

Pareamiento

Junio de 2017

Oficina de Planificación Estratégica y Efectividad en el Desarrollo



I. Intuición

II. Definición

- Supuestos
- Metodología

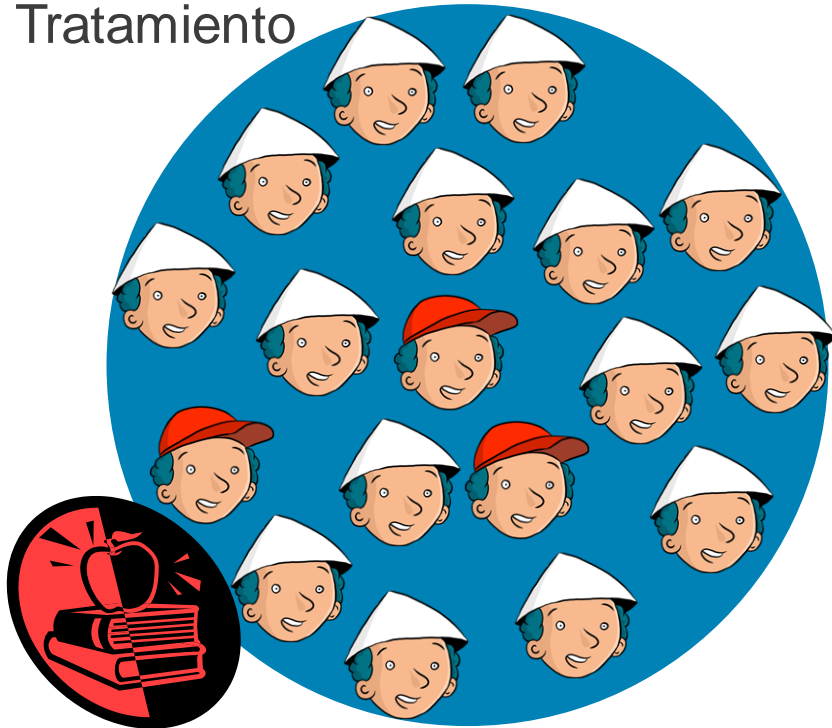
III. Ejercicio

Extra: Selección de técnica

I. INTUICIÓN

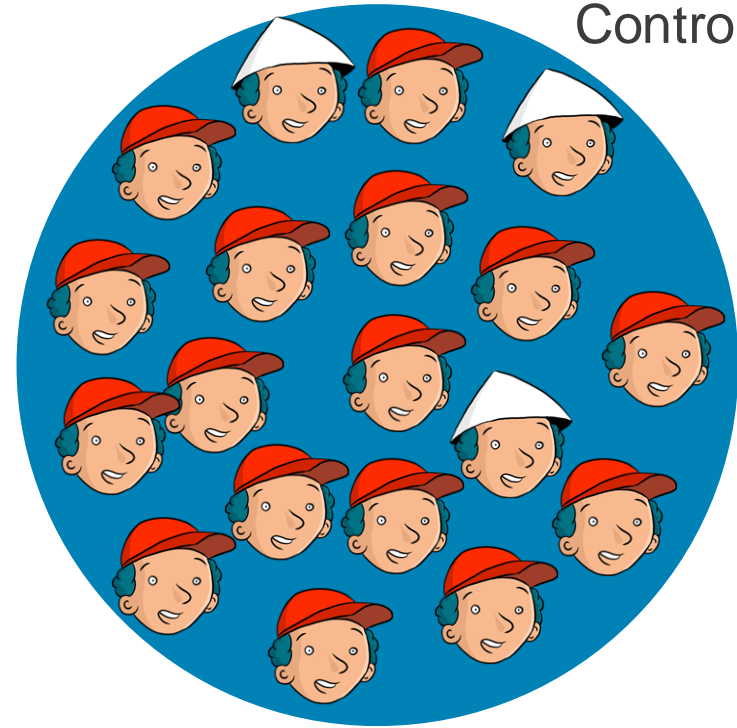
Problema de evaluación

Tratamiento



$$E[Y_i|T] = 630$$








Control





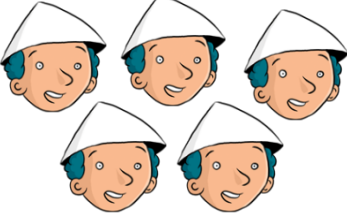

$$\hat{E}[Y_i|C] = 600$$

$$\text{IMPACTO} = 630 - 600 = 30$$

Idea General

	Tratamiento	Comparación	Impacto Efecto tratamiento sobre los tratados (ETT)
Edad 7-9 15%	  610	  600	
Edad 10-12 80%	 635	 600	
Edad 13-15 5%	 640	?	

Idea General

	Tratamiento	Comparación	Impacto ETT
Edad 7-9 20%	 610	 600	10
Edad 10-12 80%	 635	 600	35
			$.2*(10)+.8*(35)$ =30

II. DEFINICIÓN

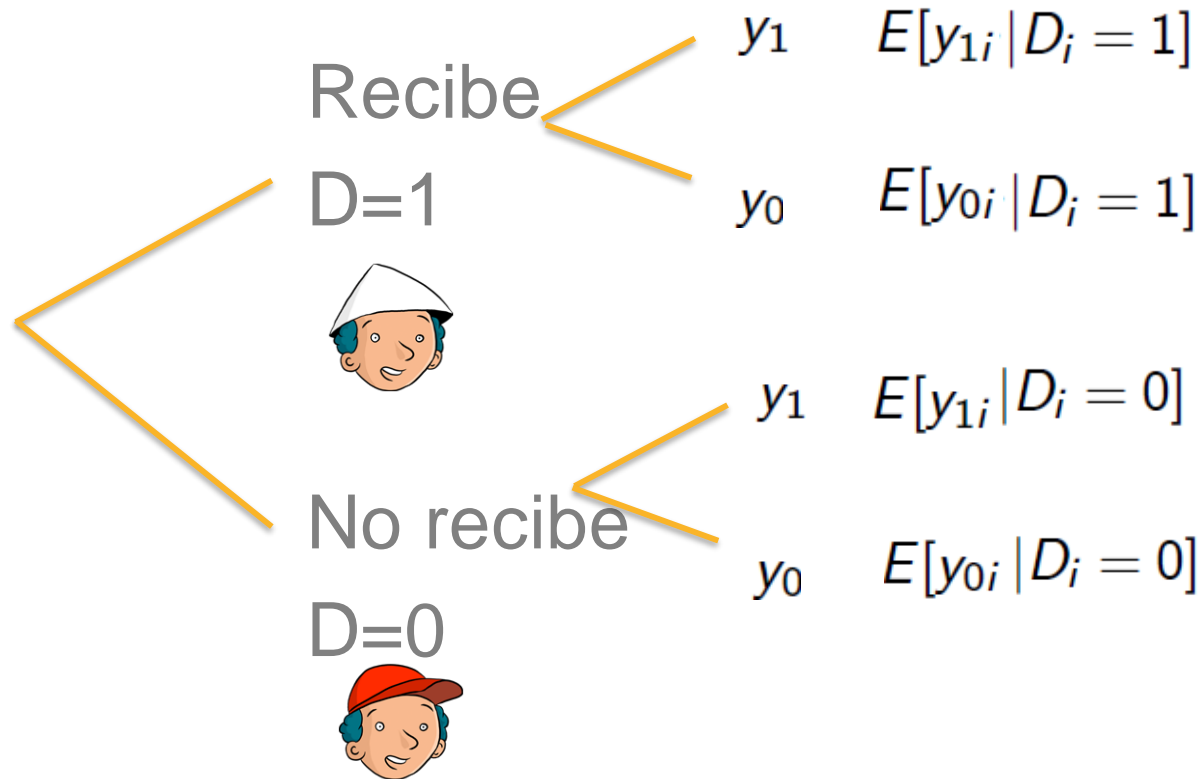
Supuestos

Estimador de pareo

$$E[y_{1i} - y_{0i} | D_i = 1] = \sum_{\mathbf{x}} (E[y_{1i} | D_i = 1, \mathbf{x}] - E[y_{0i} | D_i = 0, \mathbf{x}]) Pr[\mathbf{x}_i = \mathbf{x} | D_i = 1]$$

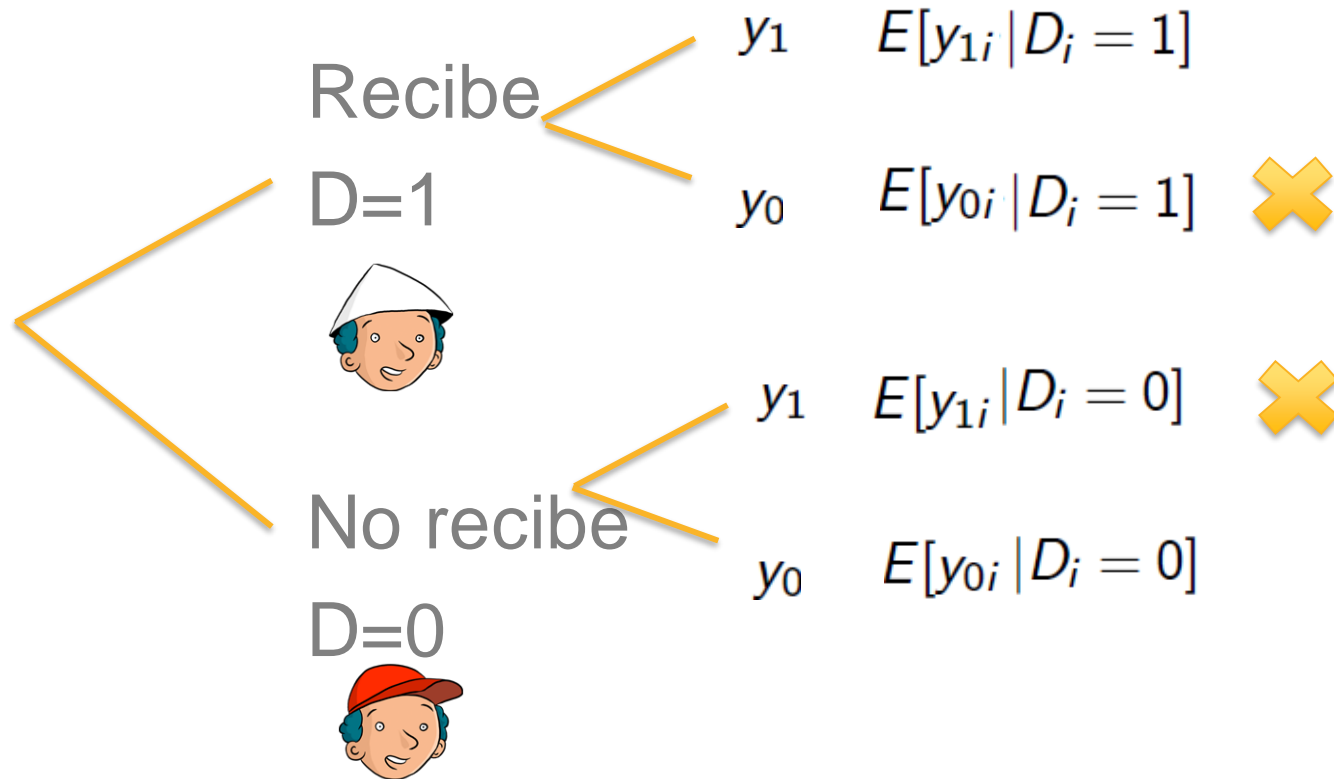
1. Independencia condicional en la media:
 $E[y_0 | D = 1, \mathbf{x}] = E[y_0 | D = 0, \mathbf{x}] = E[y_0 | \mathbf{x}]$
2. Soporte común: $0 < Pr[D = 1 | \mathbf{x}] < 1$
3. SUTVA

ATE vs ATET



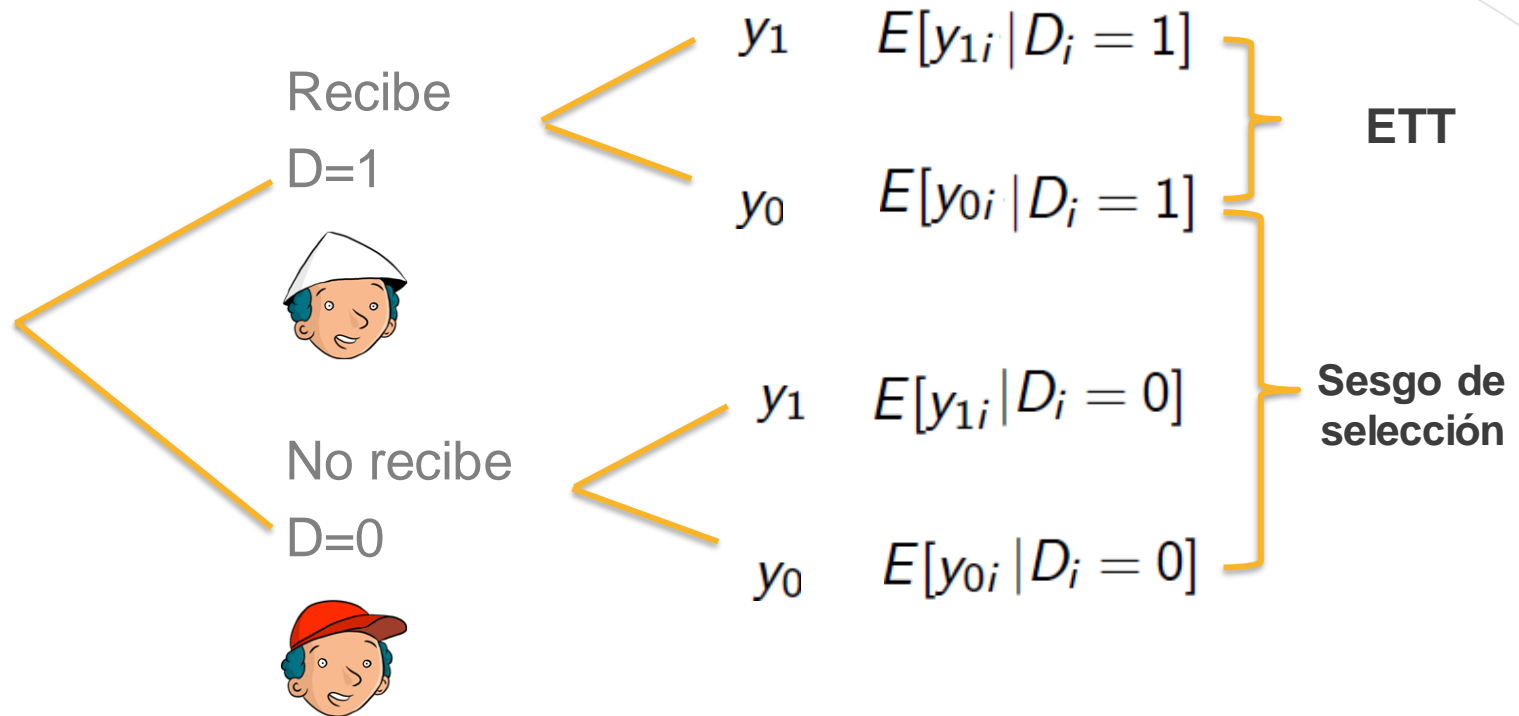
$$ATE = E[y_i | D_i = 1] - E[y_i | D_i = 0]$$

ATE vs ATET



$$ATE = E[y_i | D_i = 1] - E[y_i | D_i = 0]$$



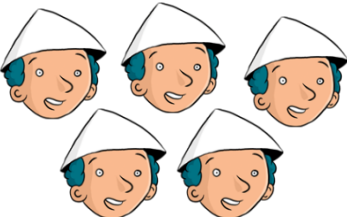

ATE vs ATET



$$E[y_i | D_i = 1] - E[y_i | D_i = 0] = E[y_{1i} - y_{0i} | D_i = 1] + E[y_{0i} | D_i = 1] - E[y_{0i} | D_i = 0]$$

$$ATE = ETT + \text{Sesgo de selección}$$

Intuición

	Tratamiento	Comparación	Impacto ATOT
Edad 7-9 20%	 610	 600	10
Edad 10-12 80%	 635	 600	35
			$.2*(10) + .8*(35)$ =30

Problema de dimensionalidad y Puntaje de propensión

(Rosenbaum and Rubin, 1983)

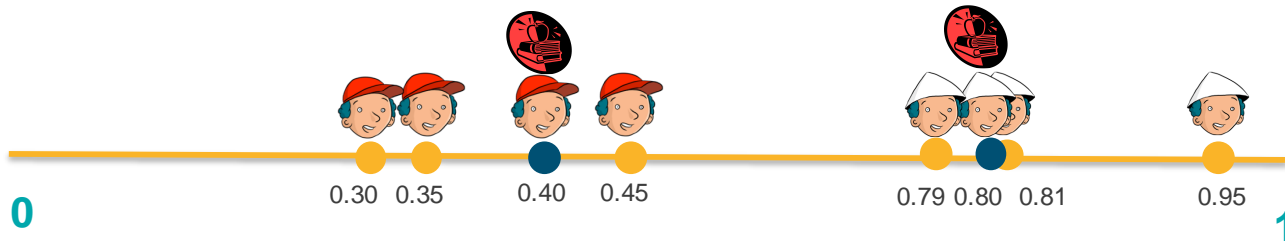
Género

Educación de los padres

Edad

Ingreso del Hogar

Puntaje de propensión = $P[\text{participar}]$



Definición

$$p(\mathbf{x}) = \Pr[D = 1 | \mathbf{X} = \mathbf{x}]$$

Rosenbaum and Rubin (1983)

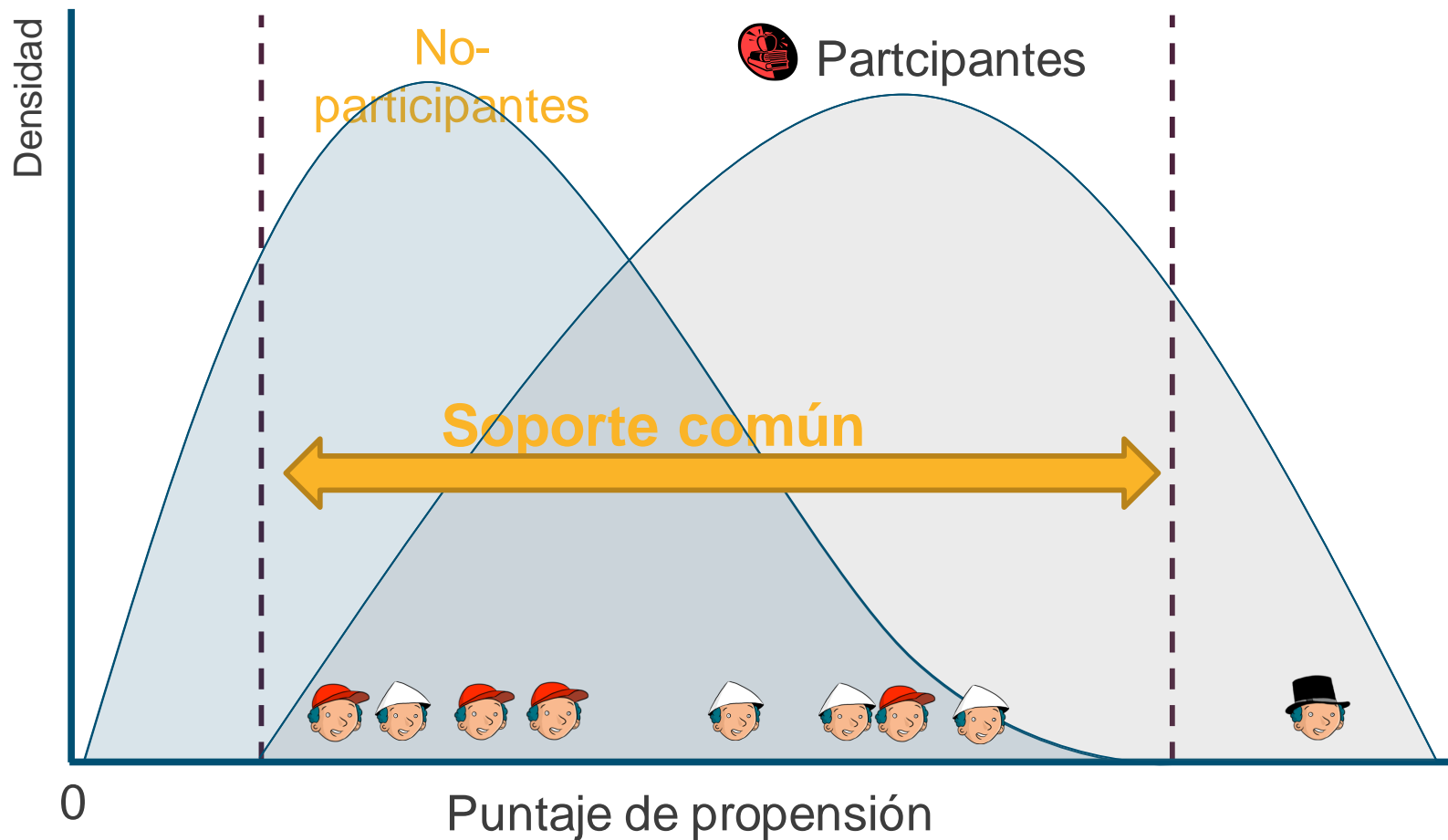
Teorema del puntaje de propensión

$$\{Y_{0i}, Y_{1i}\} \perp D_i \parallel X_i \implies \{Y_{0i}, Y_{1i}\} \perp D_i \parallel p(X_i)$$

Estimador de pareamiento

$$E[y_{1i} - y_{0i} | D_i = 1] = \sum_{\mathbf{x}} (E[y_{1i} | D_i = 1, \mathbf{x}] - E[y_{0i} | D_i = 0, \mathbf{x}]) \Pr[\mathbf{x}_i = \mathbf{x} | D_i = 1]$$

Densidad en puntajes de propensión



Definición

$$p(\mathbf{x}) = \Pr[D = 1 | \mathbf{X} = \mathbf{x}]$$

Teorema de puntaje de propensión

$$\{Y_{0i}, Y_{1i}\} \perp D_i \parallel X_i \implies \{Y_{0i}, Y_{1i}\} \perp D_i \parallel p(X_i)$$

Estimador de pareamiento

$$E[y_{1i} - y_{0i} | D_i = 1] = \sum_{\mathbf{x}} (E[y_{1i} | D_i = 1, \mathbf{x}] - E[y_{0i} | D_i = 0, \mathbf{x}]) \Pr[\mathbf{x}_i = \mathbf{x} | D_i = 1]$$

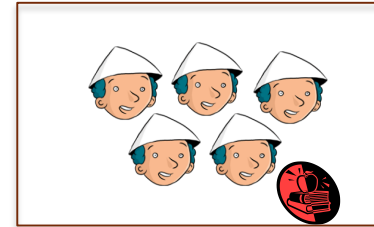
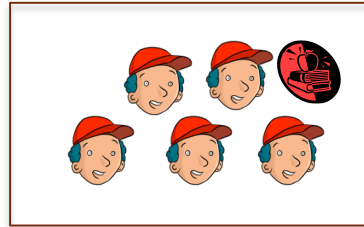
$$\frac{1}{N_T} \sum_{i \in D=1} [y_{1i} - \sum_j w(i, j) y_{0j}]$$

III. EJERCICIO

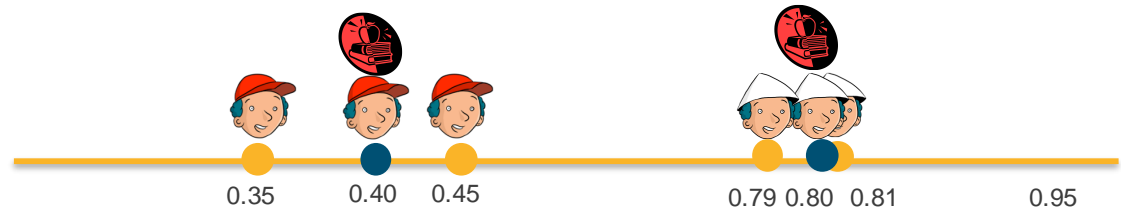
EXTRA. SELECCIÓN DE TÉCNICAS

Tipos de pareamiento

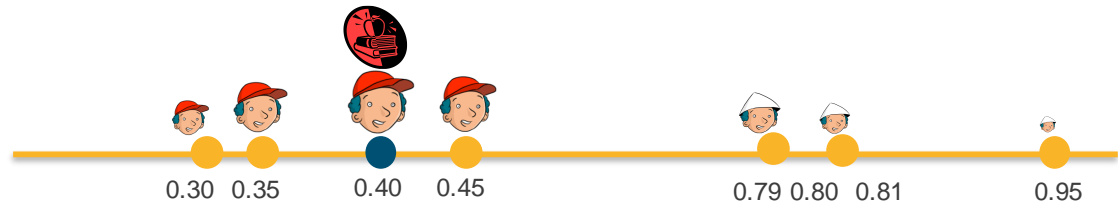
Por grupos



Por vecindario



Ponderado



Grupos

Exacto

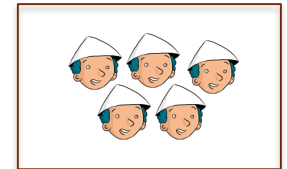
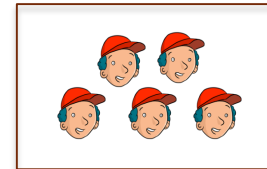
Burdo

Completo

Estratificado

Subclasificación

$$w_i = \begin{cases} 1 & X_i = X_j \\ 0 & \text{Si no} \end{cases}$$



Vecinos

Vecino más
cercano

Radio

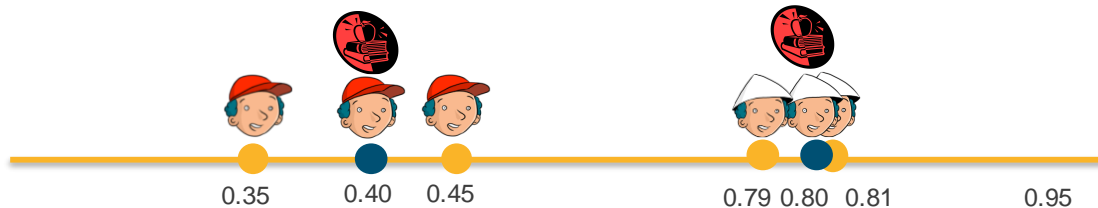
1:1

1:k

$$w_i = \begin{cases} 1 & j \in A_i \\ 0 & \text{Si no} \end{cases}$$

$$A_i(\mathbf{x}) = \{j \mid \min_j \|\mathbf{x}_i - \mathbf{x}_j\|\}$$

$$A_i(p(\mathbf{x})) = \{p_j \mid \min_j \|p_i - p_j\|\}$$



Vecinos

Vecino más
cercano



1:1



1:k

radio

$$x_1 < x_2 < \dots < x_0 < \dots < x_N,$$

$$w_{i0,k} = \frac{1}{k} \times \mathbf{1} \left(|i - 0| < \frac{k-1}{2} \right)$$

Vecinos



Vecinos

Vecino más
cercano

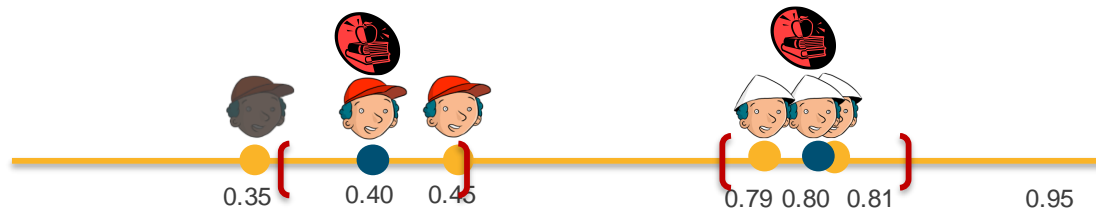
Radio

1:1

1:k

$$A_i(p(\mathbf{x})) = \{p_j \mid \|p_i - p_j\| < r\}$$

w a promediar en el conjunto A



Vecinos

Vecino más
cercano



1:1



1:k

radio

$$A_i(p(\mathbf{x})) = \{p_j \mid \|p_i - p_j\| < r\}$$

w a promediar en el conjunto A

Optimo

Métricas

Mahalanobis

$$D_{ij} = (X_i - X_j)' \Sigma^{-1} (X_i - X_j)$$

$\Sigma = I$

Euclidiano

Σ es diagonal

Estandarizado

Completo

Puntaje de propensión

$$\|p_i - p_j\|$$

Probit

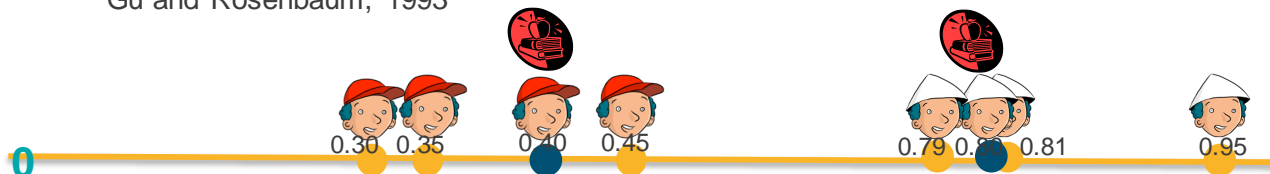
Logit

Lineal

Puntaje de propensión lineal

$$\|\text{logit } p_i - \text{logit } p_j\|$$

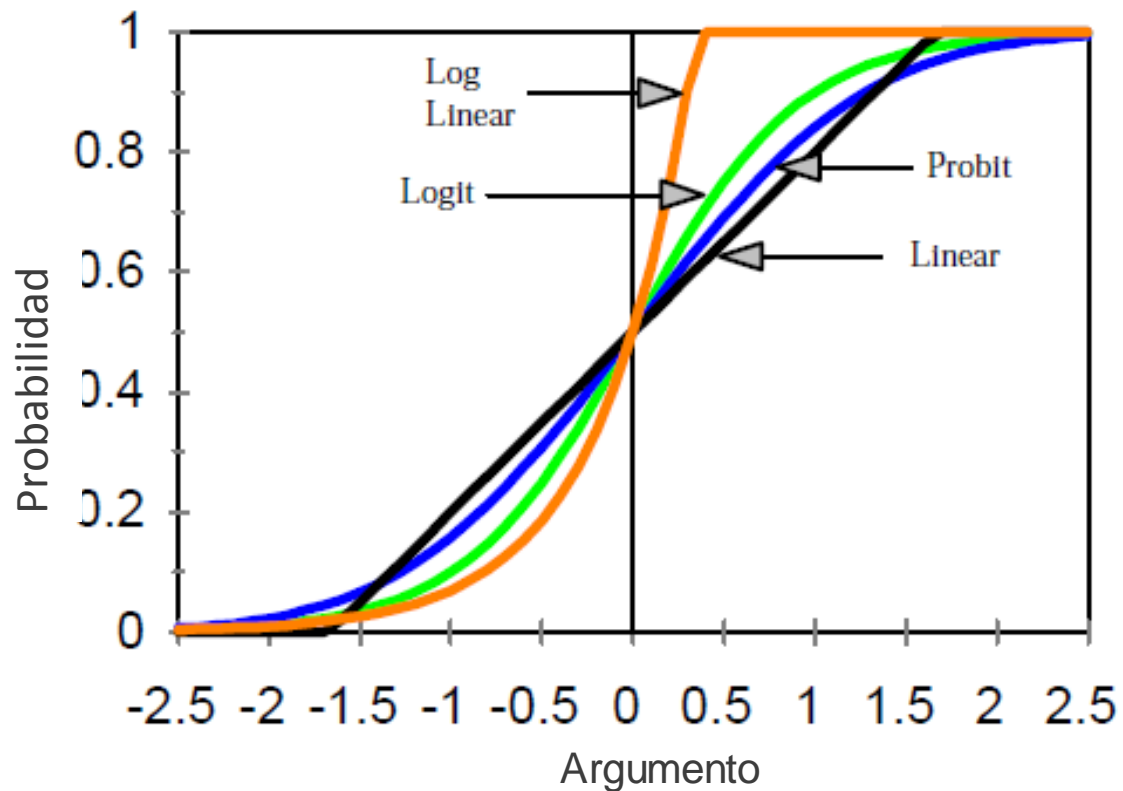
Gu and Rosenbaum, 1993



1

Cómo elegir psm

Modelos de respuesta binomial



Ponderaciones

Kernel

Normal, biweight, epanechnikov, uniforme, tricubo

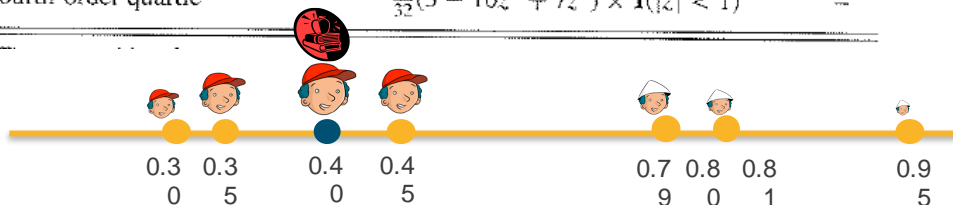
$$W_j(P(X_i)) = \frac{K\left(\frac{P(X_i) - P(X_k)}{h_n}\right)}{\sum_{\substack{k=1 \\ \{D_k=0\}}}^{n_0} K\left(\frac{P(X_i) - P(X_k)}{h_n}\right)}$$

Kernel	Kernel Function $K(z)$	δ
Uniform (or box or rectangular)	$\frac{1}{2} \times \mathbf{1}(z < 1)$	1.3510
Triangular (or triangle)	$(1 - z) \times \mathbf{1}(z < 1)$	—
Epanechnikov (or quadratic)	$\frac{3}{4}(1 - z^2) \times \mathbf{1}(z < 1)$	1.7188
Quartic (or biweight)	$\frac{15}{16}(1 - z^2)^2 \times \mathbf{1}(z < 1)$	2.0362
Triweight	$\frac{35}{32}(1 - z^2)^3 \times \mathbf{1}(z < 1)$	2.3122
Tricubic	$\frac{70}{81}(1 - z ^3)^3 \times \mathbf{1}(z < 1)$	—
Gaussian (or normal)	$(2\pi)^{-1/2} \exp(-z^2/2)$	0.7764
Fourth-order Gaussian	$\frac{1}{2}(3 - z^2)(2\pi)^{-1/2} \exp(-z^2/2)$	—
Fourth-order quartic	$\frac{15}{32}(3 - 10z^2 + 7z^4) \times \mathbf{1}(z < 1)$	—

Funciones de kernel:

- i) Simétricas
alrededor de 0
- ii) Suman uno, el valor
esperado es cero
- iii) Están limitadas y
tienen momentos
finitos

Ver Trivedi (2005) para
más detalles.



Qué método es mejor?

- Depende de tus datos
 - \uparrow # vecinos \rightarrow \uparrow sesgo \downarrow eficiencia
- Y de qué dimensiones puedes observar
- **Usa el método que te de el mejor balance**(Stuart, 2010)
 - Recuerda estandarizar las diferencias!
 - Sobre el mayor número de covariables
 - A unas cuantas variables de prognosis
 - El menor número estandarizado 0.25
 - El valor-p menor de Kolmogorov-Smirnov

Y los errores?

- Default (STATA):
 - (i) observaciones independientes,
 - (ii) ponderaciones fijas,
 - (iii) homostedasticidad dentro de los grupos de T y C,
 - (iv) La varianza de Y no depende del puntaje de propensión
- Pareameinto de VMC: Abadie e Imbens (2006) sugieren una corrección (iv). [Los intervalos de bootstrap divergen aún en el caso de un control y pareamiento exacto]
- Otros: Bootstrap.

Algunos consejos prácticos

- Muestra que tus supuestos posiblemente se mantienen
 - Independencia condicional
 - Balance
 - Análisis de sensibilidad
 - Otros métodos
 - Placebo
 - Soporte común
- Haz pruebas con diferentes técnicas de pareamiento

Resumen

- Pareamiento:

Crea control : promedio de individuos sin tratamiento pero con las mismas características observables

- Supuesto 1: Independencia condicional en la media
- Supuesto 2: Soporte común



Referencias

ABADIE, A. and IMBENS, G.W. (2006) “Large Sample Properties of Matching Estimators for Average Treatment Effects,” *Econometrica* , vol. 74, no. 1, 235-267.

ABADIE, A. and IMBENS, G.W. (2008) “On the Failure of the Bootstrap for Matching Estimators,” *Econometrica* , vol. 76, no. 6, 1537-1558.

Cameron A. C. and Trivedi P. K. (2006) Microeconometrics. Methods and Applications. Cambridge University Press

Diamond, A. and Sekhon, J. S. (2013) “Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies” *The Review of Economic and Statistics*. Vol. 95, No. 3, Pages 932-945

Referencias

- Dehejia, Rajeev. 2005. “Practical Propensity Score Matching: A Reply to Smith and Todd.” *Journal of Econometrics* 125 (1–2): 355–364.
- Dehejia, Rajeev and Sadek Wahba. 1999. “Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs.” *Journal of the American Statistical Association* 94 (448): 1053–1062.
- Dehejia, Rajeev H. and Sadek Wahba. 2002. “Propensity Score Matching Methods for Nonexperimental Causal Studies.” *Review of Economics and Statistics* 84 (1): 151–161.
- Diprete, Thomas A. and Henriette Engelhardt. 2004. “Estimating Causal Effects With Matching Methods in the Presence and Absence of Bias Cancellation.” *Sociological Methods & Research* 32 (4): 501–528.

References

- Guo S. and Fraser M. W. (2010) Propensity Score Analysis: Statistical Methods and Applications. SAGE Publications.
- Heckman, James J., Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. “Characterizing Selection Bias Using Experimental Data.” *Econometrica* 66 (5): 1017–1098.
- Heinrich, C., Maffioli, A., and Vazquez, G. (2010) “A Primer for Applying Propensity-Score Matching”, IDB-TN-161. Inter-American Development Bank.
- LaLonde, Robert. 1986. “Evaluating the Econometric Evaluations of Training Programs with Experimental Data.” *American Economic Review* 76 (September): 604–20.
- Mebane, Walter R. Jr. and Jasjeet S. Sekhon. 1998. “GENetic Optimization Using Derivatives (GENOUD).” Software Package. <http://sekhon.berkeley.edu/rgenoud/>.
- Stuart, E. (2010) “Matching Methods for Causal Inference: A Review and a Look Forward”, *Stat Sci.* February 1; 25(1): 1-21.



 rosangelab@iadb.org

