



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ciencias Matemáticas

Escuela Académico Profesional de Estadística

**El modelo de respuesta no aleatorizada transversal
para estudiar el comportamiento de los estudiantes de
la FCM frente a preguntas sensibles**

TESIS

Para optar el Título Profesional de Licenciado en Estadística

AUTOR

Jhon Wilver GALVÁN TAPIA

ASESOR

Olga Lidia SOLANO DÁVILA

Lima, Perú

2016



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Galván, J. (2016). *El modelo de respuesta no aleatorizada transversal para estudiar el comportamiento de los estudiantes de la FCM frente a preguntas sensibles*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Académico Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.

1043



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
(Universidad del Perú, DECANA DE AMÉRICA)
FACULTAD DE CIENCIAS MATEMÁTICAS



ESCUELA ACADÉMICO PROFESIONAL DE ESTADÍSTICA

ACTA DE SUSTENTACIÓN DE TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE LICENCIADO EN ESTADÍSTICA

13/2
94

En la Ciudad Universitaria, Facultad de Ciencias Matemáticas, siendo las 16:10. horas del día 16. de junio del año 2016 se reunieron los docentes designados como miembros del Jurado:

- Mg. Ana María Cárdenas Rojas (Presidente)
- Mg. Rosa Ysabel Adriazola Cruz (Miembro)
- Mg. Olga Lidia Solano Dávila (Miembro Asesor)

Para la sustentación de la Tesis intitulada "EL MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL PARA ESTUDIAR EL COMPORTAMIENTO DE LOS ESTUDIANTES DE LA FCM FRENTE A PREGUNTAS SENSIBLES", presentada por el Bachiller Jhon Wilver Galvan Tapia, para obtener el Título Profesional de Licenciado en Estadística.

Luego de la exposición de la Tesis, la Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los miembros del jurado, el expositor mereció la aprobación... *sobresaliente*... con un calificativo promedio de *dieciocho (18)*.....(letras y números).

A continuación los miembros del jurado, dan manifiesto que el participante Bachiller Jhon Wilver Galvan Tapia, en virtud de haber aprobado la sustentación de su tesis, será propuesto para que se le otorgue el Título Profesional de Licenciado en Estadística.

Siendo las 16:55. horas, se levantó la Sesión, firmando para constancia la presente Acta en tres (3) copias originales.

Rosa Ysabel Adriazola Cruz
Mg. Rosa Ysabel Adriazola Cruz
Miembro

Olga Lidia Solano Dávila
Mg. Olga Lidia Solano Dávila
Miembro Asesor

Ana Cárdenas Rojas
Mg. Ana María Cárdenas Rojas
Presidente

FICHA CATALOGRÁFICA

GALVAN TAPIA JHON WILVER

El Modelo de Respuesta No Aleatorizada Transversal para estudiar el comportamiento de los estudiantes de la FCM frente a preguntas sensibles (Lima-2016).

xiii, 94 p, 29.7 cm (UNMSM, Licenciado, Estadística,2016)

Facultad de Ciencias Matemáticas. Estadística.

I. UNMSM/FACULTAD DE CIENCIAS MATEMÁTICAS

DEDICATORIA

A Dios porque siempre estuvo conmigo,

a mi madre,

a todos mis seres queridos,

y a todos mis amigos.

AGRADECIMIENTO

A DIOS porque me ha dado la vida, mi familia, amigos y todo lo que tengo.

A MIS PADRES por su gran apoyo que me dieron en todos estos años de estudios universitarios para ser lo que ellos quisieron que sea un verdadero profesional.

A MI FAMILIA que con sus experiencias y motivación me inclinaron a ser lo que ahora soy, muchas gracias a ellos también.

A MI ASESORA: gracias a sus conocimientos, experiencia profesional y su tiempo en la cual me ayudo a cumplir con éxito mi investigación y así cumplir con unas de mi metas trazadas en mi vida personal, agradezco a la profesora Mg Olga Solano Dávila.

A MIS PROFESORES de la Facultad de Ciencias Matemáticas, especialmente a mis profesores de la Escuela Académica Profesional de Estadística con sus enseñanzas me ayudaron a formarme profesionalmente.

RESUMEN

EL MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL PARA ESTUDIAR EL COMPORTAMIENTO DE LOS ESTUDIANTES DE LA FCM FRENTE A PREGUNTAS SENSIBLES

BR. JHON WILVER GALVAN TAPIA

Junio– 2016

ASESORA

Mg. Olga Solano Dávila

TITULO OBTENIDO

Licenciado en Estadística

El Modelo de Respuesta no Aleatorizada Transversal propuesto por (Wu, Liang, & Lai, 2008), se aplicó un estudio en la Facultad de Ciencias Matemáticas de la UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS, para investigar el comportamiento de los estudiantes sobre preguntas sensibles como: copia en los exámenes, apuntes para rendir un examen, consumo drogas para rendir un examen y copia de párrafos obtenidos de otros trabajos. Los resultados se compararon con el Método de Entrevista Directa. La población comprende a todos los alumnos de la FCM matriculados en el semestre 2015-I.

PALABRAS CLAVES: MODELO DE RESPUESTA NO ALEATORIZADA, MODELO TRANSVERSAL, PREGUNTA SENSIBLE

ABSTRACT**THE NON-RANDOMIZED RESPONDE CROSSWISE MODEL TO STUDY THE
BEHAVIOR OF THE STUDENTS OF THE FCM COMPARED TO SENSITIVE
QUESTIONS****BR. JHON WILVER GALVAN TAPIA**

June 2016

ADVISOR

Mg Olga Solano Dávila

TITLE OBTAINED

Degree in Statistics

The Non-Randomized Response Crosswise Model proposed by (Wu, Liang, & Lai, 2008), to be applied in a study at the Faculty of Mathematical Sciences of the NATIONAL UNIVERSITY OF SAN MARCOS, to investigate the behavior of students on sensitive questions such as; cheating on exams, notes to copy on exams, drug consumption to take an exam and copy of paragraphs obtained from other Works. The results were compared with the method of direct interview. The population includes all the students of the FCM enrolled in the semester 2015-I.

KEY WORDS: NON-RANDOMIZED RESPONSE MODEL, CROSS MODEL, SENSITIVE QUESTION

ÍNDICE GENERAL

RESUMEN	vi
ABSTRACT.....	vii
CAPÍTULO I: MARCO REFERENCIAL.....	2
1.1 INTRODUCCIÓN	2
1.2 ANTECEDENTES.....	3
1.2.1 ANTECEDENTE NACIONAL.....	4
1.2.2 ANTECEDENTE INTERNACIONAL.....	4
1.3 CONCEPTOS BÁSICOS.....	7
1.4 JUSTIFICACIÓN.....	8
1.5 VIABILIDAD Y LIMITACIÓN.....	9
1.6 OBJETIVOS.....	9
1.6.1 OBJETIVO GENERAL.....	9
1.6.2 OBJETIVOS ESPECÍFICOS	9
CAPÍTULO II: EL MODELO DE RESPUESTA NO ALEATORIZADA	12
2.1 BASE TEÓRICA	12
2.1.1 EL DISEÑO DE LA ENCUESTA PARA EL MODELO DE WARNER.....	13
2.1.2 EFICIENCIA RELATIVA	15
2.1.3 GRADO DE PROTECCIÓN DE LA PRIVACIDAD	17
2.2 DISEÑO DE ENCUESTA PARA EL MODELO TRANSVERSAL.....	21
2.2.1 CONEXIÓN CON EL MODELO DE WARNER	22
2.2.2 VENTAJAS DEL MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL EN COMPARACIÓN CON EL MODELO WARNER.....	24
2.2.3 EL INTERVALO DE CONFIANZA DE WALD.....	24
2.2.4 INTERVALO DE CONFIANZA DE WILSON	25
2.3 TAMAÑO DE MUESTRA CON EL MÉTODO DEL ANÁLISIS DE LA POTENCIA DE PRUEBA	26
2.3.1 MÉTODOS DE ANÁLISIS DE PRECISIÓN Y POTENCIA.....	27
2.3.2 ANÁLISIS DE LA PRECISIÓN.....	28
2.3.3 ANÁLISIS POTENCIA	29

2.4 EL MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL DE UNA MUESTRA.....	29
2.4.1 PRUEBA DE HIPÓTESIS UNILATERAL.....	30
2.4.2 PRUEBA DE HIPÓTESIS BILATERAL.....	32
2.4.3. COMPARACIÓN DEL PODER ASINTÓTICO Y EL PODER EXACTO DE LA PRUEBA	32
2.4.4. RELACIÓN ENTRE EL TAMAÑO DE MUESTRA DEL MRNA TRANSVERSAL Y MÉTODO TRADICIONAL.....	33
CAPÍTULO III: APLICACIÓN DEL MODELO DE RESPUESTA NO ALEATORIZADA, TRANSVERSAL, PARA ESTUDIAR EL COMPORTAMIENTO DE LOS ESTUDIANTES EN LA FCM Y COMPARACIÓN CON EL MÉTODO DE ENTREVISTA DIRECTA	37
3.1 MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL	37
3.1.1 ELECCIÓN DE LAS VARIABLES	37
3.1.2 PLAN DE MUESTREO DE LA APLICACIÓN	39
3.1.3 TAMAÑO DE MUESTRA	40
3.2 RESULTADOS DE LA APLICACIÓN DEL MODELO DE RESPUESTA NO ALEATORIZADA, EL MODELO TRANSVERSAL.....	42
3.2.1 RESULTADOS DE LA APLICACIÓN DEL MODELO TRANSVERSAL.....	42
3.2.2 RESULTADOS DE LA APLICACIÓN DEL MÉTODO DE ENTREVISTA DIRECTA.....	49
3.2.3 COMPARACIÓN ENTRE EL MODELO TRANSVERSAL Y EL MÉTODO DE ENTREVISTA DIRECTA	54
CONCLUSIONES Y RECOMENDACIONES	57
REFERENCIA BIBLIOGRÁFICA	60
ANEXO.....	62

ÍNDICE DE TABLAS

Tabla N° 2.1: Eficiencia Relativa para varias combinaciones de π y p	17
Tabla N° 2.2: El MRNA Transversal y sus correspondientes probabilidades.....	21
Tabla N° 2.3: Tamaño de la muestra n_c para la prueba $H_0: \pi = \pi_0$ versus $H_1: \pi = \pi_1 < \pi_0$ con él $\alpha = 0.05$ y el 0.8 de la potencia y la relación n_c/n_D	34

ÍNDICE DE GRÁFICOS

Gráfico 2.1 El modelo de Warner.....	15
Gráfico 2.2 Grado de protección del entrevistado con $\pi = 0.10$	18
Gráfico 2.3 Grado de protección del entrevistado con $\pi = 0.50$	19
Gráfico 2.4 Grado de protección del entrevistado con $\pi = 0.70$	19
Gráfico 2.5 Grado de protección del entrevistado con $\pi = 0.90$	20

ÍNDICE DE CUADROS

Cuadro N° 1: Probabilidades del Error de Tipo I y tipo II.....	27
Cuadro N° 2: Preguntas y Enunciados utilizados en el MRNA Transversal.....	37
Cuadro N° 3: Distribución de alumnos matriculados según EAP y sexo de la FCM-UNM (Semestre 2015-I).....	40
Cuadro N° 4: Distribución de la muestra para el MRNA Transversal de los alumnos matriculados según EAP y sexo de la FCM-UNMSM (Semestre 2015-I).....	40
Cuadro N° 5: Distribución de la muestra para el MED de los alumnos matriculados según EAP y sexo de la FCM-UNMSM (Semestre 2015-I).....	40
Cuadro N° 6: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación en porcentaje de estudiantes que han copiado en los exámenes.....	53
Cuadro N° 7: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación en porcentaje de estudiantes que han utilizados notas en los exámenes.....	53
Cuadro N° 8: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación en porcentaje de estudiantes que han utilizado droga para mejorar el rendimiento académico en los exámenes.....	54
Cuadro N° 9: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación en porcentaje de estudiantes que han presentado un trabajo intencionalmente adoptado del trabajo de otra persona.....	54

ÍNDICE DE ANEXOS

ANEXO N°1 Estimación de la proporción λ mediante el método de Máxima Verosimilitud en el MRNA Transversal.....	61
ANEXO N°2 El estimador de la proporción π mediante el método de Máxima Verosimilitud en el MRNA Transversal.....	62
ANEXO N°3 Procedimiento de estimación de la varianza del estimador π	63
ANEXO N°4 Resultados de la encuesta realizada a los estudiantes matriculados en el semestre 2015-I de la Facultad de Ciencias Matemáticas de la Universidad Mayor de San Marcos mediante el MRNA Transversal.....	66
ANEXO N°5 Resultados de la estimación de parámetros de la encuesta realizada en los estudiantes matriculados en el semestre 2015-I de la Facultad de ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos mediante el MRNA Transversal.....	75
ANEXO N°6 Resultados de la encuesta realizada a los estudiantes matriculados en el semestre académico 2015-I de la Facultad Ciencias Matemáticas de la UNMSM mediante el Método de Entrevista Directa.....	78
ANEXO N°7 Estimación de los parámetros de los resultados obtenidos de la encuesta obtenida en los estudiantes de la Facultad de Ciencias Matemáticas de la UNMSM el semestre académico 2015-I mediante el Método de la Entrevista Directa.....	85
ANEXO N° 8 Realizando códigos en el programa R Project.....	87
ANEXO N° 9 “Cuestionario 1: El Modelo de Respuesta No Aleatorizada Transversal”.....	90
ANEXO N°10 “Cuestionario 2: Sobre El Método de Entrevista Directa”.....	93

CAPÍTULO I:
MARCO REFERENCIA

CAPÍTULO I: MARCO REFERENCIAL

1.1 INTRODUCCIÓN

En estudios de encuestas por muestreo, incluso de aquellas realizadas por el Instituto de Estadísticas Oficiales como el Instituto Nacional de Estadística e Informática (INEI), los problemas de no respuestas y aquellos entrevistados que no contestan con veracidad las preguntas, son problemas frecuentes que se les presentan en sus encuestas. Estos dos problemas son las razones principales que generan sesgos de muestreo en cualquier tipo de estudio de muestreo probabilístico. Por lo cual la importancia de estudiar un diseño que mejore el problema del sesgo de respuesta, que vendría a ser el Modelo de Respuesta Aleatorizada propuesto por (Warner, 1965) quien utilizó un mecanismo aleatorio, en la cual trataría de no comprometer al entrevistado con respecto a preguntas sensibles, para así disminuir el sesgo de respuesta sobre dichas preguntas. Pero en esta investigación se estudiara el Modelo de Respuesta No Aleatorizada (MRNA) propuesto por (Wu, Liang, & Lai, 2008) partida del diseño de Warner con el propósito de disminuir el sesgo de respuesta.

En los últimos años existen países que están aplicando el Modelo de Respuesta No Aleatorizada para analizar el comportamiento de las personas sobre preguntas sensibles por lo que el MRNA es un método especialmente diseñado para asegurar privacidad a los entrevistados en los estudios de temas sensibles, delicados o embarazosos, con el propósito de disminuir el sesgo de respuesta. Según (Trujillo & Gonzàles, 2012) “los problemas de no respuesta y aquellos que se generan cuando los entrevistados no contestan con veracidad las preguntas, son bastante frecuentes debido al tratamiento de aspectos sensibles o confidenciales para la mayoría de personas en la población”. Se intenta con ello evitar el sesgo de los entrevistados hacia la respuesta socialmente más deseable de ciertas conductas. Se utilizan para analizar temas como copiar en los exámenes, insolvencia, fraudes, haber sido arrestado, conducir bajo los efectos del alcohol, tener un hijo fuera del matrimonio, aborto, etc.

El presente trabajo de investigación involucra preguntas sensibles o delicadas, como el comportamiento al plagio de exámenes, uso de notas para el plagio de exámenes, uso de drogas

ilegales y uso de trabajos adoptados de otras personas. Los entrevistados pueden negarse a responder las preguntas o dar respuestas falsas para proteger su privacidad, es tan común en la práctica, que es difícil medir y controlar. Para hacer frente a esta dificultad, los enfoques de respuesta no aleatorizada permiten a los profesionales de encuesta por muestreo y estadísticos, aplicaciones para proteger la privacidad de los entrevistados y analizar adecuadamente los datos recogidos.

El primer capítulo comprende el marco teórico, el cual contiene los antecedentes de investigación, conceptos básicos, justificación y objetivos del estudio de investigación.

El segundo capítulo comprende el Modelo de Respuesta no Aleatorizada (MRNA) que partirá de una base teórica del Modelo de Respuesta Aleatorizada (MRA) propuesta por (Warner, 1965), para luego, explicar el MRNA y hacer una conexión entre estos modelos, una vez estudiado la teoría de MRA y MRNA se estudiará el diseño de encuesta para el MRNA Transversal.

En el tercer capítulo se aplica el MRNA Transversal propuesto por (Wu, Liang, & Lai, 2008), en los estudiantes matriculados en la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos. La población de estudio serán todos los estudiantes matriculados en el semestre 2015-I. También se aplicará el Método de Entrevista Directa (MED) a una muestra asignada, el cual nos permitirá comparar con el MRNA Transversal.

Y finalmente se presenta las conclusiones del estudio de investigación así como recomendaciones que permitirán entender el MRNA Transversal aplicado a los estudiantes de la Facultad de Ciencias Matemáticas.

1.2 ANTECEDENTES

En los últimos años se está aplicando el Modelo de Respuesta No Aleatorizada para estudiar el comportamiento de las personas sobre preguntas sensibles y estas han demostrado su efectividad, debido a que muchos entrevistados tienden a mentir sobre alguna pregunta sensible

(comprometedora), por lo que, existen muchos estudios e investigaciones realizadas recientemente del MRNA Transversal con respecto a preguntas sensibles, con el propósito de buscar el anonimato del entrevistado para disminuir el sesgo de respuesta.

A continuación mostramos antecedentes del MRNA Transversal a nivel nacional e internacional

1.2.1 ANTECEDENTE NACIONAL

No se registra un estudio del MRNA Transversal en la Universidad Nacional Mayor de San Marcos, ni en el Perú.

1.2.2 ANTECEDENTE INTERNACIONAL

Existen una multitud de estudios realizados para mejorar la calidad y veracidad de las respuestas obtenidas sobre temas sensibles, como consumo de drogas, relaciones sexuales con más de dos personas, consumo de alcohol, copia de exámenes, etc.

Frecuentemente los investigadores obtienen resistencia a la participación en estos estudios u obtienen respuestas falsas por parte de los entrevistados. Warner (1965), Realizo la primera propuesta para obtener respuestas validas ente preguntas embarazosas basándose en la realización de dos preguntas mutuamente excluyentes (por ejemplo “A: Usted consume drogas para dar un examen”; B: Usted no consume drogas para dar un examen”) posteriormente, se desarrollaron otros métodos basados en (Warner, 1965).

Investigaciones recientes se mencionaran algunas que han proporcionado la primera evidencia empírica de que el MRNA Transversal supera al Método de Entrevista Directa sobre preguntas sensibles (sesgo de respuesta) (Coutts, Jann, Krumpal, & Naeher, 2011) .

Khosravi (2015), realizo un estudio sobre la prevalencia del uso de drogas ilícitas en estudiantes matriculados en las universidades de “Shahround”. El instrumento de medición fue

diseñado en base al MRNA Transversal donde se utilizaron seis preguntas sensibles con otras seis preguntas no sensibles las principales preguntas sensibles que consideraron fueron el uso de drogas ilícitas y residuos de opio.

La muestra utilizada en este estudio fue de 1644 alumnos, el 49.5% (814) fueron varones y el 50.5% (830) fueron mujeres, los resultados más importantes fueron:

- El 19% de estudiantes utilizaron al menos un tipo de drogas ilícitas.
- El 14.9% usaron residuos de opio una vez en su vida.
- El 3.5% de estudiantes utilizaron drogas.
- El 3% usaron residuo de opio durante el último mes.

También se estudió la comprensión del diseño del cuestionario, 40.27% de los estudiantes reportaron que ellos comprendían las instrucciones del cuestionario, 38.09% de estudiantes comprendieron las instrucciones de forma parcial, 12.02% de los estudiantes reportaron poca comprensión, y 9.59% de estudiantes no comprendieron el cuestionario.

En la investigación realizada por Jann, Jerke, & Krumpal (2011), sobre el plagio de trabajos de estudiantes de nacionalidad Alemana y Suiza se utilizó el MRNA Trasversal. La encuesta fue realizada entre los meses de junio y julio de 2009 entre 474 estudiantes universitarios alemanes y suizos. Los cuestionarios se entregaron en cada casa de estudio, Universidad de Leipzig (273), la ETH de Zúrich (111), y de la LMU Múnich (90), los estudiantes fueron asignados aleatoriamente de uno a dos condiciones experimentales, utilizando una muestra en relación de tres a uno para el MRNA Transversal(n= 358) y para el Método de Entrevista Directa(n=116) respectivamente.

De acuerdo con preguntas de entrevista directas, la prevalencia de plagio parcial es de 7.3%, frente al 22.3% para el MRNA Transversal, por lo tanto, podemos confirmar la expectativa de que, en el MRNA Transversal las estimaciones son más altas que las estimaciones del Método de Entrevista Directa, sin embargo, las estimaciones de plagio severo no parecen diferir entre los dos métodos de medición. Las estimaciones son de un 1% para el Método de Entrevista Directa y el 1.6% para el MRNA Transversal.

En la investigación realizada por Kundt(2014), tiene como objetivo de ver la proporción de personas que evaden impuestos mediante una encuesta por línea en Alemania en el 2014. Los entrevistados se les pregunto mediante un cuestionario en el correo electrónico, a través de correo de lista de universidad y en negocios así como las redes sociales. En una declaración introductoria, los entrevistados fueron informados sobre el contenido del cuestionario. Se les aseguró el anonimato completo y que los datos proporcionados por ellos serían utilizadas únicamente con fines científicos.

Para evaluar los beneficios de estudiar la evasión de impuestos se utilizó el MRNA Transversal. En este modelo, en la pregunta no sensible se aplicó una distribución logarítmica. Newcomb-Benford utiliza “La ley de Benford” que fue introducida con éxito para los MRA, pero no para el MRNA Transversal. Demostrando que la prevalencia de evasión de impuestos para el MRNA Transversal es significativamente mayor en comparación con el Método de Entrevista Directas. El análisis Multivariante revela que este resultado es robusto. Además utilizando el “Benfor Ilusion”, se encontró que la varianza del estimador para el MRNA Transversal sigue siendo bastante bajo. Además, los resultados presentados sugieren que el MRNA Transversal reduce significativamente la falta de respuesta.

Korndorfer, Krumpal, & Schmukle (2014), estos autores presentaron un artículo de investigación sobre evasión de impuesto aplicando el MRNA Trasversal, realizaron una encuesta en línea (correo electrónico) donde compararon el MRNA Transversal donde se asignaron aleatoriamente a los participantes con una muestra para el Método de Entrevista Directa (1650) y para el MRNA Transversal (5000), todos los participantes recibieron un correo electrónico con garantía de anonimato. Con respecto a la situación específica de las tasas de respuesta, de las 1650 personas que originalmente fueron invitados a participar al estudio mediante la Método de Encuesta Directa, se obtiene 305 personas (18.48%) proporcionó información sobre su comportamiento de evasión de impuestos directamente. De las 5000 personas que fueron invitadas por el MRNA Transversal, 862 personas (17.24%) dieron la necesaria información confidencial a través del MRNA Trasversal. En el Método de Entrevista Directa el 93.56% proporcionó información sobre su evasión de impuestos, mientras que en el MRNA Transversal

un 89.98% dio la respuesta necesaria. Así, en este análisis, las tasas de respuesta fueron de hecho ligeramente inferiores en la MRNA Transversal que en el de Entrevista Directa.

Los autores de esta investigación señalan que el MRNA Transversal es específicamente utilizable para encuestas en línea donde las personas carecen de un dispositivo aleatorio y lo más importante es que presenta un gran nivel de confianza y más comprensible, por lo cual el MRNA Trasversal es el más adecuado para encuesta en línea (correo electrónico).

1.3 CONCEPTOS BÁSICOS

ENCUESTA POR MUESTREO

Es el método que permite la selección de una muestra que proviene de una población, utilizando técnicas de muestreo, obteniéndose una enumeración muestral para luego recoger información pertinente en investigaciones de diversas disciplinas.

SESGO DE RESPUESTA

El sesgo de respuesta es un indicador estadístico que muestra el error que se detecta en los resultados de un estudio y que se debe a factores en la recolección, análisis, interpretación de los datos.

PREGUNTA SENSIBLE

Pregunta que el entrevistado puede considerar comprometedora, delicado, confidencial, etc... afectando la veracidad de su respuesta.

MODELO DE RESPUESTA ALEATORIZADA (MRA)

Modelo que busca la privacidad del entrevistado con el uso de mecanismos aleatorios, frente a preguntas sensibles.

MODELO DE RESPUESTA NO ALEATORIZADA (MRNA)

Modelo que busca la privacidad del entrevistador sin el uso de mecanismos aleatorios, frente a preguntas sensibles, por ejemplo el MRNA Transversal se utiliza una “tabla de contingencia”.

1.4 JUSTIFICACIÓN

Las tareas, ejercicios, investigaciones, reflexiones y ensayos son prácticas imprescindibles para que los alumnos pasen de simplemente “entender” el discurso del profesor, a realmente comprender y apropiarse del conocimiento y habilidades que la educación desea que ellos desarrollen. Sin embargo, los fraudes y plagios de ciertos alumnos en sus trabajos y tareas académicas son un problema muy grave en nuestra universidad y el resto de instituciones educativas del país; esto impide el aprendizaje del alumno, crea la falsa imagen de aprovechamiento. Se han estimado en estudio de investigación mediante encuestas realizadas donde el número de estudiantes que han copiado en un examen es alta en Universidades de Alemania (Höglinger, Jann, & Diekmann, 2014) ; como también la percepción de estudiantes de Guatemala, de primero a quinto ciclo, los alumnos de quinto ciclo mostraron mayor grado de seguridad que cuando parrafean no reconocen al autor y esto hace ver como plagio grave (Schlosser Montes, 2014).

Por este motivo, la importancia de realizar este estudio. Los resultados del presente estudio servirán como evidencia para incentivar programas de educación, escenario sumamente

importante para promover comportamientos adecuados que beneficien el desarrollo integral del alumno.

1.5 VIABILIDAD Y LIMITACIÓN

La investigación es viable, ya que cuenta tanto con recurso económico, como humanos, para llevar a cabo.

Se contara con el apoyo de la Dirección Académica de la Facultad de Ciencias Matemáticas y de todas sus escuelas profesionales seleccionadas, para realizar el estudio de investigación, así mismo se contara con el apoyo de la Escuela Académica Profesional Estadística, en brindar su apoyo para lo cual facilitara la recolección de datos.

La investigación tomara un tiempo de 11 meses.

Existe una limitación para conseguir más recursos para poder estudiar en toda la Universidad Nacional Mayor de San Marcos y así poder comparar los resultados con las demás Facultades, pero se espera que en un futuro próximo se pueda realizar dicha investigación.

1.6 OBJETIVOS

1.6.1 OBJETIVO GENERAL

Estudiar la teoría del Modelo de Respuesta no Aleatorizada Transversal y aplicarlo para investigar el comportamiento de los estudiantes de la FCM frente a preguntas sensibles.

1.6.2 OBJETIVOS ESPECÍFICOS

- Estudiar el MRNA Transversal.
- Aplicar el MRNA Transversal a los estudiantes de la FCM matriculados en el semestre 2015-I para estudiar el comportamiento de los alumnos frente a preguntas sensibles.

- Aplicar el Método de Entrevista Directa a los estudiantes de la FCM matriculados en el semestre 2015-I, para estudiar el comportamiento de los alumnos frente a preguntas sensibles.

CAPÍTULO II:
EL MODELO DE RESPUESTA NO
ALEATORIZADA

CAPÍTULO II: EL MODELO DE RESPUESTA NO ALEATORIZADA

2.1 BASE TEÓRICA

Esta investigación trata sobre una encuesta de muestreo con el objetivo de obtener información para visualizar una necesidad. La encuesta por muestreo nos permitirá encontrar la muestra representativa de nuestra población en estudio.

Según Raj (1968), la encuesta de muestreo permite se tomen decisiones que tienen en cuenta los factores significativos de los problemas que se procura resolver. Como instrumento para descubrir hechos, la encuesta por muestreo no se ocupa principalmente de la interpretación económica o sociológica de los hechos que demuestra, aunque debería proporcionar información adecuada para esas interpretaciones. Más bien ocupa de la adecuada representación de los hechos individuales registrados y de su recopilación y resumen.

En muchas investigaciones con respecto a preguntas sensibles, mediante entrevista directa, generalmente los entrevistados no contesten con veracidad, por lo cual a lo largo de las últimas décadas, los investigadores realizaron métodos para no comprometer a los entrevistados buscando aumentar el anonimato sobre las preguntas sensibles. Diversos estudios se han realizado proveniente del Modelo de Respuesta Aleatorizada de (Warner, 1965) para aumentar la protección de la privacidad en la que combina preguntas sensibles con preguntas no sensibles. Por ejemplo, presentando las siguientes preguntas a los entrevistados:

- (a) ¿Usted consume drogas? (es decir, que pertenece a la clase A).
- (b) ¿Usted no consume drogas? (es decir, que no pertenece a la clase A).

Utilizando un mecanismo aleatorio se determina al azar la pregunta a contestar, por lo que el entrevistador no es consciente del resultado y no puede observar directamente.

El componente aleatorio es controlado por el entrevistador provocando una desconfianza en el entrevistado, por lo que el Modelo de Respuesta Aleatorizada (MRA) todavía deja a los

entrevistados la oportunidad de elegir una estrategia de autoprotección al ignorar el mecanismo de MRA y simplemente contestan no, en cualquier caso.

Para superar este problema, esta investigación emplea un nuevo modelo que viene hacer el Modelo de Respuesta No Aleatorizada (MRNA) el cual fue propuesto por (Wu, Liang, & Lai, 2008) quienes realizaron varios diseños de este modelo.

En esta investigación se describirá el Modelo de Respuesta No Aleatorizada Transversal y su conexión con el Modelo de Respuesta Aleatorizada de (Warner, 1965). Este modelo mejora la protección de la privacidad del entrevistado; porque los entrevistados tienen dos opciones para responder en forma conjunta; este modelo parece más adecuado para estudiar temas sensibles en comparación con el MRA, por lo tanto, con el fin de evitar el uso de dispositivos de aleatorización, recientemente otros autores desarrollaron el Modelo de Respuesta No Aleatorizada (Wu, Liang, & Lai, 2008). Y han demostrado que estos MRNA generalmente obtienen mejores resultados que los correspondientes al MRA en el grado de protección de la privacidad del entrevistado.

El MRNA, el cual utiliza una variable no sensibles (ejemplo, fecha de nacimiento del entrevistado), combinado con una variables sensibles (ejemplo, consumo de droga del entrevistado) para formar una tabla de contingencia incompleta, obtiene indirectamente de los resultados de las preguntas sensibles (Wu, Liang, & Lai, 2008); (Tan, Tian, & Tang, 2009), por lo cual el modelo no requiere de mecanismos de aleatorización como lo hacía el diseño de Warner.

2.1.1 EL DISEÑO DE LA ENCUESTA PARA EL MODELO DE WARNER

Cada entrevistado está dirigido a responder la pregunta (a) o (b) en privado por medio de un dispositivo de aleatorización como un dado o un spinner, sin indicar al entrevistador que pregunta respondió. Por lo tanto, el entrevistador recibe una respuesta (por ejemplo, 'sí' o 'no') del entrevistado sin conocer de qué pregunta se trata. Esto alivia la privacidad y reduce el número de preguntas negativas a responder con respuestas falsas.

Sea p el elemento que denota la probabilidad de asignar la pregunta (a) a un entrevistado por el dispositivo (aleatorio) una asignación al azar; esto es,

$$p = P \{ \text{seleccionar el enunciado (a) por el dispositivo de aleatorización} \}.$$

El mecanismo de aleatorización es totalmente controlado por el investigador. Es decir, la probabilidad p es elegido por el investigador como parte del diseño además debe ser conocido.

Supongamos que queremos estimar la proporción (π) de la población que pertenece a la pregunta sensible, a la clase A, es decir $\pi = P(A)$. Sea n' el número de preguntas que respondieron "sí" obtenidas a partir de n cuestionarios. El estimador de máxima verosimilitud (EMV) de π es:

$$\hat{\pi}_w = \frac{p - 1 + (n'/n)}{2p - 1}, \quad p \neq 0.5, \quad (2.1)$$

Siempre que $\hat{\pi}_w \in [0,1]$, donde el subíndice 'w' se refiere al 'Modelo de Warner'. Warner (1965) mostró que $\hat{\pi}_w$ es insesgado con varianza igual a:

$$Var(\hat{\pi}_w) = Var(\hat{\pi}_D) + \frac{p(1-p)}{n(2p-1)^2} \quad (2.2)$$

donde,

$$Var(\hat{\pi}_D) = \frac{\pi(1-\pi)}{n} \quad (2.3)$$

Se denota la varianza de $\hat{\pi}_D$ correspondiente al Método de Entrevista Directa (MED) para cualquier π fijo, observamos que:

$$nVar(\hat{\pi}_w) = \pi(1-\pi) + \frac{p(1-p)}{(2p-1)^2}, \quad (2.4)$$

es una función creciente de p cuando $0 < p < 0,5$, que se acerca rápidamente hasta el infinito cuando $p \rightarrow 0,5$, y luego se convierte en una función decreciente de p cuando $0.5 < p < 1$ (ver Gráfico 2.1).

Gráfico 2.1 El modelo de Warner

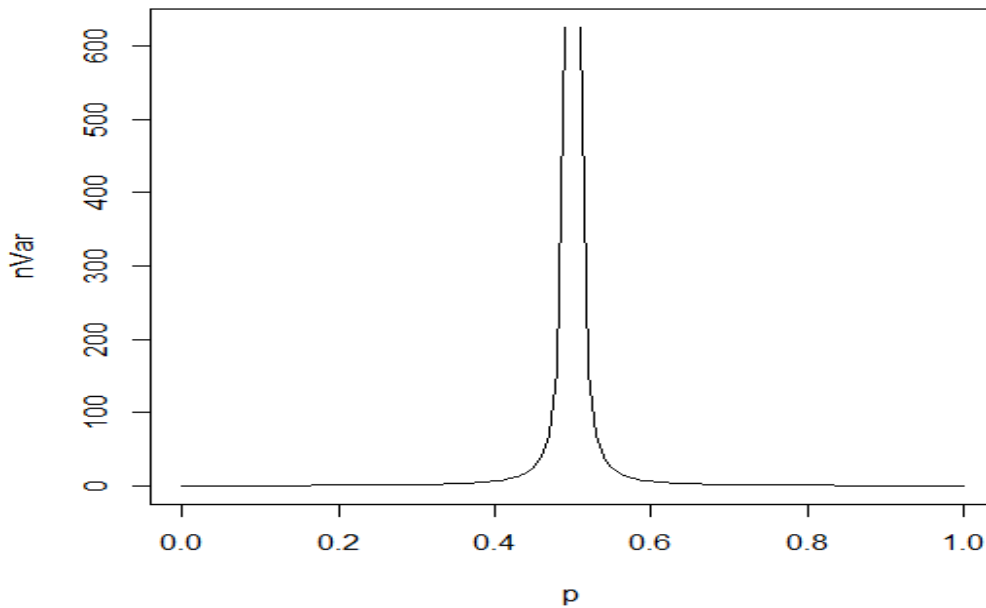


Gráfico 2.1. Muestra la relación entre la $nVar(\hat{\pi}_w)$ definida en (2.4) vs p con $\pi = 0.40$ para el Modelo de Warner.

2.1.2 EFICIENCIA RELATIVA

La limitación del modelo Warner es quizás su ineficiencia cuando se comparan con el MED. Tenga en cuenta que se introduce el segundo término de (2.2) debido al dispositivo de la aleatorización. Cuando $p = 0$ o $p = 1$, el modelo de Warner se reduce a la MED.

Se define la eficiencia relativa (ER) del diseño de Warner respecto al MED con

$$ER_{W \rightarrow D}(\pi, p) \cong \frac{Var(\hat{\pi}_w)}{Var(\hat{\pi}_D)} = 1 + \frac{p(1-p)/(2p-1)}{\pi(1-\pi)} \quad (2.5)$$

Que no depende del tamaño de la muestra (es decir, n), y permite que n_w y n_D sean los tamaños de muestra necesarios para el diseño de Warner y MED, respectivamente. Para lograr la misma precisión de estimación, tenemos:

$$\frac{\pi(1-\pi)}{n_w} + \frac{p(1-p)}{n_w(2p-1)^2} = \frac{\pi(1-\pi)}{n_D} \quad (2.6)$$

A partir de (2.5), se deduce que

$$\frac{n_w}{n_D} = ER_{W \rightarrow D}(\pi, p) \quad (2.7)$$

Por lo tanto, la eficiencia relativa está directamente relacionada con la proporción los tamaños de la muestra de los dos diseños.

La Tabla 2.1 muestra algunos valores de $ER_{W \rightarrow D}(\pi, p)$ para diversas combinaciones de π y p . Por ejemplo, cuando $\pi = 0.30$ y $p = 0.38$, tenemos,

$$ER_{W \rightarrow D}(0.15, 0.20) = 4.4858$$

Lo que sugiere que el tamaño de muestra requerido para el diseño Warner se trata de 4 veces la requerida para la MED con el fin de lograr la misma estimación de precisión. Para reducir la variabilidad, tenemos que aumentar el tamaño de la muestra n , resultando un incremento posiblemente significativo en costo (Tian & Tang, 2013).

TABLA 2.1: Eficiencia Relativa para varias combinaciones de π y p

π	P					
	0.20	0.30	0.34	0.38	0.42	0.46
0.05	10.3567	28.6316	47.1349	87.1111	201.3289	818.1053
0.10	5.9383	15.5833	25.3490	46.4475	106.7292	432.2500
0.15	4.4858	11.2941	18.1875	33.0806	75.6324	305.4118
0.20	3.7778	9.2031	14.6963	26.5642	60.4727	243.5781
0.25	3.3704	8.0000	12.6875	22.8148	51.7500	208.0000
0.30	3.1164	7.2500	11.4353	20.4775	46.3125	185.8214
0.35	2.9536	6.7692	10.6326	18.9792	42.8269	171.6044
0.40	2.8519	6.4688	10.1309	18.0428	40.6484	162.7188
0.45	2.7957	6.3030	9.8542	17.5264	39.4470	157.8182
0.50	2.7778	6.2500	9.7656	17.3611	39.0625	156.2500

Elaboración por (Wu, Liang, & Lai, 2008)

2.1.3 GRADO DE PROTECCIÓN DE LA PRIVACIDAD

Intuitivamente, el óptimo grado de protección de la privacidad (GPP) se alcanza cuando $p = 0.5$, que corresponde al caso de la varianza infinita (véase el gráfico 2.1).

Cuando p es demasiado pequeño o demasiado grande, la privacidad de los entrevistados no puede ser suficientemente protegida, por lo tanto, los investigadores se ven obligados a seleccionar un valor de p dentro de algún sub-intervalo de $(0, 0.5)$, donde $n\text{Var}(\pi_w)$ es una función simétrica en $p = 0.5$ y por lo tanto adopta un método de aleatorización desigual (por ejemplo, moneda sesgada) (Tian & Tang, 2014). Dado que la información sensible de un entrevistado respecto a su pertenencia a la clase sensible A se caracteriza a través de $P(A | \text{sí})$ y $P(A | \text{no})$.

Definimos.

$$GPP_{SI}(\pi, p) \cong P(A|si) = \frac{\pi p}{\pi p + (1 - \pi)(1 - p)} \quad (2.8)$$

y

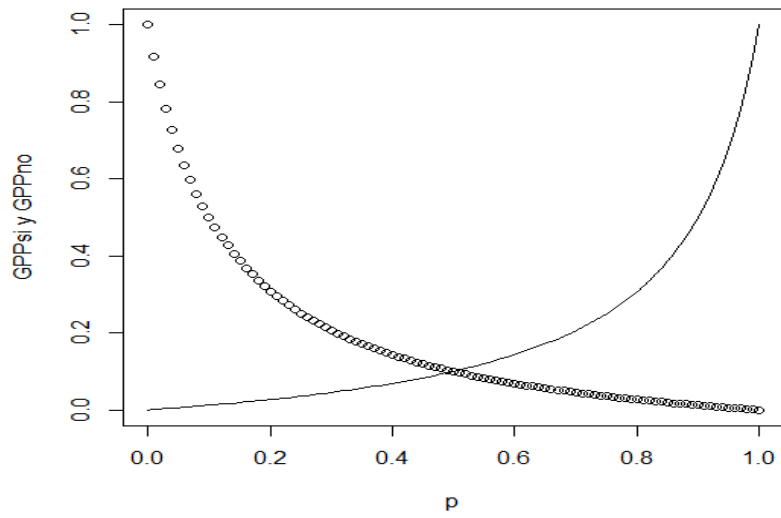
$$GPP_{NO}(\pi, p) \cong P(A|no) = \frac{\pi(1 - p)}{\pi(1 - p) + (1 - \pi)p} \quad (2.9)$$

para medir la información privada con el modelo de Warner.

El Gráfico 2.2 muestra que para un π fijo, $GPP_{SI}(\pi, p)$ es una función monótona creciente de p , mientras $GPP_{NO}(\pi, p)$ es una función monótona decreciente de p . En particular, para cualquier $\pi \in (0, 1)$, cuando $p = 0.5$, tenemos

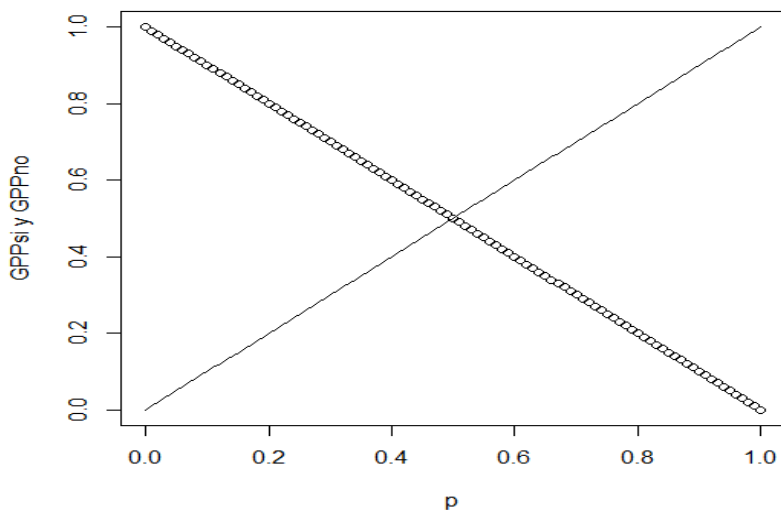
$$GPP_{SI}(\pi, 0.5) = GPP_{NO}(\pi, 0.5) = \pi$$

Gráfico 2.2: Grado de protección del entrevistado con $\pi = 0.10$



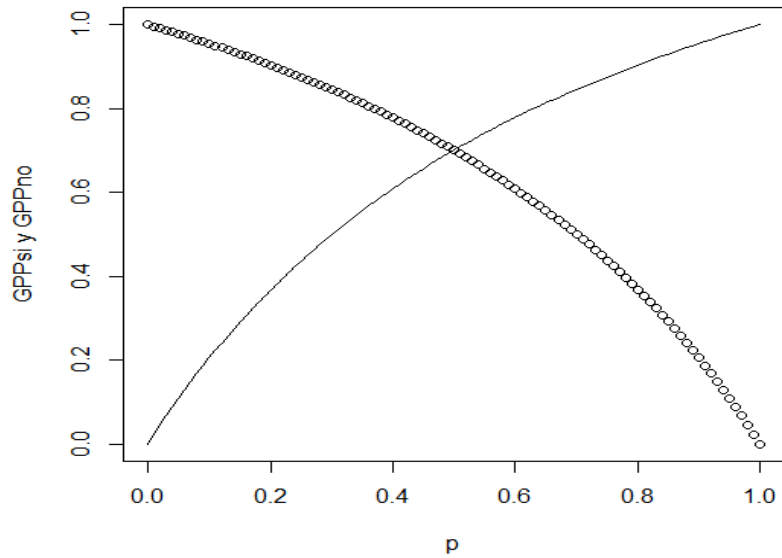
Elaboración por (Tian & Tang, 2014) ver Anexo5

Gráfico 2.3: Grado de protección del entrevistado con $\pi = 0.50$



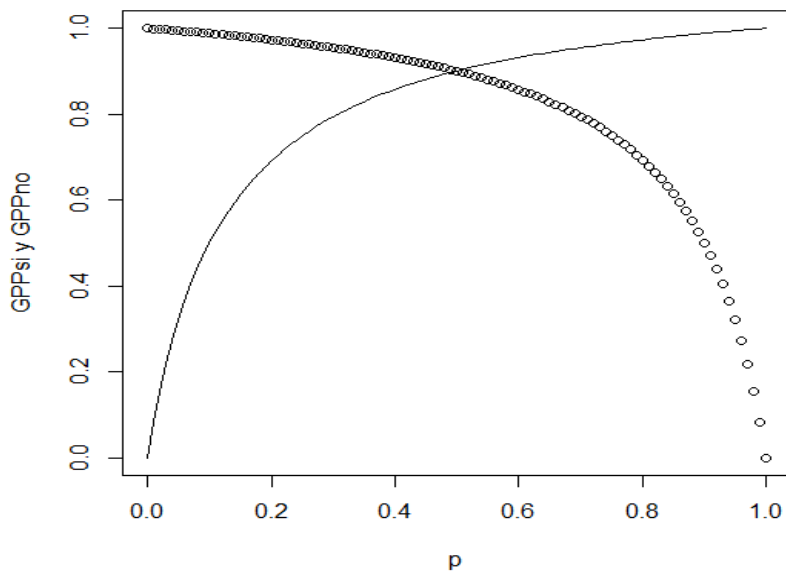
Elaboración por (Tian & Tang, 2014) ver Anexo4

Gráfico 2.4: Grado de protección del entrevistado con $\pi = 0.70$



Elaboración por (Tian & Tang, 2014) ver Anexo5

Gráfico 2.5: Grado de protección del entrevistado con $\pi = 0.90$



Elaboración por (Tian & Tang, 2014) ver Anexo4

En los Gráficos 2.2, 2.3, 2.4 y 2.5 los diagramas de $GPP_{si}(\pi, p)$ -línea continua- y $GPP_{no}(\pi, p)$ -línea de puntos- versus p en el modelo Warner y para un π fijo.

(i) $\pi = 0,10$; (ii) $\pi = 0.50$; (iii) $\pi = 0.70$; (iv) $\pi = 0.90$,

lo que implica que en este caso,

$$P(A|si) = P(A|no) = P(A).$$

En otras palabras, si el encuestado pertenece a la clase sensible no depende de la respuesta "sí" o "no", llegando así al máximo GPP. Sin embargo, cuando $p \in [0, 0.50]$, tenemos $GPP_{SI}(\pi, p) < \pi$, mientras $GPP_{no}(\pi, p) \geq \pi$.

2.2 DISEÑO DE ENCUESTA PARA EL MODELO TRANSVERSAL

Este diseño de investigación propuesto por (Wu, Liang, & Lai, 2008), con el propósito de mejorar la protección de privacidad del entrevistado con respecto al uso de preguntas sensibles sin el uso del componente aleatorio. Este diseño de encuesta por muestreo permitió conocer la importancia de MRNA, e inspiró su aplicación a los problemas sobre preguntas sensibles.

TABLA 2.2: El MRNA Transversal y sus correspondientes probabilidades

Categoría	W = 0	W = 1	Categoría	W = 0	W = 1	Marginal
Y = 0	○	□	Y = 0	$(1 - \pi)(1 - p)$	$(1 - \pi)p$	$(1 - \pi)$
Y = 1	□	○	Y = 1	$\pi(1 - p)$	πp	π
			Marginal	1 - p	p	1

Fuente: Adaptado de la tabla de Yu, Tian y Tang (2008)

El MRNA Transversal puede ser visto como una versión no aleatoria del original MRA de Warner. Se denota $\{Y = 1\}$ la clase de la gente que posee una característica sensible (por ejemplo el uso de drogas) y $\{Y = 0\}$ denota la clase complementaria. Dejamos que W sea una variable no sensible dicotómica y debe ser independiente de Y . El entrevistador debe seleccionar una variable W conveniente de modo que la proporción:

$$p = P[W = 1]$$

debe ser conocido en adelante o puede ser estimado fácilmente. Por ejemplo, podemos definir $\{W = 1\}$ si una persona ha nacido entre agosto y diciembre y $\{W = 0\}$ de la otra manera. De ahí, es razonable asumir que $p = 5/12 = 0.41667$. Nuestro objetivo es estimar la proporción.

$$\pi = P[Y = 1]$$

El entrevistador puede diseñar un cuestionario en el formato como muestra en la Tabla 2.1 y pedir a cada entrevistado que ponga una cruz en el círculo superior (es decir, $\{Y = 0, W = 0\}$) si él/ella pertenece a uno de los dos círculos o poner una cruz en el cuadrado superior (es decir, $\{Y = 0, W = 1\}$) si él/ella pertenece a uno de los dos cuadrados, por lo que el entrevistado no está siendo expuesto si la cruz esta puesta en la parte superior del circulo o cuadrado, Note que tanto $\{Y = 0, W = 0\}$ como $\{Y = 0, W = 1\}$ son clases no sensibles. Así, un entrevistado que pertenece a la clase sensible (es decir, $\{Y = 1\}$) no está siendo expuesto si marca en el círculo/cuadrado superior. Las probabilidades de las celdas correspondientes son dadas en el lado derecho de la Tabla 2.1.

2.2.1 CONEXIÓN CON EL MODELO DE WARNER

Supongamos que hay un total de n cruz (correspondiente a n entrevistados en la encuesta sin contar la respuestas en blanco) con n' cruz puestas en el círculo superior y $n - n'$ cruz de ser puesto en el cuadrado superior. Los datos observados, según (Vakilian, Mousavi, & Keramat, (2014, p. 03) se denotan por:

$$Y_{obs} = \{n; n', n - n'\}.$$

La probabilidad de poner un aspa en el círculo superior es entonces:

λ : La probabilidad de marcar el círculo superior siendo

$$\lambda = \frac{n'}{n}$$

$$n' \sim \text{Binomial}(n, \lambda),$$

$$(1 - \pi)(1 - p) + \pi p.$$

Dejando

$$\lambda = (1 - \pi)(1 - p) + \pi p,$$

tenemos

$$\pi = \frac{p - 1 + \lambda}{2p - 1}, \quad (2.10)$$

donde $p = P(W = 1) \neq 0.5$ es conocido. La función de verosimilitud es proporcional a $\lambda^{n'}(1 - \lambda)^{n-n'}$ (ver anexo) de modo que el EMV de λ está dada por $\hat{\lambda} = n' / n$. Por lo tanto, el EMV de π es:

$$\hat{\pi}_C = \frac{p - 1 + \hat{\lambda}}{2p - 1}, \quad p \neq 0.5, \quad (2.11)$$

siempre que $\hat{\pi}_C \in [0,1]$, donde el subíndice "C" se refiere al modelo 'transversal'.

Para encontrar la EMV de π desde

$$n' \sim \text{Binomial}(n, \lambda),$$

Tenemos

$$E(n') = n\lambda \quad \text{y} \quad \text{Var}(n') = n\lambda(1 - \lambda)$$

Por lo tanto, $E(\hat{\pi}_C) = \pi$, es decir, $\hat{\pi}_C$ es un estimador insesgado de π , y

$$Var(\hat{\pi}_c) = \frac{\lambda(1-\lambda)}{n(2p-1)^2} = \frac{\pi(1-\pi)}{n} + \frac{p(1-p)}{n(2p-1)^2} \quad (2.12)$$

2.2.2 VENTAJAS DEL MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL EN COMPARACIÓN CON EL MODELO WARNER

El EMV y la varianza son idénticos al modelo Warner respectivamente. Por lo tanto, (Wu, Liang, & Lai, 2008) consideran al MRNA Transversal como una versión no aleatorizado del modelo Warner.

En comparación con el modelo de Warner aleatorizado original, este no aleatorizado modelo transversal tiene las siguientes ventajas:

- Reducción de costos (ya que no se requiere un dispositivo de aleatorización).
- Reproducibilidad (Es decir, se espera que el mismo entrevistado de la misma respuesta por si el diseño de encuesta se presenta al entrevistado de nuevo).
- Una mejor cooperación (ya que la pregunta no sensible no se controla por los entrevistadores y los entrevistados son propensos a confiar en los entrevistadores).
- Una desventaja del MRNA Transversal y del MRA de Warner no aplicable en un cuestionario con muchas preguntas.

2.2.3 EL INTERVALO DE CONFIANZA DE WALD

El intervalo de confianza Wald se basa en la distribución asintótica del estimador muestral $\hat{\lambda}$. Es fácil demostrar que un estimador de $Var(\hat{\pi}_c)$ viene dada por («Intervalos de confianza para credibilidad», s. f., p. 214)

$$\widehat{Var}(\hat{\pi}_c) = \frac{\hat{\lambda}(1-\hat{\lambda})}{(n-1)(2p-1)^2} = \frac{\hat{\pi}_c(1-\hat{\pi}_c)}{n-1} + \frac{p(1-p)}{(n-1)(2p-1)^2} \quad (2.13)$$

Por el Teorema del Límite Central, cuando $n \rightarrow \infty$, la distribución de $\hat{\pi}_C$ se distribuye así:

$$\hat{\pi}_C - \pi / \sqrt{\widehat{\text{Var}}(\hat{\pi}_C)} \sim N(0,1), \quad \text{cuando } n \rightarrow \infty$$

A $(1 - \alpha)$ 100% de confianza de Wald para π viene dada por:

$$[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}] = \left[\hat{\pi}_C - z_{\alpha/2} \sqrt{\widehat{\text{Var}}(\hat{\pi}_C)}, \hat{\pi}_C + z_{\alpha/2} \sqrt{\widehat{\text{Var}}(\hat{\pi}_C)} \right], \quad (2.14)$$

donde Z_α es el α -ésimo cuantil superior de la distribución normal estándar.

Observación 2.1: En la práctica el límite inferior del Intervalo de confianza Wald dado por (2.14) es menor que cero, si el valor verdadero de π es cercano a cero ò el límite superior son mayores que uno, si el valor verdadero de π es cercano a uno, dando un intervalo de confianza no útil. Incluso $[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}]$ está dentro del intervalo de la unidad $[0,1]$, su rendimiento se demostró (Agresti y Coull, (1998)), al ser inferior al intervalo de confianza de Wilson, que se introduce a continuación.

2.2.4 INTERVALO DE CONFIANZA DE WILSON

Un inconveniente para el intervalo de confianza de Wald (2.14) es que el límite inferior puede ser menor a cero, cuando el verdadero valor de π esta cercano a cero, mientras que el límite superior puede estar más allá de uno cuando el verdadero valor π está cercano a uno.

Se construye el intervalo de confianza π de Wilson al $(1 - \alpha)$ 100% basado en

$$1 - \alpha = P \left\{ \left| \frac{\hat{\pi}_C - \pi}{\sqrt{\widehat{\text{Var}}(\hat{\pi}_C)}} \right| \leq z_{\alpha/2} \right\}$$

$$\begin{aligned}
&= P\{(\hat{\pi}_c - \pi)^2 \leq z_{\alpha/2}^2 \text{Var}(\hat{\pi}_c)\} \\
&= P\left[\hat{\pi}_c^2 - 2\hat{\pi}_c\pi + \pi^2 \leq z_{\alpha/2}^2 \left\{\pi(1 - \pi) + \frac{p(1 - p)}{(2p - 1)^2}\right\}\right] \\
&= P\{(1 + z_*)\pi^2 - (2\hat{\pi}_c + z_*)\pi + \hat{\pi}_c^2 - z_*\partial \leq 0\} \tag{2.15}
\end{aligned}$$

donde

$$z_* \hat{=} \frac{z_{\alpha/2}^2}{n} \text{ y } \partial \hat{=} \frac{p(1 - p)}{(2p - 1)^2},$$

Al resolver la desigualdad cuadrática dentro de la probabilidad (2.15) se obtiene el siguiente intervalo de confianza de Wilson de π :

$$[\hat{\pi}_{c,WSI}, \hat{\pi}_{c,WSS}] = \frac{2\hat{\pi}_c + z_* \pm \sqrt{(2\hat{\pi}_c + z_*)^2 - 4(1 + z_*)(\hat{\pi}_c^2 - z_*\partial)}}{2(1 + z_*)} \tag{2.16}$$

El cual, en general, está incluido en el intervalo $[0,1]$

2.3 TAMAÑO DE MUESTRA CON EL MÉTODO DEL ANÁLISIS DE LA POTENCIA DE PRUEBA

El cálculo del tamaño de la muestra es un componente esencial para la encuesta diseñada sobre preguntas sensibles. En esta sección, vamos a discutir el cálculo del tamaño de la muestra para el MRNA Transversal, basado en el método de análisis de potencia para una muestra. En la Sección 2.3.1, se introduce, brevemente dos métodos (es decir, el método de análisis de la precisión y el análisis de método de la potencia) para los cálculos del tamaño de la muestra).

2.3.1 MÉTODOS DE ANÁLISIS DE PRECISIÓN Y POTENCIA

Uno de los pasos más importantes en las encuestas por muestreo es la determinación del número de participantes. En la práctica, los tamaños de muestra se pueden determinar basados en el método de análisis de precisión o el método de análisis de potencia que está estrechamente relacionado con el concepto de error tipo I, error tipo II y la función potencia en la prueba de hipótesis.

2.2.3.1 EL ERROR DE TIPO I, EL ERROR DE TIPO II Y LA POTENCIA

Supongamos que X_1, \dots, X_n es una muestra aleatoria de una población con función de probabilidad $f(x; \theta)$, donde $\theta \in \Theta$ y Θ denota el espacio paramétrico.

Sea $X = (X_1, \dots, X_n)^T$ y $x = (x_1, \dots, x_n)^T$ denota su realización. Consideremos las siguientes hipótesis:

$$H_0: \theta \in \theta_0$$

$$H_1: \theta \in \theta_1$$

C y C' denotan respectivamente la región crítica y la región de aceptación de una prueba para contrastar la hipótesis nula H_0 frente a la alternativa H_1 .

- (1) El rechazo de la hipótesis nula H_0 cuando es verdadera se llama error tipo I. La probabilidad de cometer un error tipo I es:

$$\begin{aligned}\alpha &= P(\text{error tipo I}) \\ &= P(\text{rechazar } H_0 \mid H_0 \text{ es verdadera}) \\ &= P(x \in C \mid \theta \in \theta_0)\end{aligned}$$

(2) La aceptación de la hipótesis nula H_0 cuando es falsa se llama error Tipo II. La probabilidad de cometer el error tipo II es:

$$\begin{aligned}\beta &= P(\text{error tipo II}) \\ &= P(\text{aceptar } H_0 | H_0 \text{ es falsa}) \\ &= P(x \in C | \theta \in \theta_1)\end{aligned}$$

(3) Potencia de una prueba se define como la probabilidad de rechazar H_0 correctamente cuando la H_0 es falsa; esto es,

$$\begin{aligned}\text{Potencia} &= 1 - \beta \\ &= P(\text{rechazar } H_0 | H_0 \text{ es falsa}) \\ &= P(x \in C | \theta \in \theta_1).\end{aligned}$$

Podemos resumir estas definiciones en la siguiente tabla.

CUADRO 1: Probabilidades del Error de Tipo I y Tipo II

	<i>Ho es verdadero ($\theta \in \theta_1$)</i>	<i>Ho es falso ($\theta \in \theta_1$)</i>
<i>Rechaza Ho ($x \in C$)</i>	α	$(1 - \beta)$ <i>Potencia</i>
<i>Acepta Ho ($x \in C$)</i>	$1 - \alpha$	β

Elaboración propia

Nota: En la práctica, se suelen utilizar “No rechazamos H_0 ” para sustituir “Aceptar H_0 ”. El nivel de confianza definido por $1-\alpha$ refleja la probabilidad o la confianza de no rechazar la hipótesis nula cuando es verdadera.

2.3.2 ANÁLISIS DE LA PRECISIÓN

El análisis de la precisión de la determinación del tamaño de la muestra, por lo general, se realiza mediante el control del error tipo I, α (llamado también nivel de significación). El nivel de confianza, $1-\alpha$, refleja la probabilidad o la confianza de no rechazar la hipótesis nula cuando es verdadera.

2.3.3 ANÁLISIS POTENCIA

El análisis de la potencia para determinar del tamaño de la muestra se realiza mediante el control de un error de tipo II, β . Un enfoque típico en pruebas de hipótesis es la de controlar un nivel aceptable y al mismo tiempo minimizar β mediante la selección de un adecuado tamaño de muestra.

2.4 EL MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL DE UNA MUESTRA

El MRNA transversal de eficiencia relativa para varias combinaciones de π y p se describe en el Tabla 2.3. Sea $Y_{obs} = Y_C$, se define una variable aleatoria de Bernoulli Y_C , sean los datos observados para n entrevistados, donde $Y_C = 1$ si el i -ésimo entrevistado pone una marca en el círculo superior, $Y_C = 0$ si el i -ésimo entrevistado pone una marca en el cuadrado superior.

Donde el subíndice “C” representa la variable de Bernoulli para el MRNA Transversal.

Por lo tanto las probabilidades:

$$P[Y_C = 1] = (1 - \pi)(1 - p) + \pi p.$$

$$P[Y_C = 0] = \pi(1 - p) + (1 - \pi)p.$$

La función de máxima verosimilitud de π está dado por

$$L_C(\pi|Y_{obs}) = \prod_{i=1}^n \{(1 - \pi)(1 - p) + \pi p\}^{y_{i,c}} \{\pi(1 - p) + (1 - \pi)p\}^{1-y_{i,c}}$$

El Estimador de Máxima Verosimilitud (EMV) de π y su varianza son dados por

$$\hat{\pi}_c = \frac{p - 1 + \hat{\lambda}}{2p - 1} \quad y \quad Var(\hat{\pi}_c) = \frac{\lambda(1 - \lambda)}{n(2p - 1)^2}$$

$$p \neq 0.5, \quad (2.17)$$

donde

$$\bar{Y}_c = \frac{1}{n} \sum_{i=1}^n Y_{ci} \quad y \quad \lambda = (1 - \pi)(1 - p) + \pi p.$$

2.4.1 PRUEBA DE HIPÓTESIS UNILATERAL

De acuerdo al Teorema del Límite Central:

$$\frac{\hat{\pi}_c - \pi}{\sqrt{Var(\hat{\pi}_c)}} = \frac{n\hat{\pi}_c - n\pi}{(\sqrt{n\lambda(1 - \lambda)})/(2p - 1)} \sim N(0,1), \quad (2.18)$$

Con el fin de contrastar la hipótesis de la proporción de la población π con la característica sensible idéntica a un valor pre-especificado π_0 por el investigador, se plantean las siguientes hipótesis.

$$H_0: \pi = \pi_0$$

$$H_1: \pi < \pi_0$$

Si la hipótesis nula H_0 es verdadera, de (2.18) se tiene

$$\frac{n\hat{\pi}_c - n\pi_0}{(\sqrt{n\lambda_0(1 - \lambda_0)})/(2p - 1)} \sim N(0,1), \quad \text{cuando } n \rightarrow \infty,$$

donde n es el tamaño de muestra del modelo cruzado para la prueba de hipótesis unilateral a la izquierda, y

$$\lambda_0 = (1 - \pi_0)(1 - p) + \pi_0 p$$

Se rechaza la hipótesis nula, con un nivel de significación alfa cuando se observa el siguiente evento

$$E = \left\{ n\hat{\pi}_c \leq n\pi_0 - \frac{z_\alpha(\sqrt{n\lambda_0(1-\lambda_0)})}{2p-1} \right\}, \quad (2.19)$$

donde z_α el α -ésimo cuantil superior de la distribución normal estándar cuando el evento, es observado, se rechaza la hipótesis nula H_0 a un nivel de significación α ,

$$\begin{aligned} \alpha &= P(\text{rechazar } H_0 / H_0 \text{ es verdadera}) \\ 1 - \beta &= P(\text{rechazar } H_0 / H_0 \text{ es falsa}) \end{aligned}$$

Bajo el supuesto que H_1 es verdadera y $\pi = \pi_1$ con $\pi_1 < \pi_0$. La potencia de la prueba puede calcularse aproximadamente por:

La potencia (en π_1) = $P(\text{Rechazar } H_0 \mid \pi = \pi_1)$

$$\begin{aligned} &= P \left\{ \frac{n\hat{\pi}_c - E_{H_1}(n\hat{\pi}_c)}{\sqrt{\text{Var}_{H_1}(n\hat{\pi}_c)}} \leq \frac{n\pi_0 - \frac{z_\alpha(\sqrt{n\lambda_0(1-\lambda_0)})}{2p-1} - n\pi_1}{\frac{(\sqrt{n\lambda_1(1-\lambda_1)})}{2p-1}} \right\} \\ &= \Phi \left(\frac{\sqrt{n}(\pi_0 - \pi_1)|2p-1| - z_\alpha\sqrt{\lambda_0(1-\lambda_0)}}{\sqrt{\lambda_1(1-\lambda_1)}} \right) \end{aligned} \quad (2.20)$$

Donde $\lambda_1 = (1 - \pi_1)(1 - p) + \pi_1 p$ y $\Phi(\cdot)$ denotada la función de distribución acumulada de la distribución normal estándar. Para una potencia dada $(1 - \beta)$, el tamaño de muestra requerido n_c se puede determinar resolviendo la siguiente ecuación

$$n_c = \left[\frac{z_\alpha\sqrt{\lambda_0(1-\lambda_0)} + z_\beta\sqrt{\lambda_1(1-\lambda_1)}}{(\pi_0 - \pi_1)(2p-1)} \right]^2 \quad (2.21)$$

donde $\lambda_i = (1 - \pi_i)(1 - p) + \pi_i p$, $i = 0, 1$ y $\pi_1 < \pi_0$. En particular, cuando $p = 1$ el MRNA Transversal es idéntico al Diseño de Entrevista Directa (n_C es igual n_D)

2.4.2 PRUEBA DE HIPÓTESIS BILATERAL

Para la prueba de dos colas, las hipótesis se especifican de la siguiente manera:

$$H_0: \pi = \pi_0$$

$$H_1: \pi \neq \pi_0$$

Teniendo en cuenta con un nivel de significación α , se observa que la relación entre el poder de la prueba, tamaño de la muestra y el tamaño del efecto es aproximadamente dado por:

$$\text{La potencia(en } \pi_1) = \Phi \left(\frac{\sqrt{n}(\pi_0 - \pi_1)|2p - 1| - \frac{z_{\alpha/2}\sqrt{\lambda_0(1 - \lambda_0)}}{2}}{\sqrt{\lambda_1(1 - \lambda_1)}} \right) \quad (2.22)$$

$$n_C = \left[\frac{z_{\alpha/2}\sqrt{\lambda_0(1 - \lambda_0)} + z_{\beta}\sqrt{\lambda_1(1 - \lambda_1)}}{(\pi_0 - \pi_1)(2p - 1)} \right]^2 \quad (2.23)$$

2.4.3. COMPARACIÓN DEL PODER ASINTÓTICO Y EL PODER EXACTO DE LA PRUEBA

La función potencia asintótica para una prueba unilateral dada en (2.20). Para derivar la formula exacta de la potencia, definimos una nueva variable aleatoria. (Ulrich, Schröter, Striegel, & Simon, 2012)

$$X_C = \sum_{i=1}^n y_{i,C},$$

tenemos $X_C \sim \text{Binomial}(n, \lambda)$ con $\lambda = (1 - \pi)(1 - p) + \pi p$.

La región de rechazo E_P dado en (2.19) puede ser reinscrita como,

$$E_P = \{X_P \leq n\lambda_0 - z_\alpha \sqrt{n\lambda_0(1 - \lambda_0)}\}$$

La potencia exacta (en π_1) para algún tamaño de la muestra n en particular, se determina con la siguiente formula:

$$\begin{aligned} \text{La potencia exacta (bajo, } \pi_1) &= \sum_{x \in E} \text{Binomial}(x|n, (1 - \pi_1)(1 - p) + \pi_1 p) \\ &= \sum_{x \in E} \binom{n}{x} \lambda_1^x (1 - \lambda_1)^{n-x} \end{aligned}$$

2.4.4. RELACIÓN ENTRE EL TAMAÑO DE MUESTRA DEL MRNA TRANSVERSAL Y MÉTODO TRADICIONAL

La Tabla 2.3 presenta algunos valores del tamaño de muestra dado en n_C con un $\alpha = 5\%$ y el 80% de la potencia en virtud de diversas combinaciones de (π_0, π_1, p) y la correspondiente potencia exacta, que se calcula con la siguiente formula:

$$\text{La potencia exacta de } \pi_1 = \sum_{x \in E_C} \binom{n_C}{x} \lambda_1^x (1 - \lambda_1)^{n_C - x},$$

La región de rechazo especificado por (2.19) puede ser reinscrita así

$$E_C = \{X_C \leq n_C \lambda_0 - z_\alpha \sqrt{n_C \lambda_0 (1 - \lambda_0)}\}$$

y

$$X_C = \sum_{i=1}^{n_C} Y_C \sim \text{Binomial}(n_C, \lambda)$$

En general, el tamaño de muestra n_C es suficiente para garantizar la potencia deseada; es decir,

$$|\text{Potencia real} - \text{potencia deseada}| \leq 5\%$$

Para un determinado par de (π_0, π_1) , está claro que n_C es una función monótona decreciente de p . Comparamos el MRNA Transversal con el Método de Entrevista Directa con la presentación de los cocientes de n_C/n_D para $p=0.75, 0.70, \dots, 0.55$, respectivamente. Por ejemplo, cuando $(\pi_0, \pi_1) = (0.4, 0.3)$ y $p=0.65$, tenemos $n_C/n_D = 12$, lo que indica que el tamaño de la muestra necesaria para el MRNA Transversal es alrededor de 12 veces más que el que se requiere para el Método de Entrevista Directa con el fin de lograr la misma potencia.

Tabla 2.3: Tamaño de la muestra n_C para la prueba $H_0: \pi = \pi_0$ versus $H_1: \pi = \pi_1 < \pi_0$ con el $\alpha = 0.05$ y el 0.8 de la potencia y la relación n_C/n_D

π_0	π_1	P= 1		P= 0.75		P= 0.7	
		n_D		n_C	n_C/n_D	n_C	n_C/n_D
0.5	0.4	153		617	4.03	964	6.30
0.5	0.35	67		274	4.09	428	6.39
0.5	0.3	37		154	4.16	240	6.49
0.4	0.35	583		2451	4.20	3829	6.57
0.4	0.3	142		611	4.30	954	6.72
0.4	0.25	61		270	4.43	422	6.92
0.3	0.25	501		2398	4.79	3747	7.48
0.3	0.2	119		596	5.01	931	7.82
0.3	0.18	81		413	5.10	645	7.96
0.2	0.16	584		3621	6.20	5657	9.69
0.2	0.13	181		1177	6.50	1838	10.15
0.2	0.1	83		574	6.92	896	10.80
0.1	0.08	1303		13819	10.61	21592	16.57
0.1	0.06	301		3441	11.43	5376	17.86
0.1	0.04	121		1523	12.59	2379	19.66

		P= 0.65		P= 0.60		P= 0.55	
π_0	π_1	n_C	n_C/n_D	n_C	n_C/n_D	n_C	n_C/n_D
0.5	0.4	1714	11.21	3856	25.20	15423	100.80
0.5	0.35	760	11.37	1709	25.51	6836	102.03
0.5	0.3	426	11.56	958	25.89	3831	103.54
0.4	0.35	6807	11.72	15316	26.27	61263	105.08
0.4	0.3	1696	12.00	3815	26.87	15257	107.44
0.4	0.25	750	12.39	1687	27.66	6747	110.61
0.3	0.25	6661	13.47	14986	29.91	59941	119.64
0.3	0.2	1655	14.14	3724	31.29	14893	125.15
0.3	0.18	1146	14.40	2579	31.84	10313	127.32
0.2	0.16	10056	17.72	22626	38.74	90501	154.97
0.2	0.13	3268	18.62	7352	40.62	29405	162.46
0.2	0.1	1593	19.85	3583	43.17	14330	172.65
0.1	0.08	38386	30.98	86367	66.28	345466	265.13
0.1	0.06	9557	33.46	21502	71.44	86005	285.73
0.1	0.04	4229	36.90	9514	78.63	38056	314.51

Fuente: elaborado por (Tian & Tang, 2014)

CAPÍTULO III:

**APLICACIÓN DEL MODELO DE
RESPUESTA NO ALEATORIZADA,
TRANSVERSAL, PARA ESTUDIAR
EL COMPORTAMIENTO DE LOS
ESTUDIANTES EN LA FCM Y
COMPARACIÓN CON EL MÉTODO
DE ENTREVISTA DIRECTA**

CAPÍTULO III: APLICACIÓN DEL MODELO DE RESPUESTA NO ALEATORIZADA, TRANSVERSAL, PARA ESTUDIAR EL COMPORTAMIENTO DE LOS ESTUDIANTES EN LA FCM Y COMPARACIÓN CON EL MÉTODO DE ENTREVISTA DIRECTA

3.1 MODELO DE RESPUESTA NO ALEATORIZADA TRANSVERSAL

La presente sección contiene la información sobre el marco metodológico que guía la investigación, se detallan los participantes, instrumentos de medición así como el procedimiento utilizado para la recogida y procesamiento de los datos.

3.1.1 ELECCIÓN DE LAS VARIABLES

La variable representa un concepto vital para esta investigación

Las variables de interés se encuentran en las cuatro preguntas sensibles con cuatro preguntas no sensibles no relacionadas, las cuales serán preguntadas a los estudiantes sobre sus comportamientos frente a estas preguntas sensibles.

A continuación en el cuadro 2 mostraremos las preguntas sensibles y las preguntas no sensibles con sus respectivos enunciados.

CUADRO 2: Preguntas y Enunciados utilizados en el MRNA Transversal

PREGUNTA			ENUNCIADOS	
1	Sensible	¿Has copiado de otros estudiantes durante un examen por lo menos una vez?	Nunca he copiado de otros estudiantes durante un examen. (Y=0)	
			He copiado de otros estudiantes durante un examen. (Y=1)	
	No Sensible	¿Nació entre agosto y diciembre?	No nací entre agosto y diciembre. (W=0)	
			Si nací entre agosto y diciembre. (W=1)	
2	Sensible	¿Has Utilizado notas para un examen?	Nunca he utilizado notas en un examen. (Y=0)	
			He utilizado notas en un examen. (Y=1)	
	No Sensible	¿Alguna vez ha interrumpido sus estudios?	Nunca interrumpí mis estudios universitarios. (W=0)	
			He interrumpido mis estudios universitarios. (W=1)	

3	Sensible	¿Has utilizado drogas para mejorar su rendimiento académico?	Nunca he utilizado drogas para mejorar mi rendimiento académico en un examen. (Y=0)	
			He utilizado drogas para mejorar mi rendimiento académico en un examen. (Y=1)	
	No Sensible	¿Utiliza la biblioteca tradicional?	No utilizo la biblioteca tradicional. (W=0)	
			Utilizo la biblioteca tradicional. (W=1)	
4	Sensible	¿Ha presentado documentos con textos adoptados intencionalmente, de otros trabajos?	Nunca he presentado un documento que contiene un párrafo intencionalmente adoptado de otro trabajo. (Y=0)	
			He presentado un documento que contiene un párrafo intencionalmente adoptado de otro trabajo. (Y=1)	
	No Sensible	¿Prefiere estudiar solo?	No me gusta estudiar solo(a). (W=0)	
			Me gusta estudiar solo(a). (W=1)	

3.1.2 PLAN DE MUESTREO DE LA APLICACIÓN

1.-Población. La población objetiva son todos los alumnos matriculados en el Semestre 2015-I de la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos.

3.-Diseño de la muestra. El tamaño de muestra para el MRNA Transversal halló mediante el análisis de potencia del error tipo II ($1-\beta$) siendo $\beta=0.2$, y con un nivel de significancia del error tipo I ($1-\alpha$) siendo $\alpha =0.05$ en el cual se obtuvo una la proporción de alumnos que copian en los exámenes es diferente a 0.26% entonces la aproximación del tamaño de muestra igual a 300, para luego repartirlo proporcionalmente a los cuatro estratos formados (Estrato 1: Ciencias de la Computación, Estrato 2: Estadística, Estrato 3: Investigación Operativa, Estrato 4: Matemática)(Ver cuadro 4). El tamaño de muestra para Método de Entrevista Directa se halló mediante el tamaño de muestra con una proporción conocida de 2.6% y un error de 5.5% en el cual se obtuvo una aproximación de tamaño de muestra igual a 200, para luego repartirlo proporcionalmente a los cuatro estratos formados (Estrato 1: Ciencias de la Computación, Estrato 2: Estadística, Estrato 3: Investigación Operativa, Estrato 4: Matemática) (Ver Cuadro 5).

4.-Tipo de investigación. Este estudio tuvo un enfoque cuantitativo, de tipo descriptivo, corresponde a un diseño no experimental; porque no se ha realizado manipulación de las variables, Transversal por que se recolectan datos en un solo momento (Hernández Sampieri, Fernández Collado, & Baptista, 2006), ya que se estudió el comportamiento de estudiantes con el propósito de describir y estimar la proporción de interés en dicha investigación.

5.-Instrumento de Recolección de datos. Se diseñó un cuestionario para el MRNA Transversal que consistía en 8 preguntas, las cuales las 4 primeras eran sobre datos generales del

alumno como la edad, nombre, escuela y ciclo de estudio, las 4 preguntas finales corresponde a las características que se van a estudiar en esta investigación sobre preguntas sensibles con 4 preguntas no sensibles(Ver anexo 9). Para Método de Entrevista Directa se realizó un cuestionario que consistía en 8 preguntas, las cuales las 4 primeras preguntas son sobre datos generales del alumno, las 4 preguntas finales corresponde a las características mediante preguntas directas (Ver anexo 10).

6.-Trabajo en campo. Para realizar el trabajo de campo contamos con el apoyo de los profesores quienes pusieron a disposición el ambiente, la oficina 211 de FCM de la UNMSM. El proceso de identificación de los alumnos pertenecientes a la muestra, se desarrolló, mostrando en una pizarra acrílica el listado de los alumnos seleccionados para el desarrollo de la investigación, invitándolos a acercarse a la oficina 211 entre las 10am y 8pm horas. La aplicación del cuestionario fue realizado por mi persona (autor de esta investigación), con el apoyo de un profesional de psicología, y compañeros de estudio, quienes realizaban las entrevistas en los horarios que yo no tenía disponibilidad, y por lo cual me siento muy agradecido.

La encuesta se realizó en los meses de mayo y junio del 2015 logrando con éxito el proceso de recolección de datos.

3.1.3 TAMAÑO DE MUESTRA

Esta investigación se realizó en la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor San Marcos, teniendo como población a los alumnos matriculados en el semestre 2015-I, el cual está dividida en 4 Escuelas Académico Profesionales, las cuales son: La Escuela de Matemáticas, Estadística, Investigación Operativa y Computación Científica.

CUADRO 3: Distribución de alumnos matriculados según EAP y sexo de la FCM-UNMSM (Semestre 2015-I)

E.A.P.	Femenino	Masculino	Total General	%
Computación C.	78	223	301	19%
Estadística	132	159	291	30%
Inv. Operativa	206	265	471	31%
Matemática	88	370	458	20%
Total General	504	1017	1521	100%

Fuente: Dirección académica de la FCM-UNMSM

CUADRO 4: Distribución de la muestra de los alumnos matriculados según EAP y sexo de la FCM-UNMSM (Semestre 2015-I)

E.A.P.	Femenino	Masculino	Total General	%
Computación C.	15	44	59	20%
Estadística	26	32	58	19%
Inv. Operativa	41	52	93	31%
Matemática	17	73	90	30%
Total General	99	201	300	100%

Elaboración Propia

CUADRO 5: Distribución de la muestra para MED de los alumnos matriculados según EAP y sexo de la FCM-UNMSM (semestre 2015-I)

E.A.P.	Femenino	Masculino	Total General	%
Computación C.	8	32	40	20%
Estadística	26	12	38	19%
Inv. Operativa	41	52	60	30%
Matemática	17	73	62	31%
Total General	99	201	200	100%

Elaboración Propia

3.2 RESULTADOS DE LA APLICACIÓN DEL MODELO DE RESPUESTA NO ALEATORIZADA, EL MODELO TRANSVERSAL

Se logró entrevistar a 300 estudiantes de la Facultad de Ciencias Matemáticas seleccionadas en la muestra, con una tasa de respuesta de 100%.

La Población entrevistada son alumnos de todas las escuelas de la FCM, de distintos ciclos académicos, ambos sexos, y que tienen edades entre los 16 y 47 años.

3.2.1 RESULTADOS DE LA APLICACIÓN DEL MODELO TRANSVERSAL

1.- Enunciado asociado a la pregunta Sensible de la Pregunta N° 1: Copia de otros estudiantes durante un examen por lo menos una vez.

Proporción

Usamos el enunciado de la pregunta no sensible “Nací entre agosto y diciembre”, con el cual se tiene la probabilidad de que el estudiante naciera entre agosto y diciembre.

$$p = 0.42$$

Por lo tanto, la proporción de los círculos marcados es:

$$\hat{\lambda} = \bar{Y}_C = \frac{1}{n} \sum_{i=1}^n Y_C = \frac{129}{300} = 0.43$$

Con estos valores hallaremos la proporción de interés, en este caso es la proporción de estudiantes que han copiado en un examen por lo menos una vez:

$$\hat{\pi}_c = \frac{p - 1 + \hat{\lambda}}{2p - 1} = \frac{0.42 - 1 + 0.43}{2 * 0.42 - 1} = 0.9$$

La proporción de los alumnos, que copiaron en un examen de otra evaluación es de 0.92.

- **Varianza**

La varianza de la proporción estimada del MRNA viene dada por:

$$\begin{aligned} \widehat{Var}(\hat{\pi}_c) &= \frac{\hat{\pi}_c(1 - \hat{\pi}_c)}{n - 1} + \frac{p(1 - p)}{(n - 1)(2p - 1)^2} \\ &= \frac{0.92(1 - 0.92)}{300 - 1} + \frac{0.42(1 - 0.42)}{(300 - 1)(2 * 0.42 - 1)^2} = 0.0321 \end{aligned}$$

- **Error Estándar**

El error estándar estimada de la proporción del MRNA viene dada por:

$$EE = \sqrt{\widehat{Var}(\hat{\pi}_c)} = \sqrt{0.0321} = 0.179$$

- **Intervalo de confianza**

Al 95% de nivel de confianza el intervalo de confianza (Wald) para π esta dada por,

$$[\hat{\pi}_{c,WL}, \hat{\pi}_{c,WU}] = \left[\hat{\pi}_c - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_c)}, \hat{\pi}_c + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_c)} \right]$$

$$\hat{\pi}_{c,WU} = \hat{\pi}_c + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_c)} = 0.92 + 1.96 \sqrt{0.0321} = 1.271$$

$$\hat{\pi}_{c,WL} = \hat{\pi}_c - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_c)} = 0.92 - 1.96 \sqrt{0.0321} = 0.569$$

$$[\hat{\pi}_{c,WL}, \hat{\pi}_{c,WU}] = [0.569, 1.271]$$

Al 95% de nivel de confianza el intervalo de confianza (Wilson) para π esta dada por

$$[\hat{\pi}_{C,WSI}, \hat{\pi}_{C,WSS}] = \frac{2\hat{\pi}_C + z_* \pm \sqrt{(2\hat{\pi}_C + z_*)^2 - 4(1 + z_*)(\hat{\pi}_C^2 - z_*\partial)}}{2(1 + z_*)}$$

$$z_* \hat{=} \frac{z^2_{\alpha/2}}{n} \quad y \quad \partial \hat{=} \frac{p(1-p)}{(2p-1)^2}$$

$$z_* = \frac{1.96}{300} = 0.0653 \quad y \quad \partial = \frac{0.42(1-0.42)}{(2 * 0.42 - 1)^2} = 9.5156$$

$$= \frac{2 * 0.92 + 0.0653 \pm \sqrt{(2 * 0.92 + 0.0653)^2 - 4(1 + 0.0653)(0.92^2 - 0.0653 * 9.5156)}}{2(1 + 0.0653)}$$

$$[\hat{\pi}_{C,WSI}, \hat{\pi}_{C,WSS}] = [0.5665, 1.2629]$$

2.- Enunciado asociado a la pregunta sensible de la Pregunta N° 2: he utilizado notas en un examen.

- **Proporción**

Usamos el enunciado de la pregunta no sensible “He interrumpido mis estudios universitarios por lo menos una vez”, con el cual se tiene la probabilidad de que el estudiante ha interrumpido sus estudios es:

$$p = 0.19$$

Por lo tanto, la proporción de los círculos marcados es:

$$\hat{\lambda} = \bar{Y}_C = \frac{1}{n} \sum_{i=1}^n Y_C = \frac{143}{300} = 0.477$$

Con estos valores hallaremos la proporción de interés, en este caso es la proporción de estudiantes que han utilizado notas, móviles calculadoras para dar un examen de evaluación:

$$\hat{\pi}_C = \frac{p - 1 + \hat{\lambda}}{2p - 1} = \frac{0.19 - 1 + 0.477}{2 * 0.19 - 1} = 0.538$$

La proporción de estudiantes que usan notas en un examen es de 0.538

- **Varianza**

La varianza estimada de la proporción estimada en el MRNA Transversal viene dada por:

$$\begin{aligned} \widehat{Var}(\hat{\pi}_C) &= \frac{\hat{\pi}_C(1 - \hat{\pi}_C)}{n - 1} + \frac{p(1 - p)}{(n - 1)(2p - 1)^2} = 0.0022 \\ &= \frac{0.538(1 - 0.538)}{300 - 1} + \frac{0.19(1 - 0.19)}{(300 - 1)(2 * 0.19 - 1)^2} = 0.0022 \end{aligned}$$

- **Error Estándar**

El error estándar estimada de la proporción del MRNA viene dada por:

$$EE = \sqrt{\widehat{Var}(\hat{\pi}_C)} = \sqrt{0.0022} = 0.0469$$

- **Intervalo de confianza**

Al 95% de nivel de confianza (Wald) para π está dada por,

$$[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}] = \left[\hat{\pi}_C - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)}, \hat{\pi}_C + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} \right]$$

$$\hat{\pi}_{C,WU} = \hat{\pi}_C + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} = 0.538 + 1.96\sqrt{0.0022} = 0.6293$$

$$\hat{\pi}_{C,WL} = \hat{\pi}_C - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} = 0.538 - 1.96\sqrt{0.0022} = 0.4467$$

$$[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}] = [0.4467, 0.6293]$$

3.- Enunciado asociado a la pregunta Sensible de la Pregunta N° 3: He utilizado drogas para mejorar mi rendimiento académico en un examen o practica por lo menos una vez.

Usamos el enunciado de la pregunta no sensible “Utilizo biblioteca tradicional”, con el cual se tiene la probabilidad de que el estudiante utilizó la biblioteca tradicional es:

$$p = 0.82$$

Por lo tanto, la proporción de los círculos marcados es:

$$\hat{\lambda} = \bar{Y}_C = \frac{1}{n} \sum_{i=1}^n Y_C = \frac{120}{300} = 0.4$$

Con estos valores hallaremos la proporción de interés en este caso es la proporción de estudiantes que han utilizado drogas para mejorar su rendimiento académico

$$\hat{\pi}_C = \frac{p - 1 + \hat{\lambda}}{2p - 1} = \frac{0.82 - 1 + 0.4}{2 * 0.82 - 1} = 0.344$$

La proporción de estudiantes que uso drogas para mejorar su rendimiento académico es 0.344.

- **Varianza**

La varianza estimada de la proporción estimada en el MRNA Transversal está dada por:

$$\widehat{Var}(\hat{\pi}_C) = \frac{\hat{\pi}_C(1 - \hat{\pi}_C)}{n - 1} + \frac{p(1 - p)}{(n - 1)(2p - 1)^2}$$

$$= \frac{0.344(1 - 0.344)}{300 - 1} + \frac{0.82(1 - 0.82)}{(300 - 1)(2 * 0.82 - 1)^2} = 0.00196$$

- **Error Estándar**

El error estándar estimada de la proporción del mRNA viene dada por:

$$EE = \sqrt{\widehat{Var}(\hat{\pi}_C)} = \sqrt{0.00196} = 0.0443$$

- **Intervalo de confianza**

Al 0.95% del intervalo de confianza (Wald) para π esta dada por,

$$[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}] = \left[\hat{\pi}_C - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)}, \hat{\pi}_C + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} \right]$$

$$\hat{\pi}_{C,WU} = \hat{\pi}_C + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} = 0.344 + 1.96 \sqrt{0.00196} = 0.4308$$

$$\hat{\pi}_{C,WL} = \hat{\pi}_C - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} = 0.344 - 1.96 \sqrt{0.00196} = 0.2572$$

$$[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}] = [0.2572, 0.4308]$$

4.- Enunciado asociado a la pregunta Sensible de la Pregunta N° 4: he presentado un documento que contiene párrafo intencionalmente adoptado del trabajo de otra persona.

Usamos el enunciado de la pregunta no sensible “Me gusta estudiar solo”, con el cual se tiene la probabilidad de que el estudiante le guste estudiar solo es:

$$p = 0.22$$

Por lo tanto, la proporción de los círculos marcados es:

$$\hat{\lambda} = \bar{Y}_C = \frac{1}{n} \sum_{i=1}^n Y_C = \frac{122}{300} = 0.407$$

Con estos valores hallaremos la proporción de interés, en este caso la proporción de estudiantes que han presentado un documento que contiene un párrafo intencionalmente adoptado del trabajo de otra persona es:

$$\hat{\pi}_C = \frac{p - 1 + \hat{\lambda}}{2p - 1} = \frac{0.22 - 1 + 0.41}{2 * 0.22 - 1} = 0.667$$

La proporción de estudiantes que uso un documento intencionalmente adoptados de otro trabajo es 0.667.

- **Varianza**

La varianza estimada de MRNA transversal es dada por

$$\begin{aligned} \widehat{Var}(\hat{\pi}_C) &= \frac{\hat{\pi}_C(1 - \hat{\pi}_C)}{n - 1} + \frac{p(1 - p)}{(n - 1)(2p - 1)^2} = 0.0026 \\ &= \frac{0.667(1 - 0.667)}{300 - 1} + \frac{0.22(1 - 0.22)}{(300 - 1)(2 * 0.22 - 1)^2} = 0.0026 \end{aligned}$$

- **Error Estándar**

El error estándar estimada de la proporción del MRNA viene dada por:

$$EE = \sqrt{\widehat{Var}(\hat{\pi}_C)} = \sqrt{0.0026} = 0.051$$

- **Intervalo de confianza**

Al 95% de nivel de confianza (Wald) para π está dada por,

$$[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}] = \left[\hat{\pi}_C - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)}, \hat{\pi}_C + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} \right]$$

$$\hat{\pi}_{C,WU} = \hat{\pi}_C + z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} = 0.667 + 1.96\sqrt{0.0026} = 0.7660$$

$$\hat{\pi}_{C,WL} = \hat{\pi}_C - z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\pi}_C)} = 0.667 - 1.96\sqrt{0.0026} = 0.5680$$

$$[\hat{\pi}_{C,WL}, \hat{\pi}_{C,WU}] = [0.5680, 0.7660]$$

3.2.2 RESULTADOS DE LA APLICACIÓN DEL MÉTODO DE ENTREVISTA DIRECTA

1.- Pregunta N° 1: ¿Has copiado de otros estudiantes durante un examen (practica) por lo menos una vez?

- **Estimador de la proporción:**

Sea p' el estimador del método de entrevista directa en la pregunta N° 1 a los alumnos matriculados de la Facultad de Ciencias Matemáticas:

$$\hat{\pi}_D = p' = \frac{n'}{n} = \frac{107}{200} = 0.535$$

Donde n' es el número de entrevistados que marcaron la respuesta “si”

Del resultado nos indica que el 53,5% de alumnos ha copiado por lo menos una vez en un examen

- **La varianza**

$$Var(\hat{\pi}_D) = \frac{p'(1-p')}{n-1} \left(1 - \frac{n}{N}\right) = \frac{0.535(1-0.535)}{200-1} \left(1 - \frac{200}{1521}\right)$$

$$Var(\hat{\pi}_D) = 0.0109$$

- **La desviación estándar**

$$DE = \sqrt{\text{Var}(\hat{\pi}_D)} = 0.033$$

- **Intervalo de confianza**

$$\hat{\pi}_D \pm Z_{1-\alpha/2} \sqrt{1 - \frac{n}{N} \sqrt{\frac{pq}{n-1}} + \frac{1}{2n}}$$

Se tiene

$$Ls = 0.535 + 1.96 \sqrt{1 - \frac{200}{1521} \sqrt{\frac{0.535(1 - 0.535)}{200 - 1}} + \frac{1}{2(200)}} = 0.602$$

$$Li = 0.535 - 1.96 \sqrt{1 - \frac{200}{1521} \sqrt{\frac{0.535(1 - 0.535)}{200 - 1}} + \frac{1}{2(200)}} = 0.473$$

$$\hat{\pi}_D = [0.473; 0.602]$$

2.- Pregunta N° 2: ¿Has utilizado notas en un examen (practica) (incluyendo notas en los teléfonos móviles, calculadoras u otras similares)?

- **Estimador de la proporción:**

Sea p' el estimador del método de entrevista directa en la pregunta N° 2, a los alumnos matriculados de la Facultad de Ciencias Matemáticas:

$$\hat{\pi}_D = p' = \frac{n'}{n} = \frac{65}{200} = 0.325$$

Donde n' es la cantidad de entrevistados que respondieron la respuesta “si”

Del resultado nos indica que el 32.5% de alumnos han utilizado notas alguna vez en un examen.

- **La varianza**

$$\begin{aligned} \text{Var}(\hat{\pi}_D) &= \frac{p'(1-p')}{n-1} \left(1 - \frac{n}{N}\right) = \frac{0.325(1-0.325)}{200-1} \left(1 - \frac{200}{1521}\right) \\ &= 0.0064 \end{aligned}$$

- **La desviación estándar**

$$DE = \sqrt{\text{Var}(\hat{\pi}_D)} = 0.08$$

- **Intervalo de confianza**

$$\hat{\pi}_D \pm Z_{1-\alpha/2} \sqrt{1 - \frac{n}{N} \sqrt{\frac{pq}{n-1}} + \frac{1}{2n}}$$

Se tiene

$$Ls = 0.325 + 1.96 \sqrt{1 - \frac{200}{1521} \sqrt{\frac{0.325(1-0.325)}{200-1}} + \frac{1}{2(200)}} = 0.388$$

$$Li = 0.325 - 1.96 \sqrt{1 - \frac{200}{1521} \sqrt{\frac{0.325(1-0.325)}{200-1}} + \frac{1}{2(200)}} = 0.267$$

$$\hat{\pi}_D = [0.267; 0.388]$$

3.- Pregunta N° 3: ¿Has utilizado drogas para mejorar el rendimiento académico en un examen o practica por lo menos una vez?

- **Estimador de la proporción:**

Sea p' el estimador del método de entrevista directa de la pregunta N° 3, a los estudiantes de la Facultad de Ciencias Matemáticas:

$$\hat{\pi}_D = p' = \frac{n'}{n} = \frac{3}{200} = 0.015$$

Donde n' es la cantidad de entrevistados que dieron la respuesta "si"

El resultado nos indica que el 1,5% de alumnos han utilizado droga alguna vez para mejorar su rendimiento académico.

- **La varianza**

$$\begin{aligned} \text{Var}(\hat{\pi}_D) &= \frac{p'(1-p')}{n-1} \left(1 - \frac{n}{N}\right) = \frac{0.015(1-0.015)}{200-1} \left(1 - \frac{200}{1521}\right) \\ &= 0.0004 \end{aligned}$$

- **La desviación estándar**

$$DE = \sqrt{\text{Var}(\hat{\pi}_D)} = 0.02$$

- **Intervalo de confianza**

$$\hat{\pi}_D \pm Z_{1-\alpha/2} \sqrt{1 - \frac{n}{N} \sqrt{\frac{pq}{n-1}} + \frac{1}{2n}}$$

Se tiene

$$Ls = 0.015 + 1.96 \sqrt{1 - \frac{200}{1521} \sqrt{\frac{0.015(1-0.015)}{200-1}} + \frac{1}{2(200)}} = 0.033$$

$$Li = 0.015 - 1.96 \sqrt{1 - \frac{200}{1521} \sqrt{\frac{0.015(1-0.015)}{200-1}} + \frac{1}{2(200)}} = 0.002$$

$$\hat{\pi}_D = [0.002; 0.033]$$

4.- Pregunta N° 4: ¿Has presentado un documento que contiene un párrafo intencionalmente adoptado del trabajo de otra persona por lo menos una vez?

- **Estimador de la proporción:**

Sea p' el estimador del método de entrevista directa de la pregunta N° 4, a los alumnos matriculados de la Facultad de Ciencias Matemáticas:

$$\hat{\pi}_D = p' = \frac{n'}{n} = \frac{54}{200} = 0.27$$

Donde n' es la cantidad de entrevistados que marcaron la respuesta “si”

El resultado nos indica que el 27% de alumnos han presentado alguna vez en un documento que contiene un párrafo intencionalmente adoptado de otro trabajo

- **La varianza**

$$Var(\hat{\pi}_D) = \frac{p'(1-p')}{n-1} \left(1 - \frac{n}{N}\right) = \frac{0.27(1-0.27)}{200-1} \left(1 - \frac{200}{1521}\right)$$

$$Var(\hat{\pi}_D) = 0.0057$$

- **La desviación estándar**

$$DE = \sqrt{Var(\hat{\pi}_D)} = 0.0755$$

- **Intervalo de confianza**

$$\hat{\pi}_D \pm Z_{1-\alpha/2} \sqrt{1 - \frac{n}{N} \sqrt{\frac{pq}{n-1}} + \frac{1}{2n}}$$

Se tiene

$$Ls = 0.27 + 1.96 \sqrt{1 - \frac{200}{1521} \sqrt{\frac{0.27(1-0.27)}{200-1}} + \frac{1}{2(200)}} = 0.330$$

$$Li = 0.27 - 1.96 \sqrt{1 - \frac{200}{1521}} \sqrt{\frac{0.27(1 - 0.27)}{200 - 1}} + \frac{1}{2(200)} = 0.215$$

$$\hat{\pi}_D = [0.215; 0.330]$$

3.2.3 COMPARACIÓN ENTRE EL MODELO TRANSVERSAL Y EL MÉTODO DE ENTREVISTA DIRECTA

CUADRO 6: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación del porcentaje de estudiantes que han copiado en los exámenes

Método Empleado	Copia en los exámenes	Error Estándar
MRNA Transversal	92%	0.1792
Entrevista Directa	53.5%	0.1044

Elaboración propia

CUADRO 7: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación del porcentaje de estudiantes que han utilizados notas en los exámenes

Método Empleado	Uso de notas en un examen	Error Estándar
MRNA Transversal	53.8%	0.0469
Entrevista Directa	32.5%	0.08

Elaboración propia

CUADRO 8: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación del porcentaje de estudiantes que han utilizado droga para mejorar el rendimiento académico

Método Empleado	Uso de drogas para mejorar el rendimiento académico	Error Estándar
MRNA Transversal	34.4%	0.0443
Entrevista Directa	1.5%	0.02

Elaboración propia

CUADRO 9: Modelo de Respuesta no Aleatorizada Transversal y Método de Entrevista Directa según la estimación del porcentaje de estudiantes que han presentado un trabajo intencionalmente adoptado del trabajo de otra persona

Método Empleado	Presentación de un trabajo intencionalmente adoptado del trabajo de otra persona	Error Estándar
MRNA Transversal	66.7%	0.0509
Entrevista Directa	27%	0.0755

Elaboración propia

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES Y RECOMENDACIONES

1. El estudio del Modelo de Respuesta No Aleatorizada Transversal permitió poner en práctica una técnica de muestreo, que resulta eficiente en encuestas con preguntas sensibles, también llamadas comprometedoras, ya que revela diferencias entre las respuestas del Método Entrevista Directa y el MRNA Transversal, siendo los del MRNA Transversal los que dan resultados más veraces como veremos a continuación.
2. Respecto a las preguntas evaluadas, la estimación del “porcentaje de estudiantes que han copiado en los exámenes por lo menos alguna vez” utilizando el MRNA Transversal es de 92%; mientras que en el Método de Entrevista Directa es de 53.5%. Observamos que casi la mayoría de estudiantes, alguna vez han copiado en los exámenes. Además, aunque las proporciones de ambos métodos son distantes, también son muy elevadas, lo que revela que esta práctica se está volviendo común entre los estudiantes de la Facultad de Ciencias Matemáticas.
3. La estimación del “porcentaje de estudiantes que usaron notas, móviles, calculadoras para rendir un examen” utilizando el MRNA Transversal es de 53.8%; mientras en el Método de Entrevista Directa es de 32.5%.
4. La estimación del “porcentaje de estudiantes que usaron drogas para mejorar su rendimiento académico” utilizando el MRNA Transversal es de 34.4%; mientras en el Método de Entrevista Directa es de 1.5%, vemos claramente que con el modelo transversal nos indica un mayor porcentaje en este caso, observamos que en la entrevista directa los alumnos tienden a dar respuestas falsas, mientras que en el modelo transversal la respuesta muestra una realidad más cierta.
5. La estimación del “porcentaje de presentación de un trabajo, adoptando intencionalmente párrafos de otros estudios, utilizando el MRNA Transversal es de 66.7%; mientras en el Método de Entrevista Directa es de 27%, obteniéndose un mayor porcentaje con el modelo transversal.

6. Se recomienda aplicar este MRNA transversal en otras facultades para poder comparar la efectividad de este modelo también utilizar preguntas más “comprometedoras” dependiendo del tema de estudio.

REFERENCIAS BIBLIOGRÁFICAS

REFERENCIA BIBLIOGRÁFICA

- Coutts, E., Jann, B., Krumpal, I., & Naeher, A. F. (2011). *Plagiarism in student papers: Prevalence estimates using special techniques for sensitive questions* (Vol. 231). Alemania: Journal of Economics and Statistics.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista, L. P. (2006). *Metodología de la investigación* (Vol. 4). Mexico: McGraw-Hill/Interamericana Editores, S.A.
- Höglinger, M., Jann, B., & Diekmann, A. (2014). Sensitive Questions in Online Surveys: An Experimental Evaluation of the Randomized Response Technique and the Crosswise Model. *University of Bern Social Sciences Working Paper*(9).
- Jann, B., Jerke, J., & Krumpal, I. (2011). Asking Sensitive Questions Using the Crosswise Model. *Oxford University*, 1-9.
- Khosravi, A., Mousavi, A. M., Chaman, R., Khosravi, F., Amari, M., & Shamsipour, M. (Junio de 2015). Crosswise Model to Assess Sensitive Issues: A Study on Prevalence of Drug Abuse Among University Students of Iran. *Zahedan University of Medical Sciences*, 1-5.
- Korndorfer, M., Krumpal, I., & Schmukle, S. C. (2014). Measuring and explaining tax evasion: Improving self-reports using the crosswise model. *Journal of Economic Psychology*, 45, 18-32.
- Kundt, T. (2014). Applying "Benford's law" to the Crosswise Model: Findings from an online survey on tax evasion.
- Raj, D. (1968). *Teoria del muestreo*. New York: McGraw-Hill.
- Schlosser Montes, K. (2014). *La percepción del plagio académico de los estudiantes y docentes de las Facultades de Arquitectura, Derecho e Ingeniería en la Universidad Rafael Landívar*. Tesis de Posgrado, Universidad Rafael Landívar, Guatemala.
- Tan, M. T., Tian, G. L., & Tang, M. L. (2009). Sample Surveys With Sensitive Questions: A Nonrandomized Response Approach. *The American Statistician*, 63, 9-16.
- Tian, G. L., & Tang, M. L. (2014). *Incomplete Categorical Data Design: Non-Randomized Response Techniques for Sensitive Questions in Surveys*. Hong Kong, China: Taylor & Francis Group, an informa business.

- Tian, G. L., Yu, J. W., Tang, M. L., & Geng, Z. (2007). A new non-randomized model for analysing sensitive questions with binary outcomes. *Statistics in Medicine*, 26, 4238-4252.
- Trujillo, L., & Gonzàles, L. M. (Diciembre de 2012). Preguntas sensibles en encuestas y metodologias alternativas para asegurar la veracidad y fidelidad de las respuestas. *ib*, 56.
- Ulrich, R., Schröter, H., Striegel, H., & Simon, P. (2012). Asking Sensitive Questions: A Statistical Power Analysis of Randomized Response Models. *American Psychological Association*, 6-8.
- Warner, S. L. (1965). *Randomized response: a survey technique for eliminating evasive answer bias*. *Journal of Applied Psychology* (Vol. 60).
- Wu, Y. J., Liang, T. G., & Lai, T. M. (2008). Two new models for survey sampling with sensitive characteristic: desing and analysis. *Metrica*, 251-263.

ANEXO

ANEXO 1: Estimación de la proporción λ mediante el método de Máxima Verosimilitud

La función de verosimilitud es proporcional a $\lambda^{n'}(1 - \lambda)^{n-n'}$ de modo que el EMV de λ está dada por $\hat{\lambda} = n' / n$.

Demostración

sea

$$LZ=L(\lambda|z) = \lambda^{n'}(1 - \lambda)^{n-n'}$$

Aplicando logaritmo a la ecuación se tiene

$$\begin{aligned}Ln(L(\lambda|z)) &= Ln(\lambda^{n'}(1 - \lambda)^{n-n'}) \\ n'Ln(\lambda) &+ (n - n')Ln(1 - \lambda)\end{aligned}$$

Derivando con respecto a λ

$$\frac{\partial(Ln(L(\lambda|z)))}{\partial(\lambda)} = \frac{\partial(n'Ln(\lambda) + (n - n')Ln(1 - \lambda))}{\partial(\lambda)}$$

$$\frac{n'}{\lambda} + \frac{n - n'}{1 - \lambda} = 0$$

$$\frac{n'}{\lambda} + \frac{n'}{1 - \lambda} - \frac{n}{\lambda} = 0$$

$$n' \left(\frac{1}{\lambda} + \frac{1}{1 - \lambda} \right) = \frac{n}{1 - \lambda}$$

$$n' \left(\frac{1 - \lambda + \lambda}{\lambda(1 - \lambda)} \right) = \frac{n}{1 - \lambda}$$

$$\frac{n'}{\lambda} = n$$

Entonces tenemos que

$$\lambda = \frac{n'}{n}$$

ANEXO 2: Estimación de la proporción π mediante el método de Máxima Verosimilitud en el MRNA transversal

La función de Máxima Verosimilitud para el MRNA Transversal es:

$$L_C(\pi|Y_{obs}) = \prod_{i=1}^n \{(1-\pi)(1-p) + \pi p\}^{y_{i,c}} \{\pi(1-p) + (1-\pi)p\}^{1-y_{i,c}}$$

$$L_C(\pi|Y_{obs}) = \{(1-\pi)(1-p) + \pi p\}^{\sum y_{i,c}} \{\pi(1-p) + (1-\pi)p\}^{\sum (1-y_{i,c})}$$

Aplicando logaritmo a la ecuación se tiene:

$$\ln(L_C(\pi|Y_{obs})) = \ln\{(1-\pi)(1-p) + \pi p\} \sum_i^n y_{i,c} + \ln\{\pi(1-p) + (1-\pi)p\} \sum_i^n (1-y_{i,c})$$

$$\frac{\partial \ln(L_C(\pi|Y_{obs}))}{\partial \pi} = \frac{(2p-1) \sum_i^n y_{i,c}}{(1-\pi)(1-p) + \pi p} + \frac{(1-2p) \sum_i^n (1-y_{i,c})}{\pi(1-p) + (1-\pi)p}$$

$$\frac{\partial \ln(L_C(\pi|Y_{obs}))}{\partial \pi} = \frac{(2p-1) \sum_i^n y_{i,c}}{(1-\pi)(1-p) + \pi p} - \frac{(2p-1) \sum_i^n (1-y_{i,c})}{\pi(1-p) + (1-\pi)p}$$

$$\frac{\partial \ln(L_C(\pi|Y_{obs}))}{\partial \pi} =$$

$$= \frac{(2p-1)\{\pi(1-p) + (1-\pi)p\} \sum_i^n y_{i,c} - [(1-\pi)(1-p) + \pi p] \sum_i^n (1-y_{i,c})}{\{(1-\pi)(1-p) + \pi p\}\{\pi(1-p) + (1-\pi)p\}}$$

$$\frac{\partial \ln(L_C(\pi|Y_{obs}))}{\partial \pi} = \frac{(2p-1)\{\sum_i^n y_{i,c} - (1-p + 2\pi p)n\}}{\{(1-\pi)(1-p) + \pi p\}\{\pi(1-p) + (1-\pi)p\}} = 0$$

$(2p-1)\{\sum_i^n y_{i,c} - (1-p + 2\pi p)n\} = 0$, dividimos entre n, entonces tenemos

$$\frac{\sum_i^n y_{i,c}}{n} - (1-p + 2\pi p) = 0$$

$$\pi = \frac{y' - 1 + p}{2p - 1}$$

ANEXO 3: Procedimiento de la estimación de la varianza del estimador π

Se da un estimador de la proporción y sabemos que n' tiene una distribución binomial

$$\lambda = (1 - \pi)(1 - p) + \pi p$$

$$\hat{\pi}_C = \frac{p - 1 + \hat{\lambda}}{2p - 1} \quad \text{donde } \hat{\lambda} = \frac{n'}{n}; n' = \text{Binomial}(n, \lambda),$$

Prueba

$\hat{\pi}_C$ es un estimador insesgado de π .

Esperanza de $\hat{\pi}_C$

$$E(\hat{\pi}_C) = E\left[\frac{\hat{\lambda} - (1 - p)}{2p - 1}\right]$$

$$E(\hat{\pi}_C) = \frac{E(\hat{\lambda}) - (1 - p)}{2p - 1}$$

$$E(\hat{\pi}_C) = \frac{\frac{E(n')}{n} - (1 - p)}{2p - 1}$$

$$E(\hat{\pi}_C) = \frac{n\lambda/n - (1 - p)}{2p - 1}$$

$$E(\hat{\pi}_C) = \frac{(1 - \pi)(1 - p) + \pi p - 1 + p}{2p - 1}$$

$$E(\hat{\pi}_C) = \frac{1 - p - \pi + \pi p + \pi p - 1 + p}{2p - 1}$$

$$E(\hat{\pi}_C) = \frac{\pi(2p - 1)}{2p - 1}$$

$$E(\hat{\pi}_C) = \pi$$

$\hat{\pi}_C$ es insesgado de π

Varianza de $\hat{\pi}_C$

$$V(\hat{\pi}_C) = V\left[\frac{\hat{\lambda} - (1 - p)}{(2p - 1)}\right]$$

$$V(\hat{\pi}_C) = \frac{1}{(2p - 1)^2} V(\hat{\lambda})$$

$$V(\hat{\pi}_C) = \frac{1}{n^2(2p - 1)^2} V(n')$$

$$V(\hat{\pi}_C) = \frac{1}{n^2(2p - 1)^2} n\lambda(1 - \lambda)$$

$$V(\hat{\pi}_C) = \frac{[(1 - \pi)(1 - p) + \pi p][1 - (1 - \pi)(1 - p) - \pi p]}{n(2p - 1)^2}$$

$$V(\hat{\pi}_C) = \frac{[(1 - p - \pi + \pi p + \pi p)(-\pi p + \pi + p - \pi p)]}{n(2p - 1)^2}$$

$$V(\hat{\pi}_C) = \frac{[1 - p + \pi(2p - 1)][p - \pi(2p - 1)]}{n(2p - 1)^2}$$

$$V(\hat{\pi}_C) = \frac{(1 - p)p + \pi(2p - 1)^2 - \pi^2(2p - 1)^2}{n(2p - 1)^2}$$

$$V(\hat{\pi}_C) = \frac{\pi(1 - \pi)}{n} + \frac{p(1 - p)}{n(2p - 1)^2}$$

ANEXO 4: Resultado de la encuesta realizada a los estudiantes matriculados en el semestre 2015- I de la Facultad Ciencias Matemáticas de la Universidad Mayor de San Marcos mediante el MRNA Transversal

Anexo4: Resultado de la encuesta del Modelo transversal

Nro.	E.A.P.	PREG 5 CIRCULO	PREG 5 CUADRA	PREG 6 CIRCULO	PREG 6 CUADRAD	PREG 7 CIRCULO	PREG 7 CUADRAD	PREG 8 CIRCULO	PREG 8 CUADRAD
1	Matemática	0		0		0		0	
2	Matemática	0			1		1	0	
3	Matemática	0		0		0		0	
4	Inv.operativa		1		1	0			1
5	Inv.operativa		1		1		1		0
6	Inv.operativa		1	0			1		1
7	Inv.operativa	0			1	0		0	
8	Inv.operativa		1		1		1		1
9	Matemática	0		0		0			1
10	Matemática	0			1	0			1
11	Matemática		1	0			1		1
12	Matemática		1		1	0		0	
13	Matemática		1		1	0			1
14	Matemática	0			1	0			1
15	Inv.operativa		1		1	0			1
16	Matemática		1		1	0			1
17	Matemática	0		0			1		1
18	Matemática	0		0			1		1
19	Matemática		1	0			1	0	
20	Matemática	0		0			1	0	
21	computación	0			1		1		1
22	computación	0		0			1		1
23	Matemática		1	0			1	0	

24	Matemática	0			1	0			1
25	computación		1	0		0			1
26	computación		1		1	0		0	
27	computación		1		1		1		1
28	Inv.operativa		1		1		1		1
29	Inv.operativa	0			1		1		1
30	Inv.operativa	0		0			1	0	
31	Inv.operativa	0			1	0		0	
32	Inv.operativa	0			1		1	0	
33	Inv.operativa	0			1		1	0	
34	Inv.operativa		1	0			1	0	
35	Inv.operativa		1	0			1		1
36	Computación		1		1		1		1
37	Matemática		1	0			1	0	
38	Matemática	0		0			1		1
39	Matemática		1	0			1		1
40	Matemática	0		0			1	0	
41	computación	0		0		0		0	
42	Matemática	0		0			1		1
43	Matemática	0		0			1		1
44	Matemática		1		1		1	0	
45	computación	0		0			1	0	
46	Matemática	0			1	0			1
47	Matemática		1		1		1		1
48	Matemática		1		1	0			1
49	Matemática	0			1		1		1
50	Matemática	0			1		1	0	
51	computación		1		1		1		1

52	Matematica	0			1		1	0	
53	computación	0			1	0			1
54	computación		1		1		1		1
55	computación		1		1	0			1
56	computación	0			1	0		0	
57	computación		1	0			1	0	
58	computación		1		1		1		1
59	Inv.operativa		1		1	0			1
60	Inv.operativa		1		1		1		1
61	Inv.operativa		1		1	0		0	
62	Inv.operativa		1		1		1	0	
63	Inv.operativa		1		1		1	0	
64	Inv.operativa	0			1		1		1
65	Inv.operativa	0			1		1		1
66	Inv.operativa		1	0		0		0	
67	Matematica		1	0			1	0	
68	Matematica	0			1		1		1
69	computación	0			1	0			1
70	Matemática		1	0			1		1
71	Matemática		1	0			1		1
72	Matemática	0		0			1		1
73	Matematica		1		1		1		1
74	Matematica	0		0		0		0	
75	Matematica	0			1	0		0	
76	Matematica	0			1		1	0	
77	Matematica	0		0		0			1
78	Matematica	0		0			1	0	
79	Inv.operativa	0			1		1		1
80	Inv.operativa	0			1		1	0	
81	Inv.operativa		1	0			1		1

82	computación		1	0		0			1
83	computación		1		1		1	0	
84	computación		1	0			1		1
85	computación		1		1		1		1
86	computación		1	0		0			1
87	computación		1	0			1		1
88	Estadística		1	0			1		1
89	Inv.operativa	0			1		1	0	
90	computación	0			1		1		1
91	Inv.operativa		1	0		0		0	
92	Inv.operativa	0		0			1	0	
93	Inv.operativa		1	0		0		0	
94	Inv.operativa		1	0			1		1
95	Inv.operativa		1		1	0		0	
96	Inv.operativa		1	0		0		0	
97	Matemática		1		1	0			1
98	Matemática	0		0			1	0	
99	Matemática	0		0			1	0	
100	Matemática	0			1		1		1
101	Matemática		1	0			1	0	
102	Matemática	0		0			1	0	
103	Matemática		1		1		1	0	
104	Matemática		1		1		1		1
105	Matemática	0		0			1		1
106	Matemática		1	0		0		0	
107	Matemática		1		1	0			1
108	Matemática		1	0		0		0	
109	Matemática		1		1		1	0	
110	Matemática		1	0			1	0	
111	Matemática	0		0			1	0	

112	Matemática		1	0			1	0	
113	Matemática		1	0		0		0	
114	Matemática		1	0			1	0	
115	Matemática	0		0			1		1
116	Matemática	0		0		0			1
117	Matemática	0		0		0			1
118	Matemática	0			1	0			1
119	Matemática		1	0			1	0	
120	Matemática	0		0			1	0	
121	Matemática		1		1		1		1
122	Matemática		1	0			1	0	
123	Matemática		1	0			1		1
124	Matemática	0			1		1	0	
125	Computación	0		0		0			1
126	Computación		1		1	0			1
127	Computación		1		1	0			1
128	Computación		1		1		1	0	
129	Computación	0		0		0			1
130	Computación		1	0			1		1
131	Computación	0		0			1	0	
132	computación	0		0			1	0	
133	computación	0		0			1	0	
134	computación	0		0			1		1
135	Estadística	0		0		0		0	
136	Estadística	0			1		1	0	
137	Estadística	0		0			1		1
138	Estadística		1	0		0			1
139	Estadística		1	0			1		1
140	Estadística	0			1	0		0	
141	Estadística	0		0			1	0	

142	Estadística		1	0		0			1
143	Estadística	0			1	0			1
144	Estadística	0			1	0		0	
145	Estadística		1		1		1		1
146	Estadística		1	0		0		0	
147	Estadística	0			1	0			1
148	Estadística	0		0		0			1
149	Estadística		1	0			1	0	
150	Estadística	0		0			1		1
151	Estadística	0			1		1	0	
152	Estadística	0			1		1		1
153	Estadística	0		0			1		1
154	Estadística		1		1		1		1
155	Estadística	0		0		0			1
156	Estadística		1		1		1		1
157	Estadística		1	0			1		1
158	Estadística	0		0			1		1
159	Estadística		1		1		1		1
160	Estadística	0			1	0		0	
161	Estadística	0			1	0		0	
162	Estadística		1		1		1	0	
163	Estadística		1	0			1		1
164	Estadística		1		1		1		1
165	Estadística	0			1	0			1
166	Estadística		1	0		0			1
167	Estadística		1		1		1		1
168	Estadística		1	0		0			1
169	Estadística	0			1		1	0	
170	Estadística	0		0		0			1
171	Estadística	0		0			1		1

172	Estadística		1	0			1		1
173	Estadística		1	0			1		1
174	Inv.operativa		1	0		0			1
175	Inv.operativa		1		1	0			1
176	Inv.operativa		1		1		1		1
177	Inv.operativa		1		1	0			1
178	Inv.operativa		1		1		1		1
179	Inv.operativa		1		1		1		1
180	Inv.operativa		1		1		1	0	
181	Inv.operativa		1		1		1		1
182	Inv.operativa	0		0		0		0	
183	Inv.operativa		1	0		0		0	
184	Inv.operativa		1		1		1	0	
185	Inv.operativa	0		0			1		1
186	Inv.operativa		1		1		1		1
187	Inv.operativa	0			1	0		0	
188	Inv.operativa		1		1	0			1
189	Computación		1	1			1		1
190	Computación	0		0		0		0	
191	Computación		1	0		0			1
192	Inv.operativa		1		1		1	0	
193	Inv.operativa	0		0			1		1
194	Inv.operativa	0		0		0		0	
195	Inv.operativa	0		0		0		0	
196	Inv.operativa		1		1		1	0	
197	Inv.operativa	0			1	0			1
198	Inv.operativa		1		1	0			1
199	Inv.operativa	0			1	0		0	

200	Inv.operativa	0		0			1		1
201	Inv.operativa	0		0		0		0	
202	Inv.operativa	0		0			1	0	
203	Inv.operativa		1	0		0		0	
204	Inv.operativa		1		1		1		1
205	Inv.operativa		1		1	0			1
206	Inv.operativa		1		1		1		1
207	Inv.operativa	0			1	0			1
208	Inv.operativa		1		1		1		1
209	Inv.operativa		1	0			1		1
210	Inv.operativa		1	0		0			1
211	Matemática	0		0		0		0	
212	Matemática		1	0		0			1
213	Matemática		1	0			1		1
214	Matemática	0			1	0			1
215	Matemática	0			1	0		0	
216	Computación		1	0			1	0	
217	Computación		1	0			1		1
218	Computación	0			1	0		0	
219	Computación		1		1	0		0	
220	Computación		1	0			1	0	
221	Computación	0		0			1		1
222	Computación		1	0			1		1
223	Computación		1		1		1		1
224	Estadística	0			1		1	0	
225	Estadística		1		1	0			1
226	Estadística	0		0		0		0	
227	Estadística	0			1		1		1
228	Matemática		1		1	0		0	
229	Matemática		1	0			1	0	

230	Matematica		1		1	0			1
231	Matematica		1		1		1		1
232	Matematica		1	0			1		1
233	Matematica		1		1	0		0	
234	Matematica		1		1	0		0	
235	Matematica		1		0		0		0
236	Matematica		1		1	0			1
237	Matematica		1		1	0			1
238	Matematica	0		0		0		0	
239	Matematica	0		0			1		1
240	Inv.operativa		1		1		1		1
241	Inv.operativa		1		1		1		1
242	Inv.operativa		1		1		1		1
243	Inv.operativa		1		1		1		1
244	Inv.operativa		1		1	0		0	
245	Inv.operativa		1		1	0		0	
246	Inv.operativa	0			1		1		1
247	Inv.operativa		1		1		1	0	
248	Inv.operativa		1		1	0			1
249	Inv.operativa		1		1	0			1
250	Estadística		1	0		0		0	
251	Matematica		1		1		1		1
252	Inv.operativa		1	0		0			1
253	Estadística	0			1	0			1
254	Estadística		1	0			1		1
255	Inv.operativa	0		0			1		1
256	Matematica		1	0			1	0	
257	Inv.operativa	0			1		1	0	
258	Inv.operativa	0			1		1		1
259	Inv.operativa	0			1		1		1

260	Inv.operativa	0		0			1	0	
261	Estadística	0		0			1	0	
262	Estadística	0		0			1	0	
263	Estadística	0			1		1		1
264	Computación	0			1	0			1
265	Computación		1	0			1		1
266	Estadística		1	0		0			1
267	Estadística		1		1		1		1
268	Estadística		1	0			1		1
269	Estadística		1		1	0		0	
270	Matemática		1		1		1	0	
271	Matemática		1		1		1		1
272	Inv.operativa		1	0			1	0	
273	Inv.operativa	0		0		0			1
274	Computación		1		1	0			1
275	Computación	0		0			1		1
276	Computación		1		1	0			1
277	Computación		1		1	0		0	
278	Computación		1		1	0			1
279	Computación		1	0			1		1
280	Computación		1		1	0			1
281	Computación	0		0		0			1
282	Computación	0		0		0			1
283	Inv.operativa		1		1		1		1
284	Inv.operativa	0		0			1	0	
285	Inv.operativa		1		1		1		1
286	Inv.operativa		1	0		0		0	
287	Inv.operativa		1		1		1		1
288	Inv.operativa	0			1	0			1
289	Inv.operativa		1		1	0			1

290	Computación		1	0			1	0	
291	Computación	0			1		1	0	
292	Computación	0			1	0			1
293	computación		1	0		0		0	
294	Estadística	0			1		1	0	
295	Estadística	0			1		1		1
296	Estadística	0		0			1	0	
297	Estadística		1	0			1		1
298	Inv.operativa		1		1		1		1
299	Inv.operativa		1	0			1		1
300	Inv.operativa		1		1		1		1

ANEXO 5: Resultados de la estimación de parámetros de la encuesta realizada en los estudiantes matriculados en el semestre 2015-I de la Facultad de Ciencias Matemáticas de la Universidad Mayor de San Marcos mediante MRNA Transversal

Los resultados fueron obtenidos mediante el programa R project a continuación se muestran los códigos.

Extraemos la base de datos mediante

```
data <- read.csv("F:/tesis/data.txt", header=FALSE)
```

Las probabilidades conocidas para cada pregunta no sensible son las siguientes:

```
p1<-5/12
```

```
p2<-0.19
```

```
p3<-0.82
```

```
p4<-0.22
```

En este siguiente paso hallaremos las preguntas de interés para calcular sus proporciones

Pregunta 5

¿He copiado de otros estudiantes durante un examen (práctica) por lo menos una vez?

```
Y1<-1-mean(data[,1])
```

```
Y1= 0.43
```

```
Preg5<-(p1-1+Y1)/(2*p1-1)
```

```
Preg5= 0.92
```

93.75% han copiado de otros estudiantes durante un examen por lo menos una vez

Pregunta 6

¿Utilizado notas en un examen (incluyendo notas en los teléfonos móviles, calculadoras u otros similares)?

```
Y2<-1-mean(data[,2])
```

```
Y2= 0.4766667
```

```
preg6<-(p2-1+Y2)/(2*p2-1)
```

```
preg6= 0.5376344
```

53.76% utilizaron notas en un examen.

Pregunta 7

¿He utilizado drogas para mejorar mi rendimiento académico en un examen o práctica por lo menos una vez?

```
Y3<-1-mean(data[,3])
```

```
Y3= 0.4
```

```
preg7<-(p3-1+Y3)/(2*p3-1)
```

```
preg7= 0.34375
```

El 34.37% de los estudiantes ha utilizado drogas para mejorar su rendimiento

Pregunta 8

¿He presentado un documento que contiene un párrafo intencionalmente adoptado del trabajo de otra persona por lo menos una vez.

```
Y4<-1-mean(data[,4])
```

```
Y4= 0.4066667
```

```
preg8<-(p4-1+Y4)/(2*p4-1)
```

```
preg8 = 0.6666667
```

El 66.67% de los estudiantes ha presentado un documento que contiene un párrafo adaptado de un trabajo

ANEXO 6: Resultados de la encuesta realizada a los estudiantes matriculados en el semestre académico 2015-I de la Facultad Ciencias Matemáticas de la UNMSM mediante el Método de Entrevista Directa

Nro.	EAP	Pregunta 1	Pregunta 2	Pregunta 3	Pregunta 4
1	Estadística	SI	NO	NO	SI
2	Computación	NO	NO	NO	NO
3	Matemática	SI	SI	NO	SI
4	Matemática	NO	NO	NO	NO
5	Matemática	SI	NO	NO	NO
6	Matemática	SI	NO	NO	SI
7	Matemática	SI	SI	NO	NO
8	Matemática	SI	NO	NO	NO
9	Matemática	SI	NO	NO	NO
10	Matemática	SI	NO	NO	NO
11	Matemática	SI	NO	NO	NO
12	Matemática	NO	NO	NO	NO
13	Inv. Operativa	SI	NO	NO	NO
14	inv. Operativa	NO	NO	NO	NO
15	inv. Operativa	SI	NO	NO	NO
16	inv. Operativa	SI	NO	NO	SI
17	inv. Operativa	SI	NO	NO	SI
18	inv. Operativa	SI	NO	NO	SI
19	inv. Operativa	SI	SI	NO	NO
20	inv. Operativa	NO	NO	NO	SI
21	inv. Operativa	NO	NO	NO	SI
22	inv. Operativa	NO	NO	NO	NO
23	inv. Operativa	NO	NO	NO	NO
24	inv. Operativa	NO	NO	NO	NO

25	inv. Operativa	NO	NO	NO	NO
26	inv. Operativa	SI	NO	NO	SI
27	Estadística	SI	SI	NO	SI
28	Matemática	NO	NO	NO	NO
29	Estadística	SI	SI	NO	NO
30	Estadística	SI	SI	NO	SI
31	Estadística	SI	NO	NO	NO
32	Estadística	NO	SI	NO	NO
33	Matemática	NO	NO	NO	SI
34	Matemática	SI	NO	NO	SI
35	Matemática	SI	SI	NO	NO
36	Matemática	NO	SI	NO	NO
37	Matemática	SI	NO	NO	NO
38	Matemática	SI	SI	NO	NO
39	Matemática	NO	NO	NO	NO
40	Matemática	NO	NO	NO	SI
41	Inv. Operativa	NO	NO	SI	NO
42	inv. Operativa	NO	NO	NO	SI
43	Inv. Operativa	NO	NO	NO	SI
44	Inv. Operativa	SI	SI	NO	NO
45	Inv. Operativa	NO	SI	NO	NO
46	Estadística	SI	SI	NO	SI
47	Estadística	SI	SI	NO	NO
48	Estadística	SI	SI	NO	SI
49	Estadística	SI	SI	NO	NO
50	Estadística	SI	SI	NO	SI
51	Estadística	SI	SI	NO	SI
52	Estadística	SI	SI	NO	NO
53	Inv.operativa	NO	NO	NO	NO
54	Computación	SI	NO	NO	SI

55	Computación	SI	SI	NO	NO
56	Computación	SI	SI	NO	NO
57	Matemática	NO	NO	NO	NO
58	Matemática	SI	NO	NO	NO
59	Matemática	SI	NO	NO	SI
60	Matemática	SI	NO	NO	NO
61	Matemática	SI	NO	NO	NO
62	Matemática	SI	NO	NO	SI
63	Matemática	SI	NO	NO	NO
64	Matemática	NO	NO	NO	NO
65	Matemática	SI	SI	NO	NO
66	Matemática	SI	NO	NO	SI
67	Matemática	SI	SI	NO	NO
68	Matemática	SI	NO	NO	NO
69	Computación	NO	NO	NO	NO
70	Inv.operativa	NO	NO	NO	NO
71	Inv.operativa	SI	SI	NO	NO
72	Computación	NO	NO	NO	SI
73	Inv.operativa	SI	SI	NO	NO
74	Estadística	NO	NO	NO	SI
75	Estadística	SI	NO	NO	NO
76	Estadística	SI	SI	NO	SI
77	Estadística	NO	NO	NO	SI
78	Estadística	SI	NO	NO	NO
79	Estadística	SI	SI	NO	SI
80	Estadística	NO	SI	NO	NO
81	Computación	NO	SI	NO	NO
82	Computación	SI	SI	NO	NO
83	Computación	NO	NO	NO	NO
84	Computación	SI	NO	NO	NO

85	Computación	NO	SI	NO	NO
86	Matemática	NO	NO	NO	NO
87	Matemática	NO	NO	NO	NO
88	Matemática	SI	NO	NO	NO
89	Matemática	NO	SI	NO	NO
90	Matemática	SI	NO	SI	SI
91	Matemática	NO	SI	NO	SI
92	Matemática	NO	NO	NO	SI
93	Estadística	SI	SI	NO	NO
94	Estadística	SI	SI	NO	SI
95	Estadística	SI	NO	NO	SI
96	Estadística	SI	SI	NO	NO
97	Estadística	SI	SI	NO	SI
98	Estadística	SI	SI	NO	NO
99	Estadística	SI	SI	NO	NO
100	Inv. operativa	SI	NO	NO	SI
101	Inv. operativa	SI	NO	NO	NO
102	Inv. operativa	NO	SI	NO	NO
103	Inv. operativa	NO	NO	NO	NO
104	Inv. operativa	SI	SI	NO	NO
105	Inv. operativa	NO	NO	NO	NO
106	Inv. operativa	SI	SI	NO	NO
107	Inv. operativa	SI	SI	NO	NO
108	Inv. operativa	SI	SI	NO	NO
109	Matemática	NO	NO	NO	NO
110	Matemática	NO	SI	NO	NO
111	Inv. operativa	NO	NO	NO	NO
112	Inv. operativa	NO	NO	NO	NO
113	Inv. operativa	NO	NO	NO	NO
114	Inv. operativa	SI	SI	NO	NO

115	Inv. operativa	NO	NO	NO	NO
116	Inv. operativa	SI	NO	NO	NO
117	Inv. operativa	NO	NO	NO	NO
118	Inv. operativa	SI	NO	NO	NO
119	Inv. operativa	SI	NO	NO	SI
120	Inv. operativa	SI	SI	NO	SI
121	Inv. operativa	NO	NO	NO	NO
122	Inv. operativa	SI	NO	NO	NO
123	Inv. operativa	SI	NO	NO	NO
124	Inv. operativa	NO	NO	NO	NO
125	Inv. operativa	NO	NO	NO	NO
126	Inv. operativa	SI	NO	NO	NO
127	Inv. operativa	NO	NO	NO	NO
128	Inv. operativa	NO	NO	NO	NO
129	Inv. operativa	NO	NO	NO	NO
130	Inv. operativa	SI	NO	NO	NO
131	Inv. operativa	SI	NO	NO	NO
132	Inv. operativa	SI	NO	NO	SI
133	Inv. operativa	SI	SI	NO	SI
134	Matemática	SI	NO	NO	NO
135	Matemática	NO	NO	NO	NO
136	Matemática	SI	NO	NO	NO
137	Matemática	SI	NO	NO	NO
138	Inv. operativa	NO	NO	NO	NO
139	Matemática	NO	NO	NO	NO
140	Matemática	NO	NO	NO	NO
141	Matemática	NO	NO	NO	NO
142	Matemática	NO	NO	NO	NO
143	Matemática	SI	SI	NO	NO
144	Computación	SI	NO	NO	NO

145	Computación	SI	SI	NO	NO
146	Computación	NO	SI	NO	SI
147	Computación	SI	NO	NO	SI
148	Computación	NO	NO	NO	SI
149	Computación	SI	NO	NO	SI
150	Computación	SI	SI	NO	NO
151	Computación	SI	NO	NO	NO
152	Computación	NO	NO	NO	NO
153	Computación	NO	NO	NO	NO
154	Estadística	NO	NO	NO	NO
155	Estadística	SI	NO	NO	SI
156	Estadística	NO	NO	NO	NO
157	Estadística	SI	SI	SI	SI
158	Estadística	SI	NO	NO	NO
159	Estadística	SI	SI	NO	SI
160	Estadística	SI	SI	NO	SI
161	Estadística	NO	NO	NO	SI
162	Estadística	NO	NO	NO	SI
163	Estadística	SI	NO	NO	NO
164	Computación	NO	SI	NO	NO
165	Computación	NO	NO	NO	NO
166	Computación	SI	SI	NO	NO
167	Computación	NO	NO	NO	NO
168	Computación	NO	NO	NO	NO
169	Computación	NO	SI	NO	NO
170	Computación	SI	SI	NO	NO
171	Computación	NO	NO	NO	NO
172	Computación	SI	SI	NO	NO
173	Computación	NO	NO	NO	NO
174	Computación	SI	SI	NO	NO

175	Inv.operativa	NO	NO	NO	NO
176	Estadística	SI	SI	NO	NO
177	Computación	NO	NO	NO	NO
178	Computación	NO	SI	NO	SI
179	Computación	SI	NO	NO	NO
180	Computación	NO	SI	NO	NO
181	Computación	SI	SI	NO	NO
182	Computación	NO	NO	NO	NO
183	Inv.operativa	SI	NO	NO	NO
184	Matemática	NO	NO	NO	NO
185	Matemática	NO	NO	NO	NO
186	Matemática	NO	NO	NO	NO
187	Computación	SI	NO	NO	NO
188	Computación	NO	NO	NO	NO
189	Matemática	NO	NO	NO	SI
190	Matemática	NO	NO	NO	NO
191	Matemática	NO	NO	NO	NO
192	Matemática	NO	NO	NO	NO
193	Matemática	NO	NO	NO	NO
194	Matemática	NO	NO	NO	NO
195	Matemática	NO	NO	NO	NO
196	Matemática	NO	NO	NO	NO
197	Inv.operativa	SI	NO	NO	SI
198	Inv.operativa	NO	NO	NO	SI
199	Inv.operativa	NO	NO	NO	NO
200	Inv.operativa	SI	NO	NO	NO

ANEXO N° 7: Estimación de los parámetros de los resultados obtenidos de la encuesta obtenida en los estudiantes de la Facultad de Ciencias Matemáticas de la UNMSM en el semestre académico 2015-I mediante Método de Preguntas Directas

Los resultados fueron obtenidos mediante el programa R project a continuación se muestran los códigos

Pregunta 5

¿He copiado de otros estudiantes durante un examen (práctica) por lo menos una vez?

```
p4=mean(PregDirecta[,1])
```

```
p4= 0.535
```

Se obtuvo que 53% de alumnos haya copiado alguna vez en un examen

Pregunta 6

¿Utilizado notas en un examen (incluyendo notas en los teléfonos móviles, calculadoras u otras similares)?

```
> preg5=mean(PregDirecta[,2])
```

```
> preg5= 0.325
```

Se obtuvo que el 32.5% de alumnos utilizo notas en un examen.

Pregunta 7

¿He utilizado drogas para mejorar mi rendimiento académico en un examen o práctica por lo menos una vez?

```
> preg6=mean(PregDirecta[,3])
```

```
> preg6= 0.015
```

Se obtuvo que el 1.5% utilizó droga para mejorar su rendimiento.

Pregunta 8

¿He presentado un documento que contiene un párrafo intencionalmente adoptado del trabajo de otra persona por lo menos una vez.

```
> preg7=mean(PregDirecta[,4])
```

```
> preg7=0.27
```

Se obtuvo que el 27% de los alumnos haya presentado un documento adoptado de otro trabajo

ANEXO N° 8: Realizando códigos en el programa R Project

Generando el grafico de Warner

```
nVar=function(p)
{
  valor=0.4*0.6+p*(1-p)/(2*p-1)^2
}
```

```
plot(nVar)
```

Gráfico sobre el grado de protección para diferentes valores de π

Para $\pi = 0.10$

```
GPPsi=function(p)
{
  valor1=0.10*p/(0.10*p+(1-0.10)*(1-p))
}
```

```
plot(GPPsi)
```

```
plot(GPPsi,type="l",ann=F)
```

```
GPPno=function(p)
```

```
{
  valor2=0.1*(1-p)/(0.1*(1-p)+(1-0.1)*p)
}
```

```
par(new=TRUE)
```

```
plot(GPPno,type="p",axes=F,ylab="GPPsi y GPPno",xlab="p")
```

Para $\pi = 0.50$

```
GPPsi=function(p)
```

```
{
  valor1=0.5*p/(0.5*p+(1-0.5)*(1-p))
}
```

```
plot(GPPsi)
```

```
plot(GPPsi,type="l",ann=F)
```



```

GPPno=function(p)
{
valor2=0.5*(1-p)/(0.5*(1-p)+(1-0.5)*p)
}
par(new=TRUE)
plot(GPPno,type="p",axes=F,ylab="GPPsi y GPPno",xlab="p")

```

Para $\pi = 0.70$

```

GPPsi=function(p)
{
valor1=0.7*p/(0.7*p+(1-0.7)*(1-p))
}
plot(GPPsi)
plot(GPPsi,type="l",ann=F)
GPPno=function(p)
{
valor2=0.7*(1-p)/(0.7*(1-p)+(1-0.7)*p)
}
par(new=TRUE)
plot(GPPno,type="p",axes=F,ylab="GPPsi y GPPno",xlab="p")

```

Para $\pi = 0.90$

```

GPPsi=function(p)
{
valor1=0.9*p/(0.9*p+(1-0.9)*(1-p))
}
plot(GPPsi)
plot(GPPsi,type="l",ann=F)

```

```
GPPno=function(p)
{
valor2=0.9*(1-p)/(0.9*(1-p)+(1-0.9)*p)
}
par(new=TRUE)
plot(GPPno,type="p",axes=F,ylab="GPPsi y GPPno",xlab="p")
```

ANEXO N° 9: “Cuestionario 1: El Modelo de Respuesta No Aleatorizada Transversal”

Universidad Nacional Mayor de San Marcos

Facultad de Ciencias Matemáticas

Estimado alumno, usted ha sido elegido al azar para participar en una investigación sobre algunas características de la conducta de los estudiantes de la FCM. Agradecemos de antemano su valiosa colaboración.

CUESTIONARIO SOBRE ALGUNAS CARACTERÍSTICAS DE LOS ALUMNOS

I. CARACTERÍSTICAS GENERALES

- 1) Edad: _____ 2) Género: 1. Femenino 2. Masculino
3) E.A.P: _____ 4) CICLO: _____

II. CARACTERÍSTICAS DEL TEMA

Por favor responda con veracidad, marque una cruz en el **círculo superior** si Ud. Pertenece a uno de los 2 círculos o marque una cruz en el **cuadrado superior** si Ud. pertenece a uno de los 2 cuadrados.

5)

Categoría	No nací entre Agosto y Diciembre (W = 0)	Nací entre Agosto y Diciembre (W = 1)
Nunca he copiado de otros estudiantes durante un examen (práctica). (Y = 0)	<input type="radio"/>	<input type="checkbox"/>
He copiado de otros estudiantes durante un examen (práctica) por lo menos una vez. (Y=1)	<input type="checkbox"/>	<input type="radio"/>

6)

Categoría	Nunca interrumpí mis estudios universitarios (W = 0)	He interrumpido mis estudios universitarios por lo menos una vez (W = 1)
Nunca he utilizado notas en un examen (incluyendo notas en los teléfonos móviles, calculadoras u otras similares) (Y = 0)	<input type="radio"/>	<input type="checkbox"/>
He utilizado notas en un examen (incluyendo notas en los teléfonos móviles, calculadoras u otras similares) (Y = 1)	<input type="checkbox"/>	<input type="radio"/>

7)

Categoría	No utilizo biblioteca tradicional (W = 0)	Utilizo biblioteca tradicional (W = 1)
Nunca he utilizado drogas para mejorar mi rendimiento académico en un examen o práctica (Y = 0)	<input type="radio"/>	<input type="checkbox"/>
He utilizado drogas para mejorar mi rendimiento académico en un examen o práctica por lo menos una vez (Y = 1)	<input type="checkbox"/>	<input type="radio"/>

Categoría	No me gusta estudiar solo(a) (W = 0)	Me gusta estudiar solo(a) (W = 1)
<p>Nunca he presentado un documento que contiene un párrafo intencionalmente adoptado del trabajo de otra persona.</p> <p>(Y = 0)</p>	<input type="radio"/>	<input type="checkbox"/>
<p>He presentado un documento que contiene un párrafo intencionalmente adoptado del trabajo de otra persona por lo menos una vez.</p> <p>(Y = 1)</p>	<input type="checkbox"/>	<input type="radio"/>

ANEXO N° 10: “Cuestionario 2: Sobre El Método Entrevista Directas”

**Universidad Nacional Mayor de San Marcos
Facultad de Ciencias Matemáticas**

**CUESTIONARIO SOBRE EL COMPORTAMIENTO DE LOS ALUMNOS DE LA FCM
FRENTE A PREGUNTAS SENSIBLES**

N° de Cuestionario

--	--	--

I. CARACTERÍSTICAS GENERALES

1. Edad: _____

Género

1. Femenino

2. Masculino

3. E.A.P: _____

4. CICLO:

1) I 2) II 3) III 4) IV 5) V 6) VI 7) VII 8) VIII 9) IX 10) X

II. CARACTERÍSTICAS DEL TEMA

¿Has copiado de otros estudiantes durante un examen (práctica) por lo menos una vez?

1. Si

2. No

¿Has utilizado notas en un examen (práctica) (incluyendo notas en los teléfonos móviles, calculadoras u otras similares)?

1. Si

2. No

¿Has utilizado drogas para mejorar tu rendimiento académico en un examen o práctica por lo menos una vez?

1. Si

2. No

¿Has presentado un documento que contiene un párrafo intencionalmente adoptado del trabajo de otra persona por lo menos una vez?

1. Si

2. No

Nombre del entrevistador: _____