

Análisis Multinivel para el Estudio de Redes Egocéntricas: Aplicaciones en el estudio de la homofilia en Chile¹

*Alejandro Plaza Reveco²
Roberto Cantillan Carrasco³
Benjamín Muñoz Rojas⁴*

Resumen:

El objetivo de este capítulo es revisar la formulación estadística y las aplicaciones sustantivas a los modelos de regresión multinivel para el análisis de redes sociales para datos egocéntricos en encuestas de opinión pública. El análisis de redes sociales en las ciencias sociales históricamente ha comprendido dos tradiciones de investigación, por un lado, la aproximación sociocéntrica la cual analiza la red completa y por otro lado la aproximación que se concentra en datos egocéntricos o de redes personales (Marsden, 1990; Marin y Wellman, 2011). En la segunda aproximación, los encuestados (ego) reportan vínculos con alteris, donde los vínculos pueden dar cuenta de distintas formas de interacción entre egos y alteris. Al respecto los datos egocéntricos exhiben una clásica estructura multinivel donde los alteres o la relación entre alteris-ego se encuentra a nivel 1, anidados jerárquicamente en egos o en redes de egos a nivel 2. En primer lugar, se reseña la conceptualización detrás de las redes personales y el diseño de éstas redes para encuestas de opinión pública. En segundo lugar, describe la estructura multinivel para redes egocéntricas y su respectiva formulación estadística. En tercer lugar, se presentan distintas hipótesis de investigación que pueden ser contrastadas a partir de estos modelos, terminando finalmente con la aplicación de estos modelos en la Encuesta Longitudinal Social de Chile (ELSOC), elaborado por el Centro de Estudios de Conflicto y Cohesión Social.

Palabras Claves: Redes Sociales, Redes Egocéntricas, Modelos Multinivel, ELSOC.

¹ Los autores agradecen a ANID Chile su apoyo a través del Centro de Estudios de Conflicto y Cohesión Social (COES, fondo ANID/FONDAP/15130009), y del Concurso Fondecyt Regular 1171426 (IR Vicente Espinoza).

² Magíster en Sociología, Pontificia Universidad Católica de Chile. Asistente de Investigación en estudio ELSOC, COES.

³Magíster en Ciencias Sociales, Universidad de Santiago de Chile

⁴ Magíster en Ciencia Política, Pontificia Universidad Católica de Chile. Asistente de Investigación en estudio ELSOC, COES

Introducción

Durante el siglo XX, los esfuerzos de los analistas de redes se enfocaron en el desarrollo de técnicas para la investigación sistemática de la interacción de los niveles individual, diádico y posicional (Wasserman & Faust, 1994). La ampliación más reciente de la perspectiva egocentrada, ha facilitado la producción y tratamiento de datos de redes personales a nivel poblacional y, junto con esto, se ha expandido el interés por el desarrollo de una perspectiva interesada en la naturaleza dinámica y multinivel de las redes personales (Crossley et al., 2015; Krivitsky & Morris, 2017; Perry et al., 2018).

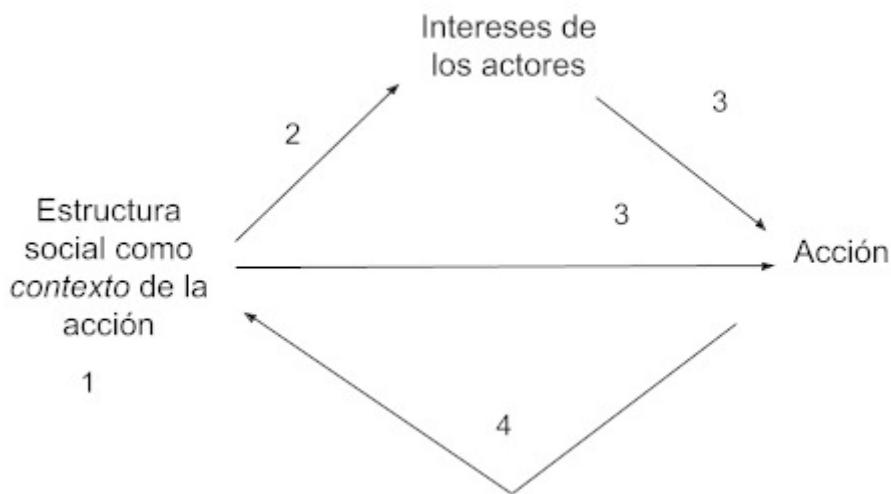
La naturaleza multinivel de las redes sociales puede abordarse con metodologías multinivel estándar o desde una perspectiva sociocéntrica (Lazega, 2015). De manera general, los modelos multinivel permiten a los investigadores modelar fenómenos complejos que se desarrollan en el intercurso de relaciones entre los niveles de ego, alteris, vínculos, y características de las redes (Wellman & Frank, 2001). Realizamos una pequeña introducción teórica a la naturaleza multinivel y dinámica de las redes personales, para luego enfocarnos en la descripción de la metodología multinivel para el análisis egocentrado. Concluimos con un ejemplo de análisis utilizando datos de la Encuesta Social Longitudinal de Chile ELSOC - COES, en su ola número 2.

Redes, contexto y acción

Los *contextos* pueden ser entendidos como estructuras multiplexadas y multinivel de interdependencias que proporcionan la base material sobre la cual se constituyen las actitudes e intereses de los individuos (Crossley, 2011; Erickson, 1988; White, 2008). De esta manera, la acción social puede ser definida como un proceso determinado y determinante, es decir, parcialmente deliberado, con capacidad para modificar el entorno y orientado por las oportunidades que brinda⁵ (Burt, 1982, 2000; Lin, 2002). Este proceso es ilustrado en la *Figura 1*.

⁵ En esta perspectiva, la acción coordinada no es solo es un medio de poder (Mann, 1991), sino que su despliegue exitoso tiende a establecerse como marco y fuente (preeexistente) más o menos formal, a la cual acuden los actores en formación para darle sentido al curso de sus intentos de control sobre el entorno (Gould, 1991, 1999; White, 2008).

Figura 1: Componentes de una teoría estructural de la acción.

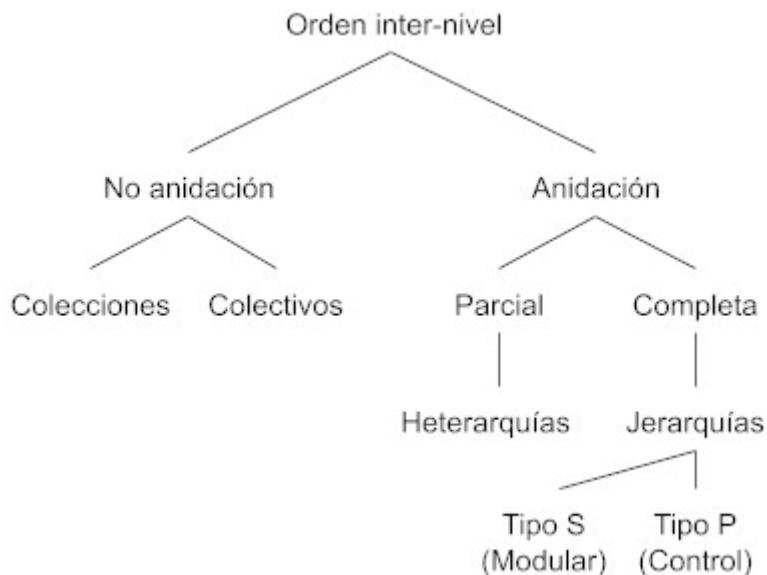


Fuente: Extraída y adaptada de Burt (1982a)⁶

La perspectiva de redes sociales considera que la comprensión del comportamiento humano exige el análisis sistemático de los contextos sociales. En este sentido, suelen incorporarse criterios pragmáticos de escalamiento, complejidad y grados de autonomía de las unidades o niveles de análisis (Kontopoulos, 1993). Una manera de ver los ordenamientos contextuales es distinguiendo dos formas básicas: *Las anidadas y no anidadas*. La *anidación* implica una sucesiva (y parcial) inclusión de niveles inferiores en los superiores en forma de partes o totalidades. Este proceso es ilustrado en la *Figura 2*. Los niveles interactúan de formas complejas y la tarea de analizar esto requiere de herramientas conceptuales y metodológicas sofisticadas (Johnson, 2008). Por ahora, lo importante es entender que cada nivel de organización-anidación posee autonomía parcial, lo cual implica propiedades particulares emergentes en cada uno de los niveles. Esta lógica de anidación parcial, o heterárquica, es propia de las lecturas actuales de las relaciones internivel. En esta lógica, las redes personales son un nivel anidación para los alteris reportados por ego. Al mismo tiempo, ego está anidado en otros niveles como barrios, comunas o ciudades.

⁶ Los números de esta figura, hacen referencia a tiempos secuenciales entre los distintos elementos.

Figura 2: Orden entre niveles estructurales.



Fuente: Adaptada de Kontopoulos, K. (2006) [1993].

Redes sociocéntricas y egocéntricas

En el análisis de redes sociales la unidad de análisis es la entidad sobre la cual se toman las medidas. De manera general, estas entidades suelen ser tres: a) el actor, b) la díada (lazos), y c) acontecimientos (cuando se analizan estructuras de afiliación) (Wasserman & Faust, 1994). Para su estudio se han desarrollado dos tipos generales de diseños muestrales: 1) los sociocéntricos y 2) los diseños egocéntricos (Krivitsky & Morris, 2017). Los análisis sociocéntricos tienen como objetivo recopilar la información completa de los egos y lazos (pueden ser sobre relaciones múltiples) presentes en una comunidad específica (factible de alcanzar), y los diseños egocéntricos se enfocan en las relaciones adyacentes a un ego, incluyendo los alteris, y las relaciones entre los alteris enunciados por el ego. El tipo habitual de datos de la red personal consiste en una muestra de encuestados, en donde para cada encuestado, se identifican todas las relaciones en su red personal o un número acotado (por ejemplo, en ELSOC COES, se solicita un máximo de 5 alteris) (Vacca, 2018). Ambos diseños han sido reconocidos como dos tradiciones históricamente distinguibles en la perspectiva del análisis de redes (Marin & Wellman, 2011; Marsden, 1990; Wasserman & Faust, 1994).

La producción de datos egocéntricos es realizada a través de "generadores de nombres"⁷ (Marin & Hampton, 2007). Con este tipo de instrumentos los egos son los entrevistados de una encuesta y los alteris no son directamente observados. Los datos de redes personales

⁷ Los generadores de nombres poseen ventajas importantes: a) pueden obtenerse a partir de muestras seleccionadas aleatoriamente, b) el número de alteris mencionados por cada ego es pequeño (por lo general con un límite de 5. ej. ELSOC-COES) lo cual hace que los datos sean muy manejables, c) son compatibles con métodos estadísticos inferenciales basados en la población (modelos de regresiones mono y multínivel) y, d) las identidades de ego y alteris pueden mantenerse anonimizadas, lo cual elimina el problema de la privacidad, además de fomentar reportes honestos (Handcock & Gile, 2010; Krivitsky & Morris, 2017; Morris, 1994; Perry et al., 2018).

pueden incluir variables respecto a egos, alteris, vínculos ego-alteris, vínculos alteris-alteris y a la red egocéntrica completa. Estas estructuras de datos exhiben una estructura multinivel. El nivel menor corresponde a los alteris o a los vínculos ego-alteris agrupadas dentro de unidades de nivel mayor, la red de los egos. Por lo general, los analistas que usan esta técnica están interesados en predecir características de los alteris o de las relaciones entre egos y sus alteris (Perry et al., 2018; T. A. B. Snijders & Bosker, 2012; Vacca, 2018).

Modelado multinivel para redes egocentradas

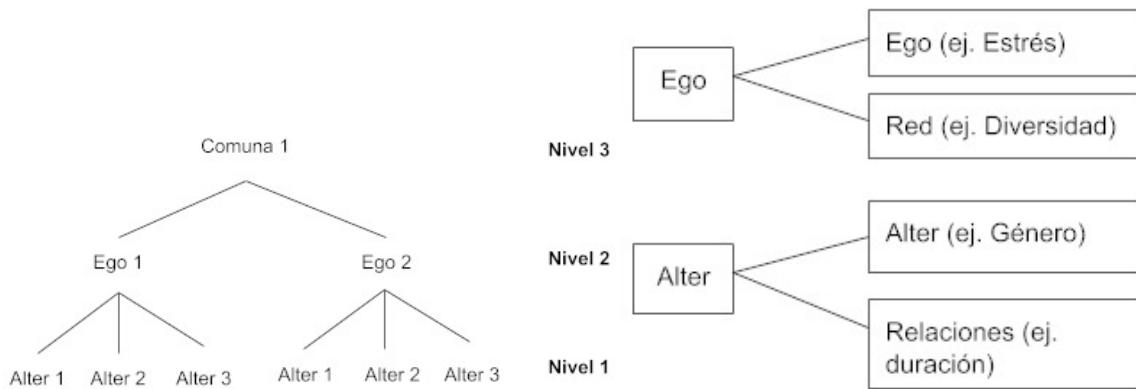
Las estructuras de datos jerárquicos contradicen el supuesto fundamental de los modelos estadísticos, esto es, que las observaciones de la variable dependiente son independientemente distribuidas. La agrupación de unidades a un nivel superior establece dependencia (por lo tanto, correlación) entre valores de la variable dependiente observada en los casos que pertenecen al mismo grupo o conglomerado.

Los modelos de efectos mixtos permiten modelar la heterogeneidad observada a partir de distintos niveles de análisis (Bell et al., 2019). El modelo lineal jerárquico se puede utilizar como un tipo de análisis de regresión para datos de redes personales si se aplican las siguientes condiciones: 1) la variable dependiente está en el nivel más bajo, es decir, el nivel de la relación o de alteris; 2) los datos no contienen superposición de redes personales de diferentes egos, o al menos esta superposición es insignificante; y 3) los datos obtenidos de diferentes egos (encuestados) son mutuamente independientes⁸ (T. Snijders et al., 1995).

De esta manera, los analistas de redes personales pueden examinar la variación de una variable dependiente de vínculo ego-alteris (por ejemplo, el soporte brindado por un alter), como una función de las características de los vínculos por sí mismo, de los encuestados que nominan a los alteris (los egos), los nominados (los alteris), y el amplio contexto en los cuales los vínculos ego-alteris están incrustados. Más raramente, los alteris han sido tratados como variables de nivel 1 y las características de los alteris como variables dependientes (T. Snijders et al., 1995; Wellman & Frank, 2001). La *Figura 4* ilustra los niveles de análisis y el ordenamiento jerárquico de los datos egocéntricos.

Figura 3: niveles de análisis y ordenamiento jerárquico para datos egocéntricos.

⁸ En Lazega et al. (2015) se discuten y ejemplifican varios modelos que relajan estos supuestos para redes egocentradas y no egocentradas.



Fuente: Adaptada de Perry et al. (2018).

Los alteris o los vínculos tienden a estar correlacionados al estar asociados al mismo ego, esto quiere decir que, son más similares entre sí en comparación con los alteris o vínculos que están asociados a un ego diferente (por ejemplo, en términos de comportamiento de fumadores, la frecuencia de contacto, el nivel de apoyo que se provee). Estas similitudes pueden ocurrir por bastantes razones que pueden ser observadas o no observables (o desconocidas) y, por lo tanto, no pueden ser incorporadas a partir de variables de control en un modelo estadístico. Algunas de estas razones pueden ser por ejemplo, la tendencia de homofilia en las redes sociales, donde los egos tienden a estar asociados a alteres similares (Blau, 1993; McPherson et al., 2001).

Los modelos jerárquicos permiten dar cuenta de la dependencia existente entre los distintos contactos de un ego al permitir uno o más de coeficiente aleatorio por las variables que dependen por cada ego. A continuación, se ejemplifica estos modelos con datos de redes personales. Usando la notación de clasificación (Browne et al., 2001; Browne & Goldstein, 2010), se indexa las unidades de nivel uno (alteris) como i , y las unidades de nivel dos (egos) como $ego(j)$. El término $ego(j)$ es la función de clasificación que regresa al ego que corresponde al alteris i . En los siguientes modelos, x_i indica una variable explicativa que mide a nivel de los alteris o de los vínculos ego-alteri; y $w_{ego(j)}$ indica una variable explicativa de variables observadas a nivel de los egos o de las redes de los egos (nivel 2). Finalmente, y_i indica la variable dependiente, la cual siempre es una característica de nivel uno de los vínculos o de los alteris.

Para recalcar las posibles preguntas de investigación que se pueden trabajar con estos tipos de modelos, presentamos las distintas hipótesis de trabajo que se pueden conjutar, adaptado de las investigaciones de Vacca (2018) en la Tabla 1.

Tabla 1 : Hipótesis para trabajar con modelo multinivel para redes Ego.

Variable explicativa	Unidad de observación	Nivel en la jerarquía	Ejemplo de variable	Ejemplo de hipótesis
Característica del vínculo	Vínculo	Nivel 1	Cercanía emocional entre Ego y alteris (Categórica)	Las personas que el ego ve como emocionalmente más cercanas tienen más probabilidades de brindar apoyo al ego.

Característica de la diada	Diada	Nivel 1	Ego y alteris son del mismo sexo (dicotómica)	Las personas que son del mismo sexo que el ego tienen más probabilidades de brindar apoyo al ego.
Característica de alteris	Alteris	Nivel 1	Edad de alteris (conteo)	Las personas mayores tienen más probabilidades de brindar apoyo al ego.
Característica de Ego	Ego	Nivel 2	Estatus marital de Ego (categórica)	Los egos casados tienden a disfrutar más apoyo social en general
Característica de composición de la red	Red	Nivel 2	Proporción de alteris que son familiares de Ego (continua)	Todos los alteris (familiares y no familiares) tienen más probabilidades de brindar apoyo al ego cuando la red está más dominada por la familia.
Característica de la estructura de la red	Red	Nivel 2	Densidad de la red personal (continua)	Es más probable que todos los alteris brinden apoyo al ego cuando son parte de redes más cohesivas (densas)
Interacción: Característica del vínculo x característica de la composición de la red	Vínculo y red	Nivel 1 x Nivel 2	Vínculo es familia (binaria) x proporción de alteris que son familiares (continua)	Es más probable que los lazos familiares sean de apoyo cuando hay una mayor proporción de miembros de la familia en la red
Interacción: Característica de alteri x característica de la estructura de la red	Alter y red	Nivel 1 x Nivel 2	Edad de alteri (Conteo) x densidad de la red personal (continua)	Las personas mayores tienen más probabilidades de proporcionar soporte en redes más cohesivas (densas)
Interacción: Característica de alteri x Característica de Ego	Alter y Ego	Nivel 1 x Nivel 2	Sexo de alteri (dicotómica) x edad de ego (conteo)	Las mujeres tienen más probabilidades de brindar apoyo cuando el ego es mayor

Fuente: Adaptada de Vacca (2018).

Modelos lineales jerárquicos

En los modelos lineales jerárquicos para redes personales, y_i , puede representar, por ejemplo, una variable continua que mide la fortaleza del vínculo entre ego y alter, u otro constructo que sintetice percepciones de cercanía emocional y confianza. El modelo puede posicionar el vínculo entre ego-alter de manera que la fortaleza dependa, de si es miembro de la familia de ego x_j , ($x_j = 1$ si alteris es miembro de la familia de ego, $x_j = 0$ de otra forma), y la edad de ego ($w_{ego(i)}$). Por ejemplo, podríamos conjeturar que los vínculos con los miembros de la familia son sistemáticamente más fuertes, y que los egos con más edad, en general tienden a reportar vínculos más fuertes dentro de sus contactos. Variables explicativas continuas en los modelos jerárquicos usualmente están centradas alrededor del promedio general (la gran media) o a la media del ego (centrado al grupo) para permitir la

interpretación de los coeficientes aleatorios y sus varianzas⁹. Por simplicidad, los modelos en esta sección sólo incluyen una variable x_j y una variable de nivel dos $w_{ego(i)}$, ellas pueden extenderse con múltiples variables de nivel 1 ($x_{1i}, x_{2i}, \dots, x_{ki}$) y de variables de nivel 2 a nivel de ego ($w_{1,ego(i)}, w_{2,ego(i)}, \dots, w_{q,ego(i)}$) como variables explicativas

A manera de ejemplo para el caso de los modelos lineales podríamos formalizar la fortaleza del vínculo de la siguiente forma¹⁰:

$$\begin{aligned}y_i &= \beta_{0,ego(i)} + \beta_1 x_i + \varepsilon_i \\ \beta_{0,ego(i)} &= \gamma_{00} + \gamma_{01} w_{ego(i)} + \mu_{0,ego(i)} \\ \varepsilon_i &\sim N(0, \sigma_e^2) \\ \mu_{0,ego(i)} &\sim N(0, \sigma_{\mu 0}^2)\end{aligned}$$

Equivalentemente si tomamos las primeras dos primeras ecuaciones, la podemos expresar del siguiente modo.

$$y_i = \gamma_{00} + \gamma_{01} w_{ego(i)} + \mu_{0,ego(i)} + \beta_1 x_i + \varepsilon_i$$

Los parámetros en este modelo son bastante similares en referencia a su contraparte en los modelos lineales de un sólo nivel. Sin embargo, el intercepto $\beta_{0,ego(i)}$ no es un parámetro fijo, sino una variable aleatoria que toma los distintos valores por cada $ego(i)$. Estos valores dependen de las características específicas de ego ($\mu_{0,ego(i)}$). Este efecto aleatorio puede ser conceptualizado como la desviación del promedio de los vínculos estrechos en la población. Entre los parámetros no aleatorios se encuentra γ_{01} el cual en este caso hipotético mide la asociación entre las características del ego $w_{ego(i)}$ (por ejemplo, la edad del ego) y las características generales para aquel ego. Por su parte, β_1 indica la asociación fija entre el alteri o el vínculo de las características de x_j (si el alteri es familiar), y la fortaleza del vínculo similar al modelo lineal estándar. De esta manera la fortaleza del lazo y_i depende de ciertas características específicas de cada ego (por ejemplo, la edad), en el cual el alteris está anidado a $ego(i)$ y en la desviación aleatoria que es unívoca para cada $ego(i)$.

Dado que todos los vínculos i están asociados al mismo $ego(i)$ comparten la misma desviación aleatoria $\mu_{0,ego(i)}$, el modelo da cuenta de la correlación entre los vínculos anidados en el mismo ego. En otras palabras, $\mu_{0,ego(i)}$ captura cualquier característica no observada de $ego(i)$ que (después de dar cuenta de las características observadas de egos y vínculos, $w_{ego(i)}$ y x_i) puede causar que los vínculos de $ego(i)$ son más parecidos en sus valores de fortaleza (es decir, que estén correlacionados). De esta manera, el modelo da cuenta que el resultado y_i puede estar potencialmente afectada por las características del ego ($w_{ego(i)}$) y un conjunto de aspectos no conocidos, no observables o no observados,

⁹ En el contexto de las regresiones multinivel o jerárquicas, el centrado de las variables puede realizarse hacia la gran media o hacia las medias dentro de los grupos. Las implicancias de los distintos tipos de centrado tienen implicancias sustantiva al momento de responder preguntas de investigación pues el centrado permite observar de manera más nítida la relación entre e intra los grupos. Para una revisión detallada revisar (Enders & Tofghi, 2007).

¹⁰ Para consideraciones de la aplicación práctica de este capítulo trabajaremos con una variable dicotómica de si el alteri tiene la misma religión que ego, la cual se verá en el apartado de modelos jerárquicos logísticos.

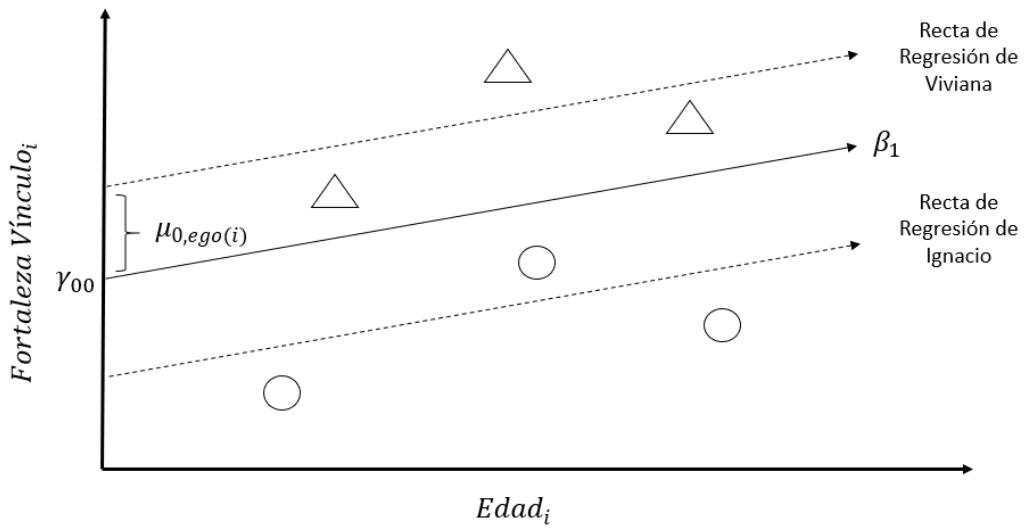
incorporados en $\mu_{0,ego(i)}$. Dado que se incluyó un coeficiente fijo (β_1) y un coeficiente aleatorio ($\beta_{0,ego(i)}$), modelos de efectos aleatorios o de efectos mixtos (Searle et al. 1992).¹¹

El efecto aleatorio a nivel de ego $\mu_{0,ego(i)}$ sigue una distribución normal con promedio cero y varianza constante, $\sigma_{\mu_0}^2$ el cual es uno de los parámetros a estimar. Además, el intercepto aleatorio también sigue una distribución normal: $\beta_{0,ego(i)} \sim N(\gamma_{00} + \gamma_{01}w_{ego(i)} + \mu_{0,ego(i)})$. La varianza $\sigma_{\mu_0}^2$ cuantifica la variabilidad entre los egos, es decir el monto de variación en y_i que es la diferencia sistemática entre egos, más allá de la variación que es explicada por los predictores incluidos en el modelo. En contraste ε_i representa la desviación aleatoria por cada vínculo i del promedio de la fortaleza promedio de los vínculos asociados al mismo $ego(i)$. El error residual de nivel 1 asume también una distribución normal con media cero y varianza constante, σ_e^2 , la cual es estimada como un parámetro del modelo. De esta forma la varianza σ_e^2 cuantifica el monto de varianza no explicada que es atribuible a las diferencias entre los vínculos en el mismo ego, en vez de que sea entre egos.

Para simplificar de manera visual el anterior modelo, en la Figura 5 se presenta un gráfico de puntos con la relación entre la edad y la fortaleza del vínculo de los alteris. Los triángulos representan los alteris reportados por un ego (o encuestada), que genéricamente llamaremos Viviana, y los círculos representan los alteris de un ego que genéricamente llamaremos Ignacio. Utilizando un modelo de regresión lineal estándar, cada encuestado tendría el mismo intercepto, no obstante, al realizar un modelo de regresión jerárquico, Ignacio y Viviana tendrán su propio intercepto aleatorio $\mu_{0,ego(i)}$ que refleja las diferencias promedio dentro de los grupos (egos) y con la gran media. En este caso la recta de regresión de Viviana esta por sobre la gran media, y esto se puede deber a características propias de Viviana (como su género, edad, entre otras) que puede apalancar a sus alteris, al compararlos con los de Ignacio.

Figura 4: Gráfico de puntos con rectas de regresión con intercepto aleatorio para la relación entre edad y fortaleza de los vínculos entre alteris

¹¹ Los modelos de efectos mixtos permiten hacer inferencia acerca de la población (en este caso los egos) y del mismo modo permiten al investigador estudiar los atributos de las unidades de nivel 2 asociadas a la varianza de nivel 1, lo que permite estudiar efectos heterogéneos o interacciones de nivel cruzado. En contraste a esta especificación existen los modelos de efectos fijos, en donde el interés teórico se centra en las variables de nivel 1, donde si bien se puede controlar por la heterogeneidad no observada a nivel 2, no es posible modelarla.



Fuente: Elaboración propia

Tal como lo señalan Snijders y Bosker (2012, p.49) los modelos multínivel asumen que los efectos aleatorios de los distintos niveles son independientes, lo que implica que la variación no explicada en el nivel 1 (alteri) del promedio del grupo no está sistemáticamente correlacionada con la desviación no explicada a nivel de grupo (ego). El supuesto de la independencia de residual entre niveles implica que en los modelos con intercepto aleatorio, el total de la variación es la suma simple de la varianza entre egos y de la varianza entre alteris:

$$Var(y_i) = \sigma_{\mu 0}^2 + \sigma_e^2$$

Esta ecuación permite descomponer la varianza en dos fuentes: a nivel dos por las diferencias sistemáticas entre los entrevistados y, a nivel 1 que representa la diferencia entre los vínculos asociadas a un mismo entrevistado. Dado esto, se puede calcular la correlación intra-clase para los distintos vínculos

$$ICC = \frac{\sigma_{\mu 0}^2}{\sigma_{\mu 0}^2 + \sigma_e^2}$$

La correlación intraclass cuantifica la proporción de la variación a nivel de vínculos (nivel 1) que se debe a la variación de los egos. Por lo general, en los modelos multínivel la correlación intraclass aporta una primera impresión de cuánta varianza está explicando los niveles de anidación.

Modelos logísticos jerárquicos

Las variables binarias son comunes en las ciencias sociales y, en el estudio de las redes sociales personales no es la excepción. Para este tipo de variable utilizamos los modelos logísticos, los cuales pueden ser derivados de los modelos lineales generalizados y los modelos logísticos jerárquicos pueden ser aplicados en base a esta misma lógica.

Un ejemplo que podemos tomar para el estudio de redes personales es si el alteris tiene la misma característica que ego (fenómeno conocido como homofilia), en donde si el alter tiene la misma característica por ejemplo en educación ($y_i = 1$) o no ($y_i = 0$). Los modelos logísticos jerárquicos asumen un evento $y_i = Bernoulli(\pi_i)$ y expresa la probabilidad de que ego y alteris tengan la misma característica. Como una función logística de un predictor lineal. Equivalentemente, el modelo expresa el logaritmo natural de las odds ($\log[\frac{\pi_i}{1-\pi_i}]$) como un predictor lineal. El predictor lineal puede ser incluido como variables a nivel del lazo o del alteri (x_i), como también de un intercepto y pendiente aleatoria que depende de las variables a nivel de ego ($w_{ego(i)}$)

$$\begin{aligned}\pi_i &= \frac{\exp(\beta_{0,ego(i)} + \beta_{1,ego(i)}x_i)}{1 + \exp(\beta_{0,ego(i)} + \beta_{1,ego(i)}x_i)} \\ \beta_{0,ego(i)} &= \gamma_{00} + \gamma_{01}w_{ego(i)} + \mu_{0,ego(i)} \\ \beta_{1,ego(i)} &= \gamma_{10} + \gamma_{11}w_{ego(i)} + \mu_{1,ego(i)}\end{aligned}$$

Las ecuaciones anteriores describen cómo ciertas características de los egos como ($w_{ego(i)}$) afectan sistemáticamente la probabilidad global de que el alter tenga la misma característica que el ego, y el impacto de x_i en esta probabilidad (a partir del $\beta_{1,ego(i)}$)

Es importante destacar que los modelos logísticos jerárquicos (como los modelos de un nivel) no incluyen un error residual de nivel 1 (ε_i), porque estos predicen la probabilidad de cierto resultado, más que el valor estimado en sí mismo. Además, la varianza de la distribución binomial depende del valor esperado, cambiando el significado y la interpretación de la varianza de nivel residual. Para derivar el modelo logístico, se considera una variable latente $y_i *$ con una escala arbitraria que determina el valor del resultado observado ($y_i = 1$ cuando $y_i * > 0$) y su error residual asumen que sigue una distribución logística estándar con varianza fija de $\sigma_e^2 = \pi^2/3$ (Snijders y Bosker, 2012, p.303). En los modelos logísticos, $\pi^2/3$ es una escala arbitraria que es atribuible a $y_i *$ para garantizar la identificación del modelo sin significado sustantivo. En la medida de que la varianza total de $y_i *$ es arbitraria es igual a la varianza de nivel 1 más cualquier varianza de nivel 2. Como resultado, los efectos fijos estimados también tienden a preservar la razón fija de la variación no explicada de $y_i *$

Tomando la varianza de nivel 1 fija en consideración, la correlación intraclase para el modelo con intercepto aleatorio se puede expresar de la siguiente forma:

$$ICC = \frac{\sigma_{\mu_0}^2}{\sigma_{\mu_0}^2 + \pi^2/3}$$

El cual mide el monto de variación en la variable de propensión continua del modelo latente (Snijders y Bosker, 2012, p. 303). Para ejemplificar el modelo anteriormente descrito llevamos a cabo un análisis multinivel logístico para evaluar la fuerza de relación de algunos predictores predictores pertinentes con la homofilia religiosa.

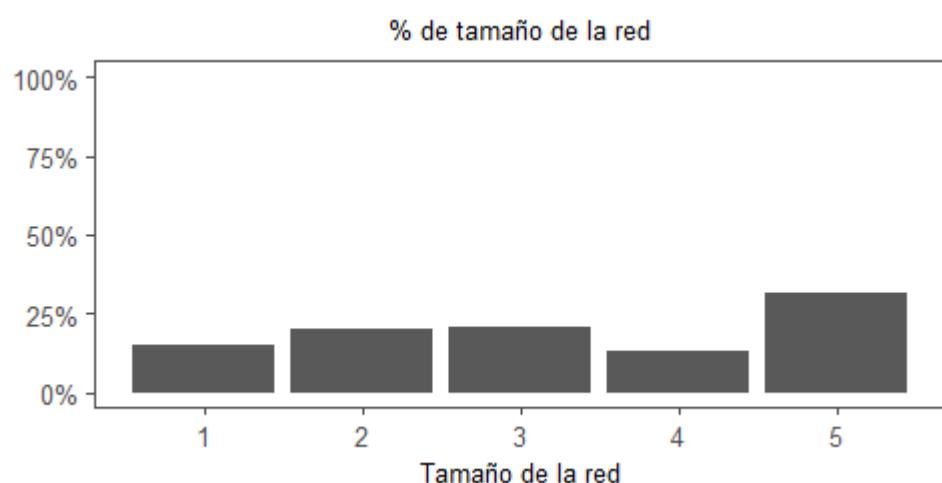
Datos y métodos

Se trabajará con la Encuesta Longitudinal Social de Chile (ELSOC) elaborada por el Centro de Estudios de Conflicto y Cohesión Social (COES). Para la primera ola de este estudio del 2016, se seleccionó una muestra aleatoria de 92 comunas (ciudades) de Chile y dentro de éstas se seleccionaron 1.067 bloques o manzanas), para luego seleccionar aleatoriamente de tres a cinco direcciones. Finalmente, se selecciona una persona dentro de la dirección entre 18 y 75 años, a las cuales se le realizó una entrevista cara a cara por medio del sistema CAPI. La muestra 2016 es representativa de la población urbana, lo que equivale al 77% de la población total del país, y al 93% de la población urbana.

Para consideraciones de este análisis utilizaremos la ola 2017 en la cual se aplicó la batería de generador de nombres, que como se mencionó anteriormente permite caracterizar los ambientes interpersonales de los encuestados (Burt, 1984; Marsden, 1990). A cada encuestado se le aplicó la siguiente pregunta “*Ocasionalmente la gente conversa de asuntos que le importan con otras personas. Pensando en los últimos seis meses, ¿quiénes son las personas con las que Ud. ha conversado?*”. Después de que los encuestados indican los nombres de pila de hasta un máximo de cinco personas con las que conversan, (sus alteris), se les consultó acerca sobre el género, edad, educación, religión, posición política y tipo de relación que tienen con estas personas. Para este caso, los entrevistados pueden mencionar hasta un máximo de cinco nombres. El límite al número máximo de posibles confidentes se tomó siguiendo las recomendaciones de Burt (1984), que busca conciliar un número suficiente para capturar la heterogeneidad de los encuestados y las restricciones metodológicas asociadas al tiempo de aplicación de la encuesta y la fatiga que esto conlleva para los encuestados.

En el Gráfico 1 se presentan los porcentajes de la cantidad de alteris reportados por los encuestados. En esta figura se puede apreciar de que el 32% de los encuestados reporta tener cinco confidentes, mientras que las personas que reportan un alteri, representan el 15% de la muestra.

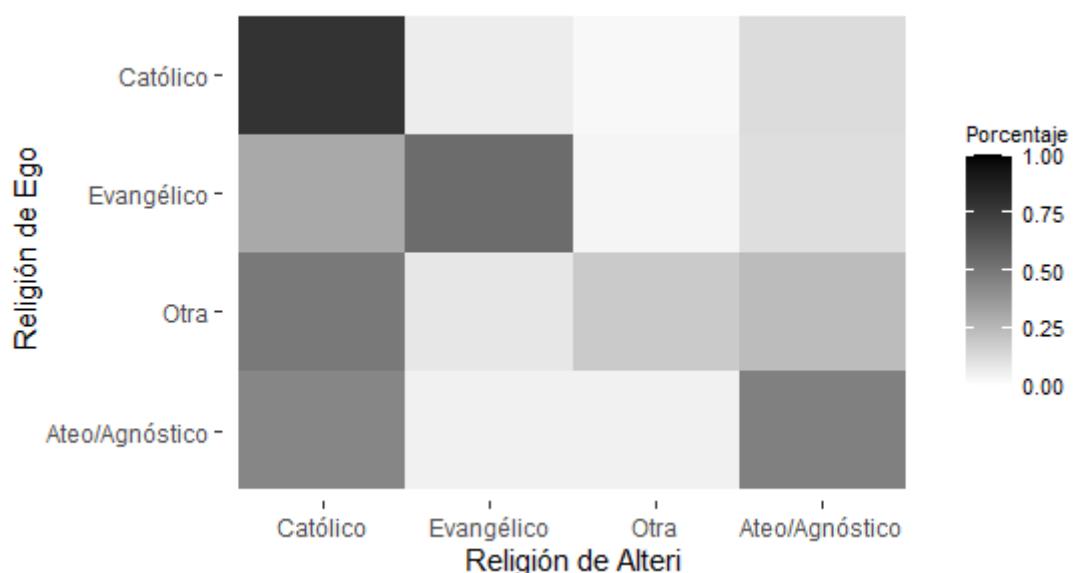
Gráfico 1: Porcentaje de Tamaño de la red de Alteris. Resultados ponderados con diseño muestral complejo



Fuente: Elaboración propia en base a ELSOC 2017

Por otro lado, en el Gráfico 2 se representa la tabla de contingencia de religión de los encuestados por cada uno de sus alteris. Cada cuadrante del gráfico representa el porcentaje de diadas ego-alteri, respecto al total encuestados observados (perfil fila). Los análisis estadísticos muestran que existen una asociación significativa entre la religión de los egos y los alteri ($\chi^2=3593,5$ (9); $p<0.001$). A nivel general se puede observar como las personas católicas tienden a reportar altos porcentajes de personas católicas entre sus alteris, mientras que se reportan menores niveles de alteres de las otras categorías. Los evangélicos y los agnósticos reportan del mismo modo porcentajes elevados de homofilia religiosa, no obstante en conjunto con las personas de otra religión de todas maneras se reporta altos niveles alteris católicos.

Gráfico 2: Diadas Ego y Alter que discuten asuntos importantes por religión. Resultados ponderados con diseño muestral complejo



Fuente: Elaboración propia en base a ELSOC 2017

En la ola 2017 de la encuesta ELSOC, se aplicó un total de 2.488 encuestas. Para poder realizar el análisis de regresión jerárquico o multínivel es necesario anidar a la cantidad de alteris nombrados por los encuestados (o los egos), es decir transmutar la base de datos de formato *wide* (o ancho) a un formato *long* (o largo). De esta forma pasamos de una base de datos donde la unidad de análisis son los alteris que nombran los encuestados pasando de 2.428 casos a 7.656 casos.¹²

En este caso analizaremos la propensión que tienen los encuestados de tener un alteri que comparten la misma religión, lo que en la literatura de redes sociales se conoce como homofilia (McPherson et al., 2001). En Norteamérica se han encontrado tasas decrecientes de homofilia para el género, estabilidad para la raza y la edad y, un débil pero significativo aumento de la homofilia educativa y religiosa (DiPrete et al., 2011; Smith et al., 2014). En Europa se ha encontrado que en regiones socio demográficamente diversas la similitud

¹² Casos completos sin valores perdidos

religiosa predice significativamente la homofilia (Adida et al., 2015). Antecedentes respecto al caso chileno (Bargsted Valdés et al., 2020) señalan que los chilenos tienden a tener interacción religiosa 2,4 veces más frecuentes en relación con un comportamiento aleatorio. Después de la edad, la religión es uno de los parámetros sociales más determinantes para entender las pautas de interacción dentro y fuera de los grupos (religiosos). A partir de esto se propone el siguiente modelo:

$$\text{logit}(\pi_i) = \gamma_{00} + \gamma_{01}\text{Religión}_{ego(i)} + \gamma_{02}\text{Edad}_{ego(i)} + \gamma_{03}\text{Sexo}_{ego(i)} + \gamma_{04}\text{Educación}_{ego(i)} \\ + \beta_1\text{Lazo}_i + \beta_2\text{Edad}_i + \beta_3(\text{Sexo}_i - \overline{\text{Sexo}_{ego(i)}}) + \beta_4\text{Educación}_i$$

Por sencillez se presenta un modelo expresado en términos de logit. Si bien puede parecer extraño que hay variables repetidas dos veces en la anterior ecuación, la razón de esto es porque unas hacen referencia a las variables asociadas a ego (con el sufijo *ego(i)*) y otras están asociadas a las características propias de los alteris (con el sufijo *i*). Para estimar la probabilidad de que un alteri tenga la misma religión que el ego, se considera la religión, la edad, el sexo y la educación del ego. Éstas serían las denominadas variables de nivel 2 (estas variables son todas aquellas con el sufijo *ego(i)*), donde el coeficiente γ expresa la pendiente de nivel 2 de estas variables. Por otro lado, a nivel 1, se considera el tipo de relación que tiene el ego con alteri (amigo/a, esposo/a, o familiar), la edad, el sexo y la educación del alteri, estas variables están siendo multiplicadas por un coeficiente β , la cual es la pendiente de nivel 1 de estas variables. Cabe a destacar que la variable sexo a nivel de alteris esta centrada al grupo, de manera que se pueda observar nítidamente el efecto dentro de los egos. Para estimar este modelo se utilizó el software estadístico R, y la librería GLMMadaptive

Resultados

En primer lugar, se calcula la correlación intra-clase (o ICC). Para llevar a cabo esto en primer lugar se estima un modelo nulo, con ningún predictor a la probabilidad de que el alteri tenga la misma religión que el ego

$$ICC = \frac{\sigma_{\mu 0}^2}{\sigma_{\mu 0}^2 + \pi^2/3} = \frac{5.55}{5.55 + \pi^2/3} = 0.63$$

Es decir, el 63% de la variación en que los alteris tengan la misma religión de los egos corresponde a su pertenencia a las unidades de nivel 2, en este caso los egos.

En la tabla 2 se presenta los resultados del modelo jerárquico logístico tomando como variable dependiente el logaritmo natural de la probabilidad de que el alteri tenga la misma religión que ego o el encuestado. En esta tabla, las primeras variables refieren a los predictores asociados al ego (nivel 2), y posteriormente a los predictores asociados a cada uno de los alteris (nivel 1)

Tomando como referencia la religión católica, se puede observar que los egos que tienen religión evangélica, ninguna/atea, o tienen otra religión se reduce bastante la probabilidad de que tengan alteris homofílicos en relación con la religión, siendo esta relación estadísticamente significativa. Al observar las variables de edad, sexo y educación, no se observan relaciones estadísticamente significativas.

Que los católicos sean más homofílicos que los otros grupos puede tener dos explicaciones plausibles. Por un lado, que este grupo por sus características específicas como grupo tienden a mantener relaciones endogrupales mucho más fuertes, generando "clausuras de grupo" en las que se reduce la interacción con otros grupos en el seno de los grupos "que hablan cosas importantes. Por otro lado, que los católicos sean mucho más homofílicos que los otros grupos, puede estar asociado a la disponibilidad demográfica de este grupo dentro de la población general. Es decir, dado que el grupo de los católicos son el grupo más numeroso en la población, la probabilidad de que los católicos se encuentren y generen relaciones de cercanía es mucho más alta que el resto de los grupos.

Tabla 2: Modelo de Regresión Logística Multinivel para Homofilia Religiosa.

	Variable	Coeficiente	Error Estándar	Valor Z	Valor p
	Intercepto	1,57	0,26	6,02	< 0,00
Predictores Ego	Ref=Católica				
	Evangélico	-4,82	0,22	-22,17	< 0,00
	Ninguna / Ateo	-1,59	0,14	-11,75	< 0,00
	Otra religión	-2,74	0,16	-16,95	< 0,00
	Ref=18-24				
	25-34	-0,22	0,21	-1,04	0,30
	35-44	-0,32	0,21	-1,52	0,13
	45-54	-0,12	0,21	-0,59	0,55
	55-64	-0,19	0,22	-0,87	0,38
	65-74	-0,07	0,24	-0,30	0,77
	75 o +	0,48	0,53	0,91	0,36
	Ref=Mujer				
	Hombre	-0,12	0,11	-1,12	0,26
	Ref=Ed. Básica o menos				
	Ed. Media incompleta	-0,10	0,18	-0,56	0,58
	Ed. Media completa	-0,05	0,16	-0,32	0,75
	Ed. Técnica superior	0,09	0,18	0,48	0,63
	Ed. Universitaria	0,11	0,19	0,61	0,54
Predictores Alteris	Ref=Amigo/a				
	Esposo/a	0,58	0,12	4,92	< 0,00
	Familiar	0,59	0,08	6,99	< 0,00
	Ref=18-24				
	25-34	0,15	0,13	1,11	0,27
	35-44	0,11	0,14	0,85	0,40
	45-54	0,16	0,14	1,18	0,24
	55-64	0,19	0,15	1,28	0,20
	65-74	0,18	0,18	1,04	0,30
	75 o +	0,59	0,24	2,50	0,01
	Ref=Mujer				
	Hombre centrado grupo	-0,17	0,08	-2,09	0,04
	Ref=Ed. Básica o menos				
	Ed. Media incompleta	-0,07	0,15	-0,51	0,61
	Ed. Media completa	-0,09	0,12	-0,73	0,46
	Ed. Técnica superior	0,15	0,15	1,00	0,32
	Ed. Universitaria	-0,04	0,14	-0,25	0,80
	N Observaciones	7656			
	N grupos	2428			
	$\sigma_{\mu 0}^2$	1,72			
	Logaritmo de Verisimilitud	-3884,88			
	AIC	7.827,76			
	BIC	7.995,81			

Fuente: *Elaboración propia en base a ELSOC 2017*

Al observar las variables asociadas a los alteris, se puede observar que el tipo de relación que tienen los egos con los alteris es una variable bastante relevante para entender la homofilia por religión, dado que si el alteris es esposo/a o familiar del ego -en comparación- de que si son amigos aumenta la probabilidad de que el alteris tenga la misma religión que ego, siendo esta relación estadísticamente significativa al 99% de confianza ($p<0.01$). Por otro lado, al observar la edad de los alteris, se puede establecer que aquellos que tengan 75 años o más tienden a tener la misma religión que ego, siendo esta relación estadísticamente significativa al 95% ($p<0.05$). Finalmente se puede establecer que los alteris que son hombres tienden a ser menos homofílicos por religión en comparación con las mujeres, siendo esta relación estadísticamente significativa al 95% de confianza ($p<0.05$).

A final de la tabla se puede observar el N por el total de observaciones correspondiente a los alteris y de grupos correspondiente a los egos, $\sigma_{\mu 0}^2$ refiere a la varianza asociada a los egos (o varianza de nivel dos), y finalmente se presenta los estadísticos de ajuste, el logaritmo de Verosimilitud, el AIC, y el BIC.

Discusión y Conclusión

Este capítulo tiene como objetivo el introducir a los lectores a la aplicación de los modelos multinivel para ego-redes en el contexto de encuestas de opinión pública utilizando las baterías de generador de nombres. Proponemos que estos modelos tienen un gran potencial para contestar preguntas de investigación en las redes sociales, al poder distinguir entre los niveles de egos y alteris, y cuantificar la proporción de variabilidad que es atribuible a diferencias sistemáticas entre los egos o los encuestados y medir el impacto de los egos y las características de estos en las relaciones.

Uno de los aspectos que no pudimos ver con detenimiento en este capítulo tiene que ver con las interacciones entre niveles de las variables de estos modelos. En los modelos multinivel (o jerárquicos) permiten tanto al intercepto como a las pendientes variar entre egos. Al aleatorizar la pendiente de variables de nivel 1 en alteris, es posible dar cuenta por otro lado de interacciones de nivel cruzado entre características del ego (nivel 2) y las características del alteris o su relación con ego (nivel 1). Así y tomando el ejemplo que se presentó en este capítulo uno podría pensar que las relaciones y la solidaridad intergeneracional es distinta por la religión del encuestado y la edad que tienen los alteris, es decir la región de los egos podría moderar de manera diferente la homofilia que se pueda manifestar en relación con las distintas edades que puedan tener los alteris

Por otro lado, y con relación a lo presentado en la Figura 4, los investigadores podrían extender el marco analítico de las regresiones jerárquicas a tres niveles considerando, alteris (nivel 1) anidados en egos (nivel 2), que a su vez estén anidados en diferentes escuelas, vecindarios, ciudades o países. Por ejemplo, Volker y Flap (2007) incluyen dentro de los egos encuestados la información de 72 municipalidades a un tercer nivel de observación, lo que les da la oportunidad de desarrollar y probar complejas teorías multinivel sobre las redes personales

Para finalizar, cabe a destacar que es posible generar una serie de indicadores relacionados a la estructura de la red, como por ejemplo la densidad, la transitividad y la centralización de las redes personales, al establecer el en conjunto con el generador de nombres una matriz alteris-alteris, que permite establecer relaciones entre los alteris mencionados (Burt, 1995; Gould y Fernandez, 1989). A nivel latinoamericano, entre las investigaciones que han aplicado una matriz alteri-alteri al generador de nombres se destacan las lideradas por Ignacio Ramos-Vidal, quien por un lado ha estudiado el rol de las redes personales para caracterizar el nivel de cohesión social en el contexto de localidades y comunidades fuertemente afectados por la violencia política en Colombia (Ramos-Vidal, Villamil y Uribe, 2019), y por otro lado analizar las redes de confianza y discusión política en organizaciones sociopolíticas (Ramos-Vidal, Castro y Palacio, 2016)

Bibliografía

- Adida, C. L., Laitin, D. D., & Valfort, M.-A. (2015). Religious Homophily in a Secular Country: Evidence from a Voting Game in France. *Economic Inquiry*, 53(2), 1187–1206. <https://doi.org/10.1111/ecin.12192>
- Bargsted Valdés, M. A., Espinoza, V., & Plaza, A. (2020). Pautas de Homofilia en Chile. *Papers. Revista de Sociología*, 105(4), 583. <https://doi.org/10.5565/rev/papers.2617>
- Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity*, 53(2), 1051–1074. <https://doi.org/10.1007/s11135-018-0802-x>
- Blau, P. M. (1993). Multilevel structural analysis. *Social Networks*, 15(2), 201–215.
- Browne, W., & Goldstein, H. (2010). MCMC Sampling for a Multilevel Model With Nonindependent Residuals Within and Between Cluster Units. *Journal of Educational and Behavioral Statistics*, 35(4), 453–473. <https://doi.org/10.3102/1076998609359788>
- Browne, W., Goldstein, H., & Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling: An International Journal*, 1(2), 103–124. <https://doi.org/10.1177/1471082X0100100202>
- Burt, R. S. (1982). *Toward a structural theory of action: Network models of social structure, perception, and action*. Academic Press.
- Burt, R. S. (1984). Network items and the general social survey. *Social Networks*, 6(4), 293–339.
- Burt, R. S. (1995). *Structural Holes*. Harvard University Press.
- Burt, R. S. (2000). The network structure of social capital. *Research in Organizational Behavior*, 22, 345–423.
- Crossley, N. (2011). *Towards relational sociology*. Routledge.
- Crossley, N., Bellotti, E., Edwards, G., Everett, M. G., Koskinen, J., & Tranmer, M. (2015). *Social network analysis for ego-nets*. SAGE Publications Ltd.
- DiPrete, T. A., Gelman, A., McCormick, T., Teitler, J., & Zheng, T. (2011). Segregation in Social Networks Based on Acquaintanceship and Trust. *American Journal of Sociology*, 116(4), 1234–1283. <https://doi.org/10.1086/659100>
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121.
- Erickson, B. (1988). The relational basis of attitudes. In *Social structures: A network approach* (NY, pp. 99–121). Cambridge University Press.
- Gould, R. V., & Fernandez, R. M. (1989). Structures of Mediation: A Formal Approach to Brokerage in Transaction Networks. *Sociological Methodology*, 19, 89. <https://doi.org/10.2307/270949>

- Handcock, M. S., & Gile, K. J. (2010). Modeling social networks from sampled data. *The Annals of Applied Statistics*, 4(1), 5–25. <https://doi.org/10.1214/08-AOAS221>
- Johnson, D. P. (2008). *Contemporary sociological theory: An integrated multi-level approach*. Springer.
- Krivitsky, P. N., & Morris, M. (2017). Inference for social network models from egocentrically sampled data, with application to understanding persistent racial disparities in HIV prevalence in the US. *The Annals of Applied Statistics*, 11(1), 427–455. <https://doi.org/10.1214/16-AOAS1010>
- Lazega, E. (2015). *Multilevel network analysis for the social sciences: Theory, methods and applications*. Springer Berlin Heidelberg.
- Lin, N. (2002). *Social capital: A theory of social structure and action* (1. paperback ed). Cambridge Univ. Press.
- Marin, A., & Hampton, K. N. (2007). Simplifying the Personal Network Name Generator: Alternatives to Traditional Multiple and Single Name Generators. *Field Methods*, 19(2), 163–193. <https://doi.org/10.1177/1525822X06298588>
- Marin, A., & Wellman, B. (2011). Social network analysis: An introduction. *The SAGE Handbook of Social Network Analysis*, 11.
- Marsden, P. V. (1990). Network Data and Measurement. *Annual Review of Sociology*, 16(1), 435–463. <https://doi.org/10.1146/annurev.so.16.080190.002251>
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Morris, M. (1994). Epidemiology and Social Networks: Modeling Structured Diffusion. In S. Wasserman & J. Galaskiewicz, *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences* (pp. 26–52). SAGE Publications, Inc. <https://doi.org/10.4135/9781452243528.n2>
- Perry, B. L., Pescosolido, B. A., & Borgatti, S. P. (2018). *Egocentric network analysis: Foundations, methods, and models*. Cambridge University Press.
- Ramos-Vidal, I.; Castro, B. y Palacio, J. (2016). «Patrones de interacción y grupos de discusión política en las redes personales de jóvenes colombianos». *América Latina Hoy*, 73 (0), 121.
- Ramos-Vidal, I.; Villamil, I. y Uribe, A. (2019). «Underlying Dimensions of Social Cohesion in a Rural Community Affected by Wartime Violence in Colombia». *International Journal of Environmental Research and Public Health*, 16 (2), 195.
- Searle, S.R., Casella, G. and McCulloch, C.E. (1992). *Variance Components* (1 edition.). Hoboken, NJ: Wiley-Interscience.
- Smith, J. A., McPherson, M., & Smith-Lovin, L. (2014). Social Distance in the United States: Sex, Race, Religion, Age, and Education Homophily among Confidants, 1985 to 2004. *American Sociological Review*, 79(3), 432–456. <https://doi.org/10.1177/0003122414531776>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publications.
- Snijders, T., Spreen, M., & Zwaagstra, R. (1995). The use of multilevel modeling for analysing personal networks: Networks of cocaine users in an urban area. *Journal of Quantitative Anthropology*, 5(2), 85–105.
- Vacca, R. (2018). Multilevel models for personal networks: Methods and applications. *Italian Journal of Applied Statistics*, 30(1), 59–97. <https://doi.org/10.26398/IJAS.0030-003>
- Volker, B. y Flap, H. (2007) Sixteen Million Neighbors: A multilevel study of the role of neighbors in the personal network of the Dutch. *Urban Affairs Review*, 43(2), 256–284. [10.1177/1078087407302001](https://doi.org/10.1177/1078087407302001)
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815478>
- Wellman, B., & Frank, K. (2001). Network capital in a multi-level world: Getting support from

- personal communities. *Social Capital: Theory and Research*, 233, e273.
- White, H. C. (2008). *Identity and control: How social formations emerge* (2nd ed). Princeton University Press.