

UNA TÉCNICA ESTADÍSTICA PARA ESTIMAR LA FRECUENCIA DE COPIA EN EXÁMENES

Beatriz Cobo Rodríguez, *Universidad de Granada*, beacr@correo.ugr.es
David Molina Muñoz, *Universidad de Granada*, dmolinam@ugr.es

RESUMEN.

En las encuestas psicológicas y sociales, el interés frecuentemente se centra en aspectos sensibles o confidenciales para las personas entrevistadas, como ejemplo podemos citar uso de drogas, evasión de impuestos, preferencias sexuales, honestidad en exámenes... Por ello muchos entrevistados rehúsan a participar en la encuesta o proporcionan respuestas falsas o respuestas condicionadas, ocasionando que la precisión y confiabilidad de los estimadores se alteren de una manera importante. La técnica de respuesta aleatoria se propone como solución a estos problemas.

Nivel educativo: Educación superior.

1. INTRODUCCIÓN.

Una encuesta es un procedimiento de investigación que se basa en interrogar a una muestra de individuos. En estudios de encuestas por muestreo el interés frecuentemente se centra en aspectos sensibles o confidenciales para las personas entrevistadas. Debido a todo esto el problema típico que surge consiste en la deseabilidad social, la cual se define como la tendencia de las personas encuestadas a responder en función de lo que es aceptable socialmente. Por lo tanto, en la búsqueda de respuestas seguras, los encuestadores se enfrentan a varias dificultades, las cuales son inherentes a toda persona encuestada: inclinaciones, tendencias, actitudes, formas de pensar y comportarnos, tiempo disponible, confianza y desconfianza, propensión a mantener la intimidad, etc.

Existe una amplia evidencia empírica de que los encuestados informan sistemáticamente sobre los comportamientos y actitudes socialmente deseables y no informan de los socialmente indeseables (Krumpal, 2013).

Para que los resultados obtenidos sean creíbles es necesario, entre otros aspectos, que el modo de encuestación tenga suficiente calidad o validez, lo que exige asumir que las respuestas de los individuos son ciertas además de otros requisitos.

La técnica de Respuestas Aleatoria (RA) introducida por Warner (Warner, 1965) es una posible solución para la protección del anonimato del entrevistado y es introducida para reducir el riesgo de evasión o no respuesta de preguntas sensitivas. Consiste en la utilización de un mecanismo aleatorio por medio del cual se selecciona una de dos preguntas complementarias: ¿pertenece al grupo con la característica A? o ¿pertenece al grupo que no tiene la característica A?, donde A es la característica sensible de interés. El entrevistado contestará sí o no y el entrevistador no tiene la posibilidad de saber qué pregunta contestó el entrevistado, protegiendo así la confidencialidad del mismo.



La afirmación más importante hecha de la RA es que produce estimaciones puntuales de comportamientos sensibles más válidas: ha habido muchos informes de que la RA logra estimaciones más precisas de la prevalencia de comportamientos socialmente indeseables que cuando se hacen preguntas sensibles directamente (Coutts and Jann, 2011). Sin embargo, el uso de RA incurre en costos adicionales, y la ventaja de usar RA, es decir, la mayor precisión de las estimaciones de población obtenidas, será solo superar estos costes si las estimaciones son sustancialmente mejores que las derivadas de los diseños de preguntas de respuestas directas (Lensvelt-Mulders et al. 2005).

2. EJEMPLO DE LA TÉCNICA DE RESPUESTA ALEATORIA USANDO EL MECANISMO B DE CHAUDHURI.

En primer lugar definimos el mecanismo B de Chaudhuri (Chaudhuri, 2011, p.100) de la siguiente manera:

Se pide a cada encuestado que seleccione una carta de una caja que contiene una proporción C de cartas marcadas correctamente y más cartas que tienen los números x_1, \dots, x_T en las proporciones q_1, \dots, q_T respectivamente. La persona informa del verdadero valor y_i si obtiene una carta codificada correctamente o el número x_j si obtiene una carta con el número x_j .

Ilustramos la utilidad de la respuesta aleatoria en el contexto de copiar en exámenes por parte de los estudiantes. El engaño académico de los estudiantes de colegios y universidades ha sido siempre una preocupación de los investigadores en la educación superior. Algunos estudios previos han demostrado que las estimaciones basadas en cuestionarios anónimos pueden informar menos severamente los niveles reales de copiar en exámenes. Scheers y Dayton (1987) encontraron que el registro de cinco formas de engaño académico varió de 39% a 83% cuando las respuestas a un cuestionario anónimo se compararon con las estimaciones obtenidas utilizando una técnica de respuesta aleatoria.

Para investigar el comportamiento de copiar en exámenes, se realizó una encuesta entre los estudiantes de primer año en nuestra universidad. Se extrajo una muestra estratificada de 11 grupos tal que estuvieron representadas las diferentes áreas de conocimiento, con afijación proporcional al número total de estudiantes. Los grupos de la muestra se dividieron al azar y en forma de pareja entre un grupo de cuestionario anónimo y un grupo de respuesta aleatoria. Todos los cuestionarios fueron administrados durante las horas de clases regulares. A los estudiantes se les aseguró el anonimato al hacer la encuesta. En este estudio, se utilizó la técnica B de Chaudhuri (2011). Se le proporcionó a cada uno de los estudiantes 10 cartas de juego para el proceso de aleatorización, y el profesor explicó cómo la técnica conserva su anonimato. Los estudiantes recibieron las siguientes instrucciones:

- contestar 0 cuando la carta seleccionada sea el número 6 ($q_1 = 0.1$)
- contestar 1 cuando la carta seleccionada sea el número 7 ($q_2 = 0.1$)
- contestar 3 cuando la carta seleccionada sea el número 10 ($q_3 = 0.1$)
- contestar 5 cuando la carta seleccionada sea el número 11 ($q_4 = 0.1$)
- contestar 8 cuando la carta seleccionada sea el número 12 ($q_5 = 0.1$)
- en otro caso, contestar la pregunta sensible verazmente ($p_1 = 0.5$)

Pregunta: ¿Cuántas veces te copiaste en un examen el año pasado?

				
0	1	2	3	5

Figura 1. Formulario de la pregunta de respuesta aleatoria

El proceso de respuesta aleatorio (mecanismo B de Chaudhuri) no provocó desconfianza y todos los estudiantes completaron con éxito los cuestionarios. Para los estudiantes que respondieron el cuestionario de forma directa, el porcentaje de no respuesta fue 5.5%.

- Cuestionario directo (muestra control): el número promedio estimado de veces que el estudiante había copiado en el último año osciló entre 0.286-1.652 (nivel de confianza del 95%).
- Cuestionario con la técnica de respuesta aleatoria (muestra aleatoria): el número promedio estimado de veces que el estudiante había copiado en el último año osciló entre 2.646-6.1904 (nivel de confianza del 95%).

Se observó una grave subestimación de la frecuencia de engaño cuando se utilizó el cuestionario directo, mientras que las estimaciones obtenidas utilizando la técnica de respuesta aleatoria eran mucho más altas. Por lo tanto, la copia en exámenes es un tema muy delicado y hay una fuerte tendencia de los estudiantes a minimizar la participación en esta práctica.

Tras ver el ejemplo podemos concluir que el mecanismo de respuesta aleatoria presenta tanto ventajas como inconvenientes, entre las ventajas podemos destacar

- Aumenta la probabilidad de contestar la verdad respecto a una pregunta directa.
- Mayor índice de respuesta.

Y entre las desventajas:

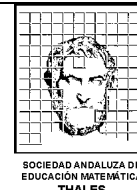
- Aumento en la complejidad de la pregunta.
- Dificultad en entender el método de aleatorización.
- Requiere de muestras de tamaños grandes.

REFERENCIAS.

- ARNAB, R.(2002).*Optimum sampling strategies under randomized response surveys*. Biometrical Journal 44, 1-16
- ARNAB, R. (2004).*Optional randomized response techniques for complex survey designs*. Biometrical Journal 46, 114-124
- BOUZA, C.(2009).*Ranked set sampling and randomized response procedures for estimating the mean of a sensitive quantitative character*. Metrika 70, 267-277
- BOUZA, C. HERRERA, C., MITRA, P. (2010). *A review of randomized responses procedures: the qualitative variable case*. Revista Investigación Operacional 31, 240-247
- CHAUDHURI, A. (2011).*Randomized response and indirect questioning techniques in surveys*. Chapman & Hall



XV CONGRESO DE ENSEÑANZA Y APRENDIZAJE DE LAS
MATEMÁTICAS: EL SENTIDO DE LAS MATEMÁTICAS.
MATEMÁTICAS CON SENTIDO



- CHAUDHURI, A., MUKERJEE, R. (1988). *Randomized response: theory and technique*. Marcel Dekker
- COUTTS, E., JANN, B.(2011). *Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT)*. *Sociological Methods & Research* 40, 169-193
- GREENBERG, B.G., KUEBLER, R.R., ABERNATHY, J.R., HORVITZ, D.G.(1971). *Application of the randomized response technique in obtaining quantitative data*. *Journal of the American Statistical Association* 66, 243-250
- KRUMPAL, I.(2013). *Determinants of social desirability bias in sensitive surveys: a literature review*. *Quality and Quantity* 47,2025-2047
- LENSVELT-MULDERS, G. J. L. M., HOX, J.J., VAN DER HEIJDEN, P.G.M, MAAS, C.J.M.(2005). *Meta-Analysis of Randomized Response Research: Thirty-Five Years of Validation*. *Sociological Methods & Research* 33, 319-348
- SCHEERS, N. J., DAYTON, C. M.(1987). *Improved estimation of academic cheating behavior using the randomized response technique*. *Research in Higher Education* 26, 61-69
- SINGH, S., KIM, J.M.(2007). *A pseudo-empirical log-likelihood estimator using scrambled responses*. *Statist. Probab. Lett.* 81, 345351
- WARNER, S. L.(1965). *Randomized response: a survey technique for eliminating evasive answer bias*. *J. Amer. Statist. Assoc.* 60, 63-69



Una técnica estadística para determinar la frecuencia de copia en exámenes

Beatriz Cobo Rodríguez^{1*} y David Molina Muñoz¹

¹ Departamento de Estadística e Investigación Operativa, Universidad de Granada

*Contacto: beacr@correo.ugr.es

Introducción

Una encuesta es un procedimiento de investigación que se basa en interrogar a una muestra de individuos.

En estudios de encuestas por muestreo el interés frecuentemente se centra en aspectos sensibles o confidenciales para las personas entrevistadas. Debido a todo esto el problema típico que surge consiste en la desabilidad social, la cual se define como la tendencia de las personas encuestadas a responder en función de lo que es aceptable socialmente. Por ello muchas entrevistadas rechazan a participar en la encuesta o proporcionan respuestas falsas o respuestas condicionadas, ocasionando que la precisión y confiabilidad de los estimadores se alteren de una manera importante.

La técnica de Respuestas Aleatorias (RA) introducida por Warner (Warner, 1965) es una posible solución para la protección del anonimato del entrevistado y es introducida para reducir el riesgo de evasión o no respuesta de preguntas sensibles.

Objetivo

Ilustramos la utilidad de la respuesta aleatoria en el contexto de copiar en exámenes por parte de los estudiantes.

El engaño académico de los estudiantes de colegios y universidades ha sido siempre una preocupación de los investigadores en la educación superior.

Algunos estudios previos han demostrado que las estimaciones basadas en cuestionarios anónimos pueden informar menos severamente los niveles reales de copiar en exámenes. Scheers y Dayton (1987) encontraron que el registro de cinco formas de engaño académico varió de 39% a 83% cuando las respuestas a un cuestionario anónimo se compararon con las estimaciones obtenidas utilizando una técnica de respuesta aleatoria.

Metodología

Para investigar el comportamiento de copiar en exámenes, se realizó una encuesta entre los estudiantes de primer año en nuestra universidad.

Se extrajo una muestra estratificada de 11 grupos tal que estuvieron representadas las diferentes áreas de conocimiento, con fijación proporcional al número total de estudiantes.

Los grupos de la muestra se dividieron al azar y en forma de pareja entre un grupo de cuestionario anónimo y un grupo de respuesta aleatoria.

Todos los cuestionarios fueron administrados durante las horas de clases regulares.

A los estudiantes se les aseguró el anonimato al hacer la encuesta.

En este estudio, se utilizó la técnica B de Chaudhuri (2011).

Técnica B de Chaudhuri (Chaudhuri, 2011, p. 100):

Se pide a cada encuestado que seleccione una carta de una caja que contiene:

- una proporción C de cartas marcadas correctamente
- Otras cartas que tienen los números x_1, \dots, x_r en las proporciones q_1, \dots, q_r . La persona informa

• el verdadero valor y , si obtiene una carta codificada correctamente

• el número x_i si obtiene una carta con el número x_i .

Se le proporcionó a cada uno de los estudiantes 10 cartas de juego para el proceso de aleatorización, y el profesor explicó cómo la técnica conserva su anonimato.

Los estudiantes recibieron las siguientes instrucciones:

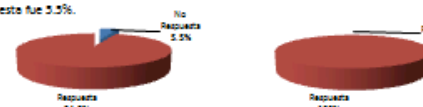
- contestar 0 cuando la carta seleccionada sea el número 6 ($q_1 = 0.1$)
- contestar 1 cuando la carta seleccionada sea el número 7 ($q_2 = 0.1$)
- contestar 2 cuando la carta seleccionada sea el número 10 ($q_3 = 0.1$)
- contestar 3 cuando la carta seleccionada sea el número 11 ($q_4 = 0.1$)
- contestar 5 cuando la carta seleccionada sea el número 12 ($q_5 = 0.1$)
- en otro caso, contestar la pregunta sensible verazmente ($p_1 = 0.5$)



Resultados

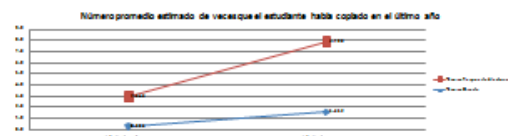
El proceso de respuesta aleatoria (mecanismo B de Chaudhuri) no provocó desconfianza y todos los estudiantes completaron con éxito los cuestionarios.

Para los estudiantes que respondieron el cuestionario de forma directa, el porcentaje de no respuesta fue 5.5%.



De los cuestionarios podemos obtener los siguientes resultados:

- Cuestionario directo (muestra control): el número promedio estimado de veces que el estudiante había copiado en el último año osciló entre 0.286-1.652 (nivel de confianza del 95%).
- Cuestionario con la técnica de respuesta aleatoria (muestra aleatoria): el número promedio estimado de veces que el estudiante había copiado en el último año osciló entre 2.646-6.1904 (nivel de confianza del 95%).



Se observó una grave subestimación de la frecuencia de engaño cuando se utilizó el cuestionario directo, mientras que las estimaciones obtenidas utilizando la técnica de respuesta aleatoria eran mucho más altas.

Por lo tanto, la copia en exámenes es un tema muy delicado y hay una fuerte tendencia de los estudiantes a minimizar la participación en esta práctica.

Conclusiones

Ventajas	Desventajas
<ul style="list-style-type: none"> • Aumenta la probabilidad de contestar la verdad respecto a una pregunta directa. • Mayor índice de respuesta. 	<ul style="list-style-type: none"> • Aumento en la complejidad de la pregunta. • Dificultad en entender el método de aleatorización. • Requiere de muestras de tamaños grandes.

Referencias

- ANNAS, K. (2002). Optimum sampling strategies under randomized response surveys. *Biometrical Journal* 44, 1-16
- SOUZA, C., HERRERA, C., MITRA, P. (2010). A review of randomized response procedures: the qualitative variable case. *Revista Investigación Operacional* 31, 240-247
- CHAUDHURI, A. (2011). *Randomized response and indirect questioning techniques in surveys*. Chapman & Hall
- GREENBERG, S.O., KUEBLER, R.K., ASERINATHY, J.K., HORVITZ, D.O. (1971). Application of the randomized response technique in obtaining quantitative data. *Journal of the American Statistical Association* 66, 245-250
- SCHEERS, N. J., DAYTON, C. M. (1987). Improved estimation of academic cheating behavior using the randomized response technique. *Research in Higher Education* 16, 61-69
- SINGH, S., KIM, J.M. (2007). A pseudo-empirical log-likelihood estimator using scrambled responses. *Statist. Probab. Lett.* 81, 345-351
- WARNER, S. L. (1965). Randomized response: a survey technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.* 60, 63-69