

**ANÁLISIS ESTADÍSTICO DE LOS RESULTADOS DE LAS OLIMPIADAS  
REGIONALES DE MATEMÁTICAS DE LA UNIVERSIDAD INDUSTRIAL DE  
SANTANDER 2011- 2012**

**MARÍA ANGÉLICA TOSCANO SALAS**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE CIENCIAS  
ESCUELA DE MATEMÁTICAS  
BUCARAMANGA  
2013**

**ANÁLISIS ESTADÍSTICO DE LOS RESULTADOS DE LAS OLIMPIADAS  
REGIONALES DE MATEMÁTICAS DE LA UNIVERSIDAD INDUSTRIAL DE  
SANTANDER 2011- 2012**

**MARÍA ANGÉLICA TOSCANO SALAS**

**Trabajo de Grado para obtener el título de  
Licenciado en Matemática**

**Directora  
TULIA ESTHER RIVERA FLÓREZ  
Mg. en Estadística**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE CIENCIAS  
ESCUELA DE MATEMÁTICAS  
BUCARAMANGA  
2013**

## Dedicatoria

*A Dios por brindarme sabiduría y entendimiento para culminar con éxito esta etapa importante de mi vida.*

*A mis padres Mara Teresa Salas, y Edgar Toscano Riveros, a mis hermanos Edgar Toscano y Danna Gabriela Toscano por su amor y apoyo incondicional durante toda mi carrera.*

*Al amor de mi vida Oscar Iván Montoya Salas por compartir su vida a mi lado, por su apoyo, su paciencia y su amor durante todos estos años.*

*María Angélica Toscano Salas*

## **Agradecimientos**

A la profesora Tulia Esther Rivera por su apoyo, paciencia y por compartir su conocimiento y experiencias académicas, que sirvieron como guía para sacar adelante este proyecto y muchos más,

A mi tía Laura Mariela Salas por estar en todos los momentos lindos y difíciles brindándome todo su apoyo.

Y finalmente a mis amigos por su cariño y apoyo incondicional durante el desarrollo del proyecto, en especial a Marisol Rueda puentes.

## TABLA DE CONTENIDO

	Pág.
INTRODUCCIÓN .....	17
1. DESCRIPCIÓN DE LA PRUEBA Y DE LA MUESTRA DE DATOS.....	19
1.1 DESCRIPCIÓN DE LA PRUEBA .....	19
1.1.1 Forma de calificar cada prueba.....	20
1.2 DESCRIPCIÓN DE LA MUESTRA DE DATOS .....	21
2. MARCO TEÓRICO .....	24
2.1 BASE DE DATOS .....	24
2.2 ESTADÍSTICA DESCRIPTIVA .....	24
2.2.1 Medidas de tendencia central .....	24
2.2.2 Medidas de dispersión .....	26
2.2.3 Medidas de localización .....	27
2.2.4 Gráfico de caja y bigotes.....	28
2.2.5. Histograma.....	29
2.2.6 Grafica de dispersión de puntos .....	30
2.2.7 Gráfico de barras de error.....	31
2.2.8 Factor.....	31
2.3 TABLAS DE CONTINGENCIA.....	32
2.3.1 Medidas de asociación.....	32
2.3.2 Distribución de frecuencia.....	33
2.4 RAZÓN DE ODDS (OR) .....	34
2.5 RIESGO RELATIVO (RR) .....	35
3. ANÁLISIS DESCRIPTIVO.....	37
3.1 NIVEL DE PARTICIPACIÓN EN LA PRUEBA .....	37
3.2 RENDIMIENTO EN LA PRUEBA.....	43
3.2.1 Rendimiento en la fase clasificatoria.....	44

3.2.2 Rendimiento en la prueba selectiva .....	48
3.2.3 Rendimiento en la fase final.....	51
3.2.4 Rendimiento por área temática:.....	53
3.2.5 Rendimiento en la prueba vs asistencia a las capacitaciones .....	55
3.3 RENDIMIENTO DE LA FASE FINAL.....	57
4. COMPARACIONES DISCRIMINANDO POR GÉNERO.....	61
4.1 NIVEL DE PARTICIPACIÓN EN LA PRUEBA .....	61
4.2 RENDIMIENTO EN LA FASE CLASIFICATORIA .....	66
4.3 RENDIMIENTO EN LA FASE FINAL.....	72
5. CONCLUSIONES .....	75
BIBLIOGRAFÍA.....	78
ANEXOS.....	79

## LISTA DE GRÁFICAS

	Pág.
Gráfica 1. Participantes por nivel (2011-2012).....	37
Gráfica 2. Participación por fase (2011).....	38
Gráfica 3. Participación por fase (2012).....	38
Gráfica 4. Participación por colegios oficiales y privados (2011-2012).....	40
Gráfica 5. Participación por colegios rurales y urbanos (2011-2012).....	41
Gráfica 6. Participación por género (2011-2012) .....	41
Gráfica 7. Mapa de santander de los municipios que participaron en las Olimpiadas Regionales de Matemáticas UIS (2011-2012) .....	43
Gráfica 8. Rendimiento en la prueba clasificatoria discriminando por año y nivel .	45
Gráfica 9. Rendimiento en la prueba selectiva discriminando por año y nivel .....	49
Gráfica 10. Rendimiento en la prueba final discriminando por año y nivel (2011- 2012).....	51
Gráfica 11. Barras de error para las variables puntaje por área temática (2011)...	53
Gráfica 12. Barras de error para las variables puntaje por área temática (2012)...	54
Gráfica 13. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Avanzado 2011) .....	57
Gráfica 14. Puntaje en la prueba final discriminando por género y por nivel (2011) .....	58
Gráfica 15. Puntaje en la prueba final discriminando por género y por nivel (2012) .....	58
Gráfica 16. Participación por Nivel Vs. Género (2011) .....	61
Gráfica 17. Participación por Nivel Vs. Género (2012) .....	62
Gráfica 18. Participación discriminando por género y zona del colegio (2011-2012) .....	63
Gráfica 19. Participación discriminando por género y status del colegio (2011-2012 .....	65

Gráfica 20. Panel con la distribución de los puntajes en la prueba clasificatoria discriminando por género nivel avanzado.....	66
Gráfica 21. Panel con la distribución de los puntajes en la prueba clasificatoria discriminando por género nivel medio .....	66
Gráfica 22. Panel con la distribución de los puntajes en la prueba clasificatoria discriminando por género nivel básico.....	67
Gráfica 23. Mediana de los puntajes en la prueba clasificatoria 2011 .....	69
Gráfica 24. Mediana de los puntajes en la prueba clasificatoria en 2012 .....	69
Gráfica 25. Gráfico de dispersión de puntos para la variable puntaje prueba final en 2011 .....	73
Gráfica 26. Gráfico de dispersión de puntos para la variable puntaje prueba final en 2012.....	74
Gráfica 27. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Avanzado 2011) .....	92
Gráfica 28. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Medio 2011.....	92
Gráfica 29. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Básico 2011.....	93
Gráfica 30. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Avanzado 2012 .....	93
Gráfica 31. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel medio 2012) .....	94
Gráfica 32. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Básico 2012) .....	94

## LISTA DE TABLAS

	Pág.
Tabla 1. Prueba de hipótesis para diferencia de proporciones .....	39
Tabla 2. Resumen de estadísticos descriptivos del rendimiento en la prueba en cada fase de la competencia (2011-2012).....	45
Tabla 3. Clasificación Nivel vs Nivel_Prueba_clasificatoria .....	47
Tabla 4. Nivel de la prueba vz nivel de desempeño en la prueba selectiva (2011-2012).....	50
Tabla 5. Nivel de la prueba vs nivel de desempeño en ña prueba final (2011-2012) .....	52
Tabla 6. Participación a las capacitaciones (2011-2012).....	55
Tabla 7. Coeficientes de Asimetría para la variable puntaje según el número de capacitaciones a las que asistió.....	56
Tabla 8. Clasificación final .....	59
Tabla 9. Prueba de hipótesis para diferencia de proporciones .....	63
Tabla 10. Prueba de hipótesis para diferencia de proporciones .....	65
Tabla 11. Mediadas de dispersión para la variable puntaje en la prueba clasificatorias .....	68
Tabla 12. Valores Prueba de hipótesis para probar diferencias de medias en el rendimiento promedio .....	70
Tabla 13. Cálculos usando los resultados de la fase clasificatoria del Nivel Avanzado (2011).....	71
Tabla 14. Colegios finalistas .....	72
Tabla 15. Participación de colegios discriminando por nivel de desempeño (2011-2012).....	80
Tabla 16. Participación por municipio (2011-2012).....	84
Tabla 17. Status del colegio vs el nivel de desempeño alcanzado durante la prueba clasificatoria .....	86

Tabla 18. Status del colegio vs el nivel de desempeño alcanzado durante la prueba clasificatoria .....	88
Tabla 19. Porcentaje de respuestas correctas en la prueba clasificatoria y selectiva (2011) .....	90
Tabla 20. Porcentaje de respuestas correctas en la prueba clasificatoria y selectiva (2012) .....	91
Tabla 21. Tabla de estadísticos descriptivos discriminando por género .....	95
Tabla 22. Cálculos usando los resultados de la fase clasificatoria del Nivel Avanzado (2012).....	96

## LISTA DE FIGURAS

	Pág.
Figura 1 Resumen de estructura en cuanto a áreas temáticas y número de preguntas.....	20
Figura 2. Distribución sesgada a la derecha.....	26
Figura 3. Ejemplo caja y bigotes.....	29
Figura 4. Histograma.....	30
Figura 5. Dispersión de puntos.....	30
Figura 6. Barras de error.....	31
Figura 7. Ejemplo tabla de contingencia.....	32
Figura 8. Ejemplo distribución de frecuencia.....	34
Figura 9. Distribución de frecuencia.....	34
Figura 10. Tabla para hallar la razón de odds.....	35

## LISTA DE ANEXO

	Pág.
ANEXO A. Ejemplo de calificación en las preguntas de tipo ensayo .....	79
ANEXO B. Participación de colegios discriminando por nivel de desempeño (2011-2012).....	80
ANEXO C. Participación por municipio (2011-2012).....	84
ANEXO D. Status del colegio vs el nivel de desempeño alcanzado durante la prueba clasificatoria .....	86
ANEXO E. Porcentaje de respuestas correctas en la prueba clasificatoria y selectiva (2011 .....	90
ANEXO F. Gráficas Comparativas del rendimiento discriminando por el número de capacitaciones a las que asistieron por niveles .....	92
ANEXO G. Tabla de estadísticos descriptivos discriminando por género.....	95
ANEXO H. Cálculos usando los resultados de la fase clasificatoria del Nivel Avanzado (2012).....	96

## Resumen

**TITULO\*:** ANÁLISIS ESTADÍSTICO DE LOS RESULTADOS DE LAS OLIMPIADAS REGIONALES DE MATEMÁTICAS DE LA UNIVERSIDAD INDUSTRIAL DE SANTANDER 2011- 2012

**AUTORES:** TOSCANO SALAS, María Angélica\*\*<sup>1</sup>

**PALABRAS CLAVE:** olimpiada matemática, diferencias por género, análisis estadístico, resultados olimpiadas, olimpiadas regionales.

### DESCRIPCIÓN

Desde 1980, en Colombia se vienen realizando las Olimpiadas Colombianas de Matemáticas competencia que es organizada por la Universidad Antonio Nariño. Regionalmente otras instituciones organizan competencias similares. Desde 2009, la Universidad Industrial de Santander, a través de EDUMAT, un grupo de investigación, adscrito a la Escuela de Matemáticas, realiza las Olimpiadas regionales de Matemáticas para niveles de básica primarios y Educación Secundaria. La competencia consta de cuatro fases: Preparatoria, Clasificatoria, Selectiva y Final y en Básica secundaria se incluyen tres niveles: Básico (6 y 7 grado), Medio (8 y 9 grado) y avanzado (10 y 11 grado). El trabajo utiliza elementos de estadística descriptiva para analizar el rendimiento de los participantes durante las competencias de 2011 y 2012.

Dentro de los resultados encontrados y sobre los niveles de participación de los estudiantes en la Prueba, se tiene que en su mayoría los participantes provienen de colegios urbanos (92% en 2011; 86% en 2012) y del sector oficial (85% en 2011; 83% en 2012). Los análisis del rendimiento en la prueba se realizaron teniendo en cuenta cuatro niveles de desempeño: Bajo, Aceptable, Superior y Alto definidos para cada fase de la competencia permitieron evidenciar que son pocos los estudiantes que obtuvieron alto rendimiento en la prueba, así en la fase clasificatoria sólo el 0,01% en 2011 y el 0, 1% en 2012. En la fase final se observa un comportamiento similar, el 1,6% en 2011 y el 0% en 2012 son solo los que alcanzan una nivel alto de desempeño entre los tres niveles de la prueba (Básico, Medio y Avanzado). En cuanto a las diferencias por género, en la fase clasificatoria las distribuciones de los puntajes para hombres y mujeres son similares, las diferencias importantes se observan sólo en el puntaje máximo alcanzado en especial en el nivel avanzado.

---

\*Trabajo de Grado

\*\*Facultad de Ciencias. Licenciatura en Matemáticas. Mg. Tulia Esther Rivera Flórez

## Abstract

**TITLE\*:** STATISTICAL ANALYSIS OF RESULTS OF MATH OLYMPICS  
REGIONAL INDUSTRIAL DE SANTANDER UNIVERSITY 2011 - 2012

**AUTHORS:** TOSCANO SALAS, María Angélica \*\*2

**Keywords:** Mathematical Olympiad, gender differences, statistical analysis, results Olympics, regional Olympics

### DESCRIPTION

Since 1980, in Colombia have been doing the Colombian Math Olympics, competition that is organized by the Antonio Nariño university and similar competitions have been organized by other institutions, regionally. Since 2009, the Universidad Industrial de Santander, through EDUMAT, a research group, attached to the Mathematics School is conducting the regional Mathematics Olympics from Elementary to Secondary Education. The competition is divided into four phases: Preparatory, Qualifying, Selective and Final in the Elementary section and in three levels: Basic (6 and 7th grade), Middle (8th and 9th grade) and advanced (10 and 11 degree) in the Secondary ones. The work uses elements of descriptive statistics to analyze the participants' performance during the competitions from 2011 to 2012.

Taking into account the findings results and the levels of students' participation in the test, the majority of students come from urban (92% in 2011, 86% in 2012) and the official sector (85% in 2011, 83% in 2012). The performance analysis of the test were done taking into account four performance levels: Low, regular, Superior and Upper defined for each phase of competition, allowed to evidence that few students who obtained a high result in the test and in the qualifying phase only 0.01% in 2011 and 0, 1% in 2012. In the final phase shows a similar behavior, 1.6% in 2011 and 0% in 2012 are only reaching a high level of performance between the three test levels (Basic, Intermediate and Advanced). As for gender differences, in the distributions qualifying scores for men and women are similar, some significant differences are observed only in the maximum score achieved especially at the advanced level.

---

\*Thesis

\*\*Faculty of Science. Bachelor of Mathematics. Mg. Tulia Esther Rivera Flórez

## INTRODUCCIÓN

La Olimpiada Internacional de Matemática es una competencia que se realiza cada año para estudiantes de secundaria y es la más antigua de las Olimpiadas Internacionales de Ciencias. La primera Olimpiada se celebró en Rumania en 1959, desde aquel tiempo se ha celebrado anualmente con excepción de 1980<sup>1</sup>. Cerca de 100 países envían equipos de un máximo de seis estudiantes junto con un líder de equipo, un tutor y observadores.

En Colombia se celebran varias competencias de matemática, algunas son organizadas por la Universidad Antonio Nariño y otras por grupos de trabajo a nivel regional. La principal de ellas, la competencia nacional llamada Olimpiada Colombiana de Matemáticas se inició en 1980 impulsada por el rector de la universidad Antonio Nariño de la época<sup>1</sup>. Esta competencia sirve como base para la selección del equipo que va a representar al país en la competencia internacional. Los mejores concursantes de la competencia son invitados a conformar el equipo nacional, y reciben entrenamiento en solución de problemas de matemáticas propios de la prueba.

A nivel de Santander, apenas en el año 2009 se realizaron las primeras olimpiadas de alcance regional, iniciativa de la Universidad Industrial de Santander a través de un equipo de trabajo perteneciente al grupo de investigación EDUMAT de la Escuela de Matemáticas. Con este proyecto la Escuela se propone generar un espacio permanente con actividades programadas a lo largo del año que puedan estimular el estudio de las matemáticas, ayudando a la formación de un pensamiento crítico y de un espíritu científico en los jóvenes, así como el desarrollo de habilidades y destrezas que les permitan un mejor desempeño en los ámbitos social, académico y laboral. Por ello, a diferencia de otras competencias cuyo objetivo es sólo identificar estudiantes talentosos, las actividades que desarrolla el Grupo de Olimpiadas de la Universidad Industrial de Santander tienen como

objetivo adicional incentivar el estudio de la matemática entre profesores y estudiantes de las provincias de Santander (Soto, Mares, García Rovira, Guanentá, Comunera y Vélez).

De otro lado, desde hace ya varios años existe un alto interés por desarrollar iniciativas a todo nivel que promuevan la equidad de género. En este sentido, al interior de la UIS se conformó un grupo llamado Fortalecimiento Equidad de Género en la Educación Superior (FEGES), cuyo objetivo es el fortalecimiento de capacidades institucionales en equidad de género e igualdad de oportunidades entre mujeres y hombres con el fin de consolidar programas académicos, procesos de formación e investigación en la educación superior en Colombia. Bajo esta perspectiva y considerando el ámbito nacional, a lo largo de la historia de las Olimpiadas Colombianas de Matemática (32 versiones), se tiene un registro de una baja participación de mujeres en la clasificación final, sólo en 1983, 2002, 2006, 2010 y 2011 ha sido una mujer la ganadora de la competencia en alguno de los tres niveles incluidos en la prueba: primer nivel, intermedio o superior. Bajo esta perspectiva y atendiendo a la invitación hecha por los integrantes del grupo FEGES, este trabajo pretende aportar evidencias sobre la existencia de diferencias debidas al género en el rendimiento en las olimpiadas a nivel regional que sirvan de insumo para la discusión y planteamiento de nuevas propuestas para el grupo de investigación.

El trabajo ha sido estructurado en cuatro capítulos, en el primero se describen aspectos generales de la competencia y la muestra de datos, en el segundo capítulo se presenta un resumen de los elementos teórico utilizados en el análisis, en el tercer capítulo aparece un análisis descriptivo del rendimiento en las diferentes pruebas aplicadas durante la competencia y en el cuarto capítulo se muestran el análisis de resultados discriminando por género.

## **1. DESCRIPCIÓN DE LA PRUEBA Y DE LA MUESTRA DE DATOS**

### **1.1 DESCRIPCIÓN DE LA PRUEBA**

Las Olimpiadas Regionales de Matemáticas UIS se llevan a cabo teniendo en cuenta cuatro niveles: Básica Primaria (incluido desde 2012), Básico (sexto y séptimo grado), Medio (octavo y noveno grado), Avanzado (décimo y undécimo grado). La competencia consta de cuatro fases: Preparatoria, Clasificatoria, Selectiva y Final. Los estudiantes son convocados a través de anuncios publicitarios (Vanguardia Liberal y medios de comunicación institucionales) adicionales a invitaciones enviadas por correo a ciertos colegios. A continuación una breve descripción de la actividad realizada en cada fase.

Fase preparatoria: de ella participan los estudiantes que aspiran hacer parte de las Olimpiadas Regionales de Matemáticas UIS. Para esta fase el grupo de Olimpiadas de EDUMAT ofrece talleres de capacitación previa a la fase clasificatoria y que son impartidos en las ciudades donde existe sede de la UIS, es decir, Socorro, Málaga, Barbosa, Barranca.

Fase clasificatoria: a esta prueba llegan los estudiantes que voluntariamente se inscriben, la prueba consta de 12 preguntas de selección múltiple con cinco opciones de respuesta. Para los estudiantes de fuera de Bucaramanga, la prueba es enviada a cada sede de la UIS, para cada institución educativa existe un responsable que coordina la realización de la prueba en cada sede de la institución donde haya participantes.

Fase selectica: En esta fase participa el 10% de los estudiantes que tuvieron los mejores puntajes de la fase clasificatoria. La prueba consta de 9 preguntas de las cuales 6 son de selección múltiple y 3 de respuesta tipo ensayo (pregunta abierta). Esta prueba es presentada en el campus universitario de la UIS.

Fase final: En esta fase sólo participan los mejores veinte puntajes de cada nivel. En caso de empate se da prioridad a los que respondieron más preguntas tipo ensayo. La prueba consta de 6 preguntas tipo ensayo, donde el estudiante debe aportar no sólo la respuesta sino una solución detallada. En el nivel avanzado se incluyen demostraciones.

En la siguiente figura se resumen la estructura en cuanto a áreas temáticas y número de preguntas de los cuestionarios utilizados en cada fase de la competencia.

Fase	Clasificatoria	Selectiva	Final
Área Temática			
Teoría de Números y Combinatoria	4 preguntas de selección múltiple	2 preguntas de selección múltiple 1 pregunta tipo ensayo	2 preguntas tipo ensayo
Geometría	4 preguntas de selección múltiple	2 preguntas de selección múltiple 1 pregunta tipo ensayo	2 preguntas tipo ensayo
Álgebra y Lógica	4 preguntas de selección múltiple	2 preguntas de selección múltiple 1 pregunta tipo ensayo	2 preguntas tipo ensayo
<b>Total preguntas</b>	12	9	6

Figura 1 Resumen de estructura en cuanto a áreas temáticas y número de preguntas

### 1.1.1 Forma de calificar cada prueba

Para obtener el puntaje final, el grupo de Olimpiadas ha definido su propia metodología y escala de medición. A continuación se describe brevemente cómo se asignan los puntajes en cada fase de la prueba.

Fase clasificatoria: todos los participantes inician con 12 puntos, si la respuesta es correcta se le otorgan 4 puntos, si no es correcta se le resta un punto, sino contesta ninguna pregunta se quedan con los 12 puntos iniciales, si contesta más de una opción no se restan puntos. El puntaje mínimo es cero y el máximo es 60.

Fase selectiva: cada estudiante inicia con 6 puntos, por cada respuesta correcta en selección múltiple se suman cuatro puntos, si la respuesta es incorrecta se resta un punto. Las preguntas tipo ensayo y las demostraciones tienen un valor de 6 puntos, aquí la puntuación está en relación con la complejidad de la solución, esto es, el grupo ha definido unos parámetros mínimos que debe contener la solución, así sobre la revisión de un procedimiento existe un valor en puntos que se otorga según la solución contenga o no cada parámetro, la suma de esos puntos da el puntaje para la pregunta (ver Anexo A). El puntaje mínimo es 0 y el máximo es 48.

Fase final, cada respuesta tiene un valor máximo de 6 puntos, las preguntas sin contestar no tienen ningún valor. Como todas las preguntas son tipo ensayo se califican como se indicó en el párrafo anterior, así el máximo puntaje posible es 36.

## **1.2 DESCRIPCIÓN DE LA MUESTRA DE DATOS**

La base de datos que servirá de soporte a este proyecto ha sido editada a partir de los registros escritos de las olimpiadas desarrolladas en 2011 y 2012 para estudiantes de Básica Secundaria. La base de datos consta de 7930 registros y 41 variables.

A continuación se hace la descripción de las variables que conforman la base de datos de los resultados de las pruebas de las olimpiadas regionales de matemáticas UIS:

- Año en que se presenta la prueba
- Identificador: nombre del estudiante
- Género: imputada de los registros de inscripción a través del nombre, hay presencia de valores faltantes (7 para 2011 y 11 para 2012)
- Municipio: lugar de procedencia
- Nombre del colegio
- Status del colegio: privado o público, imputada de la las bases de datos del ICFES
- Ubicación del colegio: rural o urbano, imputada de la base de datos del ICFES
- Puntaje\_Prueba\_Clasificatoria, Puntaje\_Prueba\_Selectiva, Puntaje\_Prueba\_Final: variables numéricas que contienen los puntajes obtenidos al presentar la prueba en cada fase. Sólo los finalistas tiene mediciones en las tres variables.
- Respuesta por pregunta: hay 18 variables categóricas que asumen valores 1,2,3,4 o 5 según sea la opción de respuesta escogida a cada pregunta de selección múltiple y 9 variables adicionales que asumen valores de 0 a 6 según la valoración de la solución a una pregunta tipo ensayo.
- Asistencia\_C1, Asistencia\_C2, Asistencia\_3: variables dicotómicas que indica si el participante asistió o no las tres capacitaciones programadas por el grupo que administra la prueba.

*Variables referentes a las pruebas:*

- Para cada pregunta se tiene registrado de la fase donde utilizada, el área temática a la que corresponde (Teoría de números y combinatoria, Geometría, Álgebra y lógica), el tipo de pregunta, y para la fase final en 2012 se tiene acceso a los cuadernillos con las soluciones.

Variables referentes a las capacitaciones:

- Identificador: nombre de los asistentes (incluye aspirantes a participar de la competencia y voluntarios)
- Fecha: presenta valores faltantes.
- Nivel: presenta valores faltantes.
- Información referente a los profesores asistentes: el nombre, institución y género.

## 2. MARCO TEÓRICO

A continuación se presenta un resumen que incluye los elementos teóricos empleados en el análisis de los datos como son definiciones y fórmulas para calcular las medidas numéricas, por ser elementos de estadística básica la descripción se hace en forma breve.

### 2.1 BASE DE DATOS

Es un arreglo rectangular de la información en el cual cada fila tiene el registro para un individuo y cada columna una variable diferente. A fin de importar la información en un programa estadístico, los datos deben no sólo preservar esta estructura sino contar con un adecuado manejo del nombre de variables, etiquetas de valores, valores perdidos, manejo de información numérica y alfanumérica.

### 2.2 ESTADÍSTICA DESCRIPTIVA

Esta rama de la estadística aporta métodos para organizar, resumir y presentar información. En esta rama de la estadística se emplean recursos como gráficas, medidas numéricas de diferentes características de la distribución de los datos como son: tendencia central, dispersión y ubicación que junto a representaciones tabulares pretenden resumir el comportamiento de un grupo de datos.

#### 2.2.1 Medidas de tendencia central

**2.2.1.1 Media muestral:** es una medida de descripción que indica el centro de la distribución de los datos.

Notación:  $\bar{x}$

Para una muestra de  $n$  observaciones,  $x_1, x_2, \dots, x_n$ , su media es

$$\bar{x} = \frac{(x_1 + x_2 + \dots + x_n)}{n}$$

O de la forma,

$$\bar{x} = \frac{1}{n} \sum x_i$$

### **2.2.1.2 La mediana: $\tilde{x}$**

En forma alternativa a la media, otra opción para medir el centro de la distribución de los datos es a través del valor central, es decir, el valor que separa la mitad de las observaciones de la muestra a su izquierda. Para hallar la mediana de una distribución se plantea el siguiente procedimiento (Moore, 1995):

Ordena todas las observaciones de la mínima a la máxima.

Ordena todas las observaciones de la mínima a la máxima.

Si el número de las observaciones  $n$  es impar la mediana  $\tilde{x}$  es la observación central de la lista ordenada. La posición de la mediana se halla contando  $(n+1)/2$  observaciones desde el comienzo de la lista ordenada.

Si el número de observaciones  $n$  es par, entonces la mediana  $\tilde{x}$  es el promedio de las dos observaciones centrales de la lista ordenada. La posición de estas dos observaciones se halla ubicando las posiciones  $n/2$  y  $(n+1)/2$  desde el comienzo de la lista.

Cuando la distribución de los datos es sesgada a la derecha se recomienda utilizar como medida de tendencia central la mediana ya que la media se ve influenciada por la presencia de valores extremos. El gráfico a continuación esboza la ubicación de las medidas de tendencia central en una distribución sesgada a derecha, como se ve la media deja de ser representativa en estos casos.

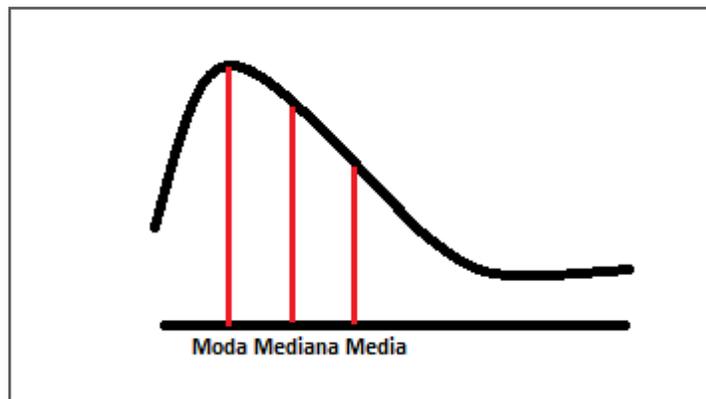


Figura 2. Distribución sesgada a la derecha

## 2.2.2 Medidas de dispersión

**2.2.2.1 Desviación estándar:** La varianza ( $s^2$ ) de un conjunto de observaciones es el promedio de los cuadrados de las desviaciones de las observaciones respecto a su media (Moore, 1995). Para una muestra de datos de  $n$  observaciones,  $x_1, x_2, \dots, x_n$ , la varianza se halla haciendo

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}$$

O de la forma más simple,

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

La desviación estándar es la raíz cuadrada positiva de la varianza  $s^2$ :

$$s = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

La desviación estándar mide la dispersión de las observaciones respecto a la media por lo que debe usarse cuándo ésta sea la medida de tendencia central escogida.

**2.2.2.2 Coeficiente de variación:** El coeficiente de variación permite comparar la dispersión de dos variables medidas en escalas diferentes (ejemplo, peso y estatura), o una misma variable medida en escalas diferentes (ejemplo, peso en kilogramos y peso en toneladas). Se calcula usando la siguiente expresión:

$$\text{Coeficiente de variación } C.V = \frac{S_x}{\bar{x}} \times 100, \bar{x} \neq 0$$

A mayor coeficiente de variación se evidencia mayor variabilidad en la muestra y menor es la representatividad de  $\bar{x}$ . El coeficiente es adimensional por lo que permite comparar las dispersiones de dos distribuciones diferentes.

### **2.2.3 Medidas de localización**

Las medidas de localización dividen la muestra de datos en partes iguales, y permiten resumir el comportamiento de la variable a través de enunciados que hablan de porcentajes de la muestra acumulados a izquierda. Entre estas medidas encontramos los cuartiles y los percentiles. Los cuartiles dividen la población en cuatro partes iguales tomando:

Q<sub>1</sub>: Primer cuartil, acumula a su izquierda 25% de la muestra de datos.

Q<sub>2</sub>: Segundo cuartil o mediana, acumula a su izquierda el 50% de la muestra de datos.

Q<sub>3</sub>: Tercer cuartil, acumula a su izquierda el 75% de la muestra de datos.

La forma de calcular los cuartiles 1 y 3 puede resumirse viendo el cuartil 1 como la mediana de la primera mitad de los datos y el cuartil 3 como la mediana de la segunda mitad de los datos ordenados en orden ascendente.

Percentiles<sup>2</sup> son noventa y nueve valores que dividen la distribución en cien partes iguales, dejando un 1% de las observaciones entre cada dos de ellos consecutivos. Los denotaremos por  $P_i$ ,  $i=1,2,3,\dots,99$ . En el caso de  $P_{75}$ , éste es un valor en el rango de variación de los datos con la característica que a su izquierda acumula el 75% de la muestra de datos. El cálculo se hace análogo al de la mediana, sustituyendo  $n/2$  por  $r.n/k$ , siendo  $r$  el orden del cuantíl y  $k$  las partes en que dicho cuantíl divide a la distribución, así en los percentiles  $k=100$  y  $r=1,2,3,\dots,99$ .

Cada programa implementa algoritmos propios para calcular valores como los cuartiles y percentiles así que puede haber diferencias entre éstos y los obtenidos al usar las fórmulas planteadas en los textos. Los resultados que se mostrarán en el análisis fueron obtenidos al procesar los datos con el programa SPSS versión 13.

#### **2.2.4 Gráfico de caja y bigotes**

Se utiliza cuando la variable es numérica continua, permite ubicar medidas de posición como la mediana, los cuartiles y valores extremos. Además en el gráfico se observan características de la distribución como simetría y sesgos que se identifican al observar la ubicación de la mediana dentro de la caja y la longitud de los bigotes, con esto también puede el experimentador darse una idea de cómo es la dispersión de los datos, se puede usar el concepto de rango intercuartílico ( $Q_3-Q_1$ ). El siguiente gráfico muestra un esquema de los elementos a observar en este recurso:

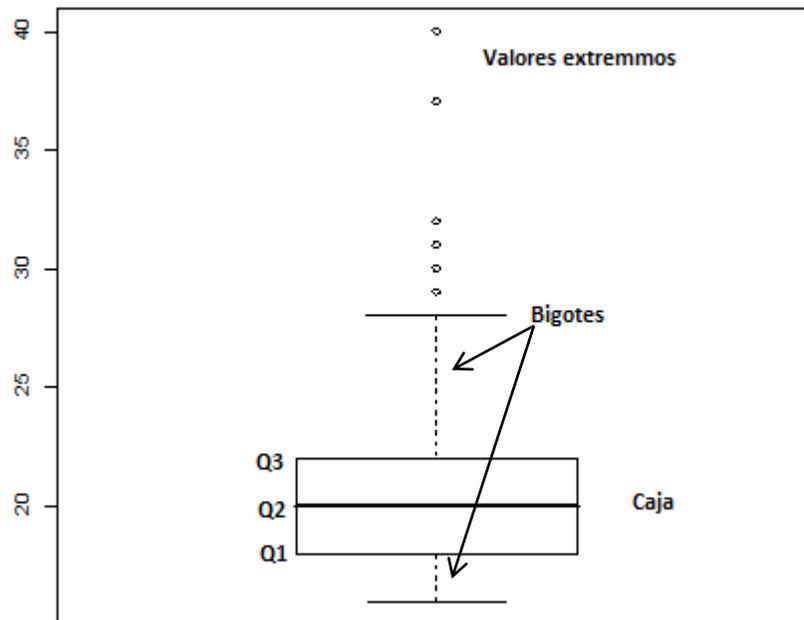


Figura 3. Ejemplo caja y bigotes

### 2.2.5. Histograma

Se utiliza con variables numéricas continuas. El histograma se construye a partir de una tabla de distribución de frecuencias que en su forma más simple tiene dos columnas, una con cierto número de subintervalos de la misma amplitud que inician en el valor más pequeño y terminan en el valor más grande y otra con las frecuencias que en este caso son el número de observaciones de la muestra que pertenecen al intervalo correspondiente. Un histograma de frecuencias se utiliza para examinar características en la distribución de los datos como son: ubicación del máximo, simetría, dispersión, vacíos en la distribución, etc. En el eje vertical se puede utilizar a cambio de las frecuencias absolutas, las frecuencias relativas o una escala de densidad.

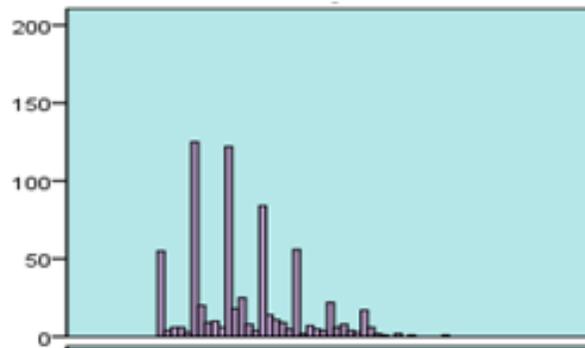


Figura 4. Histograma

### 2.2.6 Grafica de dispersión de puntos

Se utiliza cuando se tienen dos variables numéricas continuas. Permite explorar la dispersión y la presencia de relaciones entre las variables (relación positiva, negativa, independencia); según la ubicación de los puntos y usando herramientas como el color y marcas se pueden hacer visibles posibles agrupamientos en los datos, también deja en evidencia candidatos a ser valores extremos. En algunos casos como el nuestro, se ha utilizado una variable categórica (Nivel en la prueba) a la que se le han asignado etiquetas numéricas para poder construir el gráfico y poder comparar el comportamiento de la variable puntaje al interior de cada subgrupo (cada uno de los tres niveles de competencia) y entre los subgrupos; además se ha usado la opción del color para diferenciar las observaciones por género.

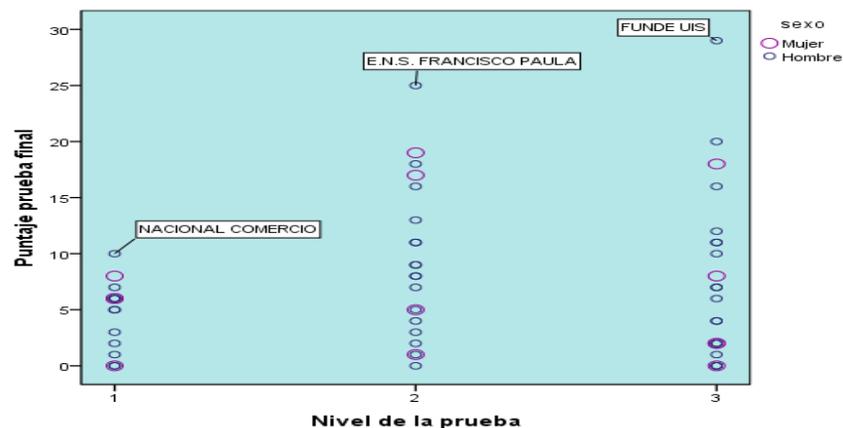


Figura 5. Dispersión de puntos

### 2.2.7 Gráfico de barras de error

Este gráfico permite representar gráficamente la ubicación de la media muestral (círculos que aparecen como punto central de la barra) para ciertos subgrupos según indique la variable en el eje horizontal. Los límites de las barras verticales corresponden a intervalos construidos alrededor de la media, pueden representar intervalos de confianza o un número especificado de errores estándar alrededor de la media. Este gráfico permite comparar la variabilidad de la medida empleada como función de resumen de los datos. En nuestro caso y como se indica en el gráfico a continuación, las barras representan intervalos a una desviación estándar de la media.

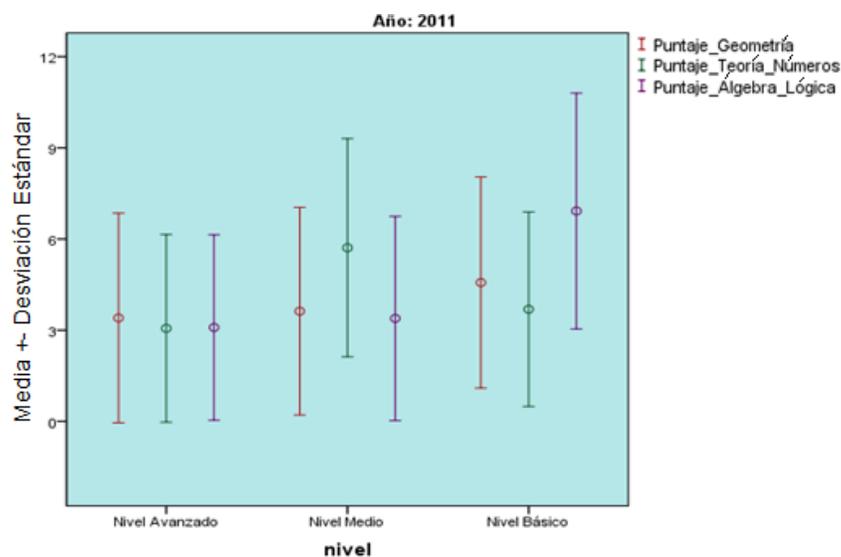


Figura 6. Barras de error

### 2.2.8 Factor

En estadística el término factor se usa para hacer referencia a una variable categórica usada para clasificar las unidades experimentales. Así son factores el género, el colegio, la participación de los estudiantes a las capacitaciones, etc. Lo usual es estudiar la relación entre dos variables una que es considerada variable independiente y otra que se supone depende de ésta a la que se llama variable dependiente. A menos que se tenga un riguroso diseño experimental, los métodos

estadísticos no pueden concluir relaciones de causalidad entre variables, por esto hablamos de asociación entre variables o entre factores y lo común es que sobre las asociaciones detectadas el investigador haga una discusión del por qué se está dando dicho comportamiento.

## 2.3 TABLAS DE CONTINGENCIA

La tabla de contingencia es una tabla que cruza dos variables categóricas o factores, en cada casilla muestra la frecuencia absoluta o número de casos o individuos que poseen el nivel de los factores que allí se interceptan. El tamaño de la tabla de contingencia se define por el número de factores que se analizan conjuntamente y el número de categorías de los mismos. Así por ejemplo el caso más simple es una tabla de contingencia 2x2.

	Factor B	
Factor A	Nivel B <sub>1</sub>	Nivel B <sub>2</sub>
Nivel A <sub>1</sub>	A	b
Nivel A <sub>2</sub>	C	d

Figura 7. Ejemplo tabla de contingencia

En igual forma se pueden tener tablas de contingencia con más de dos factores, y factores con más de dos categorías.

### 2.3.1 Medidas de asociación

Para explorar la relación entre dos variables se usan medidas numéricas y pruebas estadísticas como la chi-cuadrado. En este trabajo no se utilizaran las pruebas chi-cuadrado ya que como requisito debe contarse con una datos provenientes de un muestreo aleatorio o en su defecto de datos poblacionales (Moore, 2005).

En ocasiones, cuando la muestra es grande y se considera que no es demasiado sesgada se implementa ésta prueba para evaluar la hipótesis de relación entre las

variables. En este sentido, debe considerarse que en la implementación de éste recurso los conteos en las celdas deben ser de por lo menos 5, en caso de ser inferiores la sugerencia es crear menos categorías. En nuestro caso, ésta estrategia resulta eficaz porque se crearon 4 niveles para la variable puntaje que discriminan el desempeño en las prueba aplicadas durante la competencia, crear menos categorías conduciría a mezclar grupos muy diferentes de participantes.

Como alternativa se puede hacer uso de análisis descriptivos como la comparación de distribuciones de frecuencias relativas condicionales de una variable para los distintos valores de la otra variable, comparar todas las frecuencias relativas condicionadas de una variable para un único valor de la otra variable con la frecuencia marginal de la primera variable y comparación de posibilidades a favor y en contra de B en cada valor de A (Cañadas, Batanero y otros, 2011). A continuación describimos cuáles de estos elementos serán incorporados en el análisis.

### **2.3.2 Distribución de frecuencia**

En el caso de una tabla de contingencia de dimensión  $r \times s$ , en cada celda además de la frecuencia absoluta (número de veces que aparece el valor) se pueden agregar tres tipos de frecuencias: por fila, por columna y sobre el total. Cada una de éstas habla de una distribución de frecuencias diferente. En este caso se utiliza  $n_{ij}$  para representar la frecuencia absoluta,  $n_i$  el total por fila,  $n_j$  indica el total por columna y  $n$  el total de datos en la muestra. Gráficamente la tabla tiene el siguiente aspecto:

		Factor B			
Factor A	Nivel B <sub>1</sub>	Nivel B <sub>2</sub>	...	Nivel B <sub>r</sub>	Total
Nivel A <sub>1</sub>	$n_{11}$	$n_{12}$	...	$n_{1r}$	$n_{1.}$
Nivel A <sub>2</sub>	$n_{21}$	$n_{22}$	...	$n_{2r}$	$n_{2.}$
	:	:	:	:	:
Nivel A <sub>s</sub>	$n_{s1}$	$n_{s2}$	...	$n_{sr}$	$n_{s.}$
Total	$n_{.1}$	$n_{.2}$	...	$n_{.r}$	$n$

Figura 8. Ejemplo distribución de frecuencia

La distribución empírica (basada en los datos de la muestra) del Nivel A<sub>1</sub> se obtiene hallando las frecuencias por fila al dividir cada frecuencia absoluta por  $n_{1.}$ , esto es:

		Factor B			
Factor A	Nivel B <sub>1</sub>	Nivel B <sub>2</sub>	...	Nivel B <sub>r</sub>	Total
Nivel A <sub>1</sub>	$n_{11} / n_{1.}$	$n_{12} / n_{1.}$	...	$n_{1r} / n_{1.}$	$n_{1.} / n_{1.}$

Figura 9. Distribución de frecuencia

Así cuando se desea comparar la distribución de cada nivel del factor A, se comparan estos valores escritos como frecuencias o como porcentajes. En este caso particular lo que se tiene es la distribución del factor B condicional a A<sub>1</sub>. En igual forma se calcularían las demás distribuciones por fila para analizar cómo se comporta el Factor B según los otros niveles del Factor A.

## 2.4 RAZÓN DE ODDS (OR)

Es una medida para estimar la fuerza de asociación entre un factor de riesgo, que en nuestro caso será el género y la presencia de un evento (acertar o no una pregunta). Se calcula haciendo el cociente entre el odds de que ocurra el evento entre quienes están expuestos al factor de riesgo y el odds correspondiente para los no expuestos, es decir que consideraremos una tabla:

	Pregunta	
Género	Respuesta incorrecta	Respuesta correcta
Mujer	A	B
Hombre	C	D

Figura 10. Tabla para hallar la razón de odds.

y el cálculo se hace a través de la expresión  $OR = \frac{a/b}{c/d}$

Interpretación:

- 1)  $OR=1$  La frecuencia de exposición es la misma “no hay asociación entre el factor de riesgo y el evento en el estudio.
- 2)  $OR>1$  La frecuencia de presentación del evento es mayor en los expuestos que en los no expuestos (el factor en estudio es factor de riesgo)
- 3)  $OR<1$  El factor en estudio es un factor protector (Lo contrario a 2)

## 2.5 RIESGO RELATIVO (RR)

Es un indicador que compara el riesgo entre dos cohorte de personas (dos poblaciones) una expuesta y otra no expuesta a algún factor determinado (factor de riesgo). Se calcula como el cociente entre el riesgo de los expuestos y los no expuestos (Woodward, 1999).

$$OR = \frac{a/(a+b)}{c/(c+d)}$$

Interpretación:

- 1)  $RR=1$  (cercano a 1) indica que el riesgo del evento es el mismo en los expuestos que en los no expuestos (no es factor de riesgo para el evento)

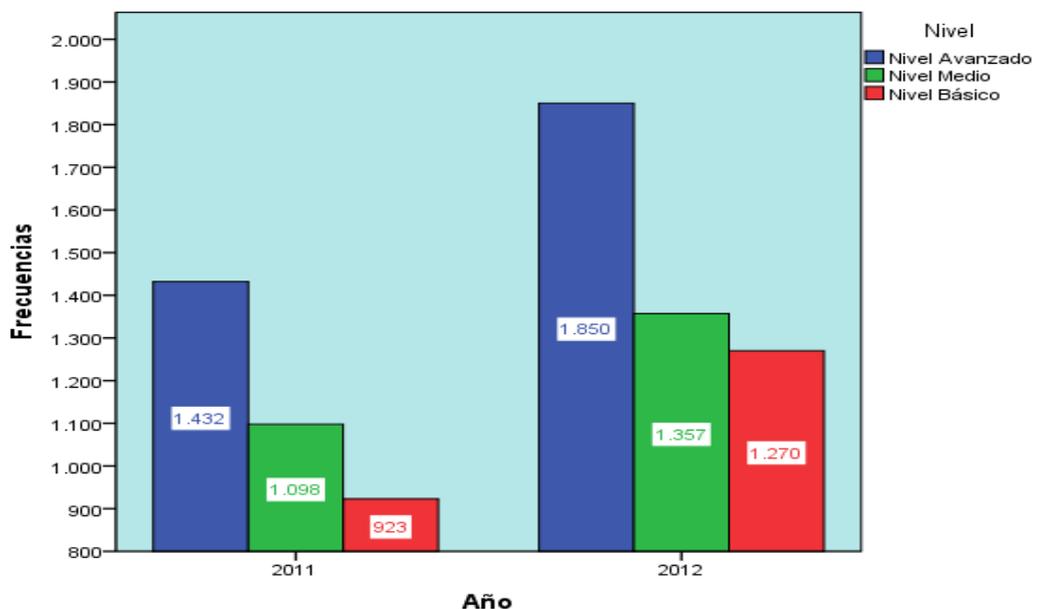
2)  $RR > 1$  indica que el riesgo en el grupo expuesto es mayor que en el grupo no expuesto.

### 3. ANÁLISIS DESCRIPTIVO

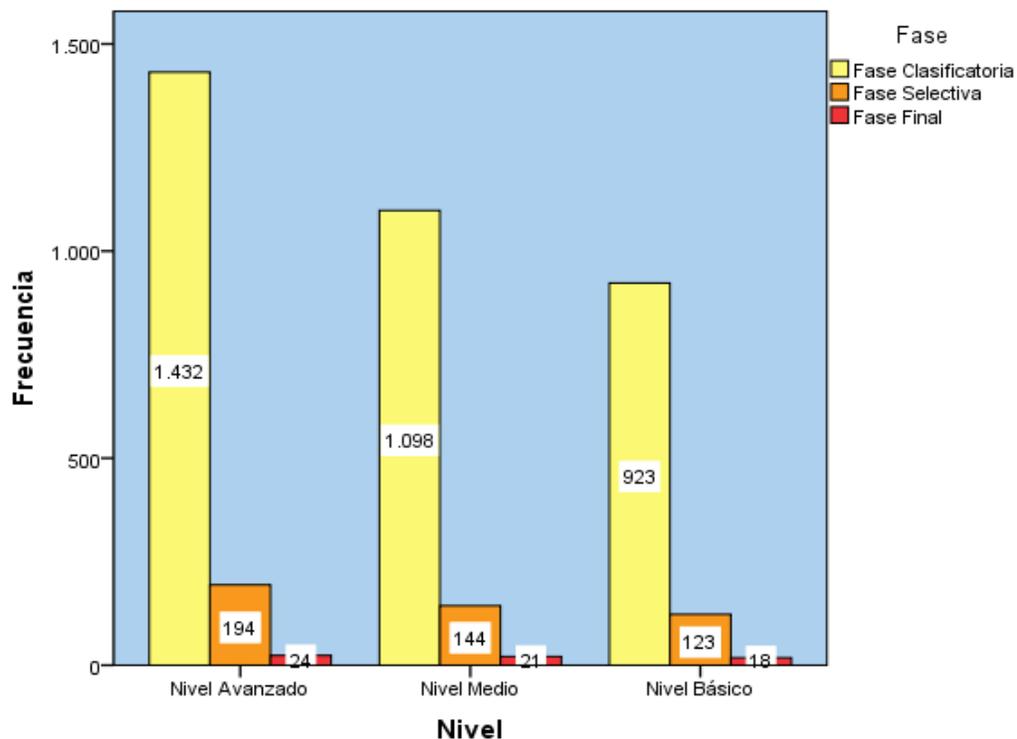
#### 3.1 NIVEL DE PARTICIPACIÓN EN LA PRUEBA

Desde la primera versión de las Olimpiadas regionales de Matemáticas en Santander realizadas en 2009 se ha observado una considerable participación en la prueba, el registro oficial indica que se inició en 2009 y 2010 con 4518 y 3013 participantes y durante el 2011 y 2012 se tuvieron 3453 y 4477 inscritos respectivamente. Los Gráficos 1, 2 y 3 ilustran la participación en 2011 y 2012 discriminando por el nivel de la competencia (Avanzado, Medio y Básico) y por fase (Clasificatoria, Selectiva y Final). De las gráficas referidas es evidente una tendencia creciente en cuanto al nivel de participación en todos los niveles de la prueba; también hay que destacar que la más alta participación, alrededor del 41%, se da en el nivel avanzado, seguido del nivel medio con 31.8% y 30.3%. En cuanto a participación por fase cabe recordar que a la primera fase se inscriben voluntariamente y de allí en adelante son los puntajes los que determinan la permanencia dentro de la competencia.

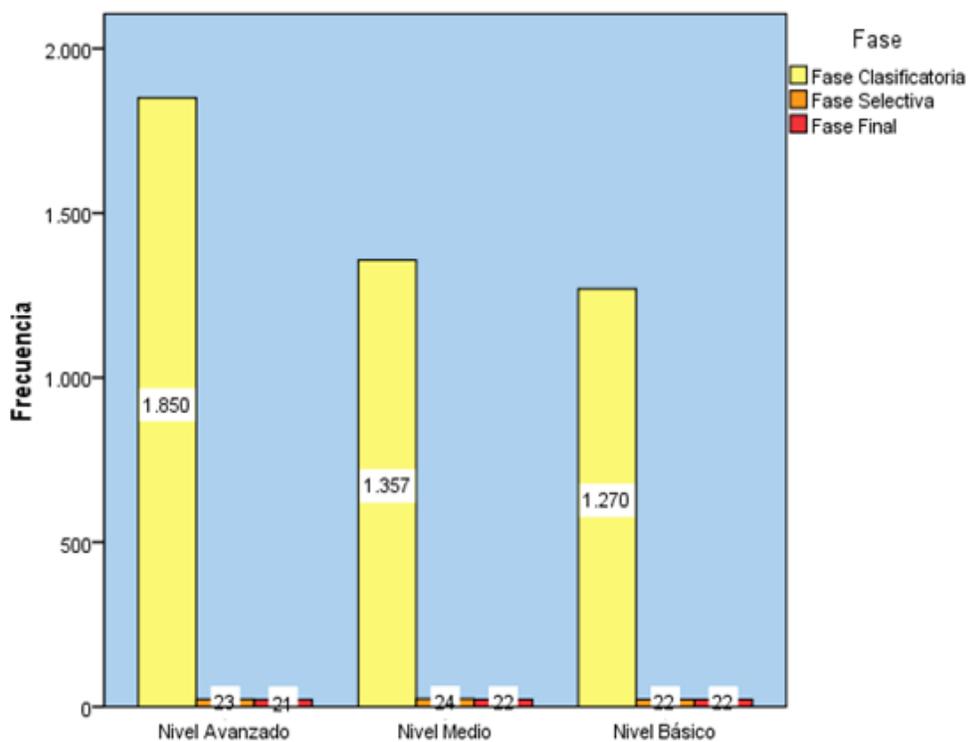
**Gráfica 1. Participantes por nivel (2011-2012)**



**Gráfica 2. Participación por fase (2011)**



**Gráfica 3. Participación por fase (2012)**



Sobre la participación según la variable status del colegio (Gráfico 4), se observa una amplia mayoría de participantes provenientes de colegios oficiales, 83% y 85% en 2011 y 2012 respectivamente. En igual forma, al discriminar la participación por zona se tiene que del total de participantes en cada año el 92% y 86% provenían de colegios urbanos (Gráfico 5). Finalmente, al considerar el género se tiene una distribución más equitativa, en 2011 se tuvo una participación femenina del 47.6% mientras que en 2012 pasó a ser del 45% (ver Gráfico 6).

En la Tabla 1 se registran los resultados que permiten concluir acerca de la significancia de las diferencias observadas en las proporciones de participación por status del colegio, por zona y por género observadas en 2011 y 2012. De ahí se concluye que hay diferencias significativas en la participación de colegios del sector oficial en el nivel avanzado y medio; también resultan significativas las comparaciones de las proporciones de participación de colegios del sector urbano y en cuanto a la diferencia de proporciones de mujeres observadas en cada nivel de la prueba en esos dos años sólo la comparación en el nivel avanzado resulta no ser significativa ( $\alpha=0,05$ ).

**Tabla 1. Prueba de hipótesis para diferencia de proporciones**

Año	Nivel	Año		Z	Valor p
		2011	2012		
$\hat{P}_{0, \text{oficial}}$	Avanzado	0,865223464	0,905945946	-3,6712046	0*
	Medio	0,802367942	0,834929993	-2,0896789	0,037*
	Básico	0,820151679	0,803937008	0,9570624	0,3421

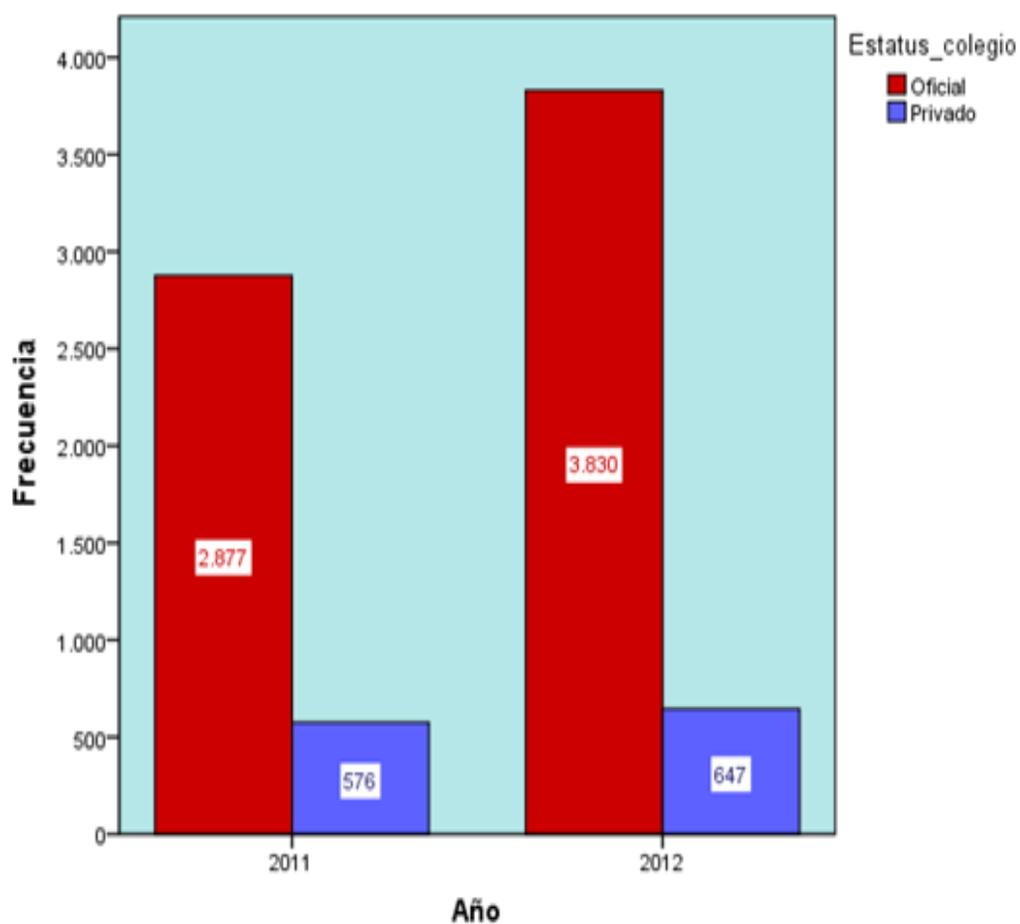
Año	Nivel	Año		Z	Valor p
		2011	2012		
$\hat{P}_{\text{Urbano}}$	Avanzado	0,88407821	0,84918919	2,89538152	0,0038*
	Medio	0,90528233	0,86366986	3,18098928	0,001*

	Básico	1	0,87480315	11,161984	0*
--	--------	---	------------	-----------	----

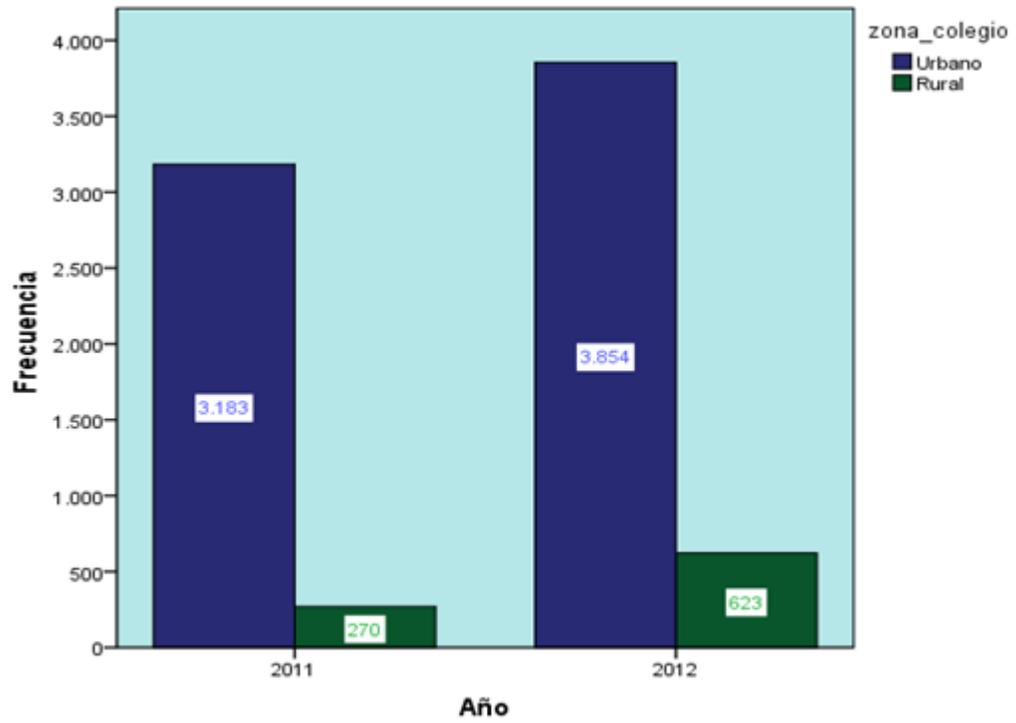
Año	Nivel	Año		Z	Valor p
		2011	2012		
$\hat{P}_{Mujeres}$	Avanzado	0,483893557	0,478850325	0,57310084	0,56
	Medio	0,464480874	0,423929099	4,04842123	0*
	Básico	0,477173913	0,44637224	2,86383519	0,004*

\*Prueba significativa al 5%.

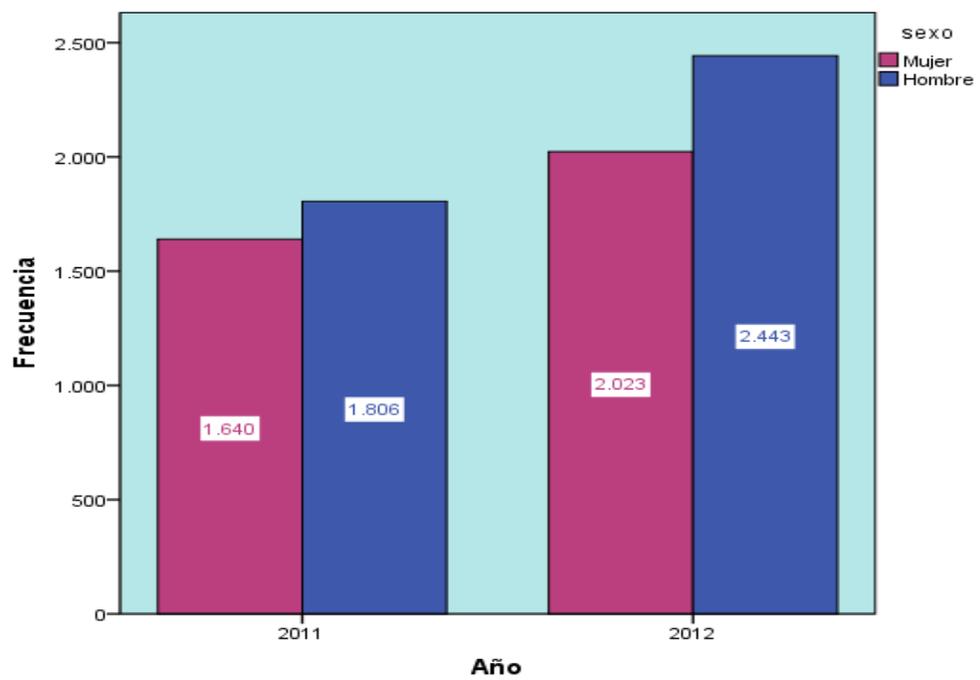
**Gráfica 4. Participación por colegios oficiales y privados (2011-2012)**



**Gráfica 5. Participación por colegios rurales y urbanos (2011-2012)**



**Gráfica 6. Participación por género (2011-2012)**



Para analizar la participación por colegios se elaboró la Tabla 15 (ver Anexo B) en la cual aparecen discriminados los colegios participantes en cada año según el nivel para el cual se inscribieron sus estudiantes. En la tabla se muestran los 55 colegios que participaron en 2011 y los 64 de 2012; en negrilla aparecen resaltados los colegios con la mayor participación observada. En el nivel avanzado los colegios que más enviaron representantes durante el 2011 son Nacional de Comercio(Bucaramanga), Colegio Técnico Microempresarial el Carmen (Florida), Colegio Centro de Comercio y Colegio Isabel Valbuena Cifuentes (Vélez) en el nivel avanzado, San José de la Salle (Bucaramanga), Nacional de comercio (Bucaramanga) y la Escuela Normal Superior de Oiba, y en el nivel básico Colegio la Presentación(Bucaramanga), Escuela Normal Superior de Oiba y Colegio Industrial del Socorro; durante el 2012 los colegios que más enviaron representantes para el nivel avanzado fueron el Colegio Nacional de Comercio (Bucaramanga), Colegio Microempresarial el Carmen (Florida), Colegio Isabel Valbuena Cifuentes(Vélez), en el nivel medio la Escuela Normal Superior de Oiba, Colegio Isabel Valbuena Cifuentes(Vélez),y el Gimnasio Superior (Florida) y para el nivel Básico el Gimnasio Superior (Florida), Colegio Nacional de Comercio (Bucaramanga) y Colegio Microempresarial el Carmen (Florida); se puede notar que los colegios que más envían participantes durante estos años para casi todos los niveles son casi siempre los mismos.

Sobre la participación por lugar geográfico destacamos que de los 87 municipios en 2011 participaron 17 y en 2012 fueron 25. En el siguiente mapa de Santander<sup>3</sup> aparece un ovalo verde para indicar el sitio de donde hubo representación en las olimpiadas (2011-2012), como se ve si comparamos por provincia es la de Mares la de menor participación. La lista completa de municipios discriminando por año aparece en la Tabla 16 (Anexo C), del listado puede notarse que además de lo esperado con Bucaramanga, Floridablanca y Barrancabermeja, son en 2011 Oiba y en 2012 Oiba y Vélez los de más alta participación.

**Gráfica 7. Mapa de Santander de los municipios que participaron en las Olimpiadas Regionales de Matemáticas UIS (2011-2012)**



### 3.2 RENDIMIENTO EN LA PRUEBA

Para analizar el rendimiento en la prueba se considerará en primera instancia los puntajes totales obtenidos en cada fase de la prueba, luego se mostrarán análisis según el nivel de desempeño alcanzado y finalmente se harán algunas comparaciones del desempeño por área temática. En todos los casos se

considera la base de datos segmentada por las variables año y nivel; también se utilizará como medida de tendencia central la mediana ( $\tilde{x}$ ) ya que las distribuciones de los puntajes son asimétricas a derecha. Como medida de variación reportaremos el coeficiente de variación (C.V) ya que la escala de medición usada en cada fase de la competencia es diferente.

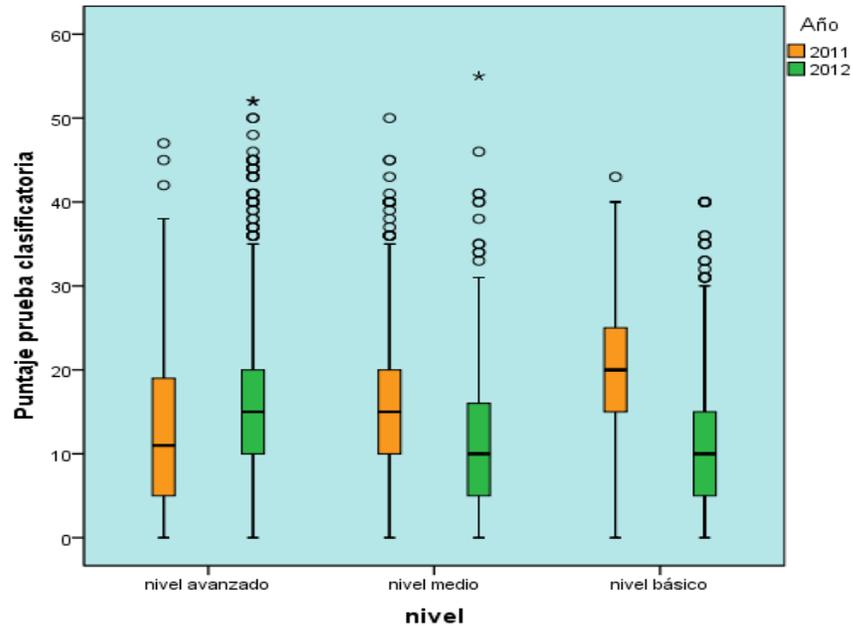
### 3.2.1 Rendimiento en la fase clasificatoria

La Gráfica 8 muestra las distribuciones de los puntajes obtenidos en la prueba clasificatoria discriminando por año. Común a los tres niveles de la prueba se observa que son pocos los participantes que alcanzan puntajes altos en la escala, recordemos que aquí los puntajes oscilan entre 0 y 60 puntos. En particular, los puntajes más altos fueron 47, 50, 43 puntos para el nivel avanzado, medio y básico respectivamente, lo que indica que el máximo número de respuestas correctas esta entre 8 y 9, en una prueba que consta de 12 preguntas de selección múltiple con respuesta única.

Para complementar lo observado en la Gráfica 8 se tiene la Tabla 2 con los estadísticos descriptivos básicos de resumen calculados para la variable puntaje en cada fase. Así, para 2011 se observa un mejor rendimiento en el nivel básico, la mediana de este nivel está por encima del percentil 75 ( $P_{75}$ ) del nivel avanzado y medio. En cambio en 2012 es el nivel avanzado quien exhibe mejoría en el rendimiento, la mediana pasa de ser 11 a ser 15, hay mayor número de valores inusuales y extremos en la cola derecha de la distribución y una leve reducción en la variabilidad ( $CV_{2011}=62.6\%$  ;  $CV_{2012}=60.55$ ); los otros dos niveles muestran disminución en el rendimiento evidenciado en la reducción del valor de la mediana de los puntajes ( $\tilde{x}_{2011\_medio} = 15$ ,  $\tilde{x}_{2012\_medio} = 10$ ,  $\tilde{x}_{2011\_básico} = 20$ ,  $\tilde{x}_{2012\_básico} = 10$ ), también hay aumento en la variabilidad de los puntajes ( $CV_{2011\_Medio} = 49.3\%$ ,  $CV_{2012\_Medio} = 61.8\%$  ;  $CV_{2011\_Básico} = 40\%$ ,  $CV_{2012\_Básico} = 65\%$ )

que se explica por la presencia de más valores inusuales en la cola derecha de la distribución, hecho positivo en este contexto pues indica que en 2012 hubo más estudiantes con rendimiento destacado que en 2011.

**Gráfica 8. Rendimiento en la prueba clasificatoria discriminando por año y nivel**



**Tabla 2. Resumen de estadísticos descriptivos del rendimiento en la prueba en cada fase de la competencia (2011-2012)**

Año	Nivel	Clasificatoria				Selectiva				Final			
		$\bar{x}$	$\tilde{x}$	$S$	$C.V$	$\bar{x}$	$\tilde{x}$	$S$	$C.V$	$\bar{x}$	$\tilde{x}$	$S$	$C.V$
2011	Avanzado	13,1	11	8,21	62,6	13,	13	7,07	51,9	7,25	5	7,50	103,5
	Medio	16,7	15	8,26	49,3	14,9	13	8,73	58,8	9,14	8	6,82	74,6
	Básico	19,7	20	7,89	39,9	12,4	11	7,89	63,5	4,61	6	2,93	63,5

2012	Avanzado	15,3	15	9,28	60,5	19,3	17	4,72	24,4	9,90	7	8,93	90,2
	Medio	12,1	10	7,52	61,8	21,47	20,5	3,94	18,3	6,14	6	5,12	83,3
	Básico	11,3	10	7,40	65	22,09	21	3,29	14,9	7,59	6,5	5,73	75,4

Para complementar el análisis anterior se definió la variable Nivel\_Prueba\_Clasificatoria que es una variable que categoriza el Puntaje obtenido de la siguiente manera:

Si  $0 \leq \text{Puntaje\_Prueba\_Clasificatoria} \leq 24$ , Nivel\_Prueba\_Clasificatoria = Bajo

Si  $25 \leq \text{Puntaje\_Prueba\_Clasificatoria} \leq 36$ , Nivel\_Prueba\_Clasificatoria = Aceptable.

Si  $37 \leq \text{Puntaje\_Prueba\_Clasificatoria} \leq 48$ , Nivel\_Prueba\_Clasificatoria = Superior.

Si  $49 \leq \text{Puntaje\_Prueba\_Clasificatoria} \leq 60$ , Nivel\_Prueba\_Clasificatoria = Alto

La distribución de frecuencias acorde al nivel de desempeño alcanzado en la prueba clasificatoria diferenciando por nivel de la competencia se muestra en la Tabla 3. Como ya mencionamos uno de los hechos más notorios es que la mayor cantidad de participantes se ubican en el nivel bajo, sólo 6 estudiantes han alcanzado niveles superiores de desempeño en esta fase de la competencia durante las dos versiones consideradas. Una posible explicación es que varios colegios no hacen una selección interna de los participantes sino que envían a todos quienes voluntariamente deseen participar. Si se comparan los porcentajes de 2012 respecto a los alcanzados en 2011 como se advirtió anteriormente sólo el

nivel avanzado muestra incrementos, aunque poco considerables, en los niveles aceptable, superior y alto.

**Tabla 3. Clasificación Nivel vs Nivel\_Prueba\_clasificatoria**

				Rendimiento en la prueba clasificatoria				
Año				Prueba clasificatoria				Total
				Bajo	Aceptable	Superior	Alto	
2011	nivel	Nivel Avanzado	Frecuencia	1248	178	6	0	1432
			%	<b>87,2%</b>	12,4%	,4%	,0%	100,0%
	Nivel Básico	Frecuencia	612	294	17	0	923	
		%	<b>66,3%</b>	31,9%	1,8%	,0%	100,0%	
	Nivel Medio	Frecuencia	885	197	15	1	1098	
		%	<b>80,6%</b>	17,9%	1,4%	,1%	100,0%	
Total		Frecuencia	2745	669	38	1	3453	
		%	79,5%	19,4%	1,1%	,0%	100,0%	
2012	nivel	Nivel Avanzado	Frecuencia	1514	298	34	4	1850
			%	<b>81,8%</b>	16,1%	1,8%	,2%	100,0%
	Nivel Básico	Frecuencia	1179	84	7	0	1270	
		%	<b>92,8%</b>	6,6%	,6%	,0%	100,0%	
	Nivel Medio	Frecuencia	1253	97	6	1	1357	
		%	<b>92,3%</b>	7,1%	,4%	,1%	100,0%	
Total		Frecuencia	3946	479	47	5	4477	
		%	88,1%	10,7%	1,0%	,1%	100,0%	

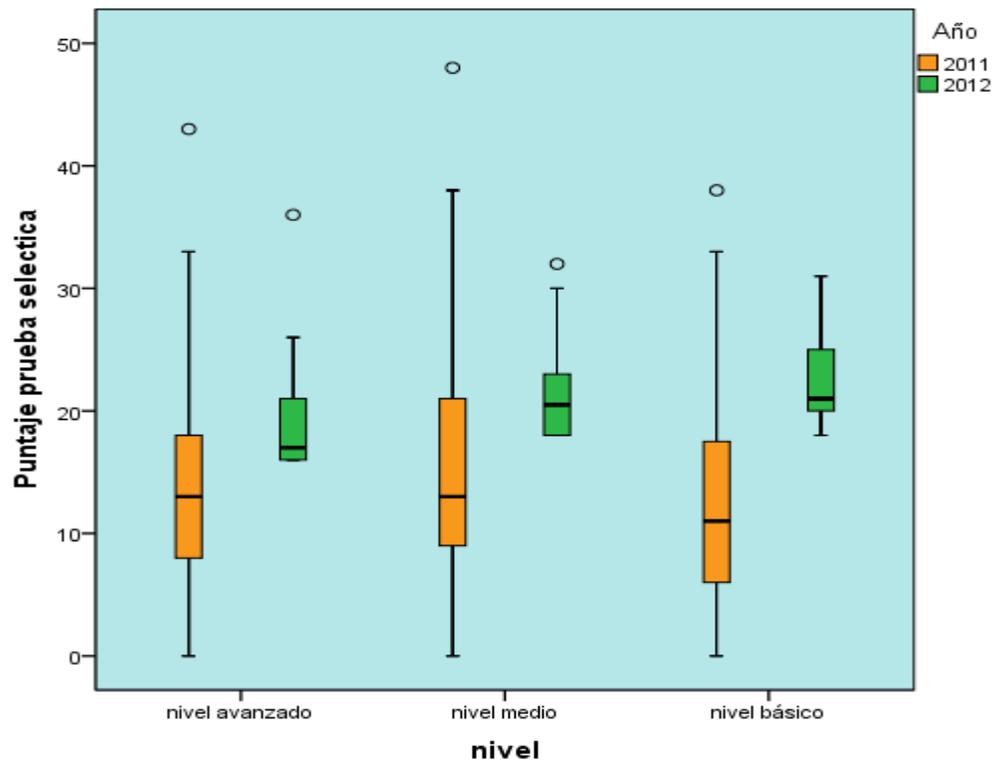
En el Anexo D aparecen las Tablas 17 y 18, éstas son tablas de contingencia donde se cruzan las variables Zona (rural o urbana) y Status del colegio (Oficial o privado) con el nivel de desempeño alcanzado en la prueba clasificatoria. Se observa un comportamiento similar en la distribución de frecuencia de cada uno de estos subgrupos al observado sin discriminar por estas variables, es decir que al parecer la zona y el status del colegio no generan diferencias en el nivel de desempeño alcanzado en la prueba clasificatoria, en todos los casos predominan los niveles aceptable y bajo, contrario a lo esperado el rendimiento de los participantes provenientes de colegios privados no es superior a los de los

oficiales, sólo en 2011 en el nivel avanzado y básico es claro un pequeño margen de error a favor de los colegios privados. La expectativa obedece a que se cree que los colegios privados si hacen una rigurosa selección de los participantes que garantice una representación exitosa.

### **3.2.2 Rendimiento en la prueba selectiva:**

En el caso de los puntajes obtenidos en la prueba selectiva (ver Gráfica 9) hay que tomar en cuenta que según el registro en 2011 se seleccionaron cerca del 13% de los mejores puntajes en la prueba clasificatoria mientras que en 2012 se seleccionaron cerca del 2% de los participantes de la fase previa para continuar en competencia. Lo anterior puede ser la explicación del por qué en 2012 se observa un considerable aumento en el rendimiento en la prueba selectiva aplicada a los tres niveles y una reducción significativa en la variabilidad (ver Tabla 2). También por resaltar tenemos que nuevamente hay pocos participantes en la zona de rendimiento destacado, en el caso del nivel avanzado en 2011 sólo el 25% de los participantes en la prueba selectiva obtuvieron puntajes superiores a 18 en una escala donde el máximo puntaje a obtener es 48, en igual forma en 2012 el percentil 75 ( $P_{75}$ ) fue de 21 puntos. Igual comportamiento se observa para el nivel medio y el básico ( $P_{75\_2011\_Medio}=21$  ,  $P_{75\_2012\_Medio}=23$ ;  $P_{75\_2011\_Básico}=18$ ,  $P_{75\_2012\_Básico}=25$ ).

**Gráfica 9. Rendimiento en la prueba selectiva discriminando por año y nivel**



Para complementar el análisis anterior se definió la variable Nivel\_Prueba\_Selectiva que es una variable categórica definida de la siguiente manera:

Si  $0 \leq \text{Puntaje\_Prueba\_Selectiva} \leq 12$ , Nivel\_Prueba\_Selectiva = Bajo

Si  $13 \leq \text{Puntaje\_Prueba\_Selectiva} \leq 24$ , Nivel\_Prueba\_Selectiva = Aceptable

Si  $25 \leq \text{Puntaje\_Prueba\_Selectiva} \leq 36$ , Nivel\_Prueba\_Selectiva = Superior

Si  $37 \leq \text{Puntaje\_Prueba\_Selectiva} \leq 48$ , Nivel\_Prueba\_Selectiva = Alto

La distribución de frecuencias acorde al nivel de desempeño en la prueba selectiva en cada nivel de la competencia se muestra en la Tabla 4. Como hecho importante notamos que la mayoría de los participantes se ubican en un nivel bajo de desempeño (puntaje menor a 12 que indica no más de una respuesta correcta

de selección múltiple o dos procedimientos parcialmente correctos en las preguntas tipo ensayo) durante el año 2011, en forma positiva se destaca que 5 estudiantes alcanzaron un nivel alto de desempeño en esta fase de la competencia (puntaje superior a 37), por el contrario durante el año 2012 se puede observar que ningún estudiante se encuentra en un nivel bajo pero tampoco hay estudiantes en el nivel alto de desempeño. Una posible explicación de este hecho puede ser la diferencia entre los niveles de dificultad de las prueba aplicadas durante el 2011 y 2012 y la presencia de preguntas tipo ensayo en la fase selectiva, preguntas que suelen ser de mayor dificultad para los participantes.

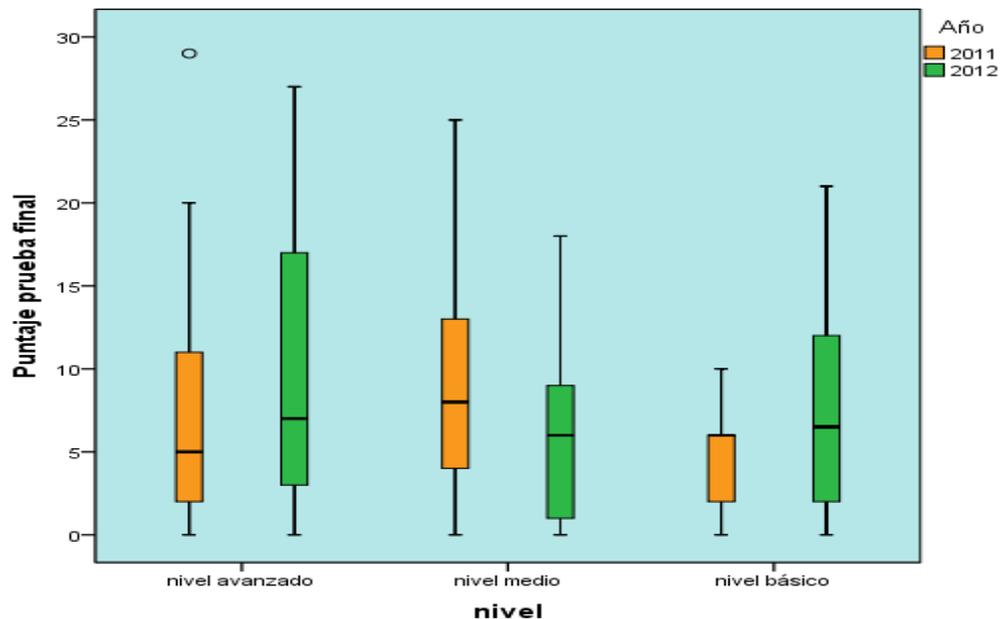
**Tabla 4. Nivel de la prueba vz nivel de desempeño en la prueba selectiva (2011-2012)**

Año				Nivel_Prueba_Selectiva				Total
				Bajo	Aceptable	Alto	Superior	
2011	Nivel Avanzado	Nivel	Frecuencia	93	88	1	12	194
			%	<b>47,9%</b>	<b>45,4%</b>	,5%	6,2%	100,0%
	Nivel Básico	Nivel	Frecuencia	72	42	1	8	123
			%	<b>58,5%</b>	<b>34,1%</b>	,8%	6,5%	100,0%
	Nivel Medio	Nivel	Frecuencia	67	57	3	17	144
		%	<b>46,5%</b>	<b>39,6%</b>	2,1%	11,8%	100,0%	
	Total		Frecuencia	232	187	5	37	461
			%	50,3%	40,6%	1,1%	8,0%	100,0%
2012	Nivel Avanzado	Nivel	Frecuencia		21		2	23
			%		<b>91,3%</b>		8,7%	100,0%
	Nivel Básico	Nivel	Frecuencia		16		6	22
			%		<b>72,7%</b>		27,3%	100,0%
	Nivel Medio	Nivel	Frecuencia		20		4	24
		%		<b>83,3%</b>		16,7%	100,0%	
	Total		Frecuencia		57		12	69
			%		82,6%		17,4%	100,0%

### 3.2.3 Rendimiento en la fase final:

En la fase final se tiene en competencia a los mejores participantes, el número oscila entre 20 y 25 en cada nivel. En el nivel avanzado y básico se observa una leve mejoría en el rendimiento en 2012 comparado con 2011, hay más participantes en la zona de alto rendimiento (ver gráfica 10). En el caso del nivel avanzado el percentil 75 se ubica en 11 mientras que este mismo valor para 2012 es 17, en cuanto a la variabilidad los coeficientes de variación indican la presencia de grupos muy heterogéneos ( $CV_{2011}=103\%$ ,  $CV_{2012}=90\%$ ), hay que resaltar que la variación es alta debido al incremento en el rango de variación en la parte derecha de la distribución (zona de los mejores puntajes). Las observaciones hechas para el nivel avanzado se dan también en el nivel básico (ver Tabla 2), por el contrario en el nivel medio toda la distribución se desplaza hacia abajo (los cuartiles de la distribución en 2012 son más bajos) y hay una alta variabilidad ( $CV_{2011}=74,6\%$ ,  $CV_{2012}=83\%$ ).

**Gráfica 10. Rendimiento en la prueba final discriminando por año y nivel (2011-2012)**



Para complementar el análisis anterior se definió la variable Nivel\_Prueba\_Final que es una variable categórica definida de la siguiente manera:

Si  $0 \leq \text{Puntaje\_Prueba\_Final} \leq 9$ , Nivel\_Prueba\_Final = Bajo

Si  $10 \leq \text{Puntaje\_Prueba\_Final} \leq 18$ , Nivel\_Prueba\_Final = Aceptable

Si  $19 \leq \text{Puntaje\_Prueba\_Final} \leq 27$ , Nivel\_Prueba\_Final = Superior

Si  $28 \leq \text{Puntaje\_Prueba\_Final} \leq 36$ , Nivel\_Prueba\_Final = Alto

La distribución de frecuencias acorde al nivel alcanzado en la prueba Final en cada nivel de la competencia se muestra en la Tabla 5. La tabla muestra que al igual que en la fase clasificatoria la mayoría de los participantes se ubican en un nivel de desempeño bajo para los años 2011 y 2012, sólo 1 estudiante alcanzó el nivel alto de desempeño en esta fase de la competencia durante el 2011. Una posible explicación a este hecho se debe a que la prueba final consta de 6 preguntas tipo ensayo por lo que el nivel de dificultad durante esta fase es mayor que en las fases anteriores y la escala de calificación inicia en cero. Igual observación se puede hacer para el nivel de desempeño superior (puntajes entre 28 y 36 puntos en una escala de 0 a 36) donde sólo se ubican 8 de los finalistas en este periodo de estudio.

**Tabla 5. Nivel de la prueba vs nivel de desempeño en ña prueba final (2011-2012)**

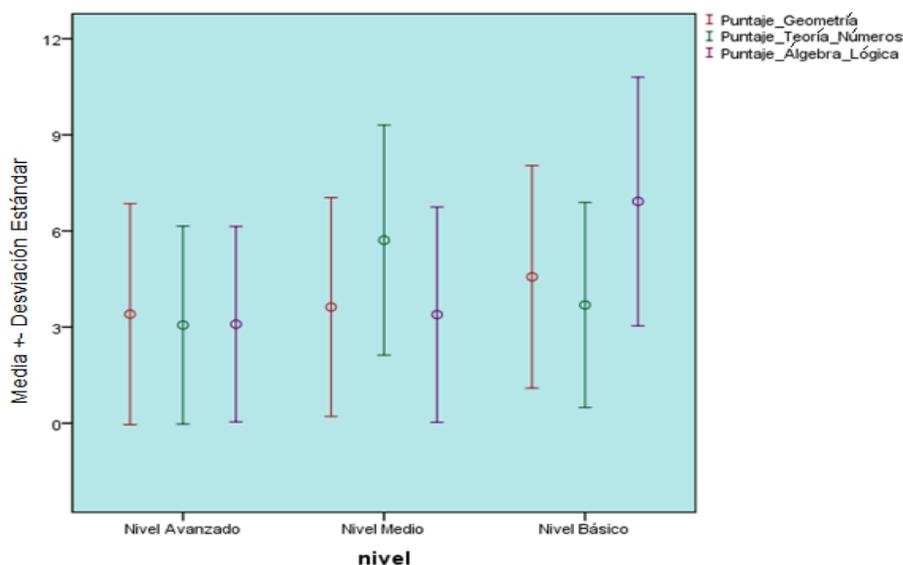
Año				Nivel_Prueba_Final				Total
				Bajo	Aceptable	Superior	Alto	
2011	Nivel Avanzado	Nivel Avanzado	Frecuencia %	16 66,7%	6 25,0%	1 4,2%	1 4,2%	24 100,0%
	Nivel Básico	Nivel Básico	Frecuencia %	17 94,4%	1 5,6%	0 ,0%	0 ,0%	18 100,0%
	Nivel Medio	Nivel Medio	Frecuencia %	13 61,9%	6 28,6%	2 9,5%	0 ,0%	21 100,0%
	Total	Total	Frecuencia %	46 73,0%	13 20,6%	3 4,8%	1 1,6%	63 100,0%
2012	Nivel Avanzado	Nivel Avanzado	Frecuencia %	12 57,1%	5 23,8%	4 19,0%		21 100,0%
	Nivel Medio	Nivel Medio	Frecuencia %	15	6	1		22

	Básico	%	68,2%	27,3%	4,5%		100,0%
Nivel	Frecuencia		17	5	0		22
Medio	%		77,3%	22,7%	,0%		100,0%
Total	Frecuencia		44	16	5		65
	%		67,7%	24,6%	7,7%		100,0%

### 3.2.4 Rendimiento por área temática:

Los resultados que se muestran a continuación fueron hallados considerando la fase clasificatoria para la cual se definieron tres nuevas variables, Puntaje\_Geometría, Puntaje\_Algebra\_Lógica y Puntaje\_Teoría de números. Es pertinente recordar que la prueba en esta fase contiene 4 preguntas de selección múltiple por cada área temática para un total de 12 preguntas, cada respuesta correcta vale 4 puntos con lo cual el mínimo puntaje en cada área es cero y el máximo 16. A continuación, las gráficas 11 y 12 muestran el gráfico de barras de error para las variables puntaje por área temática, como se lee del título que acompaña al eje vertical la gráfica representa los intervalos construidos alrededor de la media usando una desviación estándar y no intervalos de confianza.

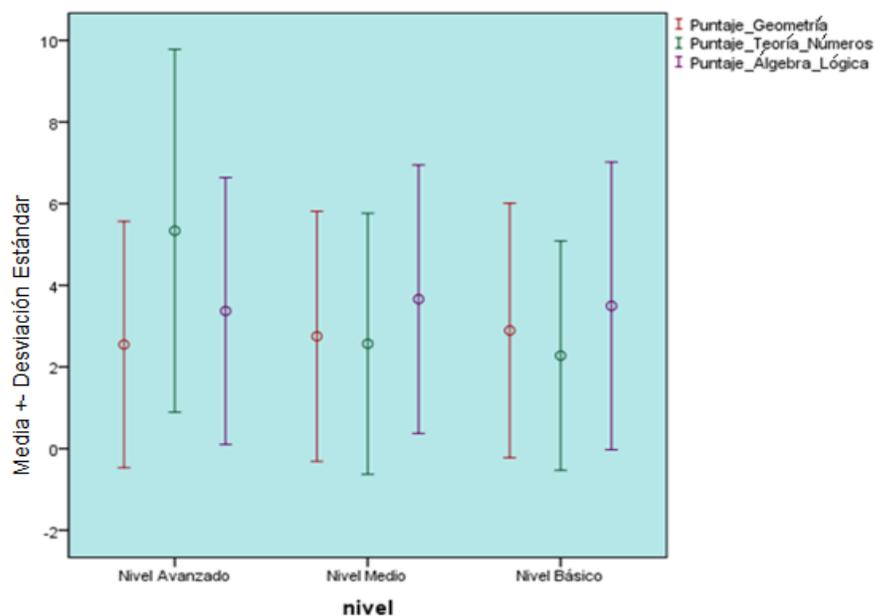
**Gráfica 11. Barras de error para las variables puntaje por área temática (2011)**



El gráfico de barras de error para 2011 determina un comportamiento similar para el nivel avanzado para las tres áreas temáticas, en el nivel medio se puede destacar el rendimiento en Teoría de Números y en el nivel básico Álgebra y lógica como ligeramente superiores. Sin embargo debe notarse que el valor de la media oscila entre 3 y 7 considerando una escala entre 0 y 16, la explicación a estos valores promedios tan bajos esta en que hay en todas las áreas hay más del 50% de los participantes con puntajes entre 0 y 4, esto significa que no hay más de una respuesta correcta en cada área temática. La variabilidad en estos puntajes es considerable, el rango de variación es de 16 unidades pero los puntajes perfectos (todas las preguntas correctas) se alcanzan en muy baja proporción.

En 2012, nivel avanzado se observa un leve incremento del puntaje promedio en Teoría de Números pero también mayor variabilidad, los otros casos podriamos decir que exhiben el mismo comportamiento que en 2011.

**Gráfica 12. Barras de error para las variables puntaje por área temática (2012)**



Complementario a los gráficos anteriores podemos observar las Tablas 19 y 20 (Anexo E) donde hemos registrado el porcentaje de respuesta correcta para cada

pregunta y a su vez las preguntas aparecen organizadas por área temática, se incluye la fase clasificatoria y selectiva. No es fácil identificar un área en la cual haya un alto número de aciertos que nos indiquen un mejor desempeño en ella; en 2011 sólo el nivel básico en su fase clasificatoria muestra altos porcentajes de acierto en Álgebra y Lógica, en 2012 en la fase selectiva se puede destacar el rendimiento en Geometría del nivel avanzado y básico.

### 3.2.5 Rendimiento en la prueba vs asistencia a las capacitaciones

El grupo a cargo de las olimpiadas ha definido una programación a lo largo de un semestre académico. Posterior a la inscripción los colegios son convocados a participar de tres sesiones de capacitación cuyo objetivo es revisar ejemplos de preguntas y conocimientos relacionados con las áreas temáticas incluidas en la competencia previo a la primera fase de la competencia, es decir la fase clasificatoria. Se recibe hasta un máximo de 10 estudiantes por institución educativa. La Tabla 6 muestra el grado de participación observado en estas actividades. Se observa que existe un bajo nivel de participación de éstas actividades que pudiera estar relacionado con la imposibilidad de los estudiantes de provincia para desplazarse a los sitios donde se realizan estas actividades (Málaga, El Socorro, Barrancabermeja, Barbosa y Bucaramanga) y para el caso de Bucaramanga el costo económico del desplazamiento local, otro factor puede ser la falta de conciencia de los profesores coordinadores de cada institución del aporte que pueden dar este tipo de actividad a la hora de presentar la prueba.

**Tabla 6. Participación a las capacitaciones (2011-2012)**

Año	Nivel	Total participantes	Asistencia Capacitación 1	Asistencia Capacitación 2	Asistencia Capacitación 3
2011	Avanzado	1432	77 5.4%	10 0.7%	71 5%
	Medio	1098	50 1,4%	5 0,1%	51 1,5%
	Básico	923	31 0,9%	2 0,1%	5 0,1%

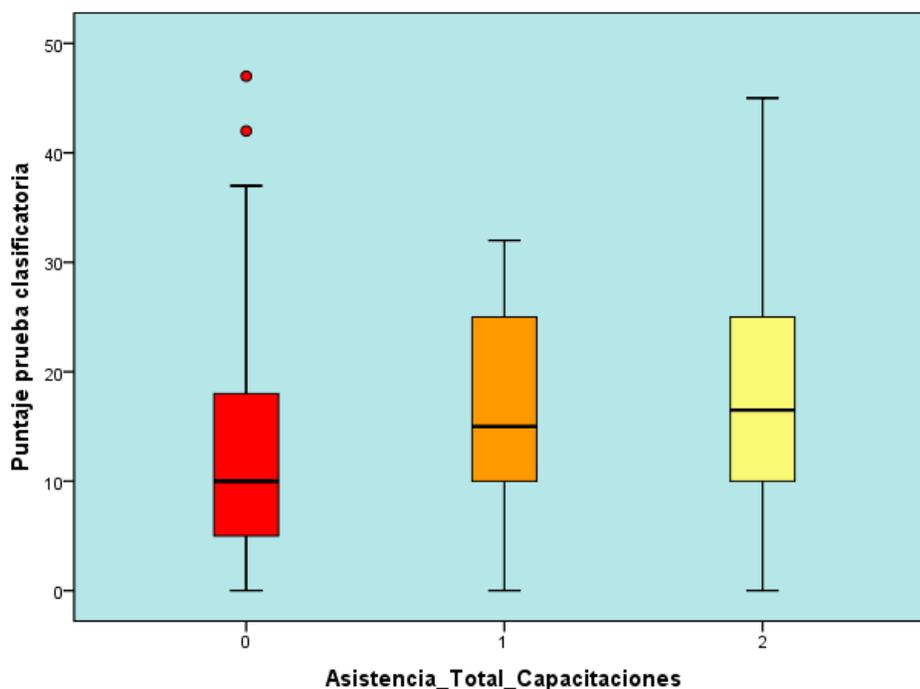
2012	Avanzado	1850	46 1,0%	59 1,3%	25 0,6%
	Medio	1357	10 0,2%	19 0,4%	18 0,4%
	Básico	1270	31 0,7%	25 0,6%	8 0,2%

Los gráficos del 27 al 32 permiten comparar el rendimiento en la prueba clasificatoria según el número de capacitaciones en las que participó el estudiante (ver Anexo F). En el nivel avanzado y básico de 2011 se observa una tendencia creciente aunque con incrementos pequeños, esto junto a la asimetría a derecha de la distribución de puntajes del nivel avanzado (ver grafica 13. caja en amarillo que corresponde a quienes asistieron a dos capacitaciones), son evidencia a favor del efecto de las capacitaciones ver tabla 7 en el rendimiento, en cuanto a la variabilidad el nivel básico luce más homogéneo. Para los datos del 2012 el nivel avanzado se vuelve más asimétrico a derecha de la distribución de los puntajes de los asistentes a dos capacitaciones, en el nivel medio aumenta la mediana y hay una considerable reducción de desviación estándar. Para resumir lo observado podríamos decir que si hay evidencia de una relación positiva, aunque parece ser débil, entre las variables puntaje en la prueba clasificatoria y número total de capacitaciones a las que asistió.

**Tabla 7. Coeficientes de Asimetría para la variable puntaje según el número de capacitaciones a las que asistió**

Año	Nivel	Coeficiente de Asimetría según el número de capacitaciones a las que asistió		
		0	1	2
2011	Avanzado	0,58	0,033	0,57
	Medio	0,47	0,41	0,21
	Básico	0,13	-0,17	-1,12
2012	Avanzado	0,80	0,59	0,19
	Medio	0,71	0,15	0,22
	Básico	0,83	0,55	2

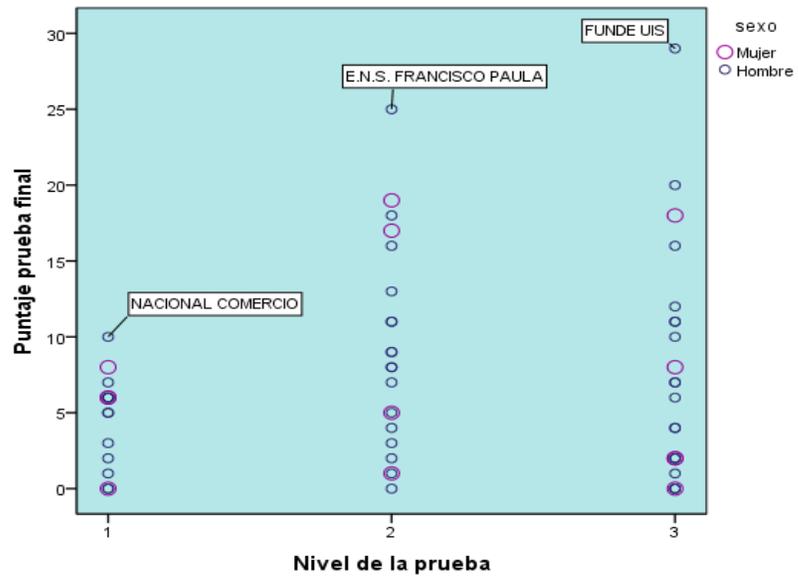
**Gráfica 13. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Avanzado 2011)**



### 3.3 RENDIMIENTO DE LA FASE FINAL

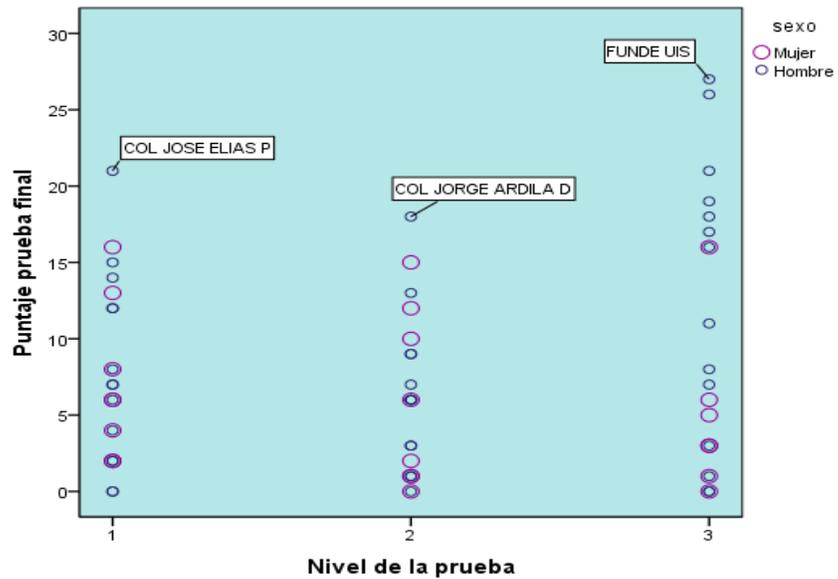
Los gráficos 14 y 15 son gráficos de dispersión por puntos contruidos sólo con los puntajes de la fase final. En los dos casos se observa una leve tendencia creciente pero a medida que aumenta el nivel de la prueba aumenta la dispersión en los valores, la escala en el eje horizontal corresponde a 1: Nivel Básico, 2: Nivel Medio y 3: Nivel Avanzado. Al parecer la prueba de la fase final resulta muy exigente para los participantes y por esto vemos tanta concentración de valores por debajo de 10 (no alcanzan a tener dos preguntas respondidas correctamente de las seis que contiene la prueba), de ahí en adelante aumenta la dispersión en la zona de los puntajes altos. Ningún estudiante alcanzó el puntaje perfecto. Las etiquetas que acompañan al último valor corresponden al nombre del colegio del estudiante ganador de la competencia en cada nivel. El color de las marcas determina el género del finalista, más adelante la variable género será considerada en relación con el desempeño en la prueba.

**Gráfica 14. Puntaje en la prueba final discriminando por género y por nivel (2011)**



1= Nivel básico, 2= Nivel medio, 3= Nivel avanzado

**Gráfica 15. Puntaje en la prueba final discriminando por género y por nivel (2012)**



1= Nivel básico, 2= Nivel medio, 3= Nivel avanzado

En las Tabla 8 se resume el cuadro final de medallería, es decir, los estudiantes que ocuparon los 3 primeros puestos en cada Nivel (2011-2012), de los resultados a lo largo de la competencia se evidencia que la mayoría de estos estudiantes mantuvieron un nivel superior o alto de desempeño en la prueba exceptuando tal vez la prueba final que como se ya mencionó ésta arroja resultados más bajos.

**Tabla 8. Clasificación final**

2011					
Nivel	Nombre	Colegio	Puntaje_ Clasificatoria	Puntaje Prueba_ Selectiva	Puntaje_ Prueba_ Final
Avanzado	Carlos Mario Sánchez	FUNDEUIS	47	43	29
	Leandro Remolina	VICENTE AZUERO	45	29	20
	Maria Camila Remolina	FUNDEUIS	42	23	18
Medio	Juan Caballero	E.N.S. FRANCISCO PAULA	31	29	25
	Andrea Chio Maldonado	COL QUINTA PUENTE	50	48	19
	Edward Gamboa	COL SEÑORA FATIMA	40	30	18
Básico	Laura Natalia Martinez R.	NACIONAL COMERCIO	40	33	10
	Camila Martinez Serrano	FUNDE UIS	35	23	8
	Geyson Sneyder Uribe P	COL DEP LUIS CARLOS G	40	29	7

2012					
Nivel	Nombre	Colegio	Puntaje_ Clasificatoria	Puntaje Prueba_ Selectiva	Puntaje_ Pr ueba_ Final
Avanzado	Carlos Vargas Sánchez	FUNDEUIS	50	16	27
	Daniel Barrera	NACIONAL COMERCIO	44	24	26
	Sebastián Sabogal	JORGE ARDILA DUARTE	50	16	21
Medio	Rafael Pinto	JORGE ARDILA DUARTE	25	20	18

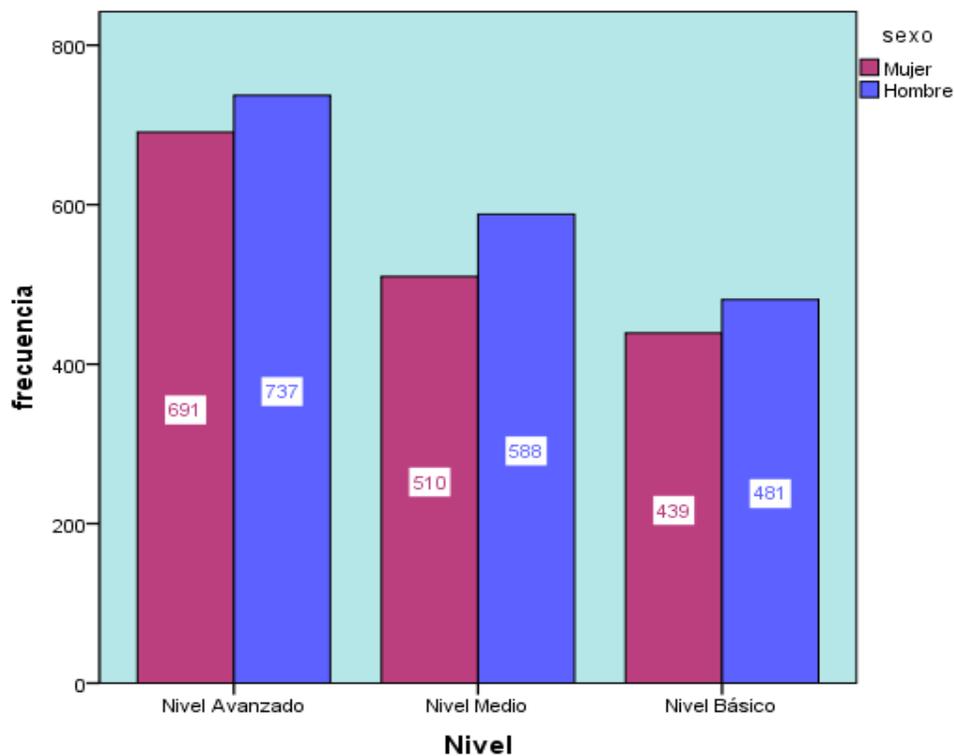
2012					
Nivel	Nombre	Colegio	Puntaje_ Clasificatoria	Puntaje Prueba_ Selectiva	Puntaje_Pr ueba_ Final
	Laura Martínez	NACIONAL COMERCIO	30	21	15
	Juan Almas	FUNDEUIS	24	20	13
Básico	Rueda Andrew	JOSE ELIAS PUYANA	40	18	21
	Meneses Juliana	GIMNASIO SUPERIOR	25	21	16
	Sergio Acelas	JOSE ELIAS PUYANA	30	31	15

## 4. COMPARACIONES DISCRIMINANDO POR GÉNERO

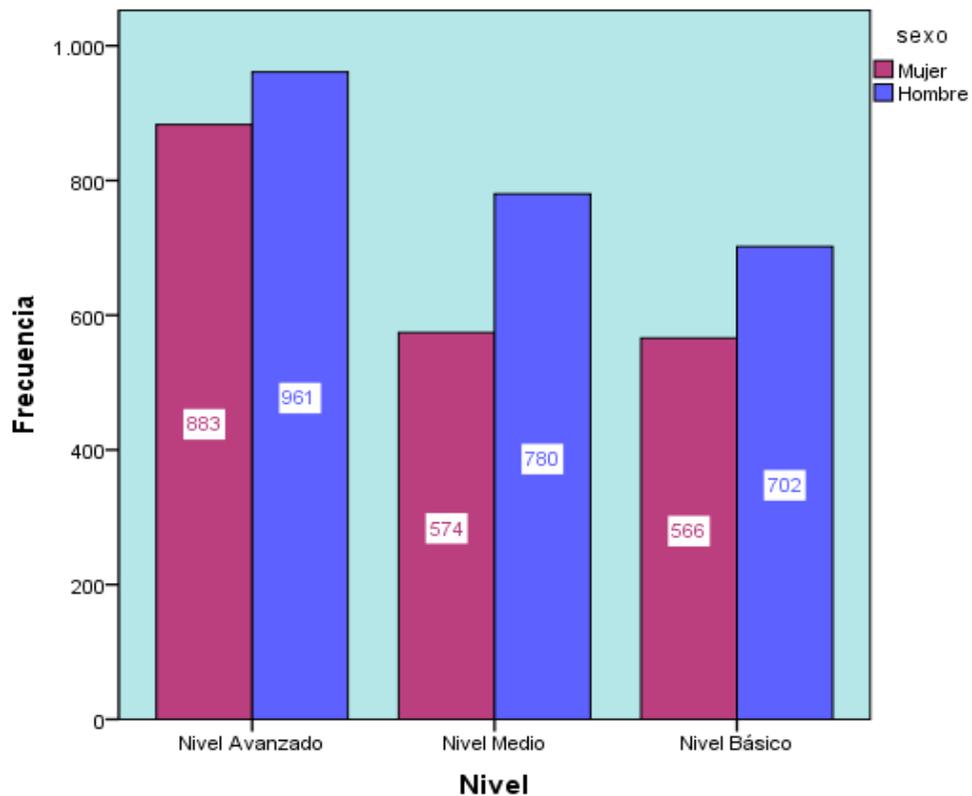
### 4.1 NIVEL DE PARTICIPACIÓN EN LA PRUEBA

En las gráficas 16 y 17 sobre la participación por género se muestra una mayor presencia de hombres en cada nivel de la competencia, sin embargo las diferencias porcentuales son bajas, por ejemplo en 2011 la participación femenina fue del 48%, 46.4% y 47.5% para los niveles avanzado, medio y básico respectivamente, éstos mismos valores en 2012 fueron 47.8%, 42.2% y 44.5%. De la Gráfica 17 es claro el aumento en el número participantes mujeres del 2011 al 2012 y que es en el nivel avanzado donde hay mayor participación femenina. Los resultados para la prueba de igualdad de proporciones aparecen en la tabla 1 y fueron comentados en la sección 3.1.

**Gráfica 16. Participación por Nivel Vs. Género (2011)**

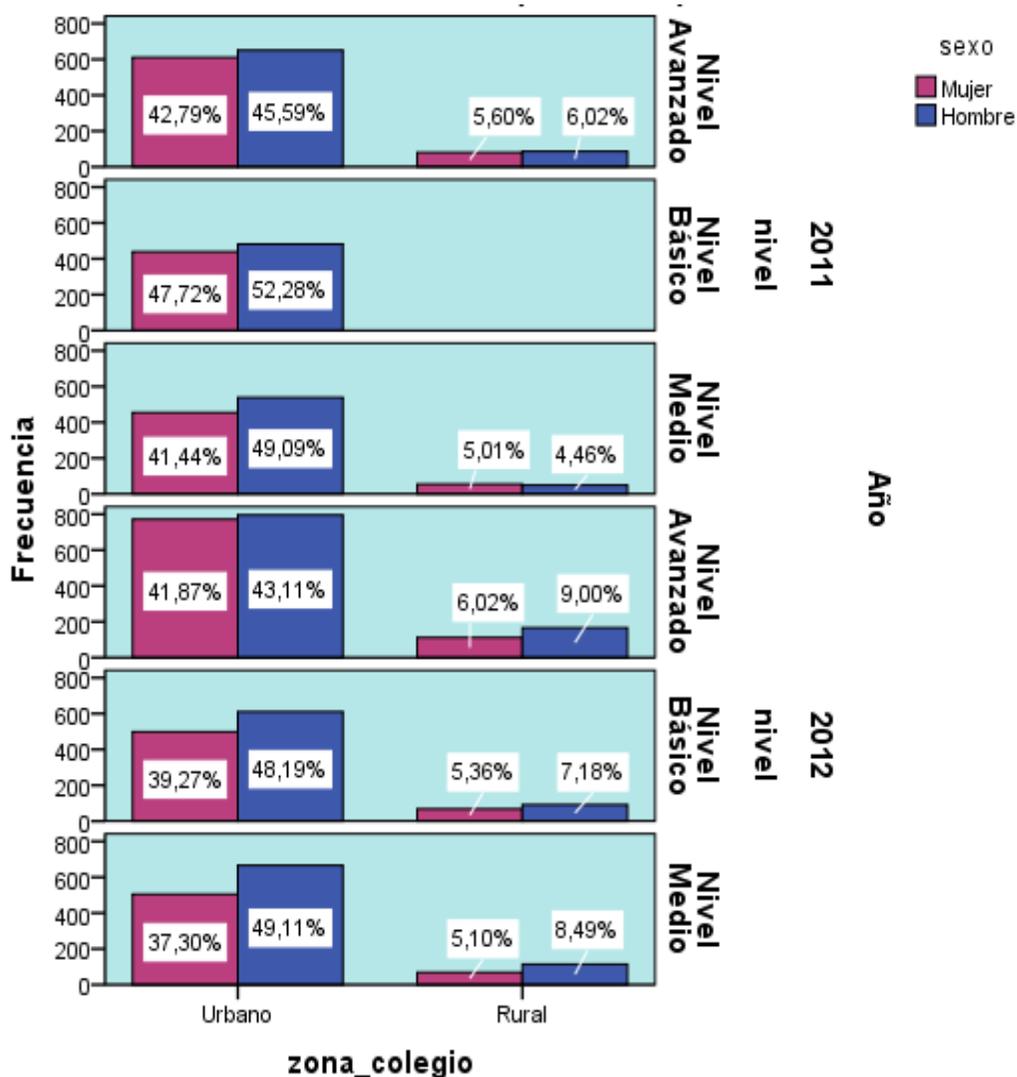


**Gráfica 17. Participación por Nivel Vs. Género (2012)**



La gráfica 18 muestra la distribución de porcentajes de participación discriminando por género y zona del colegio. Se observa que hay una baja participación tanto de hombres como de mujeres en los colegios ubicados en las zonas rurales de Santander, de otro lado se evidencia la diferencia en la participación entre los hombres y las mujeres tanto en los colegios rurales y urbanos pero es una diferencia poco considerable ya que los porcentajes de participación mostrados en la gráfica son muy cercanos. En la tabla 9 se registra los valores que permiten comparar proporciones observadas en 2011-2012 discriminando por género y zona; de las comparaciones realizadas se observan diferencias estadísticamente significativas en la proporción en el nivel de participación de mujeres del nivel medio, sector urbano. La participación de hombres del sector rural en cada uno de los niveles que contiene la prueba presenta diferencias estadísticamente significativas.

**Gráfica 18. Participación discriminando por género y zona del colegio (2011-2012)**



**Tabla 9. Prueba de hipótesis para diferencia de proporciones**

Nivel	Zona del colegio	Proporciones muestrales		z	Valor p
		$\hat{P}_{Mujeres\_2011}$	$\hat{P}_{Mujeres\_2012}$		
Avanzado	Urbano	0,42787115	0,4186551	0,529259915	0,60
	Rural	0,05602241	0,06019523	-0,50491491	0,61
Medio	Urbano	0,4143898	0,37296898	2,089629964	0,03*
	Rural	0,05009107	0,05096012	-0,09765847	0,92

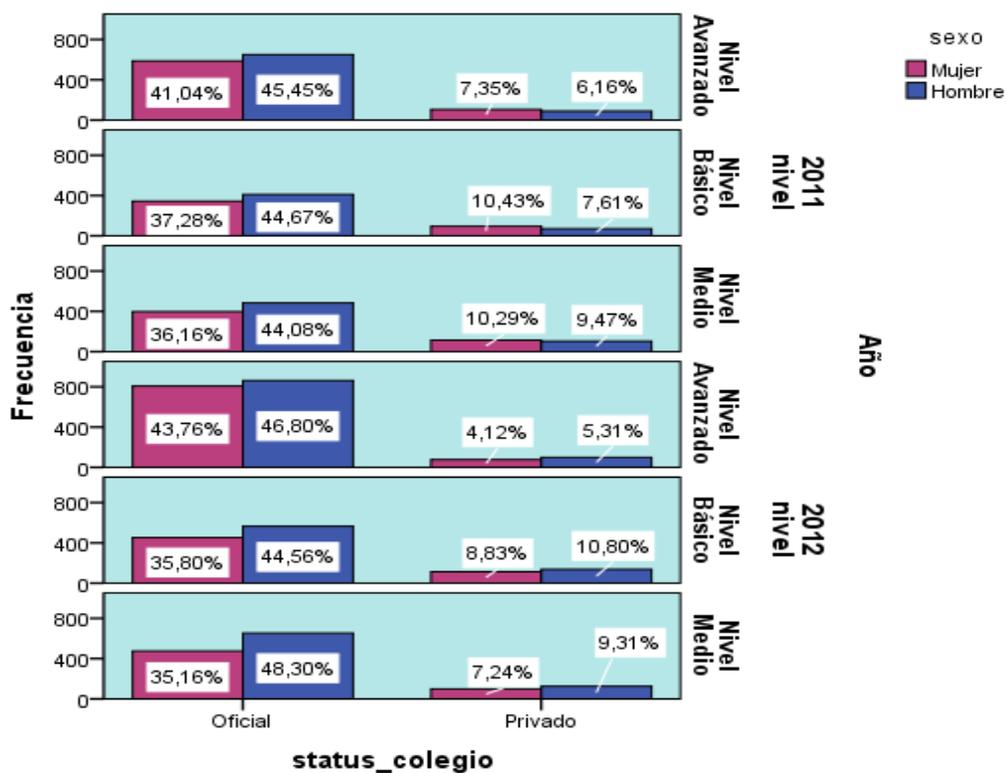
Básico	Urbano	0,47717391	0,39274448	3,939785927	0
	Rural	0	0,05362776	-7,13582961	0

Nivel	Zona del colegio	Proporciones muestrales		z	Valor p
		$\hat{P}_{Hombres\_2011}$	$\hat{P}_{Hombres\_2012}$		
Avanzado	Urbano	0,45588235	0,43112798	1,414062352	0,15
	Rural	0,06022409	0,09002169	-3,17050752	0*
Medio	Urbano	0,49089253	0,49113737	-0,01205953	0,99
	Rural	0,04462659	0,08493353	-3,97283558	0*
Básico	Urbano	0,52282609	0,4818612	1,891787612	0,06
	Rural	0	0,07176656	-8,3000255	0*

\*Prueba significativa al 5%.

De la gráfica 19 se puede resaltar que en los colegios oficiales es levemente superior el porcentaje de hombres en todos los niveles de la prueba durante 2011 y 2012, la gráfica nuevamente deja evidencia que el nivel de mayor participación es el avanzado lo que quiere decir que los estudiantes que más participan de las olimpiadas son estudiantes cursando décimo o undécimo grado. Un hecho importante de resaltar es que sólo en los colegios privados en 2011 hay más participación de mujeres que de hombres en todos los niveles. En la tabla 10 se registra los valores que permiten comparar proporciones observadas en 2011-2012 discriminando por género y status del colegio; de las comparaciones realizadas se observan diferencias estadísticamente significativas en la proporción en el nivel de participación de mujeres del nivel avanzado y medio, sector privado. La participación de hombres del sector oficial y privado en los niveles medio y básico presenta diferencias estadísticamente significativas.

**Gráfica 19. Participación discriminando por género y status del colegio (2011-2012)**



**Tabla 10. Prueba de hipótesis para diferencia de proporciones**

Nivel	Status del colegio	Proporciones muestrales		Z	Valor p
		$\hat{P}_{Mujeres\_2011}$	$\hat{P}_{Mujeres\_2012}$		
Avanzado	Oficial	0,41036415	0,43763557	-1,56466065	0,11
	Privado	0,07352941	0,04121475	4,01016301	0*
Medio	Oficial	0,36156648	0,35155096	0,51504626	0,61
	Privado	0,10291439	0,07237814	2,68116965	0,022*
Básico	Oficial	0,37282609	0,35804416	0,70927623	0,48
	Privado	0,10434783	0,08832808	1,26115562	0,20

Nivel	Status del colegio	Proporciones muestrales		Z	Valor p
		$\hat{P}_{Hombres\_2011}$	$\hat{P}_{Hombres\_2012}$		
Avanzado	Oficial	0,45448179	0,46800434	-0,76944474	0,44

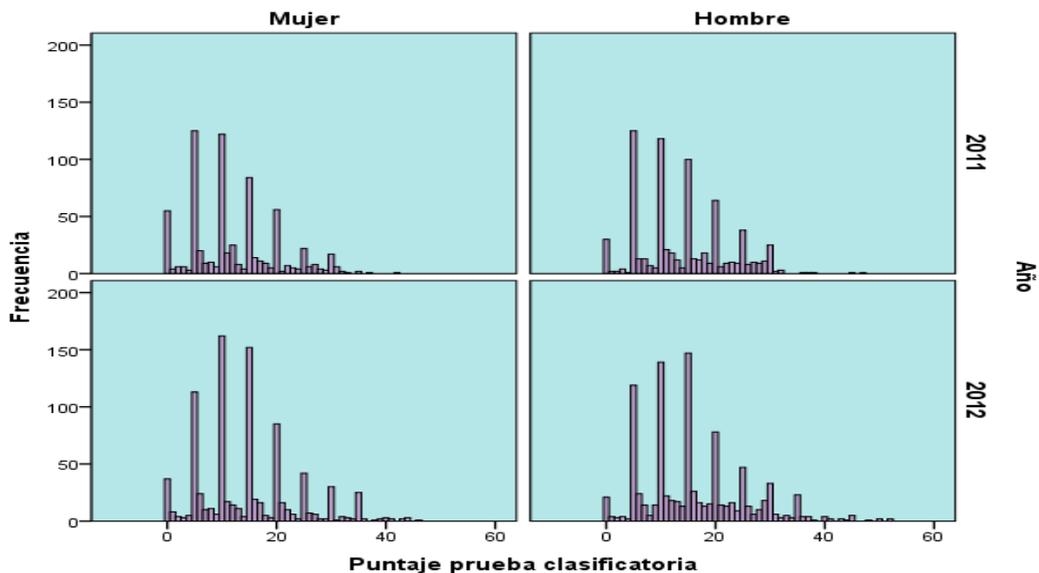
	Privado	0,06162465	0,05314534	1,03886108	0,30
Medio	Oficial	0,44080146	0,48301329	-2,08418455	0,03*
	Privado	0,09471767	0,09305761	0,14020353	0,88
Básico	Oficial	0,44673913	0,4455836	0,05367645	0,96
	Privado	0,07608696	0,10804416	-2,52126568	0,01*

\*Prueba significativa al 5%.

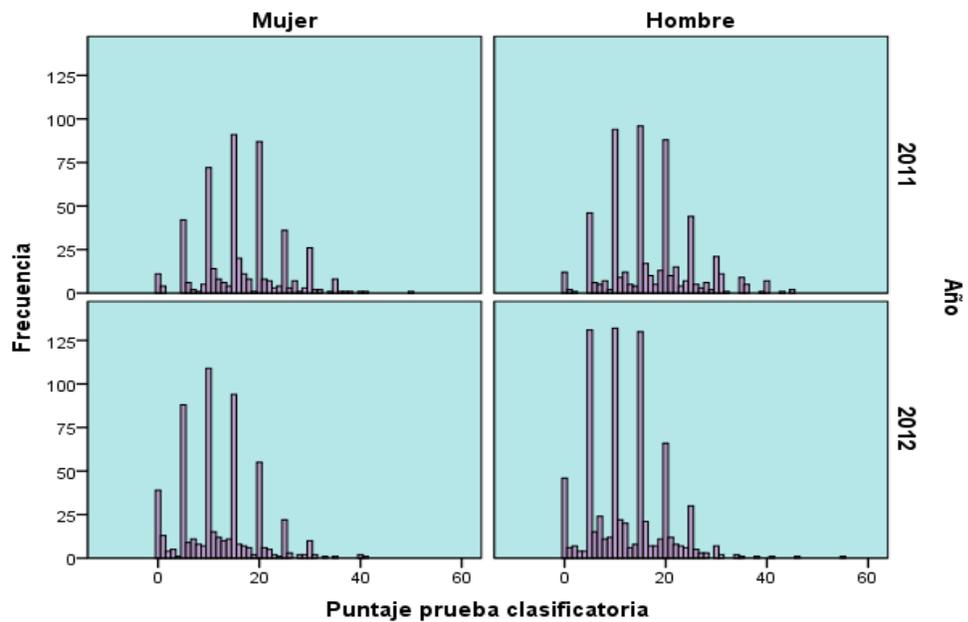
## 4.2 RENDIMIENTO EN LA FASE CLASIFICATORIA

Para la fase clasificatoria de 2011 se inscribieron 3453 participantes de los cuales el 47.5% fueron mujeres, en 2012 se tienen 4477 participantes y de estos 45.1% son mujeres. De las gráficas 20 al 22 se observa que la distribución de los puntajes en la prueba clasificatoria es asimétrica a derecha, por esta razón las gráficas 23 y 24 se construyeron para comparar el rendimiento por género usando como medida de tendencia central la mediana.

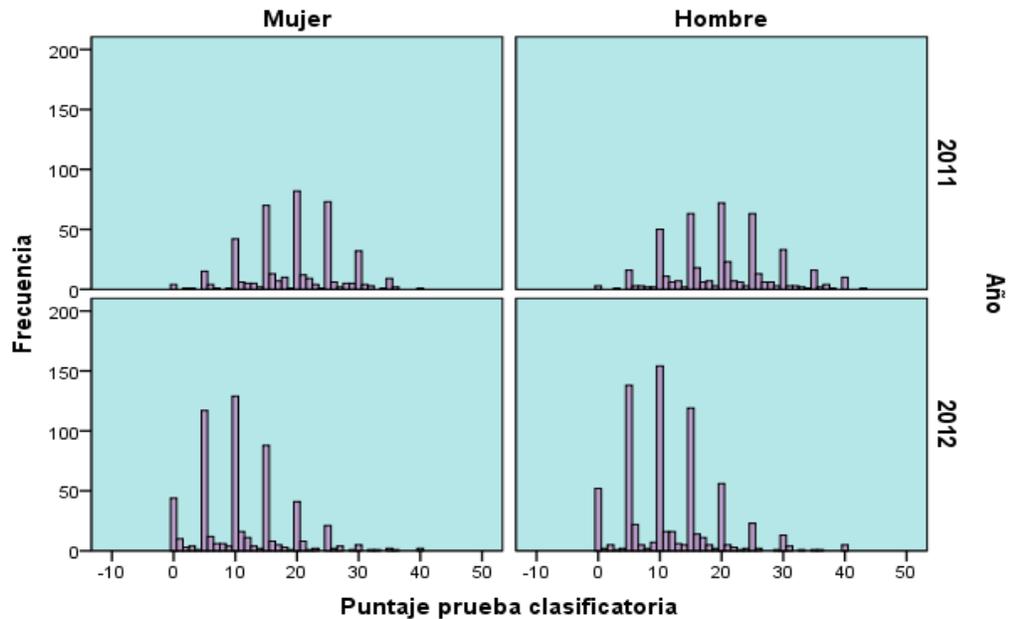
**Gráfica 20. Panel con la distribución de los puntajes en la prueba clasificatoria discriminando por género nivel avanzado**



**Gráfica 21. Panel con la distribución de los puntajes en la prueba clasificatoria discriminando por género nivel medio**



**Gráfica 22. Panel con la distribución de los puntajes en la prueba clasificatoria discriminando por género nivel básico**



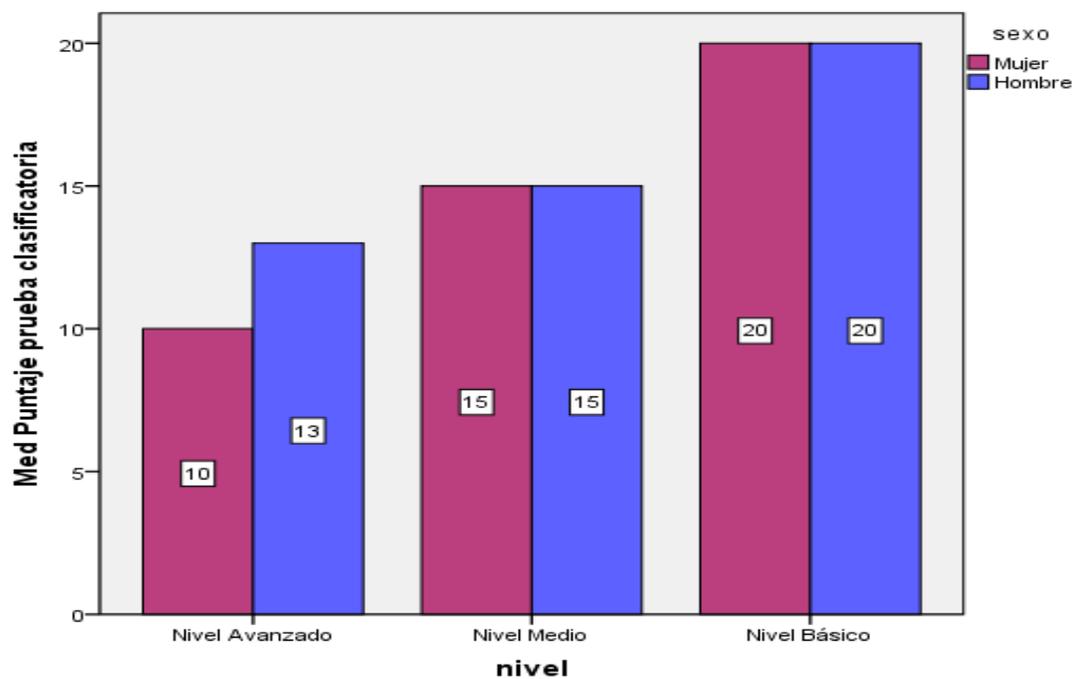
En 2011, solo se ven diferencias en el nivel avanzado aunque la diferencia es de sólo tres puntos; para 2012 y como hecho muy interesante las medianas en cada subgrupo coinciden, esto indica que en cada nivel, el 50% de los puntajes se acumula a izquierda del mismo valor (ver gráfica 23 y 24), volviendo a los histogramas observamos que esa primera parte de la distribución es igual para

hombres y mujeres y es en la segunda mitad de la distribución donde si podremos hablar de diferencias en cuanto a variabilidad de los puntajes ver tabla 11.

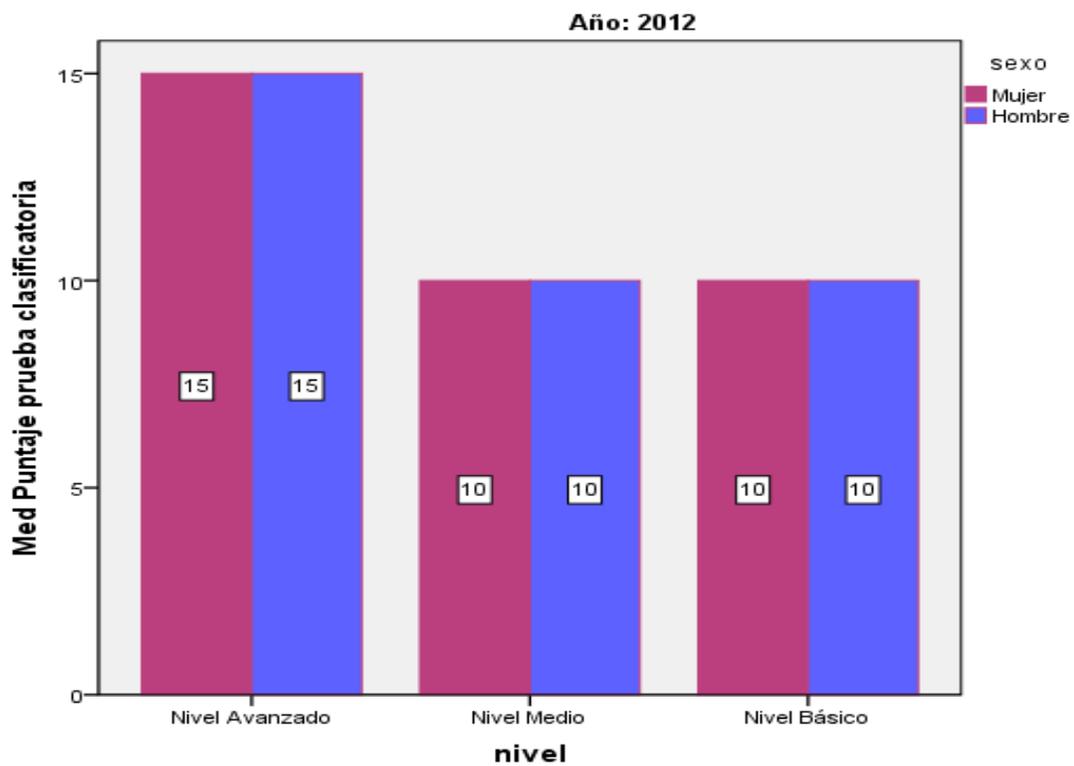
**Tabla 11. Medias de dispersión para la variable puntaje en la prueba clasificatorias**

Año	Nivel	Género	Desviación estándar	Rango de variación	Rango de variación 25% superior Máx – Q <sub>3</sub>
2011	Avanzado	Mujer	8,07	42	26
		Hombre	8,25	47	27
	Medio	Mujer	8,043	50	30
		Hombre	8,457	45	24
	Básico	Mujer	7,415	40	15
		Hombre	8,319	43	18
2012	Avanzado	Mujer	9,023	46	26
		Hombre	9,464	52	31
	Medio	Mujer	7,561	41	26
		Hombre	7,496	55	39
	Básico	Mujer	7,329	40	25
		Hombre	4,461	40	25

Gráfica 23. Mediana de los puntajes en la prueba clasificatoria 2011



Gráfica 24. Mediana de los puntajes en la prueba clasificatoria en 2012



La Tabla 21 del Anexo G permite comparar las distribuciones desde otros parámetros de la distribución como son la media, desviación estándar, los percentiles y el puntaje máximo. La comparación del rendimiento promedio y descriptores como la desviación estándar, los percentiles 25, 50 y 75 nos permiten concluir que no hay diferencias por género, sin embargo hay que ir más a la derecha de la distribución a la zona de los altos puntajes, sólo allí existen diferencias importantes por género en el nivel avanzado y medio, en el nivel básico los valores extremos son muy similares.

En la tabla 12 se presentan los estadísticos de resumen de las pruebas de hipótesis que nos permiten probar si hay diferencia en el rendimiento promedio al discriminar por género; sólo en el nivel avanzado se observa diferencias significativas en el rendimiento promedio entre hombres y mujeres.

**Tabla 12. Valores Prueba de hipótesis para probar diferencias de medias en el rendimiento promedio**

Año	Nivel	$\bar{x}_M$	$\bar{x}_H$	Valor P prueba F	T	Grados de libertad	Valor p	IC para la diferencia de medias
2011	Avanzado	12,13	14,03	<sup>a</sup> 0,292	-4,375	1426	0*	(-2,741,-1,044)
	Medio	16,52	16,92	<sup>a</sup> 0,185	-0,793	1096	0,428	(-1,378,0,585)
	Básico	19,37	20,05	0,039	-1,313	917,49	0,19	(-1,699,0,337)
2012	Avanzado	14,55	16,05	<sup>a</sup> 0,09	-3,466	1842	0,001*	(-2,342,-0,649)
	Medio	12,06	12,23	<sup>a</sup> 0,943	-0,417	1352	0,677	(-0,984,0,639)
	Básico	11,08	11,65	<sup>a</sup> 0,616	-1,357	1266	0,175	(-1,388,0,253)

<sup>a</sup>Prueba F de homogeneidad de varianzas sugiere asumir varianzas iguales.

\*Prueba significativa al 5%.

Para estudiar el comportamiento por área temática se construyeron las Tablas 19 y 20 del Anexo E, aquí nuevamente como ya se mencionó en el capítulo anterior en los resultado globales, es difícil identificar un área con un desempeño destacado, sólo Álgebra y Lógica en el caso del nivel básico en 2011.

A continuación se calculan los odds ratio y riesgos relativos de no acierto para las preguntas de la prueba en la fase clasificatoria. Estos valores son medidas

alternativas a las probabilidades ampliamente utilizados en epidemiología, se utilizan para medir la fuerza con la que un determinado evento está asociado a un factor, desde el punto de vista estadístico lo que se evalúa es la diferencia observada entre el número de veces que ocurre el evento relativo al número de veces que no ocurre. Como factor de exposición se considera el género y el evento de interés no acertar una pregunta. Como ya se dijo en el marco teórico, también es común reportar el riesgo relativo que es el cociente entre el riesgo de presentar el evento entre quienes tiene el factor y el riesgo de presentar el evento y no tienen el factor de riesgo. Los siguientes resultados (Tabla 13) corresponden al nivel avanzado 2011. También es común reportar el riesgo relativo que es el cociente entre el riesgo de presentar el evento entre quienes tiene el factor y el riesgo de presentar el evento y no tienen el factor de riesgo (Woodward, 1999).

**Tabla 13. Cálculos usando los resultados de la fase clasificatoria del Nivel Avanzado (2011)**

Pregunta	Odds Ratio de no acierto			Riesgo relativo de no acierto
	Mujeres	Hombres	Odds Ratio	
1	459/232 = 1,97	439/298 = 1,47	1,97 / 1,47 = 1,34	(459/691) / (439/737) = 1,11
2	4,85	3,30	1,46	1,07
3	6,51	6,08	1,06	1
4	6,85	5,58	1,22	1,02
5	1,89	1,53	1,23	1,08
6	5,58	5,82	0,95	0,99
7	4,11	4,26	0,96	0,99
8	7,12	7,01	1,01	1
9	5,39	4,66	1,15	1,02
10	5,33	6,37	0,83	0,97
11	15,06	12,16	1,23	1,01
12	2,59	1,80	1,44	1,12

Como se indicó la interpretación se hace considerando si el valor de odds ratio o riesgo es 1, superior a 1 o inferior a 1. Según los valores calculados, sólo para las preguntas 1 y 12 el riesgo de no acertar es un poco mayor para las mujeres que para los hombres, en los otros casos el riesgo de no acertar la pregunta es el mismo para hombres y mujeres. La Tabla 22 (Anexo H) presentan el cálculo de odds ratio y de riesgos relativos calculados para la fase clasificatoria del nivel avanzado de 2012 en la que tampoco se observan diferencias por género.

#### 4.3 RENDIMIENTO EN LA FASE FINAL

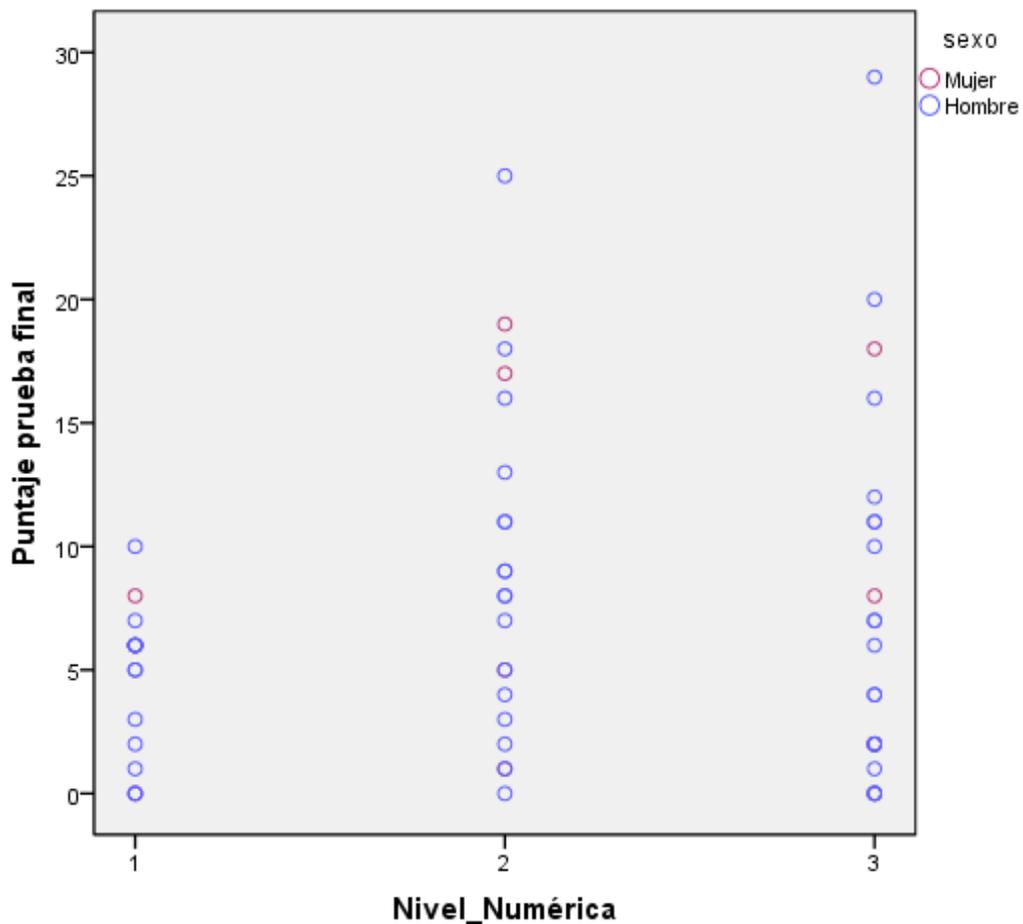
El grupo de finalistas se compone de la siguiente manera:

**Tabla 14. Colegios finalistas**

<b>Nivel</b>	<b>2011</b>	<b>2012</b>
Avanzado	5 Mujeres y 19 Hombres 21 de colegio urbano, 3 de colegio rural 14 de colegio oficial, 10 de colegio privado	7 Mujeres y 14 Hombres 21 de colegio urbano 16 de colegio oficial, 5 de colegio privado
Medio	4 Mujeres y 17 Hombres 19 de colegio urbano, 2 de colegio rural 13 de colegio oficial, 8 de colegio privado	8 Mujeres y 14 Hombres 20 de colegio urbano, 2 de colegio rural 17 de colegio oficial, 5 de colegio privado
Básico	4 Mujeres y 14 Hombres 18 de colegio urbano 10 de colegio oficial, 8 de colegio privado	8 Mujeres y 14 Hombres 20 de colegio urbano, 2 de colegio rural 17 de colegio oficial, 5 de colegio privado
<b>Total Mujeres finalistas</b>	<b>13 entre 63 finalistas</b>	<b>23 entre 65 finalista</b>

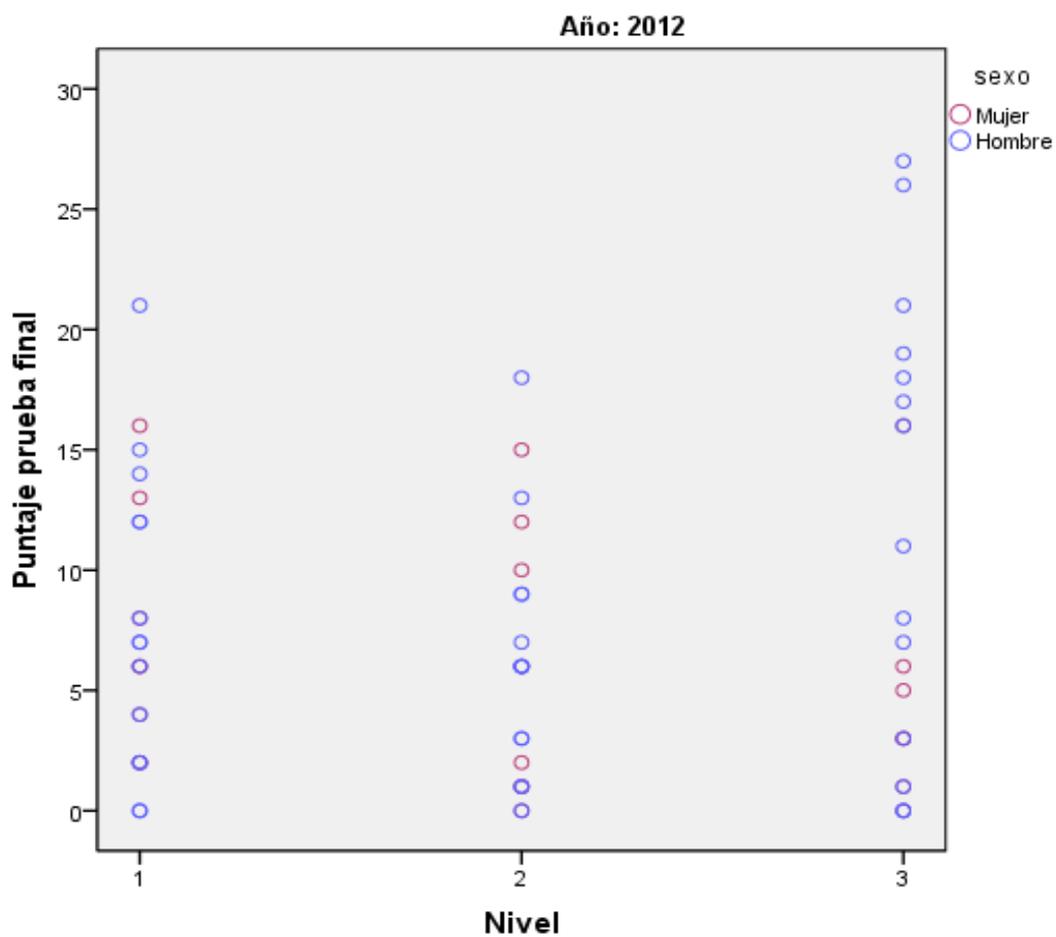
Los gráficos 25 y 26 a continuación dejan en evidencia el rotundo dominio del género masculino en la fase final en cuanto a número de participantes por nivel. En 2011 se destaca que en cada nivel en los tres primeros lugares hay al menos una mujer. En 2012 aumenta el número de mujeres finalistas, pero en el nivel avanzado no queda ninguna entre los tres primeros lugares, en el nivel básico y avanzado ocupan el segundo lugar. Es bueno recalcar que en todos los casos se observa un ganador con una ventaja importante sobre el segundo y tercero.

**Gráfica 25. Gráfico de dispersión de puntos para la variable puntaje prueba final en 2011**



1= Nivel Básico, 2= Nivel Medio, 3= Nivel Avanzado

Gráfica 26. Gráfico de dispersión de puntos para la variable puntaje prueba final en 2012



1= Nivel Básico, 2= Nivel Medio, 3= Nivel Avanzado

## 5. CONCLUSIONES

Sobre el nivel de participación en la competencia hay que destacar la tendencia creciente en el número de participantes desde el inicio de la Olimpiadas Regionales en Santander; también llama la atención que es el nivel avanzado (décimo y undécimo grado) el que presenta los mayores porcentajes de participación en la competencia. Como hecho importante al observar la representación por provincias se destaca la baja participación de la provincia de Mares, sólo los municipios de Barrancabermeja y Puerto Wilches han hecho presencia en la prueba, sería pertinente estudiar los factores que pueden estar motivando esta escasa participación.

Respecto a la composición de la muestra de individuos que han participado en las competencias de 2011 y 2012 se observan hechos esperados como son la mayor participación sector urbano y de colegios del sector público. Llama la atención que en la fase clasificatoria, inicio de la competencia, los tres niveles exhiben casi igual participación del género femenino que el masculino. Del listado de participación por colegios hay varios que se destacan no sólo por participar con un número importante de estudiantes sino por ubicar varios de ellos entre los finalistas, es conveniente destacar: Colegio Jorge Ardila, Colegio José Elias Puyana, Colegio Cooperativo CONFENALCO, Escuela Normal Superior Francisco de Paula (Málaga), FUNDEUIS, Colegio Nacional de Comercio, Colegio San José de la Salle, Gimnasio Cantillana

Del rendimiento en la prueba hay que destacar los bajos puntajes, son pocos los estudiantes que alcanzan un nivel superior o alto de desempeño. Las distribuciones de los puntajes son asimétricas, comportamiento usual en este tipo de competencias tras haber hecho una selección previa y la presencia de estudiantes talentosos que se distancian del resto. Al comparar el rendimiento por

niveles se tiene que es el nivel avanzado el que logra mejores puntajes aunque en la fase final son grupos muy heterogéneos (alta variabilidad), el avanzado es también el único nivel del que se muestran algunas evidencias de avance en el desempeño en 2012 comparado con 2011.

La comparación por rendimiento por área temática no arrojo resultados concluyentes, no hay altos porcentajes de respuesta para una misma área ni un comportamiento sólido en un nivel que se mantenga en el tiempo.

El estudio de la asociación entre asistencia a las capacitaciones y rendimiento en la fase clasificatoria no muestra una relación positiva en todos los casos y cuando se observa es muy débil (nivel avanzado y nivel básico en 2011, nivel avanzado y nivel medio en 2012). Aquí hay que aclarar que no todos los participantes tienen la oportunidad de participar de estas actividades pues se ofertan sólo diez cupos por institución educativa.

Sobre la conformación del grupo de finalistas se destaca que la mayoría de integrantes provienen de colegios urbanos de Bucaramanga, hay una mayor participación de colegios del sector oficial de lo cual se puede inferir que hay un trabajo importante de los profesores de Matemáticas de estos colegios y que los colegios privados podrían estar haciendo una selección interna antes de inscribirse en la prueba.

El estudio de diferencias por género mostró que en la mayor parte de la distribución de los puntajes los parámetros coinciden y que las diferencias si se dan pero sólo en la zona de alto rendimiento. Es de destacar que aunque el porcentaje de mujeres en la ronda final es muy inferior al de hombres en todos los niveles, ha habido mujeres ocupando los lugares de honor. En general se observa bajo rendimiento en la prueba final, en el caso de las mujeres el máximo puntaje obtenido no supera los 19 puntos en una escala de 0 a 36 lo cual indica que es

poco probable que se hayan respondido todas las preguntas. Finalmente resaltar que son los niveles básicos y medio donde mejor desempeño tienen las finalistas.

## BIBLIOGRAFÍA

**CAÑADAS, G., BATANERO, C. y otros. (2011).** *Estrategias en el estudio de la asociación en tablas de contingencia por estudiantes de psicología. Educación Matemática*, vol. 23, núm. 2, pp. 5-31. Grupo Santillana México.

<sup>3</sup>**Mapa división política de Santander** [Recuperado el 31 julio de 2013 de. [http://4.bp.blogspot.com/-3l\\_Mv7qxdag/T9Tb90zDKXI/AAAAAAAAAOY/-FH\\_sokQTgk/s1600/En+Santander+2.jpg](http://4.bp.blogspot.com/-3l_Mv7qxdag/T9Tb90zDKXI/AAAAAAAAAOY/-FH_sokQTgk/s1600/En+Santander+2.jpg)].

<sup>2</sup>**Medidas de localización** [Recuperado el 20 de agosto de 2013 de <http://biplot.usal.es/DOCTORADO/3CICLO/BIENIO-06-08/SPSS%20Estadistica.pdf>]

**MOORE, David S. (1995).** *Estadística Básica Aplicada*. Antoni Bosch, editor. W.H. Freeman and Company.

**WOODWARD, M. (1999).** *Epidemiology Study desing and data analysis*. Chapman & Hall/CRC

<sup>1</sup>**Universidad Antonio Nariño** [Recuperado el 7 de Mayo de 2013 de <http://olimpia.uan.edu.co/olimpiadas/public/frameset.jsp>]

## ANEXOS

### ANEXO A. Ejemplo de calificación en las preguntas de tipo ensayo

#### NIVEL MEDIO

4. Jacobo pidió 4 pares de medias negras y algunos pares de medias azules. El precio de cada par medias negras es 2 veces el de cada par azules. Hizo el pedido y encontró que el número de pares de los dos colores han sido cambiado. Esto aumentó la cuenta en un 50%. ¿Cuál es la relación entre el número de pares de medias negras y el de azules en el pedido original?

Solución:

Sea  $x$  el número de parejas de medias azules y el precio de un par de medias azules. Entonces  $2y$  es el precio de un par de medias negras  $\rightarrow$  (2 puntos)

$$x \cdot 2y + 4y = \frac{3}{2} (4 \cdot 2y + x \cdot y) \rightarrow (1 \text{ punto})$$

Se divide por  $y$  y se resuelve la ecuación para  $x$ .

$$2x + 4 = \frac{3}{2} (8 + x) \Rightarrow 4x + 8 = 24 + 3x$$
$$x = 16 \rightarrow (1 \text{ punto})$$

$$4 : x = 4 : 16 = 1 : 4$$

Rta: La relación entre el número de pares de medias negras y el de azules en el pedido original es  $1 : 4$   $\rightarrow$  (2 punto)

**ANEXO B. Participación de colegios discriminando por nivel de desempeño  
(2011-2012)**

**Tabla 15. Participación de colegios discriminando por nivel de  
desempeño (2011-2012)**

Frecuencia

Año		nivel			Total	
		Nivel Avanzado	Nivel Medio	Nivel Básico		
2011	colegio	AQUILEO PARRA	37	45	12	94
		CAMACHO CARREÑO	17	8	15	40
		CIUDADELA EDUCATIVA	8	8	33	49
		COL AMERICAS	22	4	6	32
		COL PRESENTACION	69	39	<b>58</b>	166
		COL QUINTA PUENTE	0	1	0	1
		COL BETHLEMITAS	6	13	12	31
		COL CASTILLO	0	24	23	47
		COL CENTRO COMERCIO	<b>79</b>	30	18	127
		COL COOPERATIVO	43	40	33	116
		COL CUSTODIO GARCIA	23	36	44	103
		COL DEP LUIS CARLOS G	15	17	20	52
		COL DIVINO AMOR	0	39	4	43
		COL DIVINO NIÑO	5	6	6	17
		COL FRANCISCO PAULA	28	23	31	82
		COL INMACULADA	43	42	45	130
		COL ISIDRO CABALLERO	19	23	9	51
		COL JESUS PRAGA	12	16	0	28
		COL JORGE ARDILA D	16	18	12	46
		COL JOSE ELIAS P	21	11	0	32
		COL JOSE FERRO	7	6	4	17
		COL LORENZO SALAZAR	39	6	14	59
		COL LUCAS CABALLERO	15	20	0	35
		COL LUIS GONZAGA	3	12	16	31
		COL MICRO EL CARMEN	<b>132</b>	38	47	217
		COL REAL MARES	22	21	39	82
		COL SEÑORA FATIMA	5	11	3	19
		COL SUP INDUSTRIAL	8	0	0	8
COL VICENTE AZUERO	13	20	25	58		

		COL VICTOR FELIX	20	0	0	20
		COL. COOP. COMFENALCO	17	12	11	40
		CRECER Y CONSTRUIR	8	14	18	40
		E.N.S MARIA AUXILIADORA	54	8	0	62
		E.N.S OIBA	33	<b>80</b>	<b>96</b>	209
		E.N.S. FRANCISCO PAULA	8	16	19	43
		E.N.S. SADY TOBON	8	0	0	8
		FUNDE UIS	10	12	15	37
		GABRIELA MISTRAL	30	35	0	65
		GIM. CANTILLANA	6	6	0	12
		GLENN DOMAN	11	12	0	23
		GUSTAVO DUARTE	53	39	0	92
		INDUSTRIAL SOCORRO	24	10	<b>58</b>	92
		INST INFANTAS	14	23	9	46
		INST MORRO	3	3	0	6
		INST TAGÜI	11	16	0	27
		INTEGRADO COMERCIO	12	0	0	12
		JOSE MARIA	9	8	3	20
		LA CUMBRE	6	5	4	15
		LICEO PATRIA	23	24	16	63
		MINCA	1	27	13	41
		NACIONAL COMERCIO	<b>269</b>	<b>93</b>	<b>87</b>	449
		ORIENTE MIRAFLORES	36	0	0	36
		POLITECNICO	3	8	3	14
		S JOSE SALLE	56	<b>51</b>	31	138
		SUPERIOR INDUSTRIAL	0	19	11	30
		Total	1432	1098	923	3,E3
2012	colegio	AQUILEO PARRA	41	37	14	92
		CARLOS GARCIA	20	27	19	66
		CIUDADELA EDUCATIVA	9	7	40	56
		COL AMERICAS	16	14	18	48
		COL CARLOS VICENTE	0	0	1	1
		COL MESA JÉRIDAS	2	4	4	10
		COL QUINTA PUENTE	1	0	0	1
		COL ROEL Y VELASCO	17	15	28	60
		COL SEÑORA ROSARIO	25	26	24	75
		COL BENITO PALERMO	21	12	14	47

COL CASTILLO	19	27	53	99
COL COOPERATIVO	41	38	5	84
COL CUSTODIO GARCIA	52	8	8	68
COL DEP LUIS CARLOS G	26	26	28	80
COL DIEGO HERNANDEZ	18	16	20	54
COL DIVINO NIÑO	5	3	3	11
COL DOMINGO SABIO	16	15	29	60
COL FRANCISCO PAULA	40	40	19	99
COL ISABEL HUNGRIA	8	6	7	21
COL ISABEL VALBUENA	<b>160</b>	<b>93</b>	15	268
COL JORGE ARDILA D	52	48	26	126
COL JOSE ELIAS P	16	16	27	59
COL JUAN CRISTOBAL M	18	24	28	70
COL LORENZO SALAZAR	40	9	5	54
COL LUIS GONZAGA	7	15	17	39
COL MARIA AUXILIADORA	0	1	1	2
COL MICRO EL CARMEN	<b>139</b>	66	<b>83</b>	288
COL MONSEÑOR EVARISTO	5	2	16	23
COL REAL MARES	16	18	17	51
COL SANTA ANA	1	0	0	1
COL SEÑORA FATIMA	8	5	1	14
COL TEC LUIS CARLOS G	19	19	11	49
COL TRINIDAD CAMACHO	13	6	4	23
COL UNIV SOCORRO	15	1	0	16
COL VICENTE AZUERO	20	35	73	128
COL VICTOR FELIX	40	0	0	40
COL. COOP. COMFENALCO	12	11	17	40
DAMASO ZAPATA	27	46	37	110
E.N.S ANTONIA SANTOS	52	0	0	52
E.N.S BUCARAMANGA	87	0	0	87
E.N.S MARIA AUXILIADORA	70	12	2	84
E.N.S OIBA	23	<b>71</b>	64	158
E.N.S. FRANCISCO PAULA	9	22	25	56
ESC INDUSTRIAL OIBA	98	53	26	177
FUNDE UIS	13	16	15	44
GIM SUPERIOR	79	<b>122</b>	<b>169</b>	370

	GIM. CANTILLANA	10	5	6	21
	GLENN DOMAN	16	16	24	56
	INST FELIPE CORDERO	17	18	27	62
	INST INDUSTRIAL	11	5	1	17
	INST ISIDRO MIRANDA	8	9	13	30
	INST MANUEL SORZANO	4	5	3	12
	INST MORRO	3	4	4	11
	INST TAGÜI	15	12	10	37
	JOSE ELIAS PUYANA - SEDE C	0	17	18	35
	JUAN JOSÉ RONDÓN	5	8	9	22
	LA CUMBRE	13	15	12	40
	NACIONAL COMERCIO	<b>230</b>	71	<b>76</b>	377
	ORIENTE MIRAFLORES	16	50	8	74
	S JOSE SALLE	71	68	8	147
	S PEDRO CLAVER	6	8	10	24
	SEM S PEDRO CLAVER	6	11	28	45
	SUPERIOR INDUSTRIAL	3	3	0	6
	<b>Total</b>	<b>1850</b>	<b>1357</b>	<b>1270</b>	<b>4,E3</b>

## ANEXO C. Participación por municipio (2011-2012)

### Tabla 16. Participación por municipio (2011-2012)

Año			Frecuencia	Porcentaje
2011	Válidos	BARBOSA	159	4,6
		BARICHARA	94	2,7
		BARRANCABERMEJA	245	7,1
		BUCARAMANGA	1301	37,7
		CERRITO	8	,2
		ENCISO	17	,5
		FLORIDABLANCA	534	15,5
		GIRÓN	80	2,3
		JESUS MARÍA	59	1,7
		MÁLAGA	119	3,4
		OIBA	209	6,1
		PALMAS DEL SOCORRO	130	3,8
		PIEDRECUESTA	199	5,8
		SAN GIL	166	4,8
		SOCORRO	92	2,7
		SUAITA	35	1,0
		VALLE DE SAN JOSÉ	6	,2
		Total	3453	100,0
		2012	Válidos	AGUADA
BARBOSA	146			3,3
BARICHARA	92			2,1
BARRANCABERMEJA	358			8,0
BUCARAMANGA	1567			35,0
CAPITANEJO	1			,0
CONCEPCION	63			1,4
CONFINES	49			1,1
FLORIDABLANCA	675			15,1
GIRÓN	144			3,2
GÜEPSA	60			1,3
JESUS MARIA	54			1,2
LOS SANTOS	10			,2
MALAGA	121			2,7
OIBA	335			7,5

		PIEDRECUESTA	105	2,3
		PUENTE NACIONAL	93	2,1
		PUERTO WILCHES	24	,5
		SAN ANDRES	84	1,9
		SAN BENITO	47	1,0
		SAN JOSE DE MIRAND	5	,1
		SAN MIGUEL	23	,5
		SOCORRO	82	1,8
		VALLE DE SAN JOSÉ	11	,2
		VÉLEZ	268	6,0
		Total	4477	100,0

**ANEXO D. Status del colegio vs el nivel de desempeño alcanzado durante la prueba clasificatoria**

**Tabla 17. Status del colegio vs el nivel de desempeño alcanzado durante la prueba clasificatoria**

Año	Nivel				Nivel_Prueba_Classificatoria			
					Bajo	Aceptable	Superior	Alto
2011	Nivel Avanzado	status_colegio Oficial	Frecuencia	1099	138	2		1239
			%	<b>88,7%</b>	11,1%	,2%		100,0%
		Privado	Frecuencia	149	40	4		193
			%	<b>77,2%</b>	20,7%	2,1%		100,0%
	Total		Frecuencia	1248	178	6		1432
			%	87,2%	12,4%	,4%		100,0%
	Nivel Básico	status_colegio Oficial	Frecuencia	522	223	12		757
			%	<b>69,0%</b>	29,5%	1,6%		100,0%
		Privado	Frecuencia	90	71	5		166
			%	<b>54,2%</b>	42,8%	3,0%		100,0%
	Total		Frecuencia	612	294	17		923
			%	66,3%	31,9%	1,8%		100,0%
Nivel Medio	status_colegio Oficial	Frecuencia	712	164	5	0	881	
		%	<b>80,8%</b>	18,6%	,6%	,0%	100,0%	
	Privado	Frecuencia	173	33	10	1	217	
		%	<b>79,7%</b>	15,2%	4,6%	,5%	100,0%	
Total		Frecuencia	885	197	15	1	1098	
		%	80,6%	17,9%	1,4%	,1%	100,0%	

2012	Nivel Avanzado	status_colegio	Oficial	Frecuencia	1368	274	31	3	1676
				%	<b>81,6%</b>	16,3%	1,8%	,2%	100,0%
		Privado	Frecuencia	146	24	3	1	174	
			%	<b>83,9%</b>	13,8%	1,7%	,6%	100,0%	
	Total		Frecuencia	1514	298	34	4	1850	
			%	81,8%	16,1%	1,8%	,2%	100,0%	
Nivel Básico	status_colegio	Oficial	Frecuencia	947	67	7		1021	
			%	<b>92,8%</b>	6,6%	,7%		100,0%	
		Privado	Frecuencia	232	17	0		249	
			%	<b>93,2%</b>	6,8%	,0%		100,0%	
	Total		Frecuencia	1179	84	7		1270	
			%	92,8%	6,6%	,6%		100,0%	
Nivel Medio	status_colegio	Oficial	Frecuencia	1047	81	5	0	1133	
			%	<b>92,4%</b>	7,1%	,4%	,0%	100,0%	
		Privado	Frecuencia	206	16	1	1	224	
			%	<b>92,0%</b>	7,1%	,4%	,4%	100,0%	
	Total		Frecuencia	1253	97	6	1	1357	
			%	92,3%	7,1%	,4%	,1%	100,0%	

**Tabla 18. Status del colegio vs el nivel de desempeño alcanzado durante la prueba clasificatoria**

Año	Nivel	Nivel_Prueba_Classificatoria				Total		
		Bajo	Aceptable	Superior	Alto			
2011	Avanzado	zona_colegio Urbano	Frecuencia	1092	169	5		1266
			%	<b>86,3%</b>	13,3%	,4%		100,0%
		Rural	Frecuencia	156	9	1		166
			%	<b>94,0%</b>	5,4%	,6%		100,0%
		Total	Frecuencia	1248	178	6		1432
			%	87,2%	12,4%	,4%		100,0%
	Básico	zona_colegio Urbano	Frecuencia	612	294	17		923
			%	<b>66,3%</b>	31,9%	1,8%		100,0%
		Total	Frecuencia	612	294	17		923
			%	66,3%	31,9%	1,8%		100,0%
Medio	zona_colegio Urbano	Frecuencia	793	185	15	1	994	
		%	<b>79,8%</b>	18,6%	1,5%	,1%	100,0%	
	Rural	Frecuencia	92	12	0	0	104	
		%	<b>88,5%</b>	11,5%	,0%	,0%	100,0%	
	Total	Frecuencia	885	197	15	1	1098	
		%	80,6%	17,9%	1,4%	,1%	100,0%	
2012	Avanzado	zona_colegio Urbano	Frecuencia	1264	272	31	4	1571
			%	<b>80,5%</b>	17,3%	2,0%	,3%	100,0%
		Rural	Frecuencia	250	26	3	0	279
		%	<b>89,6%</b>	9,3%	1,1%	,0%	100,0%	
		Total	Frecuencia	1514	298	34	4	1850
			%	81,8%	16,1%	1,8%	,2%	100,0%
	Básico	zona_colegio Urbano	Frecuencia	1024	81	6		1111
			%	<b>92,2%</b>	7,3%	,5%		100,0%
		Rural	Frecuencia	155	3	1		159
		%	<b>97,5%</b>	1,9%	,6%		100,0%	
	Total	Frecuencia	1179	84	7		1270	
		%	92,8%	6,6%	,6%		100,0%	
	Nivel	zona_colegio Urbano	Frecuencia	1076	89	6	1	1172

Medio		%	<b>91,8%</b>	7,6%	,5%	,1%	100,0%
	Rural	Frecuencia	177	8	0	0	185
		%	<b>95,7%</b>	4,3%	,0%	,0%	100,0%
		Frecuencia	1253	97	6	1	1357
		%	<b>92,3%</b>	7,1%	,4%	,1%	100,0%

**ANEXO E. Porcentaje de respuestas correctas en la prueba clasificatoria y selectiva (2011)**

**Tabla 19. Porcentaje de respuestas correctas en la prueba clasificatoria y selectiva (2011)**

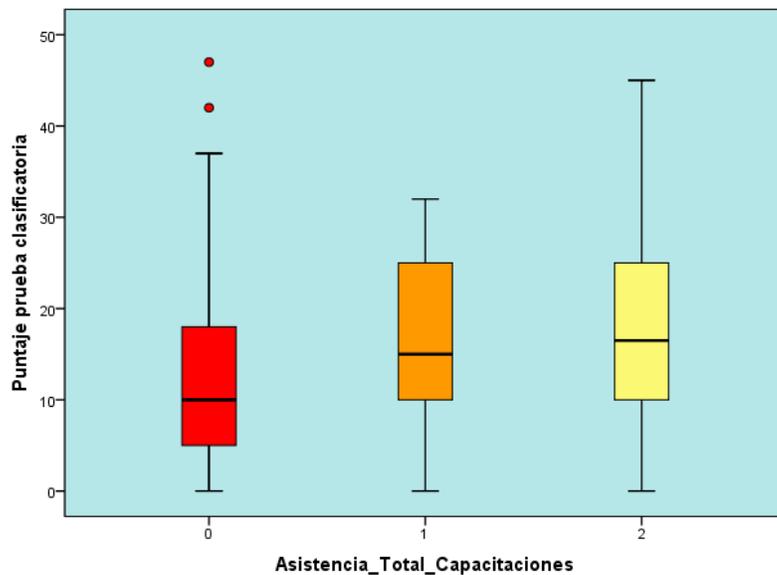
Resultados Olimpiada Regional 2011										
Fase	Área Temática	Nivel Avanzado			Nivel Medio			Nivel Básico		
		% Rta Correcto			% Rta Correcto			% Rta Correcto		
		pregunta	Total	Femenino	pregunta	Total	Femenino	pregunta	Total	Femenino
Clasificatoria	Teoría de números y combinatoria	pc5	37,1	16,7	pc1	<b>75,6</b>	33,8	pc7	32,1	13,9
		pc2	20,20	8,3	pc4	24,5	11,2	pc1	28,2	12,2
		pc8	12,4	6	pc7	24,3	11,1	pc4	23,9	12,1
		pc11	6,9	3	pc10	18,4	8,4	pc10	8,2	3,2
	Geometría	pc1	37,1	16,2	pc8	26,9	13,6	pc5	<b>53,3</b>	26,3
		pc7	19,3	9,5	pc2	21,9	11,5	pc8	28,5	14,1
		pc10	14,6	7,6	pc5	19,7	8,7	pc11	16,5	7,6
		pc4	14	6,2	pc11	18,2	8,9	pc2	15,9	7,9
	Álgebra y lógica	pc12	31,9	13,4	pc12	44,5	20	pc6	<b>68,8</b>	<b>34</b>
		pc9	16,7	7,6	pc3	15,8	7,2	pc9	<b>45,4</b>	<b>22,2</b>
		pc6	14,9	7,4	pc6	13,8	7,7	pc3	<b>42,2</b>	<b>18</b>
		pc3	13,7	6,4	pc9	10,5	4	pc12	16,5	7,1
Selectiva	Teoría de números y combinatoria	ps2	18,1	6,7	ps2	36,8	14,6	ps2	35	12,2
		ps5	11,9	1,6	ps6	38,9	13,2	ps5	28,5	8,9
	Geometría	ps3	<b>48,7</b>	15,5	ps3	31,9	12,5	ps3	32,5	12,2
		ps6	9,8	3,6	ps5	11,8	6,9	ps6	8,9	1,6
	Álgebra y lógica	ps1	<b>85,5</b>	30,6	ps1	<b>61,1</b>	20,8	ps4	<b>43,1</b>	14,6
		ps4	15,5	6,2	ps4	<b>58,3</b>	19,4	ps1	5,7	2,4

Tabla 20. Porcentaje de respuestas correctas en la prueba clasificatoria y selectiva (2012)

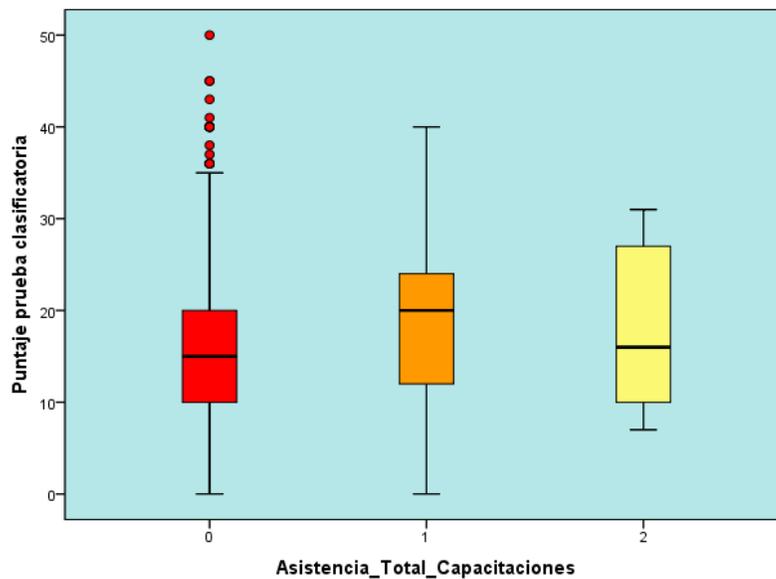
Resultados Olimpiada Regional 2012										
Fase	Área Temática	Nivel Avanzado			Nivel Medio			Nivel Básico		
		% Rta Correcto			% Rta Correcto			% Rta Correcto		
		pregunta	Total	Femenino	pregunta	Total	Femenino	pregunta	Total	Femenino
Clasificatoria	Teoría de números y combinatoria	pc11	42	18,7	pc8	20	8,4	pc11	25	10,4
		pc5	34,7	15,4	pc5	15,3	6	pc2	13,2	5,1
		pc2	29,8	13,6	pc11	14,9	5,8	pc5	9,6	4,3
		pc8	26,9	11,8	pc2	14,3	5,8	pc8	9,2	4,3
	Geometría	pc1	25,3	11,8	pc1	25,9	11	pc7	22,8	9,9
		pc4	16,5	7,8	pc4	22,4	9,8	pc1	22,5	10,6
		pc10	12,1	5,7	pc10	10,2	3,8	pc4	15,3	6,9
		pc7	9,8	4,3	pc7	10,1	4,7	pc10	11,8	4,7
	Álgebra y lógica	pc12	27,7	13,2	pc6	31,9	13,7	pc12	35,2	14,5
		pc6	26,9	12,6	pc12	27,8	12,2	pc9	23,2	10,2
		pc9	15,1	8	pc3	17,8	8,1	pc6	20	8,6
		pc3	14,8	7,2	pc9	13,8	5,7	pc3	9,1	4,3
selectiva	Teoría de números y combinatoria	ps2	<b>78,3</b>	30,4	ps5	<b>62,5</b>	20,8	pc2	<b>54,5</b>	18,2
		ps5	39,1	13	ps2	33,3	16,7	pc5	<b>54,5</b>	18,2
	Geometría	ps6	<b>73,9</b>	26,1	ps3	<b>79,2</b>	33,3	pc3	<b>72,7</b>	27,3
		ps3	<b>56,5</b>	21,7	ps6	16,7	4,2	pc6	<b>50</b>	18,2
	Álgebra y lógica	ps4	<b>47,8</b>	13	ps4	41,7	20,8	pc1	<b>90,9</b>	<b>36,4</b>
		ps1	13	0	ps1	45,8	16,7	pc4	36,4	9,1

## ANEXO F. Gráficas Comparativas del rendimiento discriminando por el número de capacitaciones a las que asistieron por niveles

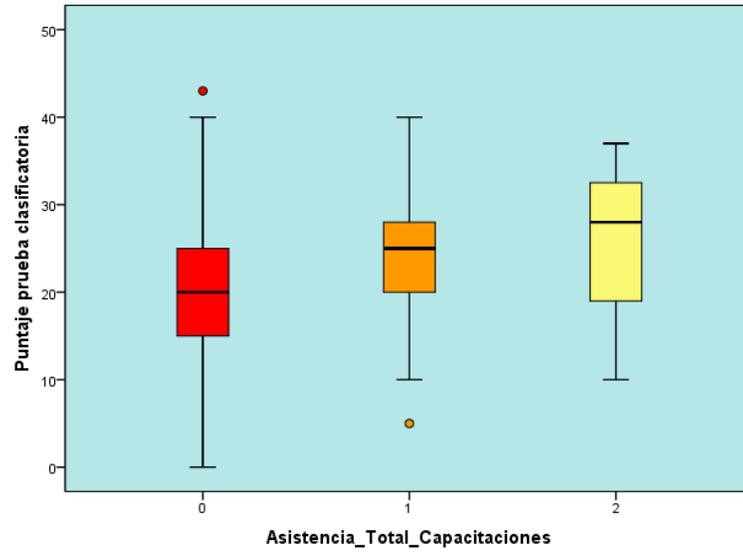
Gráfica 27. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Avanzado 2011)



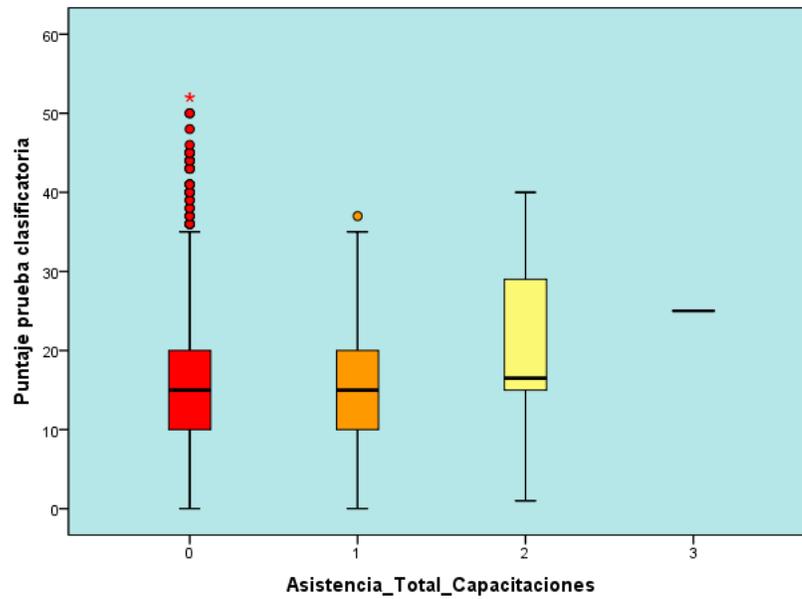
Gráfica 28. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Medio 2011)



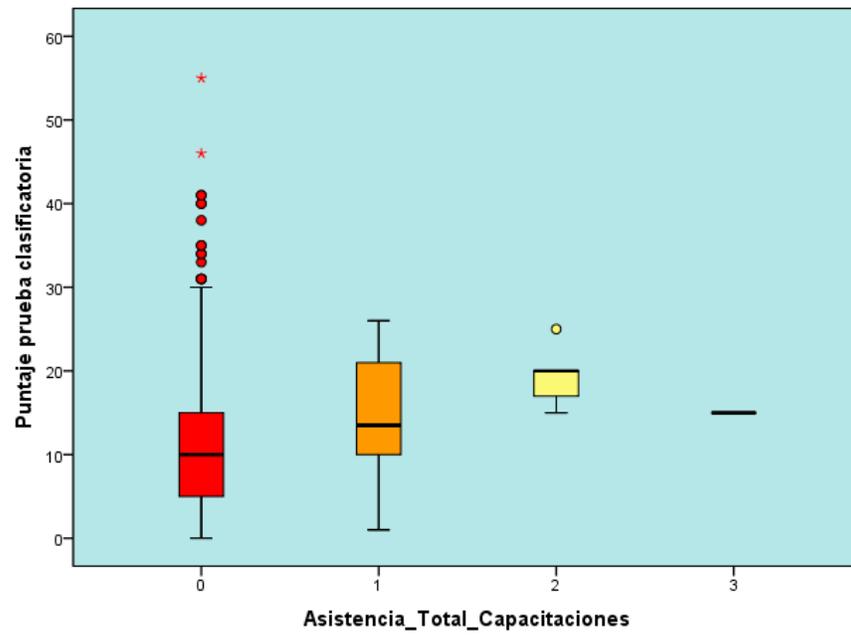
**Gráfica 29. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Básico 2011)**



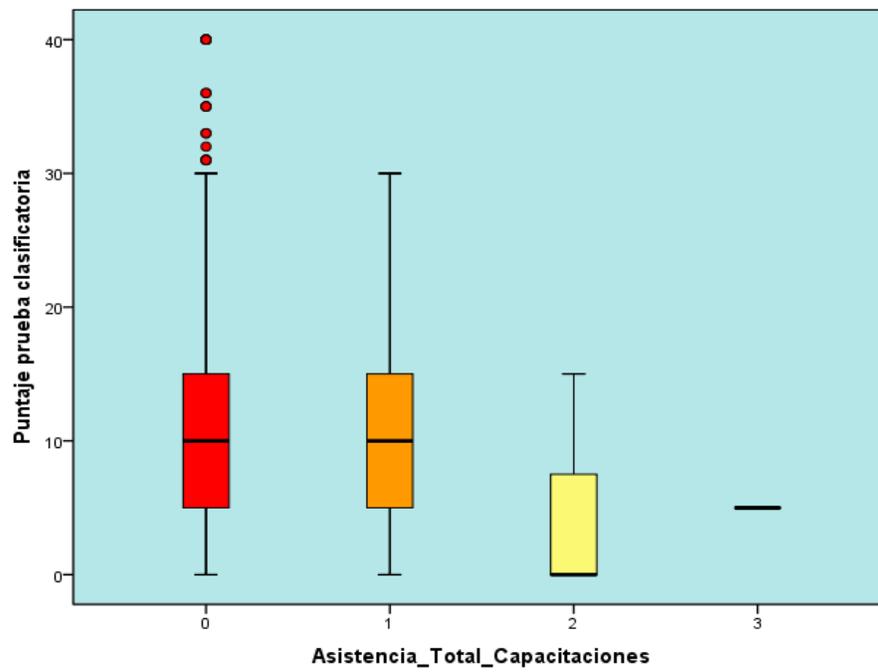
**Gráfica 30. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Avanzado 2012)**



**Gráfica 31. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel medio 2012)**



**Gráfica 32. Comparativo del rendimiento discriminando por el número de capacitaciones a las que asistieron (Nivel Básico 2012)**



## ANEXO G. Tabla de estadísticos descriptivos discriminando por género

**Tabla 21. Tabla de estadísticos descriptivos discriminando por género**

Año	Nivel	Fase Clasificatoria							Fase Selectiva							Fase Final						
		$\bar{x}$	$\tilde{x}$	<i>S</i>	<i>CV</i>	<i>P.25</i>	<i>P.75</i>	<i>Max</i>	$\bar{x}$	$\tilde{x}$	<i>S</i>	<i>CV</i>	<i>P.25</i>	<i>P.75</i>	<i>Max</i>	$\bar{x}$	$\tilde{x}$	<i>S</i>	<i>CV</i>	<i>P.25</i>	<i>P.75</i>	<i>Max</i>
2011	Avanzado	12.13	10	8.07	66,53	5	16	42	12.01	12	6.22	51,79	6.75	16	29	6	2	7.34	122,33	25.1	13	18
		14,03	13	8,254	58,83	7	20	47	14,65	13	7,416	50,62	9	19	43	7,58	6	7,712	101,74	2	11	29
	Medio	16.52	15	8.04	48,67	10	20	50	12.21	11	8.82	72,24	5	17	48	10.50	11	8.85	84,29	2	18.5	19
		16,92	15	8,457	49,98	10	21	45	16,95	16	8,184	48,28	11	22	38	8,82	8	6,55	74,26	3,5	12	25
	Básico	19.37	20	7.41	38,26	15	25	40	10.61	10	6.54	61,64	6	14.25	26	5	6	3,464	69,28	1,5	7,5	8
		20,05	20	8,319	41,49	15	25	43	13,48	12	8,464	62,79	6	19,5	38	4,50	5,50	2,902	64,49	1,75	6	10
2012	Avanzado	14,55	15	9,023	62,01	9	20	46	19,25	18,5	3,808	19,78	16	21,75	26	4,86	3	5,336	109,79	1	6	16
		16,05	15	9,464	58,97	10	21	52	19,33	17	5,273	27,28	16	20	36	12,43	13,5	9,436	75,91	2,50	19,5	27
	Medio	12.06	10	7.561	62,69	5	15	41	21,67	21	2,915	13,45	19,50	23	28	5,88	4	5,793	98,52	1	11,50	15
		12,23	10	7,49	61,24	5	16	55	21,33	20	4,546	21,31	18	23	32	6,29	6	4,921	78,24	2,5	9	18
	Básico	11,08	10	7,329	66,15	5	15	40	21,25	20	3,059	14,40	19,25	24,75	26	7,13	6	5,055	70,90	2,50	11,75	16
		11,65	10	7,461	64,04	5	15	40	22,57	21	3,435	15,22	20	25,25	31	7,86	7	6,249	79,50	2	12,5	21

**ANEXO H. Cálculos usando los resultados de la fase clasificatoria del Nivel Avanzado (2012)**

**Tabla 22. Cálculos usando los resultados de la fase clasificatoria del Nivel Avanzado (2012)**

Pregunta	Odds ratio de no acierto		
	Mujeres	Hombres	Odds ratio
1	4,06	3,84	1,05
2	3,53	3,21	1,09
3	6.68	6,86	0,97
4	6,13	5,96	1,02
5	3,10	2,70	1,14
6	3,78	3,66	1,03
7	11.03	9,51	1,15
8	4,06	3,44	1,18
9	5,96	7,33	0,81
10	8.33	8,21	1,01
11	2,56	2,22	1,15
12	3,29	3,58	0,91