

OPERACIÓN DEL PROGRAMA PARA EL ANÁLISIS GRÁFICO DE ÍTEMS (PAGI®), PARA HABILITAR A DOCENTES EN SERVICIO EN LA MEJORA DE SUS PRUEBAS DE AULA

GUADALUPE DE LOS SANTOS LÁZARO
Universidad Pedagógica Nacional

RESUMEN: El trabajo docente implica realizar una programación y evaluación de los contenidos, por ello los docentes elaboran sus reactivos en reuniones de colegiado o con base en su experiencia. Sin embargo, pocas ocasiones desarrollan ítems con validez y calidad técnica adecuada. En esta investigación se propone un curso-taller que tiene como objetivo principal actualizar a los docentes para mejorar las características que definen la calidad técnica de ítems de opción múltiple que utilizan en sus pruebas, mediante el análisis gráfico de ítems. Se realizan dos participaciones en Baja California, en las ciudades de Tijuana y Mexicali, donde se habilita a 33 docentes del nivel básico, medio superior y superior, en un curso-taller de cuatro horas para el manejo del Programa para el Análisis

Gráfico de Ítems (PAGI). Los materiales empleados son un manual que contiene aspectos conceptuales acerca de la técnica del análisis gráfico de ítems (AGI), la aplicación PAGI y una guía que permite al participante analizar cada ítem de su base de datos y emitir un juicio razonado sobre la calidad del ítem. Los principales resultados del estudio son proponer una herramienta a los docentes para elaborar ítems de buena calidad que cuente con validez, confiabilidad y discriminación adecuada para evaluar los procesos de aprendizaje. En consecuencia, se considera que el programa representa una aportación relevante para la formación y la actualización de profesores en el área de la evaluación del aprendizaje, además los docentes se interesaron por la herramienta y la pusieron en práctica desde su aula.

PALABRAS CLAVE: PAGI, desarrollo de pruebas de aula, análisis gráfico de ítems

Introducción

Como parte del seguimiento del trabajo titulado Diseño, operación y evaluación del Programa para el Análisis Gráfico de Ítems (PAGI®), para habilitar a docentes en servicio en la mejora de sus pruebas de aula. Se realiza dos participaciones como parte de la Semana Estatal de Evaluación de Baja California en Tijuana y en las Jornadas de evaluación bajacalifornianas en Mexicali. Donde se obtuvo información relevante de la operación del curso-taller, además se observa la funcionabilidad del programa denominado PAGI.

La evaluación es considerada de gran importancia en la práctica escolar, es auxiliar de diferentes ciencias como la economía, psicología, física, pedagogía; su intención es la medición que agrega un valor al aprendizaje esperado (Tamez, 2013). La OCDE (2011) menciona que los resultados de aprendizaje con medidas confiables y válidas de desempeño escolar pueden servir como base para políticas y programas educativos, con ello aumenta la confianza de los actores del sistema educativo.

Por ello, resulta urgente fomentar y desarrollar en las instituciones de educación superior de México una cultura de la evaluación, para contribuir al desarrollo sostenido de la educación donde todos los involucrados adquieren un compromiso para la toma de decisiones (Tamez 2003). En lo que respecta a la Educación Básica, se ha implementado el Catálogo Nacional de actualización 2009 – 2010; en donde se oferta una serie de cursos que permiten a los docentes actualizarse en diferentes área temáticas. Sin embargo, no se menciona cursos o posgrados enfocados a diseñar instrumentos de evaluación con los elementos mínimos de calidad técnica como la validez, confiabilidad y valor agregado (SEP, 2010).

Los docentes cuando evalúan dentro del salón de clase deben de diseñar un plan de evaluación que incluya instrumentos para identificar el desempeño escolar. Estos instrumentos deben reconocer los resultados obtenidos de forma objetiva, válida, confiable y transparente para que al estudiante le quede claro su proceso de aprendizaje (Frade, 2008).

La evaluación debe ser congruente con las actividades instrumentadas para el logro de ciertos aprendizajes, ésta busca vincular lo conceptual, procedimental y actitudinal para entender el desempeño de los estudiantes en distintos contextos y situaciones. Sin embargo, la construcción de instrumentos de evaluación centrados en el desempeño requiere de conocimientos de diseño técnico como lo es la validez y confiabilidad. En México, lo que se ha hecho es la implementación de un manual para la elaboración de reactivos según su tipo: opción múltiple, falso – verdadero, correlacion, etc. enfocado en el nivel superior. Sin embargo, solo se hace mención de los criterios que implementan para realizarlos, pero no comprueba la calidad técnica de los ítems del instrumento (SEP, 2008).

Así, se desarrolló un programa de cómputo que facilitara a los docentes adquirir y utilizar conocimientos y habilidades elementales de psicometría que les permitiera saber si los ítems de tipo objetivo que utilizan en sus pruebas tienen dificultad apropiada, buena discriminación y un funcionamiento adecuado de los distractores. Sin embargo, la construcción de instrumentos de evaluación centrados en el desempeño requiere de conocimientos de diseño técnico como lo es la validez y confiabilidad.

Por tal motivo, se considera que los docentes pueden analizar cada ítem de manera visual por medio de la técnica del Análisis Gráfico de Ítems (AGI) desarrollada por Batenburg y Laros (2002), esta técnica permite al docente analizar de manera simple, directa y visual cómo responden sus estudiantes ítems de respuesta seleccionada. Se considera que una representación gráfica permite observar variaciones en las tendencias de la respuesta correcta y los distractores en un ítem de respuesta seleccionada, respecto a la calificación en la prueba (Haladyna, 2004). Así, el análisis de ítems resulta más significativo e interpretable y pueden comprenderlo fácilmente personas ajenas a la psicometría, sin antecedentes estadísticos necesarios para calcular los índices tradicionales de los parámetros del ítem.

En consecuencia, el objetivo general del estudio fue operar y evaluar un programa de cómputo que permitiera a profesores en servicio, mejorar sus pruebas de aula mediante el uso de la técnica del análisis gráfico de ítems.

Cabe señalar que el análisis gráfico de ítems parte del supuesto de que quienes obtienen una puntuación máxima en una prueba, tienen alta probabilidad de seleccionar la opción correcta en cada ítem, y que la elección de opciones falsas o distractores disminuye al aumentar la puntuación total. Lo anterior puede observarse en la Figura 1.

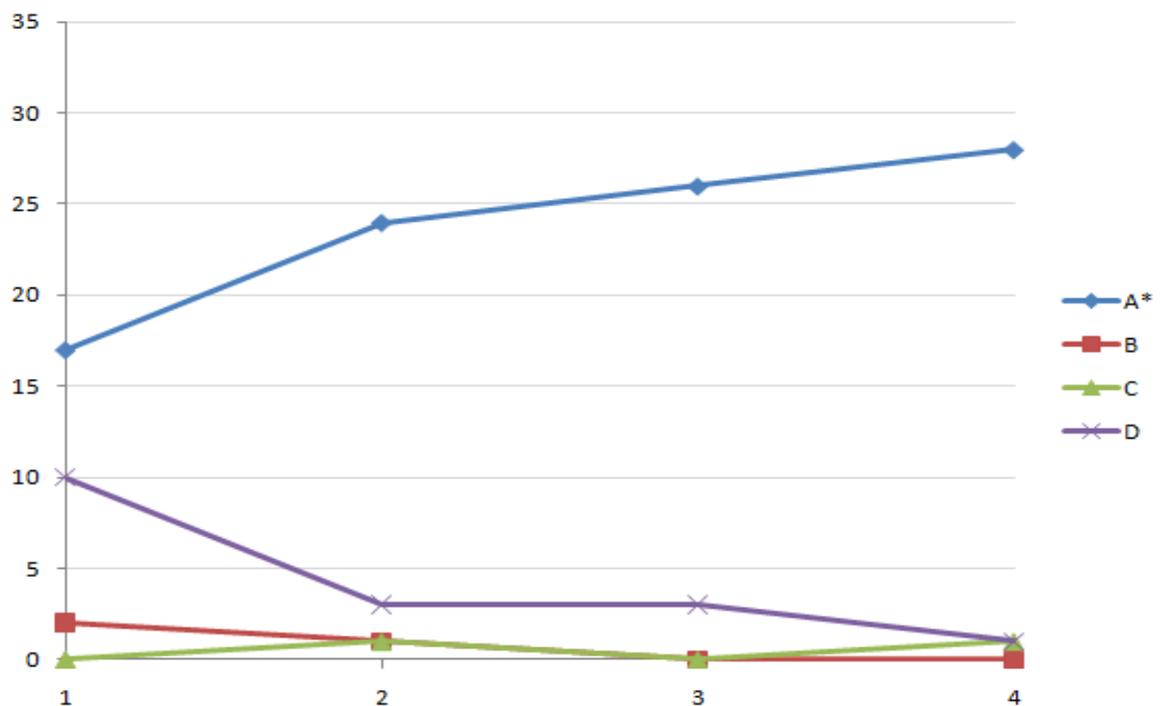


Figura 1. Frecuencias de respuestas a la opción correcta y a los distractores, por grupos de habilidad, en un ítem de opción múltiple con cuatro opciones.

En la figura 1, se observa que el eje Y muestra la frecuencia de examinados que respondieron el ítem. En el eje X se observan cuatro grupos de habilidad que se formaron a partir de la calificación que obtuvieron los alumnos en la prueba: grupo 1, bajo; grupo 2, medio bajo; grupo 3, medio alto; y grupo 4, alto (con las calificaciones más altas en el examen). En este caso, en cada grupo de habilidad hay 30 examinados, porque el ítem fue aplicado a una muestra de 120 estudiantes.

Las líneas que se observan en el eje de coordenadas son las curvas de respuestas al ítem y se observa que la curva de la respuesta correcta (A) aumenta a medida que incrementa la habilidad. Por otro lado, las líneas que están marcadas como B, C y D son llamados distractores, y para ser efectivos sólo deben elegirlos los alumnos que no dominan el contenido.

El ítem que se ilustra en la figura 1 se considera idealmente bueno, porque la curva de la respuesta correcta aumenta conforme incrementa la habilidad y la elección de los distractores disminuye a medida que incrementa la habilidad.

Método

Participantes

Inicialmente se realizó un piloteo donde participaron 21 profesores en servicio de secundaria de Ensenada, Baja California (2010), con diversas especialidades, y 12 profesores de bachillerato y de universidades de Tampico y la Ciudad de México. Como parte de la Primera Semana Estatal de Evaluación de Baja California (2011) con sede en la ciudad de Mexicali, B.C. se tuvo la participación de 52 profesores en servicio: 28 profesores del nivel medio superior, 17 profesores de educación secundaria y 7 profesores del nivel primaria. También como parte de los trabajos de las Jornadas bajacalifornianas en materia de Evaluación (2011) con sede en la ciudad de Tijuana, B.C. se tuvo la participación de 17 profesores: 1 del nivel primaria, 7 del nivel secundaria y 9 de educación media superior. Cabe destacar que de todos los participantes, 17 tienen posgrado desde especialidad hasta doctorado. Para participar en el taller los docentes debieron cumplir con los siguientes requisitos: traer una laptop y un archivo en Excel© de Microsoft con las respuestas brutas que dieron 32 o más alumnos a por lo menos 10 ítems de una prueba de aula aplicada previamente.

Materiales y procedimientos

Para habilitar a los docentes con la técnica del análisis gráfico de ítems, se desarrollaron y aplicaron los siguientes materiales e instrumentos:

- Programa para el Análisis Gráfico de Ítems (PAGI®). Para utilizar la técnica del análisis gráfico de ítems en pequeña escala, resulto necesario hacer

adecuaciones que permitieran identificar las características y condiciones de la evaluación en el aula. El problema fue garantizar que con pocos estudiantes se observa claramente los cambios de tendencia en la curvas de la respuesta correcta y de los distractores. Fue necesario utilizar una base de datos auténtica derivada de la aplicación a gran escala de la prueba EFCyE elaborada por la Universidad Autónoma de Baja California. Para obtener un buen resultado se hicieron variar sistemáticamente el número de grupos de habilidad (2, 3, 4 y 5) en que podrían dividirse las puntuaciones en dicha prueba, así como el número total de examinados (10, 15, 20, 25, 30, 35, 40, 45, 50) que se podrían asignar a cada grupo de habilidad. Tras los correspondientes ensayos se observó que, cuando se representaron las puntuaciones en la prueba en 4 grupos de habilidad, y se asignó un mínimo de ocho examinados a cada grupo de habilidad (32 casos en total), fue posible observar con claridad las tendencias de las respuestas al ítem. Posteriormente se desarrolla el programa de cómputo PAgI® para que los docentes utilicen dicha técnica en sus evaluaciones de aula.

- Manual del participante. Contiene información conceptual y metodológica sobre aspectos principales de la teoría clásica de los tests, el análisis gráfico de ítems y la interacción del programa de cómputo PAgI®.

Se instrumentaron cuatro cursos – taller, cada uno de ellos en una sesión de cuatro horas, la dinámica del curso consistió en revisar conceptos necesarios para la interpretación gráfica de los ítems y para entender la interface gráfica de PAgI®. Después cada participante utilizó el programa y analizó las gráficas que produjo PAgI® sobre sus ítems, hasta que fue capaz de registrar en la guía un dictamen razonado sobre la calidad técnica de cada uno de sus ítems.

Resultados

El resultado principal del estudio consistió en analizar la funcionalidad de implementar un curso –taller, para que los docentes obtengan conocimientos básicos de la técnica del análisis gráfico de ítems, con la finalidad de que mejoren sus pruebas de aula. Durante el desarrollo del curso – taller los docentes contestaron una guía que permitió el análisis de ítems de la base de datos que llevaron. Esta guía permitió identificar ítems que tienen la discriminación adecuada, además de observar cuáles son las opciones que se tienen que modificar por no cumplir con los parámetros de validez y discriminación.

Con base en los juicios que asignaron los docentes en la guía para interpretar el comportamiento de sus ítems, se analizó su correspondencia con cada gráfico que produjo PAGO®. En general, se observó que los docentes formularon comentarios significativos y congruentes con las gráficas que produjo el software. En particular, fueron congruentes la identificación de distractores poco funcionales, la estimación de la dificultad y el poder discriminativo del ítem, así como el dictamen general sobre eliminar el ítem o sobre qué aspectos modificar en el ítem a fin de mejorarlo. Como se muestra en la Tabla 1, donde se presentan los comentarios que formularon algunos de los participantes sobre la ejecución de sus estudiantes, en uno de los ítems evaluados.

Tabla 1. Interpretaciones de los participantes de algunos ítems que analizaron

Docente	Ítem	Opción	Comentario
5	1	A	Funciona adecuadamente
		B	Resulta confundir a algunos expertos
		C*	Lo seleccionan los alumnos que si dominan el contenido
		D	Es un distractor muy atractivo para todos
		JUICIO FINAL SOBRE EL ÍTEM	El ítem es muy difícil y se necesita modificar algunos distractores.
10	1	A	Si cumple aumenta la habilidad y discrimina.
		B	Discrimina en un nivel medio bajo
		C*	Tiene una dificultad alta. Solo contesto de los que no dominan el contenido.
		D	Distrae y confunde a los de mayor habilidad.
		JUICIO FINAL SOBRE EL ÍTEM	Es de dificultad alta, no queda claro quienes dominan el contenido.
15	5	A	Es un distractor que funciona adecuadamente
		B	Logra distraer a los alumnos que no dominan el contenido
		C*	La respuesta correcta la seleccionan los alumnos que dominan el contenido.
		D	Es un distractor que funciona adecuadamente.
		JUICIO FINAL SOBRE EL ÍTEM	Es un ítem idealmente bueno que discrimina a los alumnos que si dominan el contenido de los que no lo dominan.
20	12	A*	Tiene buena comportamiento, resulta ser una pregunta difícil.
		B	Es un buen distractor separa a los expertos de los que no dominan el contenido.
		C	Es un buen distractor, presenta un comportamiento adecuado.
		D	No funciona, se debe de cambiar porque nadie lo contesta.
		JUICIO FINAL SOBRE EL ÍTEM	Se tiene que modificar un distractor para que sea funcional.
25	21	A	No discrimina a los grupos de habilidad alta de la baja.
		B*	Si cumple su función ya que lo eligieron más lo de habilidad baja.
		C	Cumple medianamente su función.
		D	Si cumple su función ya que si discrimina.
		JUICIO FINAL SOBRE EL ÍTEM	Es de dificultad media, si discrimina de los examinados con mayor habilidad.

* Se trata de la respuesta correcta

Las argumentaciones de los docentes corresponden a las gráficas que se produjeron en el PEGI® sobre los ítems que analizaron. alguna de ellas, se muestra en la Figura 2, donde el ítem debe de modificar algunos distractores porque resultan ser atractivos para algunos alumnos que si dominan el contenido. A diferencia de la gráfica obtenida de la Figura 3 del ítem 5 del participante 15, se observa que tiene una discriminación adecuada porque solo seleccionan la respuesta correcta los alumnos que dominan el contenido y los distractores son seleccionados por aquellos que les falta un dominio sobre el contenido.

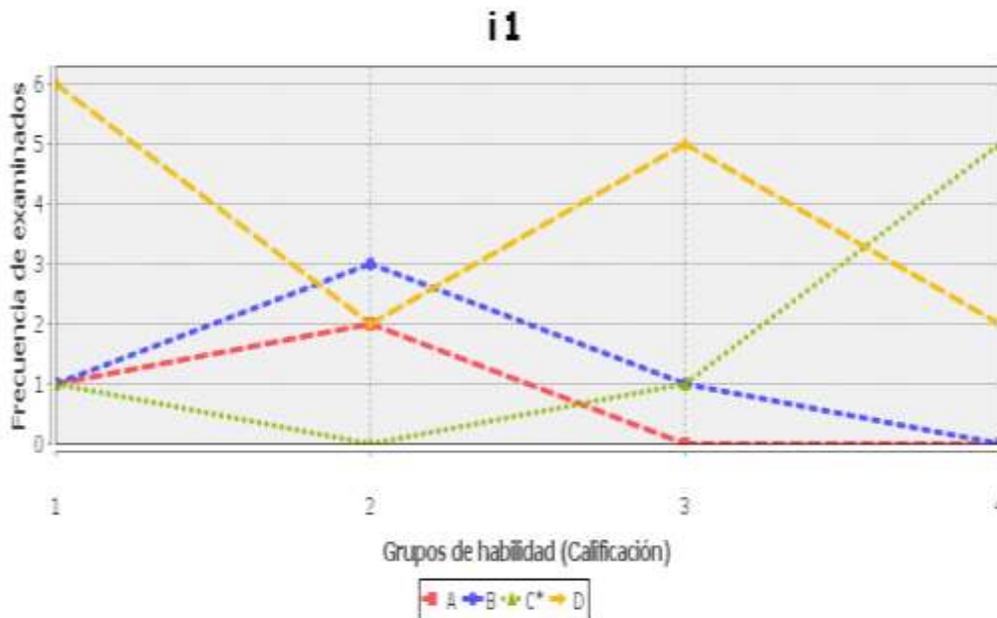


Figura 2. Gráfica producida por PEGI® a partir de la respuesta al ítem 1 del participante 5.

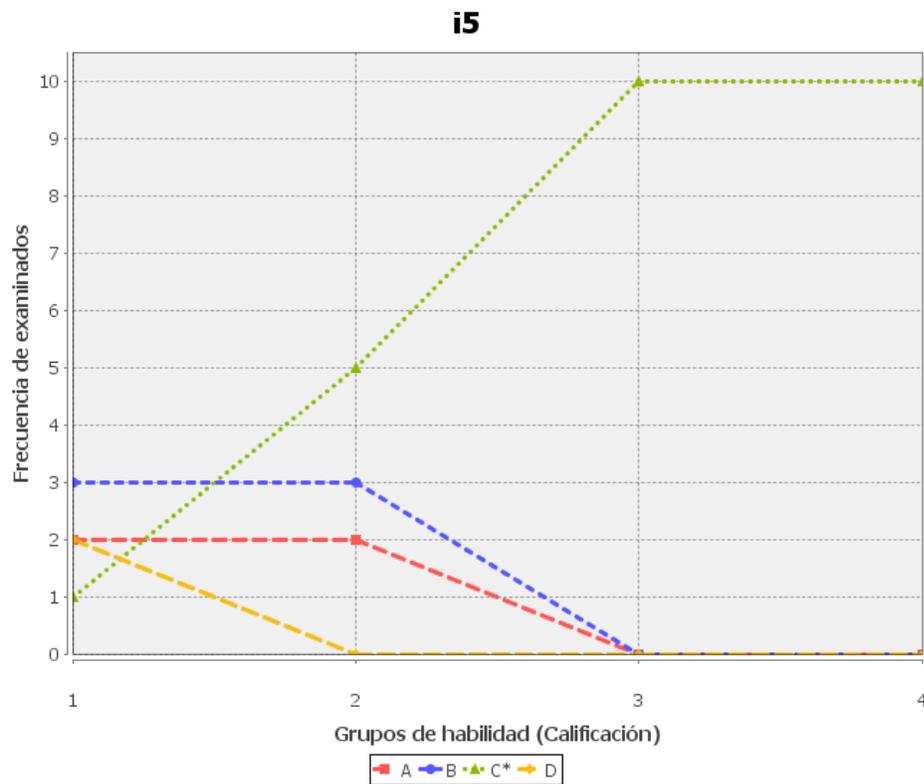


Figura 3. Gráfica producida por PAGI® a partir de las respuestas al ítem 5 del participante

15

Conclusiones y recomendaciones

Los principales resultados obtenidos permiten concluir que se logró cumplir con los objetivos propuestos. Así, fue posible actualizar a los docentes por medio del programa de cómputo, de modo que con ello mejoran sus pruebas objetivas mediante el análisis gráfico de ítems. Los siguientes argumentos apoyan esta conclusión:

- Los participantes opinaron que PAGI® les permitió identificar fortalezas y debilidades de sus ítems, según los indicadores psicométricos que definen su calidad técnica, así como formas específicas de mejorarlos. Sus participaciones durante la capacitación y los productos que generaron tras analizar sus ítems y llenar la guía, confirman esta opinión.

- La guía de evaluación arrojó datos que muestran su calidad psicométrica y por ello se puede afirmar que en opinión de los docentes los materiales diseñados propiciaron aprendizajes significativos que fueron pretendidos.
- Los docentes demostraron aprendizajes significativos de habilidades básicas de psicometría, que son necesarios para emitir un juicio razonado acerca del ítem que aplico en su prueba de aula.

Se sugiere utilizar el programa de cómputo en profesores de otros niveles y modalidades educativos, antes de pensar en la posibilidad de utilizarlo de manera más extensiva tanto en la formación inicial, como en la actualización de los profesores en servicio.

Bibliografía

Batenburg, T. V. & Laros, J. A. (2002). Graphical Analysis of Items. *Education research and evaluation*, 8(3), 319 – 333.

De Los Santos G. (2010) Desarrollo, operación y evaluación de un módulo para capacitar a docentes en servicio para que mejoren sus exámenes de opción múltiple mediante el análisis gráfico de ítems. Tesis de maestría publicada. Instituto de Investigación y Desarrollo Educativo, universidad Autónoma de Baja California.

Frade L. (2008) Evaluación por competencias. En *Inteligencia educativa* (2da Ed) (Pág 19-26) México

Haladyna, M. (2004). Developing and validating multiple choice test items.

Consultado el 15 de Noviembre de 2009 en: <http://books.google.com.mx/books>

Lozano, M. E. (2007). Formación docente: El servicio de asesoría académica a las escuelas y el trayecto formativo. En *Memorias del IX Congreso Nacional de Investigación Educativa* [Disco compacto] Yucatán, México: Universidad Autónoma de Yucatán.

Martínez, M., Maya, C. y Zenteno, E. (1996). Hacia una evaluación institucional de la Educación Normal. En *Memorias del II Foro de Evaluación Educativa*. Saltillo, Coahuila, México: Centro Nacional de Evaluación para la Educación Superior.

OCDE (2011), La medición del aprendizaje de los alumnos: Mejores prácticas para evaluar

el valor agregado de las escuelas, OECD Publishing. Consultado en:

<http://dx.doi.org/10.1787/9789264090170-es>

Shepard, L. (2006). La evaluación en el Aula. En Robert I. Brennan (Eds), *Identification of Mild Handicaps* (pp545- 572). Bolder, Colorado, E.E.U.U. (Reimpreso en México por Instituto Nacional para la Evaluación de la Educación, Ed. 2006, México. Consultado el 30 de octubre de 2008 en: <http://www.inee.edu.mx/>

Secretaría de Educación, Pública [SEP] (2008). Hacia la construcción de un sistema nacional de evaluación de la educación media superior. Sep.gob.mx. Consultado el 17 de Abril en: <http://www.reformaiems.sems.gob.mx/>

Tamez S. (2003) La evaluación como factor de cambio en la educación. En Revista Cultura, Estadística y Geografía. 22. Pág. 64-72