

APÉNDICE

El modelo de Rasch en la medición de la inseguridad alimentaria: conceptos básicos, parámetros y estadísticas.

El modelo de medición Rasch provee un conjunto de herramientas analíticas para estimar lo apropiado de conjuntos de ítems en un cuestionario en la construcción de escalas, para crear una escala a partir de los ítems, y para comparar el funcionamiento del conjunto de ítems en diferentes poblaciones y contextos. Este apéndice presenta los conceptos básicos y las matemáticas subyacentes al modelo de Rasch y describe los parámetros del modelo y los estadísticos de ajuste de los ítems usados comúnmente para evaluar datos de encuestas en seguridad alimentaria. Información más detallada esta disponible en Wright (1977; 1983), Bond y Fox (2001), Fischer y Molenaar (1995); Baker (1992) y Hambleton et al. (1991), y en la página de Internet del laboratorio psicométrico MESA de la Universidad de Chicago en www.rasch.org. Sugerimos empezar revisando las publicaciones de Bond y Fox o de Fischer y Molenaar. Información respecto a las aplicaciones de los métodos de Rasch para el desarrollo y evaluación de escalas de seguridad alimentaria están disponibles en Hamilton et al. (1997a; 1997b), Bickel et al. (2000); Nord y Bickel (2002), y Nord (2002b); Nord (2003).

Conceptos básicos: severidad de los ítems y severidad del hogar

Una característica esencial de las escalas de seguridad alimentaria es que los ítems que las constituyen varían a lo largo de un amplio rango de severidad de la seguridad alimentaria. El nivel de severidad preciso de cada ítem (el "calibre del ítem" o "puntaje del ítem" discutidos más adelante) se estima de manera empírica a partir del patrón general de respuesta a los ítems en la escala por parte de los hogares entrevistados. Sin embargo, el rango de severidad de las condiciones identificadas por cada uno de los ítems es evidente solo de manera intuitiva al examinar el contenido de los ítems.

Por ejemplo, el ítem niños no comieron por todo un día es una manifestación más severa de inseguridad alimentaria que el ítem adultos redujeron el tamaño de

sus comidas o se saltaron comidas, y asimismo este último indica un nivel más severo de inseguridad alimentaria que el ítem esta preocupado que se le termine la comida antes de tener más dinero para comprar más. Estas diferencias en severidad son observadas de dos maneras en patrones de respuesta de hogares encuestados. Primero, los ítems más severos son contestados afirmativamente con menor frecuencia que ítems menos severos. Segundo, un hogar que afirma un ítem de mediana severidad es propenso a afirmar todos los ítems que son menos severos y de manera similar, un hogar que niega un ítem de mediana severidad es propenso a negar todos los ítems que son más severos. Estos patrones de respuesta típicos no son universales, pero son predominantes y entre los hogares que se desvían de estos patrones típicos la extensión de esta desviación tiende a ser leve.

En esencia, el modelo de Rasch (nombrado así en honor al matemático danés Georg Rasch) formaliza este concepto de el orden de severidad de los ítems y provee métodos estadísticos estándar para estimar la severidad relativa de cada ítem y para estimar en que medida los patrones de respuesta observados en una base de datos son consistentes con el concepto del orden de severidad. El modelo de Rasch fue desarrollado en principio en el campo de las pruebas educativas, donde múltiples ítems correctos/incorrectos, de diversa dificultad, son usados para medir el nivel de conocimiento o habilidad de un individuo. De manera más general, el modelo puede ser usado para evaluar la ubicación de un individuo u hogar a lo largo de un continuo— en nuestro caso, un continuo de severidad de inseguridad alimentaria— combinando información de múltiples ítems dicotómicos (si/no) que varían por cada punto en el continuo que cada uno de ellos refleja de manera única. Esto corresponde exactamente al carácter del constructo de medición de inseguridad alimentaria/hambre. No existe un lenguaje común usado para describir el continuo completo de inseguridad alimentaria y hambre. ¿No se pregunta a las personas, "en una escala de uno a diez, que tan inseguro en alimentos es su hogar?" Pero las personas hablan abiertamente de experiencias específicas, tales como quedarse sin dinero para alimentos, o respecto a comportamientos y condiciones resultantes tales como

verse forzados a reducir la calidad y cantidad de alimentos. Por ello, información sobre esas experiencias, comportamientos y condiciones puede ser obtenida a través de preguntas de encuesta bien diseñadas.

Matemáticas del modelo de Rasch

El uso del modelo de Rasch para crear una medición de inseguridad alimentaria y hambre asume que tanto los ítems indicadores que componen la escala y los hogares que responden a los ítems pueden ser ubicados en un mismo continuo subyacente de severidad alimentaria. Además asume que la probabilidad de que un hogar específico conteste de manera afirmativa un ítem específico depende de la diferencia entre el nivel de severidad de ese hogar y la severidad se el ítem en cuestión. El modelo uni-paramétrico de Rasch, el cual es usado para crear la escala de seguridad alimentaria, asume específicamente que el logaritmo de la posibilidad (odds - cociente entre la probabilidad de que un evento suceda y la probabilidad de que no suceda) de que un hogar vaya a afirmar un ítem es proporcional a la diferencia entre el nivel de severidad "verdadero" del hogar y el nivel de severidad "verdadero" del ítem. Por ello, la posibilidad (odds) de que un hogar al nivel de severidad h afirme un ítem al nivel de severidad i es:

$$(1) P_{h,i}/Q_{h,i}=e^{(h-i)}$$

donde P es la probabilidad de afirmar el ítem, Q es $1-P$, es decir, la probabilidad de que el hogar vaya a negar el ítem, y e es la base del logaritmo natural. Resolviendo la ecuación (1) para $P_{h,i}$, la probabilidad de que el hogar vaya a afirmar el ítem puede ser expresada como:

$$(2) P_{h,i} = e^{(h-i)} / (1 + e^{(h-i)})$$

Entonces, la severidad de un ítem puede concebirse como el nivel de severidad de los hogares que están justo al borde de afirmar o negar aquel ítem. La posibilidad (odds) de que un hogar vaya a afirmar un ítem que tiene el mismo nivel de severidad que el hogar es 1, correspondiendo a una probabilidad de 0.5. La posibilidad (odds) de que un hogar vaya a afirmar un ítem con una puntaje de severidad menor en una unidad que el del hogar es e^1 , o alrededor de 2.7, correspondiendo a una probabilidad de 0.73 [o sea, $2.7/(1+2.7)$]. La probabilidad de que un hogar vaya a

afirmar un ítem dos unidades más abajo que su propio nivel de severidad es 0.88, y para un ítem tres unidades más abajo es 0.95.

Metricidad de la escala y discriminación los ítems

Dado que es la diferencia entre el puntaje de severidad del hogar y el del ítem la que determina la probabilidad de una respuesta afirmativa, está claro que esta medida puede ser transformada adicionando una constante a ambos puntajes sin cambiar los atributos de la escala. Es decir, el tamaño de los intervalos en la escala conlleva información significativa, pero el punto cero es arbitrario. El proyecto de medición de la seguridad alimentaria en los Estados Unidos ha adoptado un estándar métrico de computo para la escala de 18 ítems basado en un puntaje de los ítems promedio de 7 para los 18 ítems con el objetivo de obtener puntajes con valores positivos para los puntajes correspondientes a todos los ítems y los hogares (Bickel et al., 2000). Esto resulta en puntajes para los hogares en un rango de aproximadamente 1.5 a 13. Otro métrico comúnmente usado asume el promedio del puntaje de los ítems en cero.

A pesar de que el tamaño de los intervalos en la escala de Rasch es inherentemente significativo, se puede ver afectado por factores tales como el error aleatorio en la medición ("ruido estadístico") de respuestas a los ítems que no forman parte fundamental del constructo de la medición. Por ello para comparar de manera significativa la severidad de los ítems entre dos encuestas es conveniente multiplicar los puntajes de los ítems de una de las escalas por una constante para así equivaler la dispersión de los puntajes de los ítems en ambas escalas. (La dispersión es medida comúnmente a través de la desviación estándar o la desviación absoluta promedio de los puntajes de los ítems). En este caso la comparación es referida como una comparación de severidades relativas de los ítems. Matemáticamente, este ajuste de la escala es equivalente a adecuar el modelo de Rasch a la ecuación (1), con la adición de un parámetro de discriminación, k , como sigue:

$$(3) P_{h,i}/Q_{h,i}=e^{k(h-i)}$$

Para una escala basada en una base de datos dada, el parámetro de discriminación es inversamente

proporcional a la desviación estándar (o cualquier otra medida lineal de dispersión tal como la desviación media absoluta) de los puntajes de los ítems en la escala. Esta relación es usada para estimar que tan bien discriminan los ítems en una encuesta, comparados con un estándar. Si se ajustan datos a examinar al modelo de Rasch con un parámetro de discriminación constreñido a 1, entonces la razón de la desviación estándar de los ítems en esos datos respecto a la desviación estándar de esos mismos ítems en una escala estandarizada permite comparar la discriminación promedio de los ítems en los datos a examinar con su discriminación promedio en la escala estandarizada. Alternativamente, si los datos se ajustan al modelo de Rasch usando una desviación estándar del puntaje de los ítems constreñida a la correspondiente al estándar y el parámetro de discriminación es estimado libremente, entonces la magnitud del parámetro de discriminación estimado corresponde a una medida de la discriminación promedio de los ítems en los datos de prueba comparada con la discriminación promedio en la escala estándar (dado que esta última es 1).

Estimación del modelo de Rasch y puntajes de severidad de los hogares

El software que implementa el modelo de Rasch inicia con una matriz de respuestas de los hogares por los ítems. Métodos de máxima verosimilitud (maximum-likelihood) son usados para estimar de manera simultánea los puntajes de severidad para los hogares y los ítems más consistentes con las respuestas observadas bajo los supuestos de Rasch. Los puntajes para los hogares resultantes representan una medida de intervalo-nivel continua de la extensión de la inseguridad alimentaria o el hambre en el hogar. Estos puntajes son apropiados para análisis de asociación tales como correlaciones y regresiones, con la salvedad de que el puntaje para los hogares que negaron todos los ítems o que fueron filtrados (eliminados) no puede ser estimado por el modelo de Rasch. El puntaje asignado a dichos hogares en algunas bases de datos es puramente arbitrario. El tamaño de el intervalo entre los hogares que negaron todos los ítems y aquellos que afirmaron solamente un ítem no puede determinarse estadísticamente a menos que se hagan supuestos

adicionales respecto a la distribución de la inseguridad alimentaria en la muestra.

Evaluando ítems individuales: estadísticas de ajuste de los ítems

El modelo de Rasch provee la base para los estadísticos de "ajuste" que evalúan que también cumple cada ítem, cada hogar y los datos en general con los supuestos del modelo de medición. Dos estadísticos usados comúnmente para estimar que tan bien las respuestas a los ítems corresponden con los supuestos del modelo de Rasch (o "se ajustan" al modelo) son INFIT y OUTFIT. Luego de que los calibres de los ítems y los puntajes de los hogares han sido estimados, la probabilidad de una respuesta afirmativa en cada celda de la matriz hogar/ítems es calculada. Los estadísticos INFIT y OUTFIT son calculados entonces comparando las respuestas observadas con las respuestas esperadas de manera probabilística en cada celda. INFIT es un estadístico ponderado para cada ítem, de tal manera que es sensible a respuestas dadas por hogares con puntajes de severidad en un rango cercano al nivel de severidad del ítem particular. OUTFIT es sensible a respuestas inesperadas de hogares con severidades mucho más altas o bajas que la del ítem en cuestión – es decir, a respuestas muy improbables (anómalas, atípicas - outliers).

Ambos estadísticos comparan desviaciones de las respuestas observadas con las desviaciones esperadas bajo los supuestos de Rasch, de tal manera que el valor esperado de estos estadísticos es 1. Desviaciones mayores a 1.0 indican ítems que están relacionados menos fuertemente o de manera menos consistente con la condición subyacente (inseguridad alimentaria) medida por el conjunto de ítems. Tal ítem tendrá una porción desproporcionada de las respuestas "fuera de lugar" (es decir, respuestas afirmativas por hogares con niveles de severidad por abajo del correspondiente al ítem o negación por hogares con puntajes de severidad por arriba del nivel de severidad del ítem en cuestión). Valores de INFIT y OUTFIT por abajo de 1.0 indican ítems que están relacionados más fuerte y consistentemente con la inseguridad alimentaria que los ítems promedio.

El modelo uni-paramétrico de Rasch, el cual es usado para crear escalas de seguridad alimentaria,

asume que todos los ítems discriminan con igual certeza, así que los valores de los estadísticos de ajuste (especialmente INFIT) que se alejan de la unidad cuestionan lo apropiado del uso del ítem en cuestión en la escala. Como regla general, para propósitos prácticos de la medición se considera que INFIT en el rango de 0.8 a 1.2 cumplen con el supuesto de discriminación igual. INFIT en un rango de 1.2 a 1.3 pudiesen no distorsionar la medición de manera inaceptable, pero deberían ser mejorados para su uso general (Linacre y Wright, 1994). INFITS abajo de 0.8 indican que un ítem está fuertemente asociado con la condición subyacente (inseguridad alimentaria) medida por el conjunto total de ítems. Incluir tal ítem puede ser aceptable en la práctica, pero la información que este provee es poco valorada en la medición de Rasch, en la cual cada ítem tiene una ponderación igual.

Estándares similares pueden ser aplicados a los estadísticos OUTFIT, pero en la práctica, los OUTFIT son muy sensibles a unas cuantas observaciones inesperadas (es decir, negaciones a ítems menos severos por hogares que afirman la mayoría de los ítems), por lo que aún entre varios miles de hogares pueden elevar el OUTFIT a 10 o 20. Más aún, estadísticos OUTFIT se distorsionan seriamente si los hogares que niegan todas las preguntas más severas son pasados por alto en los restantes ítems más severos. Tales filtros son implementados comúnmente para reducir la carga de preguntas en los hogares que son completamente seguros. Interpretados de manera cuidadosa, los estadísticos OUTFIT pueden ayudar a identificar ítems que presentan problemas de comprensión o que tiene

significados idiosincrásicos en poblaciones pequeñas, pero no hay guías estándar para su evaluación.

Escala de seguridad alimentaria: aplicaciones y estándares

En los Estados Unidos se han recabado datos de seguridad alimentaria anualmente desde 1995 a través de una encuesta representativa a nivel nacional patrocinada por el Departamento de Agricultura (USDA) y conducida por la Oficina del censo de los Estados Unidos (Hamilton et al., 1997a; Bickel et al., 1999; Andrews et al., 2000; Nord et al., 2006). La Escala de Seguridad Alimentaria del Hogar de los Estados Unidos fue desarrollada y probada inicialmente usando datos del Suplemento de Seguridad Alimentaria en la Encuesta Continua de Población (Current Population Survey – CPS) de abril de 1995 (Hamilton et al., 1997b). La escala de seguridad alimentaria ha sido re-estimada en los años subsiguientes y se ha encontrado esencialmente estable (Ohls et al., 2001; Cohen et al., 2002). Los puntajes para los ítems basados en la encuesta de seguridad alimentaria de agosto de 1998 se presentan en la Guía para Medir la Seguridad Alimentaria del Hogar, versión revisada 2000 (Guide to Measuring Household Food Security, Revised 2000; Bickel et al., 2000) y son la base para los métodos descritos en esta guía para asignar puntajes a los hogares en los Estados Unidos, un método que no requiere del software para Rasch. Los puntajes estándar en este documento son también la base de puntajes para hogares en archivos de datos de uso público de la Encuesta Continua de Población iniciando en 1998.