en el período total en que se realiza el trabajo de campo.

En conclusión, esta metodología tiene una eficiencia aceptable y un bajo coste económico, puesto que el trabajo de campo se puede realizar con entrevista personal y con un reducido número de encuestadores, sin perder el control del error de muestreo; el coste sólo depende del número de UDTs muestrales y del número de desplazamientos de los encuestadores. Además, la metodología expuesta permite realizar inferencias tanto en la población de usos como de usuarios y, si es posible conocer el tamaño de la población de usos, una vez concluido el trabajo de campo, se consigue una notable reducción en el error de estimación del tamaño de la población de usuarios.

Respecto a las limitaciones de la metodología expuesta, se debe prestar especial atención al cálculo del número de UDTs muestral propuesto mediante la utilización de las fórmulas habituales de estimación de una proporción considerando como población las UDTs. El tamaño obtenido de UDTs y, en consecuencia, de clientes, será, habitualmente, superior al que obtendríamos si realizáramos directamente el cálculo de tamaño de muestra en la población de clientes. Para poder realizar este cálculo se precisa el conocimiento, previo al trabajo de campo, del tamaño de la población de clientes; aunque es muy difícil que en la práctica se disponga de esta información en poblaciones móviles, en la sección 2.7 se hace una breve introducción a la forma de actuar en el cálculo del tamaño de muestra en este caso. Por tanto, la elección del número de UDTs para la muestra sistemática debe realizarse con la intención de tomar UDTs que representen convenientemente los distintos días y unidades de tiempo del marco decidido y, además, permitan ejecutar el trabajo de campo de tal forma que un único equipo de encuestación pueda realizar todas las encuestas en todos los servicios ubicados en un mismo espacio.

Parte II

Análisis de eficiencia con outputs estimados

Objetivos y Motivación

Si en el primer bloque de la tesis abordábamos la metodología muestral en unidades productivas que prestan un servicio en un horario con objeto de resolver las cuestiones de cómo y a cuantos clientes encuestar para disponer de su opinión, en este segundo bloque nos planteamos el análisis de la eficiencia técnica de un conjunto de unidades productivas homogéneas, que prestan un mismo servicio, utilizando la opinión de sus clientes (outputs) y una información, conocida, que afecta o influye en dicha opinión (inputs). Como ya hemos comentado, los responsables de los servicios públicos o privados destinados a una población de usuarios tienen un objetivo claro, lograr los mejores resultados con los recursos disponibles y, desde un punto de vista empresarial, es obvio el interés en lograr la optimización de los recursos en beneficio de los usuarios, lo cual se trasladará en una mejor opinión del cliente sobre los servicios. En esta línea, las encuestas para analizar la calidad de los servicios prestados, tipo SERVQUAL (Parasuramun y otros 1988 [108]), son cada vez más frecuentes y, a través de ellas, es relativamente sencillo obtener (estimar) uno o más índices de satisfacción. Estos índices ofrecen una medida de calidad de los servicios. Si a estos índices de satisfacción (outputs) añadimos una información determinista del servicio que pueda influir en la opinión-satisfacción de sus clientes (inputs), nos encontramos ante una situación donde es posible intentar obtener medidas de eficiencia que permitan analizar los servicios que, con recursos comparables, obtienen mejor opinión-satisfacción de sus clientes. La imposibilidad, en la gran mayoría de los casos, de disponer de la opinión de todos los usuarios conduce a la necesidad de estimarla lo que implica cometer un error en la evaluación de la eficiencia con respecto a usar un censo de opinión. Motivados por esta situación, el objetivo de este bloque de la tesis es desarrollar una metodología estadística que permita estimar con garantías la eficiencia técnica de un conjunto homogéneo de unidades productivas que prestan un servicio con usuarios tomando como output su opinión, estimada con una muestra probabilística de clientes.

Este bloque se organiza de la siguiente forma: en el primer capítulo se hace un breve repaso a los aspectos generales del DEA y del DEA en presencia de inputs y/o outputs con variaciones estocásticas, DEA estocástico; puesto que el objetivo es facilitar la comprensión de los resultados de los apartados de aportaciones contenidos en el segundo capítulo, no se hará una exposición exhaustiva de estos aspectos. En el segundo capítulo se define el índice de eficiencia poblacional: suponemos un conjunto de inputs deterministas que explican la respuesta u opinión de una población de individuos; esta respuesta se sintetiza en uno o más valores en cada unidad productiva (DMU), considerados como outputs poblacionales, y, utilizando la técnica de programación lineal DEA, se obtiene el índice de eficiencia poblacional. Este índice va a dar lugar a la introducción de la mayor parte de los resultados que forman parte de las aportaciones teórico-prácticas más importantes de la tesis: ante la situación de outputs estimados a partir de una muestra, nos planteamos la determinación del tamaño de muestra necesario para estimar el índice de eficiencia poblacional de manera puntual, por intervalo de confianza y por región de confianza con la calidad (error de estimación) deseada por el experimentador. Sólo consideraremos la situación de outputs estocásticos atendiendo a lograr el objetivo general de esta memoria, utilizar la opinión-satisfacción de una muestra de usuarios de un conjunto de unidades que prestan un servicio para intentar ordenar las unidades según un criterio de eficiencia.

Capítulo 3

Preliminares

3.1. Análisis envolvente de datos

La eficiencia y su evaluación

De un modo muy general, puede decirse que el objetivo de la mayoría de las empresas e instituciones públicas es la obtención de unos resultados a partir de unos recursos. En esta línea, se considera que una empresa o institución aumenta su eficiencia si consigue “obtener más [resultados] usando menos [recursos]”. Un concepto de asignación de recursos eficiente es el de optimalidad en el sentido de Pareto, es decir, cuando no es posible reasignar los recursos existentes de forma que algún output mejore sin que otro empeore o, de otro modo, cuando no es posible obtener más usando lo mismo o cuando no es posible obtener lo mismo usando menos.

Al hablar de eficiencia productiva se distingue entre la eficiencia técnica y la eficiencia precio. La eficiencia técnica mide la capacidad de una DMU para producir la mayor cantidad posible de bienes usando para ello la menor cantidad posible de recursos y admite dos orientaciones, input, producir una cantidad fija de bienes usando para ello de la menor cantidad posible de recursos, y output, producir la mayor cantidad posible de bienes usando para ello una cantidad fija de recursos. La eficiencia precio mide la capacidad de una DMU para producir bienes con un valor total máximo usando para ello recursos con el mínimo coste posible. La diferencia entre estas dos eficiencias es que la eficiencia técnica es un concepto que incide en los procesos productivos, centrándose en el análisis de las cantidades y no en los valores de esas cantidades. En la eficiencia precio con orientación output el valor de lo producido es el único output y en la eficiencia precio con orientación input el coste de los recursos usados es el único input. En cambio, en la eficiencia técnica puede haber varios inputs y varios outputs.

Un repaso histórico de la medición de la eficiencia nos conduce a Koopmans (1951) [62], que considera un vector input-output técnicamente eficiente si y sólo si se cumplen dos condiciones, la primera, fijados los inputs, el aumento de algún output sólo es posible disminuyendo algún otro output (eficiencia con orientación output), y, la segunda condición, fijados los outputs, la disminución de algún input sólo es posible aumentando algún otro input (eficiencia con orientación input). Si representamos por T el conjunto de todos los vectores (x,y) tales que con $x \in \mathbb{R}^m$ se puede producir $y \in \mathbb{R}^s$, podemos formalizar la idea anterior del modo siguiente. Un vector input-output (x,y) se considera técnicamente eficiente si se verifican estas dos propiedades:

- A. No existe $y' \in \mathbb{R}^s$ tal que $y < y'$ (todas las componente del vector y' son mayores o iguales que las correspondientes componente del vector y y, al menos una componente, es estrictamente mayor) y , además, $(x,y') \in T$; es decir, con los mismos inputs no puedo producir más outputs.
- B. No existe $x' \in \mathbb{R}^m$ tal que $x' < x$ y, además, $(x',y) \in T$; es decir, no puedo producir los mismos outputs con menos inputs.

Matemáticamente, las propiedades A. y B. son equivalentes a:

A. No existe $s^+ = (s_1^+, \dots, s_s^+) \in \mathbb{R}^s$ tal que $s^+ > 0$ y además $(x, y + s^+) \in T$; es decir,

$$\max \left\{ \sum_{r=1}^s s_r^+ / s^+ > 0, (x, y + s^+) \in T \right\} = 0$$

B. No existe $s^- = (s_1^-, \dots, s_m^-) \in \mathbb{R}^m$ tal que $s^- > 0$ y además $(x - s^-, y) \in T$; es decir,

$$\max \left\{ \sum_{j=1}^m s_j^- / s^- > 0, (x - s^-, y) \in T \right\} = 0$$

En el mismo año (1951) Debreu [45] propuso una medida radial de la eficiencia técnica que analiza la máxima reducción equiproporcional de todos los inputs compatible con el mismo nivel de outputs (eficiencia con orientación input) o la máxima ampliación equiproporcional de todos los outputs compatible con el mismo nivel de inputs (eficiencia con orientación output). Más concretamente, para un vector input-output $(x, y) \in T$ la medida de la eficiencia desde los inputs es el mínimo del conjunto $\{\lambda \in [0, 1] : (\lambda x, y) \in T\}$, mientras que, para ese mismo vector, la medida de la eficiencia desde los outputs es el inverso del máximo del conjunto $\{\mu \geq 1 : (x, \mu y) \in T\}$; como $\mu \geq 1$, se calcula el inverso para que la eficiencia resulte un número en el intervalo $[0, 1]$.

Farrell (1957) [55] establece las bases conceptuales a partir de las cuales se desarrollarían posteriormente los modelos que tienen como objetivo la medición de la eficiencia en la producción usando métodos de programación matemática. Farrell diferencia entre eficiencia técnica y eficiencia precio o asignativa, introduce el concepto de función de producción y señala los problemas de su estimación. Además, propone estimar la función de producción a partir de la observación de unidades reales de producción. Esto da lugar a la función de producción empírica o implícita y al concepto de eficiencia relativa.

Charnes-Cooper-Rhodes (1978) [24] son los grandes impulsores del DEA. Por medio de la Programación Lineal, desarrollan una técnica para modelar la eficiencia de Farrell (procedimientos computacionales para el cálculo de eficiencias), plantean un nuevo enfoque del DEA basado en ideas fraccionales, introducen el nombre de Análisis Envolvente de Datos (Data Envelopment Analysis) y amplían las ideas de Farrell para trabajar con múltiples inputs y múltiples outputs.

Färe-Grosskopf-Lovell (1994) [54] consideran que el principal atractivo de las medidas radiales es que son independientes de las unidades de medida. Estos autores señalan que con la definición de Debreu una unidad productiva puede ser eficiente y, sin embargo, no resultar eficiente con la definición de Koopmans (al tener holguras en alguno de sus recursos) y plantean la cuestión de obtener una medida con la que sólo se considere eficiente una unidad productiva cuando satisfaga la definición de Koopmans.

En general, la medición empírica de la eficiencia ha sido abordada desde dos enfoques, el paramétrico y el no paramétrico. En ambos se trata de buscar una relación funcional $y = f(x)$ que determine la frontera eficiente de la producción. El concepto de función de producción frontera se interpreta como el límite máximo de producción que sirve de referencia para el cálculo de la ineficiencia del resto de las unidades observadas. En el enfoque paramétrico, la función se supone conocida salvo por un número finito de parámetros estimables utilizando técnicas estadísticas. Las dos formas funcionales más habituales son la Cobb-Douglas y la Translog (Lovell 1993 [95]). Dentro de las fronteras paramétricas se distinguen entre las de carácter determinista y estocástico. Los deterministas atribuyen toda la desviación de la frontera a la ineficiencia técnica y la función de producción se define como $y = f(x) - u$ donde u es una perturbación aleatoria que mide la distancia de cada unidad a la frontera de producción, a la que se impone una forma particular de su distribución (Greene 1980 [61]). Las fronteras estocásticas consideran que las DMUs pueden verse afectadas

por otros factores, al margen de la propia ineficiencia del productor, y la función de producción pasa a tener la forma $y = f(x) + \varepsilon$; $\varepsilon = v - u$; el componente aleatorio v representa sucesos no controlables por la unidad productiva mientras que u representa una medida de la ineficiencia. La principal ventaja del modelo estocástico es que permite separar, para la muestra en su conjunto, el impacto de la ineficiencia técnica y de los fenómenos aleatorios sobre el valor de los residuos, mientras que su debilidad reside en que los resultados obtenidos están fuertemente condicionados por la forma funcional especificada supuesta, no puede obtenerse una estimación de la ineficiencia en que incurre cada unidad productiva de manera particularizada y en las dificultades de aplicar esta técnica en un contexto de más de un output (Martínez-Cabrera 2003 [99]).

En el enfoque no paramétrico, la frontera eficiente se estima mediante técnicas de programación lineal. En este caso, no resulta necesario especificar una forma funcional para la función de producción, basta exigir al conjunto de posibilidades de producción una serie de propiedades, como por ejemplo, libre disponibilidad de inputs y outputs, convexidad y rendimientos constantes o variables de escala.

Dentro del enfoque no paramétrico son posibles dos perspectivas. La primera, a la que denominaremos perspectiva envolvente, se apoya en las ideas y conceptos de Farrell. La segunda, a la que denominaremos perspectiva fraccional, se basa en las ideas expuestas en el trabajo de Charnes, Cooper and Rhodes en 1978 [24]. Si bien conceptualmente ambas son distintas, sin embargo desde el punto de vista operativo, ambas se apoyan en la programación lineal. De un modo simplificado se puede decir que los autores antes citados desarrollaron una técnica, denominada Análisis Envolvente de Datos (DEA), que, a través de la programación lineal, consiguen modelizar el método propuesto por Farrell para medir la eficiencia.

La principal técnica no paramétrica para medir la eficiencia es el DEA. Si bien tiene el inconveniente de ser una técnica determinística, por lo que la presencia de observaciones atípicas puede sesgar las medidas de eficiencia, tiene las ventajas de ser más flexible, de evitar errores de especificación al no ser necesario especificar ninguna forma funcional, y de ofrecer información particularizada para cada unidad productiva. El criterio empleado por el DEA para definir la frontera de producción empírica es la noción paretiana de eficiencia. Así, una DMU se considera eficiente si no existe ninguna otra DMU en la muestra que produzca más de alguno de los outputs sin producir menos de algún otro, y sin utilizar más de alguno de los recursos productivos, inputs; o bien, si no existe ninguna DMU que produzca los mismos outputs con menos de algún factor productivo y no más de los restantes. La medida de eficiencia obtenida es relativa, es decir, primero se identifican las DMUs con los mejores resultados, que son las que van a definir la frontera de posibilidades de producción, y a continuación se mide la eficiencia de las restantes DMUs evaluando su desviación con respecto a las DMUs que logran optimizar su rendimiento. La medición de la eficiencia en términos relativos explica la especial importancia que tiene disponer de una muestra de DMUs con un alto grado de homogeneidad, esto es, han de producir los mismos outputs a partir de los mismos inputs, y han de tener los mismos objetivos globales.

Otro tipo de información que suministra el DEA sobre cada DMU son el grupo de referencia para las unidades ineficientes y los objetivos de recursos y de resultados óptimos.

Perspectiva envolvente de la metodología DEA

En la perspectiva envolvente, se considera que, además de las DMUs reales observadas, existen otras DMUs virtuales que envuelven a las primeras, de ahí el nombre de Análisis Envolvente de Datos. Todas estas DMUs conforman T , el conjunto de posibilidades de producción. Las distintas formas en que las DMUs envuelven a las DMUs de la muestra dan lugar a los distintos modelos DEA.

Desde la perspectiva envolvente, la metodología DEA sigue tres pasos (Álvarez-Pinilla 2001 [110]):

Paso 1. Caracterización de la tecnología. Se trata de describir, por medio de un conjunto de supuestos, las propiedades de la tecnología de producción o funcionamiento del sistema. En general, la

tecnología no es conocida, ni observable, ni deducible. A partir de los supuestos tecnológicos admitidos y de los datos de actividad realmente observados, se delimita el conjunto de planes de producción que se consideran factibles o realizables (conjunto de posibilidades de producción). Tal conjunto debe contener o envolver al conjunto de planes realmente observados. De aquí el nombre DEA. Además, los distintos conjuntos de supuestos admitidos darán lugar a diferentes formas de envolver los datos observados.

Paso 2. Definición de los índices de eficiencia. Existen distintos índices que informan de diferentes aspectos de la eficiencia (y, por tanto, de la ineficiencia) de las unidades de producción consideradas.

Paso 3. Elaboración de procedimientos de cálculo. Se trata de elaborar un programa matemático capaz de calcular el índice definido en el paso 2 bajo las restricciones asociadas al paso 1. Es decir, este índice se obtiene por comparación con el resto de las unidades observadas, así como con aquellas otras unidades virtuales que se obtienen a partir de las primeras y de los supuestos admitidos en el paso 1.

En la caracterización de la tecnología, si x es un vector de inputs, y es el correspondiente vector de outputs y $T = \{(x,y); x \text{ puede producir } y\}$ es el conjunto de posibilidades de producción, alguno de los supuestos básicos son:

S1. Libre disponibilidad de inputs:

$$(x,y) \in T, \tilde{x} \geq x \implies (\tilde{x},y) \in T$$

S2. Libre disponibilidad de outputs:

$$(x,y) \in T, \tilde{y} \leq y \implies (x,\tilde{y}) \in T$$

S3. Convexidad:

$$(x,y), (x',y') \in T, \lambda \in [0,1] \implies \lambda(x,y) + (1-\lambda)(x',y') \in T$$

S4. Rendimientos constantes de escala:

$$(x,y) \in T, \alpha \geq 0 \implies \alpha(x,y) \in T$$

Los supuestos que se admiten dan lugar a distintos modelos DEA. En esta memoria sólo vamos a considerar los modelos CCR y BCC:

- Modelo CCR (Charnes-Cooper-Rhodes), supuestos S1, S2, S3 y S4.
- Modelo BCC (Banker- Charnes-Cooper), supuestos S1, S2 y S3.

En la definición de los índices de eficiencia consideraremos las orientaciones input y output:

- Índice de eficiencia con orientación input

$$\theta^* = \min \{ \theta \in (0,1] : (\theta x, y) \in T \}$$

- Índice de eficiencia con orientación output

$$\left(\frac{1}{\omega^*} \right) \text{ donde } \omega^* = \max \{ \omega \geq 1 : (x, \omega y) \in T \}$$

Los modelos DEA CCR y BCC pueden escribirse, en términos generales de forma fraccional, multiplicativa y envolvente. La eficiencia técnica DEA en forma fraccional, en el caso particular del modelo CCR, donde las medidas de eficiencia técnica de orientación input y output coinciden, se define como el cociente entre la suma ponderada de los outputs y la suma ponderada de los inputs. El conjunto óptimo de pesos que determinan la ponderación están sujetos a la restricción de que ninguna DMU puede tener una puntuación de eficiencia mayor que uno usando estos mismos pesos. Por tanto, la expresión del modelo fraccional CCR con orientación input es:

$$\begin{aligned} \max_{u_r, v_j} \quad & \delta_o = \frac{\sum_{r=1}^s u_r y_{ro}}{\sum_{j=1}^m v_j x_{jo}} \\ \text{s.t.} \quad & \\ & \frac{\sum_{r=1}^s u_r y_{ri}}{\sum_{j=1}^m v_j x_{ji}} \leq 1; \quad i = 1, \dots, n \\ & u_r, v_j \geq \varepsilon \quad \forall r, j \end{aligned} \quad (3.1)$$

donde

- Se consideran n DMUs, cada uno de los cuales utiliza los mismos inputs (en diferentes cantidades) para obtener los mismos outputs (en diferentes cantidades).
- x_{ji} ($x_{ji} \geq 0$) representa la cantidad de j th input ($j = 1, \dots, m$) consumido por la i th DMU ($i = 1, \dots, n$).
- x_{jo} representa la cantidad de j th input ($j = 1, \dots, m$) consumido por la DMU evaluada, DMU_o .
- y_{ri} ($y_{ri} \geq 0$) representa la cantidad observada de r th output ($r = 1, \dots, s$) producido por la i th DMU.
- y_{ro} representa la cantidad observada de r th output ($r = 1, \dots, s$) producido por la DMU evaluada, DMU_o .
- u_r ($r = 1, \dots, s$) y v_j ($j = 1, \dots, m$) representan los pesos; los pesos son diferentes entre las distintas DMUs.

La transformación lineal de Charnes and Cooper (1962) [25] del modelo fraccional da lugar al problema lineal equivalente, conocido como modelo en forma multiplicativa

$$\begin{aligned} \max \quad & w_o = \sum_{r=1}^s \mu_r y_{ro} \\ \text{s.t.} \quad & \\ & \sum_{j=1}^m \delta_j x_{jo} = 1 \quad (\text{restricción de normalización}) \\ & \sum_{r=1}^s \mu_r y_{ri} - \sum_{j=1}^m \delta_j x_{ji} \leq 0; \quad i = 1, \dots, n \\ & \mu_r, \delta_j \geq \varepsilon \end{aligned} \quad (3.2)$$

Para todo programa lineal original (programa primal) existe otro programa lineal asociado, programa dual, que puede ser utilizado para determinar la solución del programa primal. Esta dualidad determina la expresión del modelo CCR con orientación input en su forma envolvente:

$$\begin{aligned} \min_{\theta, \lambda} \quad & \theta \\ \text{s.t.} \quad & \\ & \sum_{i=1}^n \lambda_i x_{ji} \leq \theta x_{jo}; \quad j = 1, \dots, m \\ & \sum_{i=1}^n \lambda_i y_{ri} \geq y_{ro}; \quad r = 1, \dots, s \\ & \lambda_i \geq 0 \end{aligned} \quad (3.3)$$

donde $\lambda = (\lambda_1, \dots, \lambda_n)$ es el vector de pesos o intensidades y θ determina la puntuación de la eficiencia técnica de la DMU_o .

Si x_o representa el vector de inputs de la DMU evaluada, DMU_o , y si θ es su eficiencia técnica, entonces $(1 - \theta)x_o$ indica la cantidad en que debe reducirse radialmente (proporcionalmente) todos los inputs de la DMU_o para convertirse en eficiente.

El modelo envolvente propuesto se resuelve en una única etapa, de forma que las variables de holgura s_r^+ (variable de holgura del rth output) y s_j^- (variable de holgura del jth input) se obtienen de forma residual, como la diferencia entre los valores input objetivo (coordenadas de la proyección sobre la frontera eficiente) y la reducción radial. Para determinar todas las posibles holguras input y output, Ali and Seiford (1993a) [2] sugieren una segunda etapa para maximizar la suma de las holguras input y output manteniendo el valor óptimo θ logrado en la primera etapa. El método para evaluar la eficiencia en el modelo CCR con orientación input en dos etapas es

$$\begin{aligned}
 & \min \theta - \varepsilon \left(\sum_{j=1}^m s_j^- + \sum_{r=1}^s s_r^+ \right) \\
 & \text{s.t.} \\
 & \sum_{i=1}^n \lambda_i x_{ji} + s_j^- = \theta x_{jo}; \quad j = 1, \dots, m \\
 & \sum_{i=1}^n \lambda_i y_{ri} - s_r^+ = y_{ro}; \quad r = 1, \dots, s \\
 & \lambda_i \geq 0, s_j^- \geq 0, s_r^+ \geq 0; \quad i = 1, \dots, n; \quad r = 1, \dots, s; \quad j = 1, \dots, m
 \end{aligned} \tag{3.4}$$

Como ya se ha comentado, la relajación de la suposición de rendimientos constantes de escala da lugar al modelo BCC. La Tabla 3.1 sintetiza el planteamiento radial de cálculo de los índices de eficiencia para los modelos CCR y BCC para las orientaciones input y output donde ε es un número pequeño no Arquimediano impuesto para evitar el ignorar algún factor en el cálculo de la eficiencia (Charnes and Cooper 1984 [26]). Una consecuencia del análisis de los modelos de eficiencia planteados es que un índice de eficiencia unitario es una condición necesaria, pero no suficiente, para alcanzar la eficiencia técnica de Pareto-Koopmans, pues además se debe exigir que las variables holgura sean nulas. Si el índice de eficiencia es uno, la DMU evaluada es eficiente, de acuerdo a la definición de Farrell, en relación con las otras DMUs. La mejora potencial de una DMU ineficiente se puede descomponer en mejora proporcional, derivada de la reducción radial, y mejora de holgura, derivada de la reducción de holgura/s input y/o aumento de holgura/s output.

Tabla 3.1: Modelos DEA CCR y BCC. Orientaciones input y output

| Modelo | Orientación | |
|------------|---|---|
| | Input | Output |
| | $ \begin{aligned} & \min \theta - \varepsilon \left(\sum_{j=1}^m s_j^- + \sum_{r=1}^s s_r^+ \right) \\ & \text{s.t.} \\ & \sum_{i=1}^n \lambda_i x_{ji} + s_j^- = \theta x_{jo}; \quad j = 1, \dots, m \\ & \sum_{i=1}^n \lambda_i y_{ri} - s_r^+ = y_{ro}; \quad r = 1, \dots, s \end{aligned} \tag{3.5} $ | $ \begin{aligned} & \max \varphi + \varepsilon \left(\sum_{j=1}^m s_j^- + \sum_{r=1}^s s_r^+ \right) \\ & \text{s.t.} \\ & \sum_{i=1}^n \lambda_i x_{ji} + s_j^- = x_{jo}; \quad j = 1, \dots, m \\ & \sum_{i=1}^n \lambda_i y_{ri} - s_r^+ = \varphi y_{ro}; \quad r = 1, \dots, s \end{aligned} \tag{3.6} $ |
| CCR | $\lambda_i \geq 0, s_j^- \geq 0, s_r^+ \geq 0; \quad i = 1, \dots, n; \quad r = 1, \dots, s; \quad j = 1, \dots, m$ | |
| BCC | Añadimos restricción $\sum_{i=1}^n \lambda_i = 1$ | |

Algunos autores, por ejemplo Coelli (1998) [32] y Tone (1999) [139], proponen un proceso de resolución DEA de más de dos etapas (multietápico) con el fin de abordar dos problemas del modelo en dos etapas:

- I. El modelo en dos etapas identifica el punto eficiente más alejado y no el más cercano, puesto que las holguras son maximizadas en lugar de minimizadas.
- II. El punto proyectado obtenido no es invariante a las unidades de medida. El componente radial de la medida de eficiencia obtenida en la primera etapa es invariante a las unidades de medida, sin embargo, la componente holgura de la segunda etapa no lo es (Lovell and Pastor 1995 [96]).

Introducción a los tópicos en la elección y ejecución de los modelos DEA básicos

En este apartado realizamos un breve resumen de los principales aspectos a tener en consideración a la hora de escoger el modelo DEA y proceder a su ejecución.

FORMA DEL CONJUNTO DE POSIBILIDADES DE PRODUCCIÓN

Si suponemos que un aumento de manera proporcional de los inputs origina un aumento en la misma proporción de los outputs (rendimientos constantes de escala) el modelo a utilizar es el CCR, mientras que, ante rendimientos variables de escala, el modelo escogido será el BCC. La introducción a los modelos de análisis de eficiencia alternativos, caso del Aditivo, del basado en medidas de holgura (SBM) o modelos de medida de Russell, no se abordan en esta sección de preliminares al no ser utilizados en el apartado considerado como de aportaciones de esta tesis.

ELECCIÓN DE LA ORIENTACIÓN: INPUT O OUTPUT

Uno de los principales propósitos de un estudio DEA es proyectar las DMUs ineficientes en la frontera producción. Para los modelos BCC y CCR hay dos direcciones, la denominada orientación-input que trata de reducir la cantidad de input todo lo posible manteniendo el nivel de outputs observado, y la orientación-output que maximiza el nivel de outputs para la cantidad de inputs observada. La determinación de una u otra orientación determina el tipo de datos utilizable. Si consideramos la orientación output, para el modelo CCR los valores de los inputs deben ser semipositivos (no negativos con al menos un valor positivo) mientras que los valores de los output pueden ser cualesquiera (negativos, ceros o positivos); en el modelo BCC, los valores de los input pueden ser cualesquiera mientras que los datos output deben ser semipositivos (Cooper, Seiford and Tone, 2002 [37]).

Ramanathan (2003) [112], al referirse a la elección de la orientación, indica que en aquellas aplicaciones en la que los factores productivos (inputs) no están completamente bajo el control del gestor, los modelos output serían los adecuados; en tanto que si los resultados del proceso (outputs) son decididos por los objetivos de los gestores antes que establecidos a partir de la mejor práctica observada sería preferible recurrir a la orientación input.

SELECCIÓN Y NÚMERO DE INPUTS Y OUTPUTS

La identificación de las variables input-output a utilizar en un análisis de evaluación de eficiencia es la primera e, indiscutiblemente, la más importante etapa en dicha evaluación. Los resultados que se obtengan dependen crucialmente de la elección realizada. Thanassoulis (2001) [138] recoge algunos aspectos a tener en cuenta en el proceso de selección de variables en un procedimiento DEA:

- Las variables se deben poder clasificar de manera exhaustiva y excluyente o en el grupo de variables input o en el grupo de variables output. La identificación de las variables exógenas es importante.
- Las variables input-output se escogerán de acuerdo al tipo de eficiencia que se desea evaluar. Por ejemplo, en centros comerciales se utilizarán distintos inputs para evaluar la eficiencia con respecto al volumen de ventas logrado o con respecto a la capacidad de atracción de clientes.
- El número de variables input-output deber ser tan pequeño como sea posible. El conjunto inicial de variables puede ser refinado utilizando procedimientos estadísticos, como por ejemplo

análisis de correlación, de varianza o regresión; en particular, si hay un único output (input), un análisis de regresión, previo al DEA, puede permitir identificar los factores que más probablemente sirven como input (output). Un análisis de sensibilidad también puede ayudar en la elección del conjunto de inputs-outputs; por ejemplo, para variables inputs u outputs consideradas como de un papel secundario, se puede analizar su impacto, en los resultados, ejecutando el análisis de eficiencia con y sin estas variables.

De manera general, si el número de DMUs (n) es menor que la suma de inputs y outputs ($m + s$), una elevada proporción de las DMUs será identificada como eficiente; si consideramos, por ejemplo, seis inputs y seis outputs, al menos 36 DMUs serán eficientes, por tanto el número total de DMUs debe ser mucho mayor que 36 para que el método DEA tenga algún valor discriminatorio (Boussofiane y otros 1991 [14]). Como regla general, para el modelo envolvente, se propone tomar (Cooper y otros 2006 [39])

$$n \geq \max\{m \times s, 3 \times (m + s)\}$$

INVARIANZA A TRASLACIONES

Por definición, un modelo DEA se dice que es invariante a traslaciones si trasladando los datos originales input y/o output a un nuevo problema, este tiene la misma solución envolvente óptima que el primero. El modelo BCC con orientación output (BCC-O) es invariante frente a traslaciones input pero no es invariante frente a traslaciones output. El modelo CCR con orientación output (CCR-O) no es invariante ni a traslaciones input ni a traslaciones output.

INVARIANZA A UNIDADES DE MEDIDA

El Teorema de Invarianza de Unidades (pp. 24 Cooper y otros 2006 [39]) garantiza que en los modelos CCR y BCC recogidos en la Tabla 3.1, los valores óptimos de maximización ($\varphi = \varphi^*$) y minimización ($\theta = \theta^*$) son independientes de las unidades en que los inputs y los outputs se miden, siempre que estas unidades sean las mismas para cada DMU.

3.2. Eficiencia DEA con datos imprecisos o estocásticos

La situación general, en DEA, supone datos con valores numéricos exactos. Sin embargo, también es posible el análisis de eficiencia DEA cuando los datos vienen dados dentro de unas cotas especificadas o cuando los datos se conocen en términos de relaciones ordinales o cuando los datos son estocásticos.

Cooper y otros (1999) [36] introducen el análisis envolvente de datos impreciso (IDEA), metodología que permite analizar la eficiencia en presencia tanto de datos imprecisos (intervalos de datos o datos categóricos ordinales) como de datos exactos. El modelo (3.7) muestra el problema no lineal (no convexo) correspondiente al análisis de eficiencia cuando los datos son imprecisos

$$\begin{aligned}
 \max_{\mu, \omega} \pi_o &= \sum_{r=1}^s \mu_r y_{ro} \\
 \text{s.t.} & \\
 \sum_{r=1}^s \mu_r y_{ri} - \sum_{j=1}^m \omega_j x_{ji} &\leq 0; \quad i = 1, \dots, n \\
 \sum_{j=1}^m \omega_j x_{jo} &= 1 \\
 x_{ji} &\in [x_{ji}^L, x_{ji}^U]; \quad j = 1, \dots, m \\
 y_{ri} &\in [y_{ri}^L, y_{ri}^U]; \quad r = 1, \dots, s \\
 \beta_r^- &\leq \frac{\mu_r}{\mu_{r+1}} \leq \beta_r^+ \\
 \alpha_j^- &\leq \frac{\omega_j}{\omega_{j+1}} \leq \alpha_j^+ \\
 \mu_r &\geq 0 \quad \forall r \\
 \omega_j &\geq 0 \quad \forall j
 \end{aligned} \tag{3.7}$$

El análisis envolvente de datos impreciso busca transformar el modelo (3.7) en un programa lineal equivalente. Para lograrlo se propone un procedimiento con dos etapas, en la primera, se realiza una

transformación de escala a los datos y a las variables. En la segunda etapa se realiza una modificación a las variables que conduce a la introducción de nuevas variables. La formalización de estas etapas es la siguiente

$$\text{Etapa 1. } \hat{y}_{ri} = \frac{y_{ri}}{\max_i y_{ri}^U}; \hat{x}_{ji} = \frac{x_{ji}}{\max_j x_{ji}^U}; \forall r, j, i$$

$$\text{Etapa 2. } Y_{ri} = \mu_r \hat{y}_{ri}; X_{ji} = \omega_j \hat{x}_{ji}; \forall r, j, i$$

Análisis envolvente de datos estocástico

Referencias de aplicaciones DEA como soluciones matemáticas, en presencia de datos con variaciones estocásticas, han sido realizadas por muchos autores, entre otros: Land et al. (1993) [90], Olesen and Petersen (1995) [107], Cooper y otros (1996 [34], 1998 [35], 2002 [37], 2004 [38]), Huang and Li (2001) [75], Bruni y otros (2009) [16], Khodabakhshi y otros (2010) [87], Khodabakhshi and Asgharian (2009) [86], Khodabakhshi (2009 [85] y 2010 [84], 2011 [83]), Wu and Lee (2010) [151] y Hoseinzadeh Lotfi y otros (2012) [74]. Teniendo en cuenta que el problema planteado, dentro de lo que consideramos las aportaciones de esta tesis, utiliza outputs con variaciones estocásticas, realizaremos una breve introducción al DEA estocástico para así poder ver las diferencias y similitudes con el planteamiento de solución a nuestro problema, basada en modelos DEA, que vamos a proponer.

El objetivo del DEA estocástico es incorporar al DEA datos inputs y/o outputs con variaciones estocásticas (Sengupta 1987 [124]). El modelo convencional DEA particiona el conjunto de DMUs en dos subconjuntos, las DMUs eficientes y las ineficientes; en presencia de variaciones estocásticas la partición de la eficiencia también es estocástica (Morita and Seiford 1999 [103]) y lo que se pretende es tratar de responder a las siguientes cuestiones: cómo de fiables son los resultados de eficiencia frente a las variaciones estocásticas? con qué probabilidad las variaciones estocásticas garantizan la eficiencia? qué input o output es más sensible a las variaciones estocásticas?. Una solución a las cuestiones planteadas consiste en modificar los modelos matemáticos para el DEA determinista planteando las restricciones en términos probabilísticos y asumiendo determinadas distribuciones de probabilidad para los inputs y los outputs. Esta técnica se denomina chance-constrained DEA (Charnes and Cooper 1959 [21] y 1963 [22], Land y otros 1993 [90], Olesen and Petersen 1995 [107], Cooper y otros 1996 [34]). La caracterización determinista “eficiente” es reemplazada por la caracterización “probablemente eficiente”. Los modelos más comunes de tipo chance-constrained son los E-modelos que obtienen una aproximación al “valor esperado” y los P-modelos que obtienen el suceso “más probable”, en el sentido de determinar si esta probabilidad es suficientemente grande. Ambos modelos incorporan restricciones en términos de probabilidad y restringen las probabilidades a que no excedan un valor $1 - \alpha$, donde α es un riesgo de violación de la restricción asociada, especificado por el usuario. De cara a introducir de manera resumida estos modelos, consideremos la notación de Cooper y otros (1996 [34], 1998 [35]) donde $\tilde{x}_i = (\tilde{x}_{1i}, \dots, \tilde{x}_{mi})'$ y $\tilde{y}_i = (\tilde{y}_{1i}, \dots, \tilde{y}_{ri})'$ representan la conducta aleatoria de los vectores input y output de la i th DMU, $i = 1, \dots, n$. Para la i th DMU, sean $x_i = (x_{1i}, \dots, x_{mi})'$ y $y_i = (y_{1i}, \dots, y_{ri})'$ los valores esperados de los vectores input y output.

Land y otros (1993) [90] introducen de manera formal un “E-modelo” para el modelo CCR chance-constrained DEA, en la unidad DMU_o de la siguiente manera

$$\begin{aligned} & \min_{\theta, \lambda} \theta \\ & s.t. \\ & P\left(\sum_{i=1}^n \lambda_i \tilde{x}_{ji} \leq \theta \tilde{x}_{jo}\right) \geq 1 - \alpha; j = 1, \dots, m \\ & P\left(\sum_{i=1}^n \lambda_i \tilde{y}_{ri} \geq \tilde{y}_{ro}\right) \geq 1 - \alpha; r = 1, \dots, s \\ & \lambda_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad (3.8)$$

siendo $(1 - \alpha)$ un valor de 0 a 1, prefijado por el experimentador, la probabilidad con la que se deben

verificar las restricciones. Si el mínimo se alcanza en el valor $\theta=1$ y todas las holguras son nulas en la solución óptima, la DMU_o se define como estocásticamente eficiente.

Si consideramos la obtención de las holguras, se obtiene la siguiente formulación del E-modelo, para la orientación output (Cooper y otros 2011 [41])

Tabla 3.2: E-modelo chance-constrained DEA. Orientación output.

| | |
|---|---|
| $\max \varphi + \varepsilon \left(\sum_{j=1}^m s_j^- + \sum_{r=1}^s s_r^+ \right)$ | |
| $s.t.$ | |
| $P \left(\sum_{i=1}^n \lambda_i \tilde{y}_{ri} - \varphi \tilde{y}_{ro} \geq s_r^+ \right) \geq 1 - \alpha; r = 1, \dots, s$ | |
| $P \left(\sum_{i=1}^n \lambda_i \tilde{x}_{ji} + s_j^- \leq \tilde{x}_{jo} \right) \geq 1 - \alpha; j = 1, \dots, m$ | |
| CCR | $\lambda_i \geq 0, s_j^- \geq 0, s_r^+ \geq 0; i = 1, \dots, n; j = 1, \dots, m; r = 1, \dots, s$ |
| BCC | Añadimos la restricción $\sum_{j=1}^n \lambda_j = 1$ |

donde P significa probabilidad y el símbolo \sim se utiliza para identificar los inputs y outputs como variables aleatorias con distribución de probabilidad conjunta conocida. La eficiencia estocástica de una DMU se alcanza si y sólo si el máximo se alcanza en $\varphi = 1$ y todas las holguras son cero, $(s_r^+ = 0 \text{ y } s_j^- = 0 \forall r, j)$

Bajo la suposición de todos los inputs y outputs estadísticamente independientes y normalmente distribuidos con medias y varianzas conocidos, el E-modelo, con orientación output, para la DMU_o, puede escribirse de manera determinista equivalente como (Cooper y otros 1998 [35]):

Tabla 3.3: E-modelo chance-constrained DEA con inputs y outputs estadísticamente independientes y normalmente distribuidos. Orientación output.

| | |
|---|---|
| $\max \varphi + \varepsilon \left(\sum_{j=1}^m s_j^- + \sum_{r=1}^s s_r^+ \right)$ | |
| $s.t.$ | |
| $\sum_{i=1}^n \lambda_i y'_{ri} - s_r^+ = \varphi y'_{ro}; r = 1, \dots, s$ | |
| $\sum_{i=1}^n \lambda_i x_{ji} + s_j^- = x_{jo}, j = 1, \dots, m$ | |
| CCR | $\lambda_i \geq 0, s_j^- \geq 0, s_r^+ \geq 0; i = 1, \dots, n; j = 1, \dots, m; r = 1, \dots, s$ |
| BCC | Añadimos la restricción $\sum_{i=1}^n \lambda_i = 1$ |

donde

$$\begin{aligned} y'_{ro} &= y_{ro} - \sigma_{ro}^O \phi^{-1}(\alpha); r = 1, \dots, s \\ y'_{ri} &= y_{ri}, i \neq o, r = 1, \dots, s \\ x'_{jo} &= x_{jo} - \sigma_{jo}^I \phi^{-1}(\alpha); j = 1, \dots, m \\ x_{ji} &= x_{ji}, i \neq o, j = 1, \dots, m \\ \sigma_{jo}^I &= \sqrt{\text{var}(x_{jo})} \\ \sigma_{ro}^O &= \sqrt{\text{var}(y_{ro})} \end{aligned}$$

La diferencia con respecto a los modelos DEA habituales es la sustitución de los valores determinista inputs x_{ji} y outputs y_{ri} por los términos x'_{ji} y y'_{ri} , donde σ_{ro}^O y σ_{jo}^I son las desviaciones típicas del rth

output y el j th input de la DMU en consideración, DMU_o , siendo ϕ la función de distribución de la ley de probabilidad normal estándar.

Un segundo modelo chance-constrained es el “P-modelo” también denominado modelo “satisficing” DEA. El concepto “satisficing” se crea en la literatura de la psicología por H.A. Simon (1957) [34] como alternativa a la suposición de conducta de optimización ampliamente utilizada en economía. Una primera versión del P-modelo que permite adaptar la definición de DEA eficiente en el contexto de la programación chance-constrained es

$$\begin{aligned} & \text{máx } P \left(\frac{\sum_{r=1}^s u_r \tilde{y}_{ro}}{\sum_{j=1}^m v_j \tilde{x}_{jo}} \geq 1 \right) \\ & \text{s.t.} \\ & P \left(\frac{\sum_{r=1}^s u_r \tilde{y}_{ri}}{\sum_{j=1}^m v_j \tilde{x}_{ji}} \leq 1 \right) \geq 1 - \alpha_i; 0 \leq \alpha_i \leq 1, i = 1, \dots, n \\ & u_r, v_j \geq 0 \forall r, j \end{aligned} \quad (3.11)$$

donde P significa probabilidad y el símbolo \sim se utiliza para identificar los inputs y outputs como variables aleatorias con distribución de probabilidad conjunta conocida. Los valores α_i son predeterminados y representan el riesgo permitido de violar las restricciones a las que están asociados. Si la distribución de las variables inputs y outputs es continua, la restricción para la unidad evaluada se puede escribir como

$$P \left(\frac{\sum_{r=1}^s u_r^* \tilde{y}_{ro}}{\sum_{j=1}^m v_j^* \tilde{x}_{jo}} \leq 1 \right) = 1 - \alpha^* \geq 1 - \alpha_o \quad (3.12)$$

y la DMU_o se define como “estocásticamente eficiente” si $\alpha^* = \alpha_o$.

Cooper y otros (2006) [39] proponen una versión del P-modelo, considerada extensión del modelo de razón CCR:

$$\begin{aligned} & \text{máx } P \left(\frac{\sum_{r=1}^s u_r \tilde{y}_{ro}}{\sum_{j=1}^m v_j \tilde{x}_{jo}} \geq \beta_o \right) \\ & \text{s.t.} \\ & P \left(\frac{\sum_{r=1}^s u_r \tilde{y}_{ri}}{\sum_{j=1}^m v_j \tilde{x}_{ji}} \leq \beta_i \right) \geq 1 - \alpha_i; 0 \leq \alpha_i \leq 1; i = 1, \dots, n \\ & u_r, v_j \geq 0 \forall r, j \end{aligned} \quad (3.13)$$

La conexión de este modelo con el concepto “satisficing” surge de interpretar el valor β_o como un “nivel de aspiración” especificado como una razón de eficiencia que se tiene que alcanzar.

Land y otros (1993) [90] y Cooper y otros (1996) [34] restringen la atención de los modelos “satisficing” a valores input deterministas donde sólo los outputs son variables aleatorias con distribución continua. Esta situación corresponde a una situación donde el administrador de las DMUs escoge los inputs sin ser capaz de controlar completamente los outputs. La expresión del P-modelo, en este caso, es

$$\begin{aligned} & \text{máx } P \left(\frac{\sum_{r=1}^s u_r \tilde{y}_{ro}}{\sum_{j=1}^m v_j x_{jo}} \geq \beta_o \right) \\ & \text{s.t.} \\ & P \left(\frac{\sum_{r=1}^s u_r \tilde{y}_{ri}}{\sum_{j=1}^m v_j x_{ji}} \leq \beta_i \right) \geq 1 - \alpha_i; 0 \leq \alpha_i \leq 1; i = 1, \dots, n \\ & u_r, v_j \geq 0 \forall r, j \end{aligned} \quad (3.14)$$

y la versión determinista equivalente, derivable de la teoría de programación, suponiendo una distribución normal multivariante de los outputs, es el siguiente problema de programación

cuadrático, donde \bar{y}_i representa el vector de medias para $\bar{y}_i = (\bar{y}_{1i}, \dots, \bar{y}_{ri})$, $i = 1, \dots, n$

$$\begin{aligned}
 \delta_o &= \max \mu^T \bar{y}_o - \beta_o \gamma^T x_o \\
 \text{s.t.} \\
 \mu^T \Sigma_o \mu &\geq 1 \\
 \mu^T \bar{y}_i - \beta_i \gamma^T x_i - \phi^{-1}(\alpha_i) \zeta_i &\leq 0; \quad i = 1, \dots, n \\
 C_{\alpha_i} \left[\zeta_i^2 - \mu^T \Sigma_i \mu \right] &\geq 0; \quad i = 1, \dots, n \\
 \mu &\geq 0, \quad \gamma \geq 0, \quad \zeta \geq 0
 \end{aligned} \tag{3.15}$$

donde

$$\begin{aligned}
 \Sigma_i &= \text{Cov}(\bar{y}_{ji}, \bar{y}_{ki}) \\
 C_{\alpha_i} &= \begin{cases} 1 & \text{si } \alpha_i < 0.5 \\ 0 & \text{si } \alpha_i = 0.5 \\ -1 & \text{si } \alpha_i > 0.5 \end{cases} \\
 \phi &\text{ función de distribución de la ley de probabilidad normal estándar.}
 \end{aligned}$$

La eficiencia estocástica de la DMU_o se alcanza en el modelo (3.15) si y sólo si $\phi(\delta_o) = \alpha_o$ (Cooper y otros 1996 [34]).

Intervalos de eficiencia

En los modelos DEA, la eficiencia se considera el máximo relativo de la razón de la suma con pesos de outputs y de inputs. Algunas de las unidades de producción clasificadas como eficientes por el DEA son DMUs peculiares más que superiores. Ya que el DEA es la evaluación del modelo desde el punto de vista optimista, Entani y otros (2002) [52] proponen la evaluación del modelo desde el punto de vista pesimista y la obtención de la eficiencia como un intervalo donde los límites superior e inferior se calculan de la maximización y minimización de problemas con una función objetivo y unas restricciones, respectivamente. Este modelo se denomina intervalo DEA o de eficiencia y admite su extensión a datos inputs y/o outputs que son intervalos. Si consideramos x_i y y_i los vectores input y output de la i th DMU ($i = 1, \dots, n$), el extremo superior del intervalo de la DMU_o se calcula, basado en la idea de que el DEA es una evaluación relativa, como

$$\begin{aligned}
 \varphi_o^U &= \max_{u,v} \frac{\frac{u}{v} \frac{y_o}{x_o}}{\max_i \frac{u}{v} \frac{y_i}{x_i}} \\
 \text{s.t.} \\
 u &\geq 0 \\
 v &\geq 0
 \end{aligned} \tag{3.16}$$

Cuando el denominador de la función objetivo se fija en uno, el problema se reduce a

$$\begin{aligned}
 \varphi_o^U &= \max_{u,v} \frac{u}{v} \frac{y_o}{x_o} \\
 \text{s.t.} \\
 \max \frac{u}{v} \frac{y_i}{x_i} &= 1; \quad i = 1, \dots, n \\
 u &\geq 0 \\
 v &\geq 0
 \end{aligned} \tag{3.17}$$

Entani y otros (2002) [52] demuestran que los valores óptimos de la anterior formulación son

equivalentes a los del modelo DEA básico

$$\begin{aligned}
 \varphi_o^U &= \max_{u,v} \frac{u' y_o}{v' x_o} \\
 \text{s.t.} \\
 \frac{u' y_i}{v' x_i} &= 1; \quad i = 1, \dots, n \\
 u &\geq 0 \\
 v &\geq 0
 \end{aligned} \tag{3.18}$$

Como extremo inferior del intervalo de eficiencia de la DMU_o se toma

$$\begin{aligned}
 \varphi_o^L &= \min_{u,v} \frac{\frac{u' y_o}{v' x_o}}{\max_i \frac{u' y_i}{v' x_i}} \\
 \text{s.t.} \\
 \frac{u' y_i}{v' x_i} &= 1; \quad i = 1, \dots, n \\
 u &\geq 0 \\
 v &\geq 0
 \end{aligned} \tag{3.19}$$

Suponiendo $\frac{u' y_i}{v' x_i} = 1, \forall i$ el valor óptimo se obtiene de la resolución de los siguientes n problemas de programación lineal

$$\begin{aligned}
 \varphi_{oi}^{L*} &= \min_{u,v} u' y_o \\
 \text{s.t.} \\
 v' x_o &= 1 \\
 u' y_i - v' x_i &= 0 \\
 u &\geq 0 \\
 v &\geq 0 \\
 (i &= 1, \dots, n)
 \end{aligned} \tag{3.20}$$

y el límite inferior del intervalo de eficiencia es

$$\varphi_o^L = \min \{1, \min_{i \neq o} \varphi_{oi}^{L*}\} \tag{3.21}$$

Una manera equivalente de obtener el extremo inferior directamente es

$$\varphi_o^L = \min_{p,r} \frac{\frac{y_{jo}}{x_{ro}}}{\max_i \frac{y_{pi}}{x_{ri}}}; \quad j = 1, \dots, m; \quad r = 1, \dots, s \tag{3.22}$$

Una vez introducida la construcción del intervalo de eficiencia en datos deterministas, Entani y otros (2002) [52] plantean el problema de la obtención del intervalo de eficiencia en la presencia de datos que son intervalos

$$x_{ji} \in [x_{ji}^L, x_{ji}^U], \quad y_{ri} \in [y_{ri}^L, y_{ri}^U]; \quad i = 1, \dots, n; \quad j = 1, \dots, m; \quad r = 1, \dots, s$$

Si los extremos inferior y superior de los intervalos input y output de la i th DMU se denotan como

$$X_i^L = (x_{1i}^L, \dots, x_{mi}^L)', \quad X_i^U = (x_{1i}^U, \dots, x_{mi}^U)',$$

$$Y_i^L = (y_{1i}^L, \dots, y_{si}^L)', \quad Y_i^U = (y_{1i}^U, \dots, y_{si}^U)'$$

el extremo superior del intervalo de eficiencia para la DMU_o se obtiene de igual forma que para datos deterministas

$$\begin{aligned} \varphi_o^U &= \max_{u,v} \frac{\frac{u}{v} \frac{y_o}{x_o}}{\max_i \frac{u}{v} \frac{y_i}{x_i}} \\ s.t. \\ u &\geq 0 \\ v &\geq 0 \end{aligned} \quad (3.23)$$

donde $\mathbf{x}_i \in [X_i^L, X_i^U]$ y $\mathbf{y}_i \in [Y_i^L, Y_i^U]$.

El valor óptimo de la formulación (3.23) es equivalente al siguiente problema de programación lineal

$$\begin{aligned} \varphi_o^U &= \max_u u' Y_o^U \\ s.t. \\ v' X_o^L &= 1 \\ u' Y_i^L - v' X_i^U &\leq 0; (i \neq o) \\ u' Y_o^U - v' X_o^L &\leq 0 \\ u &\geq 0 \\ v &\geq 0 \end{aligned} \quad (3.24)$$

Respecto al extremo inferior del intervalo de eficiencia de la DMU_o, éste se obtiene del siguiente problema de minimización

$$\begin{aligned} \varphi_o^L &= \min_{u,v} \min_{\mathbf{x}_i, \mathbf{y}_i} \frac{\frac{u}{v} \frac{y_o}{x_o}}{\max_i \frac{u}{v} \frac{y_i}{x_i}} \\ s.t. \\ u &\geq 0 \\ v &\geq 0 \end{aligned} \quad (3.25)$$

cuyo valor óptimo es equivalente al obtenido en

$$\begin{aligned} \varphi_o^L &= \min_{u,v} \frac{u' Y_o^L}{v' X_o^L} \\ s.t. \\ \max \left(\frac{\max_{i \neq o} \frac{u' Y_i^U}{v' X_i^L}, u' Y_o^L}{v' X_o^U} \right) &= 1 \\ u &\geq 0 \\ v &\geq 0 \end{aligned} \quad (3.26)$$

Una manera equivalente de obtener el extremo inferior directamente es

$$\varphi_o^L = \min_{p,r} \frac{\frac{y_{op}^L}{x_{or}^U}}{\max_j \frac{y_{jp}^U}{x_{jr}^L}} \quad (3.27)$$

La Tabla 3.4 sintetiza el método optimista/pesimista de Entani y otros (2002) [52] de obtención del intervalo de eficiencia de la DMU_o en presencia de datos que son intervalos

Tabla 3.4: Método optimista/pesimista de Entani y otros (2002) [52] de obtención del intervalo de eficiencia

| Extremo inferior | Extremo Superior |
|---|---|
| $\varphi_o^L = \min_{p,r} \frac{\frac{y_{op}^L}{x_{or}^U}}{\max_j \frac{y_{jp}^U}{x_{jr}^L}}$ | $\begin{aligned} \varphi_o^U &= \max_u u' Y_o^U \\ s.t. \\ v' X_o^L &= 1 \\ u' Y_i^L - v' X_i^U &\leq 0; \quad (i \neq o) \\ u' Y_o^U - v' X_o^L &\leq 0 \\ u &\geq 0 \\ v &\geq 0 \end{aligned}$ |

El punto de vista optimista/pesimista de obtención de intervalos de eficiencia, en presencia de datos que son intervalos, admite otras variantes (Despotis and Smirlis 2002 [46], Wang y otros 2005 [146]). Ante la presencia de datos imprecisos, Despotis y Smirlis proponen, a diferencia del IDEA (Cooper y otros 1999 [36]), una transformación de las variables, realizada sobre el conjunto de datos original, sin aplicar ninguna transformación de escala en los datos. Para formular el modelo propuesto por estos autores, supongamos n DMUs que utilizan m inputs para producir s outputs tales que

$$x_{ji} \in [x_{ji}^L, x_{ji}^U], y_{ri} \in [y_{ri}^L, y_{ri}^U]; i = 1, \dots, n; j = 1, \dots, m; r = 1, \dots, s$$

siendo los extremos superior e inferior de todos los intervalos valores estrictamente positivos. En esta situación, el siguiente modelo CCR

$$\begin{aligned} \text{máx } \theta_o &= \sum_{r=1}^s u_r y_{ro} \\ s.t. \\ \sum_{j=1}^m v_j x_{jo} &= 1 \\ \sum_{r=1}^s u_r y_{ri} - \sum_{j=1}^m v_j x_{ji} &\leq 0 \\ u_r, v_j &\geq \varepsilon \quad \forall r, j \end{aligned} \tag{3.28}$$

es no lineal (no convexo), donde, además de las variables $\{u_r\}_{r=1, \dots, s}, \{v_j\}_{j=1, \dots, m}$, los niveles de inputs x_{ji} y de outputs y_{ri} también son variables cuyos valores exactos tienen que ser estimados.

El objetivo es transformar este modelo en uno equivalente lineal, para lo cual se aplica la siguiente transformación a las variables x_{ji} e y_{ri}

$$x_{ji} = x_{ji}^L + s_{ji} (x_{ji}^U - x_{ji}^L); \quad 0 \leq s_{ji} \leq 1$$

$$y_{ri} = y_{ri}^L + t_{ri} (y_{ri}^U - y_{ri}^L); \quad 0 \leq t_{ri} \leq 1;$$

Con esta transformación, las variables x_{ji} e y_{ri} del modelo (3.28) son reemplazadas por las nuevas

variables s_{ji} y t_{ri} obteniendo el siguiente programa lineal

$$\begin{aligned}
 \text{máx } \theta_o &= \sum_{r=1}^s u_r y_{ro}^L + p_{ro} (y_{ro}^U - y_{ro}^L) \\
 \text{s.t.} \\
 \sum_{r=1}^s u_r y_{ri}^L + p_{ri} (y_{ri}^U - y_{ri}^L) - \sum_{j=1}^m v_j x_{ji}^L + q_{ji} (x_{ji}^U - x_{ji}^L) &\leq 0 \\
 p_{ri} - u_r &\leq 0; \quad i = 1, \dots, n; r = 1, \dots, s \\
 q_{ji} - v_j &\leq 0; \quad i = 1, \dots, n; j = 1, \dots, m \\
 u_r, v_j &\geq \varepsilon \quad \forall r, j \\
 p_{ri} = u_r t_{ri} &\geq 0, q_{ji} = v_j s_{ji} \geq 0 \quad \forall r, j, i
 \end{aligned} \tag{3.29}$$

La eficiencia obtenida por la DMU_o con el modelo (3.29) se puede demostrar que coincide con la cota superior para la eficiencia calculada con el modelo

$$\begin{aligned}
 \text{max } \theta_o^U &= \sum_{r=1}^s u_r y_{ro}^U \\
 \text{s.t.} \\
 \sum_{j=1}^m v_j x_{jo}^L &= 1 \\
 \sum_{r=1}^s u_r y_{ro}^U - \sum_{j=1}^m v_j x_{jo}^L &\leq 0 \\
 \sum_{r=1}^s u_r y_{ri}^L - \sum_{j=1}^m v_j x_{ji}^U &\leq 0; \quad i \neq o \\
 u_r, v_j &\geq \varepsilon \quad \forall r, j
 \end{aligned} \tag{3.30}$$

El modelo (3.30) es un modelo DEA con datos exactos, donde los niveles de inputs y outputs se ajustan en favor de la unidad evaluada y lo más en contra para el resto de las DMUs. Despotis and Smirlis (2002) [46] ponen de manifiesto en su artículo diversos beneficios de usar el modelo (3.29) en lugar del modelo (3.30) y proponen el cálculo de un intervalo de eficiencia para la DMU_o utilizando, para determinar la cota superior, el modelo (3.30) y, para obtener la cota inferior, el modelo

$$\begin{aligned}
 \text{max } \theta_o^L &= \sum_{r=1}^s u_r y_{ro}^L \\
 \text{s.t.} \\
 \sum_{j=1}^m v_j x_{jo}^U &= 1 \\
 \sum_{r=1}^s u_r y_{ro}^L - \sum_{j=1}^m v_j x_{jo}^U &\leq 0 \\
 \sum_{r=1}^s u_r y_{ri}^U - \sum_{j=1}^m v_j x_{ji}^L &\leq 0; \quad i \neq o \\
 u_r, v_j &\geq \varepsilon \quad \forall r, j
 \end{aligned} \tag{3.31}$$

Wang y otros (2005) [146] tratan de atender diversos problemas detectados en la aplicación de los modelos de Entani y otros (2002) [52] y de Despotis and Smirlis (2002) [46]. El modelo de Entani y otros utiliza sólo datos de un input y un output para calcular la cota inferior de cada DMU sin tener en cuenta cuantos datos input y output participan en el modelo, lo cual conduce a una pérdida de información, y, además, los conjuntos de restricciones usados son diferentes de una DMU a otra. Una característica de los intervalos de eficiencia obtenidos con el modelo de eficiencia DEA de Entani y otros es el ser más amplios que los obtenidos por los otros modelos de intervalos de eficiencia DEA. En el modelo de Despotis y Smirlis también los conjuntos de restricciones usados son diferentes de una DMU a otra e, incluso, también son diferentes los conjuntos de restricciones para obtener los extremos inferior y superior de una misma DMU. Por ejemplo, el conjunto de restricciones utilizado para medir la cota de eficiencia superior (3.30) de la DMU_o se basa en los datos

$$\{(x_{jo}^L, y_{ro}^U), (x_{ji}^U, y_{ri}^L), i = 1, \dots, n; i \neq o; j = 1, \dots, m; r = 1, \dots, s\} \tag{3.32}$$

mientras que el conjunto de restricciones para la cota de eficiencia inferior (3.31) se basa en los datos

$$\{(x_{jo}^U, y_{ro}^L), (x_{ji}^L, y_{ri}^U), i = 1, \dots, n; i \neq o; j = 1, \dots, m; r = 1, \dots, s\} \tag{3.33}$$

La principal desventaja de utilizar diferentes conjuntos de restricciones es la falta de comparabilidad entre las eficiencias debido a las diferentes fronteras adoptadas en el proceso de medir las eficiencias.

El modelo de intervalo fraccional DEA propuesto por Wang y otros, basado en aritmética de intervalos, hace uso siempre del mismo conjunto de restricciones que conforman una única, y fija, frontera de producción y conducen a la determinación de los extremos superior e inferior de eficiencia para la unidad evaluada, DMU_o . Supongamos n unidades que utilizan m inputs para producir s outputs tales que

$$x_{ji} \in [x_{ji}^L, x_{ji}^U], y_{ri} \in [y_{ri}^L, y_{ri}^U]; x_{ji}^L > 0, y_{ri}^L > 0, i = 1, \dots, n; j = 1, \dots, m; r = 1, \dots, s$$

Los modelos de programación fraccional propuestos por Wang y otros para medir las cotas superior e inferior de eficiencia de la DMU_o son

$$\begin{aligned} \text{máx } \theta_o^U &= \frac{\sum_{r=1}^s u_r y_{ro}^U}{\sum_{j=1}^m v_j x_{jo}^L} \\ \text{s.t.} & \\ \theta_i^U &= \frac{\sum_{r=1}^s u_r y_{ri}^U}{\sum_{j=1}^m v_j x_{ji}^L} \leq 1; i = 1, \dots, n \\ u_r, v_j &\geq \varepsilon \quad \forall r, j \end{aligned} \quad (3.34)$$

$$\begin{aligned} \text{máx } \theta_o^L &= \frac{\sum_{r=1}^s u_r y_{ro}^L}{\sum_{j=1}^m v_j x_{jo}^U} \\ \text{s.t.} & \\ \theta_i^L &= \frac{\sum_{r=1}^s u_r y_{ri}^L}{\sum_{j=1}^m v_j x_{ji}^U} \leq 1; i = 1, \dots, n \\ u_r, v_j &\geq \varepsilon \quad \forall r, j \end{aligned} \quad (3.35)$$

Usando la transformación Charnes-Cooper, los modelos (3.34) y (3.35) de programación fraccional se pueden simplificar en los siguientes modelos equivalentes de programación lineal

$$\begin{aligned} \text{max } \theta_o^U &= \sum_{r=1}^s u_r y_{ro}^U \\ \text{s.t.} & \\ \sum_{j=1}^m v_j x_{jo}^L &= 1 \\ \sum_{r=1}^s u_r y_{ri}^U - \sum_{j=1}^m v_j x_{ji}^L &\leq 0; i = 1, \dots, n \\ u_r, v_i &\geq \varepsilon \quad \forall r, i \end{aligned} \quad (3.36)$$

$$\begin{aligned} \text{máx } \theta_o^L &= \sum_{r=1}^s u_r y_{ro}^L \\ \text{s.t.} & \\ \sum_{j=1}^m v_j x_{jo}^U &= 1 \\ \sum_{r=1}^s u_r y_{ri}^L - \sum_{j=1}^m v_j x_{ji}^U &\leq 0; i = 1, \dots, n \\ u_r, v_i &\geq \varepsilon \quad \forall r, i \end{aligned} \quad (3.37)$$

donde θ_o^U corresponde a la mejor eficiencia relativa posible de la DMU_o cuando todas las demás DMUs permanecen en la mejor actividad de producción, mientras que θ_o^L corresponde a la cota inferior de la mejor eficiencia relativa posible de la DMU_o . El mejor intervalo de eficiencia relativa lo constituye $[\theta_o^L, \theta_o^U]$. Un aspecto interesante, observable en el modelo (3.35), es que tanto la mejor como la peor actividad productiva posible de la DMU_o son tenidas en cuenta en el modelo, en concreto, la mejor actividad productiva juega el papel de pertenecer a la lista de referencia del modelo. Por tanto, se crea la confusión de que dos datos diferentes de la DMU_o son utilizados en el modelo, uno, corresponde al dato virtual DMU_o' obtenido al tomar el máximo de inputs de la DMU_o para obtener el mínimo de outputs posible. Este dato virtual se compara con los n datos de las DMUs, incluido el de la DMU_o , correspondientes a tomar la mejor actividad productiva de todas ellas (el mínimo posible de los inputs y el máximo posible de outputs), lo cual conduce, automáticamente, a que la eficiencia de la unidad virtual sea inferior a uno. Una de las características más relevantes de los modelos (3.36) y (3.37), con los que se obtienen las cotas superior e inferior del intervalo de eficiencia de Wang y otros, es que ambos tienen la misma frontera de producción correspondiente a

tomar la mejor actividad productiva de todas las DMUs. Si consideramos $[H_o^L, H_o^U]$ el intervalo de eficiencia obtenida por los modelos (3.30) y (3.31) de Despotis y Smirlis, y $[\theta_o^L, \theta_o^U]$ el intervalo de eficiencia obtenida por los modelos (3.36) y (3.37), Wang y otros demuestran dos teoremas que ponen de manifiesto las relaciones $\theta_o^L \leq H_o^L$ y $\theta_o^U \leq H_o^U$. Una vez obtenido el intervalo de eficiencia de la DMU_o, esta unidad se define como eficiente DEA si $\theta_o^U = 1$; si $\theta_o^U < 1$ la unidad se define ineficiente DEA. Por tanto, en función de la definición y de las relaciones demostradas entre intervalos de eficiencia, el modelo de Wang y otros reduce el número de unidades eficientes DEA en comparación con el modelo de Despotis y Smirlis.

3.3. Aplicaciones del DEA: Análisis de eficiencia en servicios prestados al público

La literatura con respecto a aplicaciones del análisis envolvente de datos es muy amplia. Desde 1978 el número de artículos, en el ámbito del DEA, supera las 5.000 referencias, según la base de datos de la ISI Web of Science (WOS). Algunos autores, como Liu y otros (2013) [93], han realizado estudios estadísticos utilizando esta base de datos con el fin de responder a cuestiones como las siguientes: cual es la proporción de artículos aplicados en la literatura DEA? exactamente, como se aplica el DEA a las realidades del mundo? cuáles son las mayores aplicaciones del DEA?. Del análisis estadístico se concluye que un 36.5% de los artículos DEA son exclusivamente metodológicos, sin incluir datos reales, mientras que el 63.5% introducen aplicaciones con datos reales. Respecto a éstas, hay 24 tipos de aplicaciones con al menos 20 artículos en la base de datos WOS. Un ranking, creado a partir del número de publicaciones, sitúa a las aplicaciones en banca en primer lugar seguido de temas sanitarios, agrícolas y ganaderos, transporte y educación. Estas cinco aplicaciones constituyen el 41% de los artículos DEA aplicados.

Los servicios públicos o privados en los que hay presencia de usuarios, siempre han recogido datos relativos a sus características (número de empleados, gastos, inversiones, etc.), pero estas tareas de recolección de datos frecuentemente se contemplaban desde una perspectiva cuantitativa y no se asociaban a los objetivos y metas de los servicios, ni se orientaban a la toma de decisiones para la mejora de los mismos. Existirían así, dos posibles enfoques en las metodologías de evaluación: el vinculado a la planificación y a la gestión de calidad y aquel que se limita a reflejar cuantitativamente la situación del servicio. Por otro lado, cada vez está más extendida la idea de que para medir el rendimiento de un servicio se precisan tanto indicadores de naturaleza estrictamente cuantitativa, como indicadores categóricos, o más cualitativos, como son los utilizados para medir la satisfacción de los usuarios y del personal; utilizar ambos tipos de indicadores conduce a obtener una descripción más completa de todos los aspectos del rendimiento de los servicios. Con esta información, la técnica DEA permite analizar los servicios desde una perspectiva de unidades productivas o unidades a analizar, inputs o recursos introducidos en el sistema y outputs o resultados de los servicios.

La multitud de trabajos de aplicación del DEA en el ámbito de servicios hace imposible dar una relación sintetizada, ilustrativa y justa de las mismas. Atendiendo a literatura reciente, el capítulo 15 del Handbook on Data Envelopment Analysis (2011) [41] se dedica a aplicaciones del DEA en el sector servicios. En concreto se muestra el uso de DEA para examinar la eficiencia relativa de las universidades, investigando el caso de la población de universidades Australianas. Como segunda aplicación se examina la productividad de los hoteles más grandes de Queensland desde el punto de vista de la eficiencia técnica; uno de los outputs propuestos es la razón entre el número de clientes satisfechos por número total de clientes del hotel (Ball y otros 1986 [7]), por tanto, nos encontramos con un claro ejemplo de lo que va a ser habitual en lo considerado aportaciones en esta tesis, el uso de la opinión de los usuarios como output. El tercer ejemplo corresponde a registros de la propiedad donde la información utilizada es recogida mediante un cuestionario. En la última aplicación se utiliza el denominado network data envelopment analysis en el ámbito de oficinas.

En referencia a la Administración Pública, Martínez-Cabrera (2003) [99] manifiesta que no resulta factible la optimización de la eficiencia en términos de la maximización del beneficio. En este sentido, la búsqueda de indicadores, que puedan sustituir la medida del beneficio en la evaluación de la eficiencia, constituye un reto para las agencias públicas si pretenden aplicar técnicas de gestión. En la evaluación del sector público en general, la evaluación de la eficiencia técnica resulta lo más apropiado, ya que tanto la eficiencia precio como la eficiencia global precisan información sobre el precio de los factores, algo habitualmente no disponible en las agencias públicas.

Capítulo 4

Análisis de eficiencia DEA en presencia de outputs estimados

Este capítulo se estructura del siguiente modo: en la segunda sección realizaremos el planteamiento del problema y en la tercera propondremos una metodología de cara a determinar el tamaño de muestra, en cada DMU, que garantice que las eficiencias obtenidas utilizando los modelos DEA CCR y BCC con orientación output, sean estimadores con buenas propiedades estadísticas. A continuación, se propondrán diversas metodologías con objeto de construir un intervalo de confianza para la eficiencia DEA obtenida si se dispusiera de un censo en todas las DMUs, eficiencia poblacional; derivada de una de estas metodologías, en la quinta sección se propone una nueva alternativa de construcción de intervalo de eficiencia. Por último, propondremos una metodología de cálculo del tamaño de muestra necesario para alcanzar la confianza que deseemos en una región simultánea para el vector de índices de eficiencia poblacional.

4.1. Planteamiento del problema

Formalmente, consideramos n DMUs fijas, cada una utilizando m inputs para producir s outputs. Denotamos $\mathbf{X}_i = (X_{1i}, \dots, X_{mi})$ el valor de los m inputs en la i th DMU ($i = 1, \dots, n$). Los outputs de la i th DMU se representan por una población finita $P_i = \{\mathbf{U}_{1i}, \dots, \mathbf{U}_{N_i i}\}$, de tamaño N_i , donde $\mathbf{U}_{ki} = (U_{k1i}, \dots, U_{ksi})$ son s diferentes medidas que representan al k th individuo de la i th DMU, por ejemplo, s índices que midan su satisfacción. Los outputs poblacionales, desconocidos, son:

$$\{\mathbf{Y}_i = f(P_i)\}_{i=1, \dots, n} = \{(Y_{1i}, \dots, Y_{si})\}_{i=1, \dots, n} \quad (4.1)$$

siendo f una función que, aplicada a la población P_i , determinan los s outputs desconocidos (Y_{1i}, \dots, Y_{si}) . La matriz de covarianzas poblacional es

$$\Sigma_i = \frac{\sum_{k=1}^{N_i} (\mathbf{U}_{ki} - \mathbf{Y}_i)^t (\mathbf{U}_{ki} - \mathbf{Y}_i)}{N_i} \quad (4.2)$$

Si en la i th población consideramos $\mathbf{V}_{1i}, \dots, \mathbf{V}_{n_i i}$ una muestra aleatoria p_i de tamaño n_i de la población finita P_i , donde \mathbf{V}_{ki} es un vector aleatorio de dimensión s , los outputs estimados son:

$$\hat{\mathbf{Y}}_i = f(p_i) = (\hat{Y}_{1i}, \dots, \hat{Y}_{si}); \quad i = 1, \dots, n \quad (4.3)$$

y el estimador de la matriz de covarianzas poblacional es

$$S_i = \frac{\sum_{k=1}^{n_i} (\mathbf{V}_{ki} - \hat{\mathbf{Y}}_i)^t (\mathbf{V}_{ki} - \hat{\mathbf{Y}}_i)}{(n_i - 1)} \quad (4.4)$$

Fijada una muestra en cada DMU, es muy conveniente distinguir entre los estimadores definidos en (4.3) y sus valores estimados representados por:

$$\hat{\mathbf{y}}_i = (\hat{y}_{1i}, \dots, \hat{y}_{si}); i = 1, \dots, n \quad (4.5)$$

Por ejemplo, en este capítulo vamos a utilizar p_i una muestra aleatoria simple sin reemplazamiento de P_i para estimar el vector de medias poblacional:

$$\left\{ \mathbf{Y}_i = f(P_i) = \frac{\sum_{k=1}^{N_i} \mathbf{U}_{ki}}{N_i} = \left(\frac{\sum_{k=1}^{N_i} U_{k1i}}{N_i}, \dots, \frac{\sum_{k=1}^{N_i} U_{ksi}}{N_i} \right) \right\}_{i=1, \dots, n} \quad (4.6)$$

$$\left\{ \hat{\mathbf{Y}}_i = f(p_i) = \frac{\sum_{k=1}^{n_i} \mathbf{V}_{ki}}{n_i} = \left(\frac{\sum_{k=1}^{n_i} V_{k1i}}{n_i}, \dots, \frac{\sum_{k=1}^{n_i} V_{ksi}}{n_i} \right) \right\}_{i=1, \dots, n} \quad (4.7)$$

en esta situación, el estimador de la varianza del estimador del output \hat{Y}_{ri} , es:

$$\widehat{\text{Var}}(\hat{Y}_{ri}) = \frac{N_i^2 \left(1 - \frac{n_i}{N_i}\right)}{n_i} S_{ri}^2; i = 1, \dots, n; r = 1, \dots, s \quad (4.8)$$

donde

$$S_{ri}^2 = \left(\frac{\sum_{k=1}^{n_i} (V_{kri} - \hat{Y}_{ri})^2}{(n_i - 1)} \right)$$

y un intervalo de confianza de nivel $1 - \alpha$ para el r th output (media poblacional) de la i th DMU es

$$[\hat{Y}_{ri}^L, \hat{Y}_{ri}^U] = [\hat{Y}_{ri} - B_i, \hat{Y}_{ri} + B_i]; i = 1, \dots, n; r = 1, \dots, s \quad (4.9)$$

donde el error de estimación B_i es

$$B_i = \phi^{-1} \left(\frac{\alpha}{2} \right) \sqrt{\widehat{\text{Var}}(\hat{Y}_{ri})}$$

siendo ϕ la función de distribución de la ley de probabilidad normal estándar.

Observada la muestra $(v_{k1i}, \dots, v_{k1i})_{k=1, \dots, n_i}$ en cada DMU, entonces

$$\hat{\mathbf{y}}_i = (\hat{y}_{1i}, \dots, \hat{y}_{si}) = \left(\frac{\sum_{k=1}^{n_i} v_{k1i}}{n_i}, \dots, \frac{\sum_{k=1}^{n_i} v_{ksi}}{n_i} \right); i = 1, \dots, n \quad (4.10)$$

$$\widehat{\text{var}}(\hat{y}_{ri}) = \frac{N_i^2 \left(1 - \frac{n_i}{N_i}\right)}{n_i} s_{ri}^2; i = 1, \dots, n; r = 1, \dots, s \quad (4.11)$$

donde

$$s_{ri}^2 = \left(\frac{\sum_{k=1}^{n_i} (v_{kri} - \hat{y}_{ri})^2}{(n_i - 1)} \right)$$

y el intervalo observado es

$$[y_{ri}^L, y_{ri}^U]; i = 1, \dots, n; r = 1, \dots, s \quad (4.12)$$

donde

$$y_{ri}^L = \hat{y}_{ri} - \phi^{-1} \left(\frac{\alpha}{2} \right) \sqrt{\widehat{\text{var}}(\hat{y}_{ri})}$$

$$y_{ri}^U = \hat{y}_{ri} + \phi^{-1} \left(\frac{\alpha}{2} \right) \sqrt{\widehat{\text{var}}(\hat{y}_{ri})}$$

La Tabla 4.1 recoge los dos modelos DEA con los que vamos a trabajar. La elección de la orientación output se justifica en el interés de observar las unidades que, manteniendo los mismos recursos (inputs), pueden mejorar la opinión-satisfacción (outputs) de los usuarios. En esta situación denotamos:

- Índice de eficiencia (técnica) poblacional $(\varphi_i)_{i=1,\dots,n}$ CCR-O y BCC-O a la maximización de los modelos (4.13) y (4.14), respectivamente, tomando los datos $(\mathbf{X}_i, \mathbf{Y}_i)_{i=1,\dots,n}$. Como \mathbf{Y}_i son desconocidos, $(\varphi_i)_{i=1,\dots,n}$ son parámetros desconocidos que pretendemos estimar.
- Abusando de la notación, representamos los estimadores $(\hat{\Phi}_i)_{i=1,\dots,n}$ de los índices de eficiencia $(\varphi_i)_{i=1,\dots,n}$ CCR-O y BCC-O a la maximización de los modelos (4.13) y (4.14), respectivamente, tomando $(\mathbf{X}_i, \hat{\mathbf{Y}}_i)_{i=1,\dots,n}$, entendiéndose que los modelos son maximizados con los datos $(\mathbf{X}_i, \hat{\mathbf{y}}_i)_{i=1,\dots,n}$ para obtener los valores estimados de los estimadores $\hat{\Phi}_i$ que denotamos por $\hat{\varphi}_i$.

Tabla 4.1: Modelos DEA, para la DMU_o, con rendimientos constantes (CCR) y variables de escala (BCC); orientación output.

| | | |
|-------|---|--|
| | $\begin{aligned} \max \quad & \varphi + \varepsilon \left(\sum_{j=1}^m s_j^- + \sum_{r=1}^s s_r^+ \right) \\ \text{s.t.} \quad & \sum_{i=1}^n \lambda_i y_{ri} - s_r^+ = \varphi y_{ro}; \quad r = 1, \dots, s \\ & \sum_{i=1}^n \lambda_i x_{ji} + s_j^- = x_{jo}; \quad j = 1, \dots, m \end{aligned}$ | |
| CCR-O | $\lambda_i \geq 0, s_j^- \geq 0, s_r^+ \geq 0; \quad i = 1, \dots, n; \quad j = 1, \dots, m; \quad r = 1, \dots, s \quad (4.13)$ | |
| BCC-O | $\sum_{i=1}^n \lambda_i = 1 \quad (4.14)$ | |

Los problemas a los que vamos a dar solución son la determinación del tamaño de muestra en cada DMU que garantice una cota razonable para el error de estimación de φ_i , la obtención de un intervalo de confianza para cada φ_i y de una región de confianza para el vector $(\varphi_1, \dots, \varphi_n)$.

4.2. Cálculo del tamaño muestral

Ante el problema de encontrar el tamaño de muestra necesario en cada DMU para estimar los outputs y con éstos estimar la eficiencia DEA, en esta sección proponemos dos planteamientos. En el primero, buscamos garantizar una precisión razonable en la estimación de los parámetros de interés (outputs); decidido un diseño muestral, la obtención del tamaño de muestra se realiza del siguiente modo: para el r th output, fijamos un error de estimación e_r y una confianza α , y obtenemos el tamaño de muestra n_{ri} para que se verifique

$$P(|\hat{Y}_{ri} - Y_{ri}| \leq e_r) = 1 - \alpha; \quad r = 1, \dots, s \quad (4.15)$$

y el tamaño de muestra en la i th DMU es

$$n_i = \max\{n_{ri}\}_{r=1,\dots,s} \quad (4.16)$$

En particular, si los outputs son la respuesta media de los individuos de la muestra tomada aleatoria simple sin reemplazamiento, el tamaño de la muestra en la i th DMU se obtiene como en (4.16) tomando (Särndal y otros 2003 [118]):

$$n_{ri} \geq \frac{n_o}{\left(\frac{n_o}{N_i} + 1\right)}; r = 1, \dots, s \quad (4.17)$$

con

$$n_o = \frac{\tau_{1-\alpha/2}^2}{(e_r)^2} \sigma_{ri}^2 \text{ y } \tau_{1-\alpha/2} = \phi^{-1}(1 - \alpha/2)$$

siendo σ_{ri}^2 la varianza poblacional del r th output en la i th DMU, es decir

$$\sigma_{ri}^2 = \left(\frac{\sum_{k=1}^{N_i} (U_{kri} - \bar{U}_{ri})^2}{N_i} \right); \bar{U}_{ri} = \left(\frac{\sum_{k=1}^{N_i} U_{kri}}{N_i} \right)$$

y ϕ la función de distribución de la ley de probabilidad normal estándar.

En el segundo planteamiento, incluido dentro de lo que podemos considerar aportaciones de esta tesis, determinaremos la relación entre tamaño de muestra y error de estimación del índice de eficiencia. Esta metodología va a permitir determinar el tamaño de muestra necesario para lograr el error de estimación de la eficiencia que deseemos. Comenzaremos introduciendo diversos resultados analíticos que dan una solución al problema en el caso de un input y un output en el modelo CCR con orientación output (CCR-O) y, mediante simulaciones, se comprobarán estos resultados en la situación de dos inputs conocidos y dos outputs estocásticos en el modelo BCC-O.

4.2.1. Modelo CCR-O con un input conocido y un output estocástico

Consideramos n DMUs con un único input con valores conocidos $(X_i)_{i=1, \dots, n}$ y un único output con valores desconocidos $(Y_i)_{i=1, \dots, n}$. Sean $(\varphi_i)_{i=1, \dots, n}$ los correspondientes índices de eficiencia poblacionales CCR-O (desconocidos). Consideramos $(\hat{Y}_i)_{i=1, \dots, n}$ y $(\hat{\Phi}_i)_{i=1, \dots, n}$ los estimadores de los outputs y de los índices de eficiencia, respectivamente, y \hat{y}_i y $(\hat{\varphi}_i)_{i=1, \dots, n}$ los correspondientes valores estimados de los outputs y de los índices de eficiencia, respectivamente. Los valores estimados de los índices de eficiencia se obtienen de la maximización de los modelos (4.13) y (4.14), tomando los datos $(X_i, \hat{y}_i)_{i=1, \dots, n}$. Sea $(Z_i = \frac{Y_i}{X_i})_{i=1, \dots, n}$, $(\hat{Z}_i = \frac{\hat{Y}_i}{X_i})_{i=1, \dots, n}$ y $(\hat{z}_i = \frac{\hat{y}_i}{X_i})_{i=1, \dots, n}$

Con esta notación, tenemos: $\varphi_i = \frac{Z_i}{\max_{j=1, \dots, n} \{Z_j\}}$, $\hat{\varphi}_i = \frac{\hat{z}_i}{\max_{j=1, \dots, n} \{\hat{z}_j\}}$

Dado $p \in (0, 1)$, se consideran los sucesos:

$$A_i = (|\hat{Y}_i - Y_i| \leq pY_i); i = 1, \dots, n$$

Lema 1 Sea m tal que $Z_m = \max_{i=1, \dots, n} \{Z_i\}$, entonces:

$$A_i = \left(\frac{\hat{Z}_i}{1+p} \leq Z_i \leq \frac{\hat{Z}_i}{1-p} \right) = (Z_i(1-p) \leq \hat{Z}_i \leq Z_i(1+p)) = \left(\frac{\hat{Z}_i}{Z_m(1+p)} \leq \varphi_i \leq \frac{\hat{Z}_i}{Z_m(1-p)} \right)$$

DEMOSTRACIÓN

La primera igualdad se obtiene despejando Y_i y dividiendo por X_i . La segunda despejando \hat{Z}_i , y la tercera dividiendo por Z_m .

Lema 2 Sean m y l tales que:

$$\begin{aligned} Z_m &= \max_{i=1, \dots, n} \{Z_i\} \\ \hat{Z}_l &= \max_{i=1, \dots, n} \{\hat{Z}_i\} \end{aligned}$$

Considerar el suceso $B = (Z_m(1-p) \leq \hat{Z}_l \leq Z_m(1+p))$, entonces:

$m = l$ implica $A_l = B$

$m \neq l$ implica $A_l \cap A_m \subset B$

DEMOSTRACIÓN

Si $m = l$, del Lema 1, se sigue $A_l = B$

Si $m \neq l$, entonces $\hat{Z}_m < \hat{Z}_l$ y $Z_l < Z_m$ y, si $A_l \cap A_m$ se verifica, se sigue:

$$Z_m(1-p) \leq \hat{Z}_m < \hat{Z}_l \leq Z_l(1+p) < Z_m(1+p)$$

Lema 3 $\forall i = 1, \dots, n$ se verifica:

$$A_i \cap B \subset \left\{ \hat{\Phi}_i \frac{(1-p)}{(1+p)} \leq \frac{\hat{Z}_i}{Z_m(1+p)} \leq \varphi_i \right\} \cap \left\{ \varphi_i \leq \frac{\hat{Z}_i}{Z_m(1-p)} \leq \hat{\Phi}_i \frac{(1+p)}{(1-p)} \right\}$$

DEMOSTRACIÓN

Dividiendo por $(1+p)$ la primera desigualdad que define el suceso B , se obtiene:

$$\frac{(1-p)}{(1+p)} \leq \frac{\hat{Z}_l}{Z_m(1+p)}$$

dado $i = 1, \dots, n$:

$$\hat{\Phi}_i \frac{(1-p)}{(1+p)} \leq \hat{\Phi}_i \frac{\hat{Z}_l}{Z_m(1+p)} = \frac{\hat{Z}_i}{Z_m(1+p)}; \text{ puesto que } \hat{\Phi}_i = \frac{\hat{Z}_i}{Z_l}$$

Dividiendo la segunda desigualdad que define el suceso B por $(1-p)$, se obtiene:

$$\frac{\hat{Z}_l}{Z_m(1-p)} \leq \frac{(1+p)}{(1-p)}$$

dado $i = 1, \dots, n$:

$$\hat{\Phi}_i \frac{\hat{Z}_l}{Z_m(1-p)} \leq \hat{\Phi}_i \frac{(1+p)}{(1-p)} \text{ y como } \hat{\Phi}_i = \frac{\hat{Z}_i}{Z_l}, \text{ se obtiene } \frac{\hat{Z}_i}{Z_m(1-p)} \leq \hat{\Phi}_i \frac{(1+p)}{(1-p)}$$

Como $A_i = \left(\frac{\hat{Z}_i}{Z_m(1+p)} \leq \varphi_i \leq \frac{\hat{Z}_i}{Z_m(1-p)} \right)$ se verifica

$$\text{Por tanto } A_i \cap B \implies \left(\hat{\Phi}_i \frac{(1-p)}{(1+p)} \leq \frac{\hat{Z}_i}{Z_m(1+p)} \leq \varphi_i \leq \frac{\hat{Z}_i}{Z_m(1-p)} \leq \hat{\Phi}_i \frac{(1+p)}{(1-p)} \right)$$

Teorema 1 Si $p(A_i) \geq 1 - \alpha \forall i = 1, \dots, n$, entonces

$$p(\varphi_i \in I_i) \geq (1 - \alpha)^3 \quad (4.18)$$

donde

$$I_i = \left[\hat{\Phi}_i \frac{(1-p)}{(1+p)}, \min \left\{ \hat{\Phi}_i \frac{(1+p)}{(1-p)}, 1 \right\} \right] \quad (4.19)$$

DEMOSTRACIÓN

$$p \left(\hat{\Phi}_i \frac{(1-p)}{(1+p)} \leq \varphi_i \leq \hat{\Phi}_i \frac{(1+p)}{(1-p)} \right) \geq p(A_i \cap B) \geq p(A_i \cap A_l \cap A_m) = (1 - \alpha)^3$$

La primera desigualdad se sigue del Lema 3, la segunda del Lema 2, y la última desigualdad se sigue porque los sucesos A_i son independientes.

Corolario 1 A la vista de la demostración, la probabilidad de que el intervalo I_i contenga a φ_i es:

- I. Si $l \neq m$ y $i \neq l, m$ la probabilidad es mayor o igual a $(1 - \alpha)^3$; si $i = l$ o $i = m$ la probabilidad es mayor o igual a $(1 - \alpha)^2$.
- II. Si $l = m$ y $i \neq m$ la probabilidad es mayor o igual a $(1 - \alpha)^2$; si $i = m$ la probabilidad es mayor o igual a $(1 - \alpha)$.

DEMOSTRACIÓN

Si $l \neq m$

$$p(A_i \cap A_l \cap A_m) = \begin{cases} (1 - \alpha)^3 & \text{para } i \neq m, l \\ (1 - \alpha)^2 & \text{para } i = m \text{ or } i = l \end{cases}$$

Si $m = l$

$$p(A_i \cap A_l \cap A_m) = p(A_i \cap A_m) = \begin{cases} (1 - \alpha)^2 & \text{para } i \neq m \\ (1 - \alpha) & \text{para } i = m \end{cases}$$

Corolario 2 Bajo las hipótesis del Teorema 1, $\widehat{\Phi}_i$ es un estimador consistente para $\varphi_i, \forall i = 1, \dots, n$.

DEMOSTRACIÓN

Si $\varepsilon \geq 1$, siempre se verifica que $|\widehat{\Phi}_i - \varphi_i| < \varepsilon$, porque $\widehat{\Phi}_i \in [0, 1]$ y $\varphi_i \in [0, 1]$

$\forall \varepsilon \in (0, 1)$ y $\alpha > 0$, entonces sea $p_0/\varepsilon = \frac{4p_0}{(1+p_0)(1-p_0)} = \frac{(1+p_0)}{(1-p_0)} - \frac{(1-p_0)}{(1+p_0)}$

Consideremos n_i tal que

$$p(|\widehat{Y}_i - Y_i| \leq p_0 Y_i) \geq \sqrt[3]{1 - \alpha}, \forall i = 1, \dots, n$$

Entonces, por Teorema 1:

$$p\left(\widehat{\Phi}_i \frac{(1-p_0)}{(1+p_0)} \leq \varphi_i \leq \widehat{\Phi}_i \frac{(1+p_0)}{(1-p_0)}\right) \geq (1 - \alpha)$$

como siempre se verifica $\widehat{\Phi}_i \frac{(1-p_0)}{(1+p_0)} \leq \widehat{\Phi}_i \leq \widehat{\Phi}_i \frac{(1+p_0)}{(1-p_0)}$, tenemos que

$$p\left(|\widehat{\Phi}_i - \varphi_i| \leq \widehat{\Phi}_i \left[\frac{(1+p_0)}{(1-p_0)} - \frac{(1-p_0)}{(1+p_0)}\right]\right) \geq (1 - \alpha)$$

y como $0 \leq \widehat{\Phi}_i \leq 1$

$$p\left(|\widehat{\Phi}_i - \varphi_i| \leq \varepsilon\right) \geq 1 - \alpha$$

Lema 4 Bajo las hipótesis del Teorema 1 si

$$\varphi_i \in \left[\widehat{\Phi}_i \frac{(1-p)}{1+p}, \min\left\{1, \widehat{\Phi}_i \frac{(1+p)}{1-p}\right\}\right]$$

entonces

$$\varphi_i \in \left[\widehat{\Phi}_i - \frac{2p}{1+p}, \min\left\{1, \widehat{\Phi}_i + \frac{2p}{1+p}\right\}\right]$$

DEMOSTRACIÓN

Si $\hat{\Phi}_i \frac{(1-p)}{(1+p)} \leq \varphi_i \implies \hat{\Phi}_i \left(1 - \frac{2p}{1+p}\right) \leq \varphi_i \implies \hat{\Phi}_i - \varphi_i \leq \frac{2p}{1+p} \hat{\Phi}_i \leq \frac{2p}{1+p}$; la última desigualdad se sigue de $0 \leq \hat{\Phi}_i \leq 1$

Si $\min \left\{1, \hat{\Phi}_i \frac{(1+p)}{(1-p)}\right\} = 1 \implies \hat{\Phi}_i \frac{(1+p)}{(1-p)} \geq 1 \implies \hat{\Phi}_i \geq \frac{(1-p)}{(1+p)} \implies \hat{\Phi}_i + \frac{2p}{1+p} \geq \frac{(1-p)}{(1+p)} + \frac{2p}{(1+p)} = 1$ y $\min \left\{1, \hat{\Phi}_i + \frac{2p}{1+p}\right\} = 1$

Si $\varphi_i \leq \min \left\{1, \hat{\Phi}_i \frac{(1+p)}{(1-p)}\right\} = \hat{\Phi}_i \frac{(1+p)}{(1-p)} \implies \varphi_i \leq \hat{\Phi}_i \left(\frac{1}{1 - \frac{2p}{1+p}}\right) \implies \varphi_i \left(1 - \frac{2p}{1+p}\right) \leq \hat{\Phi}_i \implies \varphi_i - \hat{\Phi}_i \leq \frac{2p}{(1+p)} \varphi_i \leq \frac{2p}{1+p}$; la última desigualdad se sigue de $0 \leq \varphi_i \leq 1$.

Teorema 2 Consideremos el modelo CCR con orientación output y sólo un input conocido y un output desconocido. Fijado $\delta \in (0, 1)$ y $\alpha \in (0, 1)$, para cada $i = 1, \dots, n$, sea n_i el tamaño muestral en la i th DMU, tal que:

$$p \left(|\hat{Y}_i - Y_i| \leq \frac{\delta}{2 - \delta} Y_i \right) \geq \sqrt[3]{1 - \alpha} \quad (4.20)$$

entonces

$$p \left(|\hat{\Phi}_i - \varphi_i| \leq \delta \right) \geq 1 - \alpha; \forall i = 1, \dots, n \quad (4.21)$$

DEMOSTRACIÓN

El resultado se sigue del Teorema 1 y del Lema 4 usando $p = \frac{\delta}{2 - \delta}$

Nota 1 Si la muestra se toma en cada DMU con muestreo aleatorio simple sin reemplazamiento, y el output es una media poblacional,

$$\left\{ Y_i = \sum_{k=1}^{N_i} U_{ki} / N_i \right\}_{i=1, \dots, n}$$

donde U_{ki} es la respuesta (satisfacción) del individuo k en la i th DMU, entonces el tamaño muestral n_i que verifica

$$p \left(|\hat{Y}_i - Y_i| \leq p Y_i \right) \geq 1 - \alpha$$

se obtiene como en (4.17).

Nota 2 En la misma situación que en la Nota 1, si el número de outputs es $s \geq 2$

$$\left\{ Y_{ri} = \sum_{k=1}^{N_i} U_{kri} / N_i \right\}_{\substack{r=1, \dots, s \\ i=1, \dots, n}}$$

y si n_{ri} es el tamaño de muestra obtenido como en la Nota 1, $r = 1, \dots, s$, entonces el tamaño de muestra en la i th DMU es

$$n_i = \max\{n_{1i}, \dots, n_{si}\} \quad (4.22)$$

Nota 3 Bajo las hipótesis del Teorema 1, el intervalo I_i nos permite contrastar:

$$H_0 : \varphi_i = 1$$

$$H_1 : \varphi_i < 1$$

4.2.2. Modelos CCR-O y BCC-O con más de un input conocido y un output estocástico

Para los modelos CCR y BCC con orientación output considerando más de un input y más de un output estocástico, los resultados se comprueban utilizando simulaciones.

UN ESTUDIO DE SIMULACIÓN

La Tabla 4.2 contiene los datos utilizados en Cooper y otros (2006) [39] correspondientes al número de médicos y enfermeras (inputs) y el número de altas e ingresos (outputs) en 12 centros sanitarios. Utilizaremos estos datos para generar el modelo poblacional de la siguiente forma: en la i th DMU, $i = 1, \dots, n$, se genera una población finita $P_i = \{U_{1i}, \dots, U_{N_i i}\}$, de tamaño N_i , $i = 1, \dots, 12$, donde las N_i son variables aleatorias independientes con distribución uniforme en $[10000, 50000]$, y

$$U_{ki} \hookrightarrow N_2 \left(\begin{pmatrix} z_{1i} \\ z_{2i} \end{pmatrix}, \begin{pmatrix} z_{1i}^2/4 & 0 \\ 0 & z_{2i}^2/4 \end{pmatrix} \right)$$

siendo (z_{1i}, z_{2i}) los valores output originales de la i th DMU, columnas 4 y 5 de la Tabla 4.2.

Tabla 4.2: Número de médicos, enfermeras, pacientes dados de alta e ingresados en 12 centros de salud

| DMU | Médicos (X1) | Enfermeras (X2) | Altas (Z1) | Ingresos (Z2) | Índice de eficiencia CCR-O | Índice de eficiencia BCC-O |
|-----|-----------------|--------------------|---------------|------------------|----------------------------------|----------------------------------|
| 1 | 2.0 | 15.1 | 10 | 9 | 1 | 1 |
| 2 | 1.9 | 13.1 | 15 | 5 | 1 | 1 |
| 3 | 2.5 | 16 | 16 | 5.5 | 0.883 | 0.925 |
| 4 | 2.7 | 16.8 | 18 | 7.2 | 1 | 1 |
| 5 | 2.2 | 15.8 | 9.4 | 6.6 | 0.763 | 0.767 |
| 6 | 5.5 | 25.5 | 23 | 9 | 0.835 | 0.955 |
| 7 | 3.3 | 23.5 | 22 | 8.8 | 0.902 | 1 |
| 8 | 3.1 | 20.6 | 15.2 | 8 | 0.796 | 0.826 |
| 9 | 3 | 24.4 | 19 | 10 | 0.960 | 0.990 |
| 10 | 5 | 26.8 | 25 | 10 | 0.871 | 1 |
| 11 | 5.3 | 30.6 | 26 | 14.7 | 0.955 | 1 |
| 12 | 3.8 | 28.4 | 25 | 12 | 0.958 | 1 |

Fuente: Tabla 1.5 Cooper y otros (2006) [39]

La Tabla 4.3 muestra el modelo poblacional simulado. En la segunda columna observamos el tamaño poblacional de los individuos de cada DMU, en las columnas 3 y 4 los inputs conocidos, los mismos que los de la Tabla 4.2. Las columnas 5 y 6 contienen los valores simulados de los outputs, Y_1 y Y_2 , que corresponden a la media poblacional:

$$(Y_{1i}, Y_{2i}) = \left(\frac{\sum_{k=1}^{N_i} U_{k1i}}{N_i}, \frac{\sum_{k=1}^{N_i} U_{k2i}}{N_i} \right); \quad i = 1, \dots, n \quad (4.23)$$

Las dos últimas columnas muestran los índices de eficiencia poblacional CCR-O y BCC-O.

Tabla 4.3: Modelo poblacional simulado

| DMU | Tamaño poblacional (N_i) | Médicos X1 | Enfermeras X2 | Y1 | Y2 | Indice de eficiencia poblacional (φ_i). CCR-O | Indice de eficiencia poblacional (φ_i). BCC-O |
|-----|------------------------------|------------|---------------|-------|-------|---|---|
| 1 | 43341 | 2.0 | 15.1 | 9.98 | 9.01 | 1 | 1 |
| 2 | 24438 | 1.9 | 13.1 | 14.98 | 5.01 | 1 | 1 |
| 3 | 45606 | 2.5 | 16 | 15.99 | 5.50 | 0.883 | 0.926 |
| 4 | 12578 | 2.7 | 16.8 | 17.96 | 7.18 | 1 | 1 |
| 5 | 19314 | 2.2 | 15.8 | 9.40 | 6.58 | 0.763 | 0.766 |
| 6 | 21782 | 5.5 | 25.5 | 22.96 | 8.97 | 0.835 | 0.957 |
| 7 | 19024 | 3.3 | 23.5 | 21.99 | 8.77 | 0.901 | 0.998 |
| 8 | 36271 | 3.1 | 20.6 | 15.17 | 8.01 | 0.797 | 0.826 |
| 9 | 30691 | 3 | 24.4 | 19.02 | 10.04 | 0.963 | 0.991 |
| 10 | 28385 | 5 | 26.8 | 24.89 | 10.01 | 0.871 | 1 |
| 11 | 28005 | 5.3 | 30.6 | 26.11 | 14.69 | 0.958 | 1 |
| 12 | 49077 | 3.8 | 28.4 | 25.09 | 11.98 | 0.960 | 1 |

La simulación lo que pretende es describir la situación observada en las aplicaciones donde los outputs son desconocidos y, en la i th DMU, los outputs son estimados como la media de la muestra observada en esta DMU. El tamaño de muestra en cada DMU se calcula, fijado δ y α , utilizando el Teorema 2 y la Nota 2. Para este cálculo se necesita una estimación inicial de la media y la varianza poblacional de los outputs en cada DMU. Los pasos del proceso de simulación para obtener los valores estimados de las eficiencias en los modelos CCR-O y BBC-O son los siguientes

- I. En la i th DMU, $i = 1, \dots, 12$, se toma una muestra previa de tamaño $n_i^{(0)}$. Con esta muestra, obtenemos $\hat{y}_i^{(0)} = (\hat{y}_{1i}^{(0)}, \hat{y}_{2i}^{(0)})$, como en (4.10), y la estimación de la varianza $\hat{\sigma}_i^{2(0)} = (\hat{\sigma}_{1i}^{2(0)}, \hat{\sigma}_{2i}^{2(0)})$, como en (4.11). El tamaño de muestra inicial $n_i^{(0)}$ debe ser pequeño con respecto al tamaño poblacional, por ejemplo, en este caso, $n_i^{(0)} = 25$.
- II. Fijados $\delta=0.2$ y 0.1 y $\alpha =0.1$, y con los valores $(\hat{y}_i^{(0)}, \hat{\sigma}_i^{2(0)})$, el tamaño muestral, n_i , se determina como en la Nota 2.
- III. En la i th DMU, se toma el tamaño de muestra n_i y se estiman las medias $\hat{y}_i = (\hat{y}_{1i}, \hat{y}_{2i})$. Con los datos $(X_{1i}, X_{2i}, \hat{y}_{1i}, \hat{y}_{2i})$, se obtienen los valores $(\hat{\varphi}_i)$, para cada DMU, maximizando los modelos (4.13) y (4.14).
- IV. Se realizan mil iteraciones del paso III. y se obtienen los valores $\{\hat{\varphi}_i^{(k)}\}_{k=1, \dots, 1000}$ y los intervalos

$$I_i^{(k)} = \left[\hat{\varphi}_i^{(k)} - \delta, \min \{ \hat{\varphi}_i^{(k)} + \delta, 1 \} \right]; \quad i = 1, \dots, 12; \quad k = 1, \dots, 1000$$
- v. Se aproxima la probabilidad $p \left(\varphi_i \in \left[\hat{\Phi}_i - \delta, \min \{ \hat{\Phi}_i + \delta, 1 \} \right] \right)$ calculando

$$C_i = \left(\frac{1}{1000} \sum_{k=1}^{1000} I_{(\varphi_i \in I_i^{(k)})} \right); \quad i = 1, \dots, 12$$

VI. La proporción de veces que la i th DMU es eficiente en las mil iteraciones se obtiene como:

$$T_i = \left(\frac{1}{1000} \sum_{k=1}^{1000} I_{(\hat{\varphi}_i^{(k)}=1)} \right); \quad i = 1, \dots, 12$$

RESULTADOS DE LA SIMULACIÓN:

La Tabla 4.4 contiene el tamaño de muestra previo y los tamaños de muestra obtenidos para cada DMU, en la última iteración, para dos valores de δ y $\alpha = 0.1$.

Tabla 4.4: Tamaño muestral obtenido para cada DMU: Previo y tomando $\delta = 0.2$ y $\delta = 0.1$

| DMU | $n_i^{(0)}$ | n_i | |
|-----------|-------------|----------------|----------------|
| | | $\delta = 0.2$ | $\delta = 0.1$ |
| 1 | 25 | 130 | 615 |
| 2 | 25 | 126 | 474 |
| 3 | 25 | 110 | 346 |
| 4 | 25 | 137 | 446 |
| 5 | 25 | 151 | 529 |
| 6 | 25 | 133 | 361 |
| 7 | 25 | 94 | 488 |
| 8 | 25 | 111 | 582 |
| 9 | 25 | 134 | 655 |
| 10 | 25 | 112 | 611 |
| 11 | 25 | 155 | 753 |
| 12 | 25 | 73 | 572 |

La Tabla 4.5 muestra la aproximación de la probabilidad

$$p \left(\varphi_i \in \left[\hat{\Phi}_i - \delta, \min \left\{ \hat{\Phi}_i + \delta, 1 \right\} \right] \right); \quad i = 1, \dots, 12$$

Tabla 4.5: Aproximación de la probabilidad: $p \left(\varphi_i \in \left[\hat{\Phi}_i - \delta, \min \left\{ \hat{\Phi}_i + \delta, 1 \right\} \right] \right)$

| DMU | C_i | | | |
|-----------|----------------|-------|----------------|-------|
| | $\delta = 0.2$ | | $\delta = 0.1$ | |
| | CCR-O | BCC-O | CCR-O | BCC-O |
| 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 0.999 | 0.999 |
| 4 | 1 | 1 | 1 | 1 |
| 5 | 1 | 1 | 1 | 1 |
| 6 | 1 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 |
| 9 | 1 | 1 | 1 | 1 |
| 10 | 1 | 1 | 1 | 1 |
| 11 | 1 | 1 | 1 | 1 |
| 12 | 1 | 1 | 1 | 1 |

La Tabla 4.6 muestra la proporción de veces que cada una de las DMUs son eficientes. Para $\delta = 0.1$, las DMUs 1 y 2 son eficientes en todas las iteraciones y para $\delta = 0.2$, sólo la DMU 1 es eficiente en todas las iteraciones.

Tabla 4.6: Proporción de veces que los valores estimados de los índices de eficiencia toman el valor 1

| DMU | φ_i | CCR-O | |
|-----|-------------|----------------|----------------|
| | | T_i | |
| | | $\delta = 0.2$ | $\delta = 0.1$ |
| 1 | 1 | 1 | 1 |
| 2 | 1 | 0.996 | 1 |
| 3 | 0.883 | 0.014 | 0 |
| 4 | 1 | 0.613 | 0.705 |
| 5 | 0.763 | 0 | 0 |
| 6 | 0.835 | 0 | 0 |
| 7 | 0.901 | 0.014 | 0 |
| 8 | 0.797 | 0 | 0 |
| 9 | 0.963 | 0.16 | 0.012 |
| 10 | 0.871 | 0 | 0 |
| 11 | 0.958 | 0.094 | 0.004 |
| 12 | 0.960 | 0.2 | 0.026 |

4.2.3. Un estudio de aplicación

En cualquier estudio social, la primera cuestión a resolver es el tamaño de muestra necesario en cada DMU para garantizar una precisión razonable de la estimación del parámetro de interés, en este caso el índice de eficiencia DEA. El Teorema 2 nos proporciona un procedimiento para calcular este tamaño de cara a cometer un error de estimación fijado a priori.

La Tabla 4.8 contiene la información de una base de datos compuesta por 15 DMUs, dos inputs (columnas 3 y 4) y un output (columna 6). Los datos inputs corresponden al número de préstamos de libros (X_1) y al número de puestos de usuarios (número de asientos y número de ordenadores ofrecidos a los usuarios) (X_2) en 15 bibliotecas, datos escalados a 0-10 en busca de lograr simplicidad. La población son los 637,100 usuarios de estas bibliotecas. En cada DMU, sólo usamos el output dado por la satisfacción media del usuario, satisfacción con su biblioteca medida en cada usuario con un valor entero de 0 (nada satisfecho) a 10 (muy satisfecho). De cara a poder aplicar el Teorema 2 y la Nota 1 y obtener el tamaño de muestra n_i en la i th DMU, se necesita tomar una muestra piloto para disponer de una estimación del output y de la varianza poblacional. Por esta razón, tomamos, en cada DMU, una muestra piloto aleatoria simple sin reemplazamiento $\{n_i^{(0)}\}_{i=1,\dots,n}$ del 0.1 % del tamaño de la población y estimamos estos dos parámetros (Tabla 4.7). Fijados $\delta=0.1$ y $\alpha=0.1$, se determina el tamaño de muestra n_i (columna 5 de la Tabla 4.8) y se obtienen los valores estimados de la satisfacción media $(\hat{y}_i)_{i=1,\dots,15}$ (columna 6 de la Tabla 4.8). Con los datos $(X_{1i}, X_{2i}, \hat{y}_i)_{i=1,\dots,15}$, se obtienen los valores estimados de los índices de eficiencia $(\hat{\varphi}_i)_{i=1,\dots,15}$ y los intervalos $\{[\hat{\varphi}_i - \delta, \min\{\hat{\varphi}_i + \delta, 1\}]\}_{i=1,\dots,15}$. Consideramos el modelo BCC porque descartamos, en este caso práctico, la suposición de variaciones constantes de escala.

Tabla 4.7: Muestra aleatoria simple previa: tamaño de muestra, estimación de la satisfacción media de los usuarios y de la varianza poblacional

| DMU | Tamaño muestral $n_i^{(0)}$ | Estimación de la satisfacción media de los usuarios | Estimación de la varianza poblacional |
|-----|-----------------------------|---|---------------------------------------|
| 1 | 89 | 5.82 | 13.79 |
| 2 | 79 | 5.91 | 12.82 |
| 3 | 64 | 4.11 | 11.12 |
| 4 | 59 | 5.51 | 16.26 |
| 5 | 57 | 6.02 | 18.00 |
| 6 | 51 | 3.88 | 13.06 |
| 7 | 42 | 5.37 | 14.72 |
| 8 | 37 | 3.62 | 11.01 |
| 9 | 35 | 4.47 | 15.25 |
| 10 | 32 | 3.46 | 10.39 |
| 11 | 24 | 6.40 | 13.82 |
| 12 | 21 | 4.68 | 16.42 |
| 13 | 19 | 4.38 | 15.03 |
| 14 | 17 | 3.56 | 12.03 |
| 15 | 13 | 6.95 | 12.75 |

Tabla 4.8: Base de datos compuesta por 15 bibliotecas: Tamaño poblacional, préstamos de libros y puestos para usuarios (inputs), tamaño muestral y estimación de la satisfacción media de los usuarios (output) con $\delta = 0.1, \alpha = 0.1$, valores estimados del índice de eficiencia BCC-O, e intervalo del Teorema 2

| DMU | Tamaño poblacional N_i | Prést. de libros X1 | Puestos de usuarios X2 | Tamaño de muestra n_i | Satisf. media de los usuarios \hat{y}_i | Índice de eficiencia estimado $\hat{\phi}_i$ | $[\hat{\phi}_i - \delta, \min\{\hat{\phi}_i + \delta, 1\}]$ |
|-----|--------------------------|---------------------|------------------------|-------------------------|---|--|---|
| 1 | 89300 | 7.04 | 7.82 | 652 | 5.85 | 0.970 | [0.870, 1] |
| 2 | 78500 | 7.81 | 6.87 | 587 | 5.04 | 0.8235 | [0.735, 0.935] |
| 3 | 64000 | 5.41 | 5.60 | 1046 | 3.48 | 0.577 | [0.477, 0.677] |
| 4 | 59100 | 2.66 | 5.18 | 850 | 5.49 | 0.933 | [0.833, 1] |
| 5 | 56500 | 3.96 | 4.95 | 789 | 5.85 | 0.975 | [0.875, 1] |
| 6 | 50700 | 3.28 | 4.44 | 1363 | 3.31 | 0.558 | [0.458, 0.658] |
| 7 | 41600 | 4.36 | 3.64 | 806 | 6.04 | 1 | [0.9, 1] |
| 8 | 37000 | 6.29 | 3.24 | 1309 | 3.49 | 0.581 | [0.481, 0.681] |
| 9 | 34600 | 5.82 | 3.03 | 1187 | 5.31 | 0.886 | [0.786, 0.986] |
| 10 | 32000 | 7.69 | 2.80 | 1340 | 3.38 | 0.565 | [0.465, 0.665] |
| 11 | 23600 | 2.61 | 2.07 | 533 | 5.51 | 0.938 | [0.838, 1] |
| 12 | 21200 | 3.61 | 1.86 | 1147 | 5.82 | 1 | [0.888, 1] |
| 13 | 18900 | 4.73 | 1.65 | 1187 | 5.36 | 0.913 | [0.813, 1] |
| 14 | 17200 | 2.12 | 1.51 | 1408 | 3.36 | 1 | [0.900, 1] |
| 15 | 12900 | 2.16 | 1.13 | 413 | 5.84 | 1 | [0.900, 1] |

Las bibliotecas {1, 4, 5, 7, 11, 12, 13, 14, 15} se pueden considerar eficientes porque el intervalo para la eficiencia del Teorema 2 contiene al valor 1.

4.3. Intervalo de confianza de eficiencia DEA (ICE)

La probabilidad de selección de los individuos de la muestra determina un error en la estimación del output y en la obtención del índice de eficiencia DEA que pasa a ser una estimación puntual. Una forma de mejorar esta estimación es determinar intervalos con una confianza razonable de contener al verdadero índice de eficiencia poblacional.

En la sección 3.2 de preliminares, hemos introducido el punto de vista optimista/pesimista como el método más habitual de obtención de intervalos de eficiencia y se han estudiado sus diversas variantes (Entani y otros 2002 [52], Despotis and Smirlis 2002 [46], Wang y otros 2005 [146]). Desde este punto de vista optimista/pesimista, vamos a proponer diversos métodos de obtención de intervalos de confianza para los índices de eficiencia poblacionales definidos en (4.13) y (4.14). Estos intervalos difieren de los intervalos de eficiencia de Despotis y Smirlis y de Wang y otros, en que estos autores consideran que los datos inputs y/o outputs son intervalos. Los datos intervalo no son intervalos de confianza y en los intervalos de eficiencia no tiene sentido medir la confianza. Asimismo, los intervalos que se van a proponer, basados en el punto de vista optimista/pesimista, difieren de los intervalos de eficiencia DEA de Entani y otros en que éstos se obtienen de la maximización (extremo superior del intervalo) y minimización (extremo inferior del intervalo) de problemas usando datos input-output exactos; los modelos DEA intervalo son extendidos a datos inputs-output intervalos. Nosotros partimos, en nuestros modelos DEA, de intervalos de confianza para los outputs.

Simar and Wilson (1998) [126] y Ferrier and Hirschberg (1997) [56] plantean un problema de obtención de intervalo de confianza DEA completamente diferente al nuestro porque para estos autores la aleatoriedad proviene de considerar las n DMUs una muestra de tamaño n de una población de DMUs; el remuestreo bootstrap de la muestra disponible permite realizar estimaciones bootstrap de la frontera de producción y obtener intervalos de confianza para las eficiencias. La aleatoriedad, en nuestro problema, viene de disponer de una muestra, en cada DMU, con la que estimamos el output(s); nosotros remuestrearemos en la muestra de cada DMU para obtener intervalos de confianza de eficiencia DEA bootstrap.

El planteamiento inicial, de los distintos métodos de construcción de intervalos de confianza de eficiencia DEA, lo vamos a realizar considerando el modelo CCR con orientación output, m inputs conocidos y s outputs estimados con el intervalo (4.9).

4.3.1. Intervalo de confianza *Tipol*

Para definir los modelos (4.24) y (4.25), que determinan las cotas inferior y superior del ICE *Tipol* del valor de eficiencia para la DMU_o en el modelo CCR con orientación output, consideramos la formulación optimista y pesimista de los modelos DEA de Despotis and Smirlis (2002) [46] que, usando datos intervalo, definen cotas de eficiencia superior e inferior para las DMUs. El modelo optimista trata de hacer la DMU tan eficiente como sea posible usando los extremos inferiores de los intervalos de sus inputs y los extremos superiores de los intervalos de sus outputs, pero justo lo opuesto para todas las otras DMUs. El modelo pesimista, sin embargo, considera la situación inversa, es decir, los extremos superiores de los intervalos de los inputs y los extremos inferiores de todos los outputs de la DMU objetivo, mientras que para el resto de DMUs se toma lo opuesto.

$$\begin{aligned}
 & \max_{u_r, v_j} D_o^L = \sum_{r=1}^s u_r \hat{Y}_{ro}^L \\
 & s.t. \\
 & \sum_{j=1}^m v_j x_{jo} = 1 \\
 & \sum_{r=1}^s u_r \hat{Y}_{ro}^L - \sum_{j=1}^m v_j x_{jo} \leq 0 \\
 & \sum_{r=1}^s u_r \hat{Y}_{ri}^U - \sum_{j=1}^m v_j x_{ji} \leq 0; i \neq o \\
 & u_r, v_j \geq \varepsilon \forall r, j
 \end{aligned} \tag{4.24}$$

$$\begin{aligned}
& \max_{u_r, v_j} D_o^U = \sum_{r=1}^s u_r \hat{Y}_{ro}^U \\
& \text{s.t.} \\
& \sum_{j=1}^m v_j x_{jo} = 1 \\
& \sum_{r=1}^s u_r \hat{Y}_{ro}^U - \sum_{j=1}^m v_j x_{jo} \leq 0 \\
& \sum_{r=1}^s u_r \hat{Y}_{ri}^L - \sum_{j=1}^m v_j x_{ji} \leq 0; i \neq o \\
& u_r, v_j \geq \varepsilon \forall r, j
\end{aligned} \tag{4.25}$$

4.3.2. Intervalo de confianza Tipo2

Los modelos (4.26) y (4.27) muestran la forma de obtener la cota inferior y superior, respectivamente, del ICE *Tipo2* para la DMU_o, en el modelo CCR con orientación output, considerando la formulación optimista/pesimista de los modelos DEA de intervalo de eficiencia de Wang y otros (2005) [146]. Los modelos que determinan el intervalo *Tipo1* utilizan conjuntos de restricciones diferentes de una DMU a otra e, incluso, para obtener los extremos inferior y superior de una misma DMU. La principal desventaja de utilizar diferentes conjuntos de restricciones es la falta de comparabilidad entre las eficiencias debido a las diferentes fronteras adoptadas en el proceso de medir las eficiencias. Los modelos DEA de intervalo de eficiencia de Wang y otros y, por tanto, del intervalo de confianza *Tipo2*, utilizan siempre el mismo conjunto de restricciones, que conforman una única y fija frontera de producción para obtener tanto la cota superior como inferior de todos los intervalos de todas las DMUs.

$$\begin{aligned}
& \max_{u_r, v_j} W_o^L = \sum_{r=1}^s u_r \hat{Y}_{ro}^L \\
& \text{s.t.} \\
& \sum_{j=1}^m v_j x_{jo} = 1 \\
& \sum_{r=1}^s u_r \hat{Y}_{ri}^U - \sum_{j=1}^m v_j x_{ji} \leq 0; i = 1, \dots, n \\
& u_r, v_j \geq \varepsilon \forall r, j
\end{aligned} \tag{4.26}$$

$$\begin{aligned}
& \max_{u_r, v_j} W_o^U = \sum_{r=1}^s u_r \hat{Y}_{ro}^U \\
& \text{s.t.} \\
& \sum_{j=1}^m v_j x_{jo} = 1 \\
& \sum_{r=1}^s u_r \hat{Y}_{ri}^U - \sum_{j=1}^m v_j x_{ji} \leq 0; i = 1, \dots, n \\
& u_r, v_j \geq \varepsilon \forall r, j
\end{aligned} \tag{4.27}$$

4.3.3. Intervalo de confianza Tipo3

Nosotros proponemos otro ICE que mantiene el punto de vista optimista/pesimista, utilizando dos fronteras de producción, una para cada extremo del intervalo, lo cual permite obtener extremos de intervalos comparables. Los modelos (4.28) y (4.29) definen los extremos inferior y superior del ICE *Tipo3*, de la DMU_o:

$$\begin{aligned}
& \max_{u_r, v_j} T_o^L = \sum_{r=1}^s u_r \hat{Y}_{ro}^L \\
& \text{s.t.} \\
& \sum_{j=1}^m v_j x_{jo} = 1 \\
& \sum_{r=1}^s u_r \hat{Y}_{ri}^U - \sum_{j=1}^m v_j x_{ji} \leq 0; i = 1, \dots, n \\
& u_r, v_j \geq \varepsilon \forall r, j
\end{aligned} \tag{4.28}$$

$$\begin{aligned}
& \max_{u_r, v_j} T_o^U = \sum_{r=1}^s u_r \hat{Y}_{ro}^U \\
& s.t. \\
& \sum_{j=1}^m v_j x_{jo} = 1 \\
& \sum_{r=1}^s u_r \hat{Y}_{ri}^L - \sum_{j=1}^m v_j x_{ji} \leq 0; \quad i = 1, \dots, n \\
& u_r, v_j \geq \varepsilon \quad \forall r, j
\end{aligned} \tag{4.29}$$

En el modelo (4.29) observamos que los pesos óptimos no pueden garantizar unos extremos superiores de los intervalos de confianza de eficiencia menores o iguales que la unidad. Sin importar los pesos que usan los modelos, las eficacias de las DMUs están todas limitadas a ser menor o igual a uno.

En los modelos (4.28) y (4.29) se consideran tanto la mejor como la peor actividad de producción de la DMU_o. Esto se puede entender desde el siguiente aspecto: La frontera de producción está determinada por la mejor (peor) actividad de producción de las n DMUs sin tener en cuenta la peor (mejor) actividad de producción. Si la mejor (peor) actividad de la producción de la DMU_o se quita del modelo, la frontera de la producción cambiaría y sería diferente de DMU a DMU lo que produciría que los límites de eficiencia obtenidos sean incomparables.

4.3.4. Intervalo de confianza Tipo4

Los tres ICEs basados en el punto de vista optimista/pesimista introducidos hasta ahora tienen, en principio, un inconveniente, la confianza de que el intervalo contenga al verdadero valor de la eficiencia depende de la calidad de la estimación del output y, salvo por simulaciones, no podemos evaluar esta confianza; no disponemos de ningún resultado analítico que ponga en relación el tamaño de la muestra en cada DMU y la confianza que se logra con el ICE. Para resolver este problema introducimos el ICE *Tipo4* basado en el Teorema 1 que, como se mostrará en la sección 4.3.4, tiene una evidente similitud con el ICE *Tipo3*. El Teorema 1 demuestra, en el caso de n DMUs con un único input con valores conocidos $(X_i)_{i=1, \dots, n}$, un único output con valores $(Y_i)_{i=1, \dots, n}$ desconocidos, $(\varphi_i)_{i=1, \dots, n}$ los índices de eficiencia poblacionales CCR-O (desconocidos), $(\hat{Y}_i)_{i=1, \dots, n}$ y $(\hat{\Phi}_i)_{i=1, \dots, n}$ los estimadores de los outputs y de los índices de eficiencia, respectivamente, y si

$$p(|\hat{Y}_i - Y_i| \leq pY_i) \geq 1 - \alpha; \quad \forall i = 1, \dots, n, \quad \text{con } p \in (0, 1),$$

entonces el intervalo

$$I_i = \left[\hat{\Phi}_i \frac{(1-p)}{(1+p)}, \min \left\{ \hat{\Phi}_i \frac{(1+p)}{(1-p)}, 1 \right\} \right]$$

tiene una probabilidad $(1 - \alpha)^3$ de contener a φ_i .

Este resultado se mantiene, por simulación, para el modelo BCC-O con más de un input y más de un output. En consecuencia, si consideramos n DMUs, m inputs conocidos, s outputs desconocidos y tomamos un tamaño de muestra n_i en la i th DMU ($i = 1, \dots, n$) para que el error de estimación de los s outputs verifiquen

$$p(|\hat{Y}_{ri} - Y_{ri}| \leq pY_{ri}) \geq 1 - \alpha; \quad \forall r = 1, \dots, s \quad \text{con } p \in (0, 1)$$

entonces proponemos como intervalo de confianza de eficiencia DEA para el índice de eficiencia poblacional φ_i , $i = 1, \dots, n$, el intervalo

$$I_i = \left[\hat{\Phi}_i \frac{(1-p)}{(1+p)}, \min \left\{ \hat{\Phi}_i \frac{(1+p)}{(1-p)}, 1 \right\} \right] \tag{4.30}$$

que, mediante simulación, se comprueba que es un intervalo de confianza $(1 - \alpha)^3$, siendo $(\hat{\Phi}_i)_{i=1, \dots, n}$ los correspondientes estimadores de los índices de eficiencia poblacionales.

Relación entre intervalos de confianza Tipo3 y Tipo4

El Teorema 3 nos permite poner en relación los ICEs Tipo3 y Tipo4.

Teorema 3 Consideremos n DMUs, un input $\{X_i\}_{i=1, \dots, n}$ conocido, un output $\{Y_i\}_{i=1, \dots, n}$, desconocido, cuyos estimadores son $\{\hat{Y}_i\}_{i=1, \dots, n}$ basado en muestras de tamaño n_i en cada DMU verificando

$$p(|\hat{Y}_i - Y_i| \leq e_i) \geq 1 - \alpha$$

donde $e_i = pY_i$ es un error de estimación, $p \in (0, 1)$, y $1 - \alpha$ la confianza. En estas circunstancias, la frontera de producción para obtener el extremo superior del ICE Tipo4 es

$$T^{supT1} = \{(X_1, \hat{Y}_1 - \hat{e}_1), \dots, (X_{i-1}, \hat{Y}_{i-1} - \hat{e}_{i-1}), (X_i, \hat{Y}_i - \hat{e}_i), (X_{i+1}, \hat{Y}_{i+1} - \hat{e}_{i+1}), \dots, (X_n, \hat{Y}_n - \hat{e}_n)\}$$

y para obtener el extremo inferior

$$T^{infT1} = \{(X_1, \hat{Y}_1 + \hat{e}_1), \dots, (X_{i-1}, \hat{Y}_{i-1} + \hat{e}_{i-1}), (X_i, \hat{Y}_i + \hat{e}_i), (X_{i+1}, \hat{Y}_{i+1} + \hat{e}_{i+1}), \dots, (X_n, \hat{Y}_n + \hat{e}_n)\}$$

siendo $\hat{e}_i = p\hat{Y}_i$, $i = 1, \dots, n$, un estimador del error de estimación con la muestra de tamaño n_i .

DEMOSTRACIÓN

El estimador del índice de eficiencia es

$$\hat{\Phi}_i = \frac{\frac{\hat{Y}_i}{X_i}}{\max_{k=1, \dots, n} \left\{ \frac{\hat{Y}_k}{X_k} \right\}}$$

y el ICE Tipo4 para la i th DMU es

$$\left[\hat{\Phi}_i \frac{(1-p)}{(1+p)}, \min \left\{ \hat{\Phi}_i \frac{(1+p)}{(1-p)}, 1 \right\} \right] = \left[\frac{\frac{\hat{Y}_i - p\hat{Y}_i}{X_i}}{\max_{k=1, \dots, n} \left\{ \frac{\hat{Y}_k + p\hat{Y}_k}{X_k} \right\}}, \min \left\{ 1, \frac{\frac{\hat{Y}_i + p\hat{Y}_i}{X_i}}{\max_{k=1, \dots, n} \left\{ \frac{\hat{Y}_k - p\hat{Y}_k}{X_k} \right\}} \right\} \right]$$

Por tanto, para obtener el extremo inferior del ICE Tipo4 de la i th DMU se toma su valor output $\hat{Y}_i - \hat{e}_i$ y los valores de la frontera de producción

$$T^{infT1} = \{(X_1, \hat{Y}_1 + \hat{e}_1), \dots, (X_{i-1}, \hat{Y}_{i-1} + \hat{e}_{i-1}), (X_i, \hat{Y}_i + \hat{e}_i), (X_{i+1}, \hat{Y}_{i+1} + \hat{e}_{i+1}), \dots, (X_n, \hat{Y}_n + \hat{e}_n)\}$$

mientras que, para obtener el extremo superior del ICE Tipo4 de la i th DMU se toma su valor output $\hat{Y}_i + \hat{e}_i$ y los valores de la frontera de producción

$$T^{supT1} = \{(X_1, \hat{Y}_1 - \hat{e}_1), \dots, (X_{i-1}, \hat{Y}_{i-1} - \hat{e}_{i-1}), (X_i, \hat{Y}_i - \hat{e}_i), (X_{i+1}, \hat{Y}_{i+1} - \hat{e}_{i+1}), \dots, (X_n, \hat{Y}_n - \hat{e}_n)\}$$

En conclusión, si consideramos la notación

$$\begin{aligned} \hat{Y}_i^L &= \hat{Y}_i - \hat{e}_i; \quad i = 1, \dots, n \\ \hat{Y}_i^U &= \hat{Y}_i + \hat{e}_i; \quad i = 1, \dots, n \end{aligned} \quad (4.31)$$

los modelos DEA para obtener los extremos del ICE *Tipo4* de la DMU_o son

$$\begin{aligned}
 & \max_{u,v} T_o^L = u \hat{Y}_o^L \\
 & \text{s.t.} \\
 & vx_o = 1 \\
 & u \hat{Y}_i^U - vx_i \leq 0; \quad i = 1, \dots, n \\
 & u, v \geq \varepsilon
 \end{aligned} \tag{4.32}$$

$$\begin{aligned}
 & \max_{u,v} T_o^U = u \hat{Y}_o^U \\
 & \text{s.t.} \\
 & vx_o = 1 \\
 & u \hat{Y}_i^L - vx_i \leq 0; \quad i = 1, \dots, n \\
 & u, v \geq \varepsilon
 \end{aligned} \tag{4.33}$$

que coinciden con los modelos DEA (4.28) y (4.29) para obtener los extremos del ICE *Tipo3* de la DMU_o en la situación de un input $(X_i)_{i=1, \dots, n}$ conocido y un output estocástico estimado con el intervalo de confianza $[\hat{Y}_i^L, \hat{Y}_i^U]$, $i = 1, \dots, n$, como en (4.31).

ESTUDIO DE SIMULACIÓN

A partir de un estudio de simulación comprobamos que los modelos propuestos para obtener los extremos del ICE *Tipo3* determinan extremos muy similares a los del ICE *Tipo4* en el caso de más de un input y más de un output. La Tabla 4.2 contiene los datos utilizados en Cooper y otros (2006) correspondientes al número de médicos y enfermeras (inputs) y del número de altas e ingresos (outputs) en 12 centros sanitarios y la Tabla 4.3 contiene el modelo poblacional simulado que vamos a utilizar.

Los pasos del proceso de simulación para comprobar la similitud entre los extremos de los intervalos *Tipo3* y *Tipo4* son:

- I. En la i th DMU, $i = 1, \dots, 12$, se toman los datos disponibles, por simulación, de media poblacional $Y_i = (Y_{1i}, Y_{2i})$ y varianza poblacional $\sigma_i^{2(0)} = (\sigma_{1i}^{2(0)}, \sigma_{2i}^{2(0)})$ y, fijados $p = 0.05$ y $\alpha = 0.1$, se determina el tamaño de muestra n_i , como en (4.17) y (4.16) para que se verifique

$$p(|\hat{Y}_{ji} - Y_{ji}| \leq pY_{ji}) \geq 1 - \alpha; \quad j = 1, 2; \quad i = 1, \dots, 12$$

- II. Se toma la muestra en cada DMU y sean $\hat{y}_i = (\hat{y}_{1i}, \hat{y}_{2i})$, $i = 1, \dots, 12$, los correspondientes valores estimados de los outputs.
- III. Con los datos $(X_{1i}, X_{2i}, \hat{y}_{1i}, \hat{y}_{2i})_{i=1, \dots, 12}$ se obtienen los ICEs, $\{I_i\}_{i=1, \dots, 12}$ *Tipo3* y *Tipo4*
- IV. Los apartados II.-III. se repiten mil veces, con lo que para cada tipo de intervalo, *Tipo3* y *Tipo4*, se obtienen $\left\{ I_i^{(k)} \right\}_{i=1, \dots, 12, k=1, \dots, 1000}$.
- V. En cada iteración, se obtiene las diferencias, en cada DMU, entre los extremos inferiores (superiores) de los intervalos *Tipo3* y *Tipo4* y representamos los boxplots de las mil diferencias entre extremos.

Las Figuras 4.1, 4.2, 4.3 y 4.4 muestran los boxplots de las diferencias entre los extremos inferiores (superiores) de los intervalos *Tipo3* y *Tipo4* en cada iteración en cada DMU, en los modelos DEA CCR-O y BCC-O, con objeto de analizar si esta diferencia es nula; esto significaría que los intervalos obtenidos en cada iteración *Tipo3* y *Tipo4* son muy similares.

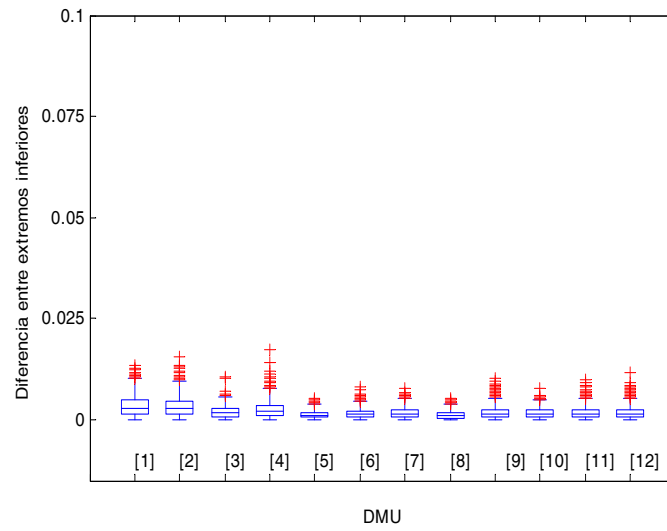


Figura 4.1: Diferencias absolutas entre extremos inferiores de ICEs *Tipo3* y *Tipo4*. Modelo CCR.

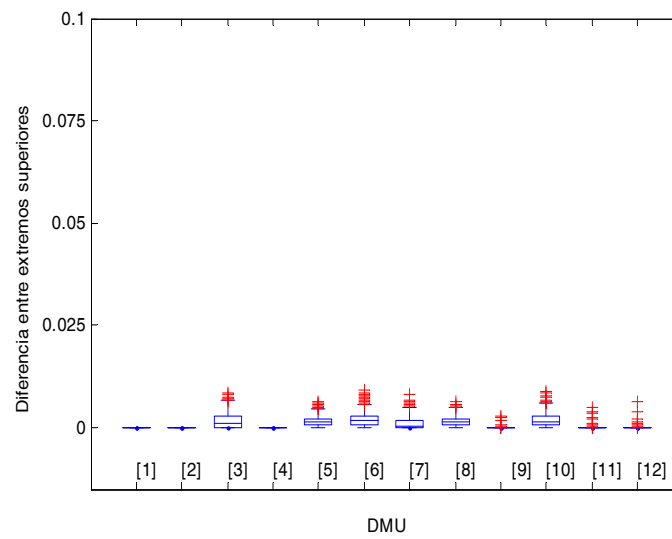


Figura 4.2: Diferencias absolutas entre extremos superiores de ICEs *Tipo3* y *Tipo4*. Modelo CCR.

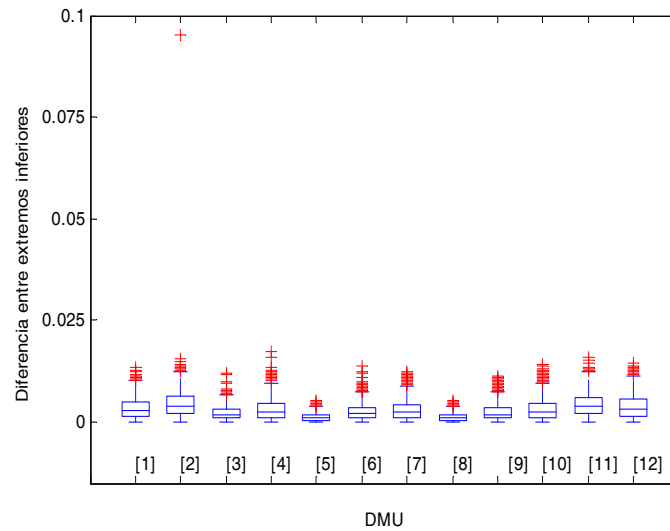


Figura 4.3: Diferencias absolutas entre extremos inferiores de ICEs *Tipo3* y *Tipo4*. Modelo BCC.

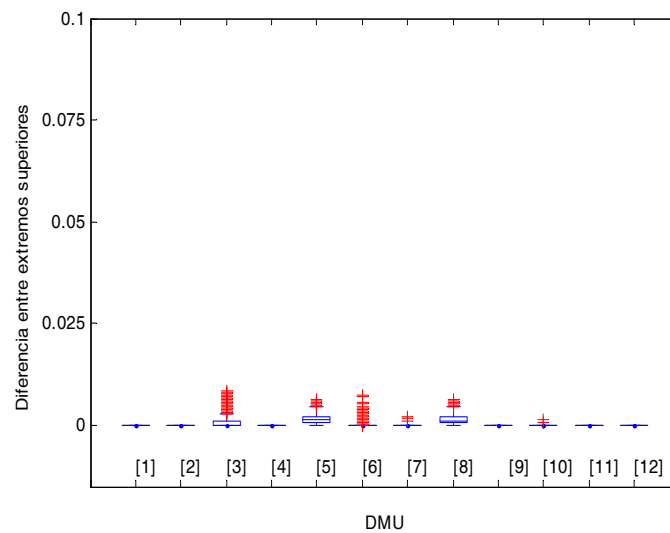


Figura 4.4: Diferencias absolutas entre extremos superiores de ICEs *Tipo3* y *Tipo4*. Modelo BCC.

Las Tablas 4.9 y 4.10 muestran un ICE *Tipo3* y *Tipo4* para cada DMU, para los modelos CCR y BCC respectivamente.

Tabla 4.9: Extremo inferior y superior de los ICEs *Tipo3* y *Tipo4* fijado $p = 0.05$ y $\alpha = 0.1$. Modelo CCR.

| DMU | $\hat{\phi}_i$ | ICE <i>Tipo3</i> | | ICE <i>Tipo4</i> | |
|-----|----------------|------------------|-----------|------------------|-----------|
| | | Ext. Inf | Ext. Sup. | Ext. Inf | Ext. Sup. |
| 1 | 1 | 0.905 | 1 | 0.910 | 1 |
| 2 | 1 | 0.905 | 1 | 0.911 | 1 |
| 3 | 0.851 | 0.770 | 0.941 | 0.771 | 0.939 |
| 4 | 1 | 0.905 | 1 | 0.905 | 1 |
| 5 | 0.791 | 0.716 | 0.874 | 0.718 | 0.872 |
| 6 | 0.806 | 0.729 | 0.891 | 0.726 | 0.894 |
| 7 | 0.925 | 0.837 | 1 | 0.841 | 1 |
| 8 | 0.778 | 0.704 | 0.860 | 0.702 | 0.861 |
| 9 | 0.945 | 0.855 | 1 | 0.854 | 1 |
| 10 | 0.879 | 0.795 | 0.972 | 0.792 | 0.976 |
| 11 | 0.948 | 0.858 | 1 | 0.855 | 1 |
| 12 | 0.962 | 0.871 | 1 | 0.873 | 1 |

Tabla 4.10: Extremo inferior y superior de los ICEs *Tipo3* y *Tipo4* fijado $p = 0.05$ y $\alpha = 0.1$. Modelo BCC.

| DMU | $\hat{\phi}_i$ | ICE <i>Tipo3</i> | | ICE <i>Tipo4</i> | |
|-----|----------------|------------------|-----------|------------------|-----------|
| | | Ext. Inf | Ext. Sup. | Ext. Inf | Ext. Sup. |
| 1 | 1 | 0.905 | 1 | 0.910 | 1 |
| 2 | 1 | 0.905 | 1 | 0.914 | 1 |
| 3 | 0.898 | 0.813 | 0.993 | 0.815 | 0.989 |
| 4 | 1 | 0.905 | 1 | 0.908 | 1 |
| 5 | 0.796 | 0.720 | 0.880 | 0.722 | 0.877 |
| 6 | 0.902 | 0.816 | 0.997 | 0.811 | 1 |
| 7 | 1 | 0.905 | 1 | 0.908 | 1 |
| 8 | 0.806 | 0.729 | 0.890 | 0.727 | 0.893 |
| 9 | 0.971 | 0.878 | 1 | 0.876 | 1 |
| 10 | 1 | 0.905 | 1 | 0.893 | 1 |
| 11 | 1 | 0.905 | 1 | 0.898 | 1 |
| 12 | 1 | 0.905 | 1 | 0.908 | 1 |

A la vista de los resultados, se confirma la similitud entre los intervalos *Tipo3* y *Tipo4*.

4.3.5. Intervalo de confianza Bootstrap

El bootstrap es un método para aproximar la distribución de estadísticos cuando esta es difícil o imposible de determinar analíticamente. La idea general es usar remuestreo para estimar una

distribución empírica para el parámetro de interés (Efron 1994 [50]). En esencia, en bootstrap se usa la potencia computacional como sustituto del cálculo.

Para el problema planteado en la sección 4.1, donde se consideran m inputs conocidos $\mathbf{X}_i = (X_{1i}, \dots, X_{mi})_{i=1, \dots, n}$ y una muestra p_i de tamaño n_i con la que estimamos s outputs $\{\mathbf{Y}_i = (Y_{1i}, \dots, Y_{si})\}_{i=1, \dots, n}$, los pasos del procedimiento de obtención del ICE bootstrap son:

- I. En la i th DMU, consideremos una muestra bootstrap de tamaño n_i con la que obtenemos la versión bootstrap de los estimadores de los outputs, $\{\hat{\mathbf{Y}}_i^* = (\hat{Y}_{1i}^*, \dots, \hat{Y}_{si}^*)\}_{i=1, \dots, n}$, y sean $\{\hat{\mathbf{y}}_i^* = (\hat{y}_{1i}^*, \dots, \hat{y}_{si}^*)\}_{i=1, \dots, n}$ los correspondientes valores estimados.
- II. Con los datos $(\mathbf{X}_i, \hat{\mathbf{y}}_i^*)_{i=1, \dots, n}$, usando los modelos (4.13) o (4.14), obtenemos las eficiencias bootstrap estimadas $(\hat{\phi}_i^*)_{i=1, \dots, n}$.
- III. Los pasos I-II. se deberían repetir B veces y las B eficiencias bootstrap estimadas para la i th DMU, $i = 1, \dots, n$, deberían ser $\{\hat{\phi}_{ib}^*\}_{b=1, \dots, B}$.
- IV. Fijada una intención de cobertura de nivel $1 - \alpha'$, el ICE bootstrap percentil observado es

$$\left(\hat{\phi}_i^{*(\alpha'/2)}, \hat{\phi}_i^{*(1-\alpha'/2)} \right); i = 1, \dots, n$$

donde $\hat{\phi}_i^{*(\alpha)}$ es el α -percentil de los B valores $\hat{\phi}_{ib}^*$.

4.3.6. Intervalo observado

Fijada una muestra en cada DMU, los intervalos (4.12) son los intervalos de confianza observados para los outputs. Si usamos estos intervalos en los modelos (4.24)-(4.25), (4.26)-(4.27) y (4.28)-(4.29), obtenemos el intervalo de confianza observado $[d_o^L, d_o^U]$ de *Tipo1*, $[w_o^L, w_o^U]$ de *Tipo2*, y $[t_o^L, t_o^U]$ de *Tipo3*, respectivamente. En esta situación, Wang y otros (2005) [146] demuestran la relación:

$$w_o^L \leq d_o^L \text{ y } w_o^U \leq d_o^U$$

Teorema 4 Sean $d_o^{L*} - d_o^{U*}$, $w_o^{L*} - w_o^{U*}$ y $t_o^{L*} - t_o^{U*}$ los valores óptimos de la función objetivo de los modelos (4.24)-(4.25), (4.26)-(4.27) y (4.28)-(4.29), respectivamente, tomando $\hat{Y}_{ri}^L = y_{ri}^L$ y $\hat{Y}_{ri}^U = y_{ri}^U$; $r = 1, \dots, s$; $i = 1, \dots, n$. Entonces

$$w_o^{L*} = t_o^{L*} \leq d_o^{L*} \text{ y } w_o^{U*} \leq d_o^{U*} \leq t_o^{U*}$$

DEMOSTRACIÓN

Teniendo en cuenta que $w_o^{L*} = t_o^{L*}$, sólo falta demostrar que $d_o^{U*} \leq t_o^{U*}$:

Sean $\{u_r^*\}_{r=1, \dots, s}$ y $\{v_j^*\}_{j=1, \dots, m}$ las soluciones óptimas al modelo (4.25). Entonces, específicamente para la DMU _{o} , tenemos

$$\sum_{r=1}^s u_r^* y_{ro}^U - \sum_{j=1}^m v_j^* x_{jo} \leq 0$$

$$\sum_{r=1}^s u_r^* y_{ro}^L - \sum_{j=1}^m v_j^* x_{jo} \leq \sum_{r=1}^s u_r^* y_{ro}^U - \sum_{j=1}^m v_j^* x_{jo} \leq 0$$

como

$$\sum_{r=1}^s u_r^* y_{ri}^L - \sum_{j=1}^m v_j^* x_{ji} \leq 0; i \neq o$$

lo que significa que $\{u_r^*\}_{r=1,\dots,s}$ y $\{v_j^*\}_{j=1,\dots,m}$ son una solución factible del modelo (4.29). En consecuencia

$$d_o^{U^*} = \sum_{r=1}^s u_r^* y_{ro}^U = t_o^U \leq t_o^{U^*}.$$

4.3.7. Evaluación de los métodos de obtención de Intervalos de confianza DEA

Con objeto de analizar la confianza y la amplitud de los intervalos y comparar las distintas metodologías de construcción de ICEs realizaremos un estudio de simulación. Las únicas confianzas que se pueden fijar a priori son la que permite calcular el tamaño de muestra necesario para estimar los outputs, α , y la confianza de los ICEs bootstrap, α' .

UN ESTUDIO DE SIMULACIÓN

Frecuentemente, la opinión de un individuo, recogida en un cuestionario, se sintetiza en uno o más índices de opinión y la media de estos índices, para una muestra de individuos de una DMU, se puede utilizar como outputs de la DMU. Por tanto, si U_{kri} es la respuesta del individuo k al r th índice de opinión de la i th DMU, el r th output de la i th DMU es la media poblacional, desconocida, de la respuesta de los N_i individuos a ese índice:

$$\left\{ Y_{ri} = \frac{\sum_{k=1}^{N_i} U_{kri}}{N_i} \right\} \begin{matrix} i = 1, \dots, n \\ r = 1, \dots, s \end{matrix} \quad (4.34)$$

Si tomamos una muestra, por ejemplo aleatoria simple sin reemplazamiento, en cada DMU, de tamaño n_i , siendo V_{kri} la respuesta del individuo k de la muestra al r th índice de opinión de la i th DMU, los outputs se estiman como

$$\left\{ \hat{Y}_{ri} = \frac{\sum_{k=1}^{n_i} V_{kri}}{n_i} \right\} \begin{matrix} i = 1, \dots, n \\ r = 1, \dots, s \end{matrix} \quad (4.35)$$

Consideremos los datos de la Tabla 4.2 utilizados en Cooper y otros (2006) [39] correspondientes al número de médicos y enfermeras (inputs) y al número de altas e ingresos (outputs) en 12 centros de salud. Estos datos se usan para generar el modelo poblacional simulado de la Tabla 4.3. La aproximación de la confianza de los intervalos DEA (ICEs), obtenidos con las distintas metodologías, así como el análisis de sus amplitudes, sigue los siguientes pasos:

- I. Fijado el error de estimación $e = 0.9$ y la confianza $1 - \alpha = 0.9$ en la i th DMU, calculamos, usando muestreo aleatorio simple sin reemplazamiento, el tamaño de muestra n_i , como en (4.17) y (4.16), considerando la varianza poblacional, σ_{ri}^2 , $r = 1, 2$, conocida para que se verifique

$$p(|\hat{Y}_{ri} - Y_{ri}| \leq e) \geq 1 - \alpha; \quad r = 1, 2; \quad i = 1, \dots, 12.$$

suponiendo \hat{Y}_{ri} definido como en (4.35).

- II. Se toma la muestra de tamaño n_i en la i th DMU y sea \hat{y}_{ri} ; $r = 1, 2$; $i = 1, \dots, 12$ el valor estimado de los outputs como en (4.10).
- III. Con los datos $(X_{1i}, X_{2i}, \hat{y}_{1i}, \hat{y}_{2i})_{i=1,\dots,12}$, se obtienen los intervalos de confianza observados para la eficiencia de cada DMU, $\{I_i\}_{i=1,\dots,12}$, *Tipo1*, *Tipo2*, *Tipo3* y bootstrap.

IV. Los pasos II-III. se repiten mil veces, y se obtiene, para cada método:

$$\left\{ I_i^{(k)} = [I_i^{(k)L}, I_i^{(k)U}] \right\} \begin{array}{l} i = 1, \dots, 12 \\ k = 1, \dots, 1000 \end{array}$$

V. Se aproxima la confianza de los intervalos con

$$C_i = \left(\frac{1}{1000} \sum_{k=1}^{1000} I_{(\varphi_i \in I_i^{(k)})} \right) \times 100$$

VI. El valor esperado de los extremos de los intervalos se aproxima con

$$E(I_i^L) = \frac{\sum_{k=1}^{1000} I_i^{(k)L}}{1000} \text{ y } E(I_i^U) = \frac{\sum_{k=1}^{1000} I_i^{(k)U}}{1000}$$

RESULTADOS DE LA SIMULACIÓN

La Tabla 4.11 contiene los tamaños muestrales obtenidos para cada DMU utilizando (4.17) y (4.16), para los valores $e = 0.9$ y $\alpha = 0.1$.

Tabla 4.11: Tamaño muestral obtenido para cada DMU fijado un error de estimación para los outputs $e=0.9$ y una confianza $1-\alpha=0.9$

| <u>n_i</u> |
|-----------------------------|
| 83 |
| 190 |
| 214 |
| 270 |
| 75 |
| 436 |
| 395 |
| 192 |
| 300 |
| 512 |
| 547 |
| <u>518</u> |

Las Tablas 4.12 y 4.13 muestran las confianzas aproximadas de los intervalos para las distintas metodologías en los modelos CCR y BCC con orientación output, fijados $e = 0.9$ y $\alpha = 0.1$. Se han seleccionado dos valores, $(1 - \alpha') = 0.98$ y $(1 - \alpha') = 0.9$, para la confianza de los ICEs bootstrap. Las confianzas simuladas de las metodologías de ICEs *Tipo1*, *Tipo3* y bootstrap, $(1 - \alpha') = 0.98$, son bastante similares. La confianza de los ICEs *Tipo3* es algo superior en todos los casos. Los intervalos *Tipo2* tienen una confianza muy baja y, para el tamaño de muestra obtenido, se descartan como opción.

Tabla 4.12: Confianza simulada de los ICEs. Modelo CCR-O.

| DMU | $C_i(\%)$ | | | | |
|-----------|--------------|--------------|--------------|--|---|
| | <i>Tipo1</i> | <i>Tipo2</i> | <i>Tipo3</i> | Bootstrap $1-\alpha'=0.98$ | Bootstrap $1-\alpha'=0.9$ |
| 1 | 100 | 100 | 100 | 100 | 100 |
| 2 | 100 | 100 | 100 | 100 | 100 |
| 3 | 99.4 | 47.3 | 99.4 | 98.2 | 91.3 |
| 4 | 100 | 50.7 | 100 | 99.4 | 97.9 |
| 5 | 99.8 | 59.3 | 99.8 | 96.9 | 89 |
| 6 | 100 | 19.3 | 100 | 98.2 | 91.5 |
| 7 | 99.8 | 19.4 | 99.8 | 97.5 | 89.3 |
| 8 | 100 | 36.3 | 100 | 98 | 90 |
| 9 | 99.9 | 21.4 | 100 | 98 | 89.8 |
| 10 | 100 | 14 | 100 | 99.2 | 89.9 |
| 11 | 99.9 | 7.5 | 99.9 | 98.3 | 91.4 |
| 12 | 100 | 12.4 | 100 | 98 | 88.6 |

Tabla 4.13: Confianza simulada de los ICEs. Modelo BCC-O.

| DMU | $C_i(\%)$ | | | | |
|-----------|--------------|--------------|--------------|--|---|
| | <i>Tipo1</i> | <i>Tipo2</i> | <i>Tipo3</i> | Bootstrap $1-\alpha'=0.98$ | Bootstrap $1-\alpha'=0.9$ |
| 1 | 100 | 100 | 100 | 100 | 100 |
| 2 | 100 | 100 | 100 | 100 | 100 |
| 3 | 99.3 | 47.5 | 99.8 | 97.3 | 88.8 |
| 4 | 100 | 86.7 | 100 | 100 | 100 |
| 5 | 99.7 | 63.6 | 99.7 | 97.2 | 88.9 |
| 6 | 99.4 | 48.5 | 99.8 | 97.8 | 90.1 |
| 7 | 99.2 | 46.8 | 99.4 | 97 | 89.6 |
| 8 | 100 | 60.2 | 100 | 98.6 | 91.4 |
| 9 | 100 | 48.1 | 100 | 98 | 91.1 |
| 10 | 100 | 76 | 100 | 99.9 | 98.9 |
| 11 | 100 | 100 | 100 | 100 | 100 |
| 12 | 100 | 100 | 100 | 100 | 100 |

Las Tablas 4.14 y 4.15 muestran la aproximación de los valores esperados de los extremos de los intervalos para las cuatro tipologías de ICEs tomando $\alpha = 0.1$ y $(1 - \alpha') = 0.98$, en los modelos CCR-O y BCC-O respectivamente. En todos los casos se observa que los extremos inferiores de los ICEs *Tipo3* son menores o iguales que los extremos inferiores de los ICEs *Tipo1*. Los valores esperados de los ICEs bootstrap son los “mejores” observándose, en todos los casos, un extremo superior igual o más pequeño y un extremo inferior igual o más grande.

De cara a utilizar la información contenida en estas dos Tablas, vamos a considerar la clasificación de n DMUs, propuesta por Despotis and Smirlis (2002), cuando se dispone de un intervalo de eficiencia

$[I_i^L, I_i^U]_{i=1,\dots,n}$, para cada DMU, en los conjuntos

$$\begin{aligned} E^{++} &= \left\{ j \in J = \{1, \dots, n\} / I_j^L = 1 \right\} \\ E^+ &= \left\{ j \in J / I_j^L < 1 \text{ and } I_j^U = 1 \right\} \\ E^- &= \left\{ j \in J / I_j^U < 1 \right\} \end{aligned} \quad (4.36)$$

El conjunto E^{++} consiste de las unidades que son eficientes en cualquier caso, el conjunto E^+ está compuesto por la unidades que son eficientes en un sentido de máximo pero que existen ajustes de input/output bajo los cuales no puede mantenerse la eficiencia, y, finalmente, el conjunto E^- compuesto por las unidades definitivamente ineficientes. Atendiendo a esta clasificación, observamos que los métodos bootstrap y *Tipo1* detectan las DMUs eficientes en cualquier caso.

Tabla 4.14: Valor esperado de los extremos de los intervalos de confianza de eficiencia ICEs, $\alpha=0.1$ y $1-\alpha'=0.98$. Modelo CCR

| DMU | <i>Tipo1</i> | | <i>Tipo2</i> | | <i>Tipo3</i> | | Bootstrap | |
|-----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | E (I_i^L) | E (I_i^U) | E (I_i^L) | E (I_i^U) | E (I_i^L) | E (I_i^U) | E (I_i^L) | E (I_i^U) |
| 1 | 1 | 1 | 0.838 | 1 | 0.838 | 1 | 1 | 1 |
| 2 | 0.997 | 1 | 0.889 | 1 | 0.889 | 1 | 0.997 | 1 |
| 3 | 0.791 | 0.982 | 0.791 | 0.885 | 0.791 | 0.982 | 0.816 | 0.968 |
| 4 | 0.904 | 1 | 0.894 | 0.988 | 0.894 | 1 | 0.942 | 1 |
| 5 | 0.639 | 0.908 | 0.639 | 0.772 | 0.639 | 0.908 | 0.676 | 0.870 |
| 6 | 0.757 | 0.914 | 0.757 | 0.818 | 0.757 | 0.914 | 0.780 | 0.882 |
| 7 | 0.813 | 0.989 | 0.813 | 0.883 | 0.813 | 0.989 | 0.843 | 0.965 |
| 8 | 0.700 | 0.902 | 0.700 | 0.788 | 0.700 | 0.902 | 0.734 | 0.857 |
| 9 | 0.857 | 1 | 0.857 | 0.941 | 0.857 | 1 | 0.898 | 0.998 |
| 10 | 0.792 | 0.951 | 0.792 | 0.851 | 0.792 | 0.951 | 0.816 | 0.917 |
| 11 | 0.861 | 1 | 0.861 | 0.923 | 0.861 | 1 | 0.895 | 0.995 |
| 12 | 0.865 | 1 | 0.865 | 0.929 | 0.865 | 1 | 0.902 | 0.997 |

Tabla 4.15: Valor esperado de los extremos de los intervalos de confianza de eficiencia ICEs, $\alpha=0.1$ y $1-\alpha'=0.98$. Modelo BCC

| DMU | <i>Tipo1</i> | | <i>Tipo2</i> | | <i>Tipo3</i> | | Bootstrap | |
|-----------|--------------|------------|--------------|------------|--------------|------------|------------|------------|
| | $E(I_i^L)$ | $E(I_i^U)$ | $E(I_i^L)$ | $E(I_i^U)$ | $E(I_i^L)$ | $E(I_i^U)$ | $E(I_i^L)$ | $E(I_i^U)$ |
| 1 | 1 | 1 | 0.838 | 1 | 0.838 | 1 | 1 | 1 |
| 2 | 1 | 1 | 0.892 | 1 | 0.892 | 1 | 1 | 1 |
| 3 | 0.827 | 0.994 | 0.827 | 0.925 | 0.827 | 0.994 | 0.839 | 0.987 |
| 4 | 0.936 | 1 | 0.904 | 0.998 | 0.904 | 1 | 0.967 | 1 |
| 5 | 0.645 | 0.911 | 0.645 | 0.780 | 0.645 | 0.911 | 0.682 | 0.874 |
| 6 | 0.885 | 0.998 | 0.884 | 0.956 | 0.884 | 0.998 | 0.892 | 0.995 |
| 7 | 0.920 | 1 | 0.910 | 0.988 | 0.910 | 1 | 0.930 | 1 |
| 8 | 0.738 | 0.921 | 0.738 | 0.831 | 0.738 | 0.921 | 0.770 | 0.879 |
| 9 | 0.902 | 1 | 0.898 | 0.986 | 0.898 | 1 | 0.937 | 1 |
| 10 | 0.950 | 1 | 0.928 | 0.997 | 0.928 | 1 | 0.953 | 1 |
| 11 | 1 | 1 | 0.934 | 1 | 0.934 | 1 | 1 | 1 |
| 12 | 1 | 1 | 0.932 | 1 | 0.932 | 1 | 1 | 1 |

Las Figuras 4.5 y 4.6 representan la amplitud simulada de los mil intervalos, para cada metodología de construcción de ICE en los modelos CCR y BCC, con orientación output, respectivamente. Si en el modelo BCC-O nos fijamos en las unidades eficientes poblacionalmente $\{1, 2, 11, 12\}$ observamos una amplitud nula en los ICEs *Tipo1* y bootstrap, por tanto, estos dos métodos detectan las unidades eficientes en cualquier caso. En estas mismas cuatro DMUs también se observa una gran similitud en las amplitudes de los ICEs *Tipo2* y *Tipo3*. Para el resto de DMUs, los ICEs *Tipo2* son los de menor amplitud y la amplitud de los ICEs bootstrap es algo menor que la amplitud de los ICEs *Tipo1* y *Tipo3*.

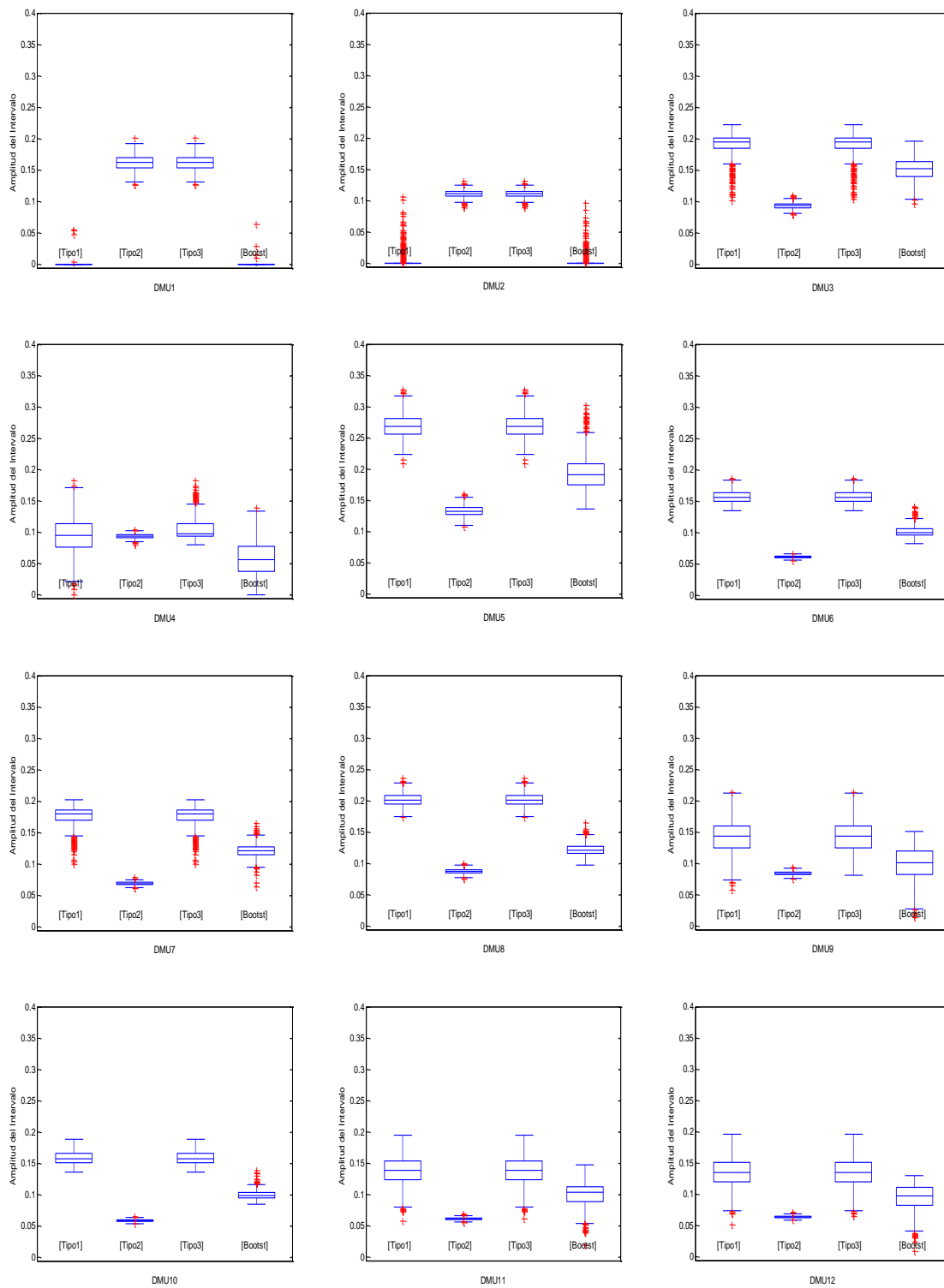


Figura 4.5: Amplitud de intervalos ICEs por tipo y DMU. Modelo CCR-O, $\alpha = 0.1$ y $1 - \alpha' = 0.98$

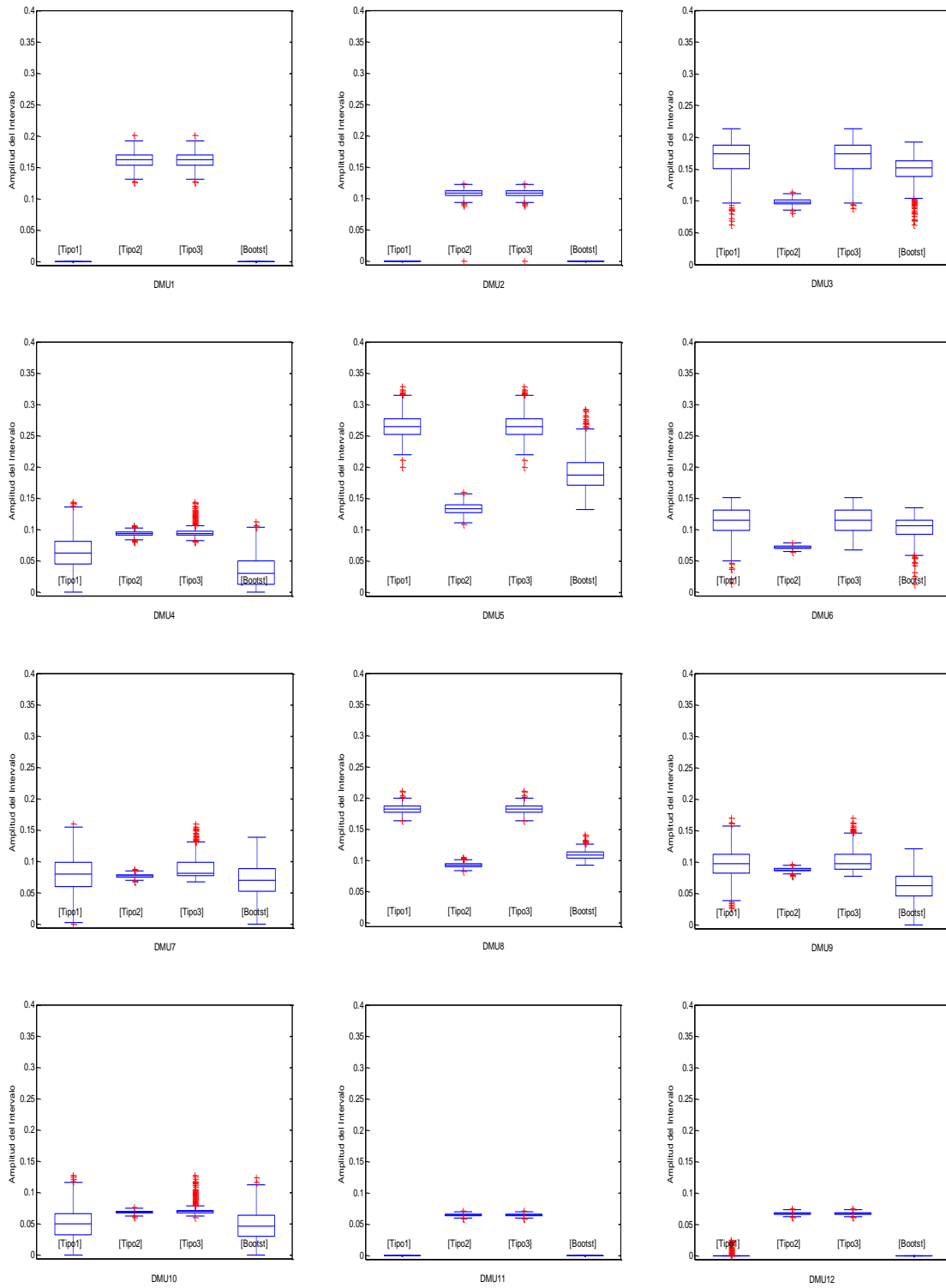


Figura 4.6: Amplitud de intervalos ICEs por tipo y DMU. Modelo BCC-O, $\alpha = 0.1$ y $1 - \alpha' = 0.98$

En conclusión, el método de ICE bootstrap tiene tres ventajas: la primera, el control a priori de α' conduce a lograr la confianza de los intervalos de eficiencia deseada por el experimentador. La segunda, permite detectar las unidades eficientes en cualquier caso y, por último, descartada la opción *Tipo2* por su baja confianza, los ICEs bootstrap son los de menor amplitud y “mejor” valor esperado de sus extremos. Comparando los métodos que utilizan el punto de vista optimista/pesimista en su construcción, los ICEs *Tipo1* y *Tipo3* tienen una confianza razonable y son muy similares, sin embargo el *Tipo3* tiene la ventaja de usar sólo dos fronteras de producción, una para obtener el extremo superior y otra para el extremo inferior.

4.3.8. Un estudio de aplicación

En este apartado, examinamos la utilidad de los intervalos de confianza de eficiencia DEA propuestos en este capítulo en un caso real: El Ministerio de Sanidad de España lleva realizando desde hace tiempo dos estadísticas, la “Estadística de Establecimientos Sanitarios con Régimen de Internado” (EESRI) y el “Barómetro Sanitario” (BS). En la primera, se recoge una amplísima información del sistema sanitario, por ejemplo, indicadores de personal vinculado al sistema sanitario español e indicadores de gasto del sistema sanitario público. En el BS, a partir de un diseño muestral, se selecciona un conjunto de individuos en cada una de las Comunidades Autónomas españolas (CCAA) y se les realiza un cuestionario para testar el sistema sanitario.

Con objeto de disponer de unos datos reales, tomamos como inputs los indicadores de personal vinculado al sistema sanitario público (Médicos por cada 1,000 habitantes, Enfermeras por cada 1,000 habitantes, Auxiliares de enfermería por cada 1,000 habitantes) y los indicadores de gasto del sistema sanitario público (Gasto por cama, Gasto por alta, Gasto por estancia) recogidos en la EESRI del año 2009. Para sintetizar la información y reducir el número de inputs, ejecutamos un análisis de componentes principales en cada uno de los grupos de variables y se obtienen las componentes ACP1 y ACP2 de la Tabla 4.16, interpretadas como tamaño de personal y de gasto, respectivamente. Como output se utiliza la respuesta media de una muestra de individuos, en cada CCAA, a la cuestión 4 del BS:

“ Y, ¿está Ud. Satisfecho/a o insatisfecho/a con el modo en que el sistema sanitario público funciona en su Comunidad Autónoma?. Para contestarme valore de 1 a 10 donde 1 significa que está Ud. muy insatisfecho y el 10 que está Ud. muy satisfecho/a”.

En la Tabla 4.16 se recoge los tamaños de muestra en cada una de las CCAA, (n_i) , y la respuesta media de los individuos de la muestra en cada CCAA, \hat{Y}_i . Los tamaños de muestra n_i se suponen determinados para cometer un cierto error de estimación con muestreo aleatorio simple sin reemplazamiento. El intervalo de confianza para la satisfacción media de la población de cada CCAA se estima utilizando (4.9) con $\alpha = 0.1$.

Tabla 4.16: Tamaño de Personal y de Gasto del Sistema Sanitario Español por CCAA en 2009 (inputs), Tamaño muestral y estimación de la satisfacción media con el Sistema Sanitario Español en 2009 (output). Índices de eficiencia estimados CCR-O y BCC-O.

| CCAA | Inputs | | Output | | Índice de eficiencia estimado CCR -O | Índice de eficiencia estimado BCC-O |
|-----------|---------------|------------|---------------------------|---|--------------------------------------|-------------------------------------|
| | Personal ACP1 | Gasto ACP2 | Tamaño muestral (n_i) | Estimación de la satisfacción media de los usuarios \hat{Y}_i | | |
| 1 | 8.58 | 6.36 | 767 | 6.29 | 0.97 | 0.97 |
| 2 | 7.26 | 10 | 336 | 6.80 | 0.87 | 0.94 |
| 3 | 7.07 | 8.75 | 320 | 7.21 | 0.96 | 1 |
| 4 | 8.60 | 7.79 | 312 | 6.60 | 0.85 | 0.93 |
| 5 | 8.66 | 6.99 | 382 | 5.39 | 0.75 | 0.78 |
| 6 | 8.56 | 8.42 | 284 | 6.27 | 0.78 | 0.87 |
| 7 | 6.44 | 7.77 | 378 | 6.79 | 1 | 1 |
| 8 | 7.80 | 7.68 | 418 | 6.81 | 0.93 | 0.96 |
| 9 | 7.80 | 6.93 | 735 | 6.02 | 0.87 | 0.88 |
| 10 | 7.96 | 6.34 | 575 | 6.44 | 0.99 | 1 |
| 11 | 5.84 | 7.54 | 315 | 6.29 | 1 | 1 |
| 12 | 6.48 | 7.96 | 440 | 5.90 | 0.86 | 0.86 |
| 13 | 9.18 | 7.85 | 642 | 6.41 | 0.80 | 0.90 |
| 14 | 10 | 7.33 | 341 | 6.21 | 0.83 | 0.89 |
| 15 | 8.05 | 8.14 | 281 | 7.18 | 0.93 | 1 |
| 16 | 7.76 | 6.62 | 399 | 6.74 | 1 | 1 |
| 17 | 7.63 | 7.75 | 254 | 6.97 | 0.95 | 0.99 |
| 18 | 8.93 | 7.33 | 503 | 6.09 | 0.81 | 0.87 |

Fuente: Estadística de Establecimientos Sanitarios con Régimen de Internado (2009) y Barómetro Sanitario (2009)

La Tabla 4.17 muestra los resultados de la estimación puntual y por ICE de la eficiencia en cada CCAA de España, considerando rendimientos variables de escala.

Tabla 4.17: Eficiencia sanitaria estimada en CCAAs de manera puntual y por intervalo utilizando datos de la Estadística de Establecimientos Sanitarios con Régimen de Internado y Barómetro Sanitario del 2009. Modelo BCC-O.

| CCAA | $\hat{\varphi}_i$ | ICE Tipo1 | | ICE Tipo2 | | ICE Tipo3 | | ICE Bootstrap 98% | |
|------|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|-------------------|-----------|
| | | Ext. Inf. | Ext. Sup. | Ext. Inf. | Ext. Sup. | Ext. Inf. | Ext. Sup. | Ext. Inf. | Ext. Sup. |
| 1 | 0.973 | 0.936 | 1 | 0.936 | 0.973 | 0.936 | 1 | 0.936 | 1 |
| 2 | 0.943 | 0.898 | 0.990 | 0.898 | 0.946 | 0.898 | 0.990 | 0.893 | 0.982 |
| 3 | 1 | 0.989 | 1 | 0.955 | 1 | 0.955 | 1 | 0.994 | 1 |
| 4 | 0.932 | 0.884 | 0.982 | 0.884 | 0.938 | 0.884 | 0.982 | 0.885 | 0.974 |
| 5 | 0.788 | 0.744 | 0.833 | 0.744 | 0.798 | 0.744 | 0.833 | 0.745 | 0.828 |
| 6 | 0.872 | 0.826 | 0.920 | 0.826 | 0.878 | 0.826 | 0.920 | 0.826 | 0.912 |
| 7 | 1 | 0.973 | 1 | 0.947 | 1 | 0.947 | 1 | 0.981 | 1 |
| 8 | 0.970 | 0.928 | 1 | 0.928 | 0.968 | 0.928 | 1 | 0.930 | 1 |
| 9 | 0.882 | 0.846 | 0.920 | 0.846 | 0.881 | 0.846 | 0.920 | 0.851 | 0.913 |
| 10 | 1 | 1 | 1 | 0.961 | 1 | 0.961 | 1 | 1 | 1 |
| 11 | 1 | 1 | 1 | 0.943 | 1 | 0.943 | 1 | 1 | 1 |
| 12 | 0.867 | 0.823 | 0.914 | 0.823 | 0.866 | 0.823 | 0.914 | 0.822 | 0.905 |
| 13 | 0.903 | 0.866 | 0.943 | 0.866 | 0.900 | 0.866 | 0.943 | 0.867 | 0.932 |
| 14 | 0.894 | 0.853 | 0.937 | 0.853 | 0.896 | 0.853 | 0.937 | 0.858 | 0.929 |
| 15 | 1 | 0.967 | 1 | 0.953 | 1 | 0.953 | 1 | 0.970 | 1 |
| 16 | 1 | 0.990 | 1 | 0.960 | 1 | 0.960 | 1 | 0.995 | 1 |
| 17 | 0.992 | 0.942 | 1 | 0.942 | 1 | 0.942 | 1 | 0.946 | 1 |
| 18 | 0.877 | 0.837 | 0.919 | 0.837 | 0.879 | 0.837 | 0.919 | 0.838 | 0.913 |

La Tabla 4.18 recoge la clasificación de las CCAAs en los conjuntos definidos por Despotis and Smirlis (4.36) en E^{++} de unidades eficientes en cualquier caso, E^+ de unidades eficientes en sentido maximal y E^- de las unidades ineficientes:

Tabla 4.18: Clasificación de las CCAAs en los conjuntos E^{++} de unidades eficientes en cualquier caso, E^+ de unidades eficientes en sentido maximal y E^- de las unidades ineficientes.

| | ICE Tipo1 | ICE Tipo2 | ICE Tipo3 | ICE Bootstrap 98% |
|----------|-------------------------------|-------------------------------------|--------------------------------|-------------------------------|
| E^{++} | 10, 11 | | | 10, 11 |
| E^+ | 1, 3, 7, 8, 15, 16, 17 | 3, 7, 10, 11, 15, 16, 17 | 1, 3, 7, 8, 10, 11, 15, 16, 17 | 1, 3, 7, 8, 15, 16, 17 |
| E^- | 2, 4, 5, 6, 9, 12, 13, 14, 18 | 1, 2, 4, 5, 6, 8, 9, 12, 13, 14, 18 | 2, 4, 5, 6, 9, 12, 13, 14, 18 | 2, 4, 5, 6, 9, 12, 13, 14, 18 |

Las CCAAs en las que se rechaza la hipótesis de ser eficientes utilizando cualquiera de las metodologías de ICE son $\{2, 4, 5, 6, 9, 12, 13, 14, 18\}$. Las CCAAs $\{10, 11\}$ se consideran eficientes en cualquier caso con los métodos *Tipo1* y bootstrap siendo eficientes en sentido maximal en los otros dos métodos.

4.4. Alternativa de Intervalo de eficiencia DEA

Los modelos para obtener los extremos del ICE *Tipo3* se han construido considerando m inputs exactos y s outputs estocásticos para los que disponemos de intervalos de confianza. Si nos fijamos como hemos trasladado los modelos que determinan los extremos del intervalo de eficiencia de Despotis and Smirlis a los modelos para obtener los extremos del ICE *Tipo1*, de igual manera podemos actuar, pero a la inversa, con los modelos que determinan los extremos del ICE *Tipo3*, lo que da lugar a la siguiente definición de intervalo de eficiencia:

Definición 1 Si consideramos los datos inputs y outputs intervalos

$$x_{ji} \in [x_{ji}^L, x_{ji}^U], y_{rj} \in [y_{ri}^L, y_{ri}^U]; i = 1, \dots, n; j = 1, \dots, m; r = 1, \dots, s$$

un intervalo de eficiencia DEA, para la DMU_o, se define por las cotas superior e inferior obtenidas mediante los modelos (4.37) y (4.38)

$$\begin{aligned} \max T_o^L &= \sum_{r=1}^s u_r y_{ro}^L \\ \text{s.t.} \\ \sum_{j=1}^m v_j x_{jo}^U &= 1 \\ \sum_{r=1}^s u_r y_{ri}^U - \sum_{j=1}^m v_j x_{ji}^L &\leq 0; i = 1, \dots, n \\ u_r, v_j &\geq \varepsilon \forall r, j \end{aligned} \quad (4.37)$$

$$\begin{aligned} \max T_o^U &= \sum_{r=1}^s u_r y_{ro}^U \\ \text{s.t.} \\ \sum_{j=1}^m v_j x_{jo}^L &= 1 \\ \sum_{r=1}^s u_r y_{ri}^L - \sum_{j=1}^m v_j x_{ji}^U &\leq 0; i = 1, \dots, n \\ u_r, v_j &\geq \varepsilon \forall r, j \end{aligned} \quad (4.38)$$

Los modelos de los extremos del intervalo de eficiencia de la Definición 1 mantienen, en todas las DMUs, la misma frontera de producción para calcular el extremo inferior del intervalo y la misma frontera para obtener el extremo superior, siendo estas dos fronteras distintas. Los modelos de los extremos del intervalo de eficiencia de Wang y otros (2005) [146] también mantienen la misma frontera, siendo la misma frontera para ambos extremos, mientras que las fronteras de producción de los modelos de los extremos del intervalo de eficiencia de Despotis and Smirlis (2002) [46] cambian para el cálculo de los extremos e incluso para el cálculo entre las distintas DMUs.

4.5. Región de confianza simultánea para el vector de índices de eficiencia poblacionales

Ante el problema de calcular el tamaño de muestra necesario en cada DMU para estimar los outputs y con éstos estimar la eficiencia DEA, en la sección 4.2 hemos propuesto una solución de este cálculo fijado un error de estimación del índice de eficiencia en cada DMU. En esta sección vamos a dar solución al problema de determinar el tamaño muestral necesario en cada DMU para estimar los outputs y lograr una región simultánea de confianza, confianza decidida por el experimentador, para el vector de eficiencias $\varphi = (\varphi_1, \dots, \varphi_n)$.

Realizaremos la construcción de la región de confianza para el vector de eficiencias en dos casos, el primero, en el que es posible demostrar analíticamente todos los resultados, corresponde al modelo CCR-O con un input conocido y un output estocástico. El segundo caso corresponde a la generalización del primero considerando más de un input conocido, más de un output estimado y tanto el modelo BCC-O como CCR-O; en este caso, las simulaciones serán la herramienta que permitan comprobar los resultados propuestos en el primer caso.

4.5.1. Modelo CCR-O con un input conocido y un output estocástico

Consideramos n DMUs con un único input con valores conocidos $(X_i)_{i=1,\dots,n}$ y un único output con valores $(Y_i)_{i=1,\dots,n}$ desconocidos. Sean $(\varphi_i)_{i=1,\dots,n}$ los correspondientes índices de eficiencia poblacionales CCR-O (desconocidos). Sean $(\hat{Y}_i)_{i=1,\dots,n}$ los estimadores de los outputs y $(\hat{\Phi}_i)_{i=1,\dots,n}$ los correspondientes estimadores de los índices de eficiencia.

Nota 4 Fijados los inputs, la eficiencia DEA de una DMU dada aumenta al aumentar el output y al disminuir los outputs del resto de DMUs. Asimismo, la eficiencia de una DMU disminuye al disminuir el output de dicha unidad y aumentar los outputs del resto de DMUs.

Lema 5 Para cada $i = 1, \dots, n$, fijado $p \in (0, 1)$, consideremos los sucesos

$$A_i = [-pY_i \leq \hat{Y}_i - Y_i \leq pY_i], \forall i = 1, \dots, n$$

y

$$B_i = \left[-\frac{2p}{1+p} \leq \hat{\Phi}_i - \varphi_i \leq \frac{2p}{1+p} \right], \forall i = 1, \dots, n$$

entonces

$$\bigcap_{i=1}^n A_i \subset \bigcap_{i=1}^n B_i$$

DEMOSTRACIÓN

Consideremos

$$\begin{aligned} \Omega_i &= \{\text{muestras } s_i \text{ de la } i\text{th DMU}\}; i = 1, \dots, n \\ \Omega &= \prod_{i=1}^n \Omega_i \\ A_i &= \{s = (s_1 \times \dots \times s_n) \in \Omega / |\hat{Y}_i(s_i) - Y_i| \leq pY_i\}; i = 1, \dots, n \\ B_i &= \{s \in \Omega / |\hat{\Phi}_i(s) - \varphi_i| \leq \frac{2p}{1+p}\}; i = 1, \dots, n \end{aligned}$$

Sea $s = (s_1 \times \dots \times s_n) \in \bigcap_{i=1}^n A_i$, por lo tanto se verifica

$$|\hat{Y}_i(s_i) - Y_i| \leq pY_i; i = 1, \dots, n$$

Consideremos la r th DMU. $\varphi_r = 1$ o $\varphi_r < 1$

- Si $\varphi_r = 1$ es porque $\frac{Y_r}{X_r} = \max_j \frac{Y_j}{X_j}$

La situación más desfavorable, para que se verifique B_r , es que el output de la r th DMU sea lo menor posible y, para el resto de DMUs, lo mayor posible, es decir:

$$Y_j^* = (1-p)Y_j I_{(j=r)} + (1+p)Y_j I_{(j \neq r)}$$

Sean φ_j^* , $j = 1, \dots, n$, las eficiencias de las DMUs obtenidas con los datos $\{(X_j, Y_j^*)\}_{j=1,\dots,n}$, por lo tanto

$$\varphi_r^* = \begin{cases} 1 & \text{si } \frac{(1-p)Y_r}{X_r} \geq \max_{j \neq r} \frac{(1+p)Y_j}{X_j} & (a) \\ \frac{(1-p)Y_r}{(1+p)\frac{Y_k}{X_k} X_r} & \text{si } (1+p)\frac{Y_k}{X_k} = \max_{j \neq r} \frac{(1+p)Y_j}{X_j} > \frac{(1-p)Y_r}{X_r} & (b) \end{cases}$$

y

$$|\varphi_r - \hat{\Phi}_r(s)| = 1 - \hat{\Phi}_r(s) \leq 1 - \varphi_r^*$$

donde la primera igualdad se obtiene porque $\varphi_r = 1$ y la segunda desigualdad se verifica porque, por la Nota 4, $\hat{\Phi}_r(s) \geq \varphi_r^*$,

entonces

$$|\varphi_r - \hat{\Phi}_r(s)| \leq \begin{cases} 0 & \text{si (a)} \\ 1 - \frac{(1-p)}{(1+p)} = \frac{2p}{1+p} & \text{si (b)} \end{cases}$$

y se verifica que $s \in B_r \forall r / \varphi_r = 1$

■ Si $\varphi_r < 1$

Sea $k \neq r / \frac{Y_k}{X_k} = \max_j \frac{Y_j}{X_j}$ y $\varphi_r = \frac{Y_r}{X_r} < 1$

Hay dos situaciones más desfavorables para que se verifique B_r

$$(I) Y_j^* = (1-p)Y_j I_{(j=r)} + (1+p)Y_j I_{(j \neq r)}$$

$$(II) Y_j^{**} = (1+p)Y_j I_{(j=r)} + (1-p)Y_j I_{(j \neq r)}$$

Caso (I) : Sean φ_j^* , $j = 1, \dots, n$, las eficiencias de las DMUs obtenidas con los datos $\left\{ (X_j, Y_j^*) \right\}_{j=1, \dots, n}$, entonces

$$\varphi_r^* = \frac{(1-p) \frac{Y_r}{X_r}}{(1+p) \frac{Y_k}{X_k}} = \frac{(1-p)}{(1+p)} \varphi_r < \varphi_r$$

Por la Nota 4

$$|\varphi_r - \hat{\Phi}_r(s)| \leq |\varphi_r - \varphi_r^*|$$

y por lo tanto

$$|\varphi_r - \hat{\Phi}_r(s)| \leq \varphi_r - \varphi_r^* = \varphi_r \left(\frac{2p}{1+p} \right) < \frac{2p}{1+p}$$

Caso (II) : Sean φ_j^{**} las eficiencias de las DMUs obtenidas con los datos $\left\{ (X_j, Y_j^{**}) \right\}_{j=1, \dots, n}$

- Si $\max \left(\frac{(1+p)Y_r}{X_r}, \max_{j \neq r} \frac{(1-p)Y_j}{X_j} \right) = \frac{(1+p)Y_r}{X_r}$
entonces $\varphi_r^{**} = 1$ y $\varphi_r \geq \frac{(1-p)}{(1+p)}$
y por lo tanto

$$|\varphi_r - \hat{\Phi}_r(s)| \leq |\varphi_r - \varphi_r^{**}| = 1 - \varphi_r \leq \frac{2p}{1+p}$$

- Si $\max \left(\frac{(1+p)Y_r}{X_r}, \max_{j \neq r} \frac{(1-p)Y_j}{X_j} \right) = \frac{(1-p)Y_k}{X_k}$ para $k \neq r$
entonces $\varphi_r < \frac{(1-p)}{(1+p)}$ y $\varphi_r^{**} = \frac{(1+p) \frac{Y_r}{X_r}}{(1-p) \frac{Y_k}{X_k}} = \frac{(1+p)}{(1-p)} \varphi_r > \varphi_r$
y por lo tanto

$$|\varphi_r - \hat{\Phi}_r(s)| \leq \varphi_r^{**} - \varphi_r \leq \frac{2p}{1-p} \varphi_r < \frac{2p}{1+p}$$

en consecuencia $s \in B_r \forall r / \varphi_r < 1$.

Teorema 5 Fijados $p \in (0, 1)$ y $\alpha \in (0, 1)$, para cada $i = 1, \dots, n$, sea n_i el tamaño de muestra en la i th DMU tal que

$$p(|\hat{Y}_i - Y_i| \leq pY_i) \geq \sqrt[n]{1 - \alpha} \quad (4.39)$$

con $p = \frac{\delta}{2 - \delta} \in (0, 1)$. Entonces, se verifica:

$$p(|\hat{\Phi}_i - \varphi_i| \leq \delta, \forall i = 1, \dots, n) \geq 1 - \alpha; \quad (4.40)$$

y, por tanto, $\bigcap_{i=1}^n [\hat{\Phi}_i \pm \delta]$ define una región de confianza $(1 - \alpha)$ simultánea para $(\varphi_1, \dots, \varphi_n)$.

DEMOSTRACIÓN:

Consideremos

$$\Omega_i = \{\text{muestras de la } i\text{th DMU}\}; i = 1, \dots, n$$

$$\Omega = \prod_{i=1}^n \Omega_i$$

$$A_i = \left\{ s = (s_1 \times \dots \times s_n) \in \Omega / |\hat{Y}_i(s_i) - Y_i| \leq \left(\frac{\delta}{2 - \delta} \right) Y_i \right\}; i = 1, \dots, n$$

$$B_i = \left\{ s \in \Omega / |\hat{\Phi}_i(s) - \varphi_i| \leq \delta \right\}; i = 1, \dots, n$$

Por el Lema 5 sabemos que

$$\bigcap_{i=1}^n A_i \subset \bigcap_{i=1}^n B_i$$

por tanto

$$p\left(\bigcap_{i=1}^n B_i\right) \geq p\left(\bigcap_{i=1}^n A_i\right)$$

como los sucesos $\{A_i\}_{i=1, \dots, n}$ son independientes

$$p\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n p(A_i) \geq \left(\sqrt[n]{1 - \alpha}\right)^n = 1 - \alpha$$

4.5.2. Modelos CCR-O y BCC-O con más de un input conocido y un output estocástico

Para los modelos CCR y BCC con orientación output considerando más de un input y más de un output estocástico, el Lema 5 y el Teorema 5 se comprueban utilizando simulaciones.

ESTUDIO DE SIMULACIÓN

Consideremos los datos de la Tabla 4.2 correspondientes al número de médicos y enfermeras (inputs) y al número de altas e ingresos (outputs) en 12 centros de salud y el modelo poblacional simulado a partir de estos datos contenido en la Tabla 4.3.

Los pasos del proceso de simulación para comprobar el Lema 5 son:

- I. Fijado p (0.1 o 0.05) se simulan valores \hat{y}_{ki} de una distribución uniforme en el intervalo

$$[(1 - p)Y_{ki}, (1 + p)Y_{ki}]; i = 1, \dots, 12; k = 1, 2$$

- II. Con los datos $(X_{1i}, X_{2i}, \hat{y}_{1i}, \hat{y}_{2i})$, los valores $(\hat{\varphi}_i)$ son obtenidos, para cada DMU, maximizando los modelos (4.13) y (4.14).

- III. Se realizan mil iteraciones de los pasos I-II. y se obtienen los valores $\left\{ \hat{\varphi}_i^{(k)} \right\}_{k=1, \dots, 1000}$ y los errores de estimación, $\left\{ E_i^{(k)} = \left| \hat{\varphi}_i^{(k)} - \varphi_i \right| \right\}_{k=1, \dots, 1000}$

La comprobación por simulación del Teorema 5 sigue los siguientes pasos:

- I. En la i th DMU, $i = 1, \dots, 12$, tomamos una muestra previa de tamaño $n_i^{(0)}$. Con esta muestra, obtenemos $\hat{y}_i^{(0)} = (\hat{y}_{1i}^{(0)}, \hat{y}_{2i}^{(0)})$, como en (4.10), y la estimación de la varianza $\hat{\sigma}_i^{2(0)} = (\hat{\sigma}_{1i}^{2(0)}, \hat{\sigma}_{2i}^{2(0)})$, como en (4.11). El tamaño de muestra inicial $n_i^{(0)}$ debe ser pequeño con respecto al tamaño poblacional, por ejemplo, nosotros tomamos $n_i^{(0)} = 25$.
- II. Fijado δ (0.15 o 0.1 o 0.05) y α (0.05 o 0.1) y con los valores $(\hat{y}_i^{(0)}, \hat{\sigma}_i^{2(0)})$ determinamos el tamaño de muestra n_i , $i = 1, \dots, 12$, como en (4.17) y (4.16) para que se verifique

$$p \left(|\hat{Y}_{ki} - Y_{ki}| \leq p \hat{y}_{ki}^{(0)} \right) \geq \sqrt[12]{1 - \alpha}; \quad i = 1, \dots, 12; \quad k = 1, 2$$

con

$$p = \frac{\delta}{2 - \delta}$$

- III. En la i th DMU, tomamos el tamaño de muestra n_i y se estiman las medias $\hat{y}_i = (\hat{y}_{1i}, \hat{y}_{2i})$ como en (4.10). Con los datos $(X_{1i}, X_{2i}, \hat{y}_{1i}, \hat{y}_{2i})$, obtenemos los valores $(\hat{\phi}_i)$, para cada DMU, maximizando los modelos (4.13) y (4.14).

- IV. Se realizan mil iteraciones del paso III., obteniéndose $\left\{ \hat{\phi}_i^{(k)} \right\}_{k=1, \dots, 1000}$

- v. Se obtienen las regiones

$$R^{(k)} = \bigcap_{i=1}^{12} \left[\hat{\Phi}_i^{(k)} \pm \delta \right]; \quad k = 1, \dots, 1000$$

- VI. Se aproxima la confianza de la región simultánea con

$$C = \left(\frac{1}{1000} \sum_{k=1}^{1000} I_{(\varphi_1, \dots, \varphi_{12}) \in R^{(k)}} \right)$$

Resultados de la Simulación para comprobar el Lema 5:

Las Figuras 4.7, 4.8, 4.9 y 4.10 muestran los boxplots de los mil errores de estimación obtenidos para cada DMU en los modelos CCR-O y BCC-O y para dos valores de p (0.1 y 0.05). También se representa la cota $\delta = 2p/(1 + p)$ con el fin de observar el comportamiento de los errores de estimación frente a esta cota.

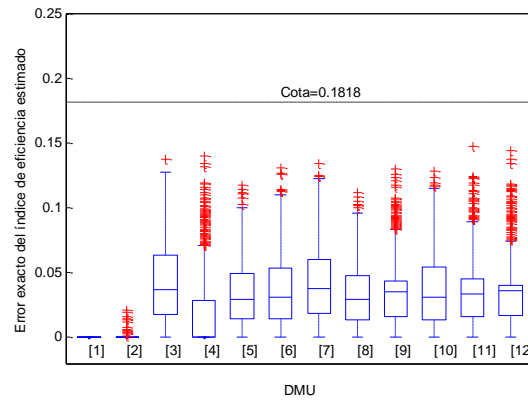


Figura 4.7: Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.1$ y $\frac{2p}{1+p} = 0.1818$. Modelo CCR-O.

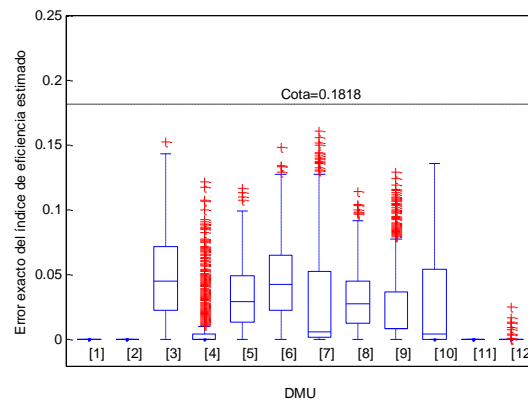


Figura 4.8: Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.1$ y $\frac{2p}{1+p} = 0.1818$. Modelo BCC-O.

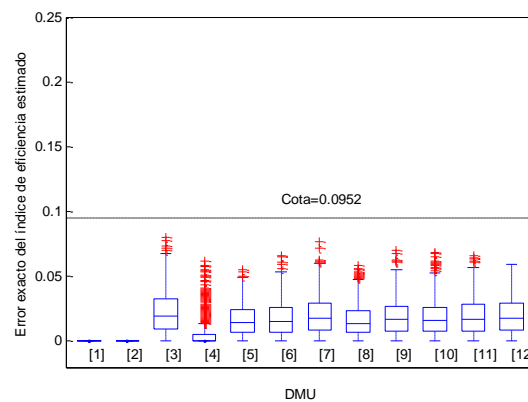


Figura 4.9: Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.05$ y $\frac{2p}{1+p} = 0.0952$. Modelo CCR-O.

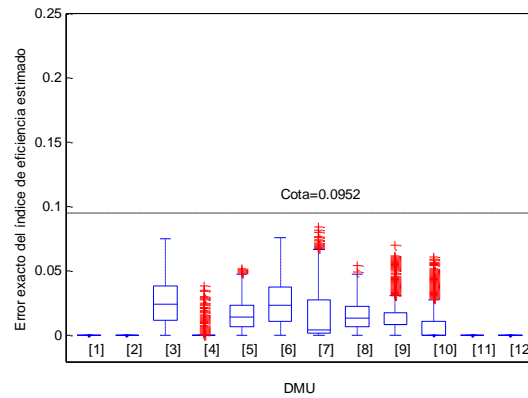


Figura 4.10: Lema 5. Boxplots de mil errores de estimación para cada DMU, $p = 0.05$ y $\frac{2p}{1+p} = 0.0952$. Modelo BCC-O.

Resultados de la Simulación para comprobar Teorema 5:

La Tabla 4.19 muestra los tamaños de muestra obtenidos en el paso II. La Tabla 4.20 contiene la confianza de la región simultánea para el vector de índices de eficiencia poblacional, para los modelos CCR-O y BCC-O y distintos valores de δ y α .

Tabla 4.19: Simulación Teorema 5. Tamaño de muestra

| DMU | n_i | | | |
|-----------|------------------------------------|-----------------------------------|------------------------------------|----------------------------------|
| | $\delta = 0.15$ $\alpha = 0.05$ | $\delta = 0.1$ $\alpha = 0.05$ | $\delta = 0.05$ $\alpha = 0.05$ | $\delta = 0.1$ $\alpha = 0.1$ |
| 1 | 555 | 985 | 3862 | 1390 |
| 2 | 449 | 674 | 2840 | 891 |
| 3 | 258 | 779 | 2514 | 1094 |
| 4 | 202 | 1078 | 2647 | 898 |
| 5 | 477 | 900 | 4554 | 765 |
| 6 | 359 | 1415 | 2871 | 988 |
| 7 | 358 | 795 | 2930 | 657 |
| 8 | 460 | 814 | 3493 | 724 |
| 9 | 385 | 765 | 3510 | 911 |
| 10 | 475 | 617 | 4184 | 845 |
| 11 | 291 | 1194 | 2449 | 881 |
| 12 | 309 | 682 | 2620 | 929 |

Tabla 4.20: Teorema 5. Confianza aproximada de la región simultánea del vector de eficiencias

| DMU | C% | | | |
|--------------|------------------------------------|-----------------------------------|------------------------------------|----------------------------------|
| | $\delta = 0.15$ $\alpha = 0.05$ | $\delta = 0.1$ $\alpha = 0.05$ | $\delta = 0.05$ $\alpha = 0.05$ | $\delta = 0.1$ $\alpha = 0.1$ |
| CCR-O | 100 | 100 | 100 | 100 |
| BCC-O | 100 | 100 | 100 | 99.9 |

Las Figuras 4.11, 4.12, 4.13, 4.14, 4.15 y 4.16 muestran los mil índices de eficiencia estimados (líneas rojas) en cada DMU, la región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) y los índices de eficiencia poblacionales (línea con símbolo \circ) para los modelos CCR-O y BCC-O y distintos valores de δ y α .

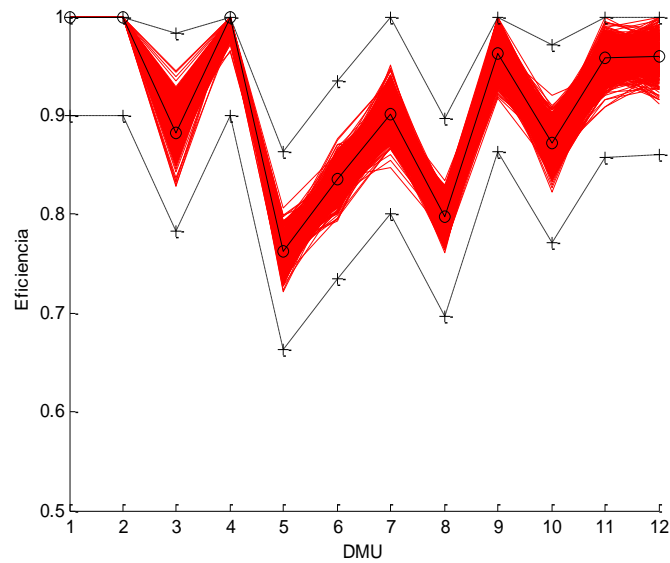


Figura 4.11: Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo CCR-O, $\delta = 0.1$ y $\alpha = 0.05$

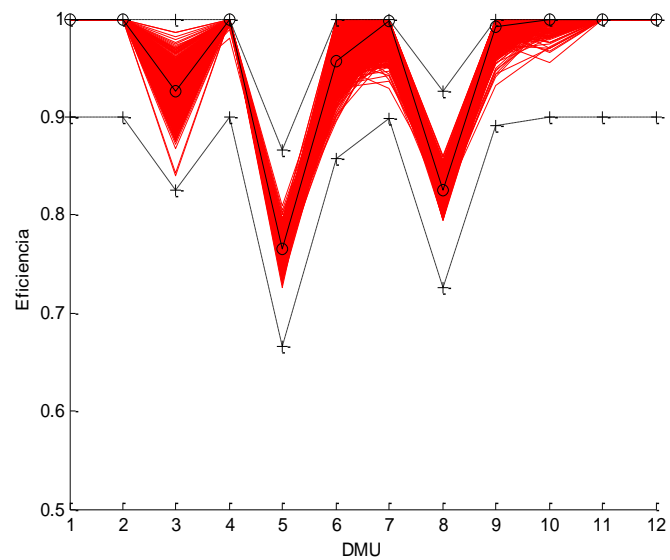


Figura 4.12: Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo \circ). Modelo BCC-O, $\delta = 0.1$ y $\alpha = 0.05$

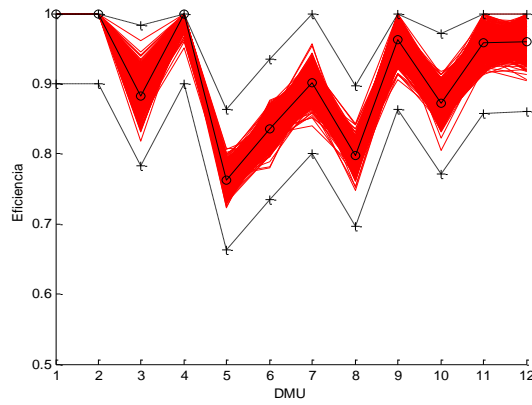


Figura 4.13: Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo o). Modelo CCR-O, $\delta = 0.1$ y $\alpha = 0.1$

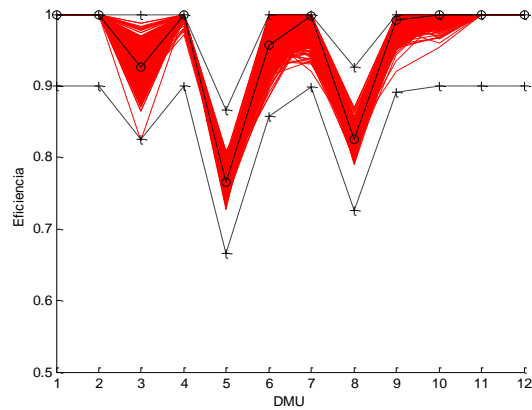


Figura 4.14: Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo o). Modelo BCC-O, $\delta = 0.1$ y $\alpha = 0.1$

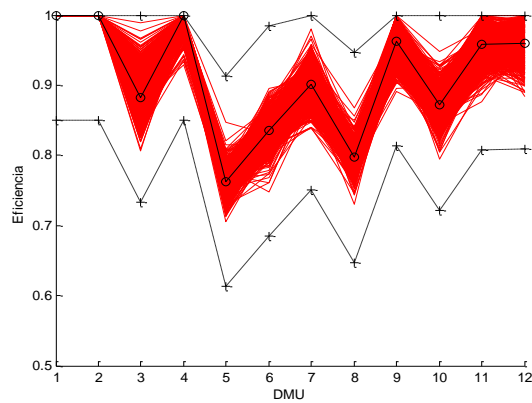


Figura 4.15: Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo o). Modelo CCR-O, $\delta = 0.15$ y $\alpha = 0.05$

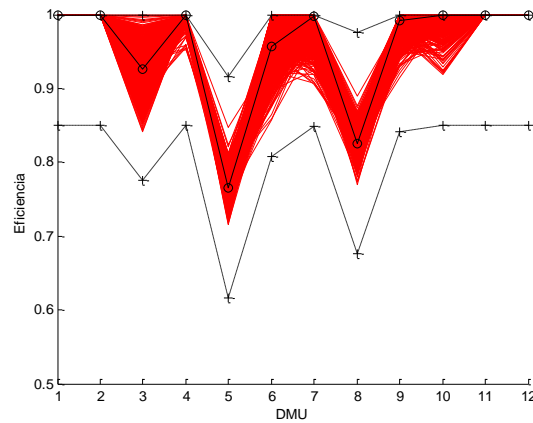


Figura 4.16: Teorema 5. Índices de eficiencia estimados (líneas rojas), región $\bigcap_{i=1}^{12} [\varphi_i \pm \delta]$ (líneas con símbolo +) e índices de eficiencia poblacionales (línea con símbolo o). Modelo BCC-O, $\delta = 0.15$ y $\alpha = 0.05$

En conclusión, el tamaño de muestra determinado por el Teorema 5 conduce a una confianza de la región simultánea superior al 99 %.

4.5.3. Un estudio de aplicación

La Tabla 4.8 contiene la información de una base de datos compuesta por 15 DMUs, dos inputs y un output. Los datos inputs corresponden al número de préstamos de libros (X_1) y al número de puestos de usuarios (número de asientos y número de ordenadores ofrecidos a los usuarios) (X_2) en 15 bibliotecas, datos escalados a 0-10 en busca de lograr simplicidad. La población son los 637,100 usuarios de estas bibliotecas. En cada DMU, sólo usamos el output dado por la satisfacción media del usuario, satisfacción con su biblioteca medida en cada usuario con un valor entero de 0 (nada satisfecho) a 10 (muy satisfecho). Los pasos para obtener la región simultánea de confianza propuesta en el Teorema 5 para el vector de eficiencias son:

- I. Tomamos, en cada DMU, una muestra piloto aleatoria simple sin reemplazamiento, $\{n_i^{(0)}\}_{i=1, \dots, n}$, del 0.1 % del tamaño de la población, y estimamos el output y la varianza poblacional (Tabla 4.21).
- II. Fijados $\delta=0.1$ y $\alpha=0.1$, determinamos el tamaño de muestra n_i para que se verifique (4.39) (columna 2 de la Tabla 4.22) y obtenemos los valores estimados de la satisfacción media $(\hat{y}_i)_{i=1, \dots, 15}$.
- III. Con los datos $(X_{1i}, X_{2i}, \hat{y}_i)_{i=1, \dots, 15}$, obtenemos los valores estimados de los índices de eficiencia $(\hat{\varphi}_i)_{i=1, \dots, 15}$ (columnas 3 y 5 de la Tabla 4.22) y la región simultánea de confianza $\bigcap_{i=1}^{12} [\hat{\varphi}_i - \delta, \hat{\varphi}_i + \delta]$ para los modelos CCR-O y BCC-O (columnas 4 y 6 de la Tabla 4.22).

Tabla 4.21: Estudio de aplicación. Resultados de la muestra aleatoria simple previa: tamaño de muestra, estimación de la satisfacción media de los usuarios y de la varianza poblacional.

| DMU | Tamaño muestral $n_i^{(0)}$ | Estimación de la satisfacción media de los usuarios | Estimación de la varianza poblacional |
|-----|-----------------------------|---|---------------------------------------|
| 1 | 89 | 6.64 | 13.09 |
| 2 | 79 | 4.98 | 16.93 |
| 3 | 64 | 3.49 | 9.18 |
| 4 | 59 | 5.71 | 14.95 |
| 5 | 57 | 6.03 | 16.41 |
| 6 | 51 | 2.67 | 7.51 |
| 7 | 42 | 6.21 | 16.28 |
| 8 | 37 | 3.38 | 10.76 |
| 9 | 35 | 4.87 | 15.71 |
| 10 | 32 | 3.38 | 8.50 |
| 11 | 24 | 5.84 | 14.69 |
| 12 | 21 | 5.31 | 18.02 |
| 13 | 19 | 4.80 | 16.80 |
| 14 | 17 | 2.50 | 10.05 |
| 15 | 13 | 7.03 | 15.41 |

Tabla 4.22: Estudio de aplicación. Tamaño de muestra, estimación del índice de eficiencia y región de confianza para la eficiencia de 15 bibliotecas, $\delta = 0.1$, $\alpha = 0.1$. Modelos BCC-O y CCR-O.

| DMU | n_i | BCC-O | | CCR-O | |
|-----|-------|----------------|---------------------|----------------|---------------------|
| | | $\hat{\phi}_i$ | Región de Confianza | $\hat{\phi}_i$ | Región de Confianza |
| 1 | 774 | 0.964 | [0.864, 1] | 0.299 | [0.199, 0.399] |
| 2 | 1750 | 0.864 | [0.764, 0.964] | 0.241 | [0.141, 0.341] |
| 3 | 1921 | 0.582 | [0.482, 0.682] | 0.234 | [0.134, 0.334] |
| 4 | 1180 | 0.879 | [0.779, 0.979] | 0.716 | [0.616, 0.816] |
| 5 | 1161 | 0.968 | [0.868, 1] | 0.533 | [0.433, 0.633] |
| 6 | 2623 | 0.562 | [0.462, 0.662] | 0.373 | [0.273, 0.473] |
| 7 | 1079 | 0.948 | [0.848, 1] | 0.474 | [0.374, 0.574] |
| 8 | 2312 | 0.552 | [0.452, 0.652] | 0.194 | [0.094, 0.294] |
| 9 | 1654 | 0.864 | [0.764, 0.964] | 0.325 | [0.225, 0.425] |
| 10 | 1841 | 0.564 | [0.464, 0.664] | 0.230 | [0.130, 0.330] |
| 11 | 1078 | 0.875 | [0.775, 0.975] | 0.726 | [0.626, 0.826] |
| 12 | 1556 | 1 | [0.900, 1] | 0.613 | [0.513, 0.713] |
| 13 | 1735 | 0.887 | [0.787, 0.987] | [0.612, 0.512] | 0.712 |
| 14 | 3386 | 1 | [0.900, 1] | 0.575 | [0.475, 0.675] |
| 15 | 770 | 1 | [0.900, 1] | 1 | [0.900, 1] |

Atendiendo a la región simultánea del modelo BCC, las bibliotecas $\{1, 5, 7, 12, 14, 15\}$ pueden ser consideradas eficientes ya que el intervalo, que define la región, en estas DMUs contiene el valor 1.

Conclusiones y Líneas futuras de investigación

En el análisis de la eficiencia DEA, la utilización de outputs estimados con una muestra probabilística, da lugar a índices de eficiencia estimados. En este bloque, los Teoremas 1 y 2 nos permiten, para el modelo CCR con sólo un input conocido y sólo un output estocástico, encontrar el tamaño de muestra necesario, en cada DMU, para controlar el error de estimación del índice de eficiencia. Las simulaciones realizadas, muestran la validez de estos resultados para más de un input conocido, más de un output estocástico y el modelo BCC.

Por otro lado, se proponen cuatro intervalos de confianza para la eficiencia DEA en presencia de outputs estimados con una muestra. La comparación, por simulación de los cuatro métodos muestra que el *Tipo2* tiene una confianza muy baja y se desaconseja su uso; los intervalos (ICEs) *Tipo1* y *Tipo3* son muy similares, considerando que es mejor el *Tipo3* porque sólo utiliza dos fronteras de producción, una para calcular todos los extremos inferiores de las ICEs y otra para calcular todos los extremos superiores de las ICEs; el intervalo de confianza DEA bootstrap es la mejor opción por dos razones: primera, porque permite que la confianza de los intervalos de eficiencia sea controlada, manteniendo una amplitud muy similar a los métodos basados en el punto de vista optimista /pesimista *Tipo1* y *Tipo3*, y, segunda razón, porque detecta las DMUs que son eficientes en cualquier caso según la clasificación (4.36).

En el Teorema 5 se determina el tamaño de muestra que nos permite obtener una región de confianza simultánea para el vector de índices de eficiencia poblacional, para el modelo CCR con sólo un input conocido y sólo un output estocástico. La validación de este resultado para más inputs conocidos, más outputs estocásticos y el modelo BCC, se realiza vía simulaciones.

Desde el punto de vista aplicado, ante la frecuente necesidad de usar como output, información obtenida de encuestas a usuarios, y evaluar la eficiencia de un conjunto homogéneo de unidades que prestan un servicio(s) a un conjunto de clientes, este bloque ofrece un método para calcular el tamaño de muestra, que se debe tomar en cada DMU. Este tamaño garantiza que la estimación del índice de eficiencia esté tan próximo a la verdadera eficiencia como deseemos (control del error de la estimación de la eficiencia) y permite estimar la eficiencia con intervalos de confianza, para cada DMU, o con una región simultánea de confianza, para todas las DMUs.

Si nos planteamos la ejecución práctica de los resultados expuestos en este bloque en una situación real de unidades que prestan un servicio a clientes a los que se les va a realizar un cuestionario para disponer de outputs basados en su opinión, la propuesta que hacemos es la siguiente: Determinaremos el tamaño de muestra en cada DMU utilizando el Teorema 2. Realizaremos la encuesta tomando la muestra probabilística de este tamaño y sintetizaremos la opinión de los individuos en uno o más outputs. Seleccionaremos un conjunto de inputs conocidos que consideremos tengan alguna influencia en la opinión recogida. El tamaño de muestra decidido nos garantiza la confianza del intervalo (4.21) que se ha definido como intervalo *Tipo4* en la sección 4.3. Los resultados de comparación de intervalos de confianza de eficiencia DEA, demostrados teóricamente o validados mediante simulaciones, permiten afirmar que, con este tamaño de muestra, también podemos obtener un ICE bootstrap-percentil de la confianza que deseemos, para el índice de eficiencia poblacional, en cada DMU. En conclusión, fijado un error de estimación para el índice de eficiencia poblacional,

δ , y una confianza $(1 - \alpha)$ obtenemos el tamaño de muestra utilizando el Teorema 2; fijamos una confianza del intervalo bootstrap-percentil, $(1 - \alpha')$, y obtenemos este ICE bootstrap, en cada DMU, según lo expuesto en la sección 4.3.5.

Respecto a posibles líneas futuras de investigación nos planteamos comprobar si los resultados de las secciones 4.2, 4.3 y 4.5 también son extrapolables a otros modelos DEA o a la situación de tanto inputs como outputs estimados con una muestra probabilística. Otros problemas a abordar son el análisis del efecto del uso de outputs estimados con una muestra probabilística en la evaluación de la súper-eficiencia, en la obtención de los valores de las holguras y en la determinación del grupo de unidades de referencia para las unidades ineficientes.

Apéndice A

Códigos de los programas MATLAB del Bloque I

A.1. Simulación de poblaciones

- I. El siguiente programa de MATLAB simula la población de la sección 2.5.1 donde todas las unidades de tiempo de atención a clientes son saturadas:

```
% Simulacion de Marco poblacional con todas las UDTs Saturadas
% Marco Tiempo: 5 dias de 8 horas cada dias 4 semanas
% UDT: Hora
clear
NI=20*8;

% Para decir el numero de usos por hora

% Si hora es no saturada numero de usos entre [1, 9] y el tiempo total de
%atencion de todos menor extrictamente que 60 minutos; se supone que se
% pueden hacer todas las encuestas 6*9<60
% Max es el numero maximo de atenciones que supuestamente se pueden atender
%en una hora
%por ejemplo cada 3 minutos una atencion
MAX=20;
%Tiempo maximo de atencion a un cliente 20 minuotos
tmax=20;

%si la hora es saturada numero de usos entre [10 y MAx] pero tiempo total de
%atencion entre [60 y 79) (suponemos 20 minutos el maximo de atencion a un
%cliente
%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
%Generamos el numero de usos en las horas saturadas, fijamos 20 como
%maximo xq se supone un tiempo minimo de atencion de 3 minutos

for i=1:8
for j=1:20
%Numero de usos entre 10 (ya no se pueden hacer las encuestas) y 20
%el maximo de atenciones
Ni(i+8*(j-1),1)=ceil(unifrnd(3,20));
end
end

%Numero de usos
N=sum(Ni);
%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
% Para cada uso hay que determinar si es usuario o no su satisfaccion y el
```

```

% tiempo de atencion
%%
% Decidir si cada individuo es usuario o no usuario (ya ha sido
% entrevistado antes)
% Probabilides de ser usuario segun pasa el tiempo

% pi: Probabilidad de que en una hora el individuo que es seleccionado no sea
% usurio (que ya haya sido encuestado)
pi=zeros(1,Ni);
for i=1:Ni
pi(1,i)=0.0005*(i-1);
end

% Determinacion de si es uso o usuraio (1=usuario 0=no usuraio)
usua=zeros(max(Ni),Ni);

% Usua valores 0=NO USUARIO 1=USUARIO
for j=1:Ni
for i=1:Ni(j,1)
clear u
u=unifrnd(0,1);
if u<pi(1,j)
usua(i,j)=0;
elseif u>=pi(1,j)
usua(i,j)=1;
end
end
end

% Proporcion de usuarios en cada hora
nusua=zeros(1,Ni);

for j=1:Ni
nusua(1,j)=sum(usua(1:Ni(j,1),j));
end
pusua=(nusua./Ni')*100;
%%

% Tiempo de atencion de cada usuario

% Primero generamos el tiempo total de atencion a todos los usos en cada
% unidad de tiempo atendiendo que la hora sea o no saturada

% Genermos el tiempo de atencion total en las horas
Ti=zeros(Ni,1);

for i=1:8
for j=1:20
% Numero de usos entre 10 (ya no se pueden hacer las encuestas) y 20
% por tanto el total estara entre 61 minutos y 74 que implica que el
% ultimo empieza a ser atendido en el minuto 59 y le atienden el
% tiempo maximo que es de 20 minutos
Ti(i+8*(j-1),1)=ceil(unifrnd(60,79));
end
end
Ti=round(Ti');

% Ahora repartimos estos tiempos entre todos los usos

```

```

% Una vez asignados el minimo de atencion a cada uso de tres minutos el
% resto del Ti hay que repartirlo entre los usos

ti=zeros(max(Ni),NI);
for j=1:NI
ti(1:Ni(j,1),j)=3;
end

% Tiempo minimo de atencion en cada hora seria
minTi=zeros(1,NI);

for j=1:NI
minTi(1,j)=sum(ti(1:Ni(j,1),j));
end

% Lo que queda de repartir de tiempo de atencion en cada hora entre los usos
resti=Ti-minTi;
u=zeros(max(Ni),NI);
for j=1:NI
clear d1
clear d2
clear d3
clear d4
if Ni(j,1)>1
if resti(1,j)>0
d1=resti(1,j)/(Ni(j,1));
u(1:Ni(j,1),j)=round(unifrnd(0, min(d1,17),Ni(j,1),1));
%para que el tiempo no sea mayor que los 20 minutos
%repartimos los minutos que aun quedan
d2=resti(1,j)-sum(u(:,j));

if d2>0 & d2<=7
d3=mas(Ni(j,1),1);
u(d3,j)=u(d3,j)+d2;
end
if d2>7 & d2<=13
d4=mas(Ni(j,1),2);
u(d4(1,1),j)=u(d4(1,1),j)+ceil(d2/2);
u(d4(2,1),j)=u(d4(2,1),j)+ceil(d2/2);
end
if d2>13
d5=mas(Ni(j,1),3);
u(d5(1,1),j)=u(d5(1,1),j)+ceil(d2/3);
u(d5(2,1),j)=u(d5(2,1),j)+ceil(d2/3);
u(d5(3,1),j)=u(d5(3,1),j)+ceil(d2/3);
end
end
else
u(1:Ni(j,1),j)=resti(1,j);
end
end

% Tiempo de atencion asignado a cada USO
ti=ti+u;
%Comprobamos que funciona
pru=sum(ti)-Ti;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% para cada usuario hay que simular respuesta de la variable de interes ,
%satisfaccion de cada usuario 0-10 con un cierto sesgo entre horas, x ejemplo

```



```

% Para decir el numero de usos por hora
% Si hora es no saturada numero de usos entre [1, 9] y el tiempo total de
% atencion de todos menor estrictamente que 60 minutos; se supone que se
% pueden hacer todas las encuestas 6*9<60
% Max es el numero maximo de atenciones que supuestamente se pueden atender
% en una hora
% Por ejemplo cada 3 minutos una atencion
MAX=20;
% Tiempo maximo de atencion a un cliente 20 minutos
tmax=20;

% Si la hora es saturada numero de usos entre [10 y MAX] pero tiempo total de
% atencion entre [60 y 79) (suponemos 20 minutos el maximo de atencion a un
% cliente

% Generamos el numero de usos en las horas no saturadas , al menos 1
Ni=zeros(NI,1);
for i=[1,2,5,6,8]
for j=1:20
% Numero de usos
Ni(i+8*(j-1),1)=ceil(unifrnd(0,19));
end
end
% Genermos el numero de usos en las horas saturadas , fijamos 20 como
% maximo porque se supone un tiempo minimo de atencion de 3 minutos

for i=[3,4,7]
for j=1:20
% Numero de usos entre 10 (ya no se pueden hacer las encuestas) y 20
% (el maximo de atenciones)
Ni(i+8*(j-1),1)=ceil(unifrnd(9,20));
end
end
%%%%

% Para cada uso hay que determinar si es usuario o no, su respuesta a la
% variable de interes , por ejemplo satisfaccion 0-10 y el tiempo de atencion

% Numero de usos
N=sum(Ni);

% Decidir si cada individuo es usuario o no usuario (ya ha contestado antes)
% probabilidades de ser usuario segun pasa el tiempo

% pi: Probabilidad de que en una hora el individuo que es seleccionado no sea
% usuario (que ya haya sido encuestado)
pi=zeros(1,NI);
for i=1:NI
pi(1,i)=0.0005*(i-1);
end

% Determinacion de si es uso o usuario (1=usuario 0=no usuraio)
usua=zeros(max(Ni),NI);

% Usua valores 0=NO USUARIO 1=USUARIO

for j=1:NI
for i=1:Ni(j,1)
clear u
u=unifrnd(0,1);

```

```

if u<pi(1,j)
usua(i,j)=0;
elseif u>=pi(1,j)
usua(i,j)=1;
end
end

end

% Proporción de usuarios en cada hora
nusua=zeros(1,NI);

for j=1:NI
nusua(1,j)=sum(usua(1:Ni(j,1),j));
end
pusua=(nusua./Ni')*100;

%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%

% Tiempo de atención de cada usuario

% Primero generamos el tiempo total de atención a todos los usos en cada
% unidad de tiempo atendiendo que la hora sea o no saturada

% Generamos el tiempo de atención total en las horas
Ti=zeros(NI,1);

for i=[1,2,5,6,8]
for j=1:20
% Número de usos mínimo de 3 minutos de atención en cada uno, lo que
% sobra habrá que repartirlo entre los usos
Ti(i+8*(j-1),1)=unifrnd(3*(Ni(i+8*(j-1),1)),min(20*Ni(i+8*(j-1),1),59));
end
end

for i=[3,4,7]
for j=1:20
% Número de usos entre 10 (ya no se pueden hacer las encuestas) y 20
% por tanto el total estará entre 61 minutos y 79 que implica que el
% último empieza a ser atendido en el minuto 59 y le atienden el
% tiempo máximo que es de 20 minutos
Ti(i+8*(j-1),1)=ceil(unifrnd(60,min(20*Ni(i+8*(j-1),1),79)));
end
end
Ti=round(Ti');
% Ahora repartimos estos tiempos entre todos los usos

% Una vez asignados el mínimo de atención a cada uso de tres minutos el
% resto del Ti hay que repartirlo entre los usos

ti=zeros(max(Ni),NI);
for j=1:NI
ti(1:Ni(j,1),j)=3;
end
% Tiempo mínimo de atención en cada hora sería de 3 minutos por uso
sumti=sum(ti);

% Lo que queda de repartir de tiempo de atención en cada hora entre los usos
resti=Ti-sumti;

```



```

u=zeros(max(Ni),NI);
Ni=Ni';
for i=1:NI
d1=0;
clear d3
clear d5
clear d6
if resti(1,i)>0
d6=floor(resti(1,i)/17);
if d6>=Ni(1,i)
u(1:Ni(1,i),i)=17;
end
if d6<Ni(1,i)
d5=mas(Ni(1,i),Ni(1,i));
for j=1:Ni(1,i)
clear d2
d3=resti(1,i)-d1;
if d3>0
d2=round(unifrnd(min(d3,2),min(d3,17)));
u(d5(j,1),i)=d2;
d1=d1+d2;
else
u(d5(j,1),i)=0;
end
end
end
end
end

% Tiempo de atencion asignado a cada USO
ti=ti+u;

% Comprobemos s las horas saturadas lo siguen siendo y las no saturadas tb)
prue2=zeros(1,NI);
sti=sum(ti)';
for i=[1,2,5,6,8]
for j=1:20
%Horas no saturadas
if sti(i+8*(j-1),1)<60
prue2(i+8*(j-1),1)=1;
end
end
end

for i=[3,4,7]
for j=1:20
% Horas saturadas
if sti(i+8*(j-1),1)>=60;
prue2(i+8*(j-1),1)=1;
end
end
end

% prue 3 coincide con el total de horas con lo que las saturadas los son y
% las otras no
prue3=sum(prue2);
% Comprobamos que funciona
pru=sum(ti)-Ti;

%%%
% Para cada usuario hay que simular Satisfaccion de cada usuario con un

```

```

% cierto sesgo entre horas, mas satisfechos cuanto mas temprano
yi=zeros(max(Ni),NI);

% Generemos satisfaccion 0-10

for j=1:160
for i=1:Ni(1,j)
if ti(i,j)>=3 & ti(i,j)<5
yi(i,j)=normrnd(8,1);
elseif ti(i,j)>=5 & ti(i,j)<10
yi(i,j)=normrnd(7,2);
elseif ti(i,j)>=10 & ti(i,j)<15
yi(i,j)=normrnd(5,2);
elseif ti(i,j)>=15 & ti(i,j)<20
yi(i,j)=normrnd(3,2);
elseif ti(i,j)>=20
yi(i,j)=normrnd(2,1);
end
end
end

% Redondeamos las satisfacciones de 0 a 10 entero
for i=1:max(Ni)
for j=1:160
if yi(i,j)>10
yi(i,j)=10;
elseif yi(i,j)<0
yi(i,j)=0;
end
end
end
yi=round(yi);

% Variable que determina si horas saturadas (1) o no saturadas (0) (se conoce
% a priori)
for i=1:8
for j=1:20
% Horas no saturadas
if sti(i+8*(j-1),1)>=60
sat(i+8*(j-1),1)=1;
else
sat(i+8*(j-1),1)=0;
end
end
end
sat=sat';

save datpobSyNS N Ni yi usua ti sat
% ti Tiempo de atencion asignado a cada USO
% sat Variable que determina si UDT saturada (1) o no saturada (0)
% usua Determinacion de si es uso o usuario (1=usuario 0=no usuario)
% yi Redondeamos las satisfacciones de 0 a 10 entero
% Ni numero de usos en las UDTs
% N total de usos

```

III. En el enlace <http://www.eio.uva.es/~tapia> se encuentran las distintas bases de datos MATLAB simuladas en el primer bloque de la tesis:

- Población simulada con todas las unidades de tiempo de uso saturado, sección 2.5.1, *datpobS.mat*.

- Población simulada con unidades de tiempo de uso saturado y no saturado, sección 2.5.2, *datpobSyNS.mat*.
- Población simulada de usuarios significativamente inferior a la población de usos, sección 2.8:
 - Todas las unidades de tiempo de usos saturado, *datpobSmenusr.mat*.
 - Unidades de tiempo de uso saturado y no saturado, *datpobSyNSmenusr.mat*.

A.2. Diseños muestrales

IV. Programa MATLAB para tomar una muestra sistemática fraccional

```
% Muestreo Sistemático fraccional
% N tamaño población
% n tamaño de muestra
function m=sistfrac(N,n)

k=N/n;
s=unifrnd(0,k);

for j=1:n
m(1,j)=s+(j-1)*k;
m(1,j)=ceil(m(1,j));
end
```

V. Programa MATLAB para tomar una muestra sistemática proporcional al tiempo de atención en cada unidad de tiempo seleccionada en primera etapa

```
% Función que toma muestra sistemática proporcional al tiempo de
% atención en cada unidad de tiempo seleccionada en primera etapa
% Suponemos tiempo en minutos
% Unidad de tiempo 60 minutos
% ti tiene el tiempo de atención de cada individuo de cada unidad de tiempo
% nt es el número de encuestas que se van a realizar en cada hora
function m=sistprop(ti,nt)
% minutos de la muestra
% como se supone que el tiempo de atención es permanente se puede tomar la
% muestra del minuto 1 al 60
m1=sistfrac(60,nt)';
[Ni,NNi]=size(ti);
T=zeros(Ni,1);
T(1,1)=ti(1,1);
for j=2:Ni
T(j,1)=T(j-1,1)+ti(j,1);
end
T=[0
T];
m=zeros(nt,1);
for j=1:nt
for i=2:(Ni+1)
if T(i-1,1)<m1(j,1) & T(i,1)>= m1(j,1)
m(j,1)=i-1;
end
end
end
```

A.3. Simulaciones de confianza de intervalos, de errores de estimación y de efecto del diseño

VI. En el enlace <http://www.eio.uva.es/~tapia> se encuentran los distintos programas que realizan las simulaciones de la sección 2.5:

- Simulación de una población con todas las unidades de tiempo de uso saturado, sección 2.5.1
 - Simulaciones de la estimación del Tamaño de la población *sec2y5y1TamPob.m*
 - Simulaciones de la estimación de la media poblacional de usos y usuarios *sec2y5y1MedPob.m*
 - Simulación del efecto del diseño *sec2y5y1effdis.m*
- Simulación de una población con unidades de tiempo de uso saturado y no saturado, sección 2.5.2
 - Simulaciones de la estimación del Tamaño de la población *sec2y5y2Tampob.m*
 - Simulaciones de la estimación de la media poblacional de usos y usuarios *sec2y5y2MedPob.m*
 - Simulación del efecto del diseño *sec2y5y2efdis.m*

Apéndice B

Códigos de los programas MATLAB del Bloque II

B.1. Simulación de poblaciones

- I. En el enlace <http://www.eio.uva.es/~tapia> se encuentran las distintas bases de datos MATLAB simuladas en el segundo bloque de la tesis:
- Modelo poblacional simulado en sección 4.2.2. a partir de datos utilizados en Cooper y otros (2006) (Tabla 4.3) *datcoopmasdat.mat*
 - Datos de población de bibliotecas del estudio de aplicación de la sección 4.2.3. (Tabla 4.8) *pob15bibmaspob.mat*
 - Datos del estudio de aplicación de la sección 4.3.8. (Tabla 4.17) de la "Estadística de Establecimientos Sanitarios con Régimen de Internado del 2009" (*MNSinput2ACP.mat*) y "Barómetro Sanitario del 2009" (*MNSoutputP4BS09.mat*)

B.2. Cálculo de tamaño de muestra

- II. El siguiente programa de MATLAB nos permite calcular el tamaño de muestra, para obtener el ICE del Teorema 2, en cada DMU, suponiendo como output una media poblacional y como diseño muestreo aleatorio simple sin reemplazamiento:

```
% Este fichero tiene por objeto determinar el tamaño de muestra necesario  
% para cada output en cada DMU, aplicando Teorema 2 suponiendo que la  
% muestra se va a tomar con m.a.s. sin  
% reemplazamiento  
% Se supone como output una media poblacional.  
% 1-alfa: Nivel de confianza deseada  
% delta: Error de estimación del índice de eficiencia deseado  
% Ni: tamaño de la población de la DMU  
% estmuout: Estimación previa de la media del output  
% varmuoutp: Estimación previa de la varianza del output  
  
function n=CalcTammas(alfa , delta ,Ni, estmuout , varmuoutp)  
a1=(1-alfa)^(1/3);  
alfa2=1-a1;  
pcemu=1-(alfa2/2);  
k=norminv(pcemu);  
p=delta/(2-delta);  
B=p*estmuout;
```

```
n0=k*k*(varmuoutp)/(B*B);
n=round(n0/((n0/Ni)+1));
end
```

- III. Programa para calcular el tamaño de muestra, en cada DMU, para obtener regiones de confianza DEA utilizando el Teorema 5 suponiendo como output una media poblacional y como diseño muestreo aleatorio simple sin reemplazamiento,:

```
% Este fichero tiene por objeto determinar el tamaño de muestra necesario
% para cada output en cada DMU, aplicando Teorema 5 suponiendo que la muestra
% se va a tomar con m.a.s. sin reemplazamiento
% Se supone como output una media poblacional.
% 1-alfa: Nivel de confianza deseada
% delta: Error de estimación del índice de eficiencia deseado
% Ni: tamaño de la población de cada DMU, vector columna
% estmuout: Estimación previa de la media del output en cada DMU (vector
% columna)
% varmuoutp: Estimación previa de la varianza del output en cada DMU (vector
% columna)
% NB: Número de DMUs

function n=tammueRC(alfa , delta , Ni, estmuout , varmuoutp , NB)
p=delta/(2-delta);
beta=1-nthroot((1-alfa),NB);
pcemu=1-(beta/2);
k=norminv(pcemu);
Bi=p*estmuout;
for b=1:NB
n0(b,1)=k*k*(varmuoutp(b,1))/(Bi(b,1)*Bi(b,1));
n(b,1)=round(n0(b,1)/((n0(b,1)/Ni(b,1))+1));
end
end
```

B.3. Índices de eficiencia DEA. Modelos BCC y CCR.

- IV. El siguiente programa permite calcular los índices de eficiencia DEA en los modelos BCC y CCR tanto para la orientación input como output.

```
% Esta es la función que genera los índices BCC y CCR para las orientaciones
% input y output
% X: Datos input (en columna)
% Y: Datos output (en columna)
function i=modelDEA(X, Y)
% Selección del modelo;
model='BCC'; % Elección del modelo 'BCC' o 'CCR' ;
orientation='oo'; % orientation = 'oo' (output) 'io' (input) ;
epsilon=.001; % epsilon (numero no arquimediano) ;
% Determinación del número de DMUs, inputs y outputs;
[n,m] = size(X);
[n,s] = size(Y);
% Computación de los resultados para el modelo;
% Los resultados para el modelo seleccionado se incluyen en la matriz Z o i;
switch model;
% Modelo BCC ;
case ('BCC')
switch orientation;
% Orientación Input ;
```

```

case ('io')
Z = zeros(n,n+m+s+1);
% Funcion objetivo para el modelo BCC-io:
%      min(0*lambda -epsilon*(s+ + s-)+theta);
f = [zeros(1,n) -epsilon*ones(1,s+m) 1];
lblambda = zeros(n,1); % Cotas inferiores para (n) lambdas;
lboutput = zeros(s,1); % Cotas inferiores para (s) outputs;
lbinput = zeros(m,1); % Cotas inferiores para (m) inputs ;
lb=[lblambda; lboutput; lbinput]; % Cotas inferiores para , outputs (s+)
% y inputs (s-);
for j=1:n
Aeq = [Y', -eye(s,s), zeros(s,m+1);
-X', zeros(m,s), -eye(m,m) X(j,:)'];
ones(1,n), zeros(1,s), zeros(1,m+1)];
beq = [Y(j,:)'; zeros(m,1);1];
z = linprog(f,[],[],Aeq,beq,lb);
Z(j,:) = z;
end
Z;
[s1, s2]=size(Z);
i=(Z(:,s2));

case ('oo')
Z = zeros(n,n+m+s+1);
% Funcion objetivo para el modelo BCC_oo:
%      max(0*lambda+epsilon*(s+ + s-)+theta);
f = [-zeros(1,n), epsilon*ones(1,s+m), 1];
lblambda = zeros(n,1); % Cotas inferiores para (n) lambdas;
lboutput = zeros(s,1); % Cotas inferiores para (s) outputs;
lbinput = zeros(m,1); % Cotas inferiores para (m) inputs ;
lb=[lblambda; lboutput; lbinput]; % Cotas inferiores para lambdas , outputs (s+)
% y inputs (s-);
for j=1:n
Aeq = [-Y', eye(s,s), zeros(s,m), Y(j,:)'];
X', zeros(m,s), eye(m,m), zeros(m,1);
ones(1,n), zeros(1,s+m+1)];
beq = [zeros(s,1);X(j,:)';1];
z = linprog(f,[],[],Aeq,beq,lb);
Z(j,:) = z;
end
Z;
[s1, s2]=size(Z);
i=1./(Z(:,s2));
end

% Modelo CCR ;
case ('CCR')
switch orientation;
% Orientacion Input ;
case ('io')
Z = zeros(n,n+m+s+1);
% Funcion objetivo para el modelo CCR-io:
%      min(0*lambda-epsilon*(s+ + s-)+theta);
f = [zeros(1,n) -epsilon*ones(1,s) -epsilon*ones(1,m) 1];
lblambda = zeros(n,1); % Cotas inferiores para (n) lambdas;
lboutput = zeros(s,1); % Cotas inferiores para (s) outputs;
lbinput = zeros(m,1); % Cotas inferiores para (m) inputs ;
lb = [lblambda; lboutput; lbinput]; % Cotas inferiores para , outputs (s+)
% y inputs (s-);
for j=1:n

```

```

Aeq = [Y', -eye(s,s), zeros(s,m+1);
-X', zeros(m,s), -eye(m,m), X(j,:)'];
beq = [Y(j,:)'; zeros(m,1)];
z = linprog(f,[],[],Aeq,beq,lb);
Z(j,:) = z;
end
Z;
[s1, s2]=size(Z);
i=(Z(:,s2));
% Orientacion Output;
case ('oo')
Z = zeros(n,n+m+s+1);
% Funcion objetivo para el modelo CCR_oo:
% max(0*lambda+epsilon*(s+ + s-)+theta);
f = [-zeros(1,n), epsilon*ones(1,s+m), 1];
lblambda = zeros(n,1); % Cotas inferiores para (n) lambdas;
lboutput = zeros(s,1); % Cotas inferiores para (s) outputs;
lbinput = zeros(m,1); % Cotas inferiores para (m) inputs;
lb = [lblambda; lboutput; lbinput]; % Cotas inferiores para, outputs (s+)
% y inputs (s-);
for j=1:n
Aeq = [-Y', eye(s,s), zeros(s,m), Y(j,:)']; ...
X', zeros(m,s), eye(m,m), zeros(m,1)];
beq = [zeros(s,1);X(j,:)'];
z = linprog(f,[],[],Aeq,beq,lb);
Z(j,:) = z;
end
Z;
[s1, s2]=size(Z);
i = 1./(Z(:,s2));
end
end

```

B.4. Intervalos de confianza de eficiencia DEA

- V. Programa para obtener el ICE *Tipol* para los modelos CCR y BCC con orientación output, en el caso de dos outputs estimados con una muestra probabilística:

```

% Esta funcion genera ICEs Tipol en el caso de 2 outputs estimados con
% intervalos de confianza
% X matriz de inputs
% icoutp1: Intervalos de confianza para el output1 (primera columna extremo
% inferior, segunda columna extremo superior)
% icoutp2: Intervalos de confianza para el output2 (primera columna extremo
% inferior, segunda columna extremo superior)

function [icT1bcc, icT1ccr]=ICETipol(X, icoutp1, icoutp2)
[NB, NCB]=size(icoutp1);
icT1bcc=zeros(NB,2);
icT1ccr=zeros(NB,2);
% ICE OPTIMISTA/PESIMISTA Tipol
for b=1:NB
clear estmulmax
clear estmulmin
clear bccinft1
clear bccsupt1
clear bccinft2
clear bccsupt2

```



```

% X matriz de inputs
% icoutp1: Intervalos de confianza para el output1 (primera columna extremo
% inferior, segunda columna extremo superior)
% icoutp2: Intervalos de confianza para el output2 (primera columna extremo
% inferior, segunda columna extremo superior)
function [icT1bcc, icT1ccr]=ICETipo3(X, icoutp1, icoutp2)

[NB, NCB]=size(icoutp1);
icT1bcc=zeros(NB,2);
icT1ccr=zeros(NB,2);

XF=X;
YFsupout1=icoutp1(:,2);
YFinfout1=icoutp1(:,1);

YFsupout2=icoutp2(:,2);
YFinfout2=icoutp2(:,1);
for b=1:NB
clear Xinf
clear Yinf
clear d1
clear d2
Xinf=[XF(b,:);
XF];
Yinfout1=[YFinfout1(b,1)
YFsupout1];

Yinfout2=[YFinfout2(b,1)
YFsupout2];

% Obtengamos extremo inferior del intervalo Tipo3
d1=bccout(Xinf,[Yinfout1,Yinfout2]);
icT1bcc(b,1)=d1(1,1);
if icT1bcc(b,1)>0.9999
icT1bcc(b,1)=1;
end
d2=ccrout(Xinf,[Yinfout1,Yinfout2]);
icT1ccr(b,1)=d2(1,1);
if icT1ccr(b,1)>0.9999
icT1ccr(b,1)=1;
end
end

for b=1:NB
clear Xinf
clear Yinf
clear d1
clear d2
Xinf=[XF(b,:);
XF];
Yinfout1=[YFsupout1(b,1)
YFinfout1];

Xinf=[XF(b,:);
XF];
Yinfout2=[YFsupout2(b,1)
YFinfout2];
% Obtengamos extremo superior del intervalo Tipo3
d1=bccout(Xinf,[Yinfout1,Yinfout2]);
icT1bcc(b,2)=d1(1,1);

```

```

if icT1bcc(b,2) > 0.9999
icT1bcc(b,2) = 1;
end
d2 = ccroun(Xinf, [Yinfout1, Yinfout2]);
icT1ccr(b,2) = d2(1,1);
if icT1ccr(b,2) > 0.9999
icT1ccr(b,2) = 1;
end
end
end

```

VIII. Programa para obtener el ICE bootstrap para los modelos CCR y BCC con orientación output, en el caso de dos outputs estimados con una muestra:

```

% Esta funcion genera ICEs bootstrap una vez que tenemos la muestra
% sobre la que remuestrear
% yiout1: contendra los datos muestrales para el primer output; suponemos
% que cada columna tiene los datos de la correspondiente DMU y la dimension
% sera columnas=numero de DMUs filas=tamano de muestra maximo de entre todas
% las DMUs, por tanto, las columnas, salvo la del maximo tamano muestral,
% tendran ceros, por tanto necesitamos el tamano de muestra para eliminar los
% ceros.
% yiout2: contendra los datos muestrales para el segundo output; suponemos
% que cada columna tiene los datos de la correspondiente DMU y la dimension
% sera columnas=numero de DMUs filas=tamano de muestra maximo de entre todas
% las DMUs, por tanto, las columnas, salvo la del maximo tamano muestral,
% tendran ceros, por tanto necesitamos el tamano de muestra para eliminar
% los ceros.
% ni: Tamanos de la muestra en cada DMU
% Ni: Tamano poblacion
% rm: Numero de remuestras boosstraap
% confic: Confianza del intervalo percentil bootstrap
% alfa: confianza de los intervalos de confianza de los outputs

function [icpercbcc, icpercocr] = ICEbootstrap(yiout1, yiout2, X, ni, Ni, rm, alfa, confic)
nmax = max(ni);

[NNb, NB] = size(yiout1);

dd = 100 - confic;

pcemu = 1 - (alfa / (2));
k = norminv(pcemu);

confic = 100 - (dd);

icpercbcc = zeros(NB, 2);
icpercocr = zeros(NB, 2);

clear estef

for j = 1:rm
clear estmuby1out1
clear yblout1
clear Byblout1
clear icsatmed95out1
clear vefrem1out1

clear estmuby1out2
clear yblout2

```

```

clear Byblout2
clear icsatmed95out2
clear vefrem1out2

% Matriz de varianzas de las rm muestras
vefrem1=zeros(NB,1);
% Matriz de varianzas de las 25 remuestras de cada
% una de las rm muestras

% Tomamos la remuestra
yb1out1=zeros(nmax,NB);
yb1out2=zeros(nmax,NB);
for b=1:NB
clear mcon

% Remuestra con reemplazamiento
mcon=mascon(ni(b,1),ni(b,1));
yb1out1(1:ni(b,1),b)=yiout1(mcon,b);

yb1out2(1:ni(b,1),b)=yiout2(mcon,b);

estmuyb1out1(b,1)=sum(yb1out1(:,b))/ni(b,1);
if estmuyb1out1(b,1)<0
estmuyb1out1(b,1)=0;
end

estmuyb1out2(b,1)=sum(yb1out2(:,b))/ni(b,1);
if estmuyb1out2(b,1)<0
estmuyb1out2(b,1)=0;
end
% error de estimacion con cada remuestra dependiendo de k

Byblout1(b,1)=k*sqrt((1-(ni(b,1)/Ni(b,1)))*var(yb1out1(1:ni(b,1),1))/ni(b,1));
icsatmed95out1(b,1)=estmuyb1out1(b,1)-Byblout1(b,1);
icsatmed95out1(b,2)=estmuyb1out1(b,1)+Byblout1(b,1);
if icsatmed95out1(b,1)<0
icsatmed95out1(b,1)=0;
end

Byblout2(b,1)=k*sqrt((1-(ni(b,1)/Ni(b,1)))*var(yb1out2(1:ni(b,1),1))/ni(b,1));
icsatmed95out2(b,1)=estmuyb1out2(b,1)-Byblout2(b,1);
icsatmed95out2(b,2)=estmuyb1out2(b,1)+Byblout2(b,1);
if icsatmed95out2(b,1)<0
icsatmed95out2(b,1)=0;
end

end
% estimamos el Indice de eficiencia con la remuestra

estefyb1(1:NB,j)=bccout(X,[estmuyb1out1,estmuyb1out2]);
estefyb2(1:NB,j)=ccrout(X,[estmuyb1out1,estmuyb1out2]);
for b=1:NB
if estefyb1(b,j)>0.9999
estefyb1(b,j)=1;
end
if estefyb2(b,j)>0.9999
estefyb2(b,j)=1;
end
end

```



```

ampTipo1bcc005=zeros (NB, NIC) ;
ampTipo3bcc005=zeros (NB, NIC) ;
ampTipo2bcc005=zeros (NB, NIC) ;

ampbootpercccr01=zeros (NB, NIC) ;
ampTipo1ccr01=zeros (NB, NIC) ;
ampTipo3ccr01=zeros (NB, NIC) ;
ampTipo2ccr01=zeros (NB, NIC) ;

ampbootpercccr005=zeros (NB, NIC) ;
ampTipo1ccr005=zeros (NB, NIC) ;
ampTipo3ccr005=zeros (NB, NIC) ;
ampTipo2ccr005=zeros (NB, NIC) ;

% Tamano de muestra
pcemu=1-(alfa/(2));
k=norminv (pcemu);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Indice de eficiencia poblacional
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Y=[medoutY1 medoutY2];

%Indice de eficiencia poblacional
bccpob=bccout (Xi, Y);
for b=1:NB
if bccpob (b, 1) >0.9999
bccpob (b, 1) =1;
end
end

ccrpob=ccrout (Xi, Y);
for b=1:NB
if ccrpob (b, 1) >0.9999
ccrpob (b, 1) =1;
end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calculamos tamano muestra en cada DMU
% Hay que fijar el error de estimacion para estimar la media
B=0.9;
for i=1:NB
ni12out1 (i, 1)=masnvar (Ni12 (i, 1), k, B, vY1 (i, 1));
ni12out2 (i, 1)=masnvar (Ni12 (i, 1), k, B, vY2 (i, 1));
ni12 (i, 1)=max (ni12out1 (i, 1), ni12out2 (i, 1));
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Variables que miden las confianzas individuales

contbootperbcc=zeros (NB, 1) ;
contTipo1bcc=zeros (NB, 1) ;
contTipo3bcc=zeros (NB, 1) ;
contTipo2bcc=zeros (NB, 1) ;

contbootpercccr=zeros (NB, 1) ;
contTipo1ccr=zeros (NB, 1) ;
contTipo3ccr=zeros (NB, 1) ;
contTipo2ccr=zeros (NB, 1) ;

for numic=1:NIC

```

```

clear em1out1
clear estmul
clear em1out2
clear estmu2
% en realidad es la proporcion de la media poblacional que estoy dispuesto a
% cometer como error es Bp*media sat poblacional
nmax=max(ni12);
% Variable que donde guardamos la muestra de cada DMU para remuestrear
clear y12out1
clear y12out2
y12out1=zeros(nmax,NB);
y12out2=zeros(nmax,NB);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Tomamos la muestra

for b=1:NB
clear mue
clear resp1
clear resp2

mue=mas(Ni(b,1), ni12(b,1));
resp1=Yi1(mue,b);
resp2=Yi2(mue,b);

% Guardemos los datos muestrales para remuestrear
y12out1(1:ni12(b,1),b)=resp1(1:ni12(b,1),1);
y12out2(1:ni12(b,1),b)=resp2(1:ni12(b,1),1);

vp1(b,1)=var(resp1);
vp2(b,1)=var(resp2);

estmul(b,1)=mean(resp1);
estmu2(b,1)=mean(resp2);
estmup1(b,1)=estmul(b,1);
estmup2(b,1)=estmu2(b,1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calculemos los intervalos de observaciones output
% error de muestreo
em1out1(b,1)=sqrt((1-(ni12(b,1)/Ni12(b,1)))*vp1(b,1)/ni12(b,1));
em1out2(b,1)=sqrt((1-(ni12(b,1)/Ni12(b,1)))*vp2(b,1)/ni12(b,1));
% error de estimacion
B1out1(b,1)=k*em1out1(b,1);
B1out2(b,1)=k*em1out2(b,1);

icsatmed95out1(b,1)=estmul(b,1)-B1out1(b,1);
if icsatmed95out1(b,1)<0
icsatmed95out1(b,1)=0;
end

icsatmed95out1(b,2)=estmul(b,1)+B1out1(b,1);

icsatmed95out2(b,1)=estmu2(b,1)-B1out2(b,1);
if icsatmed95out2(b,1)<0
icsatmed95out2(b,1)=0;
end

icsatmed95out2(b,2)=estmu2(b,1)+B1out2(b,1);

```



```

contTipo3bcc(b,1)=contTipo3bcc(b,1)+1;
end
end

for b=1:NB
if bccpob(b,1)>=icpercbcc(b,1)& bccpob(b,1)<=icpercbcc(b,2)
contbootpercbcc(b,1)=contbootpercbcc(b,1)+1;
end
end

for b=1:NB
if ccrpob(b,1)>=icTipo2ccr(b,1)& ccrpob(b,1)<=icTipo2ccr(b,2)
contTipo2ccr(b,1)=contTipo2ccr(b,1)+1;
end
end

for b=1:NB
if ccrpob(b,1)>=icTipo1ccr(b,1)& ccrpob(b,1)<=icTipo1ccr(b,2)
contTipo1ccr(b,1)=contTipo1ccr(b,1)+1;
end
end

for b=1:NB
if ccrpob(b,1)>=icTipo3ccr(b,1)& ccrpob(b,1)<=icTipo3ccr(b,2)
contTipo3ccr(b,1)=contTipo3ccr(b,1)+1;
end
end

for b=1:NB
if ccrpob(b,1)>=icpercccr(b,1)& ccrpob(b,1)<=icpercccr(b,2)
contbootpercccr(b,1)=contbootpercccr(b,1)+1;
end
end

ampbootbcc01(:,numic)=ampbootpercbcc;
ampTipo1bcc01(:,numic)=ampTipo1bcc;
ampTipo3bcc01(:,numic)=ampTipo3bcc;
ampTipo2bcc01(:,numic)=ampTipo2bcc;

ampbootpercccr01(:,numic)=ampbootpercccr;
ampTipo1ccr01(:,numic)=ampTipo1ccr;
ampTipo3ccr01(:,numic)=ampTipo3ccr;
ampTipo2ccr01(:,numic)=ampTipo2ccr;

end
contbootpercbcc01=(contbootpercbcc/NIC)*100;
contTipo1bcc01=(contTipo1bcc/NIC)*100;
contTipo3bcc01=(contTipo3bcc/NIC)*100;
contTipo2bcc01=(contTipo2bcc/NIC)*100;

contbootpercccr01=(contbootpercccr/NIC)*100;
contTipo1ccr01=(contTipo1ccr/NIC)*100;
contTipo3ccr01=(contTipo3ccr/NIC)*100;
contTipo2ccr01=(contTipo2ccr/NIC)*100;

save confbcc01 contbootpercbcc01 contTipo1bcc01 ...
contTipo3bcc01 contTipo2bcc01
save ampbcc01 ni12 ampbootbcc01 ampTipo1bcc01 ...
ampTipo3bcc01 ampTipo2bcc01

```

```

save confccr01      contbootpercccr01      contTipo1ccr01 ...
contTipo3ccr01      contTipo2ccr01
save ampeccr01      ni12      ampbootccr01      ampTipo1ccr01 ...
ampTipo3ccr01      ampTipo2ccr01

```

- X. Programa para obtener los resultados del estudio de aplicación de los ICEs, con los datos del Ministerio de Sanidad, de la sección 4.3.8.

```

% Ejemplo de aplicacion 4.3.8 de obtencion de ICEs con datos del Ministerio

clear
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Confianza 90%
alfa=0.1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Numero de DMUs
NB=18;

% Remuestras bootstrap
rm=1000;

% Cargamos datos outputs
load outputP4BS09
% Y (datos) Ni (tamano poblacion) medout (satisfaccion media) varout
% (varianza de la satisfaccion media)

% inputs
load input2ACP
% dat2ACP tiene los inputs (persona , gastos)
X=dat2ACP;
N=sum(Ni);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Los datos se toman como muestrales
ni=Ni;
yi=Y;
k=1.96;

% Indice de eficiencia estimado
% BCC
bccest=bccoutp(X, medout);
% CCR
ccrest=ccroutp(X, medout);

for b=1:NB
if bccest(b,1)>0.9999
bccest(b,1)=1;
end
end

for b=1:NB
if ccrest(b,1)>0.9999
ccrest(b,1)=1;
end
end

% Obtenemos los ICEs para p=0.1
p=0.1;

```



```

% Numero de iteraciones , cuantos errores de estimacion de cada DMU vamos
% a generar
NIC=1000;

% Cargamos datos poblacionales
load datcoopmasdat
% save datcoop Ni Yi1 Yi2 Xi medoutY1 medoutY2 vY1 vY2

% Tamano de muestra para confianza raiz 12 de 1-ALFA
pcemu=1-(alfa/2);
k=norminv(pcemu);

N=sum(Ni);

%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
Y=[medoutY1 medoutY2];

% Indice de eficiencia poblacional
% BCC
bccpob=bccout(Xi,Y);
for b=1:NB
if bccpob(b,1)>0.9999
bccpob(b,1)=1;
end
end
% CCR
ccrpob=ccrout(Xi,Y);
for b=1:NB
if ccrpob(b,1)>0.9999
ccrpob(b,1)=1;
end
end

icinfbcc=zeros(NB,NIC);
icsupbcc=zeros(NB,NIC);
cont005bcc=zeros(NB,1);
cont01bcc=zeros(NB,1);
icinfbccr=zeros(NB,NIC);
icsupccr=zeros(NB,NIC);
cont005ccr=zeros(NB,1);
cont01ccr=zeros(NB,1);

% Guardamos en una matriz la estimacion de la media y la varianza que
% utilizamos para el calculo del tamano de la muestra
estmup=zeros(NB,NIC);
vp=zeros(NB,NIC);

% Para verificar Lema 5
indbcc=zeros(NB,NIC);
indccr=zeros(NB,NIC);

%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
% Tomamos muestra piloto de 25 individuos
for i=1:NB
ni0(i,1)=25;
end

% Tomemos la muestra y verifiquemos el error de estimacion del indice de
% eficiencia
contbcc=zeros(NB,1);

```



```

estmup2(b,h)=estmu2(b,1);
vp2(b,h)=var(resp2);

end

%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
% Indice de eficiencia estimado con la muestra
% BCC
estefbcc=bccout(Xi,[estmu1 estmu2]);
for b=1:NB
if estefbcc(b,1)>0.9999
estefbcc(b,1)=1;

end
if estefbcc(b,1)==1
indbcc(b,h)=1;
end
end
% CCR
estefccr=ccrout(Xi,[estmu1 estmu2]);
for b=1:NB
if estefccr(b,1)>0.9999
estefccr(b,1)=1;

end
if estefccr(b,1)==1
indccr(b,h)=1;
end
end
%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
% Error de estimacion exacto del indice de eficiencia exacto con cada muestra
errestidebcc(:,h)=(bccpob-estefbcc);
errestideccr(:,h)=(ccrpob-estefccr);

end

erridebcc=abs(errestidebcc);
cocerrideyidebcc=zeros(NB,NIC);
for i=1:NIC
cocerrideyidebcc(:,i)=erridebcc(:,i)./bccpob;
end

errideccr=abs(errestideccr);
cocerrideyideccr=zeros(NB,NIC);
for i=1:NIC
cocerrideyideccr(:,i)=errideccr(:,i)./ccrpob;
end

for s=1:NB
perc95bcc(s,1)=prctile(cocerrideyidebcc(s,:),95);
end

for s=1:NB
perc95ccr(s,1)=prctile(cocerrideyideccr(s,:),95);
end

for i=1:NB
for j=1:NIC
if erridebcc(i,j)<=cota;

```

```

contbcc(i,1)=contbcc(i,1)+1;
end
if errideccr(i,j)<=cota;
contccr(i,1)=contccr(i,1)+1;
end
end
end
% Proporción de veces que se verifica la cota del Lema 5 de regiones de
% confianza
contccr=(contccr/NIC)*100;
contbcc=(contbcc/NIC)*100;

save lema1005 ni01 contccr cocerrideyideccr errideccr contbcc ...
cocerrideyidebcc erridebcc

```

XII. Programa para validar el Teorema 5 en la sección 4.5.2.

```

% Fichero construido para verificar teorema 5 de regiones de confianza

clear

% Error de estimacion del indice de eficiencia
delta=0.1;

% correspondiente error de estimacion de la media poblacional
pp=delta/(2-delta);

% Confianza
alfa=0.1;

% Numero de items
NB=12;

% Numero de iteraciones, cuantos errores de estimacion de cada DMU vamos
% a generar
NIC=1000;

% Cargamos datos poblacionales
load datcoopmasdat
% Ni Yi1 Yi2 Xi medoutY1 medoutY2 vY1 vY2

% Tamano de muestra para confianza raiz 12 de 1-ALFA

beta=1-nthroot((1-alfa),NB);
pcemu=1-(beta/2);
k=norminv(pcemu);

N=sum(Ni);

%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
% Outputs poblacionales
Y=[medoutY1 medoutY2];

% Indice de eficiencia poblacional
% BCC
bccpob=bccout(Xi,Y);
for b=1:NB
if bccpob(b,1)>0.9999
bccpob(b,1)=1;
end

```



```

end
% CCR
ccrpob=ccrout (Xi ,Y) ;
for b=1:NB
if ccrpob (b, 1) >0.9999
ccrpob (b, 1) =1;
end
end
% Errores de estimacion del indice de eficiencia guardados en una matriz
icinfbcc=zeros (NB, NIC) ;
icsupbcc=zeros (NB, NIC) ;
cont005bcc=zeros (NB, 1) ;
cont01bcc=zeros (NB, 1) ;
icinfccr=zeros (NB, NIC) ;
icsupccr=zeros (NB, NIC) ;
cont005ccr=zeros (NB, 1) ;
cont01ccr=zeros (NB, 1) ;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Tomamos muestra piloto de 25 individuos
for i=1:NB
ni0 (i, 1) =25;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for ps=1
p=pp (ps, 1) ;

% Guardamos en una matriz la estimacion de la media y la varianza que
% utilizamos para el calculo del tamano de la muestra
estmup=zeros (NB, NIC) ;
vp=zeros (NB, NIC) ;

%Para verificar teorema 5
indbcc=zeros (NB, NIC) ;
indccr=zeros (NB, NIC) ;
cont=zeros (NB, 1) ;

for h=1:NIC
for b=1:NB
clear mue
clear resp1
clear resp2
mue=mas (Ni (b, 1) , ni0 (b, 1)) ;
resp1=Yi1 (mue, b) ;
resp2=Yi2 (mue, b) ;

vp1 (b, 1) =var (resp1) ;
vp2 (b, 1) =var (resp2) ;

estmu1 (b, 1) =mean (resp1) ;
estmu2 (b, 1) =mean (resp2) ;
estmup1 (b, 1) =estmu1 (b, 1) ;
estmup2 (b, 1) =estmu2 (b, 1) ;

end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calculamos el tamano de muestra que vamos a usar en todas las iteraciones
for b=1:NB
clear mue
clear resp1

```



```

%RC BCC para el vector de indices de eficiencia
icinfccr(:,h)=estefccr-delta;
icsupccr(:,h)=estefccr+delta;

end

for i=1:NB
for j=1:NIC
if icinfbcc(i,j)<=bccpob(i,1) & icsupbcc(i,j)>=bccpob(i,1);
cont01bcc(i,j)=1;
end
if icinfccr(i,j)<=ccrpob(i,1) & icsupccr(i,j)>=ccrpob(i,1);
cont01ccr(i,j)=1;
end

end
end
sumacont01bcc=sum(cont01bcc);
contbcc=0;
for i=1:NIC
if sumacont01bcc(1,i)==NB
contbcc=contbcc+1;
end
end

sumacont01ccr=sum(cont01ccr);
contccr=0;
for i=1:NIC
if sumacont01ccr(1,i)==NB
contccr=contccr+1;
end
end

% Proporción de veces que las regiones de confianza contienen
% a todas las eficiencias poblacionales
RCpbcc=(contbcc/NIC)*100;
RCpccr=(contccr/NIC)*100;

end

save th1delta01alfa01 RCpbcc RCpccr NB NIC ni01...
delta p alfa beta vp estmup bccpob ccrpob icinfccr...
icsupccr icinfbcc icsupbcc

```


Bibliografía

- [1] Malcolm Abbott and Chris Doucouliagos. The efficiency of australian universities: a data envelopment analysis. *Economics of Education review*, 22(1):89–97, 2003.
- [2] Agha Iqbal Ali and Lawrence M Seiford. The mathematical programming approach to efficiency analysis. *The measurement of productive efficiency*, pages 120–159, 1993.
- [3] Per Andersen and Niels Christian Petersen. A procedure for ranking efficient units in data envelopment analysis. *Management science*, 39(10):1261–1264, 1993.
- [4] Pascal Ardilly and David Le Blanc. Sampling and weighting a survey of homeless persons: a french example. *Survey methodology*, 27(1):109–118, 2001.
- [5] Hossein Azizi. The interval efficiency based on the optimistic and pessimistic points of view. *Applied Mathematical Modelling*, 35(5):2384–2393, 2011.
- [6] Hossein Azizi and Rasul Jahed. Improved data envelopment analysis models for evaluating interval efficiencies of decision-making units. *Computers & Industrial Engineering*, 61(3): 897–901, 2011.
- [7] Stephen D Ball, Keith Johnson, and Paul Slattery. Labour productivity in hotels: an empirical analysis. *International Journal of Hospitality Management*, 5(3):141–147, 1986.
- [8] Rajiv D Banker. Hypothesis tests using data envelopment analysis. *Journal of productivity analysis*, 7(2-3):139–159, 1996.
- [9] Rajiv D Banker. Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. *Management science*, 39(10):1265–1273, 1993.
- [10] Rajiv D Banker, Abraham Charnes, and William Wager Cooper. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9): 1078–1092, 1984.
- [11] Authella M Bessent, E Wailand Bessent, Abraham Charnes, William W Cooper, and Nellie Carr Thorogood. Evaluation of educational program proposals by means of DEA. *Educational Administration Quarterly*, 19(2):82–107, 1983.
- [12] J Enrique Bigné Alcañiz, Miguel Ángel Moliner Tena, Teresa Vallet Bellmunt, and Javier Sánchez García. Un estudio comparativo de los instrumentos de medición de la calidad de los servicios públicos. 1997.
- [13] D Birch. Public participation in local government a survey of local authorities. 02lg00871, local and regional government research unit, odpm, london, 2002.
- [14] Aziz Boussofiane, Robert G Dyson, and Emmanuel Thanassoulis. Applied data envelopment analysis. *European Journal of Operational Research*, 52(1):1–15, 1991.

- [15] J Michael Brick. The future of survey sampling. *Public Opinion Quarterly*, 75(5):872–888, 2011.
- [16] ME Bruni, D Conforti, P Beraldi, and E Tundis. Probabilistically constrained models for efficiency and dominance in DEA. *International Journal of Production Economics*, 117(1): 219–228, 2009.
- [17] E. Carrillo y M. Tamayo. Valoración del consumo ciudadano de los servicios públicos locales: las encuestas sobre satisfacción con servicios públicos en las ciudades. In *Retos e innovación en gestión de servicios públicos-Servicios locales y colaboración pública-privada organizado por la Universidad Internacional Menéndez Pelayo*, 2008.
- [18] Albert Caruana. Service loyalty: The effects of service quality and the mediating role of customer satisfaction. *European journal of marketing*, 36(7/8):811–828, 2002.
- [19] Mehmet E Ceyhan and James C Benneyan. Handling estimated proportions in public sector data envelopment analysis. *Annals of Operations Research*, 221(1):107–132, 2014.
- [20] David W Chapman. To substitute or not to substitute that is the question. *The Survey Statistician*, 48:32–34, 2003.
- [21] Abraham Charnes and William W Cooper. Chance-constrained programming. *Management science*, 6(1):73–79, 1959.
- [22] Abraham Charnes and William W Cooper. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations research*, 11(1):18–39, 1963.
- [23] Abraham Charnes, William Wager Cooper, and MJL Kirby. Chance-constrained programming: an extension of statistical method. Technical report, DTIC Document, 1971.
- [24] Abraham Charnes, William W Cooper, and Edwardo Rhodes. Measuring the efficiency of decision making units. *European journal of operational research*, 2(6):429–444, 1978.
- [25] Abraham Charnes and William W Cooper. Programming with linear fractional functionals. *Naval Research logistics quarterly*, 9(3-4):181–186, 1962.
- [26] Abraham Charnes and William Wager Cooper. Preface to topics in data envelopment analysis. *Annals of Operations research*, 2(1):59–94, 1984.
- [27] Abraham Charnes. *Data Envelopment Analysis: Theory, Methodology, and Applications: Theory, Methodology and Applications*. Springer Science & Business Media, 1994.
- [28] Arijit Chaudhuri and Horst Stenger. *Survey sampling: theory and methods*. CRC Press, 2005.
- [29] CIS. Calidad de los servicios públicos (vi), 2011.
- [30] Timothy J Coelli, DS Prasada Rao, Christopher J O’Donnell, and George E Battese. Data envelopment analysis. *An Introduction to Efficiency and Productivity Analysis*, pages 161–181, 2005.
- [31] Antón Costas Comesaña. De consumidor a ciudadano: el papel de la satisfacción del ciudadano en la sostenibilidad de los mercados de servicios públicos. *Información Comercial Española, ICE: Revista de economía*, (836):33–50, 2007.
- [32] Tim Coelli. A multi-stage methodology for the solution of orientated DEA models. *Operations Research Letters*, 23(3):143–149, 1998.

- [33] Wade D Cook and Larry M Seiford. Data envelopment analysis (DEA)—thirty years on. *European Journal of Operational Research*, 192(1):1–17, 2009.
- [34] WW Cooper, Zhimin Huang, and Susan X Li. Satisficing DEA models under chance constraints. *Annals of operations research*, 66(4):279–295, 1996.
- [35] William W Cooper, Zhimin Huang, Vedran Lelas, Susan X Li, and Ole B Olesen. Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA. *Journal of Productivity Analysis*, 9(1):53–79, 1998.
- [36] William W Cooper, Kyung Sam Park, and Gang Yu. IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. *Management Science*, 45(4):597–607, 1999.
- [37] William W Cooper, Honghui Deng, Zhimin Huang, Susan X Li, et al. Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. *Journal of the Operational Research Society*, 53(12):1347–1356, 2002.
- [38] William W Cooper, Honghui Deng, Zhimin Huang, and Susan X Li. Chance constrained programming approaches to congestion in stochastic data envelopment analysis. *European Journal of Operational Research*, 155(2):487–501, 2004.
- [39] William W Cooper, Lawrence M Seiford, and Kaoru Tone. *Introduction to data envelopment analysis and its uses: with DEA-solver software and references*. Springer Science & Business Media, 2006.
- [40] William W Cooper, Lawrence M Seiford, and Kaoru Tone. *Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software*. Springer Science & Business Media, 2007.
- [41] William W Cooper, Lawrence M Seiford, and Joe Zhu. *Handbook on data envelopment analysis*, volume 164. Springer Science & Business Media, 2011.
- [42] J Joseph Cronin Jr and Steven A Taylor. Measuring service quality: a reexamination and extension. *The journal of marketing*, pages 55–68, 1992.
- [43] Eloy Cuéllar Martín, Eloísa del Pino Matute, Joaquín Ruíz López, et al. Guía para la evaluación de la calidad de los servicios públicos. 2009.
- [44] Anand Desai, Samuel J Ratick, and Arie P Schinnar. Data envelopment analysis with stochastic variations in data. *Socio-Economic Planning Sciences*, 39(2):147–164, 2005.
- [45] Gerard Debreu. The coefficient of resource utilization. *Econometrica: Journal of the Econometric Society*, pages 273–292, 1951.
- [46] Dimitris K Despotis and Yiannis G Smirlis. Data envelopment analysis with imprecise data. *European Journal of Operational Research*, 140(1):24–36, 2002.
- [47] DGE. Encuesta sobre la satisfacción de los ciudadanos con los servicios prestados por la administración de la comunidad de castilla y león, 2009.
- [48] Don A Dillman and Leah Melani Christian. Survey mode as a source of instability in responses across surveys. *Field methods*, 17(1):30–52, 2005.
- [49] Bradley Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, pages 1–26, 1979.

- [50] Bradley Efron and Robert J Tibshirani. *An introduction to the bootstrap*. CRC press, 1994.
- [51] Ali Emrouznejad, Barnett R Parker, and Gabriel Tavares. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-economic planning sciences*, 42(3):151–157, 2008.
- [52] Tomoe Entani, Yutaka Maeda, and Hideo Tanaka. Dual models of interval DEA and its extension to interval data. *European Journal of Operational Research*, 136(1):32–45, 2002.
- [53] Subdirección General de Estudios Tributarios. Opiniones y actitudes fiscales de los españoles en 2009. 2008.
- [54] Rolf Färe, Shawna Grosskopf, Mary Norris, and Zhongyang Zhang. Productivity growth, technical progress, and efficiency change in industrialized countries. *The American economic review*, pages 66–83, 1994.
- [55] Michael James Farrell. The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, pages 253–290, 1957.
- [56] Gary D Ferrier and Joseph G Hirschberg. Bootstrapping confidence intervals for linear programming efficiency scores: With an illustration using italian banking data. *Journal of Productivity Analysis*, 8(1):19–33, 1997.
- [57] Christopher A Field and Alan H Welsh. Bootstrapping clustered data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(3):369–390, 2007.
- [58] Carlos Figuera, José Luis Rojo-Álvarez, Inmaculada Mora-Jiménez, Alicia Guerrero-Curienes, Mark Wilby, and Javier Ramos-López. Time-space sampling and mobile device calibration for wifi indoor location systems. *Mobile Computing, IEEE Transactions on*, 10(7):913–926, 2011.
- [59] Ronald D Fricker and Matthias Schonlau. Advantages and disadvantages of internet research surveys: Evidence from the literature. *Field methods*, 14(4):347–367, 2002.
- [60] Ellen Garbarino and Mark S Johnson. The different roles of satisfaction, trust, and commitment in customer relationships. *the Journal of Marketing*, pages 70–87, 1999.
- [61] William H Greene. Maximum likelihood estimation of econometric frontier functions. *Journal of econometrics*, 13(1):27–56, 1980.
- [62] Tjalling C Koopmans. Analysis of production as an efficient combination of activities. *Activity analysis of production and allocation*, 13:33–37, 1951.
- [63] J. García-Laguna, J. Sáez, y J.A. Tapia. Modelos de análisis envolvente de datos para el análisis de factores de eficiencia. aplicación en items de las universidades de Castilla y León. Proyecto financiado por la Consejería de Hacienda de la Junta de Castilla y León. (Ref. OTP/04/03).
- [64] J. García-Laguna, J. Salvador y J.A. Tapia. Estudio estadístico del sistema bibliotecario público en Castilla y León. Proyecto financiado por la Consejería de Hacienda de la Junta de Castilla y León. (Ref. OTP/06/01)., 2006.
- [65] Irene Gijbels, Enno Mammen, Byeong U Park, and Leopold Simar. On estimation of monotone and concave frontier functions. *Journal of the American Statistical Association*, 94(445):220–228, 1999.
- [66] Daniela Golinelli, Joan S Tucker, Gery W Ryan, and Suzanne L Wenzel. Strategies for obtaining probability samples of homeless youth. *Field Methods*, 27(2):131–143, 2014.

- [67] Madelyn S Gould, Sylvan Wallenstein, and Marjorie Kleinman. Time-space clustering of teenage suicide. *American Journal of Epidemiology*, 131(1):71–78, 1990.
- [68] Evangelos Grigoroudis and Yannis Siskos. *Customer satisfaction evaluation: methods for measuring and implementing service quality*, volume 139. Springer Science & Business Media, 2009.
- [69] Shawna Grosskopf. Statistical inference and nonparametric efficiency: A selective survey. *Journal of Productivity Analysis*, 7(2-3):161–176, 1996.
- [70] Peter Hall and Léopold Simar. Estimating a changepoint, boundary, or frontier in the presence of observation error. *Journal of the American statistical Association*, 97(458):523–534, 2002.
- [71] Bob E Hayes. *Measuring customer satisfaction and loyalty: survey design, use, and statistical analysis methods*. ASQ Quality Press, 2008.
- [72] Nigel Hill, Jim Alexander, and Asociación Española de Normalización y Certificación. *Manual de satisfacción del cliente y evaluación de la fidelidad*. Aenor, 2003.
- [73] Kristen Himelein, Stephanie Eckman, and Siobhan Murray. Sampling nomads: A new technique for remote, hard-to-reach, and mobile populations. *Journal of Official Statistics*, 30(2):191–213, 2014.
- [74] L Hosseinzadeh, N Nematollahi, MH Behzadi, M Mirbolouki, and Zohreh Moghaddas. Centralized resource allocation with stochastic data. *Journal of Computational and Applied Mathematics*, 236(7):1783–1788, 2012.
- [75] Zhimin Huang and Susan X Li. Stochastic DEA models with different types of input-output disturbances. *Journal of Productivity Analysis*, 15(2):95–113, 2001.
- [76] GR Jahanshahloo and S Abbasian-Naghneh. Data envelopment analysis with imprecise data. *Applied Mathematical Sciences*, 5(61-64):3089–3106, 2011.
- [77] Jill Johnes. Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review*, 25(3):273–288, 2006.
- [78] Graham Kalton. Practical methods for sampling rare and mobile populations. In *Proceedings of the Annual Meeting of the American Statistical Association*, pages 5–9, 2001.
- [79] Graham Kalton. Sampling flows of mobile human populations. *Survey Methodology*, 17(2):183–194, 1991.
- [80] Graham Kalton. Models in the practice of survey sampling. *International Statistical Review/Revue Internationale de Statistique*, pages 175–188, 1983.
- [81] Chiang Kao. Interval efficiency measures in data envelopment analysis with imprecise data. *European Journal of Operational Research*, 174(2):1087–1099, 2006.
- [82] Chiang Kao and Shiang-Tai Liu. Stochastic data envelopment analysis in measuring the efficiency of taiwan commercial banks. *European Journal of Operational Research*, 196(1):312–322, 2009.
- [83] Mohammad Khodabakhshi. Super-efficiency in stochastic data envelopment analysis: An input relaxation approach. *Journal of computational and applied mathematics*, 235(16):4576–4588, 2011.

- [84] Mohammad Khodabakhshi. An output oriented super-efficiency measure in stochastic data envelopment analysis: Considering iranian electricity distribution companies. *Computers & Industrial Engineering*, 58(4):663–671, 2010.
- [85] M Khodabakhshi. Estimating most productive scale size with stochastic data in data envelopment analysis. *Economic Modelling*, 26(5):968–973, 2009.
- [86] M Khodabakhshi and M Asgharian. An input relaxation measure of efficiency in stochastic data envelopment analysis. *Applied Mathematical Modelling*, 33(4):2010–2023, 2009.
- [87] M Khodabakhshi, M Asgharian, and Greg N Gregoriou. An input-oriented super-efficiency measure in stochastic data envelopment analysis: Evaluating chief executive officers of us public banks and thrifts. *Expert Systems with Applications*, 37(3):2092–2097, 2010.
- [88] Alois Kneip, Byeong U Park, and Léopold Simar. A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Econometric theory*, 14(06):783–793, 1998.
- [89] Chuen Tse Kuah, Kuan Yew Wong, and Wai Peng Wong. Monte carlo data envelopment analysis with genetic algorithm for knowledge management performance measurement. *Expert Systems with Applications*, 39(10):9348–9358, 2012.
- [90] Kenneth C Land, CA Lovell, and Sten Thore. Chance-constrained data envelopment analysis. *Managerial and Decision Economics*, 14(6):541–554, 1993.
- [91] Erich Leo Lehmann and George Casella. *Theory of point estimation*, volume 31. Springer Science & Business Media, 1998.
- [92] Paul S Levy and Stanley Lemeshow. *Sampling of populations: methods and applications*. John Wiley & Sons, 2013.
- [93] John S Liu, Louis YY Lu, Wen-Min Lu, and Bruce JY Lin. A survey of DEA applications. *Omega*, 41(5):893–902, 2013.
- [94] John S Liu, Louis YY Lu, Wen-Min Lu, and Bruce JY Lin. Data envelopment analysis 1978–2010: A citation-based literature survey. *Omega*, 41(1):3–15, 2013.
- [95] CA Knox Lovell. Production frontiers and productive efficiency. *The measurement of productive efficiency: techniques and applications*, pages 3–67, 1993.
- [96] CA Knox Lovell and Jesús T Pastor. Units invariant and translation invariant DEA models. *Operations Research Letters*, 18(3):147–151, 1995.
- [97] Duncan A MacKellar, Kathleen M Gallagher, Teresa Finlayson, Travis Sanchez, Amy Lansky, and Patrick S Sullivan. Surveillance of hiv risk and prevention behaviors of men who have sex with men a national application of venue-based, time-space sampling. *Public Health Reports*, 122(Suppl 1):39, 2007.
- [98] Gordon Mansergh, Sathapana Naorat, Rapeepun Jommaroeng, Richard A Jenkins, Supaporn Jeeyapant, Kamolset Kangarnrua, Praphan Phanuphak, Jordan W Tappero, and Frits Van Griensven. Adaptation of venue-day-time sampling in southeast asia to access men who have sex with men for hiv assessment in bangkok. *Field Methods*, 18(2):135–152, 2006.
- [99] Marcelino Martínez Cabrera. *La medición de la eficiencia en las instituciones de educación superior*. Fundacion BBVA/BBVA Foundation, 2003.
- [100] Peter McCullagh. Resampling and exchangeable arrays. *Bernoulli*, pages 285–301, 2000.

- [101] Georgios Mihelis, Evangelos Grigoroudis, Yannis Siskos, Yannis Politis, and Yiannis Malandrakis. Customer satisfaction measurement in the private bank sector. *European Journal of Operational Research*, 130(2):347–360, 2001.
- [102] MSC Ministerio de Sanidad y Consumo. Barómetro sanitario 2009, 2010.
- [103] Hiroshi Morita and Lawrence M Seiford. Characteristics on stochastic DEA efficiency: Reliability and probability being efficient. *Journal of the Operations Research Society of Japan*, 42(4):389–404, 1999.
- [104] Farzana B Muhib, Lillian S Lin, Ann Stueve, Robin L Miller, Wesley L Ford, Wayne D Johnson, Philip J Smith, Community Intervention Trial for Youth Study Team, et al. A venue-based method for sampling hard-to-reach populations. *Public health reports*, 116(Suppl 1):216, 2001.
- [105] N Krishnan Namboodiri. *Survey sampling and measurement*. Elsevier, 2013.
- [106] Ole Bent Olesen. Comparing and combining two approaches for chance constrained DEA. *Journal of Productivity Analysis*, 26(2):103–119, 2006.
- [107] Ole Bent Olesen and NC Petersen. Chance constrained efficiency evaluation. *Management Science*, 41(3):442–457, 1995.
- [108] Arun Parasuraman, Valarie A Zeithaml, and Leonard L Berry. Servqual. *Journal of retailing*, 64(1):12–40, 1988.
- [109] F Pedraja, FJ Salinas, y Javier Suárez. La medición de la eficiencia en el sector público. In *La medición de la eficiencia y la productividad*, pages 243–268. Ediciones Pirámide, 2001.
- [110] Antonio Álvarez Pinilla. Concepto y medición de la eficiencia productiva. In *La medición de la Eficiencia y la Productividad*, pages 19–40. Ediciones Pirámide, 2001.
- [111] Antonio Álvarez Pinilla. *La medición de la eficiencia y la productividad*. Ediciones Pirámide, 2014.
- [112] Ramu Ramanathan. *An introduction to data envelopment analysis: a tool for performance measurement*. Sage, 2003.
- [113] Danielle E Ramo, Christian Grov, Kevin Delucchi, Brian C Kelly, and Jeffrey T Parsons. Typology of club drug use among young adults recruited using time–space sampling. *Drug and alcohol dependence*, 107(2):119–127, 2010.
- [114] Jon NK Rao and CFJ Wu. Resampling inference with complex survey data. *Journal of the american statistical association*, 83(401):231–241, 1988.
- [115] Subhash C Ray. *Data envelopment analysis: theory and techniques for economics and operations research*. Cambridge University Press, 2004.
- [116] L Rubio. La percepción del ciudadano en la mejora de la calidad de los servicios públicos. *Pensar lo público*, pages 353–379, 2000.
- [117] A Ruiz, C Hidalgo, MJ Jiménez, y J Ruiz. Guía de orientación para la realización de estudios de análisis de la demanda y de encuestas de satisfacción. *Ministerio de Administraciones Públicas*, 2006.

- [118] Carl-Erik Särndal, Bengt Swensson, and Jan Wretman. *Model assisted survey sampling*. Springer Science & Business Media, 2003.
- [119] Reza Farzipoor Saen. Developing a new data envelopment analysis methodology for supplier selection in the presence of both undesirable outputs and imprecise data. *The International Journal of Advanced Manufacturing Technology*, 51(9-12):1243–1250, 2010.
- [120] Priscilla Salant and Don A Dillman. *How to conduct your own survey*. Wiley New York, 1994.
- [121] Richard Scheaffer, William Mendenhall III, R Ott, and Kenneth Gerow. *Elementary survey sampling*. Cengage Learning, 2011.
- [122] Lawrence M Seiford. A bibliography for data envelopment analysis (1978-1996). *Annals of Operations Research*, 73:393–438, 1997.
- [123] Lawrence M Seiford. Data envelopment analysis: the evolution of the state of the art (1978–1995). *Journal of Productivity Analysis*, 7(2-3):99–137, 1996.
- [124] Jati K Sengupta. Data envelopment analysis for efficiency measurement in the stochastic case. *Computers & operations research*, 14(2):117–129, 1987.
- [125] Leopold Simar and Paul W Wilson. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, 136(1):31–64, 2007.
- [126] Leopold Simar and Paul W Wilson. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management science*, 44(1):49–61, 1998.
- [127] Léopold Simar and Paul W Wilson. A general methodology for bootstrapping in non-parametric frontier models. *Journal of applied statistics*, 27(6):779–802, 2000.
- [128] Léopold Simar and Valentin Zelenyuk. Stochastic FDH/DEA estimators for frontier analysis. *Journal of Productivity Analysis*, 36(1):1–20, 2011.
- [129] Ravindra Singh and Naurang Singh Mangat. *Elements of survey sampling*, volume 15. Springer Science & Business Media, 1996.
- [130] Nevena Stancheva and Vyara Angelova. Measuring the efficiency of university libraries using data envelopment analysis. *University of Economics-Vane*, 2004.
- [131] Ann Stueve, Lydia N O'Donnell, Richard Duran, Alexi San Doval, and Juliana Blome. Time-space sampling in minority communities: results with young latino men who have sex with men. *American journal of public health*, 91(6):922, 2001.
- [132] Seymour Sudman. Improving the quality of shopping center sampling. *Journal of Marketing research*, pages 423–431, 1980.
- [133] RA Sugden and TMF Smith. Domains of study and poststratification. *Journal of statistical planning and inference*, 136(9):3307–3317, 2006.
- [134] Jesus Alberto Tapia and Jose Antonio Menéndez. Probabilistic web survey methodology in education centers: An example in spanish schools. *Survey Practice*, 8(3), 2015.
- [135] Jesus Alberto Tapia, E. Cubo, y J. Espinosa. Encuesta de calidad de las bibliotecas de la universidad de Valladolid. Proyecto de investigación financiado por el rectorado de la UVA 2004

- [136] Jesus Alberto Tapia y S. Sacristán. Encuesta de satisfacción de usuarios de las bibliotecas de la universidad de Valladolid. Proyecto realizado en el contexto de las becas de colaboración estudiantiles de la Universidad de Valladolid, 2003.
- [137] Madjid Tavana, Rashed Khanjani Shiraz, and Adel Hatami-Marbini. A new chance-constrained DEA model with birandom input and output data. *Journal of the Operational Research Society*, 65(12):1824–1839, 2013.
- [138] Emmanuel Thanassoulis. *Introduction to the theory and application of data envelopment analysis*. Springer, 2001.
- [139] K Tone. An extension of the two phase process in the CCR model. In *Proceedings of the 1999 Fall National Conference of the Operations Research Society of Japan*, 1999.
- [140] Kaoru Tone. A slacks-based measure of efficiency in data envelopment analysis. *European journal of operational research*, 130(3):498–509, 2001.
- [141] Panagiotis Tziogkidis. The Simar and Wilson’s bootstrap DEA approach: a critique. Technical report, Cardiff University, Cardiff Business School, Economics Section, 2012.
- [142] Gregg G Van Ryzin and Eloísa Del Pino. Cómo escuchar, cómo aprender y cómo responder: las encuestas ciudadanas como una herramienta para la reinención del gobierno. *Papeles de evaluación y calidad*, 9:1–26, 2009.
- [143] Terry G Vavra. *Cómo medir la satisfacción del cliente según la ISO 9001: 2000*. FC Editorial, 2003.
- [144] Vasja Vehovar. Field substitutions redefined. *Survey Statistician*, 48:35–37, 2003.
- [145] Miguel Beltrán Villalva. El problema de la calidad en los servicios públicos. In *Evaluación y calidad en las organizaciones públicas*, pages 23–36. Instituto Nacional de Administración Pública, 2000.
- [146] Ying-Ming Wang, Richard Greatbanks, and Jian-Bo Yang. Interval efficiency assessment using data envelopment analysis. *Fuzzy sets and Systems*, 153(3):347–370, 2005.
- [147] Ning-Shen Wang, Rong-Hua Yi, and Wei Wang. Evaluating the performances of decision-making units based on interval efficiencies. *Journal of Computational and Applied Mathematics*, 216(2):328–343, 2008.
- [148] YM Wang, KS Chin, and JB Yang. Measuring the performances of decision-making units using geometric average efficiency. *Journal of the Operational Research Society*, 58(7):929–937, 2007.
- [149] Brady T West, Frauke Kreuter, and Ursula Jaenichen. Interviewer effects in face-to-face surveys: A function of sampling, measurement error, or nonresponse?, 2013.
- [150] Kirk Wolter. *Introduction to variance estimation*. Springer Science & Business Media, 2007.
- [151] Desheng Dash Wu and Chi-Guhn Lee. Stochastic DEA with ordinal data applied to a multi-attribute pricing problem. *European Journal of Operational Research*, 207(3): 1679–1688, 2010.
- [152] Feng Yang, Sheng Ang, Qiong Xia, and Chenchen Yang. Ranking DMUs by using interval DEA cross efficiency matrix with acceptability analysis. *European Journal of Operational Research*, 223(2):483–488, 2012.