

# MODELOS REFERIDOS AL FENÓMENO DEL DESEMPLEO A PARTIR DE DATOS DE LA ENCUESTA PERMANENTE DE HOGARES: ALCANCES Y LIMITACIONES\*

Director del proyecto: Servy, E.

Integrantes: Hachuel, L.; Boggio, G.; Wojdyla, D; Marí, G; Cuesta, C.; Méndez, F.

Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística de la Facultad de Ciencias Económicas y Estadística. Universidad Nacional de Rosario.

Estimación y comparación de las tasas de desocupación de los aglomerados de la EPH en octubre 2002 utilizando modelos a efectos fijos y aleatorios.

## Para estimación de tasas: Modelo de efectos aleatorios

**Modelo sin offset** →  $\text{logit}(\pi_i) = \alpha + u_i, i=1, \dots, I$

$\pi_i$ : probabilidad de desocupación del  $i$ -ésimo aglomerado

$\{u_i\}$  los efectos aleatorios independientes con distribución  $N(0, \sigma^2)$ .

**Modelo con offset** →  $\text{logit}(\pi_i) = \text{logit}(q_i) + \alpha + u_i, i=1, \dots, I$

el término  $\text{logit}(q_i)$  es conocido y se denomina "offset".

## Para comparación de tasas: Modelo de efectos fijos

**Modelo sin offset** →  $\text{logit}(\pi_i) = \alpha + \beta_i, i=1, \dots, I$

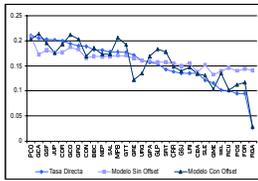
$\beta_i$  representa un efecto fijo y permite determinar si hay diferencias en las chances de desocupación entre dos áreas distintas.

**Modelo con offset** →  $\text{logit}(\pi_i) = \text{logit}(q_i) + \alpha + \beta_i, i=1, \dots, I$

$\text{logit}(q_i)$  conocido ("offset")

## RESULTADOS

### Estimación de tasas de desocupación



Las estimaciones a partir del modelo de efectos aleatorios sin offset presentan valores más "estables" o similares para todos los aglomerados, mientras que las estimaciones obtenidas con el modelo de efectos aleatorios con offset si bien son más irregulares, logran "captar" mejor valores extremos.

### Comparación de tasas de desocupación

De acuerdo al modelo sin offset, el efecto correspondiente a los aglomerados resulta significativo ( $p < 0.001$ ) lo cual indica que la chance de desocupación en octubre de 2002 no es la misma en todos los aglomerados.

De acuerdo al modelo con offset, el efecto aglomerado no resulta significativo ( $p = 0.9833$ ): igual razón de odds de desocupación en octubre 2002 vs. mayo 2002 en todos los aglomerados. Ese valor constante representa una reducción en la chance de desocupación en octubre 2002 de aproximadamente un 23% respecto a mayo de ese año.

## INVESTIGACIÓN FUTURA

Determinar bajo qué situaciones particulares resulta preferible la elección de modelos de efectos aleatorios como metodología apropiada para la obtención de buenas estimaciones.

Estudio de los cambios de condición laboral en el aglomerado Gran Rosario durante los años 2001 y 2002 (4 ondas) a través de modelos de transición.

## MODELO DE TRANSICION

controla dependencia entre las respuestas condicionando sobre el estado ocupacional previo

$$\text{logit } \pi_{i(t|t-1)} = x'_{i(t-1)}\beta + y_{i(t-1)}x'_{i(t-1)}\alpha$$

$\pi_{i(t|t-1)} = P(Y_{it}=1 | Y_{i(t-1)}=y_{i(t-1)})$ ,  $Y_{it}=1$  si el individuo  $i$ -ésimo estuvo desocupado en la onda  $t$ , 0 en caso contrario,

$x'_{i(t-1)}$  incluye variables explicativas, variable indicadora de la transición y posibles interacciones entre ellas,

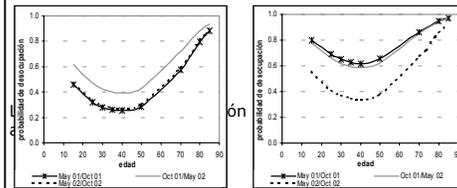
$y_{i(t-1)} x'_{i(t-1)}$  incluye la respuesta previa, posibles interacciones entre ella y las variables explicativas.

## RESULTADOS

La edad, nivel de ingreso, la rama de actividad y la transición influyen sobre los cambios de estado ocupacional de una onda a la siguiente.

Probabilidad de desocupación para varones de la rama construcción, con ingresos bajos según transición

estado previo ocupado    estado previo desocupado



Razones de odds entre estado ocupacional y nivel de ingreso

Nivel de ingreso	Razón de odds
Medio vs Bajo	0.60
Alto vs Bajo	0.42

## CONCLUSIÓN

El estado ocupacional previo es un fuerte predictor del estado ocupacional actual con variaciones según la transición que se trate y afectando en forma diferente la probabilidad de desocupación en cada onda según la rama de actividad.

Razones de odds entre estado ocupacional y estado ocupacional en el período previo según transición y rama de actividad

Rama de actividad	May 01/Oct 01	Oct 01/May 02	May 02/Oct 02
Construcción	4.75	9.03	5.11
Manufactura	9.37	17.82	10.08
Ss comerciales	28.89	54.92	31.06
Int Financiera	48.17	91.58	51.80
Adm Pub y def + Inst Pub y ss salud	17.91	34.05	19.26
Otras activ de ss	13.26	25.22	14.26

Evaluación mediante un estudio de simulación de la distorsión de las propiedades de los estimadores clásicos de la probabilidad de transición, cuando se los aplica a una muestra de estructura compleja.

## ESTIMADOR CLÁSICO (Anderson y Goodman)

$p_{ij}(t-1, t)$  probabilidad condicional de estar en el estado  $j$  en el tiempo  $t$  dado que estaba en el estado  $i$  en el tiempo  $(t-1)$ .

Estimador máximo verosímil

$$\hat{p}_{ij}(t-1, t) = \frac{n_{ij}(t)}{n_i(t-1)} \rightarrow n_{i*}(t) = \sum_{j=1}^k n_{ij}(t) = n_i(t-1)$$

Para una cadena de Markov con  $\sum_{j=1}^k p_{ij}(t) = 1$  probabilidades de transición estacionarias las estimaciones de las probabilidades de transición son

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{i*}} = \frac{\sum_{t=1}^T n_{ij}(t)}{\sum_{t=1}^T n_i(t-1)}$$

## SIMULACIÓN

Generación de matrices  $3 \times k$  (conglomerado) usando el modelo MMCC (Servy et al), considerando una cadena estacionaria de Markov de primer orden

$$p(x_2, x_3 | x_1) = p(x_1) p(x_2 | x_1) p(x_3 | x_2)$$

- Presencia de correlación intra-conglomerado
- Parámetros de la simulación  $p(x_2 | x_1) \quad p(x_3 | x_2)$

$p(x_2   x_1)$	$x_2$	
$x_1$	0	0.50
$x_1$	1	0.55
0	$p(0 0)$	0.60
1	$1-p(0 0)$	0.65
	$p(0 0)$	0.75

Correlaciones intra-conglomerado ( $\rho$ ): baja, media y alta

- Tamaño muestral ( $n$ ): 15, 30, 70, 100, 500, 1000
- Tamaños de conglomerados ( $k$ ): 2, 3, 4, 5
- Repeticiones: 1000
- Evaluación de los estimadores:

$$\text{sesgo}(\hat{p}(0|0)) = \frac{1}{4 * 6} \sum_k \sum_n \frac{(\hat{p}_{kn}(0|0) - p(0|0))}{p(0|0)} * 100$$

$$\text{sesgo}(\text{var}_{\text{mas}}) = \frac{1}{4 * 6} \sum_k \sum_n \frac{\text{var}_{\text{mas},kn} - \text{var}_{r,kn}}{\text{var}_{r,kn}} * 100$$

$\rho$	% de Sesgo relativo (en valores absoluto) de $\hat{p}(0 0)$ y de $\text{var}(\hat{p}(0 0))$					
	Tamaño de Muestra					
	15	30	70	100	500	1000
Baja	0.306 (24.6 6)	0.367 (22.6 0)	0.146 (21.8 6)	0.117 (23.2 6)	0.084 (24.0 9)	0.068 (22.6 9)
Medi	0.387 (49.7 4)	0.462 (52.8 8)	0.211 (48.6 6)	0.120 (50.4 9)	0.039 (51.3 0)	0.068 (49.1 8)
Alta	0.534 (65.0 5)	0.435 (59.1 1)	0.229 (64.4 9)	0.189 (54.7 1)	0.100 (64.6 0)	0.092 (64.5 9)

Sesgo relativo de  $\hat{p}(0|0)$  pequeño. Disminuye en valor absoluto cuando  $n$  aumenta. Aumenta en valor absoluto cuando las correlaciones entre los individuos del mismo conglomerado aumentan.

Verdadera variancia del estimador subestimada cuando se ignora el carácter complejo de la muestra. Esta subestimación aumenta cuando el tamaño de la correlación intra-clase sube, y no disminuye cuando se incrementa el tamaño muestral.

## INVESTIGACIÓN FUTURA

Determinar distorsiones de los errores de tipos I y II en el test de estacionariedad propuesto por Anderson y Goodman, bajo muestreo complejo. También se realizarán estudios sobre los estimadores clásicos que incluyen variables explicativas.