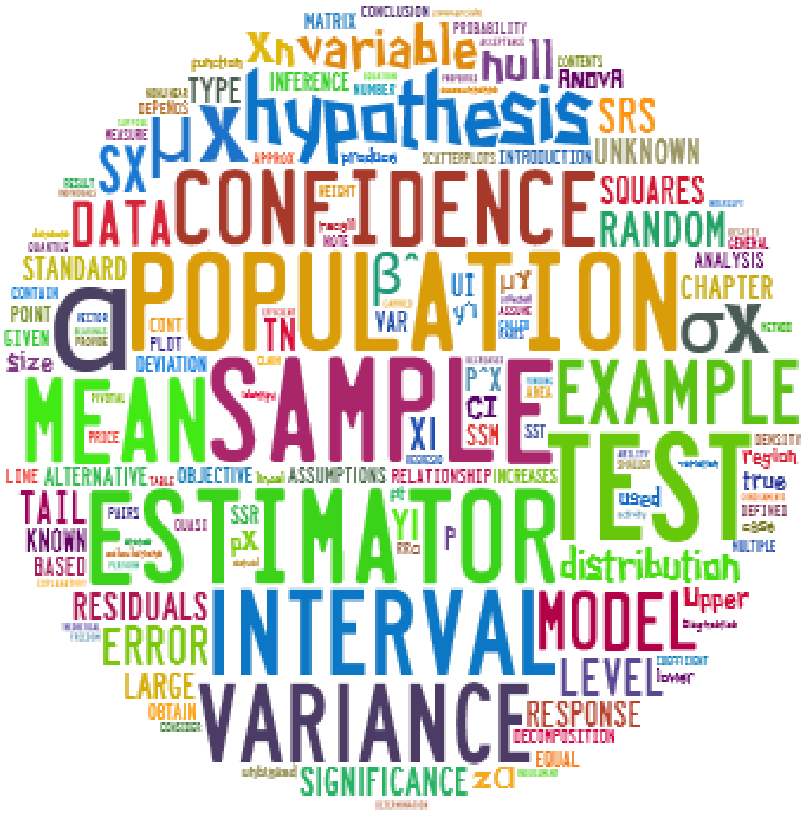


Evaluación de la eficiencia con modelos no paramétricos

PRÁCTICA 1

Análisis de eficiencia y Productividad

Enunciado

Para la realización de la práctica utilizaremos la base de datos "`rice producers.xlsx`" disponible en el CAMPUS VIRTUAL UMH. Además, también se utilizarán los siguientes 3 *toy datasets* (A, B y C, respectivamente):

Store	A	B	C	D	E	F	G	H
x_1: employee	2	3	3	4	5	5	6	8
y_1: sale	1	3	2	3	4	2	3	5

Store	A	B	C	D	E	F	G	H	I
x_1: employee	4	7	8	4	2	5	6	5.5	6
x_2: floor area	3	3	1	2	4	2	4	2.5	2.5
y_1: sale	1	1	1	1	1	1	1	1	1

Store	A	B	C	D	E	F	G
x_1: employee	1	1	1	1	1	1	1
y_1: customers	1	2	3	4	4	5	6
y_2: sales	5	7	4	3	6	5	2

También se ha simulado una base de datos (D) con 50 DMUs generada como se muestra a continuación:

$$x_1 \sim U(a=1, b=10)$$

$$u \sim |N(\mu=1, \sigma=0.4)|$$

$$y_D = \ln(x_1) + 3$$

$$y_1 = y_D - u$$

Preguntas teóricas

En el contexto del Análisis de Eficiencia se desea evaluar la eficiencia de una muestra de n unidades llamadas *Decision-Making Units* [DMUs], donde cada DMU_i, $i=1,...,n$, consume $\boldsymbol{x}_i = (x_{i1},\ldots,x_{ij},\ldots,x_{im}) \in \mathbb{R}^m_+$ *inputs* para la producción de $\boldsymbol{y}_i = (y_{i1},\ldots,y_{ir},\ldots,y_{is}) \in \mathbb{R}^s_+$ *outputs*.

Se asume que las DMUs son generadas a partir de un Proceso Generador de Datos (PGD). En el caso de considerar un único *output*, el PGD es una función desconocida, monótona no decreciente y generalmente cóncava:

$$f(\boldsymbol{x}) : \mathbb{R}^m_+ \to \mathbb{R}_+$$

Este PGD $\left[f(\boldsymbol{x}) \right]$ se conoce como la frontera teórica de producción y mide cuál es el máximo output producible dando cierto nivel de recursos. Por ejemplo, ¿cuál es el máximo número de zapatos (y_1) que se pueden fabricar dado cierto número de trabajadores $(x_1 = 5)$?

La estimación de esta frontera de producción (llamada frontera de Mejores Prácticas en algunos contextos) y la medición de la eficiencia de las unidades de la muestra puede llevarse a cabo bajo dos metodologías bien diferenciadas: enfoques paramétricos y enfoques no paramétricos.

- Un modelo es considerado paramétrico cuando el número de parámetros a estimar es fijo y determinado a priori.
- Un modelo es considerado no paramétrico cuando el número de parámetros a estimar no es fijo y viene determinado por la muestra de datos, los hiperparámetros que definen el modelo, etc.

La principal diferencia entre ambas metodologías es la presunción previa de una forma funcional del PGD. Por ejemplo, bajo un enfoque paramétrico, podemos considerar que $f(\boldsymbol{x})$ es una función de producción de tipo Cobb-Douglas (generalmente, es su forma log-lineal):

$$y_D = \alpha_0 \cdot x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot x_3^{\alpha_3}$$

$$\ln(y_D) = \ln\left(\alpha_0 \cdot x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot x_3^{\alpha_3}\right)$$

$$\ln(y_D) = \ln(\alpha_0) + \alpha_1 \cdot \ln(x_1) + \alpha_2 \cdot \ln(x_2) + \alpha_3 \cdot \ln(x_3),$$

y estimar, entonces, el vector de coeficientes $\boldsymbol{\alpha}=(\alpha_0,\alpha_1,\alpha_2,\alpha_3)$ que hace que la expresión $\ln(\alpha_0)+\alpha_1\cdot\ln(x_1)+\alpha_2\cdot\ln(x_2)+\alpha_3\cdot\ln(x_3)$ se ajuste lo mejor posible a los datos disponibles (al *output* observado). En el enfoque paramétrico, el número de parámetros a estimar suele ser relativamente pequeño en relación al número de *inputs*. Además, este tipo de enfoques suelen permitir interpretaciones claras de los modelos, por ejemplo, en el caso lineal, α_1 representa el cambio marginal del máximo *output* que se puede producir si modificamos el primer *input* dejando constante el resto de los *inputs*. Si estamos ante un modelo log-log (como el de arriba) aumentar $x_1 \to e^1\cdot x_1$ producirá un aumento en el valor esperado de $y \to e^{\alpha_1}\cdot y$. En el enfoque no paramétrico, a priori, no se asume ninguna forma funcional concreta para $f(\boldsymbol{x})$.

Finalmente, en cuanto a los *outputs* observados $(\boldsymbol{y}_i)$, cabe resaltar que son traslaciones (verticales) de este PGD:

$$\boldsymbol{y}_i = f(\boldsymbol{x}_i) - u_i + \varepsilon_i,\ i=1,\dots,n$$

donde u mide la ineficiencia técnica de una DMU y ε mide cierto error aleatorio.

- Los modelos paramétricos generalmente asumen que $u \sim |N(0,\sigma_u)|$ y $\varepsilon \sim N(0,\sigma_\varepsilon)$. Debido a que partimos de la existencia de error aleatorio, se les denomina modelos estocásticos. Estos supuestos permiten estimar los parámetros del modelo mediante métodos como el de máxima verosimilitud y realizar inferencia estadística sobre los mismos: intervalos de confianza y/o contrastes de hipótesis sobre su significatividad.
- Los modelos no paramétricos únicamente asumen que $u_i \geq 0$, $i=1,\dots,n$ y no consideran error aleatorio, es decir, $\varepsilon_i = 0$, $i=1,\dots,n$. Precisamente, debido a que no se considera la existencia de error aleatorio, se dice que este tipo de modelos son deterministas. Por lo tanto, bajo este enfoque, cierta DMU_i será técnicamente eficiente cuando $u_i = 0$.

Además, en el caso del Análisis Envolvente de Datos (modelo no paramétrico en el que nos centraremos), siempre hablaremos de eficiencia técnica "relativa", dado que dependerá exclusivamente de la muestra de datos utilizada.

En la asignatura de Análisis de Eficiencia y Productividad nos centramos en el estudio y aplicación de modelos no paramétricos, desde su enfoque tradicional hasta los recientes avances desde el campo del Aprendizaje Automático.

Ejercicio 1. Dado el siguiente conjunto de datos:

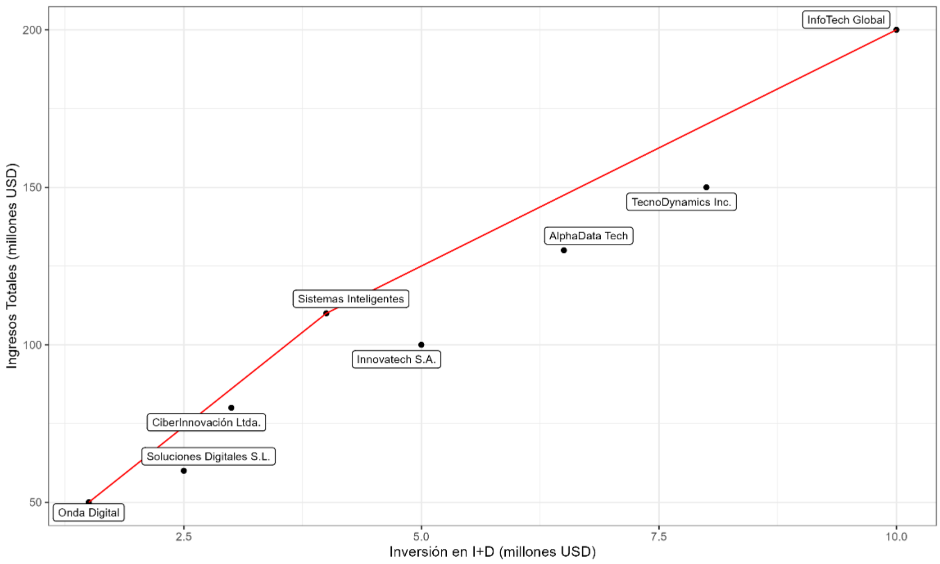
Empresa	Inversión en I+D	Empleados	Ingresos Totales	Patentes
Innovatech S.A.	5.0	105	100	4
Soluciones Digitales S.L.	2.5	125	60	7
TecnoDynamics Inc.	8.0	50	150	5
CiberInnovación Ltda.	3.0	80	80	6
InfoTech Global	10.0	160	200	3
Onda Digital	1.5	150	50	5
Sistemas Inteligentes	4.0	140	110	4
AlphaData Tech	6.5	90	130	6

a) ¿Qué variables son de tipo *input* y cuáles son de tipo *output*? Justifica tu respuesta.

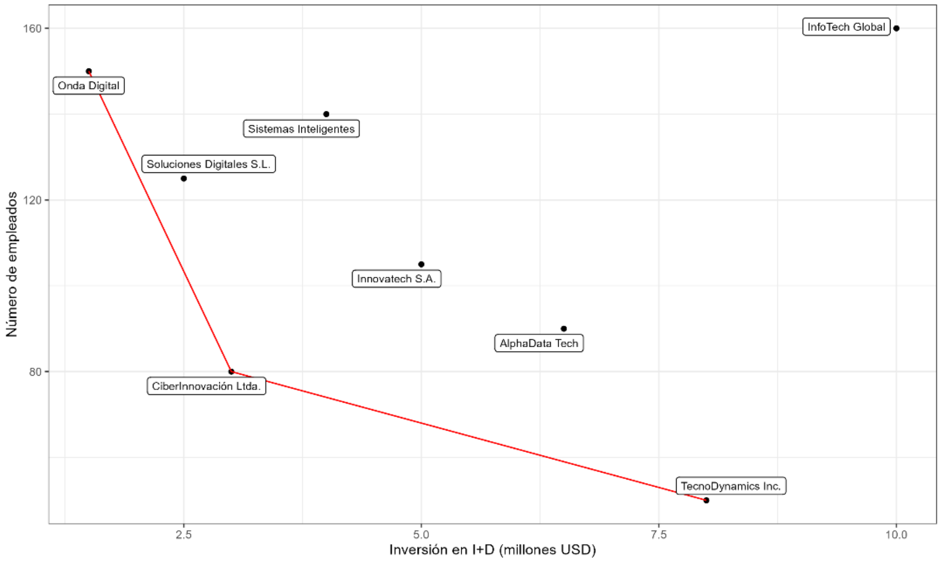
La “inversión en I+D” y el número de “empleados” especializados son variables de tipo *input*, ya que representan recursos utilizados por la empresa en su proceso productivo. Se espera que ambas tengan una relación *monótona no decreciente* con la producción, es decir, que un aumento en estos insumos no reduzca la capacidad productiva de la empresa.

Por otro lado, los “ingresos totales” y el número de “patentes” registradas son variables de tipo *output*, dado que reflejan los resultados obtenidos a partir de los recursos invertidos. Un incremento en los insumos debería, en principio, conducir a mayores ingresos y una mayor producción de innovación, expresada en el número de patentes.

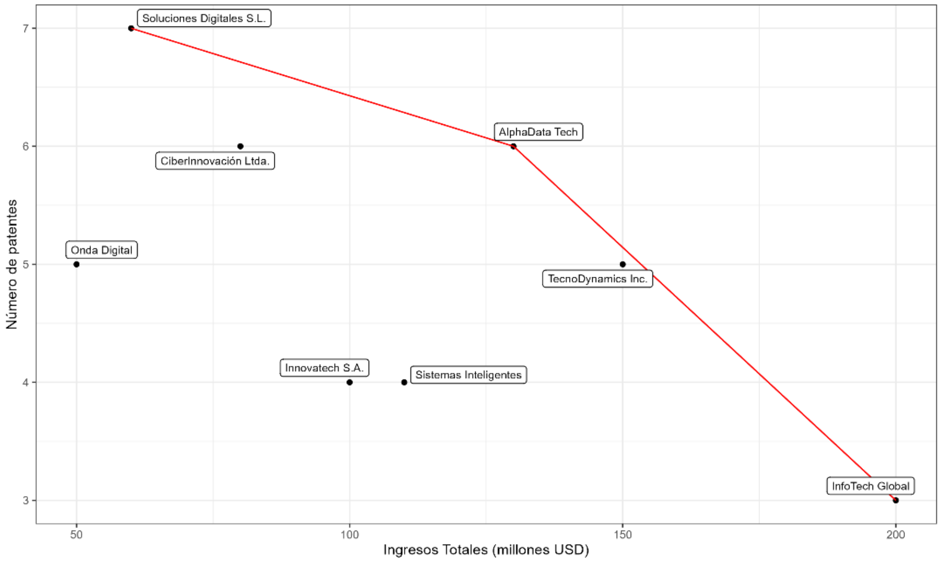
b) Dadas las variables “inversión en I+D” e “ingresos totales”, identifica, mediante una representación gráfica, que unidades de la muestra son técnicamente eficientes.



c) Dadas las variables “inversión en I+D” y “empleados”, identifica, mediante una representación gráfica, que unidades de la muestra son técnicamente eficientes.



d) Dadas las variables “ingresos totales” y “patentes”, identifica, mediante una representación gráfica, que unidades de la muestra son técnicamente eficientes.



e) Proporciona una expresión que permita calcular la eficiencia de las unidades

$$\theta = \frac{\mu_1 \cdot \text{Ingresos Totales} \ + \ \mu_2 \cdot \text{Patentes}}{v_1 \cdot \text{Inversion en ID} \ + \ v_2 \cdot \text{Empleados}}$$

¿De qué dos maneras pueden calcularse los pesos de la expresión anterior?

- Mediante métodos de análisis de eficiencia.
- Mediante análisis de expertos.

Ejercicio 2. Menciona que ventajas y/o inconvenientes presentan cada una de las metodologías frente a la otra.

Ventajas de los métodos paramétricos de análisis de la eficiencia

- Permiten distinguir la parte atribuible al ruido aleatorio de la correspondiente a la ineficiencia técnica.
- Permiten incorporar tests de bondad de ajuste y/o realizar inferencia sobre los parámetros estimados.
- Ofrecen una clara interpretabilidad de los parámetros del modelo.
- En caso de especificar correctamente el modelo, se adaptan bien incluso a tamaños muestrales reducidos.

Inconvenientes de los métodos paramétricos de análisis de la eficiencia

- La especificación de una forma funcional para el PGD puede resultar demasiado restrictiva; una mala elección puede sesgar negativamente los resultados.
- Una distribución inadecuada para el término de ineficiencia puede conducir a resultados erróneos.
- Presentan dificultades al trabajar con escenarios multi-*output*.
- Son computacionalmente más costosos, pues se resuelven mediante técnicas como máxima verosimilitud o emparejamiento de momentos.

Ventajas de los métodos no paramétricos de análisis de la eficiencia

- No requieren asignar pesos previos a los inputs y outputs en escenarios multi-*input* y multi-*output*, lo que permite tratar el multi-*output* como una extensión natural del mono-*output*.
- No es necesario especificar una forma funcional para el PGD, lo que favorece la flexibilidad de las técnicas utilizadas.
- Impone propiedades axiomáticas (monotonía, libre disponibilidad, etc.) desde un enfoque no paramétrico, brindando rigor, claridad (base sólida para el razonamiento), consistencia, demostraciones, teoremas y generalizaciones.
- Son computacionalmente eficientes, principalmente debido a que se resuelven mediante modelos de optimización lineal.

Desventajas de los métodos no paramétricos de análisis de la eficiencia

- Alta sensibilidad a los *outliers*.
- Tienden a sobreajustar la muestra de datos analizada; por ello, solo miden eficiencia relativa y los resultados dependen estrictamente de la muestra considerada.
- Sensibles a bases de datos "rectangulares" (pocas DMUs y muchas variables), lo que puede derivar en el conocido problema de la "maldición de la dimensionalidad", causando valoraciones excesivamente optimistas.
- No incorporan medidas de error aleatorio (por ejemplo, condiciones meteorológicas adversas o huelgas), de manera que cualquier desviación de la frontera de producción se asume como ineficiencia técnica.
- No permiten incorporar tests de bondad de ajuste ni inferencia sobre los parámetros estimados.

Ejercicio 3. Consideramos una muestra de n DMUs, para las cuales se desea evaluar la eficiencia técnica. Cada DMU consume $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ij}, \ldots, x_{im}) \in \mathbb{R}_+^m$ *inputs* para producir $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{ir}, \ldots, y_{is}) \in \mathbb{R}_+^s$ *outputs*. Para medir la eficiencia (relativa) de cada DMU, es necesario definir un conjunto tecnológico común T compartido por todas las DMUs de la muestra. Desde una perspectiva más amplia, esta tecnología puede expresarse como:

$$T = \left\{ \left(\boldsymbol{x}, \boldsymbol{y} \right) \in \mathbb{R}_+^{m+s} : \boldsymbol{x} \text{ puede producir } \boldsymbol{y} \right\}.$$

Existen tres tipos de tecnologías de producción:

Nombre	Dataset	Convexidad	Libre disponibilidad	Rendimientos
CCR	*Toy dataset A*	✓	✓	Constantes
BCC	*Toy dataset B*	✓	✓	Variables
FDH	*Toy dataset C*	×	✓	Variables

A continuación, se define formalmente cada una de estas tecnologías. A partir de su definición, proporciona un punto que pertenezca a cada una de las tecnologías para cada una de las bases de datos que se proporcionan en la tabla anterior.

Tecnología DEA (CCR)

La tecnología CCR (Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444.) es una tecnología convexa que satisface el principio de libre disponibilidad bajo rendimientos constantes a escala:

$$\hat{T}_{CRS}=\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathbb{R}_{+}^{m+s}: x_j\geq\sum_{i=1}^{n}\lambda_i x_{ij},\ j=1,\ldots,m,\ y_r\leq\sum_{i=1}^{n}\lambda_i y_{ir},\ r=1,\ldots,s,\ \lambda_i\geq 0,\ i=1,\ldots,n\right\}.$$

$$x_1\geq\begin{pmatrix}\lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 & \lambda_6 & \lambda_7 & \lambda_8\end{pmatrix}\cdot\begin{pmatrix}2\\3\\3\\4\\5\\5\\6\\8\end{pmatrix}\rightarrow x_1\geq 2\lambda_1+3\lambda_2+3\lambda_3+4\lambda_4+5\lambda_5+5\lambda_6+6\lambda_7+8\lambda_8$$

$$y_1\leq\begin{pmatrix}\lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 & \lambda_6 & \lambda_7 & \lambda_8\end{pmatrix}\cdot\begin{pmatrix}1\\3\\2\\3\\4\\2\\3\\5\end{pmatrix}\rightarrow y_1\leq\lambda_1+3\lambda_2+2\lambda_3+3\lambda_4+4\lambda_5+2\lambda_6+3\lambda_7+5\lambda_8$$

$$\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5,\lambda_6,\lambda_7,\lambda_8\geq 0$$

El vector $\boldsymbol{\lambda}=(0.5,1,0,0,0,0,0,0)$ proporciona una posible solución al problema, con el cual se obtiene un conjunto de DMUs virtuales que satisfacen:

$$x_1\geq 2\cdot 0.5+3\cdot 1=4$$

$$y_1\leq 1\cdot 0.5+3\cdot 1=3.5$$

Por lo tanto, una posible DMU virtual podría ser $DMU=(5,3)$.

Anexo

Modelo radial input en formato de ratio con tecnología CCR:

$$\max_{\boldsymbol{v},\boldsymbol{\mu}} \frac{\sum_{r=1}^{s} \mu_r \cdot y_{0r}}{\sum_{j=1}^{m} v_j \cdot x_{0j}}$$

subject to

$$\begin{array}{rcll} \dfrac{\sum_{r=1}^{s} \mu_r \cdot y_{ir}}{\sum_{j=1}^{m} v_j \cdot x_{ij}} & \leq & 1, & i = 1,\dots,n \\ v_j & \geq & 0, & j = 1,\dots,m \\ \mu_r & \geq & 0, & r = 1,\dots,s \end{array}$$

Modelo radial output en formato de ratio con tecnología CCR:

$$\min_{\boldsymbol{v},\boldsymbol{\mu}} \frac{\sum_{j=1}^{m} v_j \cdot x_{0j}}{\sum_{r=1}^{s} \mu_r \cdot y_{0r}}$$

subject to

$$\begin{array}{rcll} \dfrac{\sum_{j=1}^{m} v_j \cdot x_{ij}}{\sum_{r=1}^{s} \mu_r \cdot y_{ir}} & \geq & 1, & i = 1,\dots,n \\ v_j & \geq & 0, & j = 1,\dots,m \\ \mu_r & \geq & 0, & r = 1,\dots,s \end{array}$$

Modelo radial input en formato de multiplicadores con tecnología DEA (CCR / BCC):

$$
\begin{aligned}
&\max_{\boldsymbol{v},\boldsymbol{\mu},w} \sum_{r=1}^{s} \mu_r \cdot y_{0r} + w \\
&\text{subject to} \\
&\qquad\qquad \begin{array}{rcll}
\displaystyle\sum_{j=1}^{m} v_j \cdot x_{0j} & = & 1 & \\
\displaystyle\sum_{r=1}^{s} \mu_r \cdot y_{ir} + w & \le & \displaystyle\sum_{j=1}^{m} v_j \cdot x_{ij}, & i = 1,\dots,n \\
v_j & \ge & 0, & j = 1,\dots,m \\
\mu_r & \ge & 0, & r = 1,\dots,s
\end{array}
\end{aligned}
$$

- Tecnología CCR: $w = 0$.
- Tecnología BCC: w libre.

Modelo radial output en formato de multiplicadores con tecnología DEA (CCR / BCC):

$$
\begin{aligned}
&\min_{\boldsymbol{v},\boldsymbol{\mu},w} \sum_{j=1}^{m} v_j \cdot x_{0j} + w \\
&\text{subject to} \\
&\qquad\qquad \begin{array}{rcll}
\displaystyle\sum_{r=1}^{s} \mu_r \cdot y_{0r} & = & 1 & \\
\displaystyle\sum_{j=1}^{m} v_j \cdot x_{ij} + w & \ge & \displaystyle\sum_{r=1}^{s} \mu_r \cdot y_{ir}, & i = 1,\dots,n \\
v_j & \ge & 0, & j = 1,\dots,m \\
\mu_r & \ge & 0, & r = 1,\dots,s
\end{array}
\end{aligned}
$$

- Tecnología CCR: $w = 0$.
- Tecnología BCC: w libre.

Modelo radial input en formato envolvente con tecnología CCR:

$$\min_{\theta,\lambda} \ \theta$$

subject to

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i \cdot x_{ij} &\leq \theta x_{0j} \qquad j=1,\ldots,m \\
\sum_{i=1}^{n} \lambda_i \cdot y_{ir} &\geq y_{0r}, \qquad r=1,\ldots,s \\
\lambda_i &\geq 0, \qquad i=1,\ldots,n
\end{aligned}$$

Tecnología BCC:

- Añadimos la restricción: $\sum_{i=1}^{n} \lambda_i = 1$

Tecnología FDH:

- Añadimos la restricción: $\sum_{i=1}^{n} \lambda_i = 1$
- Añadimos la restricción: $\lambda_i \in \{0,1\}$, $i=1,\ldots,n$.

Modelo radial output en formato envolvente con tecnología CCR:

$$\max_{\phi,\lambda} \ \phi$$

subject to

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i \cdot x_{ij} &\leq x_{0j} \qquad j=1,\ldots,m \\
\sum_{i=1}^{n} \lambda_i \cdot y_{ir} &\geq \phi y_{0r}, \qquad r=1,\ldots,s \\
\lambda_i &\geq 0, \qquad i=1,\ldots,n
\end{aligned}$$

Tecnología BCC:

- Añadimos la restricción: $\sum_{i=1}^{n} \lambda_i = 1$

Tecnología FDH:

- Añadimos la restricción: $\sum_{i=1}^{n} \lambda_i = 1$
- Añadimos la restricción: $\lambda_i \in \{0,1\}$, $i=1,\ldots,n$.

Modelo función distancia direccional en formato envolvente con tecnología CCR:

$$
\begin{aligned}
&\max_{\beta,\lambda} \ \beta \\
&\text{subject to} \\
&\qquad \sum_{i=1}^{n} \lambda_i \cdot x_{ij} \quad \leq \quad x_{0j} - \beta G_{x_j} \qquad j = 1,\dots,m \\
&\qquad \sum_{i=1}^{n} \lambda_i \cdot y_{ir} \quad \geq \quad y_{0r} + \beta G_{y_r}, \quad r = 1,\dots,s \\
&\qquad\qquad \lambda_i \quad \geq \quad 0, \qquad\qquad i = 1,\dots,n
\end{aligned}
$$

`direction = “mean”:` $G_{x_j} = \overline{x}_j$ y $G_{y_r} = \overline{y}_r$.

`direction = “briec”:` $G_{x_j} = x_{0j}$ y $G_{y_r} = y_{0r}$.

Tecnología BCC:

- Añadimos la restricción: $\sum_{i=1}^{n} \lambda_i = 1$

Tecnología FDH:

- Añadimos la restricción: $\sum_{i=1}^{n} \lambda_i = 1$
- Añadimos la restricción: $\lambda_i \in \{0,1\}$, $i = 1,\dots,n$.