

# Estimación Doble Afectación

## ➤ Resumen

El presente documento describe la metodología utilizada para el cálculo de la población con doble afectación en el país para los años 2016, 2017 y 2018.

## ➤ Palabras Clave

Desastres, desplazamiento, doble afectación.

## ➤ Introducción

Para efectos del presente documento, la doble afectación ocurre cuando una persona o comunidad, sufre simultáneamente del efecto del desplazamiento, y de un desastre natural. Este ejercicio busca identificar la población colombiana con doble afectación a través de una metodología que permita comprender el comportamiento de estos dos eventos.

## ➤ Antecedentes

En años anteriores, se han utilizado diversas metodologías para analizar la doble afectación. El año inmediatamente anterior, se analizó la información de desastres de la Unidad Nacional de Gestión del Riesgo en Desastres (UNGRD) y de desplazamiento por expulsión de la Unidad para la Atención y Reparación Integral a las Víctimas (UARIV), y se tuvieron en cuenta los siguientes criterios:

- En los municipios en los que no se presentaron desastres o no se presentaron desplazamientos por conflicto, la “doble afectación” fue igual a cero.
- En el resto de los casos, la “doble afectación” fue igual al número total de personas afectadas por desastres.

De ello, se identificaron un total de 1.068.000 personas con doble afectación. Sin embargo, con esta metodología se identificaron algunas dificultades: 1) no tiene en cuenta el número de personas desplazadas por expulsión, 2) no hace relación entre los dos eventos, 3) asume que todos los afectados por desastres fueron también desplazados (100%), 4) no tiene en cuenta la población en donde se presentaron los eventos (población en riesgo) y, 5) la población que se calculó con doble afectación fue tres veces la población identificada como desplazada por expulsión. Estos antecedentes, plantearon la necesidad de buscar otra forma de estimar la doble afectación y por ello se planteó esta propuesta metodológica.

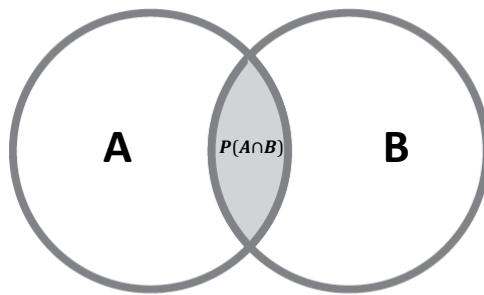
### ➤ Conceptos

A continuación, se describen algunos conceptos utilizados para la metodología de doble afectación.

#### Teoría de conjuntos

Un evento es simplemente un conjunto, así que las relaciones y resultados de la teoría elemental de conjuntos pueden ser utilizados para estudiar eventos. Para estimar la población con doble afectación, se utilizó el concepto de intersección de la teoría de conjuntos.

La intersección de dos eventos A y B, denotada por  $A \cap B$  y leída “A y B”, es el evento que consiste en todos los resultados que están tanto en A como en B. La región sombreada es  $(A \cap B)$ .



#### Probabilidad

El objetivo de la probabilidad es asignar a cada evento A un número  $P(A)$ , llamado la probabilidad del evento A, el cual dará una medida precisa de la oportunidad de que A ocurra.

#### Probabilidad condicional

La probabilidad condicional describe cómo afecta la información de que “un evento B ha ocurrido” a la probabilidad asignada a un evento A. Se utilizará la notación  $P(A|B)$  para representar la probabilidad condicional de A dado que el evento B haya ocurrido. B es el “evento condicionante”. La probabilidad condicional de A dado B es proporcional a  $P(A \cap B)$ .

Para dos eventos cualesquiera A y B con  $P(B) > 0$ , la probabilidad condicional de A dado que B ha ocurrido está definida por:  $P(A|B) = P(A \cap B) / P(B)$

#### Independencia

En algunos casos, la probabilidad de que ocurra o haya ocurrido A no se ve afectada por el conocimiento de que B ha ocurrido, así que  $P(A | B) = P(A)$ . Es entonces natural considerar a A y B como eventos independientes, es decir que la ocurrencia o no ocurrencia de un evento no afecta la probabilidad de que el otro ocurra.

Cuando dos eventos son independientes,  $P(A \cap B)$  se calcula:

$$P(A \cap B) = P(A) * P(B)$$

### **Modelo logit**

Es un modelo de regresión en el cual la variable dependiente o de respuesta es de naturaleza cualitativa. Para explicar el comportamiento de una variable dicotómica es preciso utilizar una función de distribución acumulativa (FDA) seleccionada apropiadamente. El modelo logit utiliza la función logística acumulativa.

#### *Características*

1. Se pueden incluir tantas variables regresoras como indique la teoría subyacente.
2. Si  $L$ , el logit, es positivo, significa que cuando se incrementa el valor de la(s) regresora(s), aumentan las posibilidades de que la regresada sea igual a 1 (lo cual indica que sucederá algo de interés).
3. Para la interpretación del modelo:  $\beta_2$ , la pendiente, mide el cambio en  $L$  ocasionado por un cambio unitario en  $X$ , es decir, dice cómo cambia el logaritmo de las posibilidades en favor  $P=1$  a medida que la variable regresora cambia en una unidad.
4. A través de los efectos marginales se puede identificar la tasa de cambio de la probabilidad por cada unidad de cambio en la regresora.

### **Modelo probit**

Para explicar el comportamiento de una variable dicotómica es preciso utilizar una función de distribución acumulativa (FDA) seleccionada apropiadamente. El modelo probit utiliza una (FDA) normal. En el modelo todas las regresoras intervienen en el cálculo de los cambios en la probabilidad. A través de los efectos marginales se puede identificar la tasa de cambio de la probabilidad por cada unidad de cambio en la regresora.

### **Diferencia entre el modelo logit y probit**

Para la mayoría de las aplicaciones, los modelos logit y probit son muy semejantes; la principal diferencia es que la distribución logística tiene colas un poco más anchas. Lo anterior significa que la probabilidad condicional  $P_i$  se aproxima a cero o a uno con una tasa menor en el modelo logit, en comparación con el probit.

### **Criterio de información de Akaike (CIA)**

Es un criterio para juzgar la bondad de un modelo de regresión y el cual permite escoger entre modelos. Al comparar dos o más modelos, se preferirá el que tenga el menor valor CIA.

#### ➤ **Metodología**

Para identificar la población con doble afectación, se asumió:

A= ser afectado por desplazamiento forzado por expulsión (UARIV)

B= ser afectado por un desastre natural (UNGRD)

Se quiere estimar el número de personas que experimentan doble afectación por eventos de desplazamientos forzado y desastres naturales en los municipios de Colombia, por lo cual, lo que queremos es:

$$P(A \cap B) = P(B) * P(A/B)$$

Para identificar la  $P(A/B)$ , se requiere identificar la relación estadística derivada de los datos, no una relación causal. Por lo cual, se utilizó la metodología propuesta por Otálara y Ayala (2010) quienes intentaron identificar la relación que existe entre diferentes variables con los municipios que son expulsores de población en calidad de desplazamiento, y para lo cual realizaron una estimación por a través de un modelo probit.

A través de la utilización de este tipo de modelo, podemos identificar la independencia o relación del desplazamiento y los desastres. Al utilizar esta metodología, se puede obtener alguno de estos dos resultados:

$$(1) P(A/B) = \beta = 0$$

$$(2) P(A/B) = \beta \neq 0$$

Si sucede (1),  $P(A/B) = \beta = 0$ , significa que, en promedio, la probabilidad que un municipio sufra eventos de desplazamiento forzado es independiente de experimente un desastre natural.

Por tanto:

$$\text{Doble afectación} = P(A) * P(B)$$

$$= \text{Probabilidad de desplazamiento} * \text{Probabilidad de desastre}$$

Si sucede 2,  $P(A/B) = \beta \neq 0$ , significa que, en promedio, la probabilidad que un municipio sufra eventos de desplazamiento forzado aumenta  $\beta\%$  (efectos marginales) cuando experimenta un desastre natural. Esta correlación estadística servirá para encontrar la doble afectación, donde  $\beta$ , es el efecto marginal identificado y  $P(B)$  es la probabilidad de que ocurra un desastre natural.

$$P(A \cap B) = P(B) * \beta$$

Por tanto:

$$\text{Doble afectación} = \text{Total personas en el municipio} * P(A \cap B)$$

Con lo anterior, para identificar la asociación entre el desplazamiento y la ocurrencia de desastres naturales, se utilizó un modelo de regresión donde la variable dependiente era binaria (Municipio con personas desplazadas por expulsión 1= si, 0=no). Se incluyeron como variables regresoras:

homicidios, personas afectadas por minas antipersonas, por actos de terrorismo, por secuestro, por amenazas, por actos de violencia sexual, reclutamiento de niños, confinamiento, número de hectáreas de coca erradicadas y la presencia de desastres naturales, por municipio. Se utilizaron los modelos logit y probit y se utilizó la prueba de los criterios de información de akaike (CIA) para validar cual era el mejor de los modelos. Además, se estimaron los efectos marginales, toda vez que estos permiten determinar el porcentaje que aumenta la probabilidad de desplazamiento toda vez que ha ocurrido un desastre natural, cuando se encontró una asociación entre los dos eventos de interés.

Se plantearon 4 escenarios con base en los datos disponibles:

1. Datos agrupados de los 3 años (2016,2017, 2018)
2. Datos agrupados de los años 2016-2017
3. Datos para cada uno de los años (año 2016 y año 2017)

### ➤ Resultados

A continuación, se presentan los resultados de los 4 escenarios planteados y las conclusiones derivadas de cada uno.

#### Escenario 1. Datos agrupados de los 3 años (2016,2017, 2018)

La probabilidad que un municipio sufra eventos de desplazamiento forzado es independiente de que experimente un desastre natural. Los resultados pueden estar influenciados por los diferentes periodos de la información de 2018.

#### Escenario 2. Datos agrupados de los años 2016-2017

Las variables de homicidios, minas, terrorismo, confinamiento, delitos sexuales, reclutamiento de menores, hectáreas de coca erradicadas, no se encontraron asociadas estadísticamente con el desplazamiento.

#### Modelo Probit

```

Probit regression                Number of obs   =    1122
                                Wald chi2(3)     =    37.06
                                Prob > chi2        =    0.0000
Log pseudolikelihood = -230.67451 Pseudo R2      =    0.5393
    
```

YDespla2016_17	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
Secues16_17	1.138265	.3712352	3.07	0.002	.4106569 1.865872
Amenaza16_17	.6261232	.1656947	3.78	0.000	.3013675 .950879
DesastreSI16_17	.3058752	.1351789	2.26	0.024	.0409294 .5708211
_cons	-.3894312	.0984993	-3.95	0.000	-.5824862 -.1963761

CIA = 469.349  
Correctamente clasificados = 91.53%



## Modelo Probit

Probit regression  
 Number of obs = 1122  
 Wald chi2(4) = 72.09  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.4905  
 Log pseudolikelihood = -329.32182

YDespla2017	Coef.	Robust		z	P> z	[95% Conf. Interval]	
		Std. Err.					
Secues2017	.6215794	.3120381	1.99	0.046	.009996	1.233163	
Amenaza17	.5946869	.1211596	4.91	0.000	.3572184	.8321554	
Sexual17	.999019	.5754278	1.74	0.083	-.1287988	2.126837	
DesastreSI17	.5696834	.1154437	4.93	0.000	.343418	.7959488	
_cons	-.535736	.0748234	-7.16	0.000	-.6823873	-.3890848	

CIA = 668.6436  
 Correctamente clasificados = 86.27%

## Efectos marginales

Average marginal effects  
 Model VCE : Robust  
 Number of obs = 1122  
 Expression : Pr(YDespla2017), predict()  
 dy/dx w.r.t. : Secues2017 Amenaza17 Sexual17 DesastreSI17

	Delta-method		z	P> z	[95% Conf. Interval]	
	dy/dx	Std. Err.				
Secues2017	.1061959	.0532874	1.99	0.046	.0017544	.2106373
Amenaza17	.1016013	.0164766	6.17	0.000	.0693077	.1338949
Sexual17	.1706808	.0984935	1.73	0.083	-.0223628	.3637245
DesastreSI17	.0973295	.0196813	4.95	0.000	.0587548	.1359042

Cuando se analiza la información del año 2017, se logró identificar que los eventos no son independientes, por ello, si se presenta un desastre, la probabilidad de un desplazamiento aumenta en un 9,7%.

### ➤ Conclusiones

De los análisis presentados con anterioridad, se logró identificar la relación entre el desplazamiento y los desastres para cada uno de los escenarios. En la tabla se presenta la consolidación de lo identificado y como podría hacerse la estimación de la doble afectación para cada uno de los años analizados.

Año	Desastre – Desplazamiento to ¿son independientes?	$\beta$	DOBLE AFECTACIÓN
2016	Si	0	$P(A) * P(B) * \text{Total personas en el municipio}$

2017	No	9,7%	Total personas en el municipio * $P(A \cap B)$ , donde $P(A \cap B) = P(B) * \beta$
2016-2017	No	3,7%	Total personas en el municipio * $P(A \cap B)$ , donde $P(A \cap B) = P(B) * \beta$
2018	??	??	??

Para el 2018, pueden considerarse varias posibilidades y con ello, cambia el número de personas con doble afectación.

1. Que el comportamiento de 2018 sea como 2016 (0)  
→ Doble afectación año 2018 = 325
2. Que el comportamiento de 2018 sea como 2017 (9,7)  
→ Doble afectación año 2018 = 17.483
3. Asignar el comportamiento de los años 2016-2017 (3,7)  
→ Doble afectación año 2018 = 6.043

Teniendo en cuenta lo anterior, el número de personas con doble afectación en el país fueron:

Año 2016 = 9.789

Año 2017 = 58.704

Finalmente, se decidió suponer que el año 2018 tendrá un comportamiento similar al año 2017, por lo cual, la población con doble afectación en el país para los años 2016-2018, se estimó en **85.976** personas, lo que corresponde al 5,4% de las personas afectadas por desastres y al 29,9% de las personas afectadas por desplazamiento.

### ➤ Bibliografía

Devore, Jay L. *Probabilidad y Estadística para Ingeniería y Ciencias*. Séptima edición. ISBN-13: 978-607-481-338-8. ISBN-10: 607-481-338-8.

Gujarati, Damodar N; Porter, Dawn C. *Econometría*. Quinta edición. ISBN: 978-607-15-0294-0.