

Simulación numérica para el desarrollo de pruebas diagnósticas objetivas

Antonio Matas Terrón

Documento de trabajo bajo licencia creative commons

Universidad de Málaga

amatas@uma.es

Resumen

En este trabajo se muestra un ejemplo de cómo utilizar la simulación numérica para obtener las tablas de tabulación para evaluar una población con la versión española de la prueba Test Your Memory. Se discute las posibilidades de la simulación numérica cuando no se dispone de una muestra o incluso una población lo suficientemente amplia como para realizar una validación de los instrumentos diagnóstica del instrumento. También se exponen sus posibilidades para analizar la utilidad de otros estudios a partir de resultados parciales.

Palabras clave: simulación numérica, utilidad diagnóstica, instrumentos de diagnóstico, Test Your Memory versión española, R cran.

Introducción

El proceso de validación de instrumentos de diagnóstico, por lo general, requiere de un alto número de datos de una muestra (Argibay, 2009). En la práctica este requisito implica:

- Disponer de una amplia red de investigadores colaboradores que puedan obtener datos representativos.
- Este equipo de de colaboradores debe estar entrenado para administrar los instrumentos eficazmente.
- Además, debe contarse con que existe una población que es suficientemente grande. Esto no sucede, por ejemplo, en el caso de las llamadas enfermedades “raras”.

Por lo general, las dos primeras condiciones requieren que los equipos de validación realicen una inversión económica importante. Sin embargo, la inversión en este tipo de estudios es escasa, situación empeoró con la crisis económica (más información acerca de esta situación en <http://www.cosce.org/> sitio web).

A pesar de ello, los investigadores, los cuidadores de la salud, psicólogos y otros profesionales de la salud necesitan seguir usando este tipo de instrumentos diagnósticos (Pinazo, y Sánchez Martínez, 2006; Muñiz y Fernández-Hermida, 2010). La consecuencia es que los estudios se están realizando con un número limitado de muestra, o sin tener en cuenta la estructura de la población de interés, o incluso, sin hacer referencia a ninguna población (ver por ejemplo, los estudios de revistas como International Journal of

Geriatric Psychiatry). Sin embargo, la utilización de estos instrumentos sin un análisis previo de sus propiedades psicométricas supone un riesgo real de errores diagnósticos.

Resulta curioso que en ocasiones, los autores de los estudios utilizan el teorema del límite central (TLC) como argumento para realizar su estudio con muestras pequeñas. Sin embargo, esta es una mala interpretación TLC (Alvarado y Retamal, 2012). El TLC sólo indica que la distribución muestral de medias de la muestra tiende a conformarse como una distribución normal cuando se cumplen una serie de condiciones (Johnson y Kubly, 1999). Básicamente, estas condiciones son que la población origen de los datos se distribuya normalmente, que el tamaño de la muestra sea lo suficientemente grande (por lo general más de 30 casos), y que la proporción de las categorías estén entre 0,1 y 0,9, preferiblemente con proporciones cercanas a 0,5.

Dado este uso indebido de las justificaciones estadísticas junto con el contexto económico expuesto anteriormente, parece necesario buscar alternativas que permitan analizar las propiedades psicométricas de estos instrumentos con un mínimo de garantías metodológicas. Es en este contexto donde la simulación numérica puede ser útil. El proceso de simulación numérica implica dos grandes fases (Revuelta y Ponsoda, 2003) por un lado la generación de datos y por otro el desarrollo del análisis estadístico:

a) Primera fase (obtener datos simulados) se estructura en los puntos siguientes:

- Se parte de datos reales, asumiendo que son escasos.
- Determinación el tipo de distribución que sustentan los datos recogidos.
- Simulación de datos teniendo en cuenta el anterior de distribución.
- Se obtiene la matriz de datos simulados y se evalúa su calidad.

b) La segunda fase consiste en desarrollar los análisis necesarios según el objetivo del estudio (contraste de hipótesis, análisis psicométricos, elaborando tablas, etc.).

El objetivo del presente estudio es utilizar la simulación numérica para construir tablas de baremación en una prueba diagnóstica de la que se tienen sólo los resultados publicados en un artículo. Un segundo objetivo es mostrar a la simulación numérica como un método adecuado para validar instrumentos objetivos de diagnósticos sobre poblaciones específicas.

Para ilustrar todo el proceso y alcanzar los objetivos anteriores se ha recurrido a un estudio de Muñoz-Neira, Enríquez, Delgado, Brown y Sclachevsky (2014), ya que utilizan una muestra relativamente pequeña aunque siguen un proceso estadístico escrupuloso y eficiente.

Método

Diseño de la investigación y el proceso

El diseño de investigación corresponde a una metodología de simulación numérica junto con un proceso de validación psicométrica, para concluir con la elaboración de tablas de baremación para personas que se correspondan con una población específica. El proceso se resume en tres fases.

- Partir de datos reales se elabora una matriz con un número mayor de datos y se evalúa su calidad.
- Análisis psicométrico de las propiedades del instrumento con los datos simulados.
- Definición de la población de generalización de los datos y elaboración de las tablas de baremación o tablas diagnósticas.

Instrumento y las variables

El Test Your Memory - Versión en español de Muñoz-Neira et al., de 2014 (a partir de ahora TYM-S) es una versión del Test Your Memory original de Brown, Pengas, Dawson, Brown y Clatworthy (2009). La prueba consiste en una hoja de papel de doble cara con espacios que el paciente debe rellenar. La prueba comprende diez tareas que miden once dominios cognitivos: orientación, copiado, conocimiento semántico (memoria retrógrada), cálculo, fluidez verbal, abstracción, denominación, habilidades visoespaciales, memoria anterógrada y función ejecutiva (capacidad para completar la prueba sin ayuda). La prueba no establece un límite de tiempo para ser realizada.

Para el presente estudio las variables de interés fueron: edad, sexo, años de formación reglada (educación y la puntuación del TYM-S. Los datos se tomaron a partir de los estadísticos presentados en el artículo de Muñoz-Neira et al., (opus cit.).

Contexto y participantes: datos originales y datos simulados

Las propiedades psicométricas del TYM-S fueron analizadas por Muñoz-Neira et al., (opus cit.) a partir de una muestra de conveniencia de hispanohablantes reclutados de la Neurología Cognitiva y Unidad de Demencias del Hospital del Salvador (Santiago de Chile, Chile). El total de los participantes fue de 74 personas, divididas en tres grupos (véase el cuadro 1). Hubo 30 participantes sin ningún trastorno preexistente (grupo de control), 30 pacientes con demencia (grupo de demencia), y 14 participantes con trastorno amnésico o multidominio (grupo MCI). Por tanto, el número de personas por grupo fue de treinta o menos. Para obtener más información acerca de los grupos ver artículo original de Muñoz-Neira et al., (opus cit.).

Tabla 1. Las características demográficas y el perfil de la muestra

	Control (n = 30)	Demencia (n = 30)	MCI (n = 14)
Edad:	Media= 71,93 (d.t.=7,06)	Media= 72,80 (d.t.=6,90)	Media= 71,71 (d.t.= 7.16)
Sexo:			
Masculino	n = 15 (50%)	N = 19 (63,33%)	N = 8 (57,10%)
Femenino	n = 15 (50%)	N = 11 (36,77%)	N = 6 (42,90%)
A ñ o s d e educación:	Media= 14,00 (d.t.=4,20)	Media= 12,70 (d.t.=3.15)	Media= 12,79 (d.t.= 4,59)
TYM-S:	Media= 43,93 (d.t.=5,55)	Media= 22,50 (d.t.=11.29)	Media= 36,50 (d.t.= 6,81)

A partir de los resultados de la tabla 1 se construyó la matriz de datos simulados según se indica más abajo.

Técnicas y análisis simulados

Los procedimientos utilizados fueron:

- Establecer de las distribuciones muestrales teniendo en cuenta los estadísticos de la tabla 1.
- Generación de los datos simulados. Hay una gran cantidad de software de estadísticas con el fin de obtener datos simulados. En este caso se utilizó R (R Core Team, 2013).
- Creación de la matriz y la evaluación de la calidad. Uno de los principales problemas en los estudios simulados es que las variables no son realmente aleatorias sino pseudoaleatorio. Esto puede llevar acarreado la aparición de correlaciones entre un valor y el siguiente. En cualquier caso, la valoración de la bondad de los datos se ha realizado teniendo en cuenta dos procedimientos. El primero de ellos fue garantizar que los datos simulados se ajustaban a los datos originales realizando un contraste de hipótesis. Para esta fase del estudio se utilizó el programa Data Desk versión 6.1 (Data Description Inc., 1996) junto con el paquete estadístico R. En segundo lugar, para comprobar realmente la falta de correlación entre los datos de una misma variable, se puede hacer un contraste de hipótesis de los datos x_i con los datos x_{i+j} (siendo i el orden de un dato y j un número natural que establece la separación con el dato inicial). Si apareciesen problemas de correlación se pueden utilizar diversas estrategias. Una de ellas es replicar la simulación tantas veces como el tamaño de la muestra y tomar la diagonal de la matriz resultante. En este proceso, los resultados no estarán correlacionados entre sí.
- La elaboración de tablas de puntuaciones por sexo y edad. Para satisfacer el objetivo principal del estudio, se realizó un análisis descriptivo organizando los

casos por sexo y edad. Para ello se usó el paquete estadístico SPSS (IBM Corp., 2012). Con el fin de ser operativos, la variable edad se categorizó en intervalos de 5 años. Debido a que la máxima puntuación en el TYM-S es de 50 puntos, los datos simulados del TYM-S fueron tomados hasta dicha puntuación. Este proceso se aplicó sólo en la elaboración de las tablas, no antes, puesto que esta circunstancia no afectaba al proceso de generación de los datos simulados.

Resultados

Distribuciones

Tomando en cuenta los estadísticos de la tabla 1, las variables se consideraron como sigue:

- La edad y años de formación reglada (educación) se consideraron normalmente distribuidas.
- Se asumió que el sexo se distribuía según la función de Bernoulli.
- Las puntuaciones del TYM-S se consideraron normalmente distribuidas.

Datos simulados

Los datos simulados se generaron con R (ver códigos de comandos en el anexo I).

Para analizar la posible correlación intra-variables se llevó a cabo el análisis correlacional entre el conjunto de datos x_i y x_{i+1} para cada variable. Para las variables normales se utilizó el coeficiente de Pearson mientras que para la variable sexo se utilizó el coeficiente de Chi-cuadrado. Los resultados mostraron una ausencia de correlación a un nivel de confianza del 95%.

Una vez generados los datos se llevó a cabo un análisis descriptivo de los mismos (tabla 2).

Tabla 2. Las características demográficas y el perfil de la muestra simulada

	Control (n = 1.000)	Demencia (n = 1.000)	MCI (n = 1000)
Edad:	Media= 72,05 (d.t.= 7,28)	Media= 72,92 (d.t.= 7.11)	Media= 71,83 (d.t.= 7,38)
Sexo:			
Masculino	N = 50,6%	N = 64,6%	N = 57%

Femenino	N = 49,4%	N = 35,4%	N = 43%
Años de educación:	Media= 14,07 (d.t.= 4,33)	Media= 12,75 (d.t.= 3,25)	Media= 12,87 (d.t.= 4,73)
TYM-S:	Media= 44,55 (d.t.= 5,73)	Media= 23,51 (d.t.= 11,67)	Media= 37,12 (d.t.= 7,03)

No se registraron diferencias significativas entre las medias de edad y años formación. No obstante, encontraron diferencias significativas entre los grupos del TYM-S (tabla 3).

Tabla 3. Diferencias significativas entre las variables

	Datos originales	Datos simulados	t	g.l.	Significatividad
Grupos control TYM-S	Media= 43,93	Media= 44,55	3,406	999	0,0007
TYM-S Grupos demencia	Media= 22,5	Media= 23,51	2,738	999	0,0063
Grupos TYM-S MCI	Media= 36,5	Media= 37,116	2,769	999	0,0057

Tablas de baremación y población de aplicación

Se elaboraron las tablas de baremación o puntajes para el TYM-S por grupos (control, demencia, MCI) en función del sexo y la edad.

En las tablas se incluyeron medias y percentiles, permitiendo tabular a cualquier usuario en base a su diagnóstico psicológico (normal, demencia o deterioro cognitivo leve), el sexo (masculino, femenino) y la edad (entre 60 y 100 años)(anexo II).

La población de aplicación de estos baremos está definida por los estadísticos de la tabla 2.

Discusión

La simulación numérica, así como la simulación de contextos, se utiliza a menudo en el mundo académico y el ámbito tecnológico (García Dunna, García Reyes, y Cárdenas,

2006). Sin embargo, estos procedimientos suelen aparecer como una parte complementaria de otro análisis, principalmente en estudios de Psicología, Sociología y en Educación (Revuelta y Ponsoda, opus cit.). Sin embargo, la simulación numérica tiene campo para ser desarrollado con otros objetivos. En este sentido, el presente trabajo ofrece un ejemplo de cómo puede ser útil en el campo de la evaluación y el diagnóstico, tomando un papel principal en el mismo, especialmente cuando no hay suficientes datos para realizar algunos análisis estadísticos.

Por otro lado, con relación al TYM-S hay tres cuestiones que han surgido durante el estudio:

- Se han encontrado algunas diferencias entre las distribuciones de los datos simulados y las distribuciones indicadas en el artículo original. Estas diferencias pueden deberse a que la variable original del TYM-S no se distribuye normalmente. En tal caso, deben tenerse en cuenta tanto las consideraciones sobre el número de datos, como las cuestiones sobre CLT que se expusieron al principio. Otro elemento de debate a tener en cuenta es que cuando se tenga en consideración una investigación como referente inicial, debe exigirse una correcta identificación de las escalas de medida utilizadas (proceso de medición adecuado). Si esto no es así, se pueden obtener resultados no interpretables (Matas, 2000).
- Por otro lado, los resultados de los datos simulados para la escala TYM-S, indican que este instrumento es útil, aunque poco sensible a las casos con un alto número de errores.
- Por último, el TYM-S parece ser más útil para evaluar los casos de demencia y casos patológicos que los casos de MCI. Dichos resultados son coherentes con los resultados de Muñoz-Neira et al. (opus cit.).

Conclusiones

La primera conclusión es que la simulación numérica es un proceso técnico que permite elaborar las tablas de baremación a partir de resultados parciales o a partir de las indicaciones de investigaciones previas, siempre que se ofrezcan los estadísticos de los datos originales. De esta forma, la simulación numérica se puede utilizar para resolver algunos problemas en el proceso de desarrollo de nuevos instrumentos de diagnóstico. Sin embargo, es necesario advertir que los resultados no se pueden generalizar directamente. En este sentido, los resultados de un estudio de simulación deben ser considerados como orientación para trabajos posteriores. En cualquier caso, cuando no es posible realizar estudios con grandes muestras, y no hay otra posibilidad, los estudios simulados son un buen punto de partida.

Referencias bibliográficas

Alvarado, H. & Retamal, L. (2012). Dificultades de comprensión del teorema central del límite en estudiantes universitarios. *Educación Matemática*, 24(3), 151-171.

- Argibay, J.C. (2009). Muestra en investigación cuantitativa. *Subjetividad y procesos cognitivos*, 13(1), 13-21.
- Brown, J., Pengas, G., Dawson, K., Brown, L.A. & Clatworthy, P. (2009). Self administered cognitive screening test (TYM) for detection of Alzheimer's disease: cross sectional study. *British Medical Journal*, 338, 1423-1431. DOI: <http://dx.doi.org/10.1136/bmj.b2030>
- Data Description Inc. Released 1996. Data Desk 6.1 for Windows. Ithaca, NY: Data Description Inc.
- García Dunna, E., García Reyes, H., & Cárdenas, L.E. (2006). *Simulación y análisis de sistemas con ProModel*. México: Pearson Educación.
- IBM Corp. Released 2012. IBM SPSS Statistics for Mac, Version 19.0. Armonk, NY: IBM Corp.
- Johnson, R., & Kubly, P (1999). *Estadística elemental: lo esencial*. México: International Thomson Editores.
- Matas, A. (2000). Análisis cuantitativo de la información. En A. Matas, E. Jiménez, J. Gil, and G. Rodríguez (Auts.), *Análisis de datos I*. Sevilla: Kronos.
- Muñiz, J. & Fernández-Hermida, J.R. (2010) La Opinión de los Psicólogos Españoles sobre el Uso de los Tests. *Papeles del Psicólogo*, 31(1), 108-122.
- Muñoz-Neira, C., Henríquez, F., Delgado, C., Brown, J., & Slachevsky, A. (2014). Test Your Memory – Spanish version (TYM-S): a validation study of a self-administered cognitive screening test. *International Journal of Geriatric Psychology*, early version on line. DOI: 10.1002/gps.4055
- Pinazo, S. & Sánchez Martínez, M. (dirs.) (2006). *Gerontología. Actualización, innovación y propuestas*. Madrid: Prentice-Hall (Pearson Educacion).
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Revuelta, J., & Ponsoda, V. (2003). *Simulación de modelos estadísticos en Ciencias Sociales*. Salamanca: La Muralla.

Anexo I

Variable edad:

```
> # Age Control group
> set.seed(100)
> age_control<-rnorm(1000, 71.93, 7.06)
> # Age Dementia group
> set.seed(100)
> age_dementia<-rnorm(1000, 72.8, 6.9)
> # Age mci group
> set.seed(100)
> age_mci<-rnorm(1000, 71.71, 7.16)
```

Variable años de formación:

```
> # Year Control group
> set.seed(100)
> year_control<-rnorm(1000, 14, 4.2)
> # Year Dementia group
> set.seed(100)
> year_dementia<-rnorm(1000, 12.7, 3.15)
> # Year MCI group
> set.seed(100)
> year_mci<-rnorm(1000, 12.79, 4.59)
```

Sexo:

```
> # Gender (getting male proportion like reference)
> gender_control<-rbinom(1000, 1, 0.5)
> gender_dementia<-rbinom(1000, 1, 0.633)
> gender_mci<-rbinom(1000, 1, 0.571)
```

Variable TYM-S:

```
> # TYMS Control group
> set.seed(100)
> TYMS_control<-rnorm(1000, 43.93, 5.55)
> # TYMS Dementia group
```

```
> set.seed(100)
> TYMS_dementia<-rnorm(1000, 22.5, 11.9)
> # TYMS MCI group
> set.seed(100)
> TYMS_mci<-rnorm(1000, 36.5, 6.81)
> # getting absolute score from round data. Others processes are possible, directly abs() for
example.
> TYMS_control<-abs(ceiling(TYMS_control))
> TYMS_dementia<-abs(ceiling(TYMS_dementia))
> TYMS_mci<-abs(ceiling(TYMS_mci))
```

- Elaboración de la matriz de datos y evaluación de su bondad de ajuste.

```
> datos<-cbind(age_control, age_dementia, age_mci, year_control, year_dementia, year_mci,
gender_control, gender_dementia, gender_mci, TYMS_control, TYMS_dementia, TYMS_mci)
> names(datos)<-c("age control", "age dementia", "age mci", "year control", "year dementia", "year
mci", "gender control", "gender dementia", "gender mci", "TYM-S control", "TYM-S dementia", "TYM-
S mci")
> write.table(datos, "datos.txt", sep="\t")
```

Anexo II

Grupo Control

Mujer

	Mean	Median	sd	Percentiles						
				5	10	25	50	75	90	95
<60	32,11	33	2,359	26,8	28	31	33	34	35	35
60-65	37,27	37	1,119	35,15	36	36	37	38	39	39
65-70	41,37	42	1,111	39	40	41	42	42	43	43
70-75	44,86	45	1,255	43	43	44	45	46	47	47
75-80	48,64	49	1,043	47	47	48	49	50	50	50
80-85	50	50	0	50	50	50	50	50	50	50
85-90	50	50	0	50	50	50	50	50	50	50
90-95	50	50	0	50	50	50	50	50	50	50
95-100	50	50	0	50	50	50	50	50	50	50

Hombre

	Mean	Median	sd	Percentiles						
				5	10	25	50	75	90	95
<60	33,33	34	1,544	30	30,8	32	34	35	35	35
60-65	37,67	38	1,216	35	36	37	38	39	39	39
65-70	41,09	41	1,278	39	39	40	41	42	43	43
70-75	44,76	45	1,158	43	43	44	45	46	46	47
75-80	48,6	49	1,056	47	47	48	49	50	50	50
80-85	50	50	0	50	50	50	50	50	50	50
85-90	50	50	0	50	50	50	50	50	50	50
90-95	50	50	0	50	50	50	50	50	50	50
95-100	50	50	0	50	50	50	50	50	50	50

Grupo Demencia

Mujer

	Mean	Median	sd	Percentiles						
				5	10	25	50	75	90	95
<60	4,94	4	4,139	0	0,7	1	4	8,75	11,6	.
60-65	5,69	6	2,505	2	2	3,25	6	8	9	9
65-70	14,03	14	2,714	10	10,9	11,75	14	16,25	18	18
70-75	22,7	22	2,481	19	19	21	22	25	26	26
75-80	30,91	31	2,316	27	28	29	31	33	34	35
80-85	39,86	40	2,263	36	37	38	40	41,5	43	44
85-90	47,22	46	2,108	44	44	46	46	49,5	50	50
90-95	50	50	0	50	50	50	50	50	50	50
95-100	50	50	0	50	50	50	50	50	50	50

Hombre

	Mean	Median	sd	Percentiles						
				5	10	25	50	75	90	95
<60	3,2	2	3,916	0	0	0	2	5,5	7,8	14,6
60-65	6,35	7	2,372	1,75	2,5	5	7	8	9	9
65-70	14,54	15	2,619	10	11	12	15	17	18	18
70-75	22,46	23	2,619	19	19	20	23	25	26	26
75-80	30,8	31	2,535	27	28	28	31	33	35	35
80-85	38,86	38,5	2,412	36	36	37	38,5	41	43	43
85-90	47,76	48	2,278	44	44,6	45	48	50	50	50
90-95	50	50	0	50	50	50	50	50	50	50
95-100	50	50	0	50	50	50	50	50	50	50

Grupo MCI

Mujer

	Mean	Median	sd	Percentiles						
				5	10	25	50	75	90	95
<60	22,17	23	2,526	16	18,7	20	23	24	25	.
60-65	28,66	29	1,356	26	27	27	29	30	30	30
65-70	33,16	33	1,398	31	31	32	33	34	35	35
70-75	37,95	38	1,457	36	36	37	38	39	40	40
75-80	42,21	42	1,362	40	41	41	42	43	44	44
80-85	46,78	47	1,298	45	45	46	47	48	49	49
85-90	50	50	0	50	50	50	50	50	Male50	50
90-95	50	50	0	50	50	50	50	50	50	50
95-100	50	50	0	50	50	50	50	50	50	50

Hombre

	Mean	Median	sd	Percentiles						
				5	10	25	50	75	90	95
<60	23,05	24	2,607	17,2	19,2	22	24	25	25,8	26
60-65	28,63	29	1,332	26	27	28	29	30	30	30,45
65-70	33,42	34	1,338	31	31	32	34	35	35	35
70-75	37,85	38	1,418	36	36	37	38	39	40	40
75-80	42,5	42	1,306	40,4	41	42	42	43	44	45
80-85	46,77	46	1,276	45	45	46	46	48	49	49
85-90	50	50	0	50	50	50	50	50	50	50
90-95	50	50	0	50	50	50	50	50	50	50
95-100	50	50	0	50	50	50	50	50	50	50