

**Evaluación de Impacto del Sistema de Apoyo a la Excelencia (SEA) en el Desempeño en  
Matemáticas de Estudiantes de Ciencias e Ingenierías Entre 2014 y 2017**

**Raúl Gabriel Ramos**

**Juan Pablo Osorio Suarez**

**Trabajo de Grado para Optar al Título de Especialista en Estadística**

**Director:**

**Gabriel Yañez Canal**

**Doctor en Ciencias Especialidad en Matematica Educativa**

**Universidad Industrial de Santander**

**Facultad de Ciencias**

**Escuela de Matemáticas**

**Especialización en Estadística**

**Bucaramanga**

**2019**

## Contenido

	<b>Pág.</b>
Introducción .....	26
1. Antecedentes .....	29
1.1 Estado Nacional de la Deserción Académica .....	29
1.2 Programa Sea .....	31
1.3 Estudios Sobre Deserción y Rendimiento Académico .....	34
2. Marco Teórico.....	38
2.1 Conceptos de Estadística.....	38
2.1.1 Prueba de hipótesis para la media, la proporción y valor p. ....	38
2.1.2 Regresión logística.....	39
2.1.3 Criterio de AIC. ....	39
2.1.4 Selección de variables y Regresión por pasos. ....	40
2.2 Técnicas de Evaluación del Impacto.....	41
2.3 Método de Pareamiento .....	49
3. Resultados y Discusión .....	53
3.1 Análisis Descriptivo.....	53
3.1.1 Estudiantes Cohorte I - 2014 / Semestre 1 a 6.....	58

3.1.1.1 Características Generales.....	58
3.1.1.2 Desempeño Académico.....	63
3.1.1.3 Pruebas SABER 11.....	79
3.1.1.4 Beneficiarios Programa SEA.....	83
3.1.1.5 Caracterización Ingreso SEA.....	95
3.1.1.6 Estudiantes Cohorte I – 2014 a 2017 / Semestre 1.....	101
3.1.1.7 Características Generales.....	101
3.1.1.8 Desempeño Académico.....	104
3.1.1.8 Pruebas Saber 11.....	113
3.1.1.9 Beneficiarios Programa SEA.....	116
3.1.1.10 Caracterización Ingreso SEA.....	121
3.2 Evaluación de Impacto del Programa Sea.....	127
3.2.1 Impacto en cohorte 2014-1.....	128
3.2.1.1 Impacto del SEA sobre Cálculo I. Cohorte 2014-1.....	128
3.2.1.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2014-1.....	143
3.2.1.3 Impacto del SEA sobre álgebra lineal. Cohorte 2014-1.....	146
3.2.1.4 Impacto del SEA sobre repitentes en álgebra lineal I. Cohorte 2014-1.....	157
3.2.1.5 Impacto del SEA sobre cálculo II. Cohorte 2014-1.....	158
3.2.1.6 Impacto del SEA sobre repitentes en cálculo II. Cohorte 2014-1.....	168
3.2.1.7 Impacto del SEA sobre cálculo III. Cohorte 2014-1.....	169
3.2.1.8 Impacto del SEA sobre ecuaciones diferenciales. Cohorte 2014-1.....	178
3.2.1.9 Síntesis de resultados para la primera cohorte del 2014.....	179
3.2.2 Impacto en cohorte 2015-1.....	180

3.2.2.1 Impacto del SEA sobre cálculo I. Cohorte 2015-1. ....	181
3.2.2.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2015-1. ....	188
3.2.2.3 Impacto del SEA sobre álgebra lineal. Cohorte 2015-1. ....	190
3.2.2.4 Impacto del SEA sobre repitentes en álgebra lineal. Cohorte 2015-1. ....	198
3.2.3 Impacto en cohorte 2016-1 .....	198
3.2.3.1 Impacto del SEA sobre Cálculo I. Cohorte 2016-1. ....	198
3.2.3.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2016-1. ....	204
3.2.3.3 Impacto del SEA sobre Álgebra Lineal. Cohorte 2016-1. ....	206
3.2.3.4 Impacto del SEA sobre repitentes en álgebra lineal. Cohorte 2016-1. ....	211
3.2.4 Impacto en cohorte 2017-1 .....	212
3.2.4.1 Impacto del SEA sobre Cálculo I. Cohorte 2017-1. ....	212
3.2.4.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2017-1. ....	220
3.2.4.3 Impacto del SEA sobre Álgebra Lineal. Cohorte 2017-1. ....	222
3.2.4.4 Impacto del SEA sobre repitentes en álgebra lineal. Cohorte 2017-1. ....	229
3.2.4.5 Análisis consolidado y comparativo de las cohortes estudiadas. ....	230
4. Conclusiones .....	236
5. Recomendaciones .....	238
Referencias Bibliograficas .....	240
Apéndices.....	247

## Lista de Figuras

	<b>Pág.</b>
Figura 1. Evaluación numérica del impacto. ....	43
Figura 2. Selección aleatoria de los grupos de contraste. ....	44
Figura 3. Ejemplo de aplicación método de regresión discontinua. ....	46
Figura 4. Ejemplo de aplicación de método de diferencia en diferencias. ....	47
Figura 5. Técnicas de evaluación de impacto y potencial para el caso SEA. ....	49
Figura 6. Rango común de la propensión. ....	51
Figura 3. Distribución por género cohortes 2014, Estudiantes de Ciencias e Ingenierías. ....	59
Figura 4. Proporción de estudiantes por género cohortes 2014, Estudiantes de Ciencias e Ingenierías. ....	60
Figura 5. Distribución por edad de ingreso cohortes 2014, Estudiantes de Ciencias e Ingenierías. ....	61
Figura 6. Proporción de estudiante que ingresaron en 2014 por edad, Estudiantes de Ciencias e Ingenierías. ....	62
Figura 7. Promedio semestral, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 semestres. ....	65
Figura 8. Promedio acumulado, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 semestres. ....	67
Figura 9. Distribución por condición académica, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 Semestres. ....	70

Figura 10. Distribución por condición académica, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 Semestres. ....	72
Figura 11. Distribución por continuidad y deserción, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 Semestres. ....	74
Figura 12. Distribución por continuidad y deserción, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014 – sexto semestre. ....	75
Figura 13. Distribución de resultados en matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	78
Figura 14. Distribución de resultados en pruebas SABER 11 Grupo 1, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	81
Figura 15. Distribución de resultados en pruebas SABER 11 Grupo 2, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	82
Figura 16. Beneficiarios del programa SEA por carrera, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	84
Figura 17. Proporción de Beneficiarios del programa SEA por carrera, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	85
Figura 19. Beneficiarios del programa SEA por carrera y asignatura, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	86
Figura 20. Proporción de Beneficiarios del programa SEA por carrera y asignatura, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. ....	88
Figura 21. Cantidad de Beneficiarios SEA por sub-programa y carrera, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	90

Figura 22. Proporción de Beneficiarios SEA por sub-programa y carrera, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	90
Figura 23. Intensidad de beneficios del programa SEA por carrera en matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	92
Figura 24. Beneficiarios del programa SEA en Matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.....	93
Figura 25. Proporción de Beneficiarios por Sub-programa SEA en matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014. ....	94
Figura 26. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera. 1ª cohorte de 2014.....	99
Figura 27. Proporción de Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera. 1ª cohorte de 2014 .....	100
Figura 28. Distribución por género, Estudiantes de Ciencias e Ingeniería, primera cohorte 2014 a 2017.....	102
Figura 29. Distribución por edad, Estudiantes de Ciencias e Ingeniería, primera cohorte 2014 a 2017.....	103
Figura 30. Promedio semestral, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	106
Figura 32. Distribución de resultados para Cálculo I, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	111
Figura 33. Distribución de resultados para Álgebra, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	112

Figura 34. Distribución de resultados en pruebas Saber 11 en matemáticas, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	115
Figura 35. Intensidad de Beneficios en Cálculo I del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	118
Figura 36. Intensidad de Beneficios en Álgebra del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	119
Figura 37. Beneficiarios en Cálculo I del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	120
Figura 38. Intensidad de Beneficios en Cálculo I del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	121
Figura 39. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2014.....	123
Figura 40. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2015.....	124
Figura 41. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2016.....	125
Figura 42. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2017.....	126
Figura 43. Beneficios del SEA en cálculo I. 1ª cohorte de 2014.....	129
Figura 44. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2014 .....	135
Figura 45. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2014 .....	136
Figura 46. Histograma de puntajes de propensión para cálculo I antes (izquierda) y después (derecha) del pareamiento. 1ª cohorte de 2014.....	137

Figura 47. Histograma bi-variado de 100 diferencias promedio en notas de cálculo I entre beneficiados y no beneficiados del SEA con pareamiento 1:1. 1ª cohorte de 2014 .....	141
Figura 48. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2014.....	143
Figura 49. Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2014.....	144
Figura 50. Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2014 .....	145
Figura 51. Beneficios del SEA en álgebra con información completa. 1ª cohorte de 2014. ....	146
Figura 52. Beneficios del SEA en álgebra con información completa excluyendo variables de riesgo. 1ª cohorte de 2014.....	147
Figura 53. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2014 .....	151
Figura 54. Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2014 .....	152
Figura 55. Histograma de puntajes de propensión para álgebra I antes y después del pareamiento. 1ª cohorte de 2014.....	153
Figura 56. Histograma de 100 diferencias promedio en notas de álgebra I entre beneficiados y no beneficiados del SEA. 1ª cohorte de 2014.....	155
Figura 57. Impacto del programa SEA sobre álgebra en función del número de horas de beneficio. 1ª cohorte de 2014.....	156
Figura 58. Diferencias en diferencias a repitentes en álgebra lineal I. 1ª cohorte de 2014 .....	158
Figura 59. Beneficios del SEA en cálculo II con información completa. 1ª cohorte de 2014....	159
Figura 60. Histograma de puntajes de propensión para cálculo II antes y después del pareamiento. 1ª cohorte de 2014.....	165
Figura 61. Impacto del programa SEA sobre cálculo II en función del número de horas de beneficio. 1ª cohorte de 2014.....	167

Figura 62. Diferencias en diferencias a repitentes en cálculo II. 1ª cohorte de 2014 .....	168
Figura 63. Beneficios del SEA en cálculo III con información completa. 1ª cohorte de 2014. .	170
Figura 64. Pareamiento por puntajes de propensión para cálculo II. 1ª cohorte de 2014.....	173
Figura 65. Histograma de puntajes de propensión para cálculo II antes y después del pareamiento. 1ª cohorte de 2014.....	174
Figura 66. Histograma bi-variado de 100 diferencias promedio en notas de cálculo III entre beneficiados y no beneficiados del SEA con pareamiento 1:1. 1ª cohorte de 2014 .....	177
Figura 67. Impacto del programa SEA sobre cálculo III en función del número de horas de beneficio. 1ª cohorte de 2014.....	178
Figura 68. Beneficios del SEA en cálculo I. 1ª cohorte de 2015.....	181
Figura 69. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2015 .....	184
Figura 70. Histograma de puntajes de propensión para cálculo I antes y después del pareamiento. 1ª cohorte de 2015.....	185
Figura 71. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2015.....	188
Figura 72. Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2015.....	189
Figura 73. Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2015 .....	190
Figura 74. Beneficios del SEA en álgebra. 1ª cohorte de 2015.....	191
Figura 75. Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2015 .....	194
Figura 76. Histograma de puntajes de propensión para álgebra I antes y después del pareamiento. 1ª cohorte de 2015.....	195
Figura 77. Impacto del programa SEA sobre álgebra I en función del número de horas de beneficio. 1ª cohorte de 2015.....	197

Figura 78. Beneficios del SEA en cálculo I. 1ª cohorte de 2016. ....	199
Figura 79. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2016 .....	202
Figura 80. Histograma de puntajes de propensión para cálculo I antes y después del pareamiento. 1ª cohorte de 2016.....	203
Figura 81. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2016.....	204
Figura 82. Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2016.....	205
Figura 83. Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2016. ....	205
Figura 84. Beneficios del SEA en álgebra I. 1ª cohorte de 2016.....	206
Figura 85. Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2016 .....	209
Figura 86. Histograma de puntajes de propensión para álgebra I antes y después del pareamiento. 1ª cohorte de 2016.....	210
Figura 87. Impacto del programa SEA sobre álgebra I en función del número de horas de beneficio. 1ª cohorte de 2016.....	210
Figura 88. Beneficios del SEA para repitentes en álgebra I. 1ª cohorte de 2016. ....	211
Figura 89. Diferencias en diferencias a repitentes en álgebra I. 1ª cohorte de 2016 .....	211
Figura 90. Beneficios del SEA en cálculo I. 1ª cohorte de 2017. ....	213
Figura 91. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2017 .....	216
Figura 92. Histograma de puntajes de propensión para cálculo I antes y después del pareamiento. 1ª cohorte de 2017.....	217
Figura 93. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2017.....	220
Figura 94. Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2017.....	221

Figura 95. Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2017 .....	221
Figura 96. Beneficios del SEA en álgebra I. 1ª cohorte de 2017.....	222
Figura 97. Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2017 .....	225
Figura 98. Histograma de puntajes de propensión para álgebra I antes y después del pareamiento. 1ª cohorte de 2017.....	226
Figura 99. Impacto del programa SEA sobre álgebra lineal en función del número de horas de beneficio. 1ª cohorte de 2017 .....	228
Figura 100. Diferencias en diferencias a repitentes en álgebra I. 1ª cohorte de 2017. ....	230

### Lista de Tablas

	<b>Pág.</b>
Tabla 1. Resumen general de la de información.....	54
Tabla 2. Codificación de Carreras. ....	56
Tabla 3. Desempeño académico global, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014.....	63
Tabla 4. Condición académica global. 1 <sup>a</sup> cohorte del 2014. Primeros 6 semestres. ....	69
Tabla 5. Desempeño académico en asignaturas de matemáticas, Estudiantes de Ciencias e Ingeniería, 1 <sup>a</sup> cohorte de 2014 .....	76
Tabla 6. Resultados globales en Pruebas SABER 11, Estudiantes de Ciencias e Ingeniería, 1 <sup>a</sup> cohorte de 2014.....	79
Tabla 7. Intensidad global de beneficios SEA en matemáticas, Estudiantes de Ciencias e Ingenierías, 1 <sup>a</sup> cohorte del 2014.....	94
Tabla 8. Categorización de puntaje en variables de caracterización del SEA (excepto cognitiva) .....	97
Tabla 9. Categorización de puntaje en dimensión cognitiva programa SEA .....	97
Tabla 10. Desempeño académico global, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1 <sup>a</sup> cohorte de 2014 a 2017.....	104
Tabla 11. Desempeño académico en Cálculo I, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1 <sup>a</sup> cohorte de 2014 a 2017.....	109

Tabla 12. Desempeño académico en Álgebra, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.....	109
Tabla 13. Resultados globales en Pruebas SABER 11, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014.....	113
Tabla 14. Resultados globales en Horas de Beneficio Cálculo I, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014.....	117
Tabla 15. Resultados globales en Horas de Beneficio Álgebra, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014.....	118
Tabla 16. Resultados globales de caracterización SEA, Semestre 1 - 1ª cohorte de 2014 a 2017. ....	122
Tabla 17. Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2014.....	130
Tabla 18. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2014 (Modelo Completo.).....	132
Tabla 19. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2014 (Modelo depurado.).....	134
Tabla 20. Resultados finales del pareamiento para beneficiados en cálculo I. 1ª cohorte de 2014 de Ciencias e Ingeniería.....	136
Tabla 21. Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2014.....	138
Tabla 22. Cantidad de beneficiarios del programa SEA en función del número de horas de tutoría recibida para Cálculo I. 1ª cohorte del 2014 .....	142

Tabla 23. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2014.....	142
Tabla 24. Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2014.....	146
Tabla 25. Comparación del grupo de Beneficiados y no Beneficiados del SEA en álgebra. 1ª cohorte de 2014.....	148
Tabla 26. Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2014 (Modelo completo).....	149
Tabla 27. Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2014 (Modelo depurado).....	150
Tabla 28. Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2014 de Ciencias e Ingeniería.....	153
Tabla 29. Comparación de estudiantes con beneficio o no del SEA en álgebra I (después de pareamiento). 1ª cohorte de 2014.....	154
Tabla 30. Estimador del impacto del programa SEA sobre repitentes de álgebra lineal I usando diferencias en diferencias. 1ª cohorte de 2014.....	158
Tabla 31. Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo II. 1ª cohorte de 2014.....	159
Tabla 32. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo II. 1ª cohorte de 2014 (Modelo completo).....	161
Tabla 33. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo II. 1ª cohorte de 2014 (Modelo depurado).....	163

Tabla 34. Resultados finales del pareamiento para beneficiados en cálculo II. 1ª cohorte de 2014 de Ciencias e Ingeniería .....	164
Tabla 35. Comparación de estudiantes con beneficio o no del SEA en cálculo II (después de pareamiento). 1ª cohorte de 2014.....	166
Tabla 36. Estimador del impacto del programa SEA sobre repitentes de cálculo II usando diferencias en diferencias. 1ª cohorte de 2014.....	169
Tabla 37. Resumen de variables cuantitativas para el modelo de propensión de cálculo III. 1ª cohorte de 2014.....	170
Tabla 38. Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo III. 1ª cohorte de 2014.....	171
Tabla 39. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo III. 1ª cohorte de 2014 (Modelo depurado).....	172
Tabla 40. Resultados finales del pareamiento para beneficiados en cálculo II. 1ª cohorte de 2014 de Ciencias e Ingeniería .....	174
Tabla 41. Comparación de estudiantes con beneficio o no del SEA en cálculo III (después de pareamiento). 1ª cohorte de 2014.....	175
Tabla 42. Diferencias promedio en notas de cálculo III entre beneficiados y no beneficiados del SEA en 100 corridas de pareamiento 1:1. 1ª cohorte de 2014.....	177
Tabla 43. Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2015.....	182
Tabla 44. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2015 (Modelo depurado.).....	183

Tabla 45. Resultados finales del pareamiento para beneficiados en cálculo I. 1ª cohorte de 2015 de Ciencias e Ingeniería.....	184
Tabla 46. Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2015.....	186
Tabla 47. Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2015.....	190
Tabla 48. Comparación del grupo de Beneficiados y no Beneficiados del SEA en álgebra I. 1ª cohorte de 2015.....	191
Tabla 49. Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2015 (Modelo depurado.).....	193
Tabla 50. Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2015 de Ciencias e Ingeniería.....	194
Tabla 51. Comparación de estudiantes con beneficio o no del SEA en álgebra I (después de pareamiento). 1ª cohorte de 2015.....	195
Tabla 52. Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2016.....	199
Tabla 53. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2016 (Modelo depurado.).....	201
Tabla 54. Resultados finales del pareamiento para beneficiados en cálculo I. 1ª cohorte de 2016 de Ciencias e Ingeniería.....	202
Tabla 55. Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2016.....	206

Tabla 56. Comparación del grupo de Beneficiados y no Beneficiados del SEA en álgebra I. 1ª cohorte de 2016.....	207
Tabla 57. Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2016 (Modelo depurado.).....	208
Tabla 58. Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2016 de Ciencias e Ingeniería.....	209
Tabla 59. Estimador del impacto del programa SEA sobre repitentes de álgebra I usando diferencias en diferencias. 1ª cohorte de 2016.....	212
Tabla 60. Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2017.....	214
Tabla 61. Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2017 (Modelo depurado.).....	215
Tabla 62. Resultados finales del pareamiento para beneficiados en cálculo I. 1ª cohorte de 2017 de Ciencias e Ingeniería.....	217
Tabla 63. Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2017.....	218
Tabla 64. Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2017.....	222
Tabla 65. Comparación del grupo de Beneficiados y no Beneficiados del SEA en álgebra I. 1ª cohorte de 2017.....	223
Tabla 66. Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2017 (Modelo depurado.).....	224

Tabla 67. Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2017 de Ciencias e Ingeniería.....	226
Tabla 68. Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2017.....	227
Tabla 69. Estimador del impacto del programa SEA sobre repitentes de álgebra I usando diferencias en diferencias. 1ª cohorte de 2017.....	230
Tabla 70. Resumen de variables significativas en modelos de propensión, cohorte 2014-1.....	231
Tabla 71. Resumen de variables significativas en modelos de propensión en cálculo I y álgebra lineal, 2014-2017.....	232
Tabla 72. Resultados relevantes de evaluación de impacto usando el método de pareamiento.2014-2017.....	234
Tabla 73. Resumen de resultados claves del uso del método de Diferencias en Diferencias en estudiantes repitentes de cálculo y álgebra lineal. 2014-2017.....	235

**Lista de Apéndices**

	<b>Pág.</b>
Apéndice A. Script principal en R .....	247
Apéndice B. Funciones en R y paquetes empleados.....	249
Apéndice C. Evaluación de impacto por pareamiento en R .....	257
Apéndice D. Evaluación de impacto por diferencias en diferencias en R .....	259

## Resumen

**Título:** Evaluación de Impacto del Sistema de Apoyo a la Excelencia (SEA) en el Desempeño en Matemáticas de Estudiantes de Ciencias e Ingenierías Entre 2014 y 2017.\*

**Autores:** Ramos, Raúl Gabriel; Osorio, Juan Pablo.\*\*

**Palabras Claves:** Evaluación de impacto, pareamiento, diferencias en diferencias, SEA, PAMRA, MIDAS.

### Descripción:

El presente trabajo evalúa el impacto del programa SEA sobre asignaturas de matemáticas vistas por estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte de cada año entre el 2014 y 2017.

El estudio se realizó empleando la técnica de pareamiento y diferencias en diferencias teniendo en cuenta las siguientes variables: las características generales del estudiante como la edad, género y programa académico al que ingresó; los resultados de las pruebas SABER 11; la caracterización realizada por el programa SEA de los estudiantes en las dimensiones social, salud, económica, académica y cognitiva; la nota final en las asignaturas de cálculo I, II y III, álgebra lineal y ecuaciones diferenciales; y las horas totales de beneficio del programa SEA recibidas directamente en cada asignatura.

La principal conclusión del estudio es que no hay suficiente evidencia estadística para atribuir un impacto significativo del programa SEA en el rendimiento de las asignaturas matemáticas de los estudiantes de ciencias e ingeniería entre 2014 y 2016. Fue solo hasta 2017 que se encontró un impacto significativo en las asignaturas de cálculo y álgebra lineal, siempre que la cantidad de horas de beneficio exceda las 24 horas en el semestre.

El estudio brinda una metodología apropiada y objetiva para evaluar el impacto del programa SEA en el desempeño académico de los estudiantes el cual se espera que sirva como referencia para estudios similares y para un análisis costo-beneficio y costo efectividad del programa SEA que permita orientar mejor los recursos del programa.

---

\* Trabajo de Grado

\*\* Facultad de Ciencias, Escuela de Matemáticas. Director: Gabriel Yañez Canal. Doctor en Ciencias Especialidad en Matematica Educativa.

### Abstract

**Title:** Impact Evaluation of the Support System For Excellence (SEA) in the Mathematics Performance of Science and Engineering Students, Between 2014 y 2017.\*

**Author:** Ramos, Raúl Gabriel; Osorio, Juan Pablo.\*\*

**Keywords:** Academic performance, Impact evaluation, matching, difference-in-difference, SEA, PAMRA, MIDAS.

#### **Description:**

In this project was performed a statistical evaluation of the impact of the SEA program on the grades mathematic subjects for the science and engineering students. The scope was limited for students entered to the UIS University in the first cohort of each year between 2014 and 2017.

The study was conducted using matching and differences-in-differences techniques taking into account the following variables: the general characteristics of the students, such as age, gender and the academic program; the results of the SABER 11 test; the SEA program characterization in the social, health, economic, academic and cognitive dimensions; the grades of mathematic subjects that includes Calculation I, II and III, Linear Algebra and Differential Equations; and the total hours of benefit of the SEA program received directly to each subject.

The main conclusion of the study is that there is not enough statistical evidence to attribute a significant impact of the SEA program on mathematic subjects performance of science and engineering students between 2014 y 2016. It was only until 2017 that a significant impact was found in Calculation I subject and Linear Algebra subject, as long as the number of hours of benefit exceeds 24 hours in the semester.

This study proposes an appropriate and objective methodology to evaluate the impact of the SEA program. Forward looking with an additional economic data is possible to perform a cost-benefit analysis that allows a better management investment of the economic funds for the SEA program.

---

\* Degree Work

\*\* Faculty of Sciences. School of Mathematics Director: Gabriel Yañez Canal. Doctor of Science Specialty in Educational Mathematics.

## Introducción

En la Universidad Industrial de Santander, UIS, se han realizado mediciones de desempeño académico con el propósito de tomar acciones correctivas para reducir la deserción y mejorar el desempeño académico de los estudiantes. Según el Sistema para la Prevención de la Deserción en las Instituciones de Educación Superior -SPADIES-, la mayor deserción se da en los primeros semestres en el denominado ciclo básico, el cual es común a las carreras de ciencias básicas e ingenierías. En particular, se ha encontrado que las asignaturas en área de matemáticas son las que más contribuyen con esta deserción.

El Ministerio de Educación Nacional (MEN) en convenio con las diferentes instituciones de educación superior, han invertido considerables recursos para reducir la deserción estudiantil a nivel universitario. Particularmente, en el año 2012 se suscribió el Convenio 901 entre la UIS y el MEN, el cual dio origen al Sistema de Apoyo a la Excelencia Académica (SEA) en 2013 en el que confluyen todos los programas –como MANSA, PAMRA, MIDAS, FCP y ASAE- que anteriormente funcionaban de manera independiente para reducir la deserción estudiantil en la Universidad (Ministerio de Educación Nacional, 2015)

A la fecha, los beneficios del programa SEA se han medido con base en análisis comparativos del desempeño académico histórico de los estudiantes semestre a semestre, tomando indicadores como el número de estudiantes que quedan en estado de condicionalidad o PFU (Por Fuera de la Universidad), el número de estudiantes a los que se otorga beneficio, el número de veces que

cursan una asignatura para aprobarla, el promedio en asignaturas específicas y el promedio semestral.

Este estudio propone aplicar una estrategia metodológica que permita evaluar con rigor estadístico el impacto del SEA en el desempeño académico de los estudiantes, teniendo en cuenta su evolución en la universidad entre los años 2014 y 2017. Aunque es limitado a las asignaturas de matemáticas del ciclo básico y a los estudiantes de ciencias e ingeniería, por su peso en el problema de la deserción global UIS, sus hallazgos pueden ser base de decisiones más eficiente en los esfuerzos de la Universidad en el ámbito general de la deserción y el rendimiento académico. La UIS y el programa SEA brindaron su apoyo suministrando la información de base de datos de información general de ingreso, rendimiento académico y de beneficio histórico del programa de los estudiantes como insumo fundamental para llevar a cabo este estudio.

El foco obedece principalmente a la alta mortalidad académica ligada a las materias de matemáticas de asignación común a una gran parte de las carreras de la UIS. Sin embargo, se considera que los principios y técnicas empleadas se pueden extenderse más allá de este trabajo, hacia otras asignaturas y/o variables de desempeño académico como el promedio general y la condicionalidad (1ª, 2ª vez y PFU) del estudiante.

Dado que no se han realizado estudios rigurosos desde el punto de vista estadístico, el presente trabajo tiene como justificación fundamental suplir esta falencia y brindar una metodología robusta y un punto de apoyo objetivo a la UIS y al SEA, para realizar ajustes en pro de la eficiencia y

mejor impacto de los recursos invertidos en la importante labor de mejorar los niveles de desempeño de los estudiantes y reducir la mortalidad académica.

El primer capítulo de esta tesis presenta los antecedentes de deserción académica y del programa SEA junto con un resumen de la técnicas de evaluación de impacto que se emplearon en el presente estudio; el capítulo 2 muestra el análisis descriptivo de los datos empleados y el uso de la técnica de pareamiento y diferencias en diferencias para evaluar el impacto del programa SEA sobre desempeño académico de estudiantes de Ciencias e Ingeniería que ingresaron a la UIS en las asignaturas de matemáticas cálculos (I, II y III), álgebra lineal y ecuaciones diferenciales. Finalmente, se presentan las conclusiones y recomendaciones para trabajos futuros y un grupo de Anexos con los script en R para los diferentes análisis los cuales puede servir como guía para futuros estudios relacionados.

## **1. Antecedentes**

### **1.1 Estado Nacional de la Deserción Académica**

La deserción de estudiantes en la educación superior ha sido una problemática ampliamente estudiada con el objetivo de formular planes, programas y estrategias que buscan asegurar la permanencia dentro del sistema (Ministerio de Educación Nacional, 2016). En este sentido, el CINDA (Centro Interuniversitario de Desarrollo, 2006) y más recientemente el (Ministerio de Educación Nacional, 2015) han realizado estudios comparativos de la deserción universitaria en diferentes países de América Latina y Europa. En su estudio la deserción es medida con dos tipos de indicadores: las tasas de deserción anual y por cohorte. La tasa de deserción anual mide el porcentaje de estudiantes que estaban matriculados un año y figuran como desertores un año después. En este indicador, Colombia registra una tasa de 10,3% al cierre de 2014, superior a la del Reino Unido, (8,6%) e inferior a la de Brasil (18%) y Estados Unidos (18.3%). La tasa de deserción por cohorte muestra el porcentaje de no culminación de estudios del total de los estudiantes que ingresan a algún programa universitario. En este indicador, Colombia alcanza una cifra de 45,8%, lo que la deja en una posición intermedia entre México con 42%, Argentina con 43%, Venezuela con 52% y Chile con 54%.

A nivel nacional, el Ministerio de Educación Nacional (MEN) ha desarrollado el Sistema para la Prevención de la Deserción en las Instituciones de Educación Superior -SPADIES- el cual

consolida y ordena información que permite hacer seguimiento a las condiciones académicas y socioeconómicas de los estudiantes que han ingresado a la educación superior en el país.

En este sentido, el MEN, con el uso del SPADIES, ha identificado los factores que más afectan la tasa de deserción e identificado las probabilidades de deserción de un estudiante matriculado en un programa de educación superior. Así por ejemplo, en el reporte de mayo 13 del 2015 (MEN, 2015), el MEN propone una serie de líneas de acción orientadas a mejorar las competencias de los estudiantes para el ingreso a la Educación Superior por medio de programas de transición y una mejor articulación entre la educación media y la superior; actualizar las prácticas pedagógicas de los docentes, brindar apoyo académico y socio-afectivo, proporcionar subsidios de sostenimiento y oportunidades de generación de ingreso para los estudiantes de más bajos recursos y profundizar en un programa masivo de orientación socio-ocupacional.

En un reporte más reciente del MEN (Ministerio de Educación Nacional, 2016) se muestran los factores más importantes que afectan la deserción en la educación superior, destacando los siguientes: el estrato, sexo, nivel educativo de los padres, ingresos económicos de la familia del estudiante, nivel de clasificación del núcleo familiar según el SISBÉN, número de personas que componen el núcleo familiar, si el joven trabajaba al momento de presentar las pruebas de estado, la clasificación según los resultados de las pruebas de estado Saber 11° y la edad de presentación de las pruebas Saber 11°. Este estudio también muestra, que a partir del 2015 se han dado notables diferencias en la tasa de deserción alcanzadas por los estudiantes de acuerdo al nivel de ingresos de las familias y que dichas diferencias se amplían cuando se discrimina la población de acuerdo con el puntaje obtenido en las pruebas Saber 11.

## 1.2 Programa Sea

Los primeros esfuerzos de la Universidad Industrial de Santander dirigidos a atender el problema de la deserción fueron liderados por Bienestar Universitario y enfocados a asuntos de salud, alimentación, becas, préstamos y consejería. Si bien estas acciones favorecían la permanencia, no obedecían a un plan integral e institucional y no tenían en cuenta la dimensión académica. Uno de los primeros programas dirigidos a enfrentar la deserción en la Institución fue el Programa de Asesoría para el Mejoramiento del Rendimiento Académico-PAMRA, dirigido por Bienestar Universitario. Este proyecto, cuyo propósito es superar las dificultades académicas del estudiante, consiste en un acompañamiento continuo por parte de un tutor par (Cardozo Ortiz, 2010), (Torrado Arenas, Manrique Hernández, & Ayala Pimentel, 2009) y (Universidad Industrial de Santander, 2014)

Posteriormente, la Vicerrectoría académica comenzó a implementar otra serie de iniciativas orientadas a la permanencia. Uno de estos programas es MIDAS modelo de intervención académica para apoyar específicamente a los estudiantes de los programas de ingenierías y ciencias, los cuales tienen los mayores niveles de deserción (Ministerio de Educación Nacional, 2015). La estrategia se basó en un acompañamiento de tutores y monitores bajo el enfoque de aprendizaje colaborativo y el aprender a aprender (Universidad Industrial de Santander SEA-UIS, 2014). El programa para el Fortalecimiento Pedagógico Cognitivo-FPC es otro de los programas de acompañamiento que actualmente funciona en la UIS. Fue concebido para ofrecer acompañamiento sicopedagógico a los estudiantes que así lo requieren, con el fin de fortalecer

procesos cognitivos y de adaptación a la vida universitaria, orientación socio-ocupacional, así como técnicas y hábitos de estudio (Universidad Industrial de Santander SEA-UIS, 2014).

Adicionalmente, la Escuela de Matemáticas ofrece un programa de Atención, Seguimiento y Acompañamiento-ASAE, para aquellos estudiantes que cursan asignaturas de matemáticas y complementan la oferta de apoyos académicas que ofrece la Universidad.

Por último, Bienestar Universitario creó el programa de Mantenimiento de la Salud-MANSA, con una perspectiva de la prevención y el autocuidado (Botello y Parada R., 2013), (Universidad Industrial de Santander SEA-UIS, 2014), (Ministerio de Educación Nacional, 2015)

Aunque estas diferentes iniciativas atendían las múltiples dimensiones del estudiante: académica, psicosocial y cognitiva; lo hacían de forma aisladas y no estaban articuladas a una estrategia integral y por consiguiente no tenían el impacto deseado (Ministerio de Educación Nacional, 2015). Para ello en el año 2012 se suscribió el Convenio 901 entre la UIS y el MEN con el objetivo de aunar esfuerzos para fomentar la permanencia y graduación estudiantiles mediante la ejecución del proyecto intervención integral con énfasis en el factor académico para disminuir la deserción académica, el cual apunta a fortalecer la capacidad de la UIS en el diseño, ejecución y evaluación de políticas y programas de fomento de la permanencia y graduación estudiantiles (Ministerio de Educación Nacional, 2012). Uno de los grandes logros alcanzados en el contexto del convenio, fue la creación del Sistema de Apoyo a la Excelencia Académica (SEA) en 2013, en el cual confluyen todos los programas anteriormente mencionados como MANSA, PAMRA, MIDAS, FCP y ASAE-, con el propósito de lograr una actuación integral y abarcar todas las

dimensiones que inciden en la permanencia del estudiante (Ministerio de Educación Nacional, 2015).

Desde el primer periodo académico de 2013, el SEA ha venido realizando la caracterización de los estudiantes que son admitidos en la universidad, lo cual permite identificar aquellos que tienen algún tipo de vulnerabilidad en las dimensiones social, biosicosocial, económica, académica y cognitiva, que pueda llegar a afectar su desempeño académico. Los estudiantes identificados con algún tipo de vulnerabilidad son invitados a participar de manera voluntaria en los diferentes programas de acompañamiento que ofrece el programa SEA. Con el mismo propósito, el SEA integra la información de los estudiantes de las diferentes dependencias de la universidad incluyendo bienestar Universitario, vicerrectoría académica, admisiones y registro, división de servicios de información y las diferentes facultades, para hacer un seguimiento continuo y completo de los estudiantes, generándose indicadores claves semestrales de desempeño como la aprobación, repitencia, cancelación y pérdida de asignaturas y promedio académico.

Con base en toda la información descrita, el SEA planea y formula estrategias dirigidas a tres momentos de la vida universitaria: antes del Ingreso, en el inicio de la vida universitaria y en el desarrollo y permanencia del estudiantes, las cuales se consolidan en acciones a través de sus subprogramas integrantes que intervienen las distintas dimensiones evaluadas: Académica (PAMRA, MIDAS, ASAE), Biosicocial (MANSA), Cognitiva (FPC) y Socioeconómica (Bienestar Universitario).

Los beneficios del programa se han medido por medio del seguimiento una serie de indicadores globales de desempeño académico como: el número de estudiantes que quedan en estado de condicionalidad (condicional 1ª o 2ª vez), número de estudiantes por fuera de la universidad (PFU) y número de veces que un estudiante tiene que cursar la misma asignatura para aprobarla. Sin embargo, estos indicadores engloban los efectos de todas las variables que pueden influir en el desempeño académico y por lo tanto no pueden ser interpretados como causales del impacto únicamente del programa SEA.

### **1.3 Estudios Sobre Deserción y Rendimiento Académico**

En la literatura se encuentran numerosos trabajos para evaluar los factores que más afectan el rendimiento académico a nivel universitario empleando técnicas estadísticas.

(García Jiménez y cols, 2000), realizan un estudio en el que comparan las técnicas de regresión lineal múltiple y regresión logística para estimar el rendimiento de estudiantes del primer semestre de psicología, en la asignatura de “Métodos y Diseños de Investigación”. Concluyen que la regresión múltiple no resulta un buen método de pronóstico del rendimiento académico, mostrando valores muy bajos de correlación ( $R^2=0.13$ ). Mientras la regresión logística realiza acertadamente clasificaciones éxito/fracaso en 70% de los casos. Concluyen además, que el factor más preponderante en el rendimiento académico futuro, es el rendimiento académico anterior. Otros buenos predictores resultan ser la motivación al iniciar la carrera y la mayor asistencia y participación activa en las clases.

(Mazza, 2007), encuentra un indicador del desempeño académico estudiantil basado en los créditos académicos, empleando un modelado estadístico multivariado a partir de variables sociodemográficas, académicas y motivacionales. Se construyen y comparan

modelos de regresión lineal y logística, así como modelos de ecuaciones estructurales. Los resultados muestran que los modelos de regresión lineal y logística, pierden potencia para detectar efectos significativos al categorizar las variables cuantitativas. Por otra parte, los modelos de ecuaciones estructurales logran contrarrestar este problema por medio de agrupaciones por ítems denominadas variables latentes.

(Miñano Pérez, 2009), presenta una revisión de varios modelos estructurales propuestos por diferentes autores para predecir el rendimiento académico y a su vez propone un modelo basado en regresión lineal y ecuaciones estructurales para evaluar el efecto de variables cognitivo-motivacionales en el desempeño académico de los alumnos.

(Olaya, Mosquera M, & Artamonova, 2009), realizan un estudio sobre los factores que influyen de manera significativa en el rendimiento académico de los estudiantes de ingenierías y tecnologías de la Universidad Tecnológica de Pereira en la asignatura Matemáticas I, en función de los resultados de las pruebas de ICFES empleando regresión logística. Este estudio confirma el buen resultado de los modelos logit, con predicciones correctas al 70%, encontrando entre los factores más influyentes, programa académico, puntaje global del ICFES, nivel de lectura y pensamiento lógico abstracto. Por otra parte factores como edad, género, estrato y tipo de colegio, riesgo de salud física y mental, fueron descartados.

(Porcel, Dapozo, & López, 2011), hacen uso de la regresión logística para estimar el rendimiento de los alumnos ingresantes a la Facultad de Ciencias Exactas y Naturales durante el primer año de carrera con las características socioeducativas de los mismos. El rendimiento fue medido por la aprobación de los exámenes parciales o finales de la primera materia de Matemática que los alumnos cursan. Entre las variables más relevantes para explicar el rendimiento académico se encuentran el título secundario obtenido, la carrera elegida y el nivel educacional alcanzado por la madre. Los mismos autores (Porcel & cols., 2011) realizan un estudio similar empleando redes neuronales artificiales (RNA) encontrando que las RNA predicen mejor del desempeño académico que la regresión logística para el mismo conjunto de datos.

(Vázquez, y otros, 2012), en un estudio de opinión, muestran los factores de mayor impacto en el rendimiento académico universitario desde la percepción de los estudiantes.

(Rico, Rodríguez Hernández y Vilalta Alonso, 2012), usan regresión logística para construir una ecuación que relaciona la calificación del estudiante en la asignatura Modelos Probabilísticos, la cual se imparte en segundo año de la carrera de Ingeniería Industrial, con sus resultados en primer año.

(Corengia, Pita, y Centeno, 2012), usan regresión múltiple y logística para predecir el éxito académico de los estudiantes de primer año de grado de una universidad privada de la Argentina.

(Mendoza Mendoza y Herrera Acosta, 2013), realizan una propuesta para predecir el rendimiento académico de los estudiantes de la Universidad del Atlántico, basado en la aplicación del análisis discriminante. Logrando una clasificación acertada del 72%, similar a la mostrada en trabajos anteriores por modelos logit.

(Cendales, Vargas-Trujillo, y Barbosa, 2013), hacen uso del análisis de conglomerados, análisis de varianza y covarianza mostrando que el perfil auto-perceptivo/actitudinal de los estudiantes tiene un efecto significativo en el rendimiento académico, y que este a su vez se encuentra influenciado por el género y programa académico.

(Osma Castellanos, Mojica Perdomo y Rivera Flórez, 2014), analizaron los resultados de las pruebas Saber Pro 2010 de los estudiantes de últimos semestres de Ingeniería Civil a nivel nacional para determinar los factores que más influyen en los puntajes obtenidos. Se estudiaron diferentes variables cualitativas y cuantitativas asociadas al puntaje obtenido mediante tablas de contingencia y análisis de correspondencia simple y múltiple. Se concluye que factores como educación de los padres y estrato socioeconómico, son un factor de gran influencia en resultado de las pruebas, mientras género, tipo de bachillerato y trabajo durante los estudios no son relevantes.

(Perez Pulido y Roa Fuentes, Hernan Arturo, 2008), hacen uso del análisis por componentes principales (PCA) y tablas de contingencia para encontrar las relaciones entre los resultados de las pruebas del ICFES y el desempeño de los estudiantes en el primer semestre de ingenierías. Con PCA se encontró que los resultados del primer semestre de ciclo básico son independientes de los puntajes del ICFES en matemáticas, física y lenguaje; mientras que, empleando tablas de contingencia, se encontró una relación entre los puntajes del ICFES en matemáticas con el rendimiento de cálculo I. Más allá de la contradicción con las dos técnicas estadísticas, los autores destacan que los resultados del ICFES en matemáticas y cálculo I están relacionados con el pensamiento matemático del individuo. Por lo que el resultado obtenido con tablas de contingencia es razonable a pesar de lo encontrado empleando PCA.

(Socha Pedraza, 2009), realiza un estudio que permite conocer y evaluar motivaciones de la deserción por parte de los estudiantes de la UIS, estableciendo premisas y categorías de análisis que conducen hacia un conjunto de estrategias para contribuir a reducir los niveles de deserción.

Con respecto al uso de algunos de los programas del SEA se encontró el trabajo de (Cardozo Ortiz, 2010) en el que se muestra desde la perspectiva de los estudiantes la experiencia en el uso del programa PAMRA.

Sin embargo, todos los anteriores trabajos no analizan la relación de algún esfuerzo específico estructurado institucional, como programas o políticas universitarias o gubernamentales, sobre el desempeño académico, sino que se concentran en entender y modelar la influencia de las variables características previas o naturales de la población estudiantil.

## 2. Marco Teórico

### 2.1 Conceptos de Estadística

**2.1.1 Prueba de hipótesis para la media, la proporción y valor p.** En general, una prueba de hipótesis es un supuesto o conjetura realizada sobre algún parámetro de la población de interés. Una prueba de hipótesis examina dos hipótesis opuestas sobre una población: la hipótesis nula, usualmente denotada como  $H_0$  y la hipótesis alternativa, denotada como  $H_1$ . La hipótesis nula por lo general establece, que "no hay efecto" o "no hay diferencia" entre los parámetros de la población estudiada mientras que la hipótesis alterna establece lo contrario. Con base en los datos de una muestra, la prueba determina si se puede rechazar la hipótesis nula o si no existe evidencia suficiente en la muestra para rechazarla.

Una prueba de hipótesis puede conducir a dos conclusiones equivocadas denominados error tipo I y error de tipo II. El error de tipo I ocurre cuando se rechaza  $H_0$  siendo correcta y el error tipo II consiste en aceptar  $H_0$  siendo falsa. La probabilidad de cometer un error tipo I, usualmente se denota con la letra griega  $\alpha$  y llamada valor p, se puede estimar por el analista a partir de la distribución del estadístico de prueba, bajo la asunción que  $H_0$  es cierta; por esta razón, rechazar  $H_0$  siempre se puede considerar una **decisión fuerte**. Por otra parte, como la probabilidad de cometer un error de tipo II, usualmente denotado por  $\beta$ , depende del verdadero parámetro de la población en el que  $H_0$  es falsa y del tamaño de la muestra, es costumbre considerar la decisión de aceptar  $H_0$  como una **decisión débil**.

En las técnicas de evaluación de impacto se realiza la comparación de la media o la proporción de entre el grupo de individuos tratados y no tratados. Por esta razón detallaremos a continuación estas dos pruebas:

**2.1.2 Regresión logística.** La regresión logística es semejante a la regresión múltiple pero con la diferencia que en la regresión logística la variable dependiente asume dos valores: 0 y 1 que en el caso del modelo de propensión empleado en las técnica de pareamiento, corresponden a la participación o no en un programa y por lo tanto, la respuesta del modelo de regresión logístico se puede interpretar como la probabilidad de participar o no el programa.

Los parámetros del modelo de regresión logística se estiman por el método de máxima verosimilitud requiriendo métodos numéricos pero que afortunadamente ya se encuentran implementados en excelentes programas estadísticos como SAS, S y R. En el presente estudio se hace uso del modelo lineal generalizado incluido en el comando glm de R.

La prueba de hipótesis en regresión logística (y en general para el modelo lineal generalizado) se basa en pruebas de cocientes de máxima verosimilitud y se puede emplear para realizar pruebas en coeficientes individuales del modelo de regresión logística a partir de un estadístico que sigue una distribución normal estándar. Esta prueba se encuentra incluido en el resumen del modelo de regresión arrojado por diferentes programas estadísticos como R.

**2.1.3 Criterio de AIC.** El criterio de información de Akaike (AIC) es una medida de la calidad relativa de un modelo estadístico, para un conjunto dado de datos que y proporciona un medio

para la seleccionar un modelo. AIC maneja el compromiso entre la bondad de ajuste del modelo y su complejidad.

AIC no proporciona una prueba de un modelo en el sentido de probar una hipótesis nula, es decir AIC puede decir nada acerca de la calidad del modelo en un sentido absoluto. El AIC se emplea para determinar el subconjunto de variables que mejor explican la variable respuesta en un modelo con muchas variables independientes candidatas.

En el caso general, el AIC es:

$$AIC = 2k - 2 \ln(L)$$

Donde  $k$  es el número de parámetros en el modelo estadístico, y  $L$  es el máximo valor de la función de verosimilitud para el modelo estimado.

Dado un conjunto de modelos candidatos para los datos, el modelo preferido es el que tiene el valor mínimo en el AIC. Por lo tanto AIC no sólo recompensa la bondad de ajuste, sino también incluye una penalidad, que es una función creciente del número de parámetros estimados. Esta penalización desalienta el sobreajuste (aumentando el número de parámetros libres en el modelo mejora la bondad del ajuste, sin importar el número de parámetros libres en el proceso de generación de datos).

**2.1.4 Selección de variables y Regresión por pasos.** Existen varios métodos para construir el modelo de regresión y seleccionar de entre todas las variables que introducimos en el modelo,

cuáles son las que necesitamos para explicarlo. El modelo de regresión se puede construir utilizando las siguientes técnicas:

- Regresión hacia adelante (Forward): consiste en ir introduciendo las variables en el modelo únicamente si cumplen una serie de condiciones hasta que no se pueda introducir ninguna más, hasta que ninguna cumpla la condición impuesta.
- Regresión hacia atrás (Backward): se introducen en el modelo todas las variables y se van suprimiendo si cumplen una serie de condiciones definidas a priori hasta que no se pueden eliminar más, es decir ninguna variable cumpla la condición impuesta.
- Regresión por pasos (Stepwise): combina los dos métodos anteriores, adelante y atrás introduciendo o eliminando variables del modelo si cumplen una serie de condiciones definidas a priori hasta que ninguna variable satisfaga ninguna de las condiciones expuestas de entrada o salida del modelo.

La técnica empleada en el presente estudio para depurar el modelo de regresión logística fue la técnica de regresión por pasos basada en el estadístico AIC.

## **2.2 Técnicas de Evaluación del Impacto**

Comprenden un conjunto de técnicas que emplean principios estadísticos fundamentales, para separar o diferenciar el efecto causado por programas o políticas sociales implementadas en un momento y lugar determinado, de otras variables circunstanciales o naturales, que también pudieron influir en la medición del resultado a evaluar (Bernal y Peña, 2016). Han sido formuladas para asegurar que, desde la planeación misma del programa, es decir antes de su ejecución

(Evaluación Prospectiva), se asegure una evaluación no sesgada o completamente causal. Sin embargo, también pueden ser aplicadas para la evaluación de programas ya iniciados o finalizados (Evaluación Retrospectiva), pero se necesita especial cuidado en la identificación de los grupos de comparación que permitirán cuantificar el impacto, dado que estos no fueron previamente concebidos ni estructurados para este propósito. Este último es el caso del programa SEA, que requerirá de la identificación y aplicación de una metodología, capaz de generar dos grupos estadísticamente no diferenciables en sus características relevantes promedio, y únicamente diferenciables en que solo uno de ellos recibió los beneficios del programa.

La razón fundamental de este requisito reside en que teóricamente la medición del impacto de un programa sobre un individuo requeriría poder evaluar el efecto de dos situaciones simultáneas sobre el mismo individuo, la de haber participado en el programa y la de no haber participado en el programa (Gertler y cols, 2017). Como es lógico, lo anterior resulta imposible, ya que una vez un individuo entra en el programa, el efecto de no haberlo recibido ya no se puede medir, esta situación ideal requeriría de un clon del individuo. Es aquí donde aparece la ayuda de la estadística que se usa con el objetivo de estimar cual habría sido el efecto no sobre el individuo, sino sobre los participantes de un programa sino hubiesen participado en él, ese efecto es denominado contrafactual el cual se define como el resultado que hubieran tenido los beneficiarios del programa si no hubieran recibido el programa. No se puede medir, se estima a partir de un grupo de control que no recibió el tratamiento y todas las técnicas de evaluación de impacto buscan la estimación de un contrafactual válido.

Así la estadística, se enfoca en las características promedio de dos grupos de comparación extraídos de la población objetivo del programa para la evaluación del resultado, un grupo que

recibe el tratamiento y un grupo de comparación que no lo recibe. La identificación de estos grupos es la esencia misma de la evaluación de impacto, pues si se logra que entre ellos la única diferencia en sus características sea el programa, el contrafactual será válido y el efecto evaluado del programa será considerado causal (Gertler y cols, 2017).

Una vez encontrado un contrafactual válido, que resulta la tarea más crítica de la evaluación, se puede proceder a la evaluación numérica del impacto, que es esencialmente simple. Se realiza calculando la diferencia del resultado objetivo promedio medido,  $E[Y_i]$  entre el grupo de tratamiento (T) y el grupo comparación (C), como se ilustra en la Figura 1. Nótese que la proporción de individuos de cada tipo de gorro (que representa las características de cada individuo) en los dos grupos es la misma, lo que ilustra que tienen en promedio las mismas características y por lo tanto el resultado de la evaluación es solo atribuible al efecto programa. En este trabajo, la medición del impacto del programa SEA, será resultado por ejemplo de la diferencia de la nota promedio en la asignatura entre el grupo de estudiantes que recibió del beneficio y el grupo de estudiantes que no lo recibió.

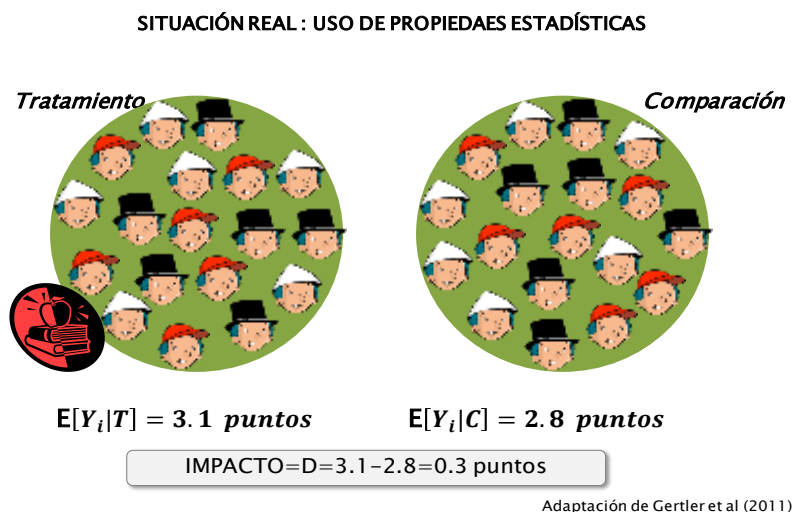


Figura 1. Evaluación numérica del impacto. Adaptado de: Gertler y cols., 2011.

Como se ilustra en la Figura 2, el procedimiento natural para obtener dos grupos de características no estadísticamente diferenciables, sería la selección aleatoria de los individuos a cada grupo, provenientes de una muestra de la población objetivo del programa; esto asegura la “validez interna” de la evaluación de impacto. A su vez la muestra puede ser seleccionada aleatoriamente de la población objetivo del programa, lo que además le daría “validez externa” a la evaluación de impacto; en este caso se pueden realizar generalizaciones de los resultados del impacto evaluado de la muestra a la población.

Aunque la aleatorización es la mejor solución para obtener un contrafactual válido, no es fácilmente aplicable, por restricciones en las características y naturaleza de los programas y políticas sociales. La mayoría de programas tiene límites económicos, éticos, políticos, logísticos e incluso de respeto a la voluntad o derecho de cada individuo de participar o no del programa, que hacen que la aleatoriedad no se pueda dar en la mayoría de los casos (Bernal y Peña, 2016)

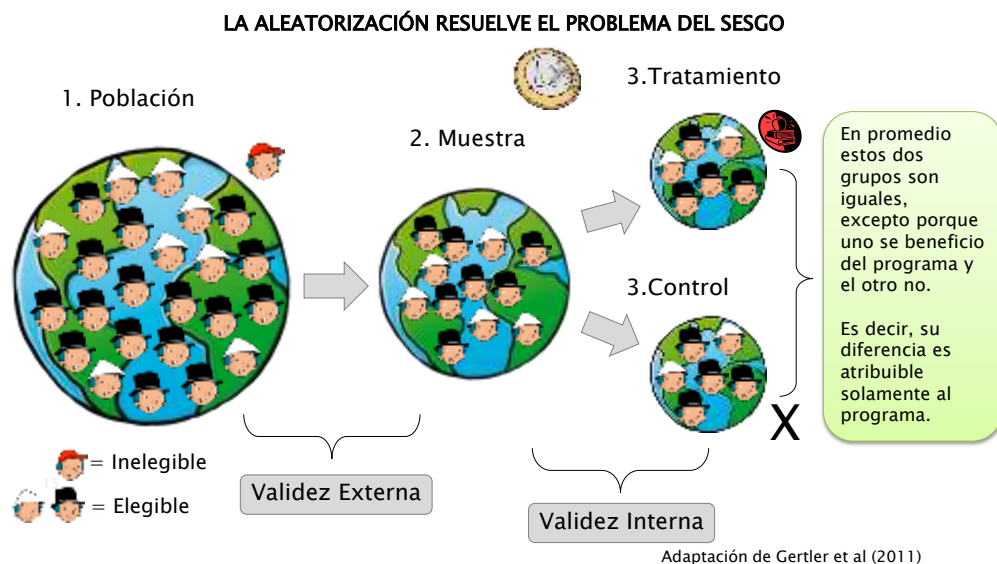


Figura 2. Selección aleatoria de los grupos de contraste. Adaptado de: Gertler y cols., 2011.

El caso del programa SEA de la UIS no es una excepción, por su complejidad y dinámica la asignación de los individuos al programa no es producto de un proceso aleatorio, sino de una situación de iniciativa misma del estudiante con una valoración previa del riesgo por parte del SEA, dado que es de carácter completamente voluntario e incluyente a toda la población universitaria.

De hecho, la forma de asignación a un programa de los individuos elegibles de una población, es tan importante que define los métodos de evaluación de impacto desarrollados con el propósito de encontrar un contrafactual válido; a continuación se describen las metodologías de evaluación de impacto principales y sus características generales de acuerdo a las descripciones de Gertler y cols. (2017).

**Asignación Aleatoria** (Gertler y cols, 2017). Se usa cuando el número de participantes potenciales es superior al número de plazas disponibles en el programa o se desea realizar una evaluación a escala menor con potencial de ampliación a toda una población. Existen dos posibilidades en este caso, la asignación aleatoria de los participantes o la oferta aleatoria. Esta última se realiza teniendo en cuenta la posible dificultad en cumplir cabalmente con la asignación aleatoria de los individuos, dada su libre capacidad de decisión de participación en el programa; se resuelve entonces seleccionando aleatoriamente un grupo de individuos, al cual se le ofrecerá de forma estructurada el programa, mientras que el grupo de complementario de esta asignación, no tendrá oferta alguna del programa.

**Regresión Discontinua** (Gertler y cols, 2017). Utiliza un índice continuo de elegibilidad, que define un umbral mínimo para decidir quienes tienen derecho a participar del programa y quiénes no. Este método tiene la desventaja de que el impacto evaluado

solo es concluyente en la cercanía del umbral mínimo del índice de elegibilidad seleccionado, punto en cual se calcula el impacto de la medición objetivo (Figura 43).

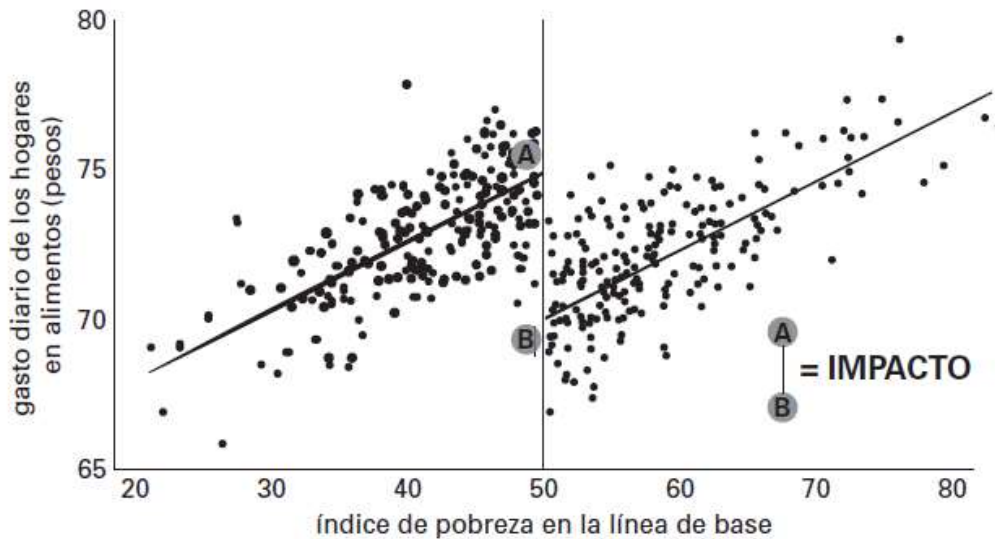


Figura 3. Ejemplo de aplicación método de regresión discontinua. Adaptado de: Gertler y cols., 2011.

**Diferencia en diferencias** (Gertler y cols, 2017): Se basa fundamentalmente en la existencia de un antes y un después del programa a evaluar, lo que significa que requiere de una medición previa de la valoración objetivo del programa o línea base. Su principal asunción es que las características entre los grupos de tratamiento y comparación son constantes en el intervalo de tiempo que comprenden el antes y el después de la aplicación del programa. Se asume entonces que, si el programa no hubiera sido implementado, se obtendría un comportamiento con la misma tendencia en diferencias de la línea base. En la Figura 4 se ilustra un ejemplo hipotético de aplicación de esta metodología al efecto del programa SEA en materias vistas de forma repetida, considerando la primera vez como la línea base de la evaluación; la línea punteada verde representa la tendencia que el grupo de tratamiento repetente de la materia (Con Beneficio) hubiera seguido de no haber participado en el programa, de forma que el impacto se evalúa como el incremento logrado respecto a

esa continuidad, y es calculado con una doble diferencia como se ilustra en la siguiente figura.

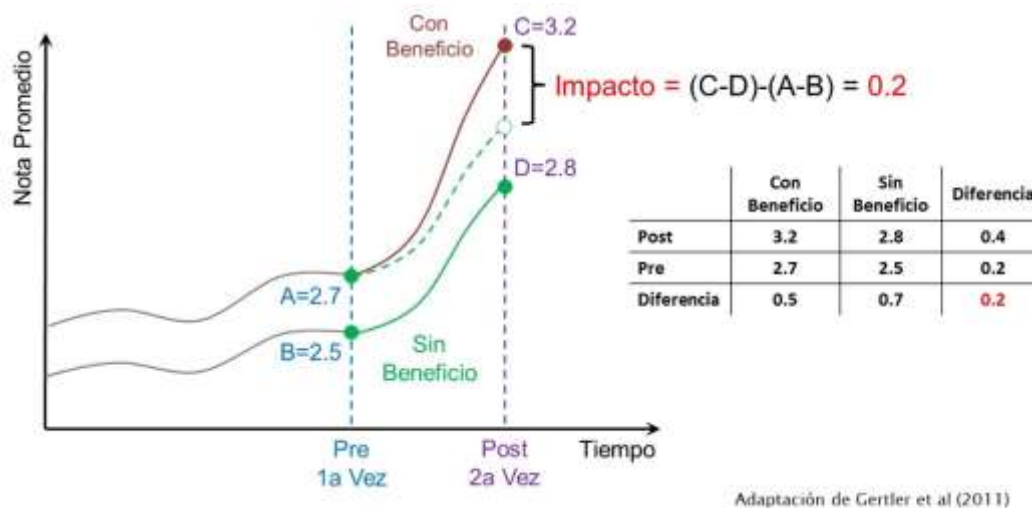


Figura 4. Ejemplo de aplicación de método de diferencia en diferencias. Adaptado de: Gertler y cols., 2011.

**Pareamiento** (Gertler y cols 2017). Es la técnica más aplicable a programas cuyas reglas de asignación no son claras y por lo tanto no conducen a un contrafactual válido. Se basan en las técnicas estadísticas de emparejamiento (matching), que usan las características observables y medibles de la población objetivo del programa, para construir un grupo de comparación artificial. Básicamente a cada individuo del grupo de tratamiento se le asigna uno o más individuos que no participaron en el programa, teniendo en cuenta que sus características sean tan similares como sea posible.

Esta metodología puede llevar a dos situaciones opuestas y extremas, por la denominada “maldición de las dimensiones”. Por un lado, si el número de características contempladas para realizar el pareamiento es muy pequeño, es muy probable que se consiga fácilmente asignar una pareja a cada individuo tratado, pero se corre el riesgo de excluir otras características importantes que lo definen como participante, en este caso la pareja asignada es solo aparentemente similar, pero realmente no lo es, dada la exclusión mencionada. Por

otra parte, si el número de características contempladas para realizar el pareamiento es muy grande se corre el riesgo de no encontrar una pareja factible para cada individuo tratado, a menos que se cuente con un gran número de mediciones. Lo anterior puede resolverse usando un modelo que condense en una única dimensión la probabilidad de participar o no del programa, denominado “modelo de propensión” (generalmente un modelo de regresión logit binomial), (Rosenbaum, 2010)<sup>1</sup>; de esta forma, el par del individuo tratado será aquel que tenga el puntaje de propensión más próximo (usando por ejemplo el método Nearest Neighbor Search).

Además, el modelo de propensión puede ser “optimizado” buscando el menor número de variables que conduzcan al mejor modelo aplicando criterios de calidad, como en el estadístico de Akaike (AIC). Realizada la construcción del modelo y el pareamiento, el impacto será calculado como la diferencia del promedio de la medición objetivo del programa entre el grupo “real” de tratamiento y el grupo “artificial” de comparación.

Por último la técnica de pareamiento, debe incluir todas las características que pueden llegar a influir en el resultado. En la medida en que características relevantes queden por fuera de la construcción del modelo de propensión, el resultado del impacto evaluado puede resultar sesgado (Bernal y Peña, 2016).

En la sección 0 se brinda una descripción más detallada de la técnica de pareamiento.

En la Figura 5 se resumen las principales técnicas de evaluación de impacto, su requerimiento fundamental y el potencial de aplicación para la evaluación de la evaluación de impacto del programa SEA.

---

<sup>1</sup> Rosenbaum, P.R. (2010). Design of Observational Studies. Springer Series in Statistics. Springer.

Metodología	Requisito	Potencial de evaluación del SEA
Modelo de diferencias	Requiere de aleatoriedad, se evalúa por diferencia simple.	<b>NO.</b> Los beneficiarios del programa SEA no se escogieron aleatoriamente
Diferencias en diferencias	Requiere de una línea base, antes y después del tratamiento.	<b>SI</b> Se evaluará si es posible aplicarla para el caso de los repitentes en cada materia.
Pareamiento	Cuando lo único que importa es lo observable y esta información la tenemos en la base de datos.	<b>SI.</b> Resulta la técnica más indicada para el SEA, dado que identifica un individuo "clon" no beneficiario a cada individuo beneficiado basado en variables observables y disponibles, suponiendo que están son todas o por lo menos las más influyentes en desempeño académico.
Variables instrumentales	Una variable que explica la participación en el programa pero no con la variable de resultados.	<b>NO.</b> Se basa en asignación o promoción aleatoria que no se uso en el programa SEA
Regresión discontinua	Cuando se tiene un umbral continuo que determina la participación del programa.	<b>NO.</b> El SEA no utilizó un único criterio de participación y a su vez este no era absoluto (ej. nivel de riesgo académico)

Figura 5. Técnicas de evaluación de impacto y potencial para el caso SEA.

### 2.3 Método de Pareamiento

La técnica de pareamiento estima el efecto o impacto de un tratamiento, política o intervención teniendo en cuenta variables observables que afectan la decisión de participar o no en el programa. La técnica busca, a cada individuo tratado, uno no tratado que tengan características similares reduciendo así el sesgo de selección y obteniendo un mejor estimando del contrafactual. Se emplea cuando el programa que se intenta evaluar no tiene reglas de asignación claras que explique por qué ciertos individuos se inscribieron en el programa y otros no lo hicieron. La técnica de pareamiento permite identificar el conjunto de individuos que no recibieron tratamiento que más se parece a los si recibieron, a partir de las características observables que ya se tienen. Estos individuos pareados posteriormente se emplean para estimar el contrafactual y evaluar el impacto del programa.

El pareamiento se realiza encontrando un “clon” de cada individuo tratado en el grupo de control (no tratado) y contrastar las variables de resultado de ambos. Donde clon se refiere a un individuo o grupo de individuos con variables observables similares como edad, sexo, educación o cualquier otra variable pertinente para el programa bajo estudio. Sin embargo, la implementación de la técnica se vuelve computacionalmente compleja de implementar si hay un gran número de variables observables con base en las cuales se realiza el pareamiento, es decir, si la dimensión de vector de variables observables,  $X$  es muy grande. A esto se le conoce como maldición de dimensionalidad (Bernal y Peña, 2016) y (Gertler y cols, 2017). En estos casos, en vez de emparejar individuos con base en un vector de características observadas  $X$ , se emparejan con base en su probabilidad estimada de participar en el programa, dadas sus características observadas  $P(X)$ , donde:

$$P(X) = P(\text{Tratado}|X)$$

Es decir, el clon para cada individuo del grupo de tratamiento será aquel del grupo de control con una probabilidad de participación más cercana. La ventaja de emparejar por  $P(X)$  en lugar de  $X$ , es que  $P(X)$  es un escalar, el pareamiento es computacionalmente más sencillo y genera estimadores consistentes del efecto del programa (Bernal y Peña, 2016). A la función  $P(X)$  se le llama puntaje de propensión (Propensity Score) y a la técnica de pareamiento por puntajes de propensión se le conoce como PSM por sus siglas en inglés (Propensity Score Matching).

Una vez se estima el puntaje de propensión de todos los individuos, el grupo de tratamiento pueden parearse con unidades en el grupo de no tratados que tienen los puntajes de propensión más cercanos. Estas unidades próximas se convierten en el grupo de comparación y se utilizan

para producir una estimación del contrafactual. El método de pareamiento por puntajes de propensión intenta imitar la asignación aleatoria a los grupos de tratamiento y de control escogiendo para el grupo de comparación aquellas unidades que tienen propensiones similares a las unidades del grupo de tratamiento.

Para que el pareamiento por puntajes de propensión produzca estimaciones del impacto de un programa para todas las observaciones tratadas, cada individuo tratado debe parearse con uno que no fue tratado. Sin embargo, en la práctica puede ocurrir que, para algunos individuos tratados, no haya individuos en el conjunto de no tratados que tengan puntajes de propensión similares. En términos técnicos, puede que se produzca una falta de rango común, o falta de superposición, entre los puntajes de propensión del grupo de tratamiento o inscrito y los del conjunto de no inscritos. Una forma de determinar si existe o no rango común es dibujando la función de densidad de probabilidad de puntajes de propensión para participantes y no participantes del programa como se muestra en la Figura 6.

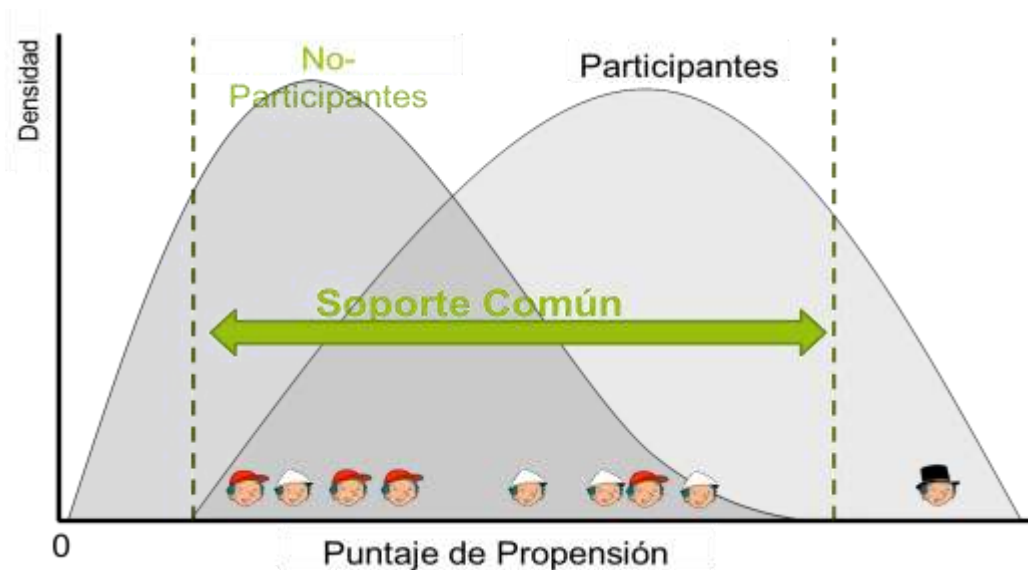


Figura 6. Rango común de la propensión. Adaptado de: Gertler y cols., 2011.

Se aprecia con gorro rojo los no participantes y con blanco los individuos que participaron del programa. También se aprecia un participante del programa con sombrero negro que no se puede aparear con ningún individuo no tratado debido a que presentan un puntaje de propensión muy alto y no se encuentra ningún individuo en el grupo de no participantes con puntaje de propensión similar. En estos casos se dice que encuentra por fuera del rango común de propensión

Según (Bernal y Peña, 2016) la técnica de PSM se puede resumir en el siguiente algoritmo:

1. Con base en las características observables para individuos tratados y no tratados estimar los parámetros del modelo de propensión. En este estudio se empleo el modelo de regresión logística que estima la probabilidad,  $P(X)$  de participar en el programa con base en las características observables,  $X$ . En el presente estudio se aplicó el algoritmo de regresión por pasos basado en el criterio de AIC para encontrar el mejor modelo de propensión. Estos conceptos se explican en la sección anterior.
2. Estimar la probabilidad de participar en el programa usando el modelo de propensión anterior tanto para individuos tratados como no tratados
3. Restringir la muestra al soporte común.
4. Realizar el pareamiento. Para cada individuo tratado se busca un individuo o grupo de individuos en el grupo de control que tenga una probabilidad de participar similar. En este estudio se empleo el método de los vecinos más cercanos.
5. Revisar que después del pareamiento, las variables observadas entre el grupo de tratamiento y de control sean similares o estén balanceadas. Paa esto se pueden comparara la distribuciones de los puntajes de propensión antes y después de pareamiento.

6. Estimar el impacto de programa como el promedio de la diferencia entre el la variable respuesta de los tratados y no tratados
7. Calcular los errores estándar e intervalos de confianza para determinar si el impacto del programa es estadísticamente significativo.

### **3. Resultados y Discusión**

#### **3.1 Análisis Descriptivo**

El análisis descriptivo que se presenta tiene dos momentos: un primer análisis relacionado con el seguimiento a los estudiantes de ingeniería y ciencias que ingresaron en la primera cohorte del 2014 durante sus primeros 6 semestres, que permite describir la dinámica de los mismos estudiantes en su paso por la universidad; y un segundo análisis relacionado con el seguimiento a los estudiantes de primer semestre de la primera cohorte de cada año entre el 2014 y 2017, el cual permite describir la dinámica de la universidad en los estudiantes recién ingresados. La no inclusión de los estudiantes de la segunda cohorte de cada año analizado, obedece únicamente a un equilibrio entre el trabajo de análisis requerido y una cobertura temporal suficientemente amplia de este trabajo. Solo se incluyó la comparación de las características generales para la primera y segunda cohorte del 2014, en la cual se observará que no existen diferencias marcadas, por lo menos para este año entre estas cohortes.

El análisis se centra en la descripción del programa SEA y su relación con el desempeño académico, a nivel general y específicamente en las asignaturas del área de matemáticas. Para esto, se cuenta con la información de los estudiantes que ingresaron a la universidad correspondientes a: información general de ingreso de los estudiantes, resultados de las pruebas SABER 11, promedio semestral y acumulado en los semestres de interés; notas finales en cálculos, álgebra y ecuaciones diferenciales; participación en las diferentes acciones de apoyo ofrecidas por la universidad e integradas en el programa SEA y, finalmente, la caracterización realizada por éste último para identificar vulnerabilidad académica en las dimensiones social, biopsicosocial, económica, académica y cognitiva. En la siguiente tabla se resume la forma como fue recibida la información para el desarrollo de la presente investigación.

Tabla 1.

*Resumen general de la de información.*

Base de datos	Descripción
Notas	Archivo MS Excel con las pruebas SABER 11, promedio semestral y acumulado, condicionalidad y notas en asignaturas de matemáticas (cálculo I, II y III, álgebra y ecuaciones diferenciales).
Beneficiarios SEA	Archivo MS Excel con la los estudiantes beneficiarios del programa SEA con el código del estudiante, sub programa empleado (ASAE, FPC, MIDAS y PAMRA) asignatura en la que recibió el beneficio, programa académico y horas de beneficio.
Caracterizados SEA	Archivos MS Excel con información similar pero no idéntica de los estudiantes caracterizados por el programa SEA en las dimensiones social, biosicosocial, económica, académica y cognitiva.

Con esta información se realizó una tarea de enlace de los datos suministrados en una base de datos única por cada cohorte de estudiantes analizada que contuviera todas las características de interés e hiciera más sencillo y reproducible los análisis objetivos de este trabajo. De esta forma se construyeron 4 bases de datos así: 2014-I, 2015-I, 2016-I y 2017-I, en donde la primera contiene información de los primeros 6 semestres y las demás solo del primer semestre académico.

Durante esta tarea de enlace y unificación de la información, se encontraron algunas inconsistencias, producto posiblemente del origen fragmentado de la información y/o de la alta dinámica de cambio de programa de algunos estudiantes. A continuación se describen estas inconsistencias y los ajustes correspondientes realizados:

- Estudiantes en la 1ª cohorte del 2014 que después de quedar PFU aparecían con notas en la misma carrera dos o tres semestres después. Según el reglamento de pregrado (División de Publicaciones UIS, 2015, pág. 12), no podrán inscribirse para iniciar nuevamente un programa académico quienes hayan perdido la calidad de estudiantes por bajo rendimiento académico en ese programa. La inconsistencia fue resuelta eliminando las notas de los estudiantes que aparecen con posterioridad al P.F.U. en la carrera original iniciada en la 1ª cohorte del 2014, entendiéndose que las notas posteriores pertenecer a un programa académico diferente y por lo tanto a otra cohorte, fuera del alcance planteado para este trabajo.
- Estudiantes con notas duplicadas en distintos programas: En la base de datos de notas se encontraron estudiantes que después de realizar cambio de carrera registraban en el nuevo programa notas idénticas para las mismas materias cursadas en el programa anterior. Se asume esta situación como producto de la homologación de materias idénticas entre carreras. A estos estudiantes se les eliminaron las notas duplicadas en las asignaturas de matemáticas y se dejaron las primeras notas pertenecientes al programa original iniciado en la primera cohorte del 2014.
- Estudiante con beneficios SEA en materias no relacionadas con el programa que adelanta. En la base de datos de beneficiarios del SEA se encontraron estudiantes recibiendo beneficios en asignaturas que no formaban parte del programa académico en el que aparecían registrados en el programa SEA, así como estudiantes en programas diferentes al registrado en la base de

datos de notas. Se identificó que estos casos se deben a estudiantes que cambiaron de carrera o que luego de quedar PFU se matricularon de nuevo en otro programa académico. En estos casos, se reasignó el beneficio respetando la pertenencia de las materias a la carrera en curso.

- Estudiantes con dos caracterizaciones SEA. En la base de datos de Caracterizados SEA se encontraron estudiantes doblemente caracterizados en periodos y programas diferentes que no coincidían con los de la base de datos de notas. Se comparó, caso por caso, con la base de datos de notas y la de beneficiarios y se ajustaron los registros con incoherencias.
- Para los estudiantes de la primera cohorte de los años 2015 y 2016, la base de datos de SABER 11 presenta una mezcla de las dimensiones evaluadas producto de su reestructuración en el año 2015. Se siguió el lineamiento dado por la universidad manteniéndose para el 2015 la anterior estructura y para el 2016 la nueva estructura de las pruebas SABER 11.

La Tabla 2 muestra las siglas empleadas en todo el documento para hacer referencia a las carreras de ciencias e ingeniería consideradas en el presente estudio.

Tabla 2.

*Codificación de Carreras.*

<b>Sigla</b>	<b>Carrera</b>
DIn	Diseño industrial
Fis	Física
Geo	Geología
Civ	Ingeniería civil
Pet	Ingeniería de petróleos
Sis	Ingeniería de sistemas
Ele	Ingeniería eléctrica
Elo	Ingeniería electrónica
Ind	Ingeniería industrial
Mec	Ingeniería mecánica
Met	Ingeniería metalúrgica
IQu	Ingeniería química
LMa	Licenciatura en matemáticas
Mat	Matemáticas
Qui	Química

En general, las variables y/o características de los estudiantes objeto de investigación se pueden agrupar, con base en la forma como fue recibida la información, en las siguientes categorías:

1. **Características Generales:** incluye variables naturales que aunque no están necesariamente ligadas al desempeño académico del estudiante, pueden ser empleadas para visualizar aspectos y características generales de forma conjunta. En este grupo de variables se encuentra el género, la edad y el programa académico del estudiante.
2. **Pruebas SABER 11:** son los resultados de las pruebas SABER 11 de los estudiantes admitidos a los diferentes programas académicos. Se evaluará si existe relación entre los resultados de las pruebas SABER 11 y el desempeño académico de los estudiantes, principalmente en el primer semestre.
3. **Desempeño Académico:** corresponde al desempeño académico general de los estudiantes semestre a semestre (promedio semestral y acumulado), así como las notas finales en las materias de cálculo (I, II, y III), álgebra y ecuaciones diferenciales. En este trabajo se busca evaluar si estas variables son o no afectadas por el uso de los beneficios que ofrece la universidad en el programa SEA.
4. **Beneficiarios Programa SEA:** variables relacionadas con el uso o no del programa SEA incluyendo intensidad horaria del beneficio por materia y semestre para cada estudiante. Se evaluará la relación entre la intensidad del beneficio recibido y el desempeño académico del grupo de estudiantes objeto del estudio.
5. **Caracterización Inicial Programa SEA:** estas variables corresponden a la caracterización realizada por el programa SEA para identificar la vulnerabilidad de los estudiantes en las dimensiones social, biosicosocial, económica, académica y cognitiva.

### 3.1.1 Estudiantes Cohorte I - 2014 / Semestre 1 a 6

#### 3.1.1.1 Características Generales.

- **Distribución por Género.**

La Figura 3 muestra el número y porcentaje de estudiantes de ciencias e ingeniería que ingresaron a la universidad en el 2014 por género y carrera. En el panel de la izquierda se presenta el total de estudiantes que ingresaron a la universidad para cada carrera en el 2014, en el panel central y en el derecho se muestran los estudiantes que ingresaron en el primero y segundo semestre del 2014 respectivamente. Se observa que las proporciones por género total y por carrera se mantienen más o menos constantes para ambas cohortes. En general, se destaca que la proporción total de hombres (67%) es algo más del doble de la de mujeres (33%).

Esta superioridad se manifiesta claramente en todas las carreras de ingeniería excepto en Diseño Industrial, Ing. Química y en Ing. Industrial donde los porcentajes son muy similares.

En ciencias las proporciones son más variadas: en Física la gran mayoría (79%) son hombres, mientras que en Química lo son las mujeres (60%) mientras que en Licenciatura en Matemáticas y Matemáticas los porcentajes son más equilibrados.

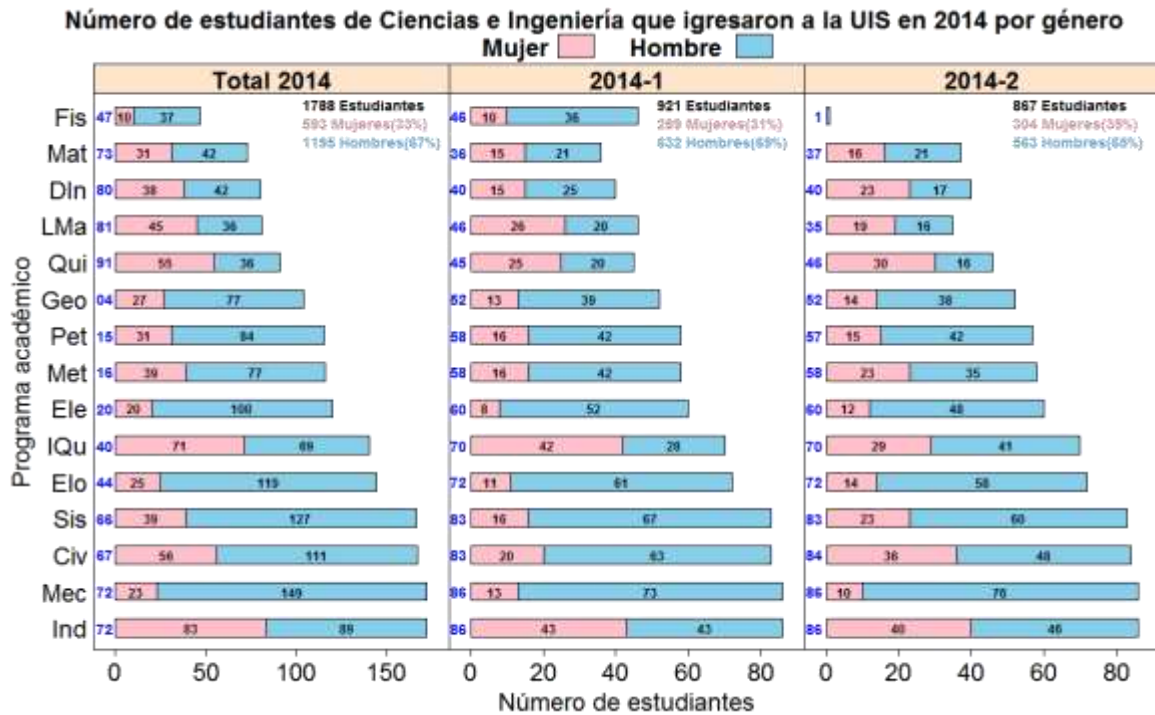


Figura 3. Distribución por género cohortes 2014, Estudiantes de Ciencias e Ingenierías.

La Figura 3 también muestra el número de estudiantes por carrera (etiqueta de color azul en el borde izquierdo de cada barra o carrera) y el total por semestre (parte superior derecha de cada panel). La información se presenta ordenada en forma decreciente de acuerdo al número de estudiantes admitidos. Las carreras con menor número de estudiante admitidos fueron: física con 47 estudiantes, seguida de matemáticas y diseño industrial con 73 y 80 respectivamente. En contraste, las carreras con mayor número de admitidos fueron ingeniería industrial e ingeniería mecánica con 172 estudiantes cada una, seguida de ingeniería civil e ingeniería de sistemas con 167 y 166 respectivamente. La población objeto de estudio se muestra en el panel central de la Figura 3, son 921 estudiantes admitidos en el primer semestre del 2014, 69% hombres y 31% mujeres.

La

Figura 4 muestra en orden decreciente, la proporción de hombres y mujeres en cada carrera; en ella se observa que ingeniería mecánica fue la carrera preferida por hombres (87%), seguida de ingeniería eléctrica y electrónica (83% cada una), física (79%), ingeniería de sistemas (77%), geología (74%) e ingeniería de petróleos (73%) mientras que Química, licenciatura en matemáticas e Ingeniería química fueron preferidas por mujeres con un 60%, 56% y 51% respectivamente.

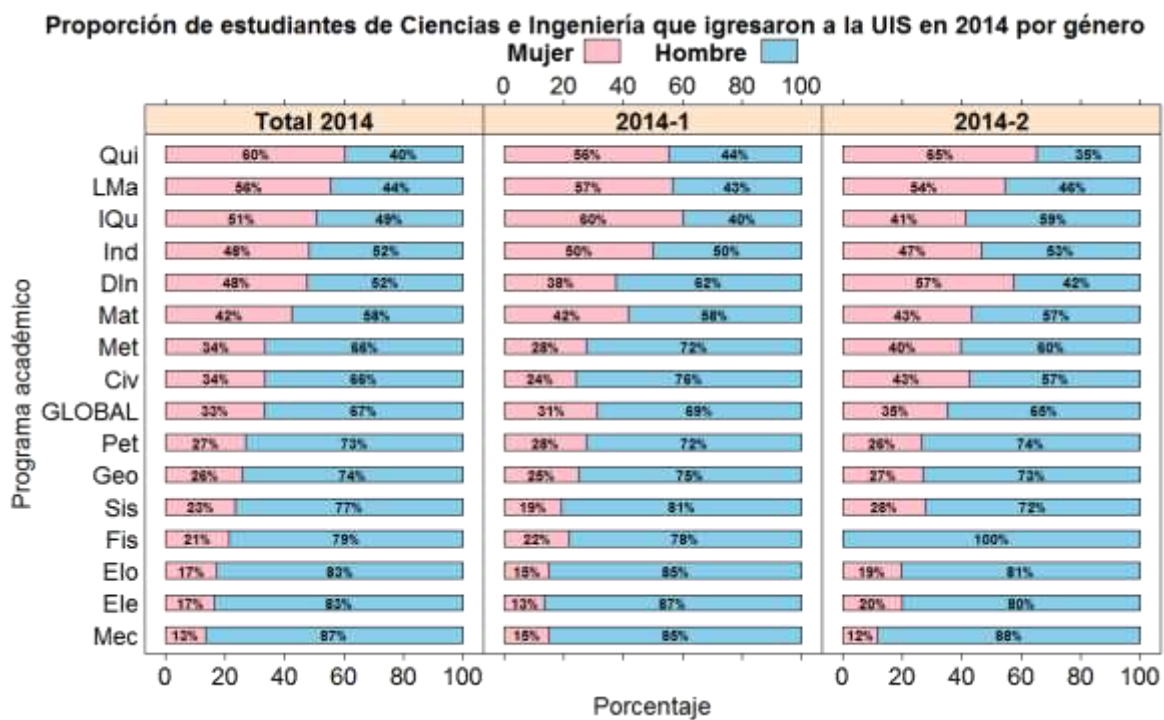


Figura 4. Proporción de estudiantes por género cohortes 2014, Estudiantes de Ciencias e Ingenierías.

- **Distribución por Edad**

La Figura 5 se construyó de forma análoga a la Figura 3 pero agrupando los estudiantes por su edad de ingreso a la universidad. Se observa que la mayor cantidad de estudiantes que ingresaron

a la universidad fueron jóvenes entre 18 y 20 años (58%) seguido de menores de edad (31%) y, en menor proporción, estudiantes mayores de 20 años (11%); proporciones que se mantienen similares en cada carrera en los dos semestres del 2014 y a nivel global. La edad es un factor que puede reflejar el nivel de madurez con la que un estudiante afronta su paso por la universidad y por lo tanto puede afectar su desempeño académico.

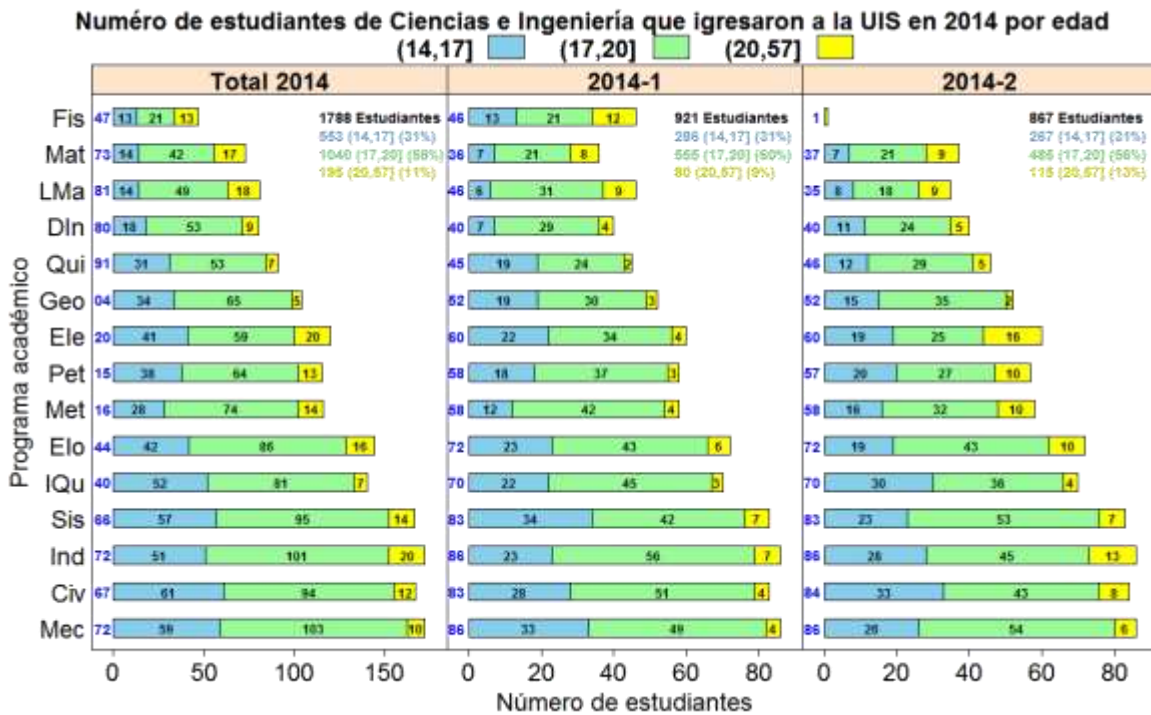


Figura 5. Distribución por edad de ingreso cohortes 2014, Estudiantes de Ciencias e Ingenierías.

La Figura 6 muestra en orden decreciente la proporción de estudiantes menores de edad en cada carrera. Las carreras con mayor proporción de estudiantes menores de edad que ingresaron en el 2014 (panel izquierdo de la Figura 6) fueron ingeniería química y civil (37% cada una) seguida de ingeniería de sistemas, ingeniería mecánica, ingeniería eléctrica y química (34%) mientras que las carreras con menor proporción de menores de edad fueron licenciatura en matemáticas (17%),

seguido de matemáticas (19%), diseño industrial (22%) e ingeniería metalúrgica (24%). Como se observa en el panel central y derecho de la Figura 6, estos porcentajes no se mantienen similares en las dos cohortes del 2014, cambian a nivel global y por carreras. En general, los admitidos con mayoría de edad aumentaron en el segundo semestre.

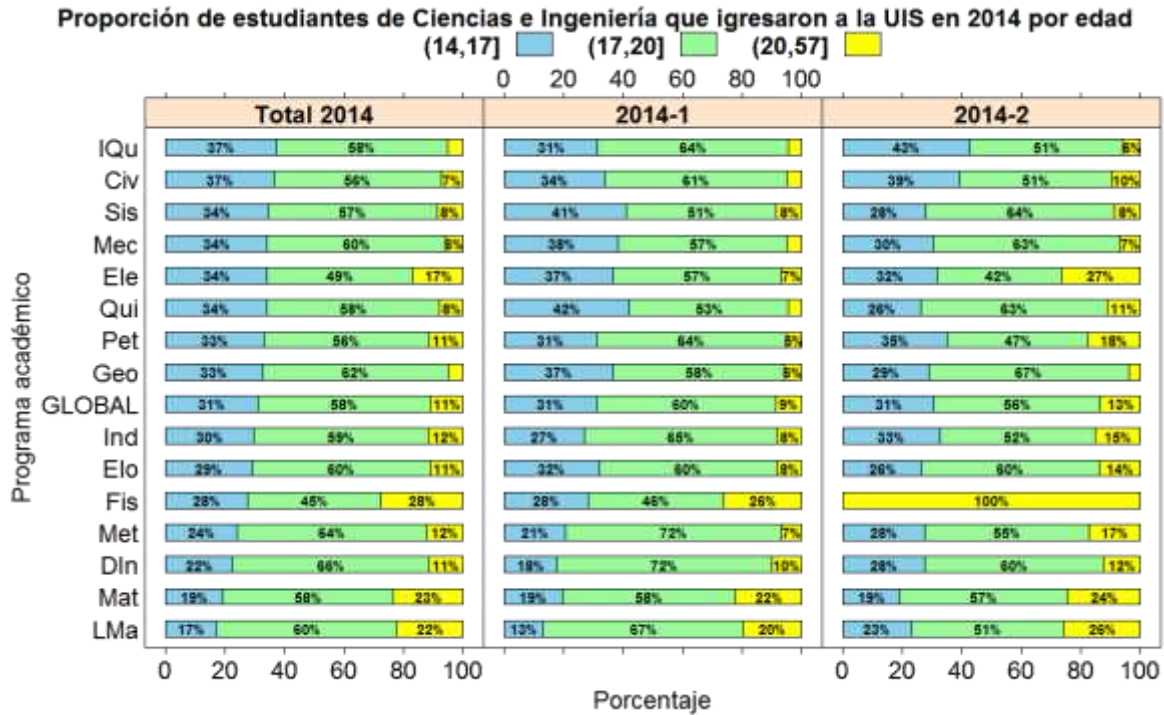


Figura 6. Proporción de estudiante que ingresaron en 2014 por edad, Estudiantes de Ciencias e Ingenierías.

También se observa que las carreras con mayor proporción de estudiantes mayores de 20 años fueron las de ciencias: física (28%) seguida de matemáticas (23%) y licenciatura en matemáticas (22%); mientras que las carreras con mayor proporción de estudiantes menores de 21 años fue geología (95%) seguido de Ingeniería química (95%) e ingeniería mecánica (94%).

**3.1.1.2 Desempeño Académico.** En esta sección se analizan los resultados académicos de los estudiantes de ciencia e Ingeniería que ingresaron a la universidad en el primer semestre del 2014 incluyendo el desempeño global (promedio semestral y acumulado) y específico en las asignaturas de matemáticas (Cálculo I, II y III; álgebra lineal y ecuaciones diferenciales). La primera sección describe el desempeño global del estudiante, la siguiente sección presenta el análisis de su condicionalidad y la última sección presenta un análisis del desempeño específico en las asignaturas del área de matemáticas (Cálculo I, II y III, álgebra lineal y ecuaciones diferenciales) vistas por los estudiantes de ciencias e ingeniería que entraron a la universidad en la primera cohorte del 2014.

**Promedio semestral y acumulado por semestre.** En este apartado se presenta un análisis comparativo del promedio académico (semestral y acumulado) de los estudiantes de Ciencias e Ingenierías que ingresaron a la UIS en el primer semestre del 2014. La Tabla 3 resume los estadísticos del seguimiento a los estudiantes de la primera cohorte del 2014 para el promedio académico semestral y acumulado en los primeros seis semestres (media,  $\bar{x}$  mediana,  $\hat{x}$  y coeficiente de variación,  $c_v$ )

Tabla 3.

*Desempeño académico global, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014.*

Semestre	Estadístico	Promedio Semestral	Promedio Acumulado
2014-1	$C_v$	0.26	0.26
	$\bar{x}$	3.23	3.24
	$\hat{x}$	3.36	3.36
2014-2	$C_v$	0.19	0.13
	$\bar{x}$	3.4	3.47
	$\hat{x}$	3.46	3.44
2015-1	$C_v$	0.18	0.11
	$\bar{x}$	3.55	3.56
	$\hat{x}$	3.64	3.54
2015-2	$C_v$	0.15	0.09

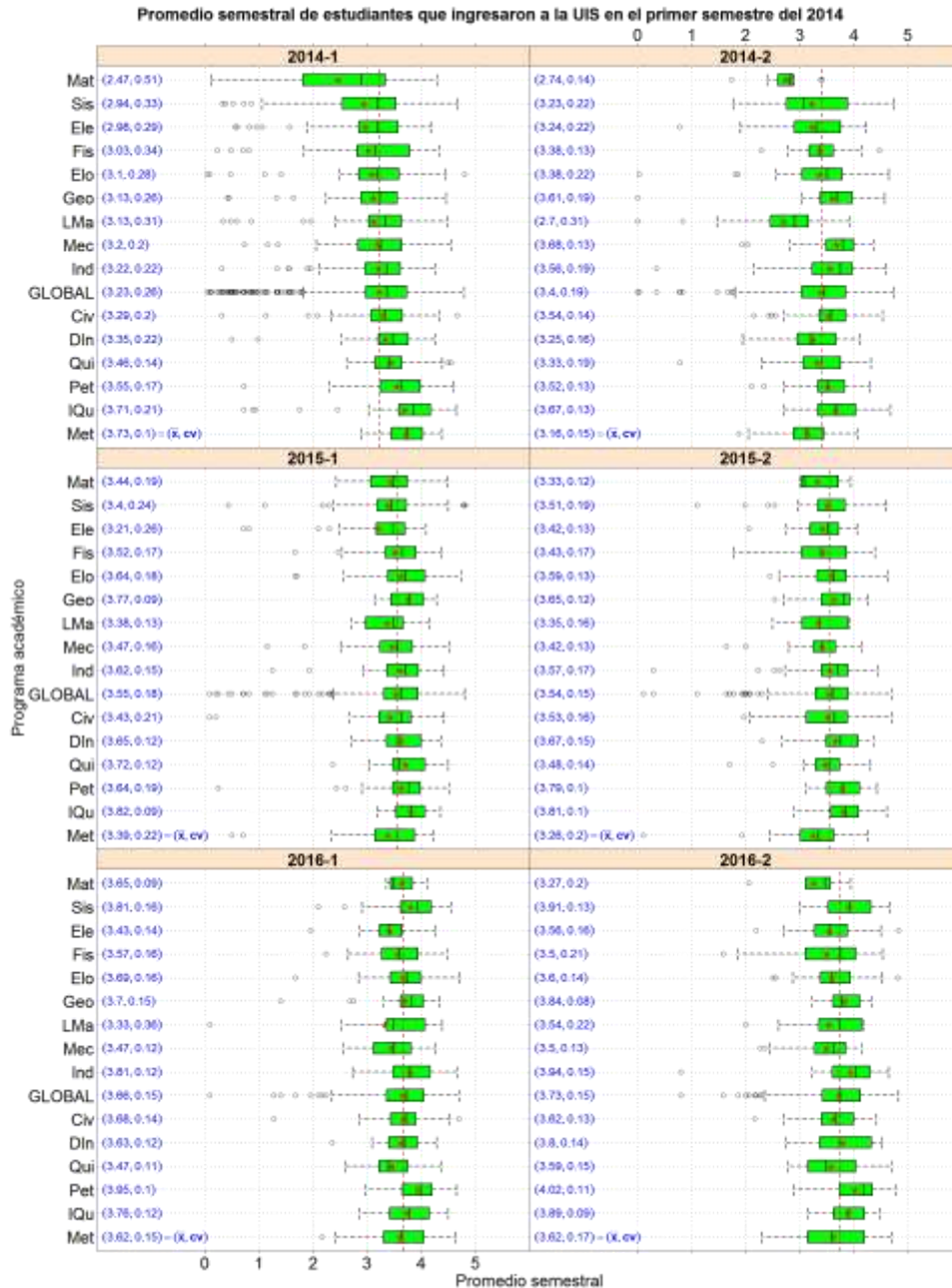
	$\bar{x}$	3.54	3.61
	$\hat{x}$	3.62	3.59
	$C_v$	0.14	0.09
2016-1	$\bar{x}$	3.66	3.64
	$\hat{x}$	3.73	3.6
	$C_v$	0.14	0.09
2016-2	$\bar{x}$	3.73	3.67
	$\hat{x}$	3.76	3.63

La Figura 7 detalla los resultados por carrera de la Tabla 3 empleando diagrama de cajas para el promedio semestral. Las carreras están ordenadas de forma descendente por su promedio en el primer semestre (2014-1) y este orden se conserva en los demás semestres, para facilitar la comparación. Se observa que la carrera de menor promedio académico fue matemáticas (2.5) seguida de ingeniería de sistemas (2.9) e ingeniería eléctrica (3.0). Entre tanto, las carreras de mayor promedio fueron ingeniería metalúrgica e ingeniería química (3.7 c/u). A nivel global la media en el primer semestre es de 3.2 (justo en el límite definido para la condicionalidad) y aumenta gradualmente en los semestres siguientes hasta 3.7 en el sexto semestre, debido principalmente a la salida de los estudiantes de menos rendimiento académico (PFU). En la Tabla 3 se observa la reducción en la variabilidad<sup>2</sup> del promedio semestral con el tiempo (los coeficientes de variación decrecen con el avance semestral), y una mayor homogeneidad entre los promedios a medida que se avanza en los semestres (los promedios por carrera son más próximos al promedio del semestre).

No se observa una distribución simétrica generalizada en las carreras (centralidad de la línea continua vertical en la caja correspondiente a la mediana) y tampoco es evidente una evolución en este aspecto al avanzar entre semestres. No obstante, el valor global muestra simetría en todos los semestres. El número de valores atípicos disminuye al avanzar entre semestres y corresponde en

<sup>2</sup> Medida como el coeficiente de variación ( $C_v = s/\bar{x}$ ,  $s = \sqrt{\sum_{k=1}^n \frac{(x_k - \bar{x})^2}{n-1}}$ ) donde s es la desviación típica y  $\bar{x}$  es la media de cada grupo.

su gran mayoría a valores atípicos bajos. Lo anterior no es extraño y solo refleja que existen más cantidad de estudiantes con promedios notablemente bajos respecto a estudiantes con promedios notablemente altos.



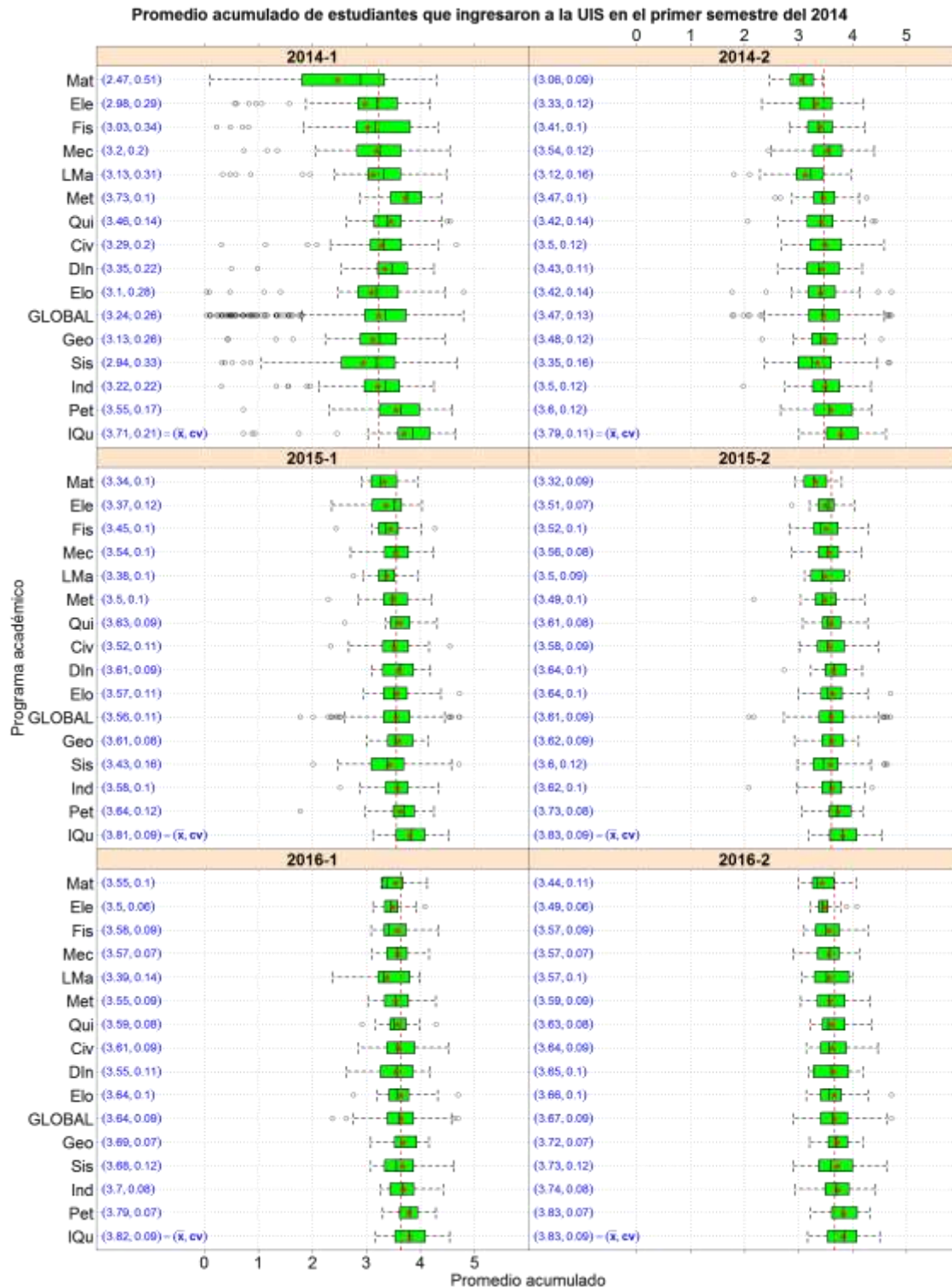
Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 7. Promedio semestral, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 semestres.

La Figura 8 detalla los resultados por carrera de la Tabla 3 empleando diagrama de cajas para el promedio acumulado. Las carreras están ordenadas de forma descendente por su promedio en el sexto semestre (2014-6) y este orden se conserva en los demás semestres.

Las carreras de mejor desempeño global después del sexto semestre fueron ingeniería química y petróleos (3.8 c/u) seguida de geología, ingeniería industrial, ingeniería de sistemas y geología (3.7 c/u); todas ellas por igual o encima de la media de la universidad (3.7). Entre tanto, las carreras de desempeño promedio más bajo fueron matemáticas (3.4) e ingeniería eléctrica (3.5).

Debido a que el promedio acumulado es menos influenciado por el promedio semestral a medida que se avanza en semestres, es evidente que existe una más clara y rápida evolución hacia la homogeneidad (en la Tabla 3 se observa que el coeficiente de variación es el mismo para los últimos 3 semestres) y simetría de las distribuciones para cada carrera (Figura 8), e incluso entre carreras al final del sexto semestre, en donde la media y la desviación son muy similares.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 8. Promedio acumulado, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 semestres.

**Condicionabilidad por semestre.** Según el reglamento académico-estudiantil de pregrado de la UIS, un estudiante que al finalizar un período académico obtenga un promedio ponderado acumulado igual o superior a dos coma setenta (2,70) pero inferior a tres coma veinte (3,20), será considerado estudiante condicional.

En caso de que su promedio sea inferior a 2.70 quedará excluido de la Universidad y en su hoja de vida se anotará “Excluido de la Universidad por bajo rendimiento académico (P.F.U)”.

La Tabla 4 resume la evolución académica de los estudiantes de ciencias e ingenierías de la primera cohorte del 2014, muestra la proporción de estudiantes en cada condición académica respecto al total de estudiantes de la cohorte (columnas 2 a 5) y la proporción de estudiante PFU respecto al total de estudiantes PFU en los tres primeros años de carrera (columna 6 y 7).

Se observa que del total de los estudiantes que quedaron PFU en 6 semestres, el 52% se dio en el primer semestre, y para el tercer semestre el acumulado llegó al 83%; al finalizar el sexto semestre, cerca del 25% del total de los estudiantes quedaron PFU; durante el sexto solo el 51% de los estudiantes están en condición normal, 47% inactivos (estudiantes sin actividad académica) y un 2% condicional. Estas cifras muestran una alta deserción y mortalidad académica en los primeros 3 semestres y respaldan que los esfuerzos de cualquier iniciativa de apoyo se concentren en el ciclo básico, especialmente en el primer semestre.

Tabla 4.

*Condición académica global. 1ª cohorte del 2014. Primeros 6 semestres.*

Respecto a:	Total Estudiantes de la Cohorte (921)				Total Estudiantes PFU en 6 Semestres (235)		
	PFU	Condicional	Normal	Inactivo	PFU Acumulado	PFU	PFU Acumulado
2014-1	13.4%	22.5%	59.7%	4.5%	13.4%	52.3%	52.3%
2014-2	2.8%	16.0%	55.7%	25.5%	16.2%	11.1%	62.4%
2015-1	5.0%	4.8%	53.4%	36.8%	21.2%	19.6%	83.0%
2015-2	2.5%	2.5%	52.8%	42.2%	23.7%	9.8%	92.8%
2016-1	1.3%	1.8%	50.8%	46.0%	25.0%	5.1%	97.9%
2016-2	0.5%	1.7%	51.1%	46.6%	25.5%	2.1%	100.0%

La Figura 9 detalla la evolución académica de los estudiantes por carrera, ilustrando las proporciones de condición normal (Verde), condicional por primera vez (Amarillo claro) o segunda vez (Amarillo oscuro), PFU (Rojo) e Inactivo (Gris).

Las carreras con mayor deserción<sup>3</sup> fueron matemáticas (83%) seguida de licenciatura en matemáticas (80%), ingeniería de sistemas (60%) e ingeniería eléctrica (58%) mientras que las carreras con mayor cantidad de estudiantes que permanecieron en situación normal al finalizar el sexto semestre fueron ingeniería química (77%), ingeniería de petróleos (71%) e ingeniería mecánica y metalúrgica (62% c/u).

<sup>3</sup> Medida como inactividad académica al finalizar el segundo semestre del 2016.

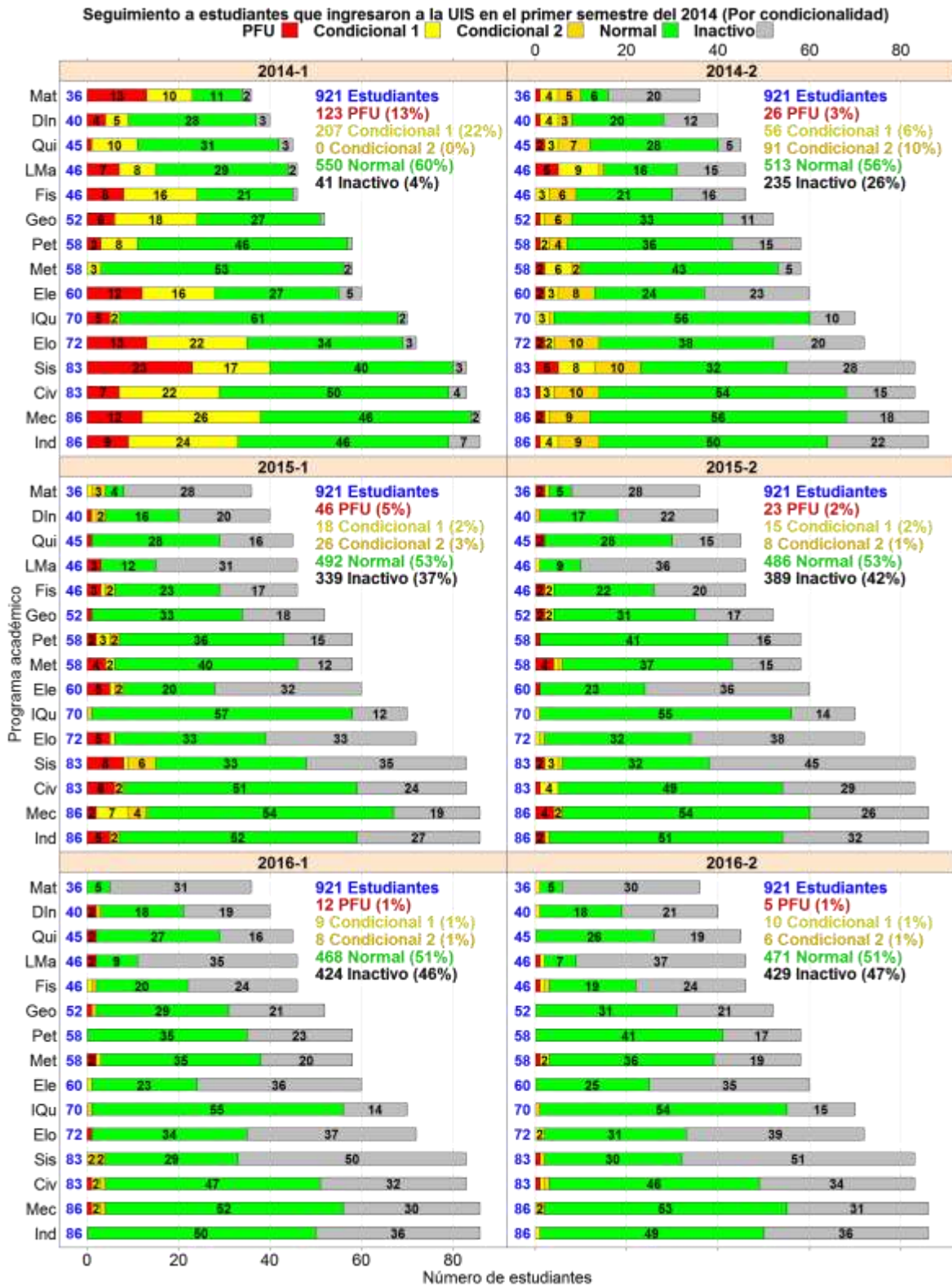


Figura 9. Distribución por condición académica, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 Semestres.

La Figura 10 muestra, ordenados de menor a mayor el porcentaje de estudiantes que finalizaron PFU el primer semestre. A nivel global (barra etiquetada como “GLOBAL” en la parte central de cada panel en la Figura 10), se observa que el 13% de los estudiantes que ingresaron a la universidad en el primer semestre, terminaron PFU, destacándose las carreras de matemáticas (36%) e ingeniería de sistemas (28%) con la mayor proporción, mientras que ingeniería metalúrgica (0%) y química (2%) son los porcentajes más bajos.

La primera condicionalidad resulta una situación de difícil solución al observarse que de los estudiantes por primera vez condicionales en el semestre 1 (en amarillo claro) el 44 % pasaron a condicionales por segunda vez (amarillo oscuro).

Este valor aumenta levemente en cada semestre hasta llegar al 67% en global y es más grave en carreras como química (70%), ingeniería metalúrgica (67%) y menor en licenciatura en matemáticas (13%).

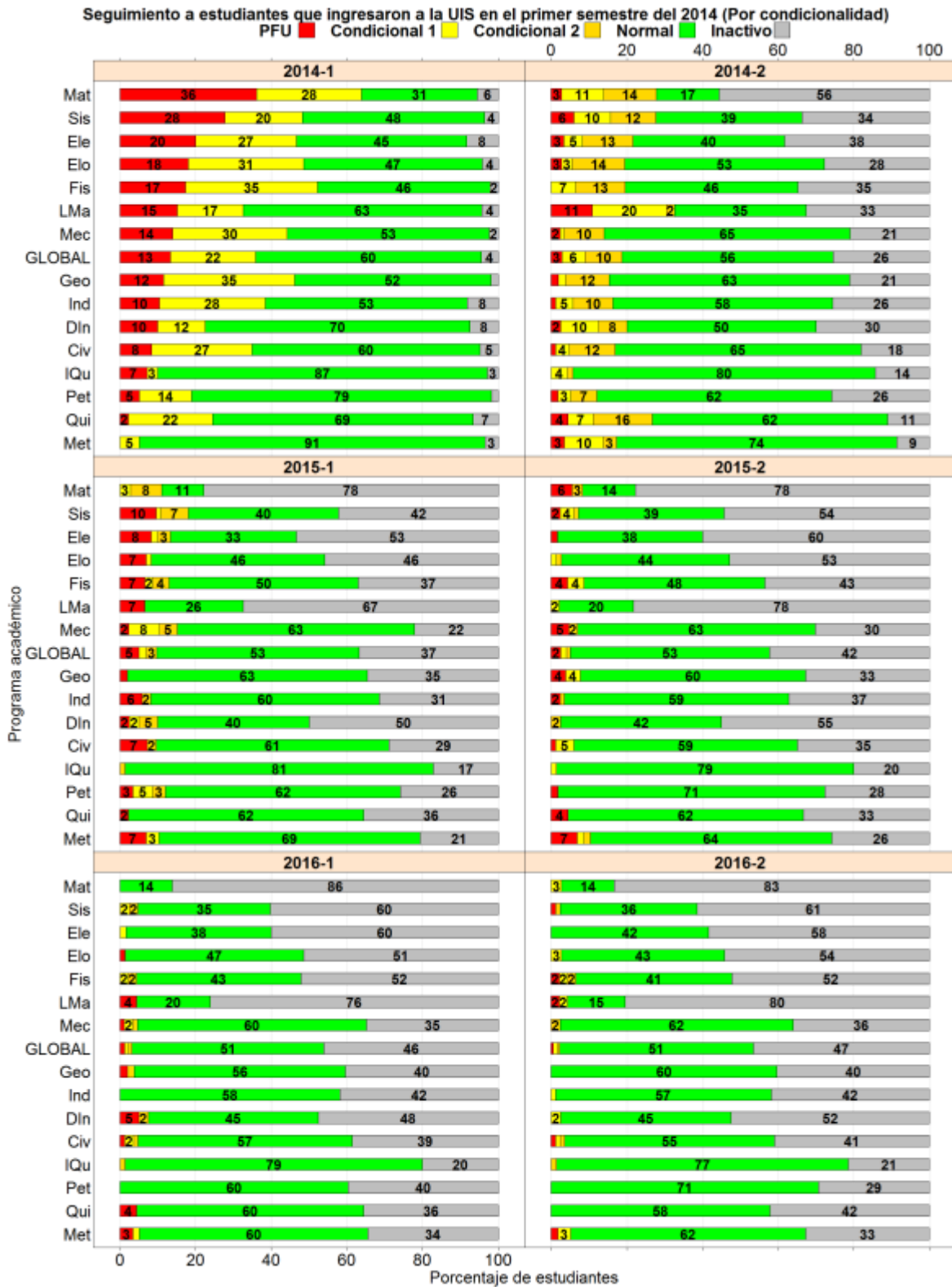


Figura 10. Distribución por condición académica, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 Semestres.

La Figura 11 muestra una vista complementaria de la deserción, clasificando los estudiantes en cuatro grupos que se acumulan hasta el semestre correspondiente: “continuo” (verde), estudiantes que presentaron actividad académica consecutiva sin interrupción ni haber quedado PFU; “discontinuos” (amarillo), como aquellos estudiantes que presentaron inactividad en algún semestre intermedio; “PFU” (rosado), como aquellos que quedaron por fuera de la universidad y, finalmente, los “Inactivos” (azul) que son aquellos que no presentan actividad académica.

La continuidad ha sido definida como la actividad académica en todos los semestres y la discontinuidad como la situación opuesta. Se observa la forma como evoluciona el grupo de estudiantes en cada categoría a nivel general (estadísticos en la parte superior derecha de cada panel) y por carrera (barras en cada panel).

Se destaca que la proporción de estudiantes que muestran actividad continua en todos los semestres decae de 82% en el primer semestre a 48% en el sexto; de igual forma, se aprecia que la proporción de retiro de un programa aumenta desde el 4% hasta un valor que se estabiliza en 22%.



Figura 11. Distribución por continuidad y deserción, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014. Primeros 6 Semestres.

En la Figura 12 se aprecia con mayor facilidad y de forma comparativa estos porcentajes por carrera al finalizar el sexto semestre. Se destaca que solo el 48% de los estudiantes finalizaron sus estudios sin interrupciones, 4% suspendieron en algún momento y reanudaron después, 26% finalizaron PFU y 22% aparecen sin actividad académica. Las carreras con menor y mayor proporción de PFU acumulado fueron Ingeniería de Sistemas (47%) e Ingeniería Química (7%) respectivamente; la mayor proporción de retirados se dio en matemáticas (44%) y licenciatura en matemáticas (43%) mientras que la carrera con mayor proporción de estudiantes que finalizaron sin interrupciones o recesos fue Ingeniería Química (77%).

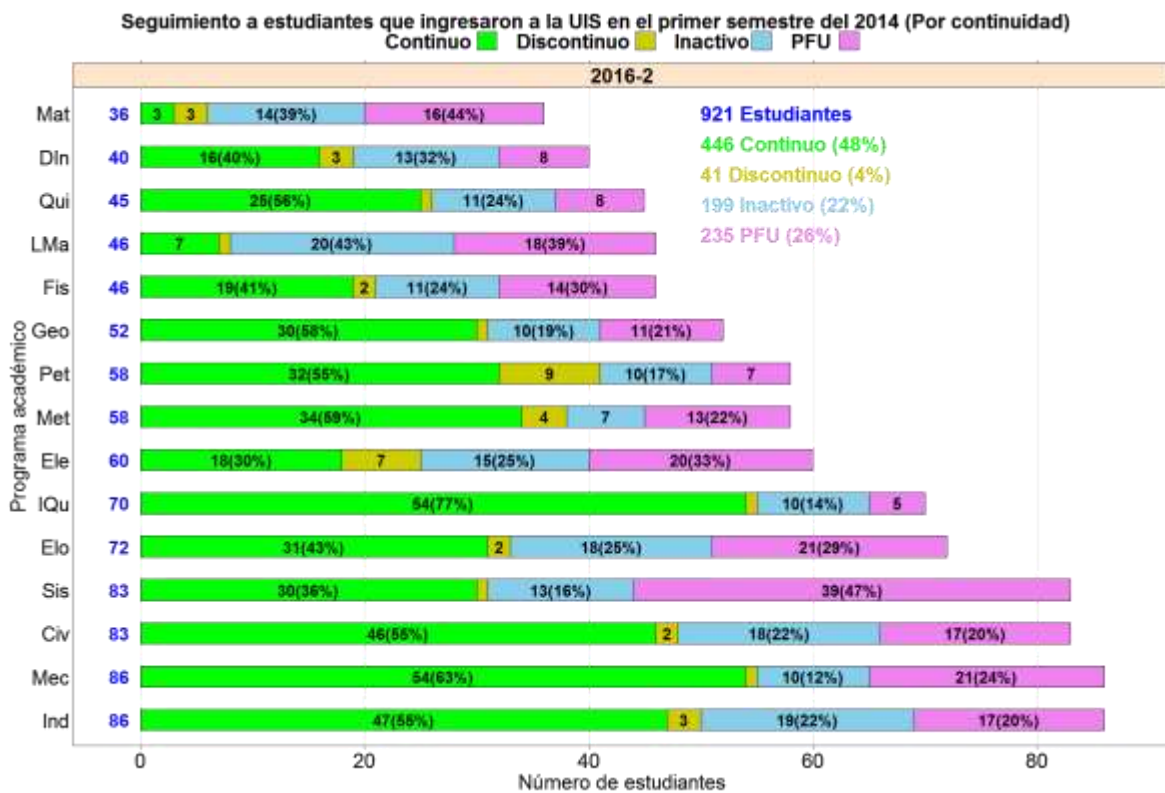


Figura 12. Distribución por continuidad y deserción, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014 – sexto semestre.

**Desempeño en asignaturas del área de matemáticas.** En esta sección se analiza la distribución de notas en el área de matemáticas de los estudiantes que ingresaron a la universidad en el primer semestre del 2014 específicamente en Cálculo I, II y III, álgebra lineal y ecuaciones diferenciales.

La Tabla 5 resume los estadísticos de las notas en asignaturas de matemáticas de los estudiantes de la primera cohorte del 2014 (media,  $\bar{x}$  mediana,  $\hat{x}$  y coeficiente de variación,  $Cv$ ) mientras que la Figura 13 detalla las notas en asignaturas de matemáticas<sup>4</sup> de los estudiantes de Ciencias e Ingeniería que ingresaron a la UIS en el primer semestre del 2014.

Tabla 5.

*Desempeño académico en asignaturas de matemáticas, Estudiantes de Ciencias e Ingeniería, 1<sup>a</sup> cohorte de 2014.*

Asignatura	Estadístico	Valor
Álgebra	$Cv$	0.33
	$\bar{x}$	2.97
	$\hat{x}$	3.15
Cálculo I	$Cv$	0.41
	$\bar{x}$	2.60
	$\hat{x}$	2.70
Cálculo II	$Cv$	0.29
	$\bar{x}$	2.90
	$\hat{x}$	3.00
Cálculo III	$Cv$	0.24
	$\bar{x}$	3.20
	$\hat{x}$	3.20
Ecuaciones	$Cv$	0.24
	$\bar{x}$	3.17
	$\hat{x}$	3.20

Se observa que el promedio de la universidad en asignaturas como cálculo I, cálculo II y álgebra se encuentra por debajo de 3.0 mientras que cálculo III y ecuaciones diferenciales está por encima

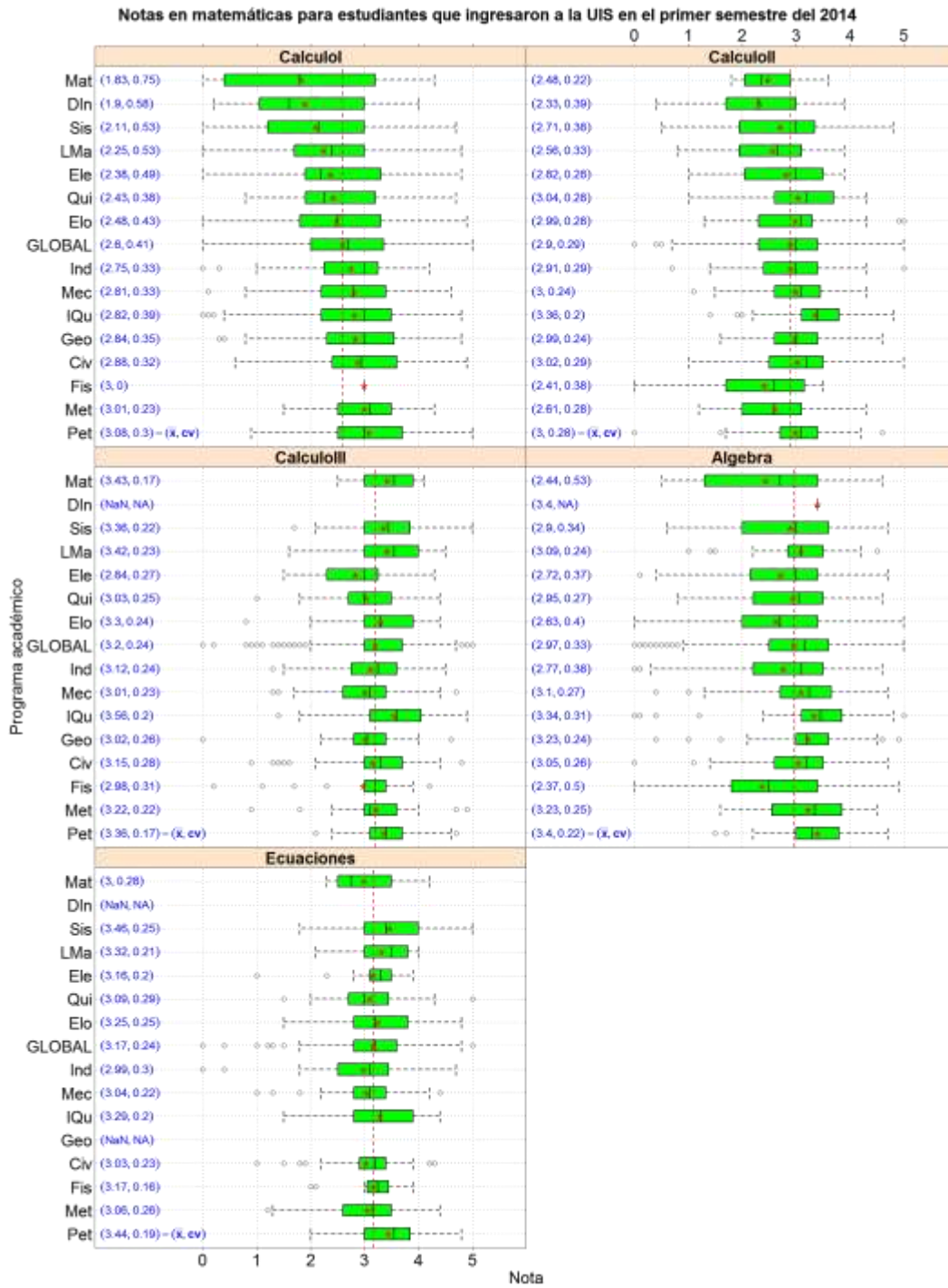
<sup>4</sup> Se considera la nota del estudiante la primera vez que aparece registrada. Es decir, no se tiene en cuenta la cancelación de la asignatura antes de finalizarla, la cual no aparece registrada en la base de datos de notas.

de 3.0. La carrera de Matemáticas presenta la nota promedio más baja en cálculo I y álgebra, y con la mayor dispersión; entre tanto, ingeniería de petróleos obtuvo el mayor promedio en estas asignaturas.

La nota promedio en los cálculos, aumenta con su avance del ciclo básico, cálculo I (2.6), cálculo II (2.9), cálculo III (3.2) y ecuaciones diferenciales (3.2); el promedio de Álgebra lineal en el primer semestre fue de 3.0.

De igual forma la dispersión de los datos es mayor en materias de primer semestre como cálculo I (1.1) y álgebra (1.0), y va disminuyendo para materias de niveles posteriores como cálculo II (0.8,) cálculo III (0.8) y ecuaciones diferenciales (0.7).

En esta evolución aparentemente positiva subyace el hecho de la alta deserción y discontinuidad académica concentrada en los primeros 4 semestres (Tabla 4) que reduce el número de estudiantes con bajo rendimiento entre cada semestre de tal manera que los estudiantes sobrevivientes se van estabilizando, reduciendo su variabilidad pero haciendo más notorios los casos atípicos.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 13. Distribución de resultados en matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

**3.1.1.3 Pruebas SABER 11.** En esta sección se presenta el análisis descriptivo de los resultados en la pruebas SABER 11 de los estudiantes de ciencias e ingeniería que ingresaron a la universidad en el primer semestre del 2014 y su relación con las notas del estudiante en el primer semestre (en matemáticas y a nivel general). La Tabla 8 resume los estadísticos de los puntajes en las distintas áreas de conocimiento en las pruebas SABER 11 de los estudiantes de la primera cohorte del 2014 (media,  $\bar{x}$  mediana,  $\hat{x}$  y coeficiente de variación, Cv).

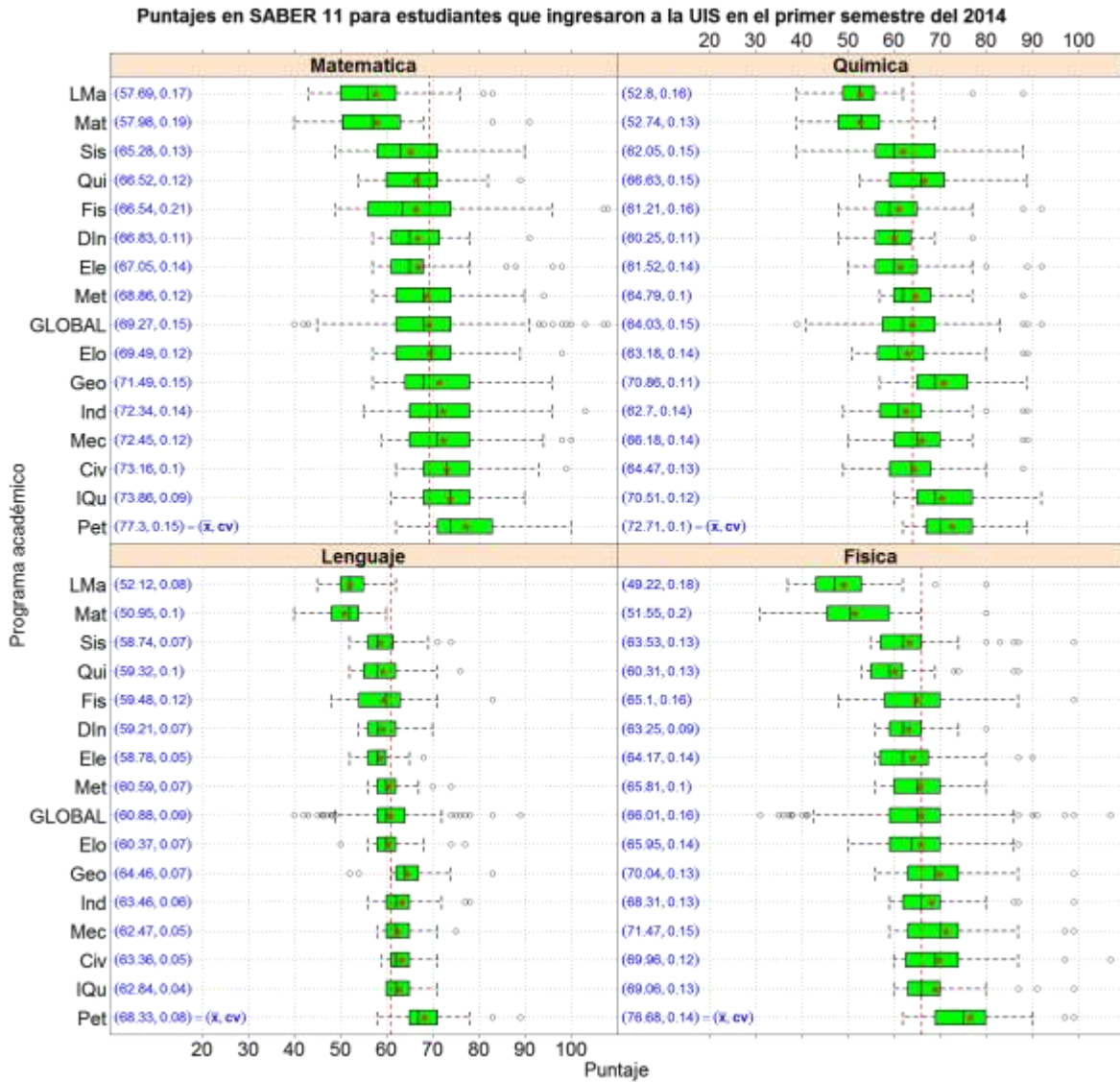
Tabla 6.

*Resultados globales en Pruebas SABER 11, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014.*

Área	Estadístico	Puntaje SABER 11
Biología	<i>Cv</i>	0.13
	$\bar{x}$	58.39
	$\hat{x}$	58.00
Filosofía	<i>Cv</i>	0.17
	$\bar{x}$	54.22
	$\hat{x}$	53.00
Física	<i>Cv</i>	0.16
	$\bar{x}$	66.01
	$\hat{x}$	65.00
Geografía	<i>Cv</i>	0.14
	$\bar{x}$	58.24
	$\hat{x}$	59.00
Historia	<i>Cv</i>	0.14
	$\bar{x}$	58.24
	$\hat{x}$	59.00
Inglés	<i>Cv</i>	0.26
	$\bar{x}$	60.86
	$\hat{x}$	57.27
Lenguaje	<i>Cv</i>	0.09
	$\bar{x}$	60.88
	$\hat{x}$	60.00
Matemáticas	<i>CV</i>	0.15
	$\bar{x}$	69.27
	$\hat{x}$	68.00
Química	<i>CV</i>	0.15
	$\bar{x}$	64.03
	$\hat{x}$	62.00

Para facilitar la interpretación de las gráficas se consideraron dos grupos de materias en las pruebas SABER 11: el primer grupo formado por Lenguaje, Física, Matemática, Química y el segundo grupo formado por Biología, Filosofía, Historia, Geografía e inglés. La Figura 14 muestra la distribución (diagrama de cajas) de los puntajes en las pruebas SABER 11 del primer grupo y en la Figura 15 se muestran para el segundo grupo. Se observa que las carreras con mayor puntaje en el área de matemáticas fueron ingeniería de petróleos e ingeniería química (por encima de la media de la universidad) mientras que las carreras de más bajo corte fueron matemáticas y licenciatura en matemáticas, característica que se repite en las demás áreas de conocimiento evaluadas.

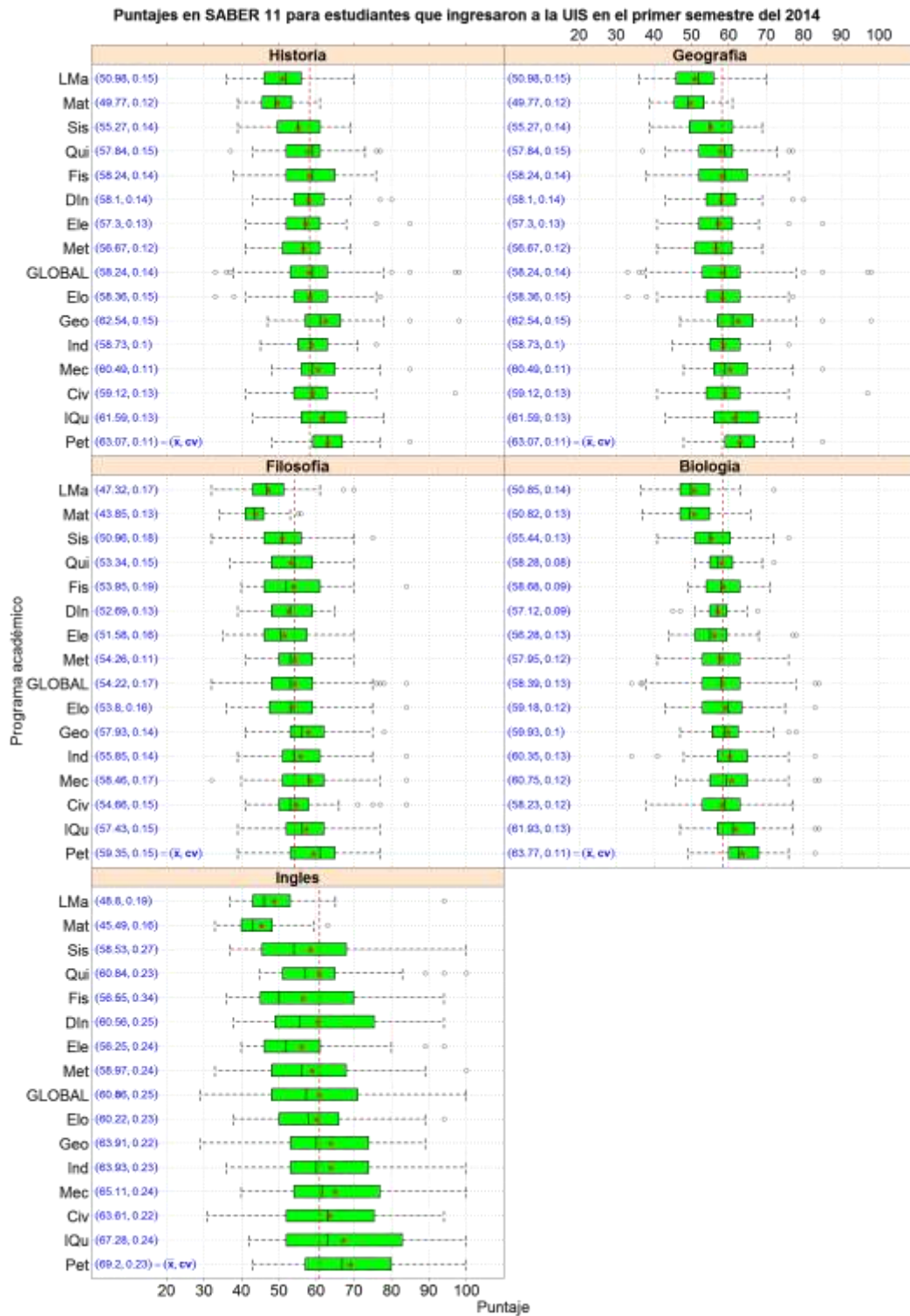
Esto coincide con los resultados observados en la sección anterior, en donde estas mismas carreras se ubicaron en los extremos correspondientes para el desempeño académico en las asignaturas de matemáticas. Sin embargo esta aparente relación entre desempeño académico y el resultado de las pruebas SABER 11 solo es clara en los extremos de la clasificación de carreras. En general, se observa una distribución poco simétrica y alta dispersión para todos los puntajes en todas las carreras así como valores atípicos elevados.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 14. Distribución de resultados en pruebas SABER 11 Grupo 1, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

La Figura 15 muestra el diagrama de cajas comparativo para las materias del segundo grupo en las pruebas Saber 11. Nuevamente se observa en los extremos las carreras de matemáticas y licenciatura en matemáticas, como las de menor puntaje e ingeniería de petróleos e ingeniería química como las de mayor puntaje.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 15. Distribución de resultados en pruebas SABER 11 Grupo 2, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

**3.1.1.4 Beneficiarios Programa SEA.** La base de datos con los beneficiarios del SEA contiene el registro de todos los estudiantes que se beneficiaron de alguno de los subprogramas del SEA desde el 2014 hasta el 2016, incluyendo el código y programa del estudiante, año y periodo en el que recibió el beneficio, el sub programa SEA usado, la asignatura en la que recibió el beneficio y las horas de beneficio. La Figura 16, muestra los beneficios en asignaturas de matemáticas del ciclo básico (Ben. Matemáticas) y en otras áreas (Otros Beneficios) que totaliza el número de beneficiados del SEA; los beneficios corresponden a la suma de beneficiarios asignatura por asignatura, sin diferenciar si un mismo estudiante recibió beneficios en varias asignaturas a la vez, por lo que el número de estudiantes totalizado no corresponde a estudiantes individuales; los estudiantes no beneficiados están divididos en estudiantes que no fueron reportados en la base de datos de beneficio del SEA y estudiantes que fueron reportados inscritos, pero totalizan cero horas de beneficio en el programa SEA; se asume que estas dos categorías corresponden a estudiantes en la misma condición que no tuvieron contacto con ningún subprograma del SEA.

La Figura 16 ilustra la distribución de beneficios del SEA para la población de estudiantes de ciencias e ingeniería global y por carrera. Se encontró que el 37% obtuvieron algún beneficio del SEA entre el 1er semestres del 2014 y el 2º semestre de 2016: un 27% en asignaturas de matemáticas (Cálculo I, II y III, álgebra y ecuaciones diferenciales) y un 10% en otras asignaturas (tanto materias del ciclo básico en física y química, como distintas asignaturas de carrera); un 40% no registran ningún beneficio del SEA y un 23% de los estudiantes que aparecen como beneficiarios no tienen horas registradas de participación. Los últimos dos grupos fueron fusionados como no beneficiarios para propósitos de este estudio, pues se asume que un estudiante sin horas registradas tiene cero horas de beneficio y por lo tanto no recibió ninguna influencia del

programa SEA en su desempeño académico. Se observa que las carreras con mayor número de estudiantes beneficiados en asignaturas de matemáticas fueron ingeniería metalúrgica e industrial (24 estudiantes cada una) mientras que diseño industrial fue la de menor número de estudiantes beneficiados (10 estudiantes).

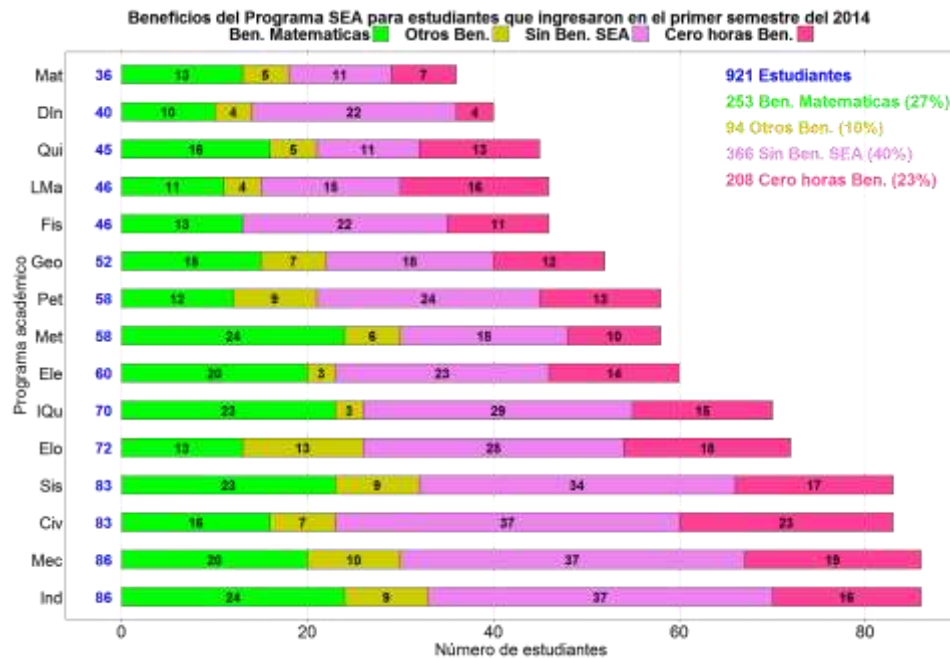


Figura 16. Beneficiarios del programa SEA por carrera, Estudiantes de Ciencias e Ingenierías, 1<sup>a</sup> cohorte de 2014.

La Figura 17 detalla la proporción de estudiantes beneficiarios del SEA por carrera. Se observa que ingeniería metalúrgica fue la carrera con mayor proporción de estudiantes beneficiados en asignaturas de matemática (41%), seguido de matemáticas y química con 36%. Por otro lado, la carrera con menos proporción de beneficiados fue ingeniería electrónica (18%) seguida de ingeniería Civil (18%) y petróleo (21%). Nótese que tomando como referencia el puntaje de las pruebas Saber en matemáticas (Figura 14 y Figura 15), es coherente que carreras como Electrónica, Civil y Petróleo teniendo puntajes altos hayan sido las de menor cobertura del programa SEA y

que Matemáticas siendo la de peor puntaje haya sido la de mayor cobertura; sin embargo este patrón no es regular, así por ejemplo carreras como Ingeniería Química y Licenciatura en Matemáticas que tuvieron puntajes altos y bajos en las pruebas Saber por el contrario recibieron una cobertura baja y alta en el programa SEA respectivamente.

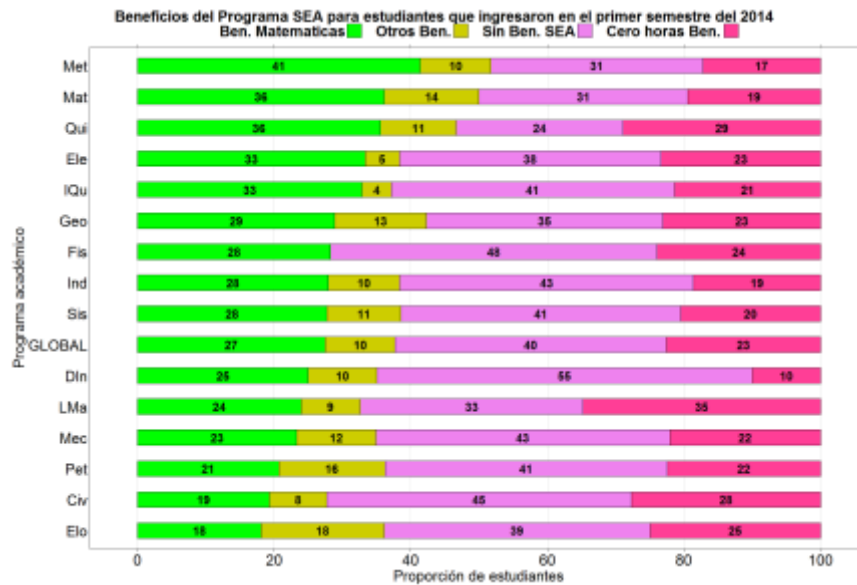


Figura 17. Proporción de Beneficiarios del programa SEA por carrera, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

La Figura 19 muestra la distribución de número de estudiantes beneficiados por asignatura y por carrera. Se observa que, de las asignaturas de matemáticas, Cálculo I es la de mayor demanda, 146 beneficiados, equivalente a un 16%; seguida de cálculo II con 95 beneficiados (10%), cálculo III con 42 (5%), álgebra con 37 (4%) y ecuaciones diferenciales con 16 estudiantes (2%). Este decrecimiento en la distribución de estudiantes del programa en las materias de cálculo, que son materias secuenciales en el ciclo básico (requisito cada una de la siguiente), está en concordancia con el comportamiento también decreciente de la proporción de estudiantes PFU con el avance de los semestres (Figura 18) y el aumento de nivel académico medio que supone la población de estudiantes que permanece (Figura 7).

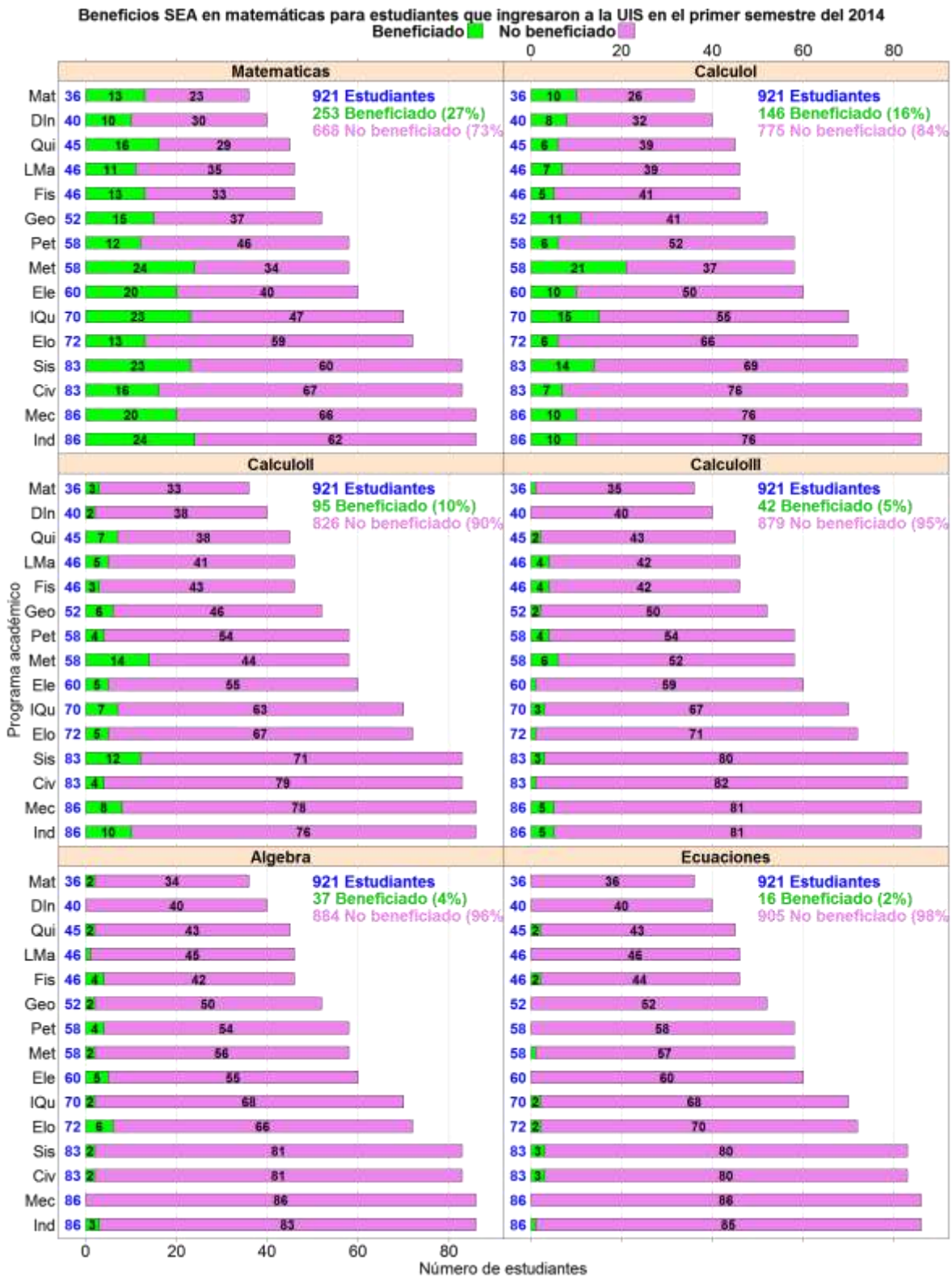


Figura 19. Beneficiarios del programa SEA por carrera y asignatura, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

En la distribución por carreras también es evidente la reducción en la cobertura del SEA en las materias de semestres elevados como ecuaciones diferenciales donde solo 8 carreras de las 15 de ciencias e ingeniería tuvieron beneficio del SEA con máximo 3 estudiantes beneficiados.

La Figura 20 brinda la misma información, pero centrando la atención en el porcentaje y no en la cantidad de estudiantes por carrera, incluyendo además el porcentaje global en cada asignatura.

Se observa que la mayoría de los beneficios en asignaturas de matemáticas se distribuyen en Cálculo I y II (porcentaje Global) mientras que las demás asignaturas no sobrepasan el 5%.

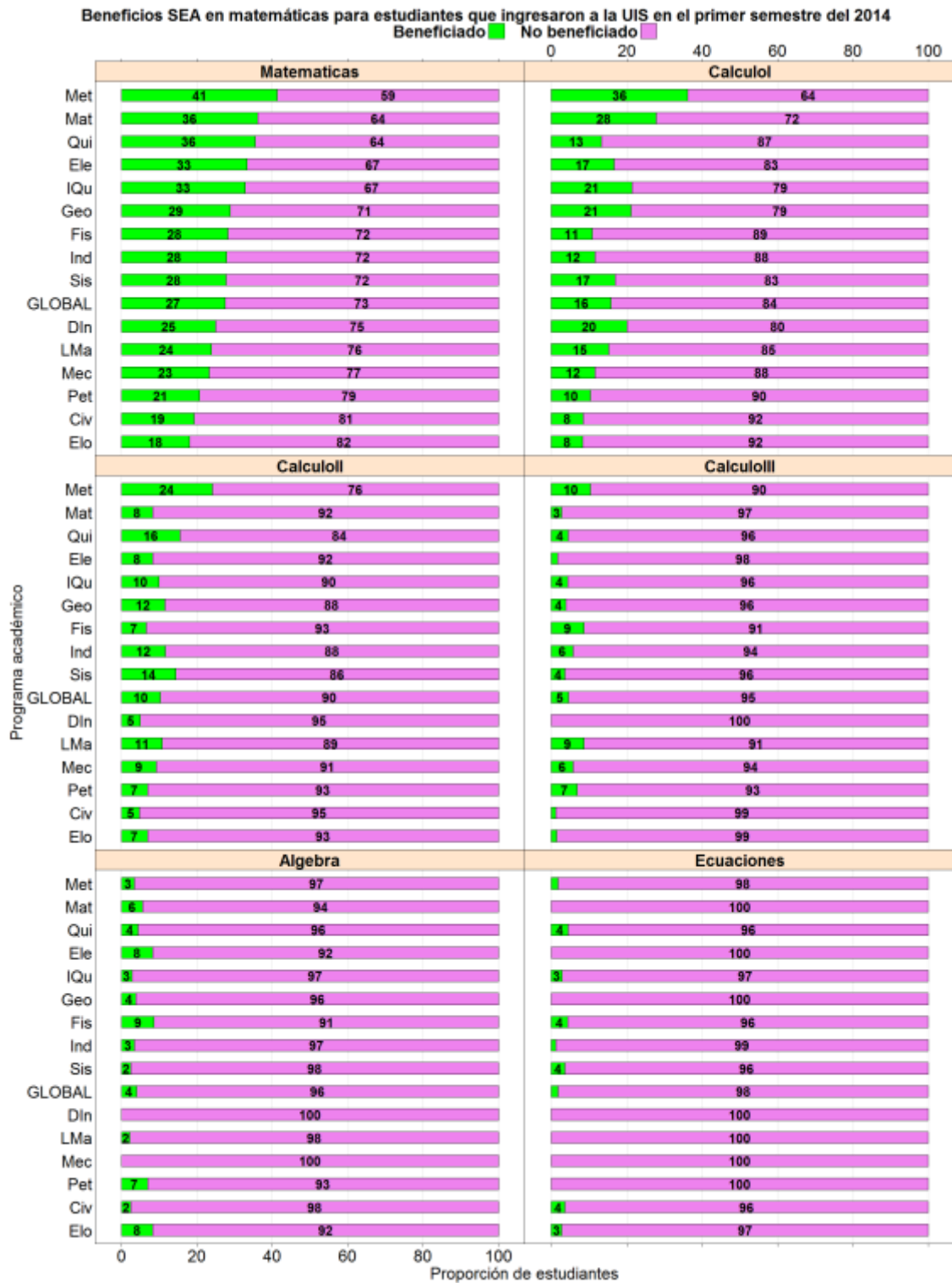


Figura 20. Proporción de Beneficiarios del programa SEA por carrera y asignatura, Estudiantes de Ciencias e Ingeniería, 1ª cohorte de 2014.

En las Figura 21 y Figura 22 se muestra la distribución de estudiantes en cada sub-programa del SEA por carrera y a nivel global. Se aprecia que el programa PAMRA fue el que tuvo mayor cantidad de estudiantes beneficiados respecto al total de estudiantes de la cohorte (25%), seguido de MIDAS (19%), ASAE (4%) y FPC (1%); las carreras con mayor proporción de beneficiados en PAMRA fueron Matemáticas (47%), ingeniería metalúrgica (40%) y licenciatura en matemáticas (30%) mientras que las de menor beneficio fueron física (11%) e ingeniería mecánica (13%). Por otro lado, en el segundo programa de mayor cobertura, MIDAS, la cobertura por carreras cambia considerablemente encontrándose que la carrera con mayor proporción de beneficios fue ingeniería metalúrgica (34%) seguida de química (31%), mientras la de menor cobertura fue matemáticas (3%).

Los programas ASAE y FPC tienen una cobertura notablemente menor, siendo las carrera con mayor cobertura diseño industrial y mecánica, con tan solo 6% y 7% respectivamente; así mismo FPC presenta solo cobertura en tres carreras de Ciencias e Ingenierías.

Las diferencias en cobertura de cada subprograma pueden estar asociadas a su evolución y madurez individual, y a un enfoque específico en materias o facultades; no es alcance de este estudio verificar el efecto individual de los subprogramas SEA, ya que las proporciones individuales de beneficiarios de los subprogramas son muy pequeñas y dificultan las aplicación de técnicas de evaluación del impacto.

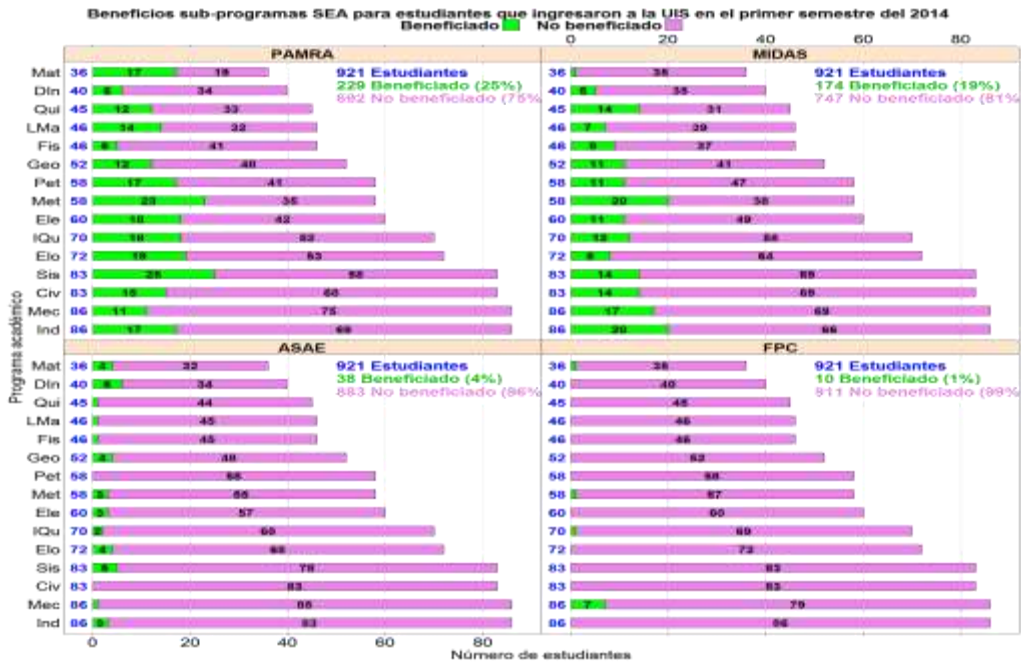


Figura 21. Cantidad de Beneficiarios SEA por sub-programa y carrera, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

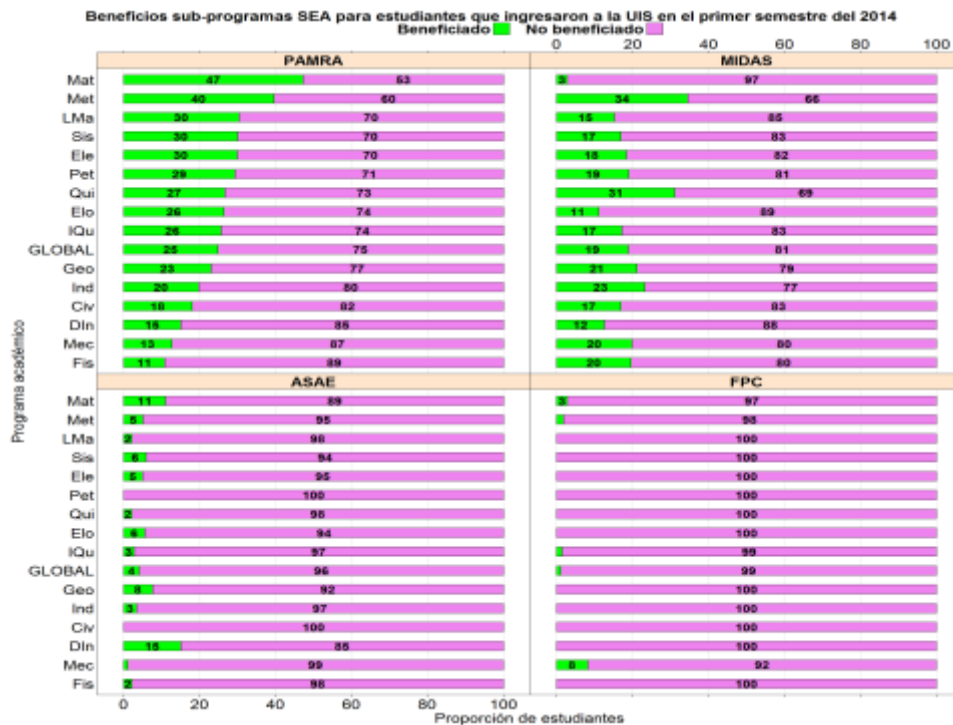
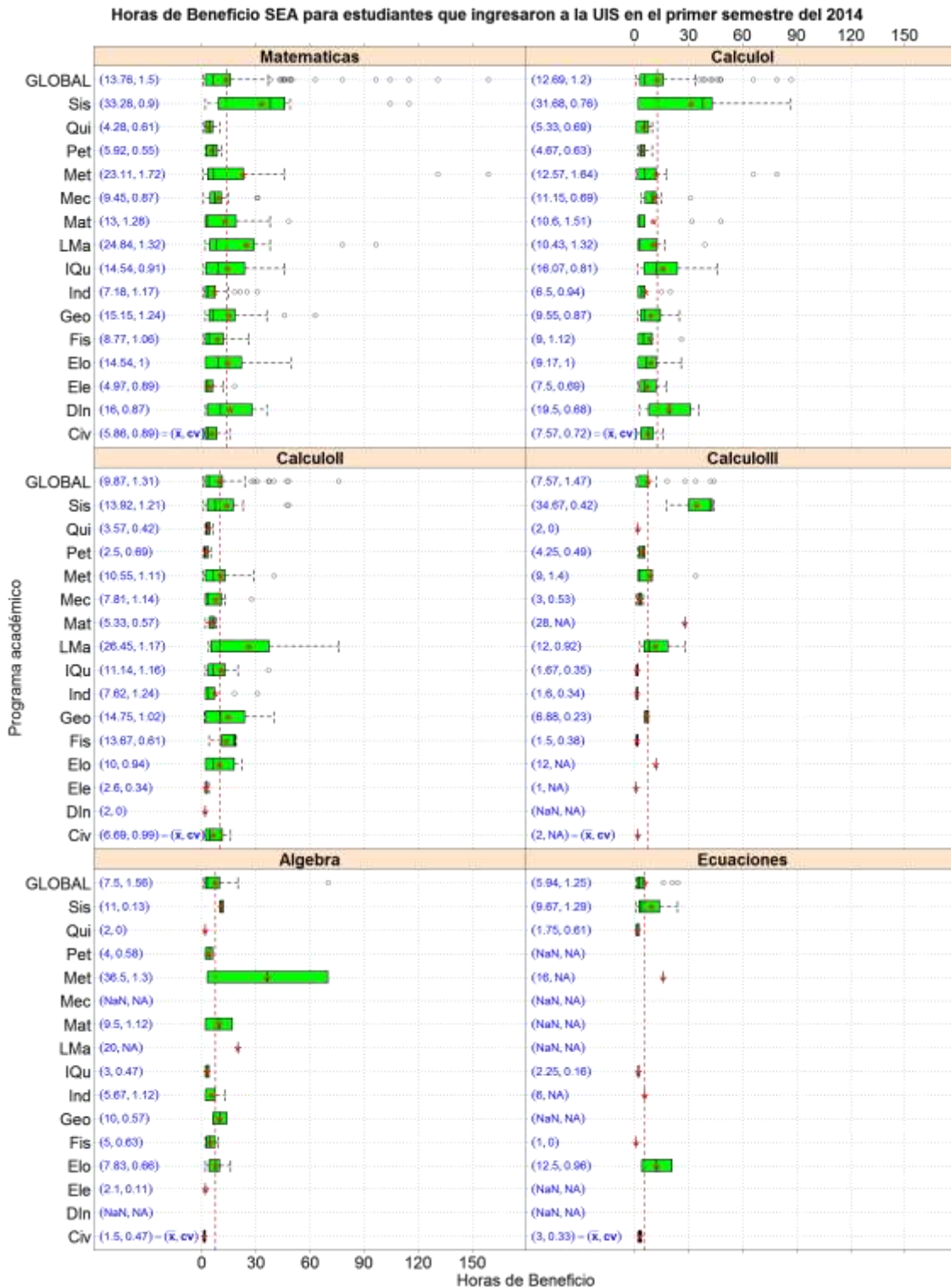


Figura 22. Proporción de Beneficiarios SEA por sub-programa y carrera, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

La Figura 23 muestra el diagrama de cajas de la intensidad horaria de beneficios recibidos por estudiantes de Ciencias e Ingeniería que ingresaron a la UIS en el primer semestre del 2014, globalmente en las asignaturas de matemáticas y de forma específica en cálculo I, II y III, álgebra y ecuaciones diferenciales.

Se observa que en promedio, los estudiantes beneficiados reportan solo aproximadamente 14 horas de asistencia al SEA en asignaturas de matemáticas, aunque con una distribución sesgada hacia la izquierda, donde se concentra la mayoría de los estudiantes; esto significa que en promedio la mayoría de los estudiantes (75%) tuvieron una intensidad de menos de 15 horas en este semestre (equivalentes a una hora de asistencia por semana-semestre).

Se observa que algunos estudiantes que recibieron un número de horas atípicamente alto a nivel global (más de 60 horas) en carreras como ingeniería metalúrgica, sistemas y licenciatura en matemáticas, aunque también específicamente solo en la asignatura de Cálculo I.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 23. Intensidad de beneficios del programa SEA por carrera en matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

Se destaca que Cálculo I fue la asignatura de mayor cobertura por intensidad horaria del programa SEA, aproximadamente 13 horas semestrales en promedio, valor muy cercano al promedio global de 14 horas. Las Figura 24 y Figura 25 muestran el número de beneficios por sub-programa y asignatura del SEA para los estudiantes de ciencias e ingeniería. Nótese que la asignatura de mayor cobertura por número de estudiantes fue Cálculo II con 201 estudiantes y no Cálculo I que le sigue con 190 estudiantes. Las demás asignaturas muestran una aplicación mucho más baja, con 52, 44 y 16 estudiantes para Cálculo III, Álgebra y Ecuaciones diferenciales respectivamente. El subprograma de mayor cobertura por número de estudiantes en las materias de primer semestre es PAMRA (Cálculo I y Álgebra), MIDAS en las materias posteriores (Cálculo II y III, y Ecuaciones Diferenciales) y ASAE se encuentra casi completamente enfocado en Cálculo I.

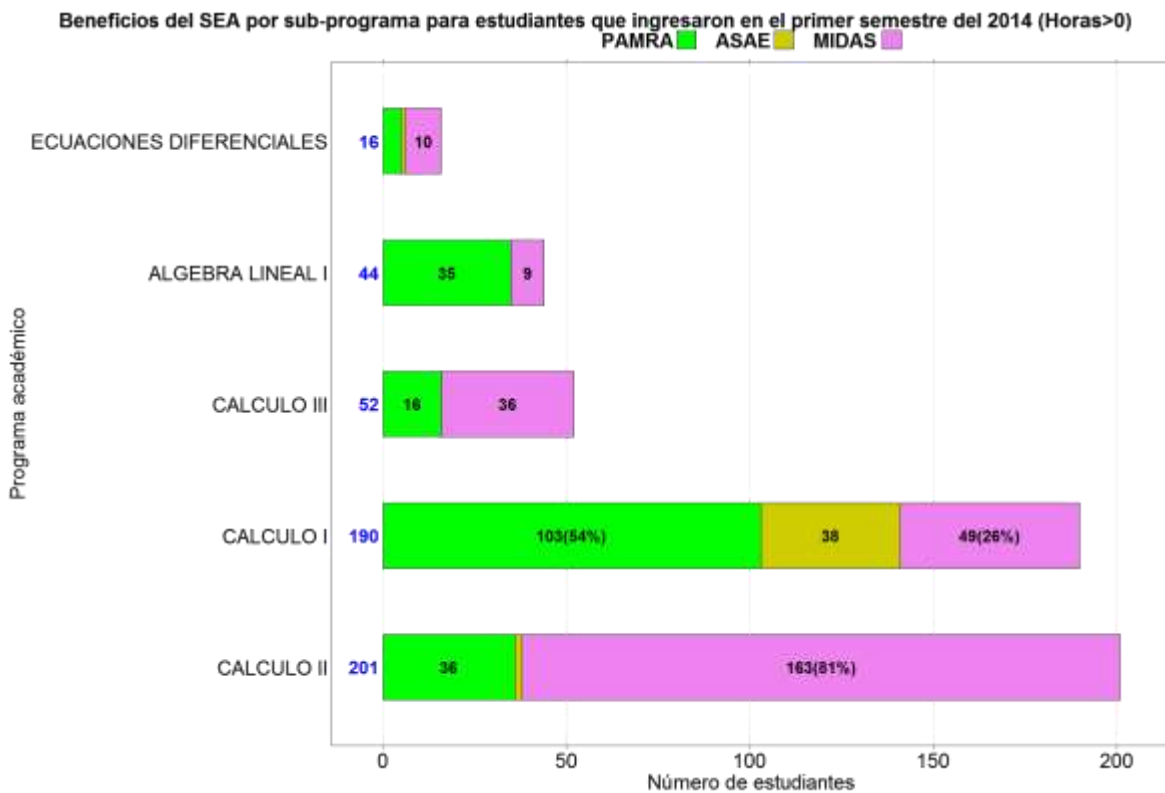


Figura 24. Beneficiarios del programa SEA en Matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014

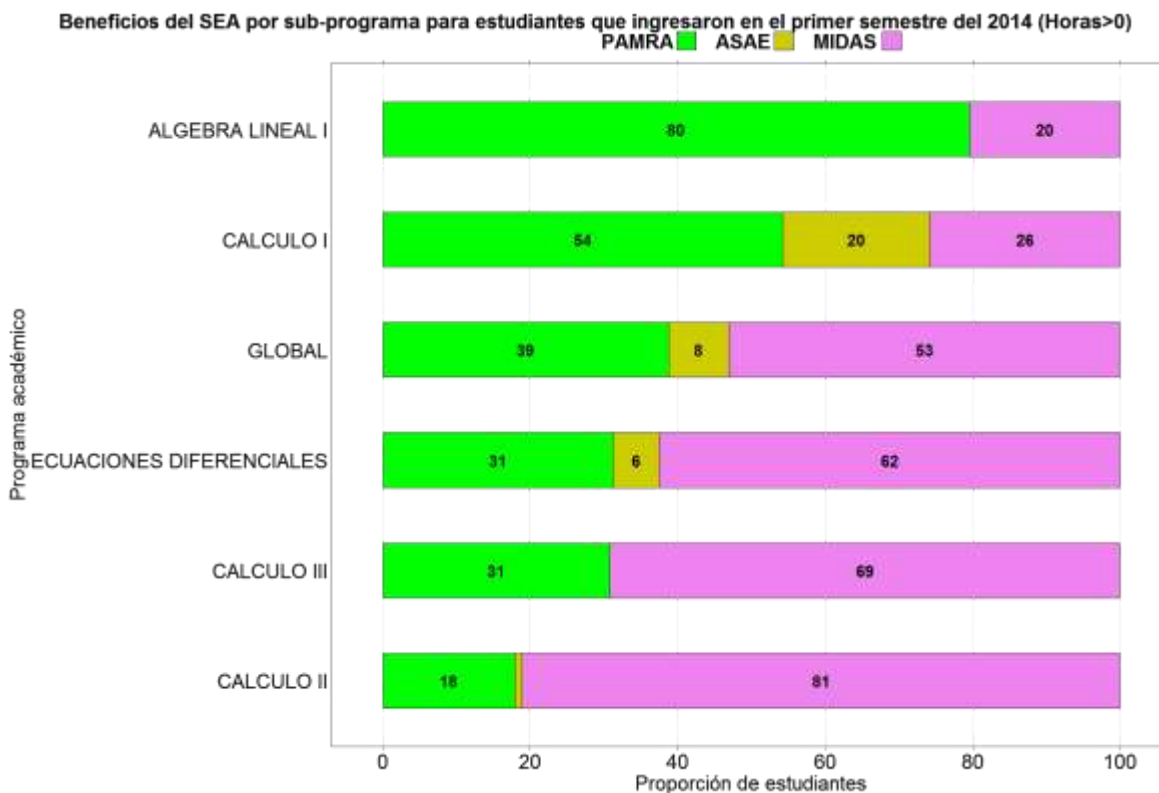


Figura 25. Proporción de Beneficiarios por Sub-programa SEA en matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte de 2014.

En la Tabla 7 se resumen los estadísticos de horas de beneficio por estudiante en cada asignatura. Es evidente que el apoyo académico del SEA ha sido focalizado en la asignaturas de Cálculo I y II, esto coincide con el hecho de que estas son las materias de peor desempeño académico para la primera cohorte del 2014 en matemáticas (Figura 11).

Tabla 7. Intensidad global de beneficios SEA en matemáticas, Estudiantes de Ciencias e Ingenierías, 1ª cohorte del 2014.

Área	Estadístico	Intensidad Beneficio SEA
Álgebra	$Cv$	1.56
	$\bar{x}$	7.50
	$\hat{x}$	3.00
Cálculo I	$Cv$	1.20

	$\bar{x}$	12.69
	$\hat{x}$	6.00
Cálculo II	$Cv$	1.31
	$\bar{x}$	9.87
	$\hat{x}$	4.00
	$Cv$	1.47
Cálculo III	$\bar{x}$	7.57
	$\hat{x}$	2.00
Ecuaciones	$Cv$	1.25
	$\bar{x}$	5.94
	$\hat{x}$	2.75
	$Cv$	1.50
Matemáticas	$\bar{x}$	13.76
	$\hat{x}$	6.00

**3.1.1.5 Caracterización Ingreso SEA.** Desde el primer periodo académico de 2013, el programa SEA ha venido realizando la caracterización de los estudiantes que son admitidos en la universidad para identificar aquellos que tienen algún tipo de vulnerabilidad en las dimensiones social, biosicosocial, económica, académica y cognitiva, que pueda llegar a afectar su desempeño académico. (Universidad Industrial de Santander SEA-UIS, 2014). Los estudiantes identificados con algún tipo de vulnerabilidad son invitados a participar, de manera voluntaria, en los diferentes programas de acompañamiento que ofrece el programa SEA.

En esta sección se presenta el análisis descriptivo de los resultados de estas pruebas de caracterización realizadas por el programa SEA para los estudiantes de ciencias e Ingeniería que ingresaron a la Universidad en el primer semestre del 2014. Para esto se cuenta con dos archivos de Excel con información complementaria y repetida que contiene los registros de los estudiantes con su código, programa, semestre y año de caracterización, en algunos casos, la condicionalidad en algún momento y los puntajes en las dimensiones social, biosicosocial, económica, académica y cognitiva junto con sus respectivas caracterizaciones de riesgo.

La revisión de los datos de caracterización muestra lo siguiente:

1. Las dos bases de datos suministradas contienen información repetida pero también complementaria. Es decir, no todos los estudiantes, objeto de la investigación en el presente estudio, se encuentran en las dos bases de datos suministradas. Por esta razón se consolidaron las dos bases de datos y posteriormente se eliminaron registros repetidos.
2. Después del proceso de consolidación de las dos bases de datos se encontraron estudiantes con el mismo código caracterizado dos veces, en periodos y programas distintos y con puntajes distintos. Esto se debe a estudiantes que reiniciaron otra carrera. Como ya se explicó anteriormente, lo que se hizo con estos casos fue tomar la caracterización SEA de estudiante con el mismo programa y periodo que se encontraba en la base de datos de notas. El número de estudiantes encontrados en esta situación fue de 106 estudiantes que corresponde al 11.5% de los estudiantes que ingresaron a la universidad en el primer semestre del 2014 (921 estudiantes).
3. No todos los estudiantes tienen todos los puntajes de caracterización en las 5 dimensiones. Es decir, faltan muchos datos; en la mayoría de los casos no se tiene la caracterización en ninguna dimensión. Sin embargo, en algunos casos hay caracterización parcial. En total se encontraron 39 registros de estudiantes que no tienen caracterización en ninguna dimensión.
4. Para cada dimensión se encuentran dos variables denominadas puntaje y el riesgo. Se encontró que el riesgo es una categorización de la variable numérica puntaje, como se resume en la siguiente tabla:

Tabla 8.

*Categorización de puntaje en variables de caracterización del SEA (excepto cognitiva)*

Puntaje		Riesgo
Min	Max	
0	0.25	ALTO
0.26	0.5	MEDIO ALTO
0.51	0.7	MEDIO
0.71	1	BAJO

5. En contraste con las demás dimensiones, en la cognitiva tanto el puntaje como el riesgo son variables categóricas con la siguiente relación:

Tabla 9.

*Categorización de puntaje en dimensión cognitiva programa SEA*

D. Cognitiva	
Puntaje	Riesgo
MUY ALTO	BAJO
ALTO	BAJO
MEDIO ALTO	BAJO
MEDIO	MEDIO
MEDIO BAJO	ALTO
BAJO	ALTO
MUY BAJO	ALTO
INDEFINIDO	INDEFINIDO

La Figura 26 y la Figura 27 muestran el número y porcentaje de estudiantes de ciencias e ingeniería que ingresaron en el primer semestre del 2014 y fueron caracterizados por el programa SEA en cada carrera. A nivel global, en la dimensión académica se encuentran la mayor proporción de estudiantes en riesgo alto (64%), 7% en riesgo bajo y 29% no fueron diagnosticados por el programa SEA (no registran puntaje en la base de datos entregada); las carreras con menor cobertura (menos caracterizados) fueron Física (54%) y Diseño industrial (48%); las carreras con mayor riesgo académico alto fueron ingeniería de sistemas (77%), licenciatura en matemáticas

(76%) e ingeniería eléctrica (72%). La segunda dimensión de mayor cantidad de estudiantes en riesgo alto a nivel global fue la cognitiva (26%), 36% en riesgo medio, 4% en bajo y 33% no evaluados; la carrera con mayor proporción de estudiantes en riesgo alto fue matemáticas (47%) mientras que la menor fue diseño industrial (15%).

Las demás dimensiones muestran una proporción de estudiantes con riesgo menor al 10%, por lo que se deduce que el riesgo académico y cognitivo deberían ser las dimensiones más relacionadas con el desempeño académico de la cohorte. Sin embargo, tomando especial atención en el riesgo académico, las carreras más y menos vulnerables no coinciden con las de mejor y peor desempeño en las asignaturas de matemáticas (sección 2.2), excepto para la carrera de matemáticas, correlación que si se encontró con los resultados de las pruebas SABER 11 en el área de matemáticas (sección 2.3).

La dimensión económica fue la de mayor cobertura (91% evaluados) con un 10% en riesgo alto, 29% en medio alto, 37% en medio y 15% en bajo. La carrera con mayor riesgo alto fue matemáticas (19%) en tanto que con menor proporción de estudiantes en riesgo bajo fue ingeniería de petróleos (29%). En la dimensión de salud, no se encuentran estudiantes con riesgo alto, 12% en riesgo medio, 78% en bajo y 10% no caracterizados. El 88% de los estudiantes de electrónica se encuentra en riesgo bajo y la carrera con mayor cantidad de estudiantes en riesgo medio fue geología (25%).

Finalmente en la dimensión social, 90% de los estudiantes fueron caracterizados: 5% en riesgo alto, 14% en medio alto, 30% en medio y 41% en bajo.

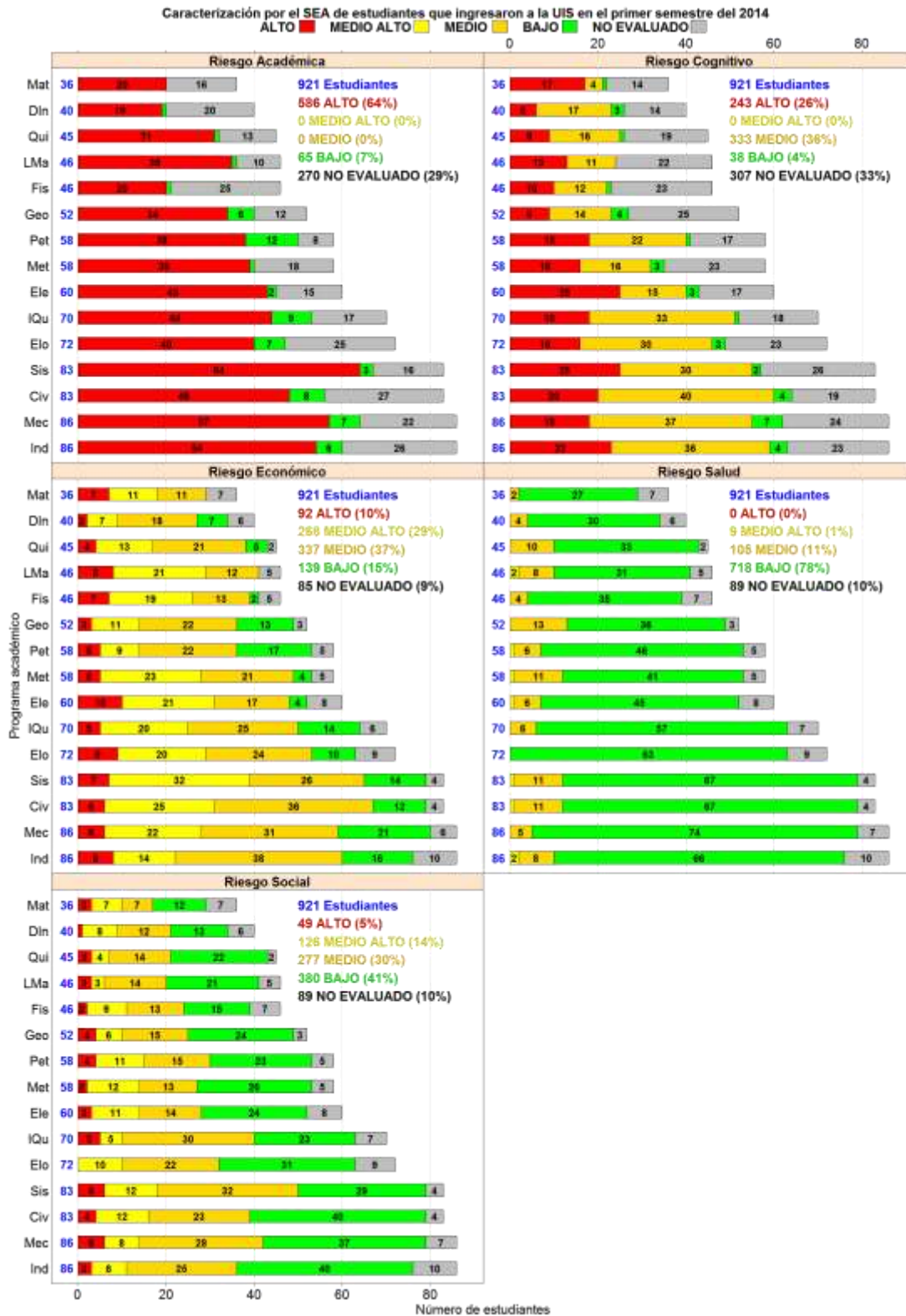


Figura 26. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera.

1ª cohorte de 2014

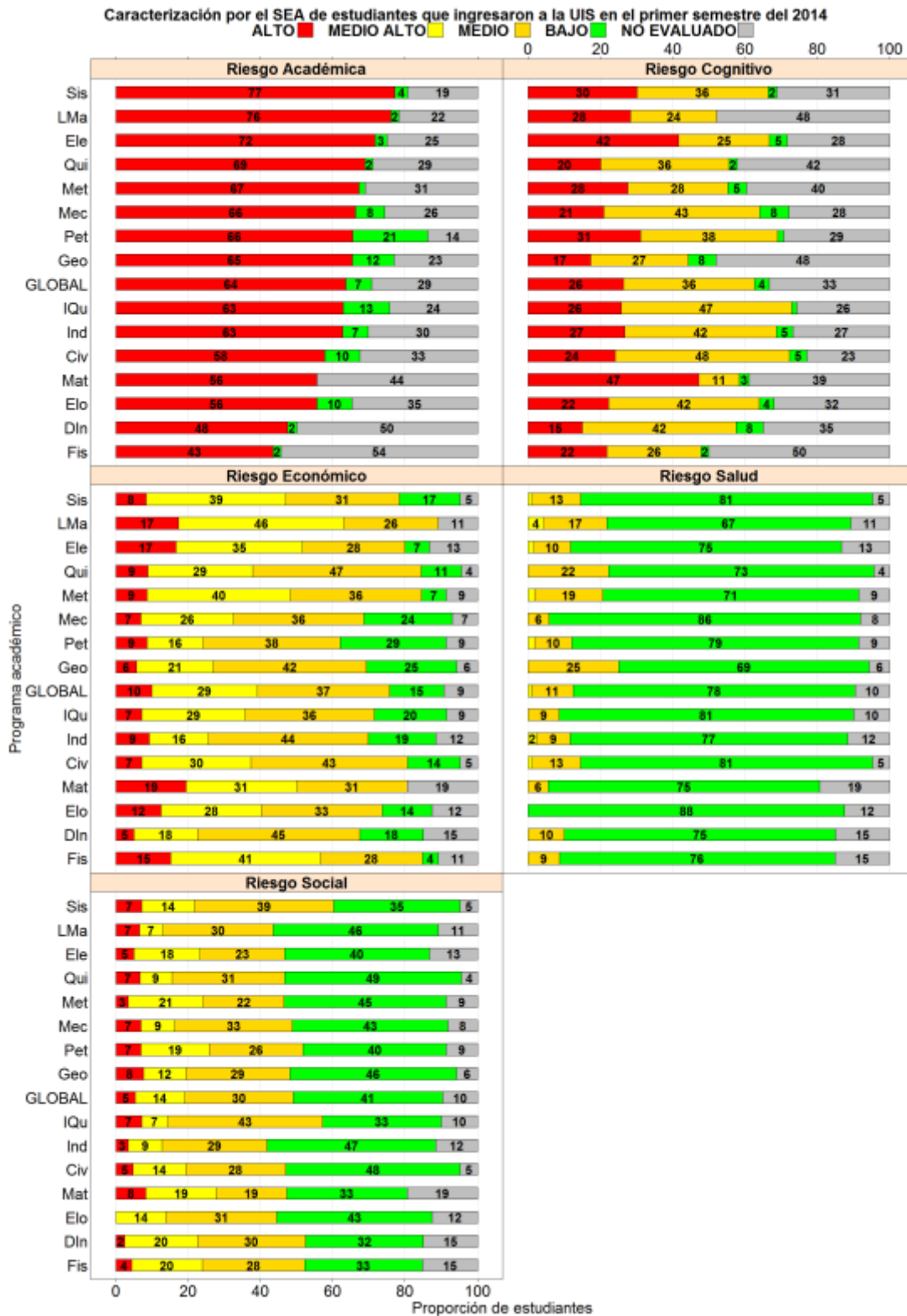


Figura 27. Proporción de Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera. 1ª cohorte de 2014

### ***3.1.1.6 Estudiantes Cohorte I – 2014 a 2017 / Semestre I***

#### ***3.1.1.7 Características Generales.***

- **Distribución por Género**

La Figura 28 muestra el número y porcentaje de estudiantes de ciencias e ingeniería que ingresaron a la universidad en la primera cohorte para los años 2014 a 2017 por género y carrera.

Es evidente que la proporción de género varió muy poco durante estos 4 años, manteniéndose la dominancia de los hombres con porcentajes entre el 68% y 70% respecto a las mujeres entre el 30% y el 32%. Resulta también notable que para cada programa académico las proporciones año a año también sean similares, lo que hace que prácticamente las mismas carreras se mantengan en el orden de preferencia por género; así por ejemplo en el extremo, Ingeniería Química (entre 48% y 66%) es la de mayor proporción de mujeres, mientras Física (entre 76% y 84%) es la de mayor proporción de hombres en todas las cohortes.

En conclusión no se observan diferencias de distribución significativas por género entre los estudiantes que ingresaron en la primera cohorte de los años 2014 a 2017.

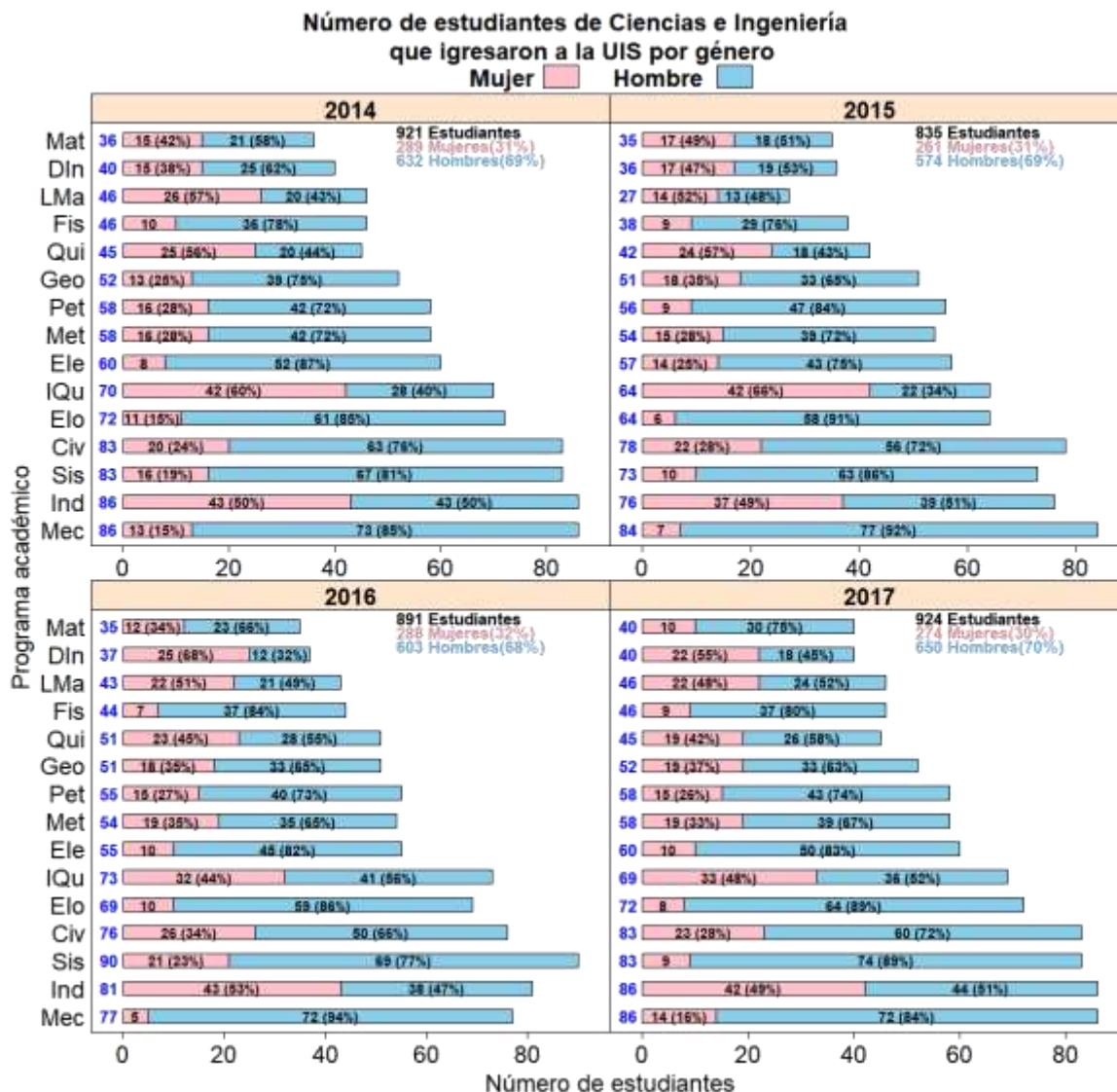


Figura 28. Distribución por género, Estudiantes de Ciencias e Ingeniería, primera cohorte 2014 a 2017.

- **Distribución por Edad**

La Figura 29 muestra el número y porcentaje de estudiantes de ciencias e ingeniería que ingresaron a la universidad en la primera cohorte para los años 2014 a 2017 por edad. Esta proporción se construye definiendo tres rangos de edad, considerando un primer rango de edad temprana ( $\leq 17$  años), de edad intermedia ( $\geq 18$  y  $\leq 20$  años) y de edad mayor ( $\geq 21$  años). Se observa una

disminución sostenida con el avance del tiempo de la proporción de estudiantes con edad intermedia (pasando de 60% a 47%) y a su vez un aumento sostenido de la proporción de estudiantes con edad temprana (pasando de 31% a 47%). De hecho en el 2017 la edad intermedia (la de mayor proporción de las demás cohortes), y la edad temprana igualan sus proporciones (47%). Por su parte la proporción de estudiantes con edad mayor (la de menor proporción en todas las cohortes) es fluctuante y no revela un comportamiento monótono, variando entre un 3% y 9%. En conclusión se observa una tendencia significativa de disminución en la edad de ingreso a la universidad para los estudiantes en la primera cohorte de los años 2014 a 2017.

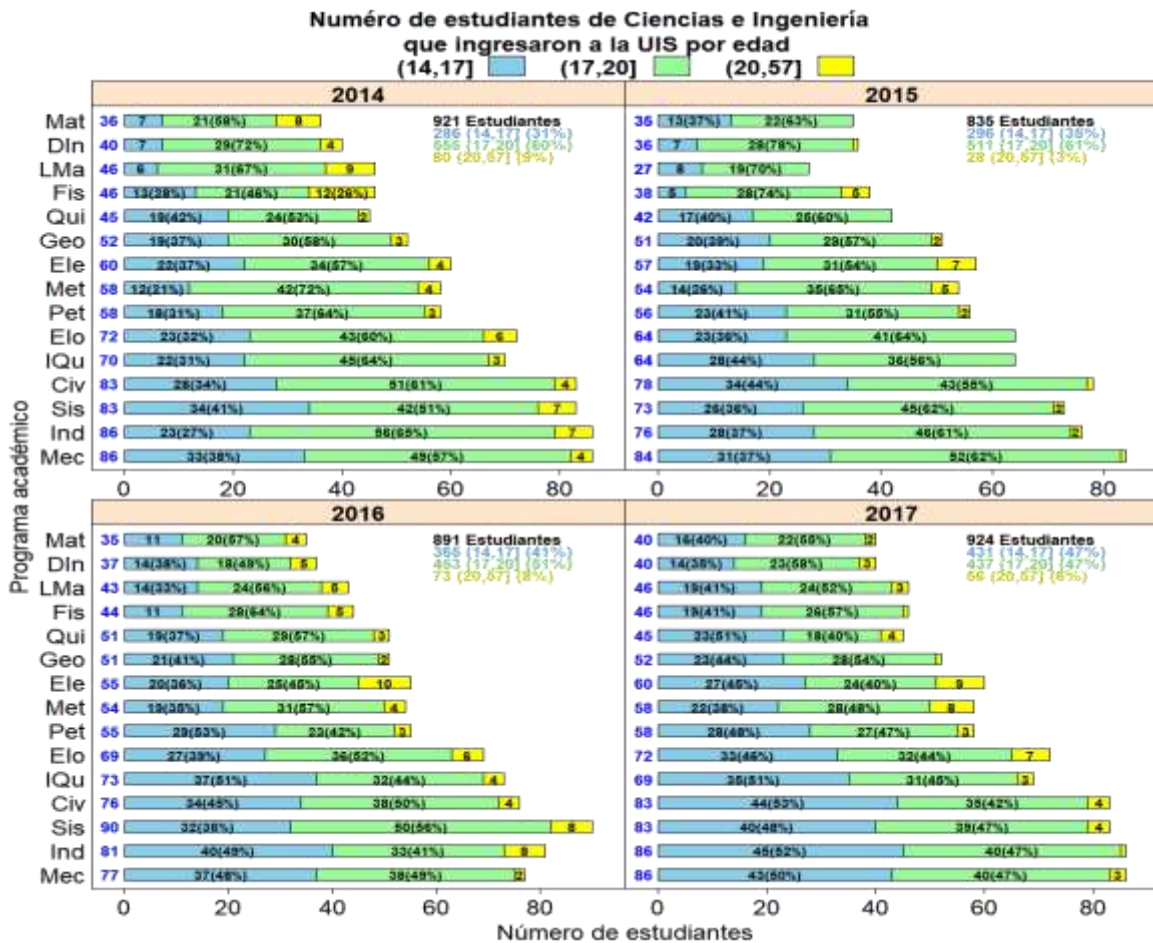


Figura 29. Distribución por edad, Estudiantes de Ciencias e Ingeniería, primera cohorte 2014 a 2017.

### 3.1.1.8 Desempeño Académico

- **Promedio Semestral**

En este apartado se presenta un análisis comparativo del promedio académico de los estudiantes de ciencias e ingeniería que ingresaron a la universidad en la primera cohorte para los años 2014 a 2017. La Tabla 10 resume los estadísticos del seguimiento a los estudiantes de la primera cohorte para el promedio académico semestral en los años analizados (media,  $\bar{x}$  mediana,  $\hat{x}$  y coeficiente de variación, Cv).

Tabla 10.

*Desempeño académico global, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.*

Cohorte	$\bar{x}$	$\hat{x}$	Cv
2014-1	3.23	3.36	0.26
2015-1	3.34	3.53	0.26
2016-1	3.38	3.50	0.22
2017-1	3.44	3.56	0.23

La Figura 30 detalla los resultados por carrera de la Tabla 3 empleando diagrama de cajas para el promedio semestral. Las carreras están ordenadas de forma descendente por su promedio para la primera cohorte (2014-1) y este orden se conserva en los demás semestres. En primer lugar se observa un aumento sostenido en el promedio Global pasando de 3.2 en el 2014 a 3.4 en el 2017, mientras la dispersión de los datos es prácticamente la misma en todas cohortes (entre 0.22 y 0.26). Los promedios de los programas en general también muestran un aumento con el tiempo, de forma que 7 programas tuvieron un promedio igual o inferior a 3.2 en el 2014, 3 en el 2015, 1 en el 2016

y 3 en el 2017. Resalta también que en todas las cohortes, los programas en el área de matemáticas tienen los promedios más bajos e ingenierías como metalúrgica y química los más altos.

La dispersión es muy similar en casi todos los programas a la dispersión global, con excepciones como la alta dispersión observada en Matemáticas y Licenciatura para el 2014 y 2015 respectivamente; además nuevamente las ingenierías metalúrgica y química resultan las de menor dispersión en su promedio.

Por último, se observan en general distribuciones no simétricas, aunque la media se encuentra relativamente cerca a la mediana en el resultado Global y en la mayoría de los programas.

En conclusión el desempeño académico en el primer semestre, de los estudiantes de Ciencia e Ingeniería de la primera cohorte entre el año 2014 y 2017, ha mejorado globalmente de forma sostenida, aunque el incremento promedio en 4 años promedio es de solo 0.2, es significativo.



Figura 30. Promedio semestral, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.

- **Estado de Condicionalidad.**

La Figura 31 muestra la distribución de estado de condicionalidad académica por programa del primer semestre de la primera cohorte entre los años 2014 y 2017, ilustrando las proporciones de estado normal (Verde), condicional (Amarillo) o PFU (Rojo) e Inactivo (Gris).

El estado de condicionalidad de los estudiantes de primer semestre, también ha venido mejorando; mostrando una reducción de un punto porcentual por año de estudiantes PFU para las cohortes I entre los años 2014 y 2016, y dos puntos porcentuales entre el 2016 y 2017; mientras la condicionalidad muestra un decrecimiento de 22% en el 2014 a 15% en el 2017.

Es extraño sin embargo, que de los estudiantes inscritos en cada cohorte, no hayan estudiantes inactivos en el 2015-I y 2016-I, como si se ve reflejado en 2014-I y 2017-I; los estudiantes inactivos son aquellos que no inician en el primer semestre en el que fueron aceptados, sea por aplazamiento temporal o decisión definitiva.

La distribución entre programas no parece tener un patrón continuo entre cohortes, y solo resalta la continua presencia de Ingeniería de sistemas entre los programas con más porcentaje de PFU en todas las cohortes.



- **Promedio Asignaturas de Matemáticas**

La Tabla 11 y Tabla 12 resume los estadísticos de las notas en asignaturas de matemáticas de los estudiantes de la primera cohorte del 2014 (media,  $\bar{x}$  mediana,  $\hat{x}$  y coeficiente de variación,  $Cv$ ).

La Figura 32 y Figura 33 detallan mediante un gráfico de cajas los resultados de la Tabla 5 para Cálculo y Álgebra respectivamente.

Tabla 11.

*Desempeño académico en Cálculo I, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.*

Cohorte	$\bar{x}$	$\hat{x}$	$Cv$
2014-1	2.61	2.7	0.41
2015-1	2.61	2.8	0.41
2016-1	2.63	2.6	0.36
2017-1	2.79	3.0	0.36

Tabla 12.

*Desempeño académico en Álgebra, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.*

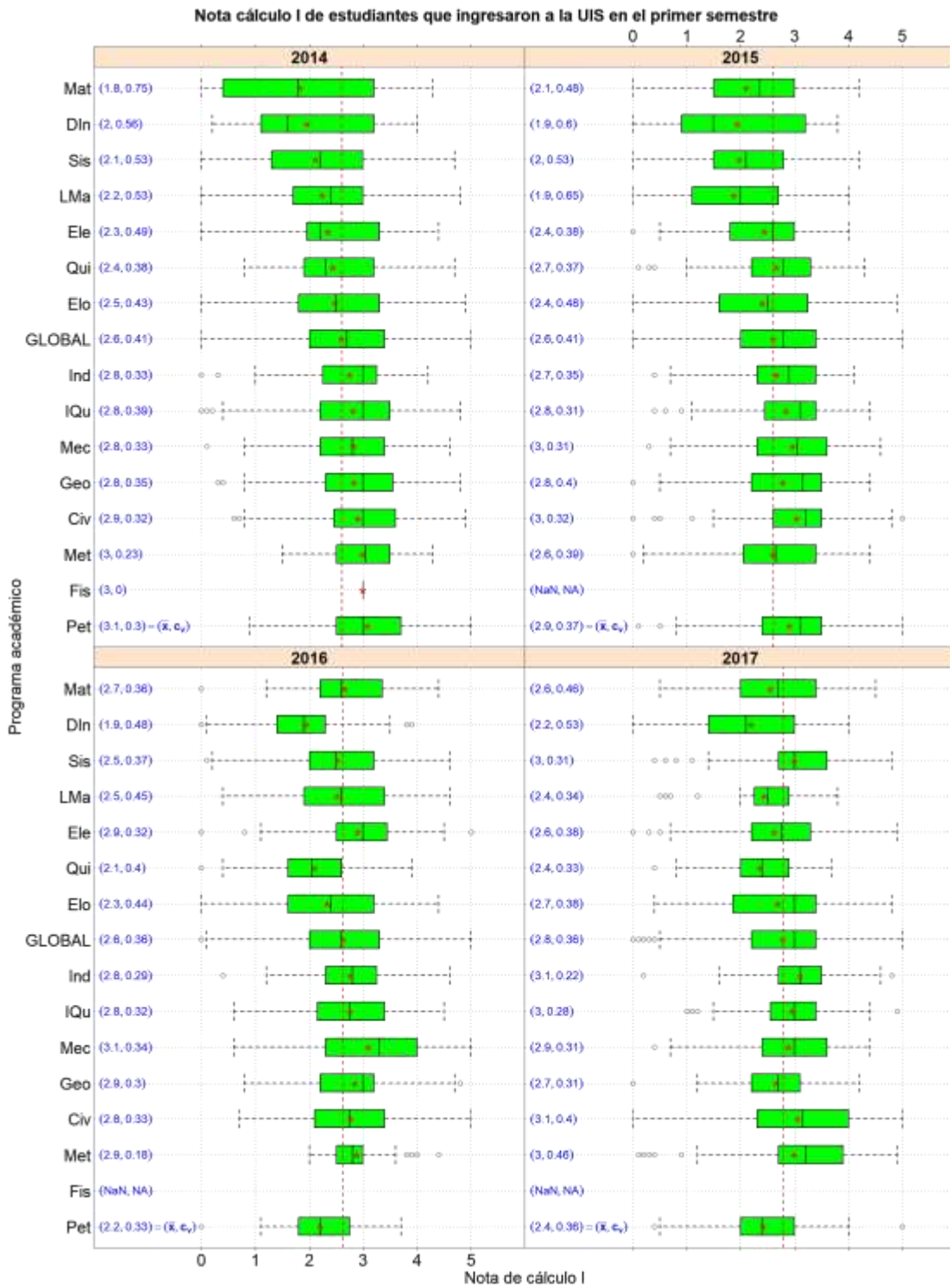
Cohorte	$\bar{x}$	$\hat{x}$	$Cv$
2014-1	2.96	3.2	0.33
2015-1	3.02	3.2	0.35
2016-1	3.14	3.3	0.31
2017-1	3.20	3.3	0.29

Se observa, al igual que en el promedio semestral, un aumento en el promedio global de las asignaturas de matemáticas de 0.2 entre el 2014 y el 2017; en cálculo se observa un promedio estable de 2.6 de 2014 a 2017, con un aumento notable a 2.8 en 2017; mientras en álgebra se observa un promedio estable de 3.0 en 2014 y 2015, con un aumento a 3.1 y 3.2 en 2016 y 2017, respectivamente. Los promedios de los programas presentan un comportamiento similar de

aumento en el promedio global; para cálculo solo 1 o 2 programas tuvieron un promedio superior o igual a 3.0 de 2014 a 2016, en comparación con 5 programas en 2017; para álgebra en 2014 y 2015, 6 y 7 programas tuvieron un promedio inferior a 3.0, en comparación con 2 y 4 programas en 2016 y 2017.

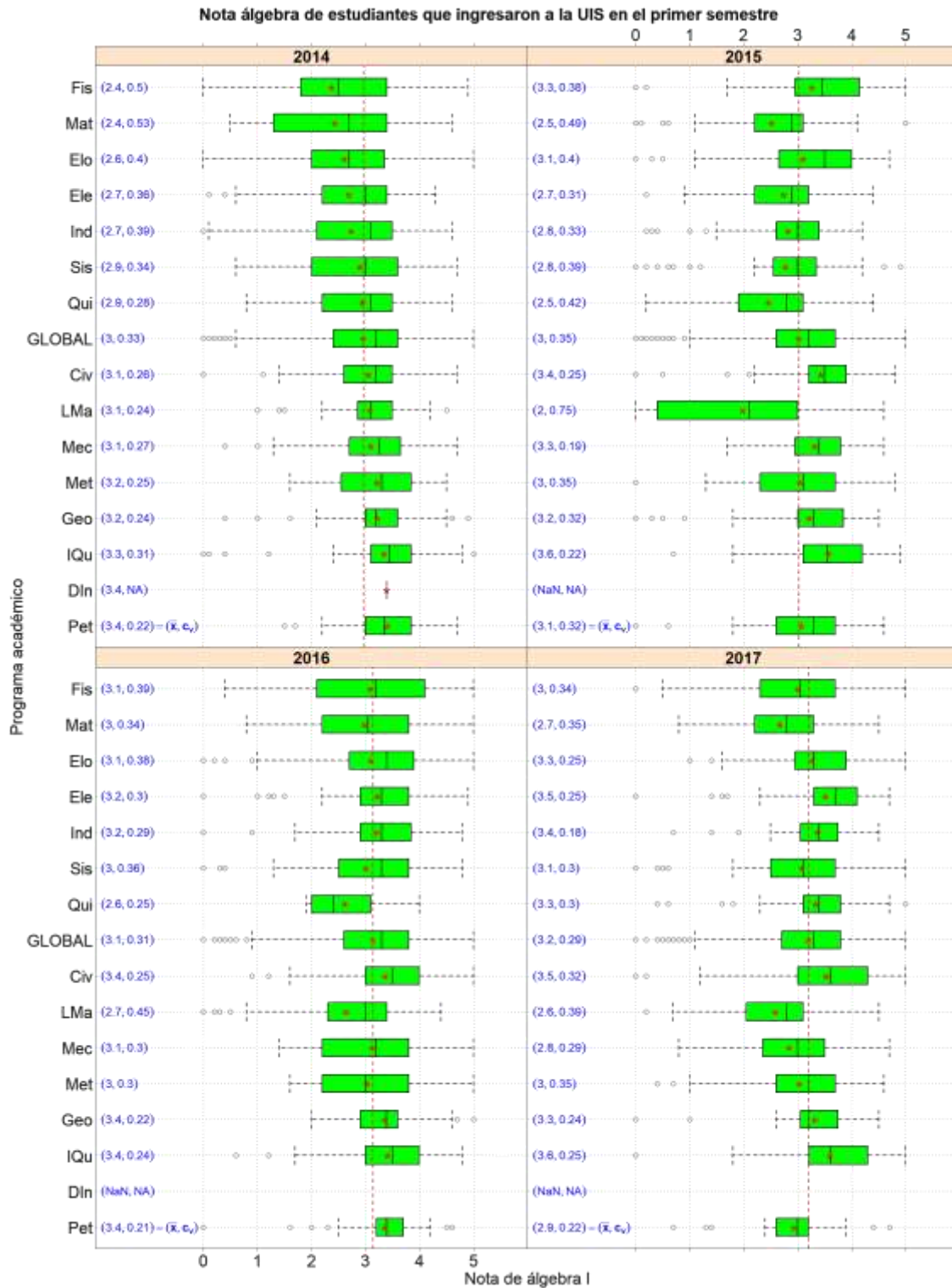
La dispersión de los datos es muy variable cohorte a cohorte para el promedio semestral entre programas, pero en global resulta muy similar (con Cv entre 0.29 a 0.35). Se observan en general distribuciones no simétricas, aunque la media se encuentra muy cerca a la mediana en el resultado Global y en la mayoría de los programas.

En conclusión el desempeño académico en las asignaturas de matemáticas en el primer semestre, de los estudiantes de Ciencia e Ingeniería de la primera cohorte entre el año 2014 y 2016 ha mejorado globalmente y notablemente en el 2017.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 32. Distribución de resultados para Cálculo I, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 33. Distribución de resultados para Álgebra, Estudiantes de Ciencias e Ingeniería, Semestre

1 - 1ª cohorte de 2014 a 2017.

**3.1.1.8 Pruebas Saber 11.** En esta sección se presenta el análisis descriptivo de los resultados en la pruebas SABER 11 de los estudiantes de ciencias e ingeniería que ingresaron a la universidad en la primera cohorte de los años 2014 a 2017.

Debido a los cambios estructurales de la prueba en 2015, se realizará únicamente la comparación en el área de matemáticas, la cual se conserva en dicha transición. La Tabla 13 resume los estadísticos de los puntajes en el área de matemáticas de las pruebas SABER 11 de los estudiantes (media,  $\bar{x}$  mediana,  $\hat{x}$  y coeficiente de variación, Cv). La Figura 34 (diagrama de cajas) muestra en detalle la información presentada en la Tabla 13.

Tabla 13.

*Resultados globales en Pruebas SABER 11, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014.*

Cohorte	$\bar{x}$	$\hat{x}$	Cv
2014-1	69.27	10.50	0.15
2015-1	75.78	9.23	0.12
2016-1	80.60	10.47	0.12
2017-1	75.62	7.10	0.10

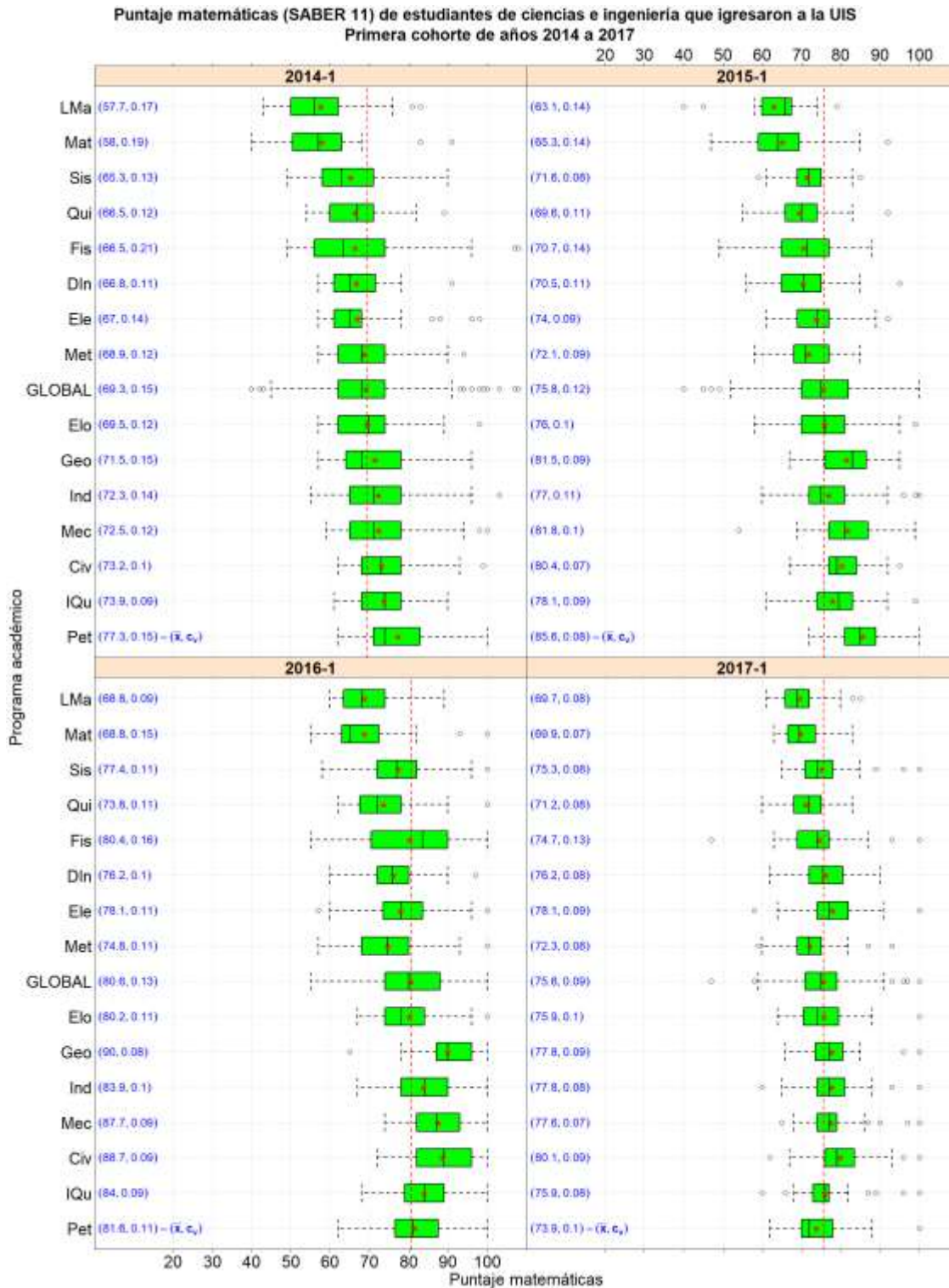
Se observa un aumento notable de 10 puntos en el promedio global entre 2014 y 2016, para luego disminuir en 5 puntos en 2017.

Este aumento en el desempeño de las pruebas Saber 11 en matemáticas coincide con el aumento de desempeño ya observado en las asignaturas de Cálculo y Álgebra en el primer semestre. También es interesante verificar que el escalafón de programas por puntaje en matemáticas es muy similar entre cada cohorte y coincide en buena medida con el escalafón observado en el promedio

semestral de los estudiantes (Figura 30), de forma que los programas de matemáticas vuelven a estar entre los puntajes más bajos e ingenierías como Química y Metalúrgica entre los más altos.

Lo anterior resulta más obvio, si se tiene en cuenta que la prueba de Saber 11 es el instrumento principal de selección para ingreso a la universidad y que cada carrera tiene distintos cortes de admisión que pueden variar entre otras por su demanda.

Por último, se observan en general distribuciones no simétricas, aunque la media se encuentra muy cerca a la mediana en el resultado Global y en la mayoría de los programas. En conclusión el puntaje en el área de matemáticas de la prueba saber 11 de los estudiantes de Ciencia e Ingeniería de la primera cohorte entre el año 2014 y 2017, ha incrementado mejorado de forma global.



Media programa ( $\bar{x}$ , \*), coeficiente de variación programa (cv), media universidad (---).

Figura 34. Distribución de resultados en pruebas Saber 11 en matemáticas, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.

**3.1.1.9 Beneficiarios Programa SEA.** En esta sección se presenta el análisis descriptivo de los beneficios recibidos específicamente en las asignaturas de matemáticas por los estudiantes de ciencias e ingeniería que ingresaron a la universidad en la primera corte de los años 2014 a 2017. En la base datos final de cada cohorte existen estudiantes no beneficiados dado que no fueron reportados en la base de datos original de beneficios del SEA y estudiantes que si fueron reportados pero con cero horas de beneficio; se asume que estas dos categorías corresponden a estudiantes en la misma condición que no tuvieron contacto con el programa SEA. La Tabla 14 y Tabla 15 resume los estadísticos de horas de beneficio recibidas por los estudiante en el programa SEA para las asignaturas de Cálculo I y Álgebra (media,  $\bar{x}$  mediana,  $\hat{x}$  y coeficiente de variación, Cv). Mientras las Figura 37 y Figura 38 detallan la distribución por horas de beneficio por programa y en global de la universidad para las asignaturas.

Los promedios de horas recibidas en Calculo I (Tabla 14 y Figura 35) por estudiante participante del programa SEA, muestran un comportamiento fluctuante (no monótono), entre los años 2014 y 2017, sin embargo la dispersión claramente se ha venido reduciendo, lo que refleja probablemente la evolución del programa SEA en su estructuración, manejo de la cobertura y aseguramiento de participación de los estudiante; este aspecto es más evidente en el 2017, cuya distribución global ya no se encuentra tan sesgada a la izquierda como en los años anteriores, no presenta valores extremos y las medias por programa son más homogéneas con fluctuación entre 14.2 a 21.9 horas, fuera de los valores promedio mínimo y máximo de Diseño Industrial (8.7) y Sistemas (24.1). No obstante lo anterior, valores de participación medios de máximo 18 horas (2015 y 2017) por semestre por estudiante logrados por SEA parecen relativamente bajos, dado

que para un semestre estándar (16 semanas) equivaldrían a una asistencia de 1.1 horas por semana por estudiante.

Por otra parte los promedios de horas recibidas en Álgebra (Tabla 15 y Figura 36) muestran una tendencia sostenida de crecimiento en intensidad horaria, pasando de 8 a 11 horas promedio por estudiante entre el 2014 y el 2017; en la máxima intensidad promedio observada de 11 horas por semestre por estudiante (2017) equivale a una intensidad en un semestre estándar (16 semanas) de 0.7 horas por semana por estudiante, lo que parece un valor bajo. La dispersión global y por carrera también disminuye significativamente con el tiempo, de forma que en el 2017, el promedio global y por carreras son más similares y con menos valores extremos, en contraste con el 2014, en donde se observan carreras con participación baja o puntual en contraste de otras con alta participación. Nuevamente esta evolución puede ser producto de la maduración del SEA en su estructura para el aseguramiento de la cobertura objetivo.

La intensidad horaria de beneficios por estudiante son naturalmente una variable clave analizar en relación al efecto del programa SEA, el análisis de impacto del siguiente capítulo proporcionará información valiosa sobre la influencia de esta intensidad horaria en el desempeño académico.

Tabla 14.

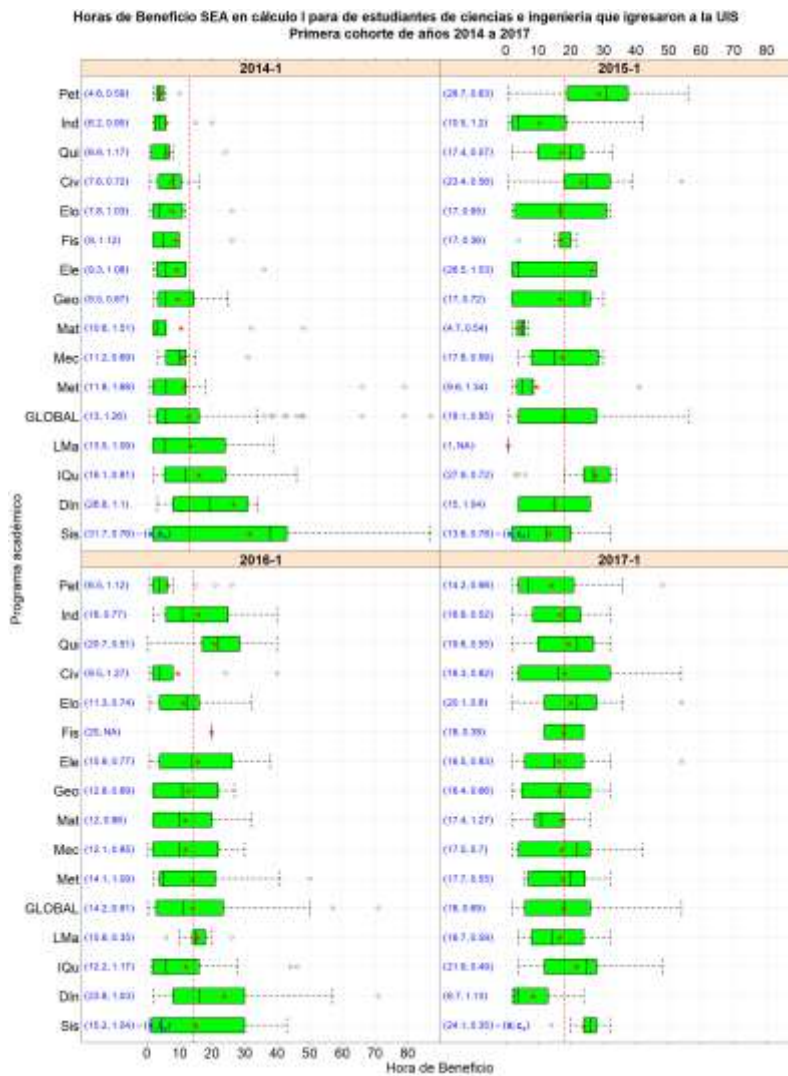
*Resultados globales en Horas de Beneficio Cálculo I, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014*

Cohorte	$\bar{x}$	$\hat{x}$	$Cv$
2014-1	12.97	6.00	1.26
2015-1	18.10	18.00	0.85
2016-1	14.15	11.25	0.90
2017-1	17.96	18.00	0.68

Tabla 15.

Resultados globales en Horas de Beneficio Álgebra, Estudiantes de Ciencias e Ingeniería, Primera cohorte del 2014

Cohorte	$\bar{x}$	$\hat{x}$	$Cv$
2014-1	7.93	3.50	1.49
2015-1	6.14	2.00	1.29
2016-1	10.06	7.50	1.10
2017-1	11.41	8.00	0.74



Media programa ( $\bar{x}$ , \*) desviación típica programa ( $Cv$ ), media universidad (---).

Figura 35. Intensidad de Beneficios en Cálculo I del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.

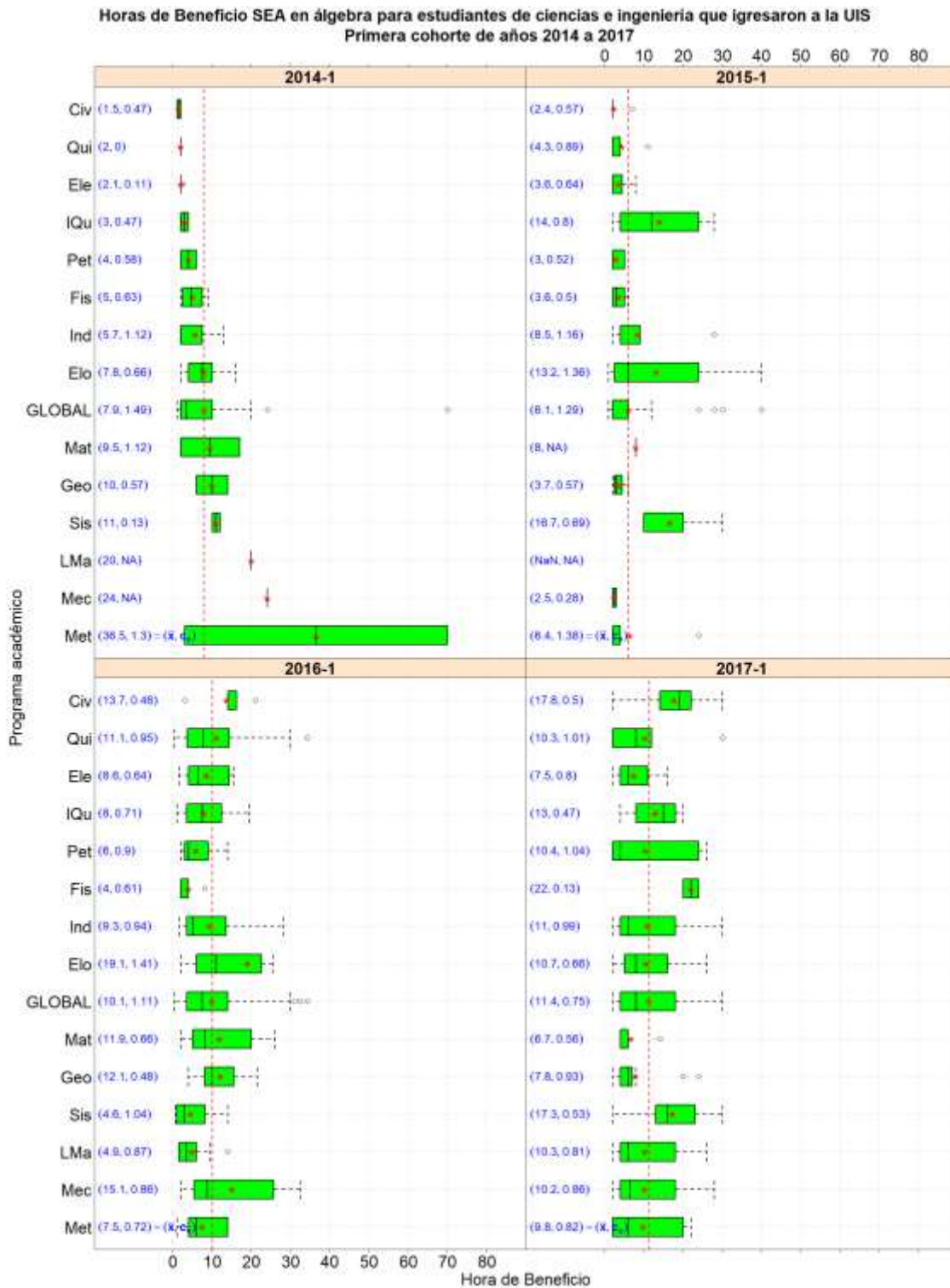


Figura 36. Intensidad de Beneficios en Álgebra del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.

En la Figura 37 y Figura 38 se ilustran el número y proporción de estudiantes beneficiados por el programa SEA para las asignaturas de Álgebra y Cálculo I. Estas figuras complementan la información ya analizada, mostrando un aumento creciente del porcentaje de participación de los estudiantes en ambas materias entre el 2014 y 2017, aumentándose la cobertura en este periodo de un 17% a 28% en cálculo y de un 4% a un 14% en álgebra. En conclusión el programa SEA ha venido aumentando su cobertura en beneficios en el área de matemáticas y a su vez ha mejorado la asistencia de horas promedio por estudiante por semestre.

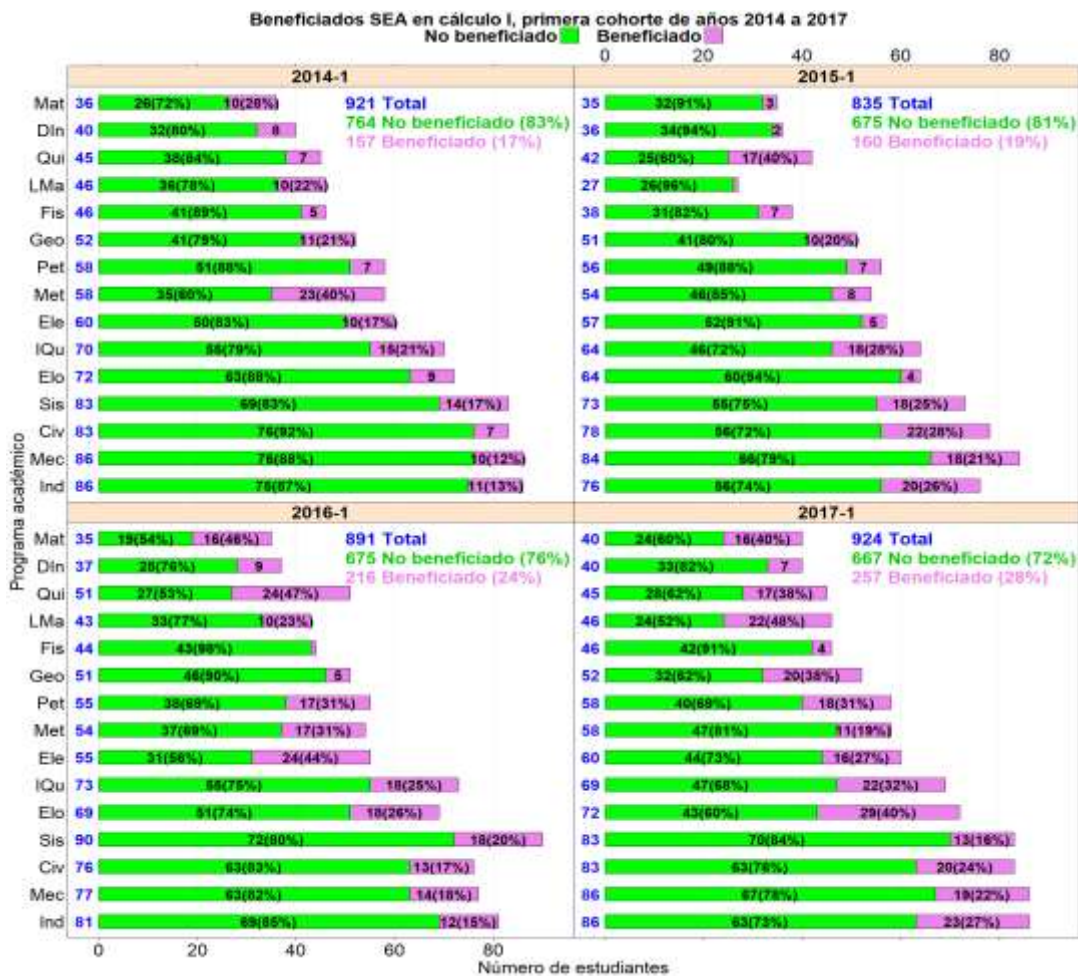


Figura 37. Beneficiarios en Cálculo I del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.

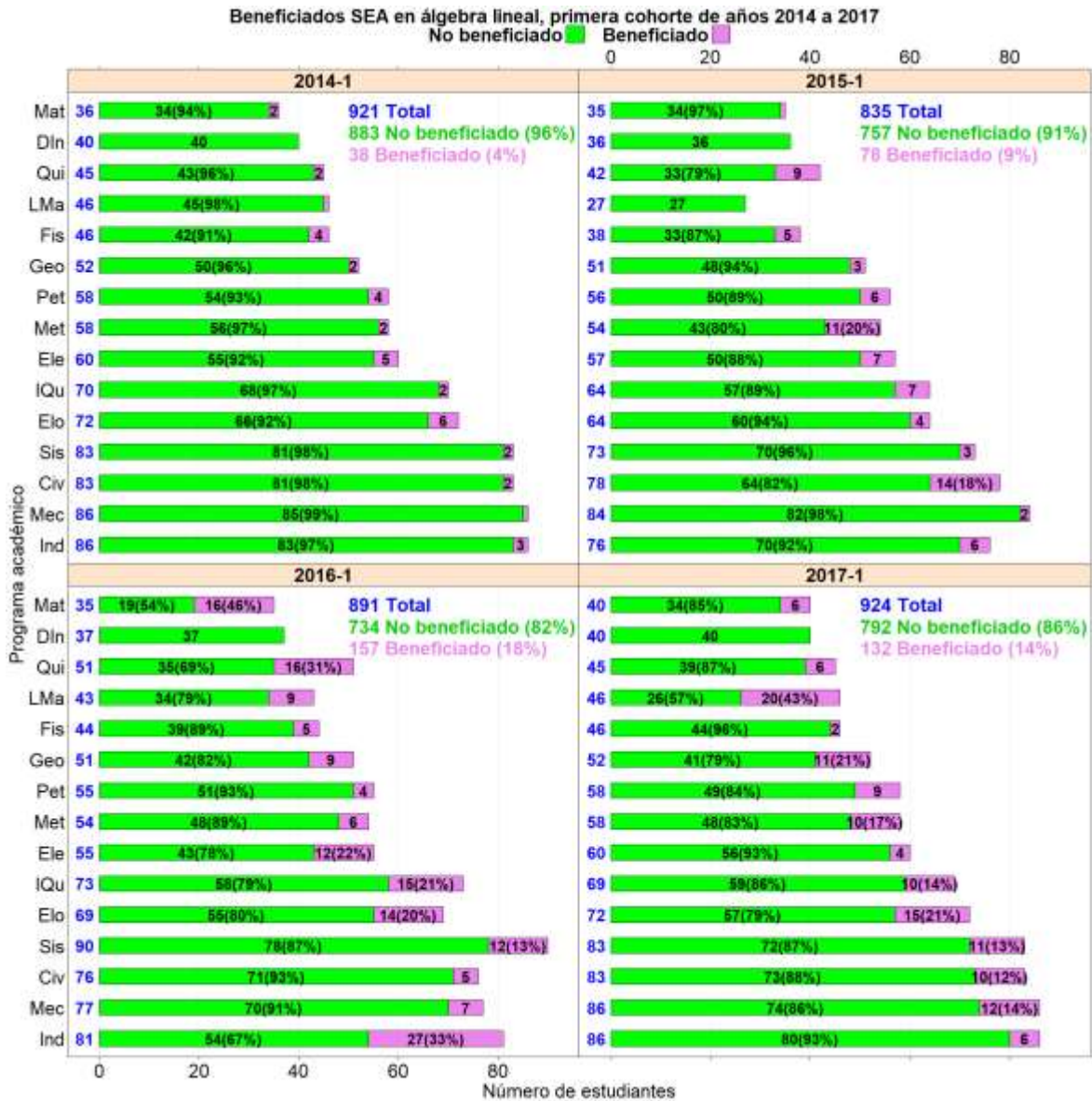


Figura 38. Intensidad de Beneficios en Cálculo I del programa SEA por carrera, Estudiantes de Ciencias e Ingeniería, Semestre 1 - 1ª cohorte de 2014 a 2017.

**3.1.1.10 Caracterización Ingreso SEA.** En la Figura 39 a Figura 42, se ilustran los resultados de la caracterización de ingreso de los estudiantes realizada por el programa SEA para la primera cohorte entre los años 2014 y 2017. Cada dimensión es usada por el SEA para realizar ajustes a la planeación de los subprogramas y los apoyos que de forma estructurada brindará a los estudiantes

por medio de subprogramas específicos. En particular la población con riesgo categorizado como “Alto” en cada dimensión evaluada es el principal objetivo del SEA. Tomando como referencia las categorías en riesgo “Alto” y “No evaluado” en cada dimensión de cada cohorte, se resume en la **¡Error! No se encuentra el origen de la referencia.** la evolución temporal en estos aspectos. Se evidencia que el SEA ha mejorado la cobertura de esta caracterización realizada a los estudiantes nuevos, pasando de un 18% a solo un 8% de estudiantes no evaluados entre el 2014-I y 2017-I. En proporción a los estudiantes identificados en riesgo alto, las dimensiones muestran el siguiente orden de importancia, Académico, Cognitivo, Económico, Social y Salud. Es evidente que el riesgo académico es dominante sobre las demás dimensiones, y solo para la cohorte 2015-I la dimensión económica resulta la de mayor proporción. Nótese también que el 2017-I tiene el menor promedio de estudiantes en riesgo alto de todas las cohortes analizadas.

Tabla 16.

*Resultados globales de caracterización SEA, Semestre 1 - 1ª cohorte de 2014 a 2017.*

<b>% Estudiantes en Riesgo Alto</b>						
<b>Cohorte / Riesgo</b>	<b>Académico</b>	<b>Cognitivo</b>	<b>Económico</b>	<b>Social</b>	<b>Salud</b>	<b>Promedio</b>
2014-I	64	26	10	5	0	21
2015-I	31	22	60	5	0	24
2016-I	63	27	11	5	1	21
2017-I	54	24	8	5	1	18
<b>% Estudiantes No Evaluados</b>						
<b>Cohorte / Riesgo</b>	<b>Académico</b>	<b>Cognitivo</b>	<b>Económico</b>	<b>Social</b>	<b>Salud</b>	<b>Promedio</b>
2014-I	29	34	9	10	10	18
2015-I	19	14	7	5	5	10
2016-I	10	6	0	15	12	9
2017-I	16	6	0	10	9	8

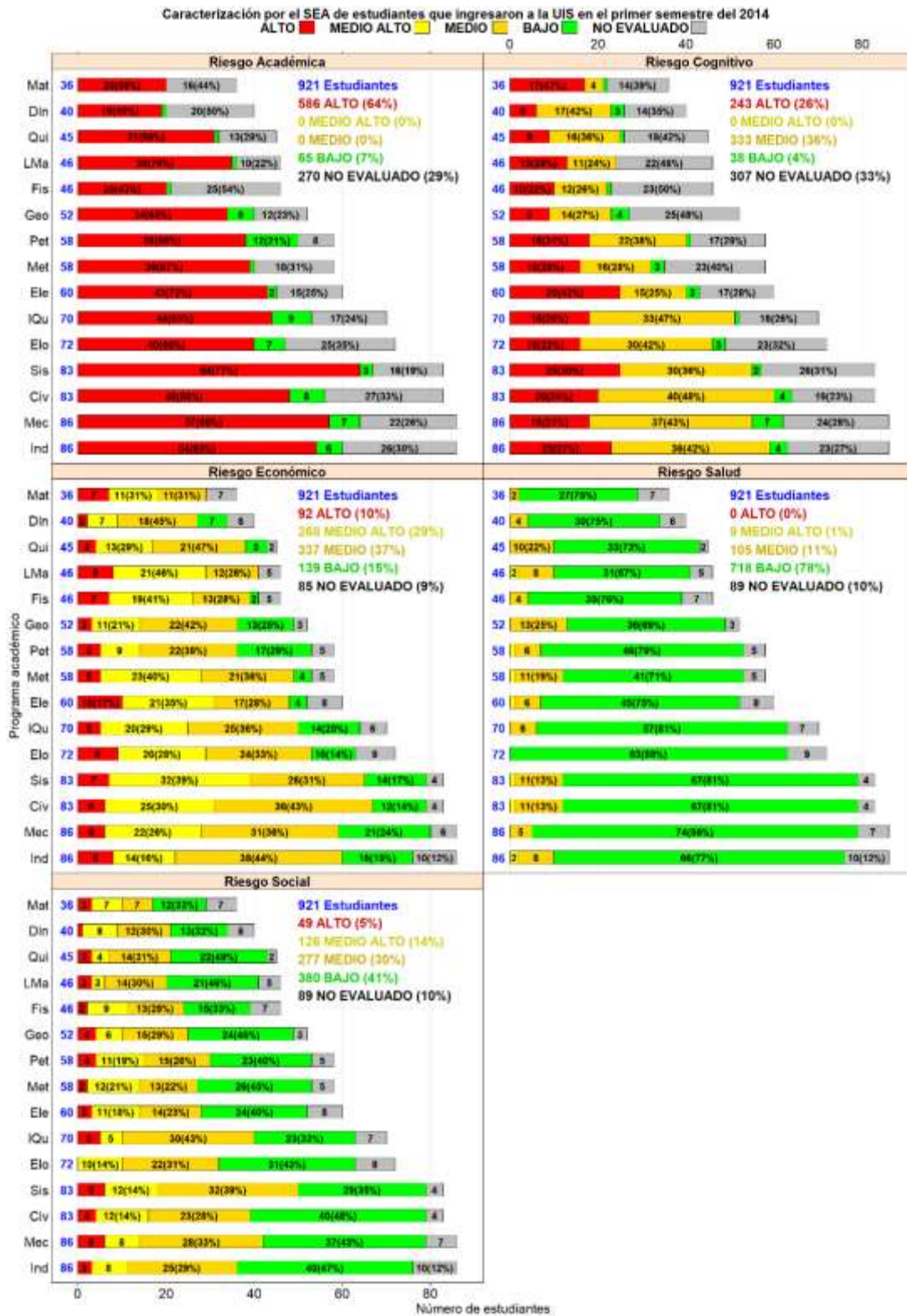


Figura 39. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2014.

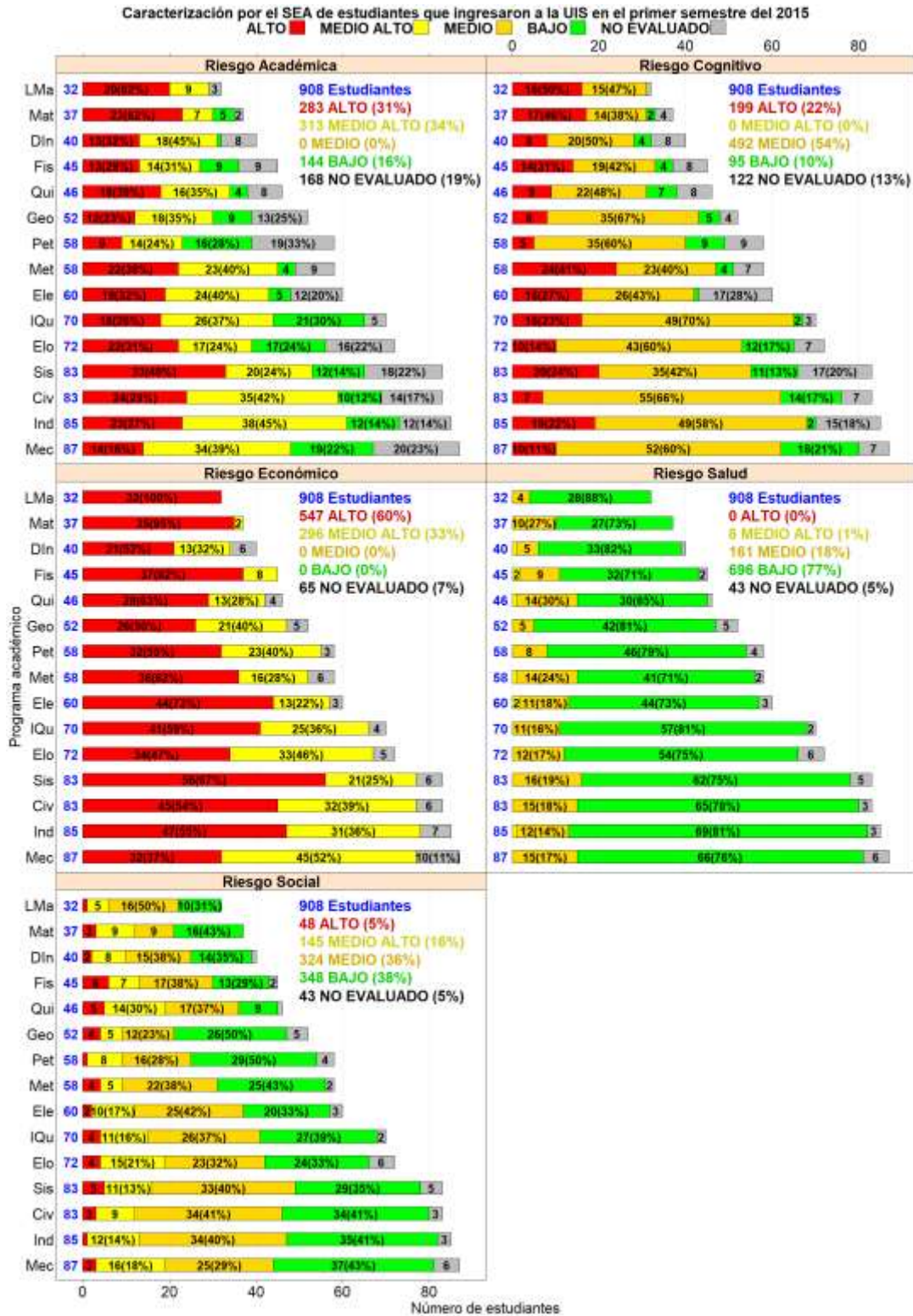


Figura 40. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2015.

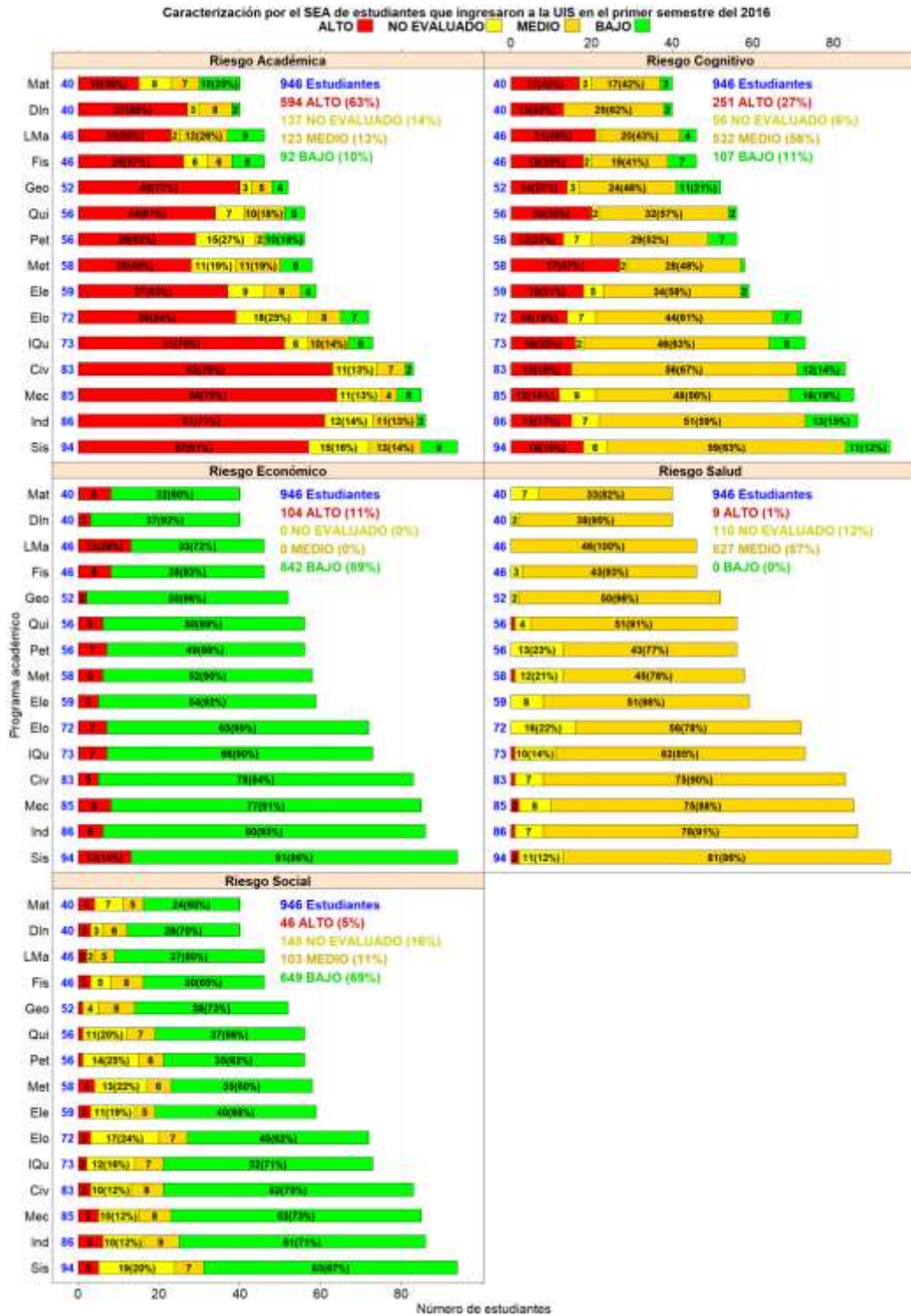


Figura 41. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2016.

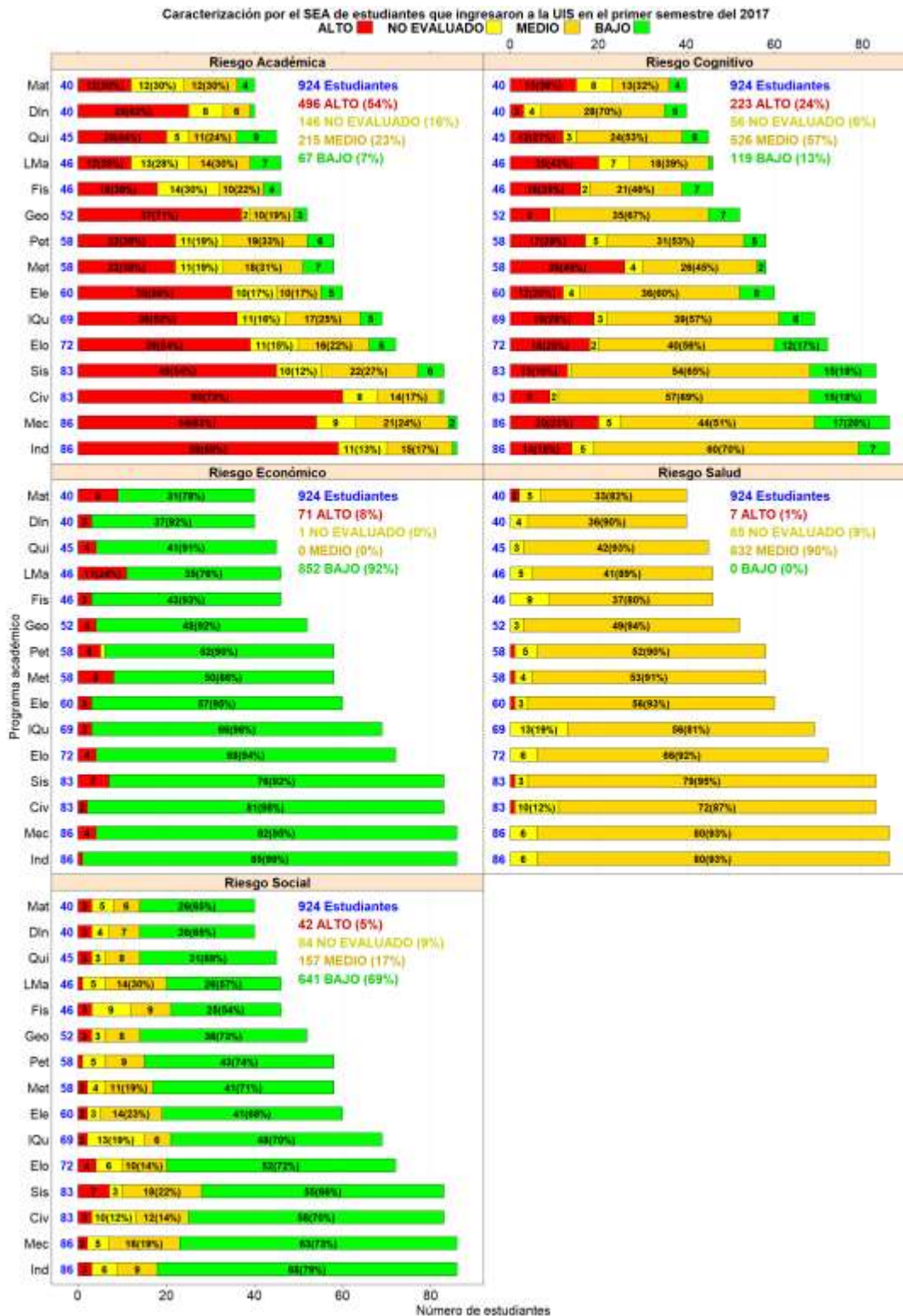


Figura 42. Estudiantes de Ciencias e Ingenierías caracterizados en el programa SEA por carrera, Semestre 1 - 1ª cohorte de 2017.

### 3.2 Evaluación de Impacto del Programa Sea

En esta sección se hace uso de las técnicas de diferencias en diferencias y pareamiento para evaluar el impacto del programa SEA sobre el desempeño académico de estudiantes de Ciencia e Ingeniería de la Universidad Industrial de Santander en las asignaturas de matemáticas.

En la sección 0 se realiza el análisis de los estudiantes de la primera cohorte del 2014 en los cálculos (I, II y III), álgebra y ecuaciones diferenciales; las secciones siguientes muestran los resultados para la primera cohorte de los años 2015 a 2017 para cálculo I y álgebra con el objetivo de mostrar la evolución de programa SEA con el tiempo en estas dos asignaturas. La técnica de pareamiento se empleó en los estudiantes que cursaron la materia por primera vez y la técnica de diferencias en diferencias para los estudiantes que la cursaron dos veces, es decir los repitentes.

Para el uso de la técnica de pareamiento, se consideraron los estudiantes con información completa<sup>5</sup> en las siguientes co-variables: el programa académico, la edad, el sexo, los resultados de las pruebas SABER 11 más influyentes en el desempeño académico (Matemática, Física, Química, Lenguaje e Inglés) y el riesgo en las 5 dimensiones consideradas por el programa SEA (económica, social, salud, académica y cognitiva). Estas co-variables se emplean en la construcción del modelo de regresión logística necesario para estimar la propensión a participar del programa SEA en cada asignatura, requerida para el uso de la técnica de pareamiento. Cada modelo de propensión fue mejorado, eliminando las variables menos significativas desde el punto de vista estadístico, empleado el algoritmo de regresión por pasos, basado en el criterio de

---

<sup>5</sup> Aquellos estudiantes que no tenían la información completa de todas las variables no fueron tenidos en cuenta en el estudio.

información de Akaike (AIC) como medida de calidad relativa del modelo. La técnica de pareamiento finaliza con la evaluación del impacto del programa SEA en función del número de horas de tutoría recibidas por el estudiante para cada asignatura en el semestre. Al final de cada sección se presentan los resultados del uso de la técnica de diferencias en diferencias sobre los repitentes de cada asignatura.

Para la primera cohorte del 2014 (sección 0), se presentan y analizan detalladamente los resultados intermedios del uso de la técnica; sin embargo, en las secciones subsiguientes se presentan los resultados de forma sintetizada para no incurrir en repeticiones que aparten al lector de los objetivos principales del estudio. Al final de cada sección se muestra una síntesis de cada cohorte y en la sección 0 se presenta un análisis consolidado y comparativo con los aspectos más importantes de los semestres considerados.

**3.2.1 Impacto en cohorte 2014-1.** La primera sección de este apartado contiene el análisis detallado del uso de las técnicas de evaluación impacto de programa SEA sobre cálculo I en estudiantes de la primera cohorte del 2014. En las secciones subsiguientes se analizan las demás asignaturas de matemáticas (álgebra, cálculo II, Cálculo III y ecuaciones diferenciales) centrandó el análisis y discusión en los aspectos más relevantes, sin detallar los resultados intermedios.

**3.2.1.1 Impacto del SEA sobre Cálculo I. Cohorte 2014-1.** La información suministrada por la Universidad para esta asignatura, contiene 430 estudiantes con información completa de un total de 921 (47%) estudiantes de Ciencias e Ingeniería que ingresaron a la Universidad la primera cohorte del 2014. Esto se debe a que en sus inicios, el programa SEA no realizó la caracterización

completa de todos los estudiantes en las 5 dimensiones de riesgo establecidas por el programa (académica, económica, cognitiva, salud y social). Como se mencionó anteriormente, la técnica de pareamiento exige que todos los estudiantes considerados tengan toda la información, por tal motivo, el presente análisis solo se continuó con los 430 estudiantes que presentan información completa. La Figura 43 muestra la cantidad de estudiantes beneficiados del programa SEA en Cálculo I en función del número de horas de beneficio recibido para la primera cohorte del 2014. Dentro de cada barra en color negro se muestra la cantidad de estudiantes en la categoría correspondiente y a la derecha se destaca en color rojo el valor acumulado. Se observa que 59 estudiantes recibieron al menos 1 hora de beneficio; la mayoría de ellos (27) entre 1 y 7 horas en el semestre; 9 de ellos recibieron más de 23 horas de tutoría, algo así como 1,5 tutoría por semana.

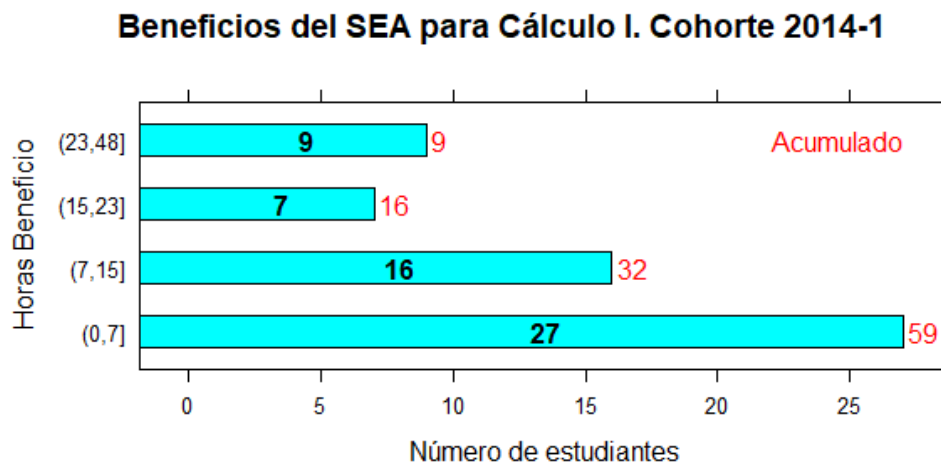


Figura 43. Beneficios del SEA en cálculo I. 1ª cohorte de 2014.

Inicialmente se realiza la evaluación del programa SEA considerando beneficiario aquel estudiante que recibió al menos una hora de beneficio en el semestre, es decir con el grupo de 59

estudiantes mostrado en la Figura 43 y posteriormente se realiza el análisis de sensibilidad del impacto en función del número de horas de beneficio recibidas en el semestre.

La Tabla 16 muestra la comparación entre el grupo de beneficiados y no beneficiados del programa SEA, indicando el valor p de la prueba estadística en la última columna antes de aplicar la técnica de pareamiento.

Tabla 16.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2014*

	No Beneficiado	Beneficiado	Valor p
n	371	59	
Nota (mean (sd))	2.86 (0.97)	2.50 (0.94)	0.009
Horas (mean (sd))	0.00 (0.00)	11.53 (10.40)	<0.001
Prog (%)			<0.001
Civ	40 (10.8)	3 (5.1)	
DIn	12 (3.2)	1 (1.7)	
Ele	26 (7.0)	5 (8.5)	
Elo	34 (9.2)	4 (6.8)	
Geo	15 (4.0)	7 (11.9)	
Ind	43 (11.6)	3 (5.1)	
IQu	32 (8.6)	9 (15.3)	
LMa	16 (4.3)	2 (3.4)	
Mat	3 (0.8)	6 (10.2)	
Mec	41 (11.1)	5 (8.5)	
Met	23 (6.2)	3 (5.1)	
Pet	36 (9.7)	1 (1.7)	
Qui	14 (3.8)	3 (5.1)	
Sis	36 (9.7)	7 (11.9)	
Edad (mean (sd))	17.95 (1.43)	18.05 (1.28)	0.616
Genero = Hombre (%)	240 (64.7)	33 (55.9)	0.249
Matemática (mean (sd))	71.40 (9.39)	64.32 (8.32)	<0.001
Lenguaje (mean (sd))	62.09 (5.33)	60.15 (6.10)	0.011
Inglés (mean (sd))	63.78 (15.42)	58.91 (14.17)	0.023
Física (mean (sd))	68.24 (10.69)	62.50 (8.76)	<0.001
Química (mean (sd))	65.97 (9.26)	62.22 (9.18)	0.004
Biología (mean (sd))	59.77 (7.60)	55.29 (6.12)	<0.001
Riesgo_Economica (%)			0.056
ALTO	29 (7.8)	6 (10.2)	
BAJO	82 (22.1)	4 (6.8)	
MEDIO	156 (42.0)	30 (50.8)	
MEDIO ALTO	104 (28.0)	19 (32.2)	

Riesgo_Social (%)			0.704
ALTO	15 (4.0)	3 (5.1)	
BAJO	164 (44.2)	30 (50.8)	
MEDIO	141 (38.0)	18 (30.5)	
MEDIO ALTO	51 (13.7)	8 (13.6)	
Riesgo_Salud (%)			0.273
BAJO	329 (88.7)	50 (84.7)	
MEDIO	41 (11.1)	8 (13.6)	
MEDIO ALTO	1 (0.3)	1 (1.7)	
Riesgo.Academico = BAJO (%)	48 (12.9)	3 (5.1)	0.129
Riesgo_Cognitiva (%)			0.483
ALTO	137 (36.9)	25 (42.4)	
BAJO	26 (7.0)	2 (3.4)	
MEDIO	208 (56.1)	32 (54.2)	
Ben (%)			<0.001
(-1,0]	371 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	27 (45.8)	
(7,15]	0 (0.0)	16 (27.1)	
(15,23]	0 (0.0)	7 (11.9)	
(23,48]	0 (0.0)	9 (15.3)	

Se destacan varios aspectos importantes en esta comparación:

- El número de estudiantes en cada grupo es muy diferente<sup>6</sup> (59 beneficiados vs 371 no beneficiados) lo que tiene dos desventajas: la prueba de comparación es más sensible al supuesto de igualdad de varianzas y la potencia estadística es menor.
- El impacto del programa SEA en cálculo I, medido como el cambio de nota promedio, es negativo: -0.36 unidades y significativo al 0.9%. Esto resulta contrario a lo esperado y se debe, en parte, a que se están comparando dos grupos heterogéneos en otras características que así como el beneficio del SEA, también pueden afectar el desempeño del estudiantes el cálculo I.
- Se destacan diferencias significativas en las pruebas SABER 11, siendo siempre superior el corte promedio en las pruebas SABER 11 de los no beneficiados que de los beneficiados. Este aspecto también explica las diferencias en la nota promedio de cálculo I entre el grupo de beneficiados y no beneficiados.

<sup>6</sup> En estadística este concepto se conoce como diseños desbalanceado y se refiere a una situación en la que el número de observaciones no es el mismo en los grupos comparados.

La técnica de pareamiento busca para cada estudiante beneficiado el estudiante más parecido en el grupo de estudiantes no beneficiados, haciendo el diseño balanceado y asegurando que las diferencias encontradas realmente se puedan atribuir al efecto del programa SEA y no a otros factores.

El modelo de regresión logística (modelo de propensión) incluyendo todas las co-variables se muestra en la Tabla 17. En la parte superior de la tabla se muestra la fórmula con las variables involucradas en la construcción del modelo y en la parte inferior se aprecian los estimados del modelo con la prueba de significancia correspondiente; se destaca que las dos variables más significativas, desde el punto de vista estadístico, fueron el puntaje en matemáticas de las pruebas SABER 11 (valor  $p=0.000215$ ) seguido del programa académico (valor  $p < 0.05$  en geología, ingeniería química y licenciatura en matemáticas) y el riesgo en Salud (valor  $p = 0.031$ ) mientras que las demás variables no resultaron significativas en el modelo de propensión al 5% de significancia.

Tabla 17.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2014 (Modelo Completo.)*

```
Beneficio ~ Prog + Edad + Genero + Matematica + Lenguaje + Ingles +
Fisica + Quimica + Biologia + Riesgo_Economica + Riesgo_Social + Ries
go_Salud + Riesgo.Academico + Riesgo_Cognitiva
```

```
Call:
glm(formula = Beneficio ~ . - Nota - Horas - Ben, family = binomial,
     data = Datos.Evaluacion2)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5704  -0.5130  -0.2956  -0.1327   2.7589
```

```
Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      7.4585102   4.0661028    1.834 0.066607 .
ProgDIn          -0.6419183   1.2836254   -0.500 0.617017
```

ProgEle	0.1029863	0.8569867	0.120	0.904346
ProgElo	0.0684500	0.8933190	0.077	0.938922
ProgGeo	2.0450130	0.8630948	2.369	0.017817 *
ProgInd	-0.0758384	0.9082573	-0.083	0.933455
ProgIQ	1.9538665	0.8053186	2.426	0.015258 *
ProgLMa	-2.6240662	1.3066936	-2.008	0.044625 *
ProgMat	1.2514938	1.1770284	1.063	0.287662
ProgMec	0.8282722	0.8349490	0.992	0.321196
ProgMet	-0.2484781	0.9670988	-0.257	0.797232
ProgPet	-0.3370095	1.2664317	-0.266	0.790155
ProgQui	-0.1662286	1.0237653	-0.162	0.871015
ProgSis	-0.1922728	0.8518227	-0.226	0.821420
Edad	0.1715134	0.1203150	1.426	0.154002
GeneroHombre	-0.3511550	0.3756641	-0.935	0.349913
Matematica	-0.1094681	0.0295783	-3.701	0.000215 ***
Lenguaje	-0.0001115	0.0430559	-0.003	0.997933
Inglés	0.0075846	0.0134351	0.565	0.572393
Física	-0.0262667	0.0270186	-0.972	0.330965
Química	-0.0133655	0.0250225	-0.534	0.593245
Biología	-0.0587195	0.0320678	-1.831	0.067085 .
Riesgo_EconomicaBAJO	-0.9239393	0.8267912	-1.118	0.263781
Riesgo_EconomicaMEDIO	0.3072945	0.6202324	0.495	0.620282
Riesgo_EconomicaMEDIO ALTO	0.2064849	0.6378454	0.324	0.746148
Riesgo_SocialBAJO	0.0229805	0.7993109	0.029	0.977064
Riesgo_SocialMEDIO	-0.5073043	0.8005955	-0.634	0.526304
Riesgo_SocialMEDIO ALTO	-0.3705547	0.8796824	-0.421	0.673582
Riesgo_SaludMEDIO	0.3747223	0.5286788	0.709	0.478455
Riesgo_SaludMEDIO ALTO	3.4396721	1.5983815	2.152	0.031400 *
Riesgo_AcademicoBAJO	-0.3762591	0.7150014	-0.526	0.598725
Riesgo_CognitivaBAJO	0.5036708	0.9028961	0.558	0.576954
Riesgo_CognitivaMEDIO	0.6621122	0.3716295	1.782	0.074807 .

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 343.88 on 429 degrees of freedom  
 Residual deviance: 261.53 on 397 degrees of freedom  
 AIC: 327.53

Number of Fisher Scoring iterations: 6

Una estrategia adecuada para simplificar el modelo consiste en eliminar las variables menos influyentes en el modelo de regresión. Sin embargo, como el efecto de cada variable en el modelo de regresión logística depende de las demás variables presentes en el modelo, lo más conveniente para depurarlo es usar una técnica sistemática que permita encontrar el subconjunto de variables más influyentes de forma conjunta.

En la presente investigación se empleó el algoritmo de regresión por pasos basado en el estadístico de Akaike (AIC) para determinar el mejor modelo de propensión que incluya el grupo

de variables que produce al estadístico AIC más bajo<sup>7</sup>; El resumen del modelo de regresión logística depurado, para estimar los puntajes de propensión o la probabilidad de que un estudiante participe en las tutorías de cálculo I, ofrecidas por el programa SEA se muestra en la Tabla 18

Tabla 18.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2014 (Modelo depurado.)*

```
Call:
glm(formula = Beneficio ~ Prog + Matematica + Biologia, family = binomial,
     data = Datos.Evaluacion2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6969  -0.5515  -0.3456  -0.1853   2.8601

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  8.42786    2.23522   3.770 0.000163 ***
ProgDIn      -0.45836    1.22467  -0.374 0.708202
ProgEle      -0.03540    0.81676  -0.043 0.965429
ProgElo      -0.03194    0.84486  -0.038 0.969845
ProgGeo       1.61961    0.78550   2.062 0.039220 *
ProgInd      -0.11707    0.87008  -0.135 0.892967
ProgIQU       1.75366    0.73782   2.377 0.017463 *
ProgLMa      -1.62150    1.10487  -1.468 0.142217
ProgMat       1.66520    1.01505   1.641 0.100900
ProgMec       0.44020    0.78696   0.559 0.575909
ProgMet      -0.11170    0.89648  -0.125 0.900845
ProgPet      -0.73159    1.20237  -0.608 0.542883
ProgQui       0.34312    0.92266   0.372 0.709980
ProgSis      -0.16937    0.78391  -0.216 0.828944
Matematica  -0.09617    0.02560  -3.756 0.000172 ***
Biologia     -0.07133    0.02789  -2.558 0.010536 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 343.88  on 429  degrees of freedom
Residual deviance: 279.10  on 414  degrees of freedom
AIC: 311.1

Number of Fisher Scoring iterations: 6
```

<sup>7</sup> Dado un conjunto de modelos candidatos, el modelo preferido es el que tiene el valor mínimo en el AIC. Por lo tanto AIC no sólo recompensa la bondad de ajuste, sino también incluye una penalidad, que es una función creciente del número de parámetros estimados, lo que desalienta el sobreajuste del modelo. (Wikipedia, La enciclopedia libre, 2017)

Las variables más significativas en el modelo de propensión depurado fueron el puntaje obtenido en matemáticas y biología en las pruebas SABER 11 y el programa académico del estudiante.

La Figura 44 muestra la distribución de puntajes de propensión estimados con el modelo de regresión logística para los grupos de beneficiados y no beneficiados del SEA en cálculo I. Se observa, una discriminación razonable y un rango común estrecho que dificulta el pareamiento apropiado de los estudiantes con altos puntajes de propensión. Se observan dos estudiantes con puntaje de propensión muy elevado para los cuales no existen parejas con puntaje de propensión similares.

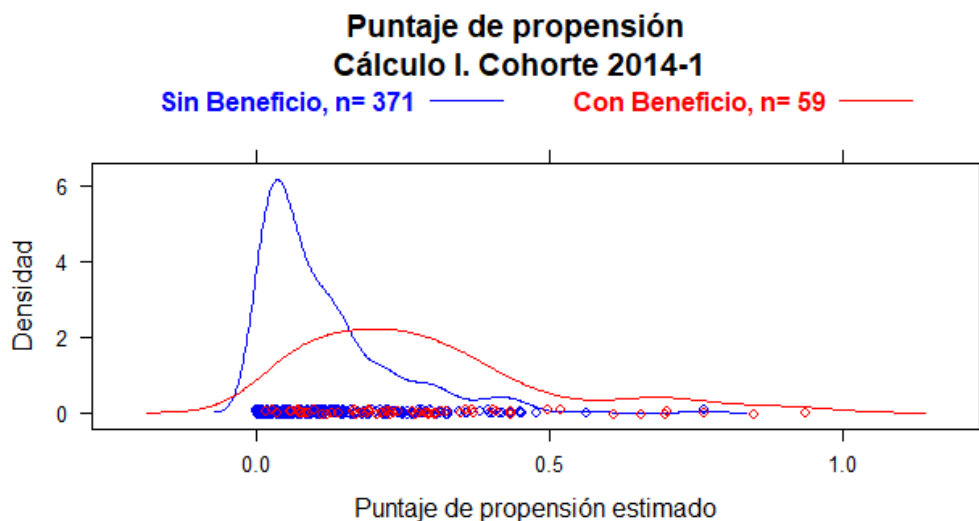


Figura 44. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2014

Las Figura 45 muestra gráficamente el diagrama de puntos con los resultados del pareamiento. En esta figura, cada punto representa el puntaje de propensión de un estudiante y los dos puntos en la parte superior derecha representa dos estudiantes con beneficio del SEA que no pudieron ser pareados debido a que se encontraban fuera del rango común. Las dos figuras del medio muestran la similitud entre los estudiantes que recibieron beneficio del programa SEA en cálculo I con su

correspondiente pareja que no recibió; 57 estudiantes de los 59 que recibieron beneficio del programa SEA en cálculo I pudieron ser pareados con estudiantes que no recibieron beneficios, dejando 2 estudiantes sin pareamiento. La parte inferior muestra los estudiantes sin beneficio que no fueron empleados en el pareamiento.

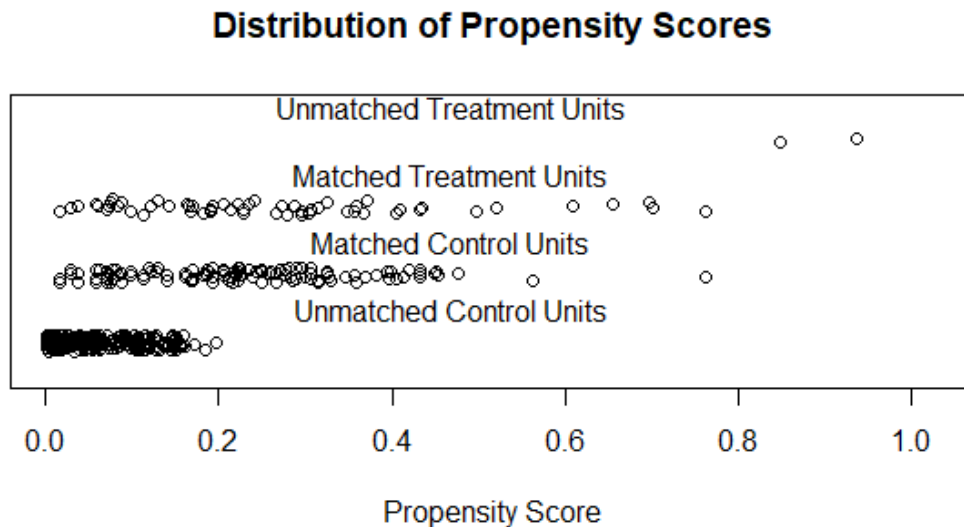


Figura 45. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2014

Los resultados finales del pareamiento se resumen en la Tabla 19, donde se aprecian dos estudiantes con beneficios que fueron descartados en el pareamiento y los 47 descartados por estar fuera del rango común que no tuvieron beneficio del SEA (Grupo de control).

Tabla 19.

*Resultados finales del pareamiento para beneficiados en cálculo I. 1ª cohorte de 2014 de Ciencias e Ingeniería.*

	Control	Treated
All	371	59
Matched	114	57
Unmatched	210	0
Discarded	47	2

La Figura 46 muestra el histograma de los puntajes de propensión para el grupo de beneficiados (superior) o no beneficiados (inferior) antes (izquierda) y después del pareamiento (derecha). Antes del pareamiento se observa una distribución considerablemente diferente entre beneficiados y no beneficiados debido a la heterogeneidad de los dos grupos antes del pareamiento, mientras que los de la derecha, después del pareamiento, son muy similares indicando un pareamiento satisfactorio.

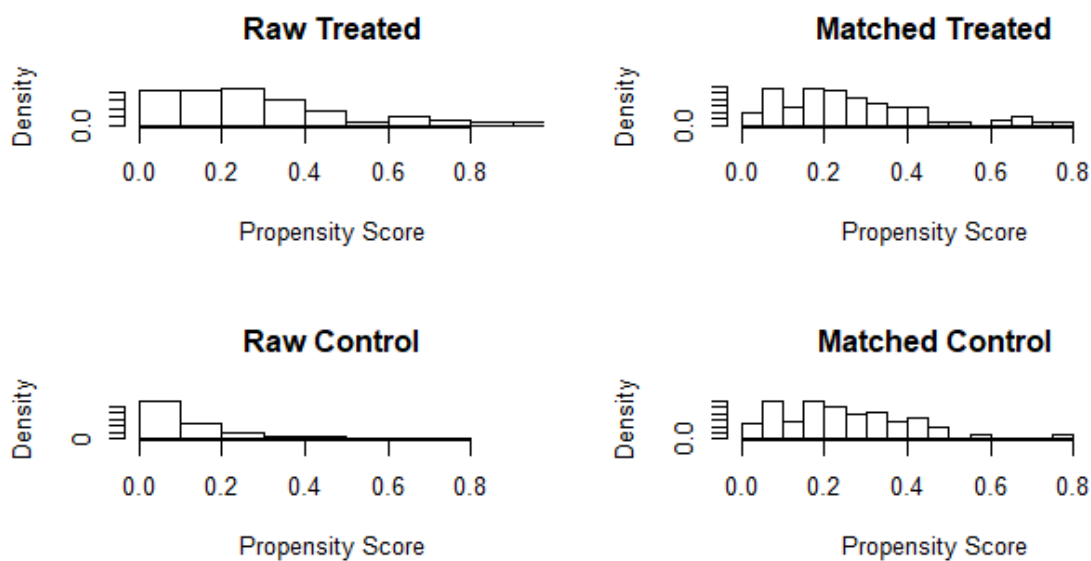


Figura 46. Histograma de puntajes de propensión para cálculo I antes (izquierda) y después (derecha) del pareamiento. 1ª cohorte de 2014.

En síntesis, tanto la información numérica como grafica indican un pareamiento satisfactorio con dos estudiantes beneficiados excluidos del pareamiento por estar fuera del rango común y una distribución similar de los puntajes de propensión, aspectos que aseguran la formación de grupos homogéneos y por lo tanto estimaciones apropiadas del contrafactual de cada estudiante beneficiado.

La Tabla 20 compara los estadísticos más importantes entre el grupo de estudiantes con beneficio y sin beneficio del programa SEA en cálculo I después de aplicar la técnica de pareamiento. Se observa que los dos grupos resultan comparables en todas las co-variables, por lo que se puede asumir que cualquier diferencia en la nota promedio de cálculo I se debe a los beneficios recibidos del programa SEA.

Tabla 20.

*Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2014*

	Sin Beneficio	Con Beneficio	p
n	57	57	
Nota (mean (sd))	2.53 (0.94)	2.50 (0.96)	0.891
Horas (mean (sd))	0.00 (0.00)	11.56 (10.49)	<0.001
Prog (%)			0.985
Civ	4 (7.0)	3 (5.3)	
DIn	0 (0.0)	1 (1.8)	
Ele	6 (10.5)	5 (8.8)	
Elo	3 (5.3)	4 (7.0)	
Geo	6 (10.5)	7 (12.3)	
Ind	3 (5.3)	3 (5.3)	
IQu	8 (14.0)	9 (15.8)	
LMa	2 (3.5)	2 (3.5)	
Mat	3 (5.3)	4 (7.0)	
Mec	2 (3.5)	5 (8.8)	
Met	3 (5.3)	3 (5.3)	
Pet	1 (1.8)	1 (1.8)	
Qui	6 (10.5)	3 (5.3)	
Sis	10 (17.5)	7 (12.3)	
Edad (mean (sd))	18.02 (1.03)	18.02 (1.29)	1.000
Genero = Hombre (%)	35 (61.4)	33 (57.9)	0.849
Matematica (mean (sd))	65.17 (7.91)	64.86 (7.93)	0.837
Lenguaje (mean (sd))	60.54 (5.40)	60.57 (5.64)	0.975
Ingles (mean (sd))	57.45 (13.21)	59.20 (14.31)	0.499
Fisica (mean (sd))	63.69 (9.87)	63.08 (8.26)	0.720
Quimica (mean (sd))	62.78 (8.83)	62.72 (8.93)	0.972
Biologia (mean (sd))	54.87 (6.37)	55.67 (5.78)	0.485
Riesgo_Economica (%)			0.600
ALTO	7 (12.3)	5 (8.8)	
BAJO	7 (12.3)	4 (7.0)	
MEDIO	23 (40.4)	29 (50.9)	
MEDIO ALTO	20 (35.1)	19 (33.3)	
Riesgo_Social (%)			0.235
ALTO	0 (0.0)	3 (5.3)	

BAJO	25 (43.9)	29 (50.9)	
MEDIO	23 (40.4)	17 (29.8)	
MEDIO ALTO	9 (15.8)	8 (14.0)	
Riesgo_Salud (%)			0.502
BAJO	51 (89.5)	48 (84.2)	
MEDIO	6 (10.5)	8 (14.0)	
MEDIO ALTO	0 (0.0)	1 (1.8)	
Riesgo.Academico = BAJO (%)	3 (5.3)	3 (5.3)	1.000
Riesgo_Cognitiva (%)			0.746
ALTO	27 (47.4)	23 (40.4)	
BAJO	2 (3.5)	2 (3.5)	
MEDIO	28 (49.1)	32 (56.1)	
Ben (%)			<0.001
(-1,0]	57 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	26 (45.6)	
(7,15]	0 (0.0)	16 (28.1)	
(15,23]	0 (0.0)	6 (10.5)	
(23,48]	0 (0.0)	9 (15.8)	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
distance (mean (sd))	0.25 (0.15)	0.27 (0.18)	0.608

En la Tabla 20 se observa una leve disminución en la nota promedio de cálculo I en el grupo de estudiantes que recibieron beneficios del SEA aunque ésta no es significativa desde el punto de vista estadístico (valor  $p = 0.891$ ). En las demás variables cuantitativas se aprecia la media, desviación y el valor  $p$  de la prueba de comparación entre el grupo de 57 estudiantes beneficiados y 57 parejas no beneficiados; en las variables cualitativas se aprecia la cantidad de estudiantes y su porcentaje con respecto al total. En general, se destaca que el grupo de estudiantes beneficiados y no beneficiados después del pareamiento es similar en todas las variables consideradas asegurando homogeneidad de los dos grupos y aumentando la potencia estadística de la comparación. La última fila muestra el puntaje de propensión promedio de dos grupos, medida de la distancia global entre el grupo de beneficiados y no beneficiados del SEA, en la cual tampoco se encuentran diferencias significativas (valor  $p = 0.608$ ).

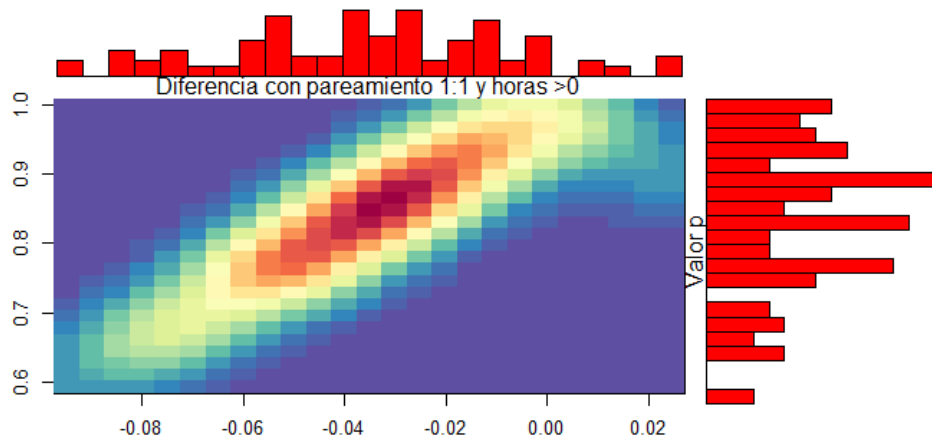
Los resultados anteriores muestran la evaluación del impacto del programa SEA en el desempeño de los estudiantes de ciencias e Ingeniería para la cohorte 2014-1 empleando la técnica

de pareamiento 1 a 1. Esta es la mejor alternativa cuando la base de datos es pequeña debido a que lleva a un diseño balanceado aislando el efecto de otros factores no deseados que también puedan afectar el desempeño en cálculo I, comparando grupos homogéneos. No obstante, y dependiendo del número de estudiantes no beneficiados, el pareamiento se puede realizar 1 a 2, 1 a 3, etc.

En el presente estudio se realizaron simulaciones para evaluar el efecto de tamaño de pareamiento en la evaluación del impacto del programa SEA sobre el desempeño de los estudiantes de cálculo I en la primera cohorte del 2014, encontrando que al aumentar el número de estudiantes en el pareamiento cerca de un valor límite, los resultados son similares a los obtenidos sin realizar pareamiento. Las diferencias disminuyen hasta un valor mínimo y posteriormente aumentan hasta valores muy cercanos al obtenido sin pareamiento, en la que los grupos no son homogéneos y por lo tanto las comparaciones no son confiables. Por tal motivo, se mantuvo un pareamiento 1 a 1 en las subsiguientes evaluaciones de impacto realizadas empleando esta técnica.

En la técnica de pareamiento, cuando el modelo de propensión contiene pocas co-variables, es muy probable que el puntaje de propensión se repita para dos o más estudiantes diferentes y por lo tanto el pareamiento no sea único. Esta situación se subsanó realizando varias corridas y considerando los valores promedio de las corridas realizadas. La Figura 47 muestra el histograma bi-variado de 100 corridas de pareamiento 1 a 1 para evaluar el impacto del SEA en cálculo I para estudiantes de ciencia e ingeniería de la primera cohorte del 2014. El eje horizontal muestra la distribución de diferencias promedio entre cada beneficiado y el promedio de los respectivos pares en el grupo de no beneficiados; en el eje vertical se aprecia la distribución del valor p de la prueba t correspondiente. Se observa que la diferencia promedio entre la nota de estudiantes beneficiados

y no beneficiados cambia entre  $-0.095$  y  $0.025$  en promedio, aunque en ningún caso, dicho cambio resulta significativo desde el punto vista estadístico (valor  $p > 0.5$ ).



*Figura 47.* Histograma bi-variado de 100 diferencias promedio en notas de cálculo I entre beneficiados y no beneficiados del SEA con pareamiento 1:1. 1ª cohorte de 2014

La evaluación de impacto del programa SEA sobre cálculo I en la primera cohorte del 2014, realizada hasta el momento, consideró como beneficiario del programa SEA aquel estudiante que recibió al menos 1 hora de beneficio. No obstante, es claro que el efecto del programa SEA depende del número de horas de beneficio recibidas en el semestre y por tanto se espera que los estudiantes que recibieron más horas de beneficio obtengan una nota mayor que los que recibieron menos horas. Debido a que el semestre académico tiene una duración de 16 semanas en promedio, y teniendo en cuenta la cantidad tan reducida de estudiantes que recibieron beneficio, se consideraron las siguientes categorías (Tabla 21) para evaluar el impacto en la nota en función del número de horas de beneficio recibidos:



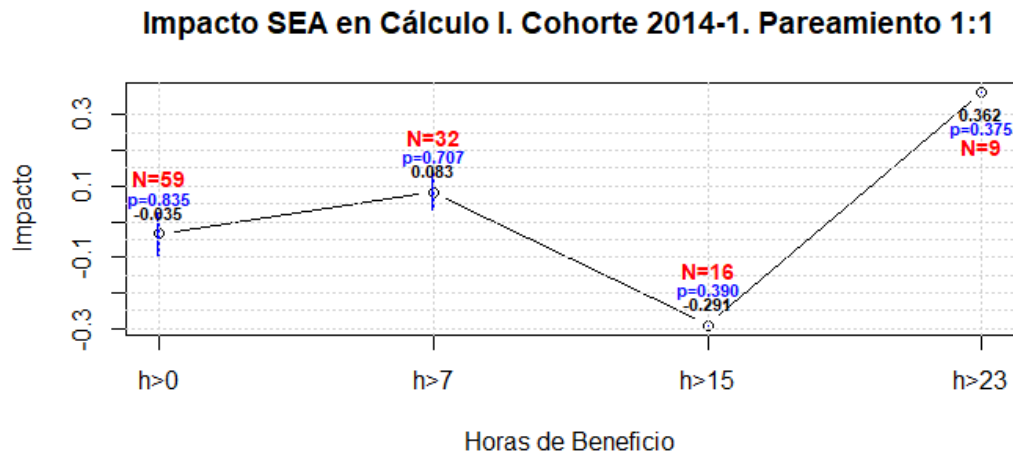


Figura 48. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2014.

En síntesis, la técnica de pareamiento no muestra un impacto significativo sobre la nota de cálculo I atribuida al uso del programa SEA a un nivel de significancia del 37.5%<sup>8</sup> en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2014.

**3.2.1.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2014-1.** Como se mencionó anteriormente, la evaluación del impacto del programa SEA sobre el desempeño académico de los estudiantes repitentes la primera vez en las asignaturas del ciclo básico de matemáticas [cálculos I, II y III, álgebra y ecuaciones diferenciales], se realizará empleando la técnica de diferencias en diferencias. Esta técnica fue descrita en la sección 0 y contrasta las diferencias en la nota de los estudiantes a lo largo del tiempo entre una población beneficiada del programa SEA y otra que no lo es. En esta sección se evaluará el impacto del programa SEA sobre los repitentes (por primera

<sup>8</sup> Menor valor de las 100 corridas con un pareamiento 1 a 1 y más de 24 horas de beneficio recibidas

vez) de cálculo I y en las próximas tres secciones se realizará sobre las demás asignaturas de matemáticas para los estudiantes que ingresaron a la UIS en la primera cohorte del 2014. De los 921 estudiantes de ciencia e ingeniería en la primera cohorte se encontró que el 35% (319) de los estudiantes cursaron cálculo I dos veces: 300 sin beneficio del SEA y 19 con beneficio.

La técnica de diferencias en diferencias supone que las tendencias de la nota son similares en los grupos de beneficiados y no beneficiados antes de la intervención y que los únicos factores que explican las diferencias en la nota entre ambos grupos, aparte del efecto del programa SEA, son constantes en el tiempo. Para asegurar la no violación del supuesto anterior, la técnica de diferencias en diferencias se debe condicionar a que el grupo de beneficiados haya recibido beneficios los dos semestres que cursó la materia y el grupo de no beneficiados no haya recibido beneficios ninguno de los dos semestre. Esto reduce el grupo de beneficiarios y no beneficiarios ya que exige que un estudiante se considera beneficiario si recibió beneficio los dos semestres que vio la materia y se considera no beneficiado si no recibió beneficio del SEA ninguno de los dos semestres. La distribución de horas de beneficio con este condicionamiento se aprecia en la Figura 49.

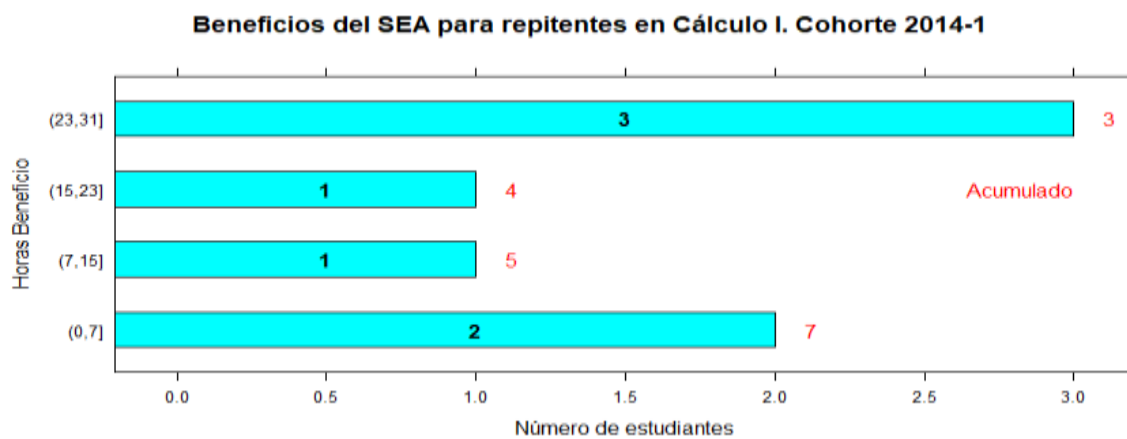
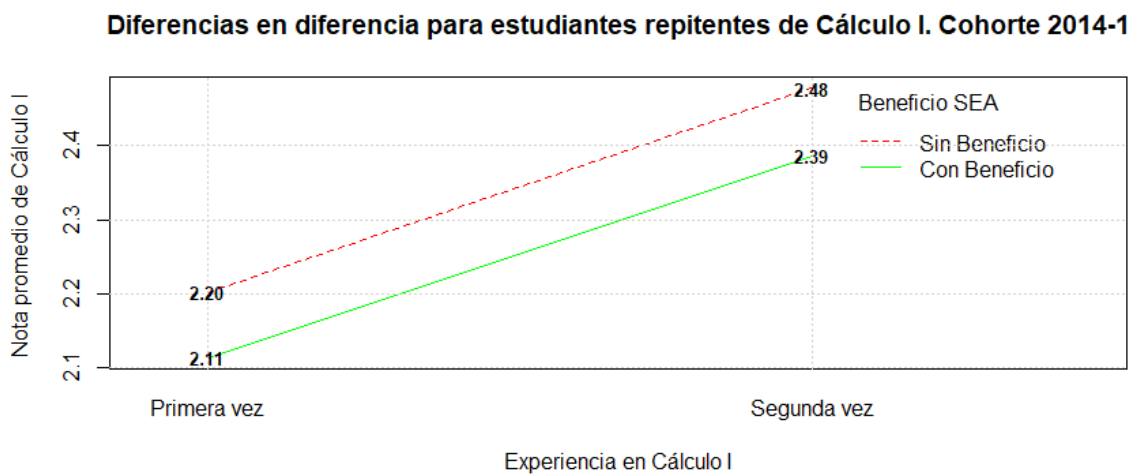


Figura 49. Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2014.

Inicialmente se considera como beneficiario aquel estudiante que recibió al menos una hora de beneficio en el semestre que cursó la asignatura por segunda vez (7 estudiantes) y posteriormente se evaluará el efecto en función del número de horas de beneficio recibidas en la medida que el número de estudiantes en cada categoría lo permita.

La Figura 50 muestra el uso del método de diferencias en diferencias para los estudiantes repitentes de cálculo I con al menos una hora de beneficio recibido en el semestre (solo 7 estudiantes). Se observa que como las dos líneas son casi paralelas, el impacto del programa SEA, estimado por el método de diferencias en diferencias es despreciable.



*Figura 50.* Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2014

Formalmente, el impacto del programa SEA en estudiantes repitentes de cálculo I se estima como la diferencia entre el cambio promedio en la nota del grupo de beneficiados ( $2.386 - 2.114 = 0.271$ ) y el cambio promedio en la nota de no beneficiados ( $2.477 - 2.202 = 0.276$ ), es decir,  $-0.004$  ( $0.271 - 0.276$ ) unidades. El estimador del impacto del programa SEA junto con la prueba estadística de nulidad se muestra en la Tabla 23.

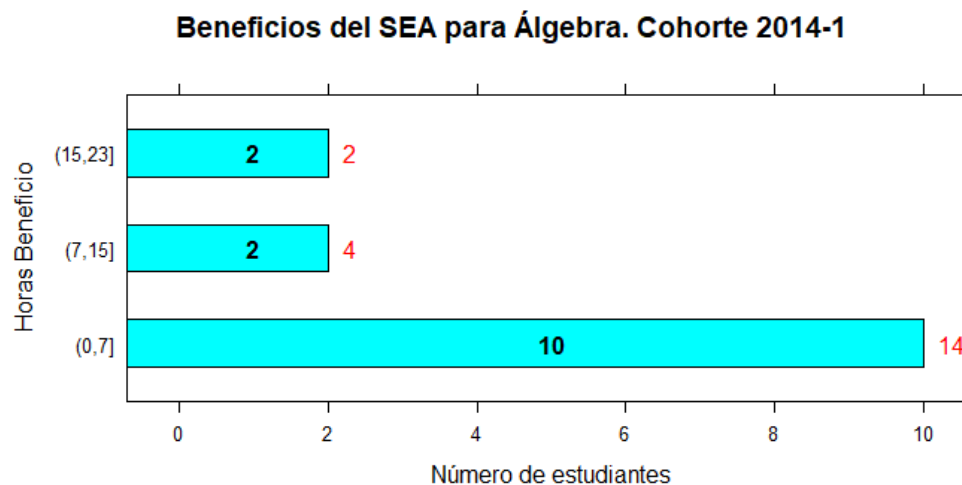
Tabla 23.

*Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2014.*

Estimate	Std. Error	t value	Pr(> t )
-0.003865546	0.373442239	-0.010351122	0.991745117

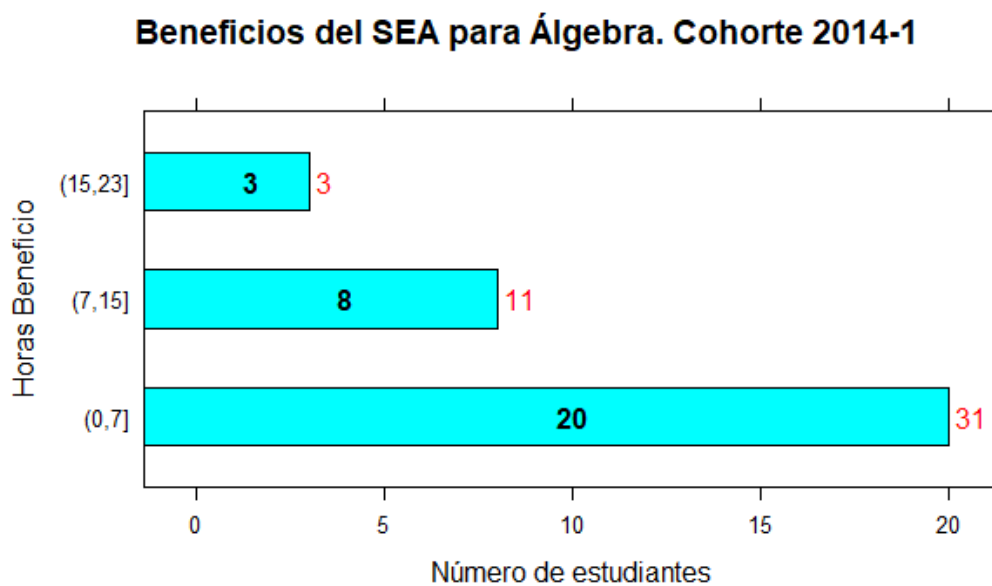
En síntesis se puede decir que el método de diferencias en diferencias arroja un impacto negativo (-0.00387) pero no significativo (valor  $p = 0.9917$ ) sobre la nota de los estudiantes repitentes de cálculo I que ingresaron a la UIS en la primera cohorte del 2014.

**3.2.1.3 Impacto del SEA sobre álgebra lineal. Cohorte 2014-1.** De los 921 estudiantes de esta cohorte, 837 (91%) estudiantes cursaron la materia, de los cuales solo 31 recibieron beneficio del SEA. Sin embargo, solo 427 (46%) reportaron información completa, 14 de los cuales con beneficio del SEA según se aprecia en la Figura 51.



*Figura 51.* Beneficios del SEA en álgebra con información completa. 1ª cohorte de 2014.

Como la depuración del modelo con los 14 estudiantes que presentaron información completa no incluyó ninguna dimensión de riesgo, se excluyeron las variables asociadas a las 5 dimensiones de riesgo y se re-planteó el modelo con los 31 individuos que presentaban información completa en las demás variables: nota de álgebra, beneficio recibido en álgebra, edad, género y los puntajes en matemáticas, física, lenguaje e inglés de las pruebas SABER 11; quedando en definitiva 806 estudiantes que no recibieron beneficio y 31 que recibieron al menos una hora, con la información completa. La distribución se aprecia en la Figura 52, encontrándose solo 3 estudiantes con más de 16 horas de beneficio, dificultando el uso de la técnica de pareamiento en esta categoría por la cantidad tan baja de estudiantes, lo que llevo a fusionarla con la categoría inferior<sup>9</sup>.



*Figura 52.* Beneficios del SEA en álgebra con información completa excluyendo variables de riesgo. 1ª cohorte de 2014.

<sup>9</sup> En adelante se fusionan categorías contiguas de menos de 5 estudiantes debido a las dificultades del uso de la técnica de pareamiento en categorías con poco número de estudiantes.

La Tabla 24 muestra la comparación entre el grupo de estudiantes beneficiados y no beneficiados del SEA en álgebra para la primera cohorte del 2014.

Tabla 24.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en álgebra. 1ª cohorte de 2014*

	No Beneficiado	Beneficiado	p
n	806	31	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
Nota (mean (sd))	2.97 (0.99)	2.88 (0.69)	0.632
Horas (mean (sd))	0.00 (0.00)	6.42 (5.33)	<0.001
Prog (%)			0.404
Civ	75 (9.3)	2 (6.5)	
DIn	1 (0.1)	0 (0.0)	
Ele	53 (6.6)	3 (9.7)	
Elo	62 (7.7)	6 (19.4)	
Fis	42 (5.2)	3 (9.7)	
Geo	48 (6.0)	2 (6.5)	
Ind	75 (9.3)	3 (9.7)	
IQu	67 (8.3)	1 (3.2)	
LMa	43 (5.3)	1 (3.2)	
Mat	31 (3.8)	2 (6.5)	
Mec	84 (10.4)	0 (0.0)	
Met	55 (6.8)	1 (3.2)	
Pet	53 (6.6)	4 (12.9)	
Qui	40 (5.0)	1 (3.2)	
Sis	77 (9.6)	2 (6.5)	
Edad (mean (sd))	18.36 (2.29)	18.45 (1.80)	0.833
Genero = Hombre (%)	561 (69.6)	15 (48.4)	0.021
Matematica (mean (sd))	69.67 (10.69)	63.52 (8.51)	0.002
Lenguaje (mean (sd))	60.98 (5.66)	61.45 (7.02)	0.654
Ingles (mean (sd))	61.28 (15.64)	55.48 (12.86)	0.042
Fisica (mean (sd))	66.27 (11.04)	63.68 (9.88)	0.199
Quimica (mean (sd))	64.51 (9.83)	61.90 (9.58)	0.148
Biologia (mean (sd))	58.64 (7.60)	56.81 (5.80)	0.186
Ben (%)			<0.001
(-1,0]	806 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	20 (64.5)	
(7,23]	0 (0.0)	11 (35.5)	

Se aprecia que la nota promedio de álgebra en el grupo de beneficiarios es inferior a la del grupo de no beneficiados, mostrando un aparente impacto negativo del programa SEA de 0.1 unidades en esta asignatura pero sin ser significativo al 5%. Entre tanto, también se destaca una diferencia significativa en el género, el puntaje en matemáticas e inglés de las pruebas SABER 11 entre los

dos grupos, aspecto que, junto con el programa SEA, también pueden explicar las diferencias en el desempeño en álgebra. La técnica de pareamiento permite comparar dos grupos similares en todos sus características, excepto por el hecho que un grupo recibió beneficios del SEA y el otro no, haciendo por tanto que las diferencias en la nota promedio de álgebra se pueda atribuir al uso del programa.

El modelo de propensión con todas las co-variables se resume en la Tabla 25, las variables más significativas fueron el puntaje en matemáticas de las pruebas SABER 11 y el género (valor  $p < 0.05$ ); mientras que las demás variables no resultaron significativas con base en el correspondiente valor  $p$ .

Tabla 25.

*Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2014 (Modelo completo)*

```
Beneficio ~ Prog + Edad + Genero + Matematica + Lenguaje + Ingles + Fisica + Quimica + Biologia
```

Call:

```
glm(formula = Beneficio ~ . - Nota - Horas - Ben, family = binomial,
     data = Datos.Evaluacion2)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-0.9165	-0.2928	-0.1899	-0.1061	3.1594

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	2.156e+00	3.700e+00	0.583	0.56003
ProgDIn	-1.765e+01	1.075e+04	-0.002	0.99869
ProgEle	3.597e-01	9.896e-01	0.364	0.71621
ProgElo	1.128e+00	8.858e-01	1.274	0.20276
ProgFis	6.740e-02	1.052e+00	0.064	0.94894
ProgGeo	1.702e-01	1.073e+00	0.159	0.87396
ProgInd	-5.714e-02	9.646e-01	-0.059	0.95277
ProgIQu	-8.378e-01	1.273e+00	-0.658	0.51051
ProgLMa	-2.004e+00	1.493e+00	-1.342	0.17973
ProgMat	-8.730e-01	1.331e+00	-0.656	0.51179

```

ProgMec      -1.583e+01  1.105e+03  -0.014  0.98857
ProgMet      -8.880e-01  1.269e+00  -0.700  0.48422
ProgPet      1.245e+00  9.774e-01  1.274  0.20261
ProgQui      -9.043e-01  1.323e+00  -0.684  0.49411
ProgSis      -6.862e-01  1.088e+00  -0.631  0.52824
Edad         3.017e-02  7.360e-02  0.410  0.68193
GeneroHombre -1.170e+00  4.215e-01  -2.776  0.00551 **
Matematica   -8.219e-02  3.204e-02  -2.565  0.01032 *
Lenguaje     3.331e-02  3.945e-02  0.844  0.39842
Ingles       -2.244e-02  1.612e-02  -1.392  0.16393
Fisica       -2.537e-03  2.848e-02  -0.089  0.92900
Quimica      -4.059e-03  2.768e-02  -0.147  0.88343
Biologia     -9.069e-04  3.438e-02  -0.026  0.97895
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)

Null deviance: 265.18 on 836 degrees of freedom
Residual deviance: 224.29 on 814 degrees of freedom
AIC: 270.29

Number of Fisher Scoring iterations: 18

```

La Tabla 26 resume los resultados del modelo de regresión logística depurado empleado para estimar los puntajes de propensión o la probabilidad de que un estudiante participen en las tutorías de álgebra, ofrecidas por el programa SEA.

Tabla 26.

*Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2014 (Modelo depurado)*

```

call:
glm(formula = Beneficio ~ Genero + Matematica + Lenguaje, family = binomial,
    data = Datos.Evaluacion2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7283  -0.3082  -0.2258  -0.1738   3.0841

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.36603    1.79950  -0.759  0.447785
GeneroHombre -0.76578    0.37205  -2.058  0.039564 *
Matematica   -0.08261    0.02398  -3.445  0.000572 ***
Lenguaje     0.06600    0.03167   2.084  0.037156 *
---

```

signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 265.18 on 836 degrees of freedom  
Residual deviance: 245.63 on 833 degrees of freedom  
AIC: 253.63

Se observa que éste excluyó todas las variables excepto el género, los puntajes en matemáticas y lenguaje de las pruebas SABER 11, resultando estas significativas al 5% con base en el valor p correspondiente para cada parámetro. Al igual que ocurrió con cálculo I, matemáticas continúa siendo la variable que mejor explica la propensión de un estudiante a tomar beneficios del SEA en álgebra, seguida de lenguaje y el género, disminuyendo significativamente la proporción de hombres que tienden a tomar beneficios, pasando de 69.6% a 48.4% como se aprecia en la Tabla 24.

La Figura 53 muestra la distribución de puntajes de propensión estimados con el modelo de regresión logística para los grupos de beneficiados y no beneficiados del SEA en álgebra observándose un rango común adecuado para el uso de la técnica de pareamiento.

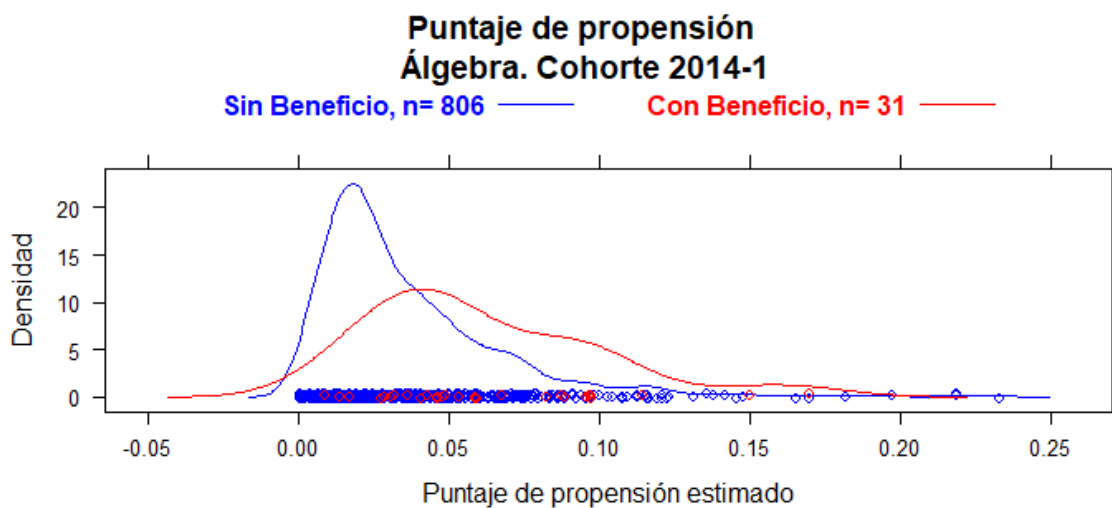
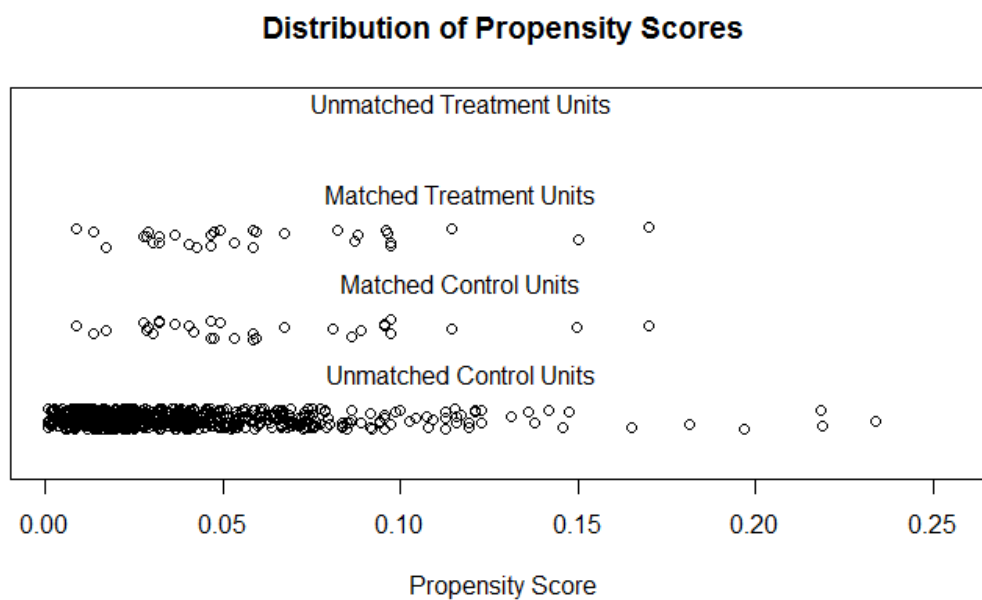


Figura 53. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2014

Las Figura 54 muestra el diagrama de puntos con los resultados del pareamiento, cada punto representa el puntaje de propensión de un estudiante. Las dos figuras del medio muestra la similitud entre los estudiantes que recibieron beneficio con su correspondiente pareja que no recibió; 31 estudiantes que recibieron beneficio del programa SEA en álgebra pudieron ser pareados con estudiantes que no recibieron beneficio. La parte inferior muestra los estudiantes sin beneficio que no fueron empleados en el pareamiento.



*Figura 54.* Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2014

Los resultados finales del pareamiento para los 31 estudiantes beneficiados por el programa SEA en álgebra se resumen en la

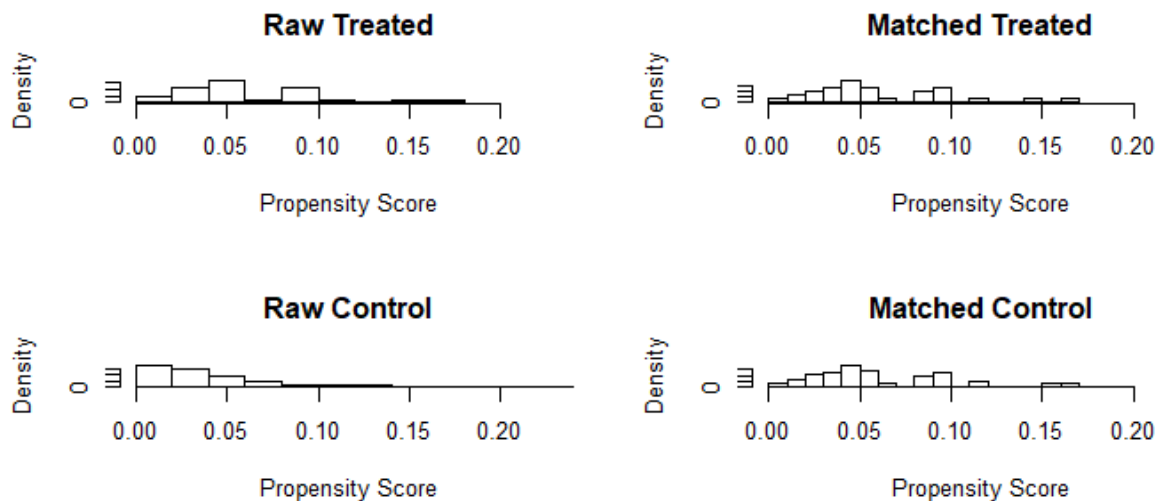
Tabla 27 destacando que no se descartó ningún estudiantes con beneficio pero si 86 que no tuvieron beneficio del SEA por estar fuera del rango común y se ignoraron los 689 restantes.

Tabla 27.

*Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2014 de Ciencias e Ingeniería*

	Control	Treated
All	806	31
Matched	31	31
Unmatched	689	0
Discarded	86	0

La Figura 55 muestra el histograma de los puntajes de propensión antes y después del pareamiento, a la izquierda se aprecian diferencias considerables en la distribución debido a la heterogeneidad de los dos grupos antes del pareamiento mientras que los de la derecha, después del pareamiento, se observa una similitud mayor entre la distribución del grupo de beneficiados y no beneficiados, lo que sugiere un pareamiento satisfactorio por puntajes de propensión y una mayor homogeneidad de los grupos comparados.



*Figura 55. Histograma de puntajes de propensión para álgebra I antes y después del pareamiento.*

1ª cohorte de 2014

En síntesis, el diagnóstico de pareamiento tanto con información numérica como gráfica indica un pareamiento satisfactorio de los 31 estudiantes que recibieron beneficio en álgebra. A continuación, se describen los resultados relevantes de dicho pareamiento.

La Tabla 28 muestra de forma comparativa los estudiantes con beneficio y sin beneficio del programa SEA en álgebra I después de aplicar la técnica de pareamiento. Se observa, que los dos grupos resultan comparables en todas las co-variables por lo que se puede asumir, que cualquier diferencia en la nota promedio de álgebra se debe a los beneficios recibidos del programa SEA.

Tabla 28.

*Comparación de estudiantes con beneficio o no del SEA en álgebra I (después de pareamiento).*

*1ª cohorte de 2014*

	Sin Beneficio	Con Beneficio	p
n	31	31	
Nota (mean (sd))	2.96 (0.81)	2.88 (0.69)	0.699
Horas (mean (sd))	0.00 (0.00)	6.42 (5.33)	<0.001
Prog (%)			NaN
Civ	1 (3.2)	2 (6.5)	
DIn	0 (0.0)	0 (0.0)	
Ele	5 (16.1)	3 (9.7)	
Elo	1 (3.2)	6 (19.4)	
Fis	1 (3.2)	3 (9.7)	
Geo	5 (16.1)	2 (6.5)	
Ind	1 (3.2)	3 (9.7)	
IQu	3 (9.7)	1 (3.2)	
LMa	4 (12.9)	1 (3.2)	
Mat	1 (3.2)	2 (6.5)	
Mec	2 (6.5)	0 (0.0)	
Met	1 (3.2)	1 (3.2)	
Pet	1 (3.2)	4 (12.9)	
Qui	3 (9.7)	1 (3.2)	
Sis	2 (6.5)	2 (6.5)	
Edad (mean (sd))	18.03 (1.28)	18.45 (1.80)	0.295
Genero = Hombre (%)	13 (41.9)	15 (48.4)	0.799
Matemática (mean (sd))	63.48 (8.03)	63.52 (8.51)	0.988
Lenguaje (mean (sd))	60.68 (6.43)	61.45 (7.02)	0.652

Ingles (mean (sd))	56.87 (12.49)	55.48 (12.86)	0.668	
Fisica (mean (sd))	62.68 (9.74)	63.68 (9.88)	0.690	
Quimica (mean (sd))	64.03 (8.15)	61.90 (9.58)	0.350	
Biologia (mean (sd))	55.58 (5.19)	56.81 (5.80)	0.384	
Ben (%)			<0.001	
(-1,0]	31 (100.0)	0 (0.0)		
(0,7]	0 (0.0)	20 (64.5)		
(7,23]	0 (0.0)	11 (35.5)		
distance (mean (sd))	0.06 (0.04)	0.06 (0.04)	1.000	

Se aprecia una leve disminución en la nota promedio de álgebra en el grupo estudiantes que reciben tutorías pero ésta no es significativa desde el punto de vista estadístico (valor  $p = 0.699$ ). En las demás variables no se observa cambio significativo al 5%

La Figura 56 muestra el impacto del programa SEA sobre álgebra, medido como la diferencia promedio entre el grupo de beneficiados y no beneficiados después de aplicar la técnica de pareamiento 100 veces. Se observa (ver eje horizontal) que el impacto oscila entre -0.13 y 0.19 sin ser significativo en ninguno de los casos (valor  $p > 0.4$  en el eje vertical).

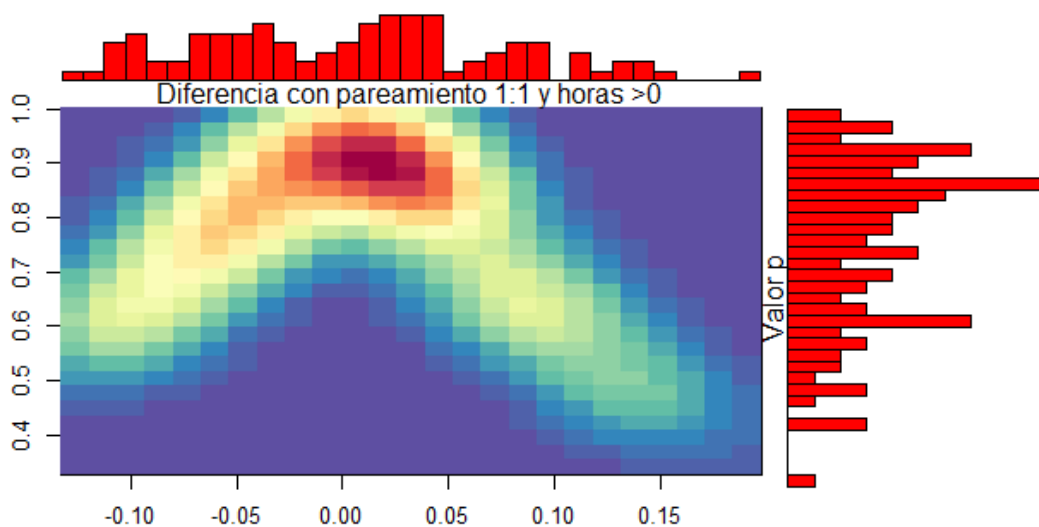
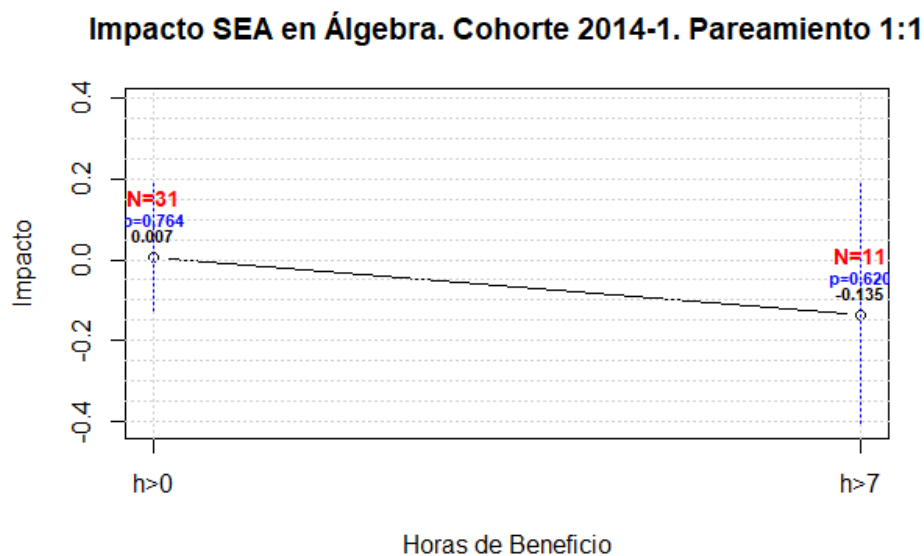


Figura 56. Histograma de 100 diferencias promedio en notas de álgebra I entre beneficiados y no beneficiados del SEA. 1ª cohorte de 2014

A pesar de los cambios de la diferencia promedio entre beneficiados y no beneficiados (a veces positivos y a veces negativos), esta diferencia no resultó significativamente distinta de cero, en ninguna de las corridas realizadas. Por lo tanto, se puede concluir que la técnica de pareamiento no muestra un impacto significativo en la nota de álgebra lineal I atribuida al uso del programa SEA a un nivel de significancia del 34% en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2014.

Finalmente, la Figura 57 muestra el impacto del programa en función del número de horas de beneficio recibidas con las dos categorías en las que fue posible emplear la técnica. A pesar de observarse una aparente tendencia decreciente, esta no resultó significativa en ninguno de los casos y los dos intervalos de variabilidad (línea punteada de color azul) se solapan, teniendo como punto central el cero.



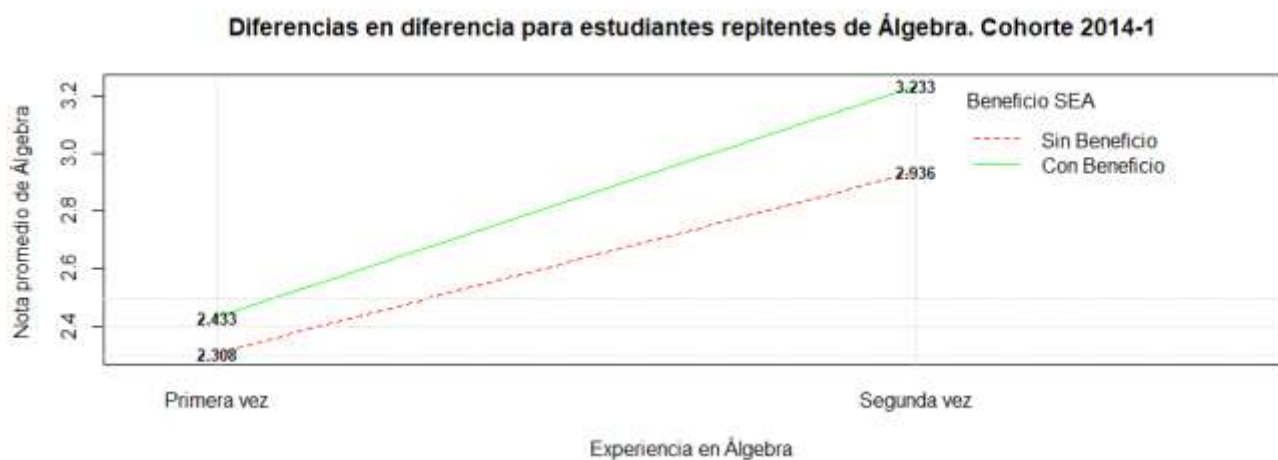
*Figura 57.* Impacto del programa SEA sobre álgebra en función del número de horas de beneficio.

1ª cohorte de 2014

En síntesis, la técnica de pareamiento no muestra un impacto significativo sobre la nota de álgebra atribuida al uso del programa SEA a un nivel de significancia del 24%<sup>10</sup> en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2014. Se observa una tendencia decreciente aunque no significativa, en una muestra de tan solo 11 estudiantes, por lo que no se puede considerar concluyente.

**3.2.1.4 Impacto del SEA sobre repitentes en álgebra lineal I. Cohorte 2014-1.** De los 921 estudiantes de ciencia e ingeniería en la primera cohorte se encontró que el 18% (168) de los estudiantes cursaron álgebra I dos veces: 157 sin beneficio, 3 con beneficio los dos semestres y 8 con beneficio la primera o la segunda vez que cursaron la materia.

La Figura 58 muestra el uso del método de diferencias en diferencias en los estudiantes repitentes de álgebra lineal I. Se observa que, como las dos líneas son casi paralelas, el impacto del programa SEA, estimado por el método de diferencias en diferencias es despreciable.



<sup>10</sup> Menor valor de las 100 corridas con un pareamiento 1 a 1 y más de 8 horas de beneficio recibidas

*Figura 58. Diferencias en diferencias a repitentes en álgebra lineal I. 1ª cohorte de 2014*

El estimador del impacto del programa SEA junto con la prueba estadística de nulidad se resume en la Tabla 29.

Tabla 29.

*Estimador del impacto del programa SEA sobre repitentes de álgebra lineal I usando diferencias en diferencias. 1ª cohorte de 2014.*

Estimate	Std. Error	t value	Pr(> t )
0.1726115	0.5520997	0.3126455	0.7547562

En síntesis se puede decir que el método de diferencias en diferencias arroja un impacto positivo (0.173 unidades) pero no significativo (valor  $p = 0.7547$ ) sobre la nota de los estudiantes repitentes de álgebra lineal que ingresaron a la UIS en la primera cohorte del 2014.

**3.2.1.5 Impacto del SEA sobre cálculo II. Cohorte 2014-1.** De los 921 estudiantes que ingresaron a la universidad a carreras de ciencias e ingeniería en el primer semestre del 2014, solamente 567 (61.6%) registran nota de cálculo II: 522 (92%) sin beneficio del SEA y 45 (8%) con beneficio. Para la evaluación del impacto del programa SEA sobre cálculo II se incluyeron como co-variables explicativas las mismas empleadas para cálculo I y álgebra así como las nota finales de cálculo I vista anteriormente por el estudiante, el promedio acumulado del semestre anterior y las horas de beneficio en cálculo I como una medida de la experiencia previa del estudiante antes de tomar el curso de cálculo II. Lo anterior con el objetivo de evaluar la forma como influyen estas variables en el modelo de propensión y realizar un mejor pareamiento o estimación del contrafactual. Incluyendo todas las variables mencionadas anteriormente solo se encontraron 305 estudiantes con información completa (33.11%): 28 beneficiados y 277 que no

recibieron beneficios del programa SEA. Con este grupo de estudiantes se construyó el modelo de regresión logística para estimar la probabilidad de inscribirse en el programa de tutorías ofrecidas por el programa SEA tal y como se aprecia en la Figura 59.

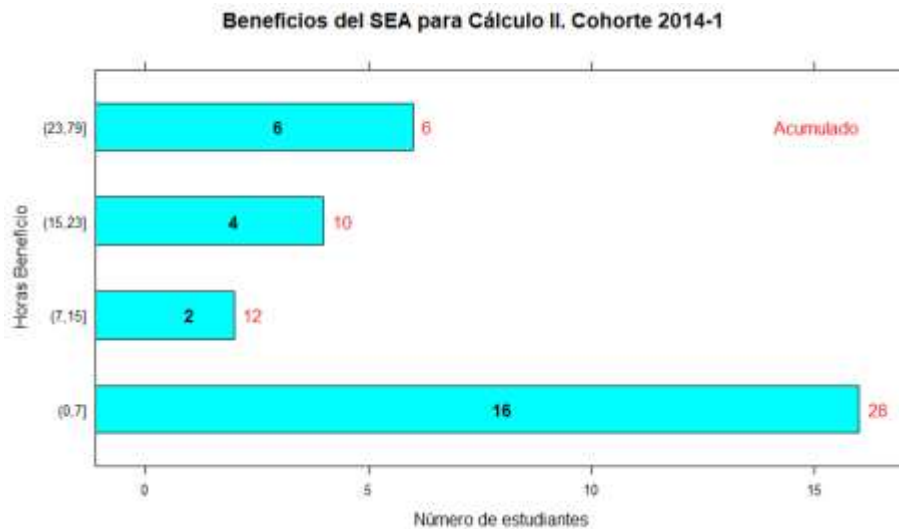


Figura 59. Beneficios del SEA en cálculo II con información completa. 1ª cohorte de 2014.

Debido a que la técnica de pareamiento no se puede realizar con una cantidad tan baja de estudiantes fue necesario fusionar las últimas tres categorías, quedando solo la categoría de (7,79] con 12 estudiantes. La Tabla 30 muestra la comparación estadística entre el grupo de beneficiados y no beneficiados antes de realizar el pareamiento.

Tabla 30.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo II. 1ª cohorte de 2014*

	No Beneficiado	Beneficiado	p
n	277	28	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
Nota (mean (sd))	3.02 (0.84)	2.86 (0.69)	0.332
Horas (mean (sd))	0.00 (0.00)	13.65 (17.38)	<0.001

Prog (%)			0.186
Civ	33 (11.9)	1 (3.6)	
DIn	8 (2.9)	1 (3.6)	
Ele	11 (4.0)	3 (10.7)	
Elo	26 (9.4)	2 (7.1)	
Geo	19 (6.9)	1 (3.6)	
Ind	31 (11.2)	5 (17.9)	
IQu	33 (11.9)	4 (14.3)	
LMa	5 (1.8)	2 (7.1)	
Mat	1 (0.4)	0 (0.0)	
Mec	36 (13.0)	0 (0.0)	
Met	16 (5.8)	4 (14.3)	
Pet	29 (10.5)	2 (7.1)	
Qui	10 (3.6)	0 (0.0)	
Sis	19 (6.9)	3 (10.7)	
Edad (mean (sd))	17.81 (1.18)	18.29 (2.02)	0.061
Genero = Hombre (%)	176 (63.5)	12 (42.9)	0.052
Matematica (mean (sd))	72.71 (9.34)	67.36 (7.68)	0.004
Lenguaje (mean (sd))	62.81 (5.12)	60.43 (5.67)	0.021
Ingles (mean (sd))	65.22 (15.37)	61.15 (14.43)	0.180
Fisica (mean (sd))	69.11 (9.95)	66.45 (11.48)	0.184
Quimica (mean (sd))	67.40 (8.91)	63.76 (10.05)	0.043
Biologia (mean (sd))	60.36 (7.62)	56.18 (6.01)	0.005
Riesgo_Economica (%)			0.702
ALTO	17 (6.1)	2 (7.1)	
BAJO	64 (23.1)	5 (17.9)	
MEDIO	124 (44.8)	11 (39.3)	
MEDIO ALTO	72 (26.0)	10 (35.7)	
Riesgo_Social (%)			0.626
ALTO	11 (4.0)	0 (0.0)	
BAJO	134 (48.4)	12 (42.9)	
MEDIO	98 (35.4)	12 (42.9)	
MEDIO ALTO	34 (12.3)	4 (14.3)	
Riesgo_Salud (%)			0.007
BAJO	246 (88.8)	24 (85.7)	
MEDIO	31 (11.2)	3 (10.7)	
MEDIO ALTO	0 (0.0)	1 (3.6)	
Riesgo.Academico = BAJO (%)	46 (16.6)	4 (14.3)	0.961
Riesgo_Cognitiva (%)			0.210
ALTO	92 (33.2)	12 (42.9)	
BAJO	24 (8.7)	0 (0.0)	
MEDIO	161 (58.1)	16 (57.1)	
Acumulado_2014_1 (mean (sd))	3.67 (0.44)	3.53 (0.38)	0.104
Nota.Max.CalculoI (mean (sd))	3.55 (0.50)	3.34 (0.38)	0.034
Horas.Total.CalculoI (mean (sd))	1.70 (5.99)	3.79 (7.40)	0.087
Ben (%)			<0.001
(-1,0]	277 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	16 (57.1)	
(7,79]	0 (0.0)	12 (42.9)	

Se observa una aparente reducción no significativa en la nota promedio de cálculo II y cambios significativos en edad, género, matemáticas, lenguaje, química y biología, así como en las en la

nota de cálculo I y los beneficios recibidos previamente en esta asignatura. En general, se destaca que los estudiantes de bajo desempeño en SABER 11 y en asignaturas previas como cálculo I son los que tienden a tomar beneficios del SEA en cálculo II. De igual forma, también se aprecia que las horas de beneficio en cálculo I difieren estadísticamente entre el grupo de beneficiados y no beneficiados de cálculo II, lo que sugiere que los estudiantes que recibieron beneficios en cálculo I anteriormente, tienden a seguir solicitando beneficios en cálculo II.

El modelo de regresión logística empleado para estimar la propensión de un estudiante a inscribirse en las tutorías ofrecidos por el programa SEA en cálculo II se muestra en la Tabla 31; en la parte superior se destaca la fórmula empleada en la construcción del modelo con la lista de co-variables consideradas, en la parte intermedia se aprecian los estimadores y el test de significancia de cada parámetro en el modelo y, al final de la Tabla 31, se encuentran los estadísticos globales del modelo de regresión logística.

Tabla 31.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo II. 1ª cohorte de 2014 (Modelo completo)*

```
Beneficio ~ Prog + Edad + Genero + Matematica + Lenguaje + Ingles + Fisica + Quimi
ca + Biologia + Riesgo_Economica + Riesgo_Social + Riesgo_Salud + Riesgo.Academico
+ Riesgo_Cognitiva + Acumulado_2014_1 + Nota.Max.CalculoI + Horas.Total.CalculoI"
```

Call:

```
glm(formula = Beneficio ~ . - Nota - Horas - Ben, family = binomial,
     data = Datos.Evaluacion2)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.36097	-0.39013	-0.18014	-0.00008	2.68819

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-2.740e+00	3.085e+03	-0.001	0.9993

ProgDIn	8.995e-01	1.789e+00	0.503	0.6152
ProgEle	2.006e+00	1.389e+00	1.445	0.1486
ProgElo	1.004e-01	1.401e+00	0.072	0.9428
ProgGeo	1.002e+00	1.540e+00	0.650	0.5154
ProgInd	1.658e+00	1.215e+00	1.364	0.1726
ProgIQu	1.928e+00	1.304e+00	1.479	0.1393
ProgLma	1.189e+00	1.825e+00	0.651	0.5148
ProgMat	-1.506e+01	1.075e+04	-0.001	0.9989
ProgMec	-1.599e+01	1.567e+03	-0.010	0.9919
ProgMet	2.103e+00	1.381e+00	1.523	0.1279
ProgPet	2.343e+00	1.488e+00	1.575	0.1152
ProgQui	-1.702e+01	2.775e+03	-0.006	0.9951
ProgSis	1.020e+00	1.441e+00	0.708	0.4789
Edad	1.991e-01	1.645e-01	1.210	0.2262
GeneroHombre	-1.298e+00	5.838e-01	-2.224	0.0261 *
Matematica	-5.771e-02	4.216e-02	-1.369	0.1710
Lenguaje	-1.307e-01	8.404e-02	-1.555	0.1200
Inglés	2.143e-02	1.870e-02	1.146	0.2517
Fisica	6.597e-02	3.698e-02	1.784	0.0744 .
Quimica	2.226e-03	3.828e-02	0.058	0.9536
Biología	-1.168e-01	4.977e-02	-2.347	0.0189 *
Riesgo_EconomicaBAJO	4.340e-01	1.270e+00	0.342	0.7325
Riesgo_EconomicaMEDIO	5.949e-01	1.216e+00	0.489	0.6248
Riesgo_EconomicaMEDIO ALTO	7.987e-01	1.255e+00	0.636	0.5245
Riesgo_SocialBAJO	1.688e+01	3.085e+03	0.005	0.9956
Riesgo_SocialMEDIO	1.760e+01	3.085e+03	0.006	0.9954
Riesgo_SocialMEDIO ALTO	1.669e+01	3.085e+03	0.005	0.9957
Riesgo_SaludMEDIO	-4.496e-01	8.138e-01	-0.552	0.5806
Riesgo_SaludMEDIO ALTO	2.299e+01	1.075e+04	0.002	0.9983
Riesgo_AcademicoBAJO	1.506e+00	8.447e-01	1.783	0.0746 .
Riesgo_CognitivaBAJO	-1.638e+01	1.713e+03	-0.010	0.9924
Riesgo_CognitivaMEDIO	3.794e-01	5.267e-01	0.720	0.4713
Acumulado_2014_1	-2.138e+00	9.518e-01	-2.246	0.0247 *
Nota_Max.CálculoI	-3.830e-01	7.687e-01	-0.498	0.6183
Horas.Total.CálculoI	2.070e-02	3.272e-02	0.633	0.5269

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 187.08 on 304 degrees of freedom

Residual deviance: 128.02 on 269 degrees of freedom

AIC: 200.02

Number of Fisher Scoring iterations: 18

Se destaca que solo 5 co-variables resultan significativas al 5% en el modelo de propensión: el género, los puntajes en física y biología en las pruebas SABER 11, el promedio acumulado y el riesgo académico. Esto indica que la gran mayoría de la co-variables consideradas no contribuyen significativamente a explicar la propensión a tomar beneficios en cálculo II y sí pueden afectar el

efecto de las demás variables que si resultaron importantes. El resumen del modelo simplificado, después de realizar la eliminación de co-variables menos significativas, empleado regresión por pasos, se muestra en la Tabla 32.

Tabla 32.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo II. 1ª cohorte de 2014 (Modelo depurado)*

```
Call:
glm(formula = Beneficio ~ Edad + Genero + Lenguaje + Fisica +
     Biologia + Riesgo_Salud + Acumulado_2014_1, family = binomial,
     data = Datos.Evaluacion2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1900  -0.4385  -0.3185  -0.1965   2.7249

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)         6.81686    4.33277   1.573  0.11564
Edad                 0.20087    0.11997   1.674  0.09406 .
GeneroHombre       -1.33443    0.45343  -2.943  0.00325 **
Lenguaje           -0.09804    0.05291  -1.853  0.06388 .
Fisica              0.04632    0.02759   1.679  0.09320 .
Biologia           -0.09367    0.03853  -2.431  0.01504 *
Riesgo_SaludMEDIO -0.23281    0.70104  -0.332  0.73983
Riesgo_SaludMEDIO ALTO 18.72622 1455.39762  0.013  0.98973
Acumulado_2014_1   -1.03452    0.56551  -1.829  0.06735 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 187.08  on 304  degrees of freedom
Residual deviance: 158.17  on 296  degrees of freedom
AIC: 176.17

Number of Fisher Scoring iterations: 14
```

El modelo depurado después de aplicar el algoritmo de regresión por pasos excluyó las co-variables poco influyentes en el modelo de propensión, manteniendo las que conducen el estadístico AIC más bajo: edad, género, el puntaje en lenguaje, física y biología en las pruebas

SABER 11, el riesgo en salud y el promedio acumulado del semestre anterior. De estas co-variables se destaca el puntaje en matemáticas el género y el puntaje en biología como las más significativas del modelo (valor  $p < 0.05$ ).

Los resultados finales del pareamiento para los 28 estudiantes beneficiados por el programa SEA en cálculo II se resumen en la Tabla 33. Como resultado del pareamiento se descartó un estudiante con beneficio y 61 que no tuvieron beneficio del SEA por estar fuera del rango común y se ignoraron los 189 restantes.

Tabla 33.

*Resultados finales del pareamiento para beneficiados en cálculo II. 1ª cohorte de 2014 de Ciencias e Ingeniería*

	Control	Treated
All	277	28
Matched	27	27
Unmatched	189	0
Discarded	61	1

La Figura 60 muestra el histograma de los puntajes de propensión antes y después del pareamiento. A la izquierda se aprecian amplias diferencias debido a la heterogeneidad de los dos grupos antes del pareamiento mientras que los de la derecha, después del pareamiento, las distribuciones son muy similares lo que sugiere un pareamiento satisfactorio por puntajes de propensión.

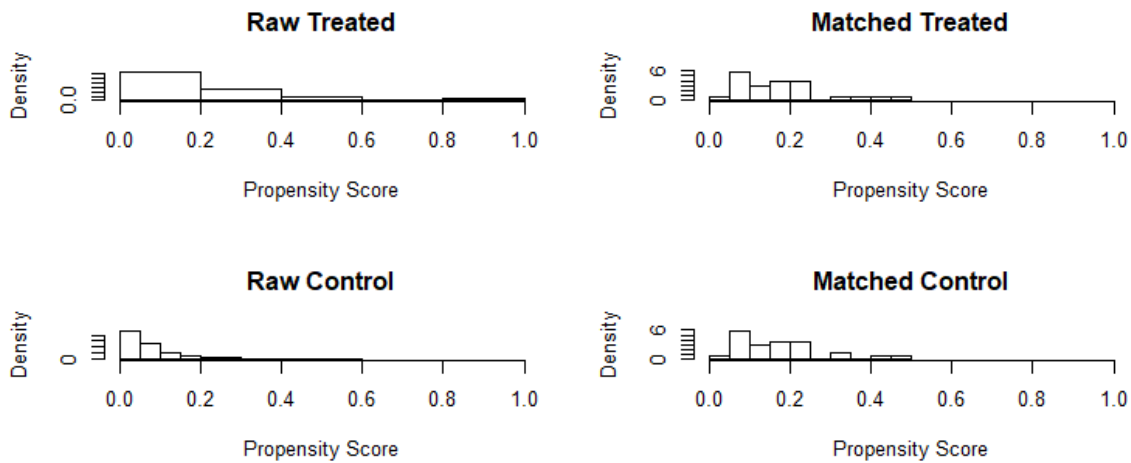


Figura 60. Histograma de puntajes de propensión para cálculo II antes y después del pareamiento. 1ª cohorte de 2014

Tanto la información numérica como gráfica indica un pareamiento satisfactorio con un estudiante beneficiado excluidos del pareamiento por estar fuera del rango común y una distribución similar de puntajes de propensión que asegura la formación de grupos homogéneos y por lo tanto estimaciones más apropiadas del contrafactual de cada estudiante beneficiado por la técnica de pareamiento.

La Tabla 34 compara los estadísticos más importantes entre el grupo de estudiantes con beneficio y sin beneficio del programa SEA en cálculo II después de aplicar la técnica de pareamiento. Se observa, que los dos grupos resultan comparables en todas las co-variables así que cualquier diferencia en la nota promedio de cálculo II se debe a los beneficios recibidos del programa SEA.

Tabla 34.

*Comparación de estudiantes con beneficio o no del SEA en cálculo II (después de pareamiento).*

*1ª cohorte de 2014*

	Sin Beneficio	Con Beneficio	p
n	27	27	
Nota (mean (sd))	2.98 (0.59)	2.86 (0.70)	0.517
Horas (mean (sd))	0.00 (0.00)	13.84 (17.68)	<0.001
Prog (%)			NaN
Civ	6 (22.2)	1 (3.7)	
DIn	0 (0.0)	1 (3.7)	
ELe	0 (0.0)	3 (11.1)	
Elo	1 (3.7)	2 (7.4)	
Geo	0 (0.0)	1 (3.7)	
Ind	7 (25.9)	5 (18.5)	
IQu	1 (3.7)	4 (14.8)	
LMa	0 (0.0)	2 (7.4)	
Mat	0 (0.0)	0 (0.0)	
Mec	4 (14.8)	0 (0.0)	
Met	2 (7.4)	3 (11.1)	
Pet	3 (11.1)	2 (7.4)	
Qui	1 (3.7)	0 (0.0)	
Sis	2 (7.4)	3 (11.1)	
Edad (mean (sd))	18.11 (1.48)	18.26 (2.05)	0.762
Genero = Hombre (%)	11 (40.7)	11 (40.7)	1.000
Matematica (mean (sd))	69.26 (8.86)	67.34 (7.83)	0.402
Lenguaje (mean (sd))	61.70 (3.07)	60.52 (5.75)	0.351
Inglés (mean (sd))	56.00 (9.25)	61.71 (14.40)	0.089
Fisica (mean (sd))	67.85 (7.42)	66.61 (11.67)	0.643
Quimica (mean (sd))	63.70 (6.50)	63.67 (10.23)	0.990
Biologia (mean (sd))	56.22 (6.23)	56.04 (6.08)	0.913
Riesgo_Economica (%)			0.349
ALTO	0 (0.0)	2 (7.4)	
BAJO	6 (22.2)	5 (18.5)	
MEDIO	14 (51.9)	10 (37.0)	
MEDIO ALTO	7 (25.9)	10 (37.0)	
Riesgo_Social (%)			0.294
ALTO	1 (3.7)	0 (0.0)	
BAJO	16 (59.3)	12 (44.4)	
MEDIO	6 (22.2)	12 (44.4)	
MEDIO ALTO	4 (14.8)	3 (11.1)	
Riesgo_Salud (%)			NaN
BAJO	23 (85.2)	24 (88.9)	
MEDIO	4 (14.8)	3 (11.1)	
MEDIO ALTO	0 (0.0)	0 (0.0)	
Riesgo.Academico = BAJO (%)	2 (7.4)	4 (14.8)	0.665
Riesgo_Cognitiva (%)			0.478
ALTO	13 (48.1)	11 (40.7)	
BAJO	1 (3.7)	0 (0.0)	
MEDIO	13 (48.1)	16 (59.3)	
Acumulado_2014_1 (mean (sd))	3.40 (0.38)	3.52 (0.39)	0.284
Nota.Max.CalculoI (mean (sd))	3.28 (0.26)	3.34 (0.38)	0.459
Horas.Total.CalculoI (mean (sd))	2.87 (6.95)	3.37 (7.20)	0.796
Ben (%)			<0.001

(-1,0]	27 (100.0)	0 (0.0)		
(0,7]	0 (0.0)	16 (59.3)		
(7,79]	0 (0.0)	11 (40.7)		
distance (mean (sd))	0.17 (0.11)	0.17 (0.12)	0.950	

Se observa una leve disminución en la nota promedio de cálculo II en el grupo estudiantes que reciben tutorías pero ésta no es significativa desde el punto de vista estadístico (valor  $p = 0.517$ ). En las demás variables no se aprecian diferencias significativas, asegurando homogeneidad de los dos grupos que se comparan y, debido a que los dos grupos tienen el mismo tamaño, se aumenta la potencia estadística de la comparación. La Figura 61 muestra en impacto en función del número de horas para las dos categorías que fue posible realizar la evaluación.

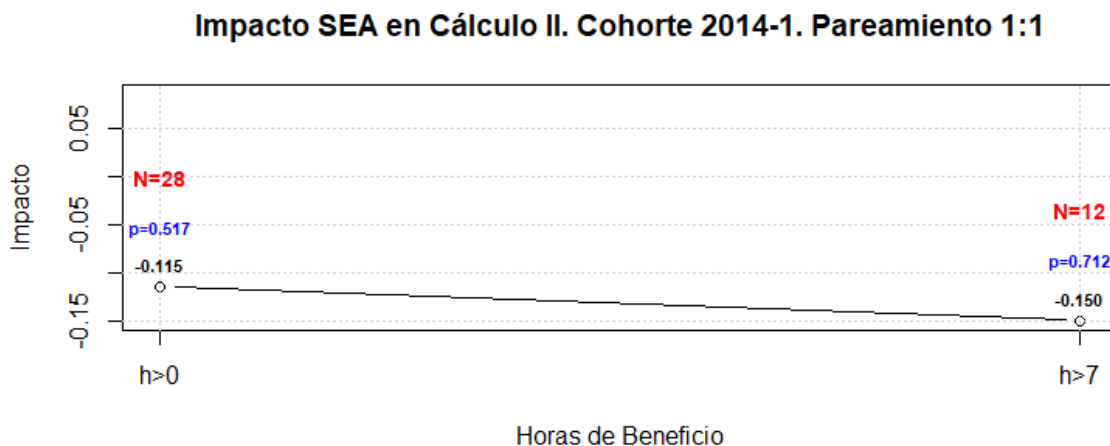


Figura 61. Impacto del programa SEA sobre cálculo II en función del número de horas de beneficio. 1ª cohorte de 2014

Se observa un impacto negativo en los dos escenarios y una tendencia decreciente aunque no significativa.

En síntesis se puede decir que la técnica de pareamiento no mostró un impacto significativo en la nota de cálculo II atribuida al uso del programa SEA a un nivel de significancia del 52% en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2014.

**3.2.1.6 Impacto del SEA sobre repitentes en cálculo II. Cohorte 2014-1.** De los 921 estudiantes de ciencia e ingeniería en la primera cohorte del 2014 se encontró que el 22% (202) de los estudiantes cursaron cálculo II dos veces: 154 sin beneficio, 7 con beneficio los dos semestres y los 41 restantes con beneficio la primera o la segunda vez que cursaron la materia.

La Figura 62 muestra el uso del método de diferencias en diferencias en los estudiantes repitentes de cálculo II. A diferencia de los casos anteriores, como las dos líneas se cruzan, el impacto del programa SEA, estimado por el método de diferencias en diferencias puede ser significativo.



Figura 62. Diferencias en diferencias a repitentes en cálculo II. 1ª cohorte de 2014

No obstante, el estimador del impacto del programa SEA junto con la prueba estadística de nulidad se resume en la Tabla 35, la cual indica que el impacto es significativo al 27%.

Tabla 35.

*Estimador del impacto del programa SEA sobre repitentes de cálculo II usando diferencias en diferencias. 1ª cohorte de 2014.*

Estimate	Std. Error	t value	Pr(> t )
-0.3896104	0.3554057	-1.0962412	0.2738027

En síntesis se puede decir que el método de diferencias en diferencias arroja un impacto negativo (0.390 unidades) significativo al 27% sobre la nota de los estudiantes repitentes de cálculo II que ingresaron a la UIS en la primera cohorte del 2014. No obstante, se debe destacar que con el número tan reducido de estudiantes beneficiados (7 estudiantes), estos resultados no se pueden considerar concluyentes. Además, por el hecho que ninguno de estos estudiantes recibió más de 7 horas de beneficio en el semestre.

**3.2.1.7 Impacto del SEA sobre cálculo III. Cohorte 2014-1.** De los 921 estudiantes matriculados en programas de ciencia e ingeniería en el primer semestre del 2014, 478 tienen nota de cálculo III: 436 con beneficio del SEA y 42 sin beneficio pero solo 265 de ellos con información completa en todas las variables: 249 sin beneficio y 15 con beneficio del SEA en cálculo III.

La Figura 63 muestra la distribución de horas de beneficio en cálculo III para los estudiantes de ciencia e ingeniería de la primera cohorte del 2014. Para poder aplicar la técnica de pareamiento, fue necesario fusionar las categorías con menos de 4 estudiantes, quedando por lo tanto una sola categoría con 31 estudiantes a la que se aplicó la metodología

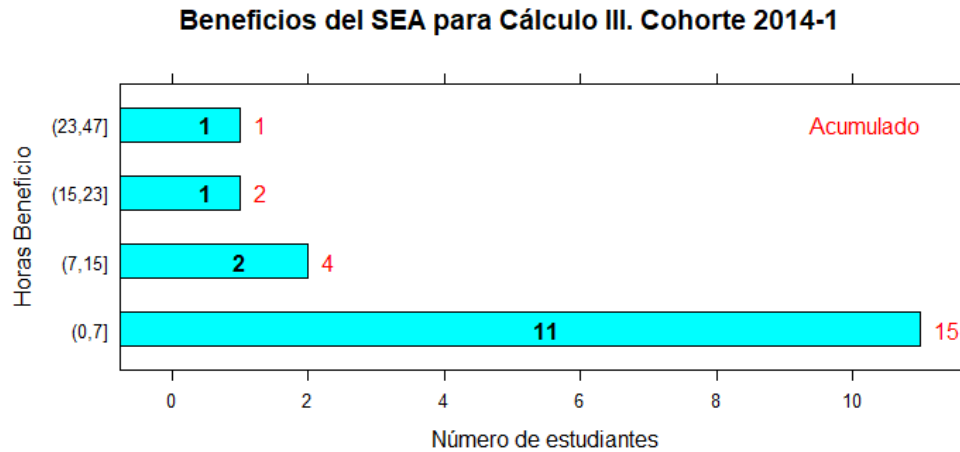


Figura 63. Beneficios del SEA en cálculo III con información completa. 1ª cohorte de 2014.

La Tabla 36 resume los estadísticos más importantes de las variables cuantitativas de los estudiantes de la primera cohorte del 2014, incluyendo la cantidad de datos no disponibles (ND's) para cada variable en la última columna, siendo la nota de cálculo I y el promedio del semestre anterior a cálculo III, las que mayor cantidad de datos faltantes presentan.

Tabla 36.

*Resumen de variables cuantitativas para el modelo de propensión de cálculo III. 1ª cohorte de 2014*

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	ND's
Cálculo III	0.0	3.0	3.2	3.2	3.7	5.0	443
Ben. Cálculo III	0.0	0.0	0.0	0.3	0.0	44.0	0
Edad	16.0	17.0	18.0	18.4	19.0	57.0	16
Matemática	40.0	62.0	68.0	69.3	74.0	114.0	40
Lenguaje	40.0	58.0	60.0	60.9	64.0	89.0	40
Inglés	7.0	48.0	57.3	60.9	71.0	100.0	7
Física	31.0	59.0	65.0	66.0	70.0	119.0	31
Química	39.0	57.5	62.0	64.0	69.0	92.0	39
Bilología	34.0	53.0	58.0	58.4	63.0	84.0	34
Promedio 2014-1	0.0	3.0	3.4	3.2	3.7	4.8	41
Cálculo I	3.0	3.1	3.3	3.5	3.7	5.0	338
Ben. Cálculo I	0.0	0.0	0.0	2.2	0.0	94.0	0
Promedio 2014-2	1.8	3.2	3.4	3.5	3.8	4.7	235
Cálculo II	3.0	3.1	3.3	3.4	3.7	5.0	388
Ben. Cálculo II	0.0	0.0	0.0	1.4	0.0	94.0	0

De los 921 estudiantes que ingresaron en la primera cohorte del 2014 solamente 265 presentaron información completa en todas las variables requeridas para la evaluación que equivale a 28.77%, de los cuales, 15 obtuvieron al menos una hora de beneficio del SEA al cursar cálculo III.

La Tabla 37 muestra la comparación de los estudiantes que recibieron beneficio del SEA en cálculo III con los que no recibieron. Además del desbalance de los dos grupos que como se ha mencionado, reduce la potencia estadística de la comparación, se destaca también diferencias significativa en co-variables como la edad (valor  $p=0.024$ ), el puntaje en biología (valor  $p=0.013$ ), el riesgo en salud (valor  $p=0.002$ ) y las horas de beneficio del SEA recibidas en cálculo II (valor  $p < 0.001$ ) que, así como el beneficio del SEA en cálculo III, también pueden impactar el desempeño del estudiante en esta asignatura.

Tabla 37.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo III. 1ª cohorte de 2014*

	No Beneficiado	Beneficiado	p
n	250	15	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
Nota (mean (sd))	3.30 (0.74)	2.97 (0.78)	0.092
Horas (mean (sd))	0.00 (0.00)	7.27 (10.71)	<0.001
Prog (%)			0.538
Civ	29 (11.6)	0 (0.0)	
Ele	13 (5.2)	1 (6.7)	
Elo	23 (9.2)	1 (6.7)	
Geo	15 (6.0)	2 (13.3)	
Ind	31 (12.4)	3 (20.0)	
IQu	34 (13.6)	2 (13.3)	
LMa	3 (1.2)	1 (6.7)	
Mat	1 (0.4)	0 (0.0)	
Mec	30 (12.0)	0 (0.0)	
Met	18 (7.2)	2 (13.3)	
Pet	28 (11.2)	1 (6.7)	
Qui	9 (3.6)	0 (0.0)	
Sis	16 (6.4)	2 (13.3)	

Edad (mean (sd))	17.76 (0.94)	18.40 (2.32)	0.024	
Genero = Hombre (%)	156 (62.4)	8 (53.3)	0.668	
Matematica (mean (sd))	72.72 (9.12)	68.94 (8.58)	0.119	
Lenguaje (mean (sd))	62.88 (4.77)	61.94 (6.09)	0.467	
Ingles (mean (sd))	65.37 (15.47)	63.55 (14.17)	0.657	
Fisica (mean (sd))	69.69 (9.96)	65.73 (8.61)	0.134	
Quimica (mean (sd))	67.66 (8.91)	66.45 (11.03)	0.612	
Biologia (mean (sd))	60.55 (7.78)	55.40 (6.59)	0.013	
Riesgo_Economica (%)			0.826	
ALTO	16 (6.4)	1 (6.7)		
BAJO	59 (23.6)	5 (33.3)		
MEDIO	105 (42.0)	6 (40.0)		
MEDIO ALTO	70 (28.0)	3 (20.0)		
Riesgo_Social (%)			0.112	
ALTO	6 (2.4)	2 (13.3)		
BAJO	124 (49.6)	7 (46.7)		
MEDIO	90 (36.0)	5 (33.3)		
MEDIO ALTO	30 (12.0)	1 (6.7)		
Riesgo_Salud (%)			0.002	
BAJO	224 (89.6)	9 (60.0)		
MEDIO	25 (10.0)	6 (40.0)		
MEDIO ALTO	1 (0.4)	0 (0.0)		
Riesgo.Academico = BAJO (%)	43 (17.2)	3 (20.0)	1.000	
Riesgo_Cognitiva (%)			0.543	
ALTO	82 (32.8)	7 (46.7)		
BAJO	20 (8.0)	1 (6.7)		
MEDIO	148 (59.2)	7 (46.7)		
Acumulado_2014_1 (mean (sd))	3.69 (0.44)	3.65 (0.55)	0.723	
Nota.Max.CalculoI (mean (sd))	3.58 (0.50)	3.47 (0.48)	0.439	
Horas.Total.CalculoI (mean (sd))	1.79 (6.15)	3.73 (6.49)	0.236	
Acumulado_2014_2 (mean (sd))	3.69 (0.38)	3.65 (0.41)	0.708	
Nota.Max.CalculoII (mean (sd))	3.52 (0.46)	3.50 (0.60)	0.879	
Horas.Total.CalculoII (mean (sd))	1.43 (5.58)	7.67 (14.51)	<0.001	
Ben = (0,47] (%)	0 (0.0)	15 (100.0)	<0.001	

El resumen del modelo de regresión depurado con las variables relevantes se presenta en Tabla 38

Tabla 38.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo III. 1ª cohorte de 2014 (Modelo depurado)*

```
Call:
glm(formula = Beneficio ~ Biologia + Riesgo_Salud + Horas.Total.CalculoII,
     family = binomial, data = Datos.Evaluacion2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1898  -0.3291  -0.2437  -0.1962   3.0214

Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	1.75668	2.54072	0.691	0.4893
Biología	-0.08765	0.04472	-1.960	0.0500 .
Riesgo_SaludMEDIO	1.44595	0.60256	2.400	0.0164 *
Riesgo_SaludMEDIO ALTO	-12.48259	1455.39759	-0.009	0.9932
Horas.Total.CalculoII	0.04926	0.02395	2.057	0.0397 *

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 115.285 on 264 degrees of freedom  
 Residual deviance: 97.704 on 260 degrees of freedom  
 AIC: 107.7

Number of Fisher Scoring iterations: 14

Las variables más significativas en el modelo de propensión depurado fueron el puntaje en biología en las pruebas SABER 11, el riesgo en salud y las horas de beneficio en cálculo II recibidas el semestre anterior.

La Figura 64 muestra el diagrama de puntos con los resultados del pareamiento. Se observan 2 estudiantes que recibieron beneficio del programa SEA en cálculo III que no pudieron ser pareados con estudiantes que no recibieron beneficios debido a que no se encuentran dentro del rango común de propensión. La parte inferior muestra los estudiantes sin beneficio que no fueron empleados en el pareamiento.

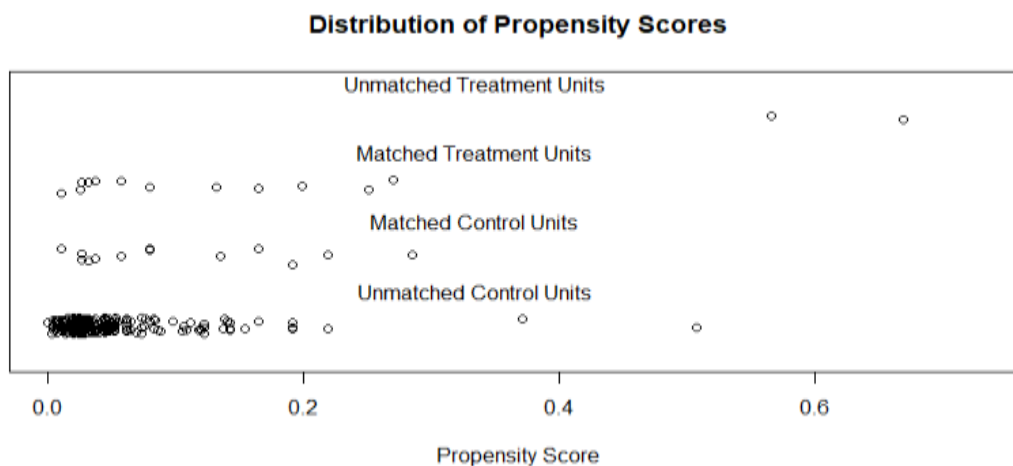


Figura 64. Pareamiento por puntajes de propensión para cálculo II. 1ª cohorte de 2014

Los resultados finales del pareamiento para los 15 estudiantes beneficiados por el programa SEA en cálculo III se resumen en la Tabla 39. Se descartan 17 estudiantes sin beneficio del SEA en cálculo III por estar fuera del rango común y se ignoraron los 220 restantes.

Tabla 39.

*Resultados finales del pareamiento para beneficiados en cálculo II. 1ª cohorte de 2014 de Ciencias e Ingeniería*

	Control	Treated
All	250	15
Matched	13	13
Unmatched	220	0
Discarded	17	2

La Figura 65 muestra el histograma de los puntajes de propensión antes y después del pareamiento. A la izquierda se aprecian las diferencias debido a la heterogeneidad de los dos grupos antes del pareamiento mientras que a la derecha, después del pareamiento, la similitud de las distribuciones indica un pareamiento satisfactorio por puntajes de propensión.

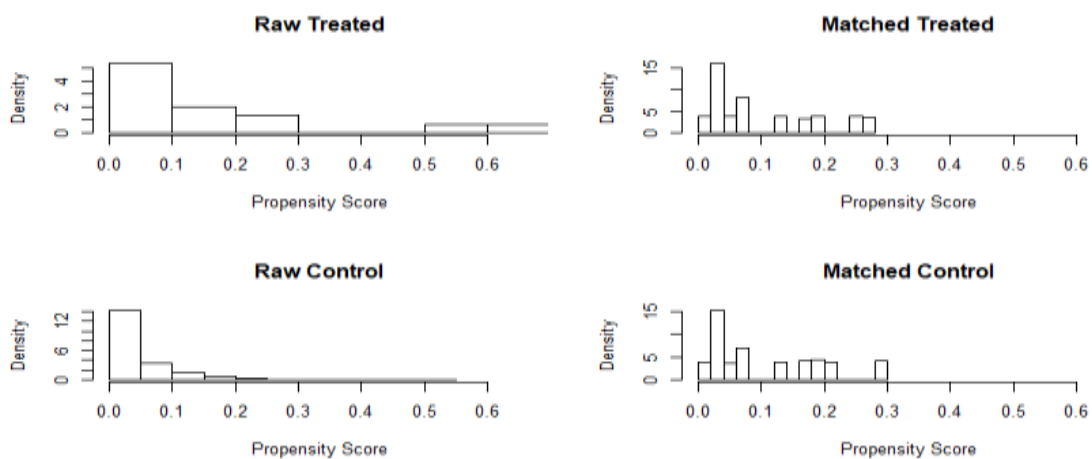


Figura 65. Histograma de puntajes de propensión para cálculo II antes y después del pareamiento.

1ª cohorte de 2014

Tanto la información numérica como gráfica indica un pareamiento satisfactorio con 13 estudiantes beneficiados que se pudieron parrear, 2 estudiantes beneficiados excluidos del pareamiento por estar fuera del rango común y una distribución similar de puntajes de propensión que asegura la formación de grupos homogéneos y por lo tanto estimaciones apropiadas del contrafactual de cada estudiante beneficiado. No obstante, es importante mencionar que debido a la cantidad tan pequeña de estudiantes beneficiados, las comparaciones derivadas del uso de la técnica no sean concluyentes.

La Tabla 40 compara los estadísticos más importantes entre el grupo de estudiantes con beneficio y sin beneficio del programa SEA en cálculo III después de aplicar la técnica de pareamiento. Se observa, que los dos grupos resultan comparables en todas las co-variables por lo que se puede asumir, que cualquier diferencia en la nota promedio de cálculo III se puede atribuir a los beneficios recibidos del programa SEA.

Tabla 40.

*Comparación de estudiantes con beneficio o no del SEA en cálculo III (después de pareamiento).*

*1ª cohorte de 2014*

	No Beneficiado	Beneficiado	p
n	13	13	
Nota (mean (sd))	3.21 (0.90)	2.98 (0.84)	0.505
Edad (mean (sd))	17.62 (1.12)	18.46 (2.50)	0.277
Genero = Hombre (%)	7 (53.8)	6 (46.2)	1.000
Matemática (mean (sd))	70.69 (11.54)	69.62 (8.67)	0.792
Lenguaje (mean (sd))	62.08 (5.36)	60.86 (5.44)	0.571
Inglés (mean (sd))	58.92 (16.21)	64.64 (14.92)	0.359
Física (mean (sd))	69.54 (12.00)	65.38 (9.21)	0.332
Química (mean (sd))	61.38 (6.45)	65.36 (11.46)	0.286
Biología (mean (sd))	55.15 (6.93)	55.69 (6.99)	0.846
Riesgo_Economica (%)			0.594
ALTO	1 (7.7)	1 (7.7)	
BAJO	2 (15.4)	5 (38.5)	

MEDIO	5 (38.5)	4 (30.8)	
MEDIO ALTO	5 (38.5)	3 (23.1)	
Riesgo_Social (%)			0.785
ALTO	2 (15.4)	1 (7.7)	
BAJO	8 (61.5)	7 (53.8)	
MEDIO	2 (15.4)	4 (30.8)	
MEDIO ALTO	1 (7.7)	1 (7.7)	
Riesgo_Salud (%)			NaN
BAJO	9 (69.2)	9 (69.2)	
MEDIO	4 (30.8)	4 (30.8)	
MEDIO ALTO	0 (0.0)	0 (0.0)	
Riesgo.Academico = BAJO (%)	0 (0.0)	2 (15.4)	0.462
Riesgo_Cognitiva (%)			0.580
ALTO	7 (53.8)	7 (53.8)	
BAJO	0 (0.0)	1 (7.7)	
MEDIO	6 (46.2)	5 (38.5)	
Acumulado_2014_1 (mean (sd))	3.52 (0.49)	3.64 (0.59)	0.600
Nota.Max.CalculoI (mean (sd))	3.41 (0.43)	3.48 (0.52)	0.682
Horas.Total.CalculoI (mean (sd))	3.12 (4.19)	4.00 (6.92)	0.697
Acumulado_2014_2 (mean (sd))	3.57 (0.34)	3.69 (0.43)	0.449
Nota.Max.CalculoII (mean (sd))	3.44 (0.42)	3.54 (0.63)	0.638
Horas.Total.CalculoII (mean (sd))	1.62 (5.82)	2.69 (6.69)	0.665
Ben = (0,47] (%)	0 (0.0)	13 (100.0)	<0.001
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
distance (mean (sd))	0.10 (0.09)	0.11 (0.09)	0.966

Se observa una leve disminución en la nota promedio de cálculo III en el grupo estudiantes que reciben tutorías pero ésta no es significativa desde el punto de vista estadístico (valor  $p = 0.505$ ).

Debido a la cantidad tan baja de estudiantes beneficiados en cálculo III y a que solo resultaron significativas en el modelo de propensión 3 variables (Biología, riesgo en salud y beneficio previo en cálculo II) el modelo de propensión no siempre lleva a un pareamiento único.

Para asegurar que las conclusiones no cambian al repetir el pareamiento, se realizaron 100 corridas para determinar el impacto del programa SEA, medido como la diferencia promedio entre el grupo de beneficiado y sus correspondientes pares no beneficiados en cálculo III. Los resultados se resumen en la Figura 66, el eje horizontal muestra la distribución de diferencias promedio entre

cada beneficiado y el promedio de los respectivos pares en el grupo de no beneficiados (impacto) mientras que en el eje vertical se aprecia la distribución del valor p de la prueba t correspondiente.

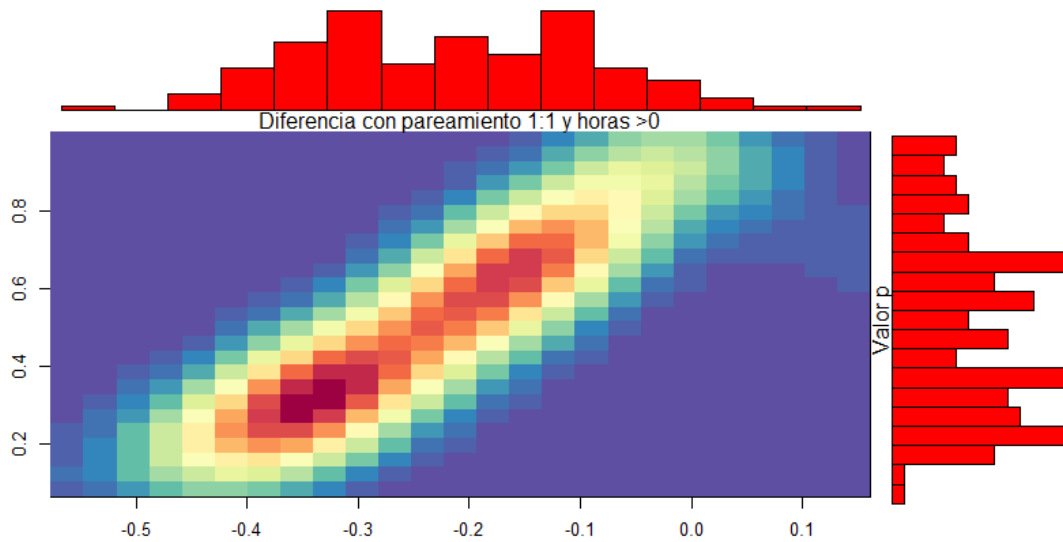


Figura 66. Histograma bi-variado de 100 diferencias promedio en notas de cálculo III entre beneficiados y no beneficiados del SEA con pareamiento 1:1. 1ª cohorte de 2014

Se observa que el 96% de las corridas realizadas, la técnica de pareamiento lleva a un impacto negativo y no significativo al 5% del programa SEA sobre cálculo III. El 4% de los casos restantes es positivo pero sin ser significativo. Los detalles numéricos de estas corridas se muestran en la Tabla 41.

Tabla 41.

*Diferencias promedio en notas de cálculo III entre beneficiados y no beneficiados del SEA en 100 corridas de pareamiento 1:1. 1ª cohorte de 2014*

Distribución de impacto	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	-0.5615	-0.3462	-0.2385	-0.2328	-0.1385	0.1462
Distribución del valor p	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	0.08473	0.33170	0.49390	0.51040	0.68780	0.98080

Debido a la cantidad tan limitada de estudiantes beneficiados de cálculo III en la primera cohorte del 2014, no fue posible evaluar el impacto del programa SEA en función del número de horas recibidas. No obstante, la Figura 67 muestra los resultados obtenidos en la única categoría que se aplicó la técnica de pareamiento.

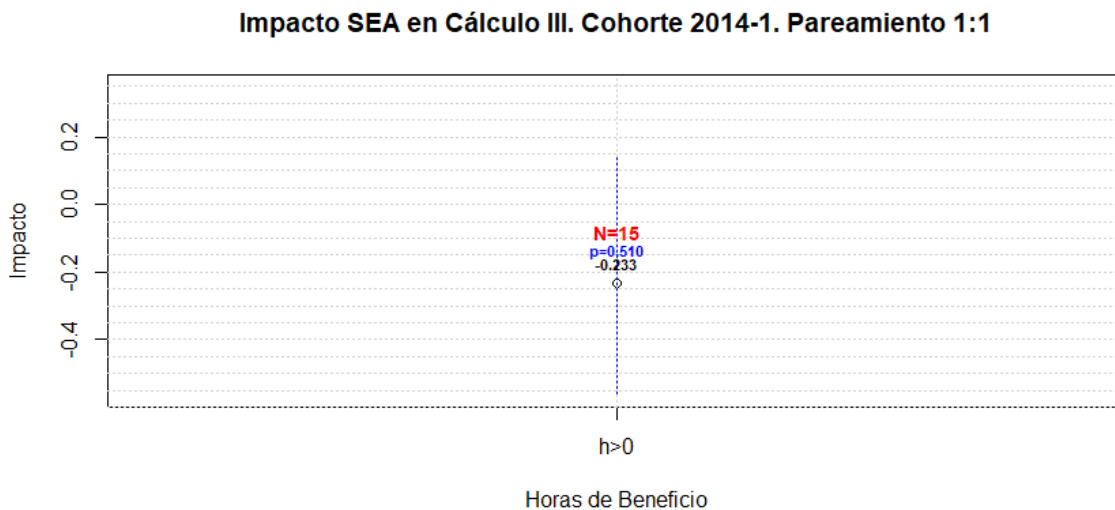


Figura 67. Impacto del programa SEA sobre cálculo III en función del número de horas de beneficio. 1ª cohorte de 2014

En síntesis, se puede decir que la técnica de pareamiento mostró un impacto negativo del programa SEA sobre cálculo III en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2014 aunque sin ser significativo al 5%.

**3.2.1.8 Impacto del SEA sobre ecuaciones diferenciales. Cohorte 2014-1.** Con información completa de ecuaciones diferenciales en la primera cohorte del 2014 solo se encuentran 93 estudiantes: 9 con beneficio del SEA: 8 estudiantes con menos de 8 horas de beneficio y uno con

23 horas de beneficio que no aprobó ecuaciones diferenciales (con una nota final de 2.3 en la asignatura).

Debido a la cantidad tan pequeña de estudiantes beneficiados, no se encontró un rango común en los puntajes de propensión entre el grupo de beneficiado y no beneficiados que permitiera hacer usos de la técnica de pareamiento. Se encontró una diferencia promedio positiva de 0.3 unidades entre el grupo de beneficiados y no beneficiados, aunque esta no es significativo desde el punto de vista estadístico (valor  $p = 0.217$ ). No obstante, se basa en una comparación desbalanceada entre 9 estudiantes beneficiados y 73 no beneficiados, lo que reduce la potencia estadística de la comparación debido al tamaño de una de las muestras y el desbalance de los dos grupos comparados. Además, debido a que se encontraron diferencias significativas en otras co-variables que también pueden afectar el desempeño en ecuaciones diferenciales, los resultados no son concluyentes y por lo tanto no se continuó con el análisis.

**3.2.1.9 Síntesis de resultados para la primera cohorte del 2014.** En síntesis se destacan las siguientes conclusiones, después de aplicar las técnicas de evaluación de impacto sobre desempeño de los estudiantes en asignaturas de matemáticas que ingresaron a la UIS en la primera cohorte del 2014:

- La técnica de pareamiento no mostró un impacto significativo sobre el desempeño académico de los estudiantes en asignaturas de matemáticas (cálculos, álgebra y ecuaciones diferenciales) de ciencia e ingeniería que ingresaron a la UIS en la primera cohorte del 2014.
- Solo fue posible evaluar el impacto del programa SEA en función del número de horas de beneficio recibidas para cálculo I y álgebra lineal, observando un leve aumento del impacto

por encima de 24 horas de beneficio recibidas en el semestre en cálculo I aunque éste no resultó significativo estadísticamente. En álgebra, la tendencia es decreciente aunque no significativa y en las demás asignaturas no fue posible realizar este análisis debido a la cantidad tan baja de estudiantes con información completa disponible para esta cohorte.

- Las co-variables con mayor cantidad de información incompleta fueron las 5 dimensiones de riesgo del SEA. Esto dificultó el uso de la técnica de pareamiento debido a que llevó a la eliminación de un porcentaje importante de datos en cada evaluación, por no presentar la información completa, con la consecuente reducción de potencia estadística en las técnicas de pareamiento.
- La técnica de diferencias en diferencias no mostró impacto significativo en la nota de los estudiantes repitentes de cálculo I, álgebra y cálculo II. No obstante, con el número tan reducido de estudiantes beneficiados en cada caso, estos resultados no se pueden considerar concluyentes razón por la cual no se pudo aplicar la técnica a las demás asignaturas matemáticas.
- Debido a la cantidad tan pequeña de estudiantes beneficiados, no fue posible evaluar el impacto del programa SEA en función del número de horas para los repitentes en las asignaturas de matemáticas (cálculos, álgebra y ecuaciones diferenciales) de la primera cohorte del 2014.

**3.2.2 Impacto en cohorte 2015-1.** En esta sección se realiza la evaluación de impacto del programa SEA sobre cálculo y álgebra para los estudiantes que ingresaron a la UIS en la primera cohorte del 2015.

**3.2.2.1 Impacto del SEA sobre cálculo I. Cohorte 2015-1.** De los 908 estudiantes de ciencia e ingeniería que ingresaron a la universidad en esta cohorte, 541 muestran información completa (60%), 116 con beneficio del SEA y 425 sin él. La distribución de beneficios por horas se muestra en la Figura 68.

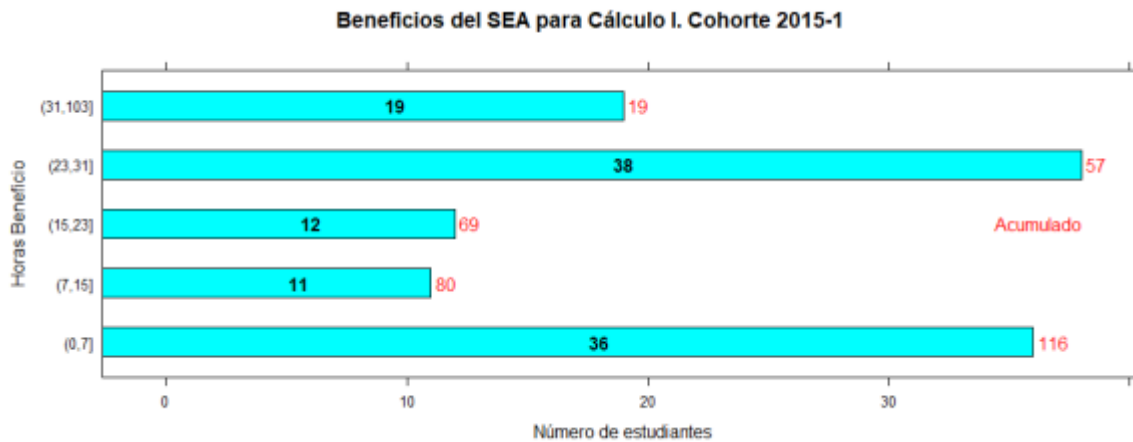


Figura 68. Beneficios del SEA en cálculo I. 1ª cohorte de 2015.

La

Tabla 42 muestra la comparación estadística entre el grupo de beneficiados y no beneficiados de SEA en cálculo I.

Se observa una diferencia significativa al 9.3% en la nota promedio de cálculo I la cual se puede atribuir al beneficio del SEA así como a otras co-variables o factores en los que también se encuentran diferencias significativas como el programa académico (valor  $p < 0.001$ ), la edad (valor  $p = 0.092$ ) y el género (valor  $p = 0.012$ ).

Tabla 42.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2015*

	No Beneficiado	Beneficiado	p
n	425	116	
Nota (mean (sd))	2.57 (1.08)	2.75 (0.83)	0.093
Horas (mean (sd))	0.00 (0.00)	19.95 (16.26)	<0.001
Prog (%)			<0.001
Civ	38 (8.9)	19 (16.4)	
DIn	21 (4.9)	1 (0.9)	
ELe	28 (6.6)	3 (2.6)	
ELo	43 (10.1)	2 (1.7)	
Geo	19 (4.5)	8 (6.9)	
Ind	38 (8.9)	15 (12.9)	
IQu	38 (8.9)	16 (13.8)	
LMa	21 (4.9)	1 (0.9)	
Mat	25 (5.9)	3 (2.6)	
Mec	42 (9.9)	13 (11.2)	
Met	33 (7.8)	5 (4.3)	
Pet	25 (5.9)	4 (3.4)	
Qui	17 (4.0)	13 (11.2)	
Sis	37 (8.7)	13 (11.2)	
Edad (mean (sd))	17.88 (1.11)	17.67 (0.83)	0.062
Genero = Mujer (%)	144 (33.9)	49 (42.2)	0.120
Matemática (mean (sd))	75.43 (8.28)	75.63 (7.08)	0.814
Lenguaje (mean (sd))	68.30 (8.24)	68.69 (6.66)	0.636
Inglés (mean (sd))	71.18 (14.11)	69.78 (13.39)	0.341
Física (mean (sd))	74.06 (7.52)	74.46 (6.79)	0.609
Química (mean (sd))	74.00 (7.53)	74.46 (6.79)	0.555
Biología (mean (sd))	74.00 (7.56)	74.46 (6.79)	0.556
Riesgo_Economica = MEDIO ALTO (%)	159 (37.4)	46 (39.7)	0.739
Riesgo_Social (%)			0.922
ALTO	19 (4.5)	6 (5.2)	
BAJO	168 (39.5)	43 (37.1)	
MEDIO	170 (40.0)	46 (39.7)	
MEDIO ALTO	68 (16.0)	21 (18.1)	
Riesgo_Salud (%)			0.326
BAJO	345 (81.2)	87 (75.0)	
MEDIO	78 (18.4)	28 (24.1)	
MEDIO ALTO	2 (0.5)	1 (0.9)	
Riesgo.Academico (%)			0.615
ALTO	117 (27.5)	39 (33.6)	
BAJO	10 (2.4)	1 (0.9)	
MEDIO ALTO	185 (43.5)	46 (39.7)	

MEDIO BAJO	76 (17.9)	19 (16.4)	
MUY ALTO	37 (8.7)	11 (9.5)	
Riesgo_Cognitiva (%)			0.311
ALTO	100 (23.5)	34 (29.3)	
BAJO	52 (12.2)	10 (8.6)	
MEDIO	273 (64.2)	72 (62.1)	
Ben (%)			<0.001
(-1,0]	425 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	36 (31.0)	
(7,15]	0 (0.0)	11 (9.5)	
(15,23]	0 (0.0)	12 (10.3)	
(23,31]	0 (0.0)	38 (32.8)	
(31,103]	0 (0.0)	19 (16.4)	

El modelo de propensión depurado se destalla en la Tabla 43, encontrando que las variables que más afectan la participación en el programa SEA fueron el programa académico del estudiante, la edad, el nivel de inglés en la pruebas SABER 11 y el riesgo cognitivo.

Tabla 43.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2015 (Modelo depurado.)*

```
Call:
glm(formula = Beneficio ~ Prog + Edad + Ingles + Riesgo_Cognitiva,
     family = binomial, data = Datos.Evaluacion2)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.2967	-0.7652	-0.5006	-0.2587	2.7947

Coefficients:

Estimate	Std. Error	z value	Pr(> z )
(Intercept)	6.267063	2.469055	2.538 0.01114 *
ProgDIn	-2.332043	1.068859	-2.182 0.02912 *
ProgEle	-1.802898	0.681948	-2.644 0.00820 **
ProgElo	-2.392442	0.780824	-3.064 0.00218 **
ProgGeo	0.032101	0.531432	0.060 0.95183
ProgInd	-0.306044	0.428790	-0.714 0.47539
ProgIQu	-0.163112	0.427107	-0.382 0.70254
ProgLMa	-2.952360	1.083449	-2.725 0.00643 **
ProgMat	-2.024913	0.705681	-2.869 0.00411 **
ProgMec	-0.297691	0.437527	-0.680 0.49626
ProgMet	-1.504921	0.579847	-2.595 0.00945 **
ProgPet	-0.911191	0.623767	-1.461 0.14407
ProgQui	0.239357	0.478517	0.500 0.61693
ProgSis	-0.498411	0.442459	-1.126 0.25997

```

Edad          -0.269148  0.125023  -2.153  0.03134 *
Ingles        -0.024403  0.009976  -2.446  0.01443 *
Riesgo_CognitivaBAJO -0.849268  0.434965  -1.952  0.05088 .
Riesgo_CognitivaMEDIO -0.474370  0.267753  -1.772  0.07645 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

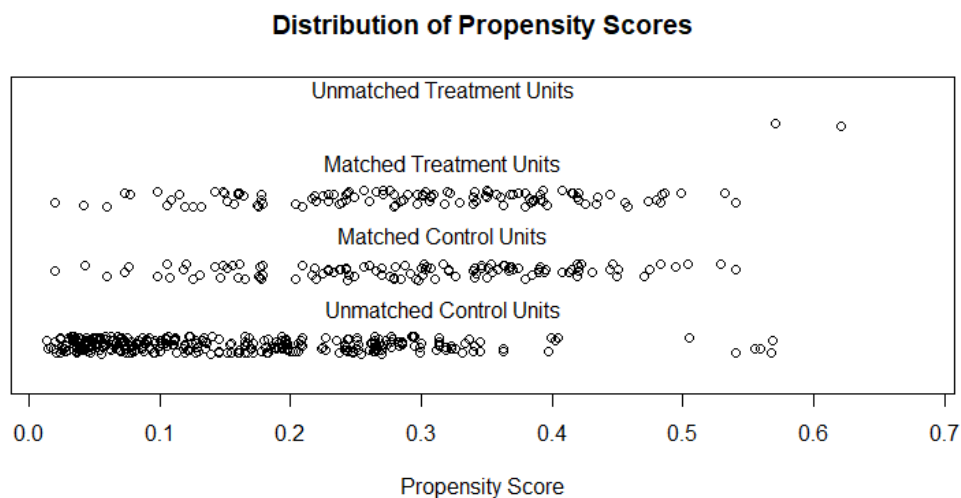
Null deviance: 562.37 on 540 degrees of freedom
Residual deviance: 502.01 on 523 degrees of freedom
AIC: 538.01

Number of Fisher Scoring iterations: 5

```

La Figura 69 muestra a dos estudiantes beneficiados que se excluyeron del proceso de pareamiento debido a que se encontraban fuera del rango común y la

Tabla 44 muestra los resultados finales del pareamiento 1:1 considerando los 116 estudiantes beneficiarios del SEA en esta asignatura.



*Figura 69.* Pareamiento por puntajes de propensión para cálculo I. 1<sup>a</sup> cohorte de 2015

Tabla 44.

*Resultados finales del pareamiento para beneficiados en cálculo I. 1<sup>a</sup> cohorte de 2015 de Ciencias e Ingeniería.*

	Control	Treated
All	425	116
Matched	114	114
Unmatched	307	0
Discarded	4	2

La Figura 70 muestra la distribución de puntajes de propensión entre el grupo de beneficiarios (arriba) y no beneficiados (abajo) del SEA en cálculo I antes (izquierda) y después (derecha) del proceso de pareamiento. Se observa una distribución muy diferente antes del pareamiento (izquierda) la cual sugiere la heterogeneidad de los dos grupos mientras que después del pareamiento (derecha) se observa que las distribuciones son más parecidas, indicando mayor homogeneidad de los dos grupos.

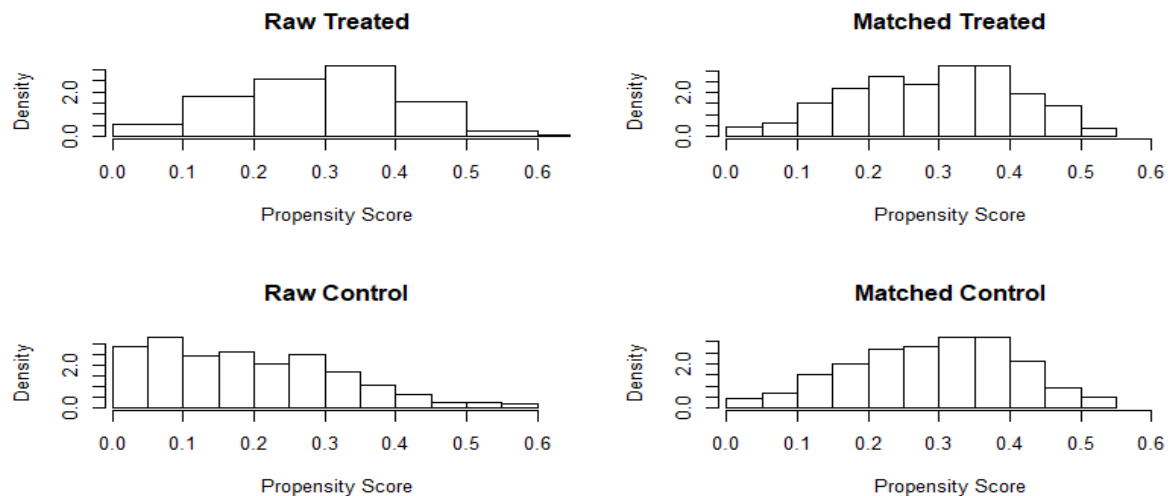


Figura 70. Histograma de puntajes de propensión para cálculo I antes y después del pareamiento.

1ª cohorte de 2015

Tanto desde el punto de vista gráfico como numérico, la técnica de pareamiento lleva a dos grupos muy similares en todas sus propiedades, excepto porque uno de ellos recibió beneficio del programa SEA en cálculo I y el otro no.

La Tabla 45 muestra la comparación de los dos grupos después del pareamiento. Se observa que no existe un impacto significativo del programa SEA y que los dos grupos resultan similares en todas las co-variables

Tabla 45.

*Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2015*

	No Beneficiado	Beneficiado	p
n	114	114	
Nota (mean (sd))	2.75 (1.07)	2.75 (0.83)	0.972
Horas (mean (sd))	0.00 (0.00)	19.86 (16.38)	<0.001
Prog (%)			0.979
Civ	16 (14.0)	19 (16.7)	
DIn	0 (0.0)	1 (0.9)	
Ele	2 (1.8)	3 (2.6)	
Elo	1 (0.9)	2 (1.8)	
Geo	8 (7.0)	8 (7.0)	
Ind	17 (14.9)	15 (13.2)	
IQu	22 (19.3)	16 (14.0)	
LMa	1 (0.9)	1 (0.9)	
Mat	4 (3.5)	3 (2.6)	
Mec	10 (8.8)	13 (11.4)	
Met	4 (3.5)	5 (4.4)	
Pet	7 (6.1)	4 (3.5)	
Qui	8 (7.0)	11 (9.6)	
Sis	14 (12.3)	13 (11.4)	
Edad (mean (sd))	17.74 (1.17)	17.68 (0.84)	0.648
Genero = Mujer (%)	44 (38.6)	47 (41.2)	0.787
Matemática (mean (sd))	76.86 (8.12)	75.78 (7.03)	0.285
Lenguaje (mean (sd))	68.80 (7.94)	68.61 (6.70)	0.850
Inglés (mean (sd))	70.83 (13.22)	70.17 (13.19)	0.703
Biología (mean (sd))	75.13 (6.81)	74.54 (6.81)	0.509
Riesgo_Economica = MEDIO ALTO (%)	39 (34.2)	45 (39.5)	0.492
Riesgo_Social (%)			0.943
ALTO	5 (4.4)	6 (5.3)	
BAJO	47 (41.2)	43 (37.7)	
MEDIO	44 (38.6)	45 (39.5)	
MEDIO ALTO	18 (15.8)	20 (17.5)	
Riesgo_Salud (%)			0.621
BAJO	92 (80.7)	86 (75.4)	
MEDIO	21 (18.4)	27 (23.7)	
MEDIO ALTO	1 (0.9)	1 (0.9)	
Riesgo_Academico (%)			0.481
ALTO	32 (28.1)	39 (34.2)	
BAJO	4 (3.5)	1 (0.9)	
MEDIO ALTO	51 (44.7)	44 (38.6)	

MEDIO BAJO	19 (16.7)	19 (16.7)	
MUY ALTO	8 (7.0)	11 (9.6)	
Riesgo_Cognitiva (%)			0.869
ALTO	35 (30.7)	33 (28.9)	
BAJO	8 (7.0)	10 (8.8)	
MEDIO	71 (62.3)	71 (62.3)	
Ben (%)			<0.001
(-1,0]	114 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	36 (31.6)	
(7,15]	0 (0.0)	11 (9.6)	
(15,23]	0 (0.0)	11 (9.6)	
(23,31]	0 (0.0)	37 (32.5)	
(31,103]	0 (0.0)	19 (16.7)	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
distance (mean (sd))	0.29 (0.12)	0.29 (0.12)	0.926

La Figura 71 muestra el uso de la técnica de pareamiento al modificar el umbral a partir del cual, un estudiante se considera beneficiado del programa SEA. El primer grupo etiquetado como “ $h>0$ ”, corresponde al caso considerado hasta el momento, es decir los 116 estudiantes beneficiados con al menos una hora de beneficio; el segundo grupo ( $h>7$ ) considera beneficiados aquellos que recibieron al menos 8 horas de beneficio en el semestre y así sucesivamente para los demás casos. Se observa que el beneficio no es significativo al 5% en ningún caso, aunque se destaca que empieza aumentar hasta casi 0.2 décimas con más de 16 horas de beneficio, para después decaer nuevamente. Esto sugiere un aparente óptimo entre 16 y 23 horas.

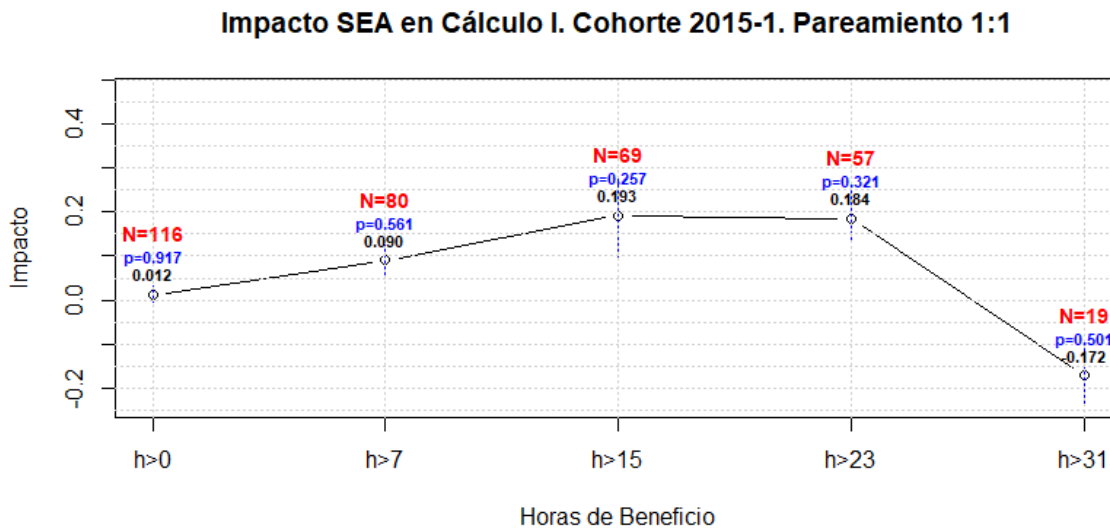
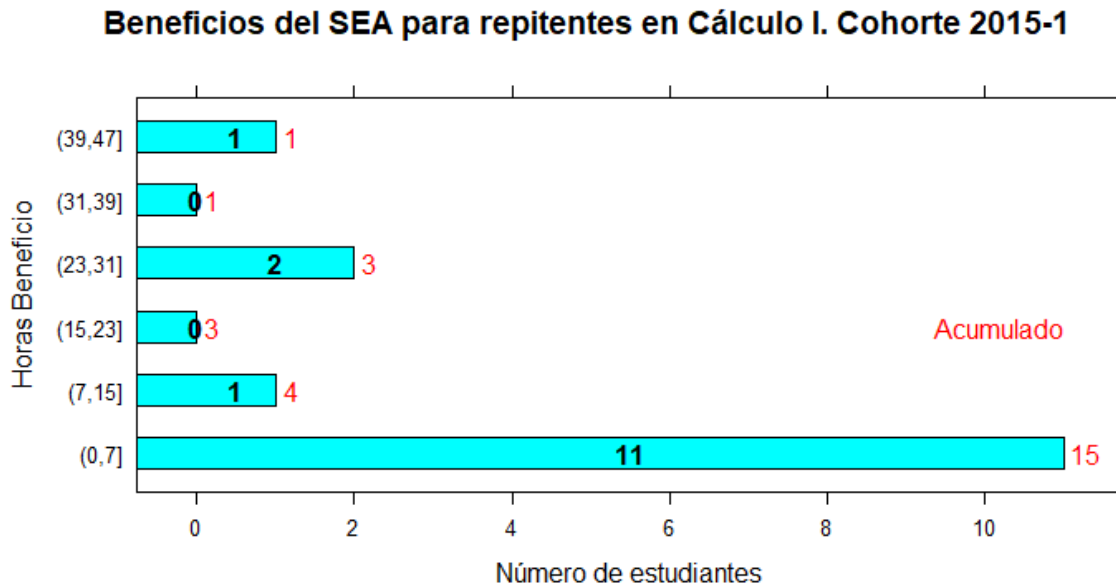


Figura 71. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2015

En síntesis, la técnica de pareamiento muestra un impacto significativo sobre la nota de cálculo I atribuida al uso del programa SEA a un nivel de significancia del 26% en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2015 si el número de horas de beneficio es superior a 16 horas en el semestre y éste es de tan solo 0.2 décimas. Para los demás casos las diferencias son menores y las significancias mucho mayores.

**3.2.2.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2015-1.** De esta cohorte, 292 estudiantes vieron cálculo I dos veces, de los cuales solo 219 cumplen los requisitos para aplicar la técnica de diferencias en diferencias descrita anteriormente (sección 0 en la página 143): 15 con beneficio del SEA y 204 sin él. La distribución de estudiantes repitentes en cálculo I en función del número de horas de beneficio recibidos la segunda vez que curso la materia se muestra en la Figura 72.



*Figura 72.* Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2015.

La Figura 73 muestra el uso de la técnica de diferencias en diferencias para evaluar el impacto del programa SEA la segunda vez que los estudiantes de la primera cohorte del 2015 cursaron cálculo I. Gráficamente se observa que, como las líneas de impacto son casi paralelas en el grupo de beneficiados o y no beneficiados, se puede concluir que el SEA no tiene un impacto diferenciador sobre los repitentes de cálculo I en esta cohorte.

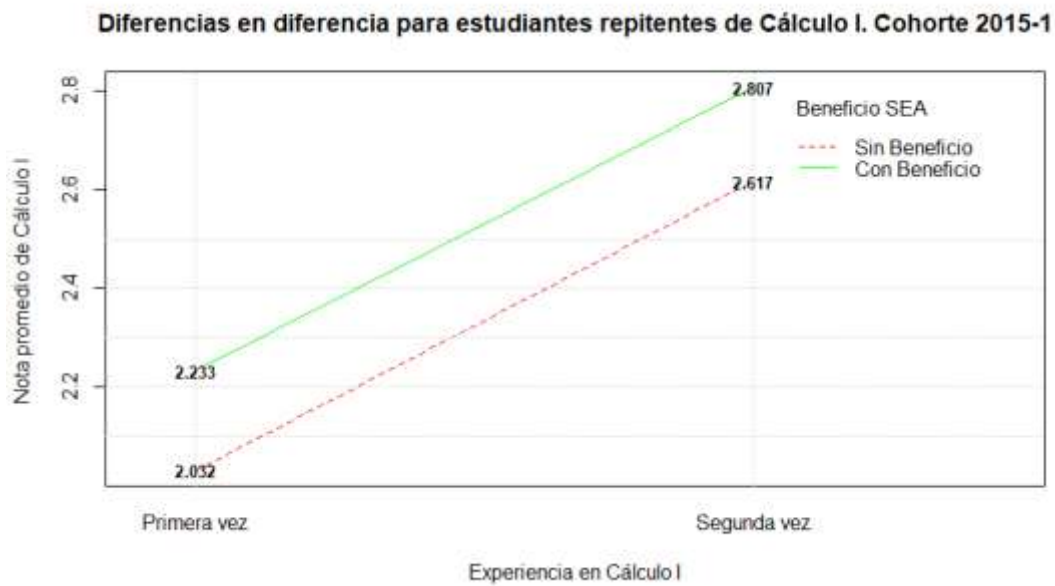


Figura 73. Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2015

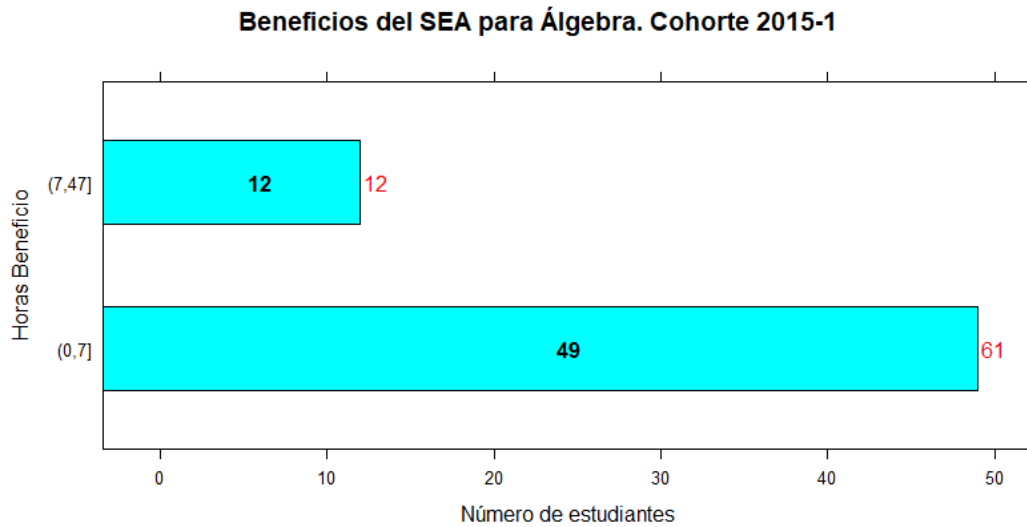
Formalmente, el estimador del impacto no resulta significativamente diferente de cero como se aprecia en la Tabla 46.

Tabla 46.

*Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2015*

Estimate	Std. Error	t value	Pr(> t )
-0.01147059	0.29270731	-0.03918791	0.96875859

**3.2.2.3 Impacto del SEA sobre álgebra lineal. Cohorte 2015-1.** De los 908 estudiantes de la primera cohorte del 2015 solo 542 tenían información completa, 61 con beneficio del SEA en álgebra. La distribución de beneficios por horas se muestra en la Figura 74.



*Figura 74.* Beneficios del SEA en álgebra. 1ª cohorte de 2015.

La Tabla 47 muestra la comparación entre el grupo de beneficiados y no beneficiados del SEA en álgebra Lineal. Se aprecia un aumento de 0.11 unidades en la nota aunque este cambio no fue significativo estadísticamente.

También se observan diferencias en el programa académico (valor  $p < 0.001$ ), el género (valor  $p = 0.043$ ), el puntaje en Lenguaje de las pruebas SABER 11 (valor  $p = 0.083$ ) y el riesgo cognitivo (valor  $p = 0.039$ ).

Tabla 47.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en algebra I. 1ª cohorte de 2015*

	No Beneficiado	Beneficiado	p
--	----------------	-------------	---

n	481	61	
Nota (mean (sd))	3.06 (1.00)	3.17 (1.00)	0.410
Horas (mean (sd))	0.00 (0.00)	5.44 (7.38)	<0.001
Prog (%)			0.002
Civ	43 (8.9)	11 (18.0)	
Ele	24 (5.0)	6 (9.8)	
Elo	39 (8.1)	3 (4.9)	
Fis	22 (4.6)	4 (6.6)	
Geo	27 (5.6)	1 (1.6)	
Ind	46 (9.6)	4 (6.6)	
IQu	49 (10.2)	5 (8.2)	
LMa	23 (4.8)	0 (0.0)	
Mat	29 (6.0)	1 (1.6)	
Mec	53 (11.0)	2 (3.3)	
Met	29 (6.0)	9 (14.8)	
Pet	25 (5.2)	4 (6.6)	
Qui	22 (4.6)	8 (13.1)	
Sis	50 (10.4)	3 (4.9)	
Edad (mean (sd))	17.87 (1.11)	17.82 (0.85)	0.728
Genero = Mujer (%)	161 (33.5)	29 (47.5)	0.043
Matemática (mean (sd))	75.49 (8.43)	74.84 (7.88)	0.565
Lenguaje (mean (sd))	67.84 (8.13)	69.77 (8.67)	0.083
Inglés (mean (sd))	70.65 (14.22)	67.93 (14.10)	0.160
Física (mean (sd))	73.95 (8.04)	73.03 (6.27)	0.392
Química (mean (sd))	73.90 (8.07)	72.87 (6.60)	0.336
Biología (mean (sd))	73.87 (8.12)	72.92 (6.48)	0.378
Riesgo_Economica = MEDIO ALTO (%)	181 (37.6)	17 (27.9)	0.177
Riesgo_Social (%)			0.802
ALTO	26 (5.4)	2 (3.3)	
BAJO	187 (38.9)	25 (41.0)	
MEDIO	190 (39.5)	26 (42.6)	
MEDIO ALTO	78 (16.2)	8 (13.1)	
Riesgo_Salud (%)			0.734
BAJO	381 (79.2)	47 (77.0)	
MEDIO	97 (20.2)	14 (23.0)	
MEDIO ALTO	3 (0.6)	0 (0.0)	
Riesgo.Academico (%)			0.134
ALTO	136 (28.3)	15 (24.6)	
BAJO	11 (2.3)	1 (1.6)	
MEDIO ALTO	192 (39.9)	34 (55.7)	
MEDIO BAJO	92 (19.1)	9 (14.8)	
MUY ALTO	50 (10.4)	2 (3.3)	
Riesgo_Cognitiva (%)			0.039
ALTO	119 (24.7)	22 (36.1)	
BAJO	58 (12.1)	2 (3.3)	
MEDIO	304 (63.2)	37 (60.7)	
Ben (%)			<0.001
(-1,0]	481 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	49 (80.3)	
(7,47]	0 (0.0)	12 (19.7)	

Los resultados del modelo de propensión depurado se aprecian en la Tabla 48, destacando que las co-variables más influyentes fueron el programa académico, el género y los riesgos, económico, académico y cognitivo.

Tabla 48.

*Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2015 (Modelo depurado.)*

```
Call:
glm(formula = Beneficio ~ Prog + Genero + Riesgo_Economica +
     Riesgo.Academico + Riesgo_Cognitiva, family = binomial, data = D
     atos.Evaluacion2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2429 -0.5293 -0.3317 -0.2010  2.6872

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      -0.9715     0.5148  -1.887  0.0591 .
ProgEle          -0.4444     0.5916  -0.751  0.4526
ProgElo         -1.0346     0.7080  -1.461  0.1439
ProgFis         -0.5341     0.6803  -0.785  0.4324
ProgGeo        -2.2345     1.0959  -2.039  0.0415 *
ProgInd        -1.5164     0.6528  -2.323  0.0202 *
ProgIQu        -1.4237     0.6175  -2.306  0.0211 *
ProgLMa       -16.7561    794.5143  -0.021  0.9832
ProgMat        -2.3621     1.0985  -2.150  0.0315 *
ProgMec        -1.9183     0.8155  -2.352  0.0187 *
ProgMet        -0.1614     0.5468  -0.295  0.7679
ProgPet        -0.5409     0.6722  -0.805  0.4210
ProgQui         0.2442     0.5749   0.425  0.6710
ProgSis        -1.5042     0.7076  -2.126  0.0335 *
GeneroMujer     0.5654     0.3191   1.772  0.0764 .
Riesgo_EconomicaMEDIO ALTO -0.5168     0.3265  -1.583  0.1135
Riesgo.AcademicoBAJO    0.4992     1.1707   0.426  0.6698
Riesgo.AcademicoMEDIO ALTO 0.5587     0.3577   1.562  0.1183
Riesgo.AcademicoMEDIO BAJO 0.1546     0.4771   0.324  0.7459
Riesgo.AcademicoMUY ALTO -1.1804     0.8077  -1.461  0.1439
Riesgo_CognitivaBAJO   -1.9336     0.7970  -2.426  0.0153 *
Riesgo_CognitivaMEDIO  -0.5417     0.3270  -1.657  0.0976 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 381.36 on 541 degrees of freedom
Residual deviance: 324.17 on 520 degrees of freedom
```

AIC: 368.17

Number of Fisher Scoring iterations: 16

En el pareamiento (ver Figura 75) se observa un estudiante beneficiado del SEA fuera del rango común y en la Tabla 49 se muestra el detalle del pareamiento tanto en el grupo de beneficiados como de no beneficiados.

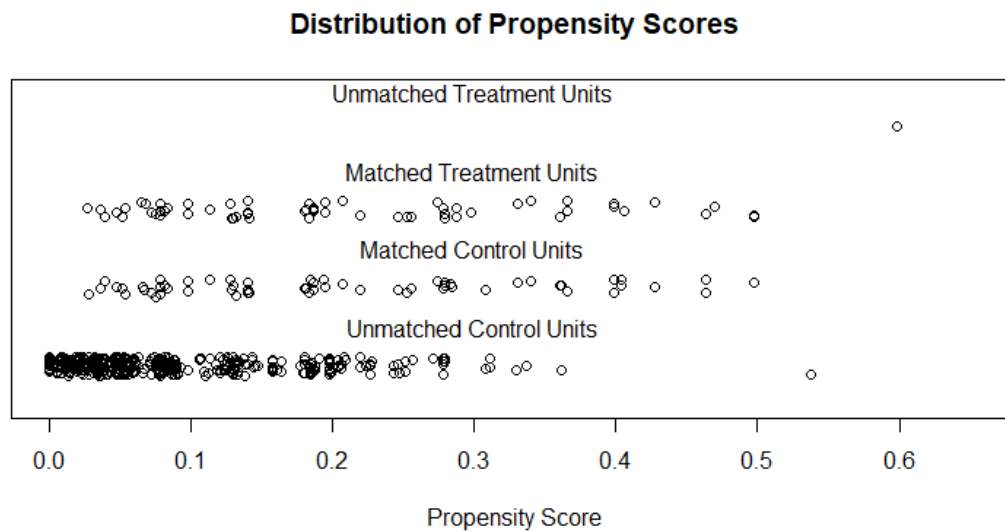


Figura 75. Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2015

Tabla 49.

*Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2015 de Ciencias e Ingeniería.*

	Control	Treated
All	481	61
Matched	60	60
Unmatched	314	0
Discarded	107	1

La efectividad del pareamiento seleccionado las parejas más parecidas se aprecia en la Figura 76.

En ella se observa que antes del pareamiento la distribución de puntajes de propensión entre los dos grupos era muy diferente (izquierda) mientras después del pareamiento las dos distribuciones

resultaron muy similares, lo que indica homogeneidad en los dos grupos comparados en sus puntajes de propensión y por ende en las co-variables que la explican.

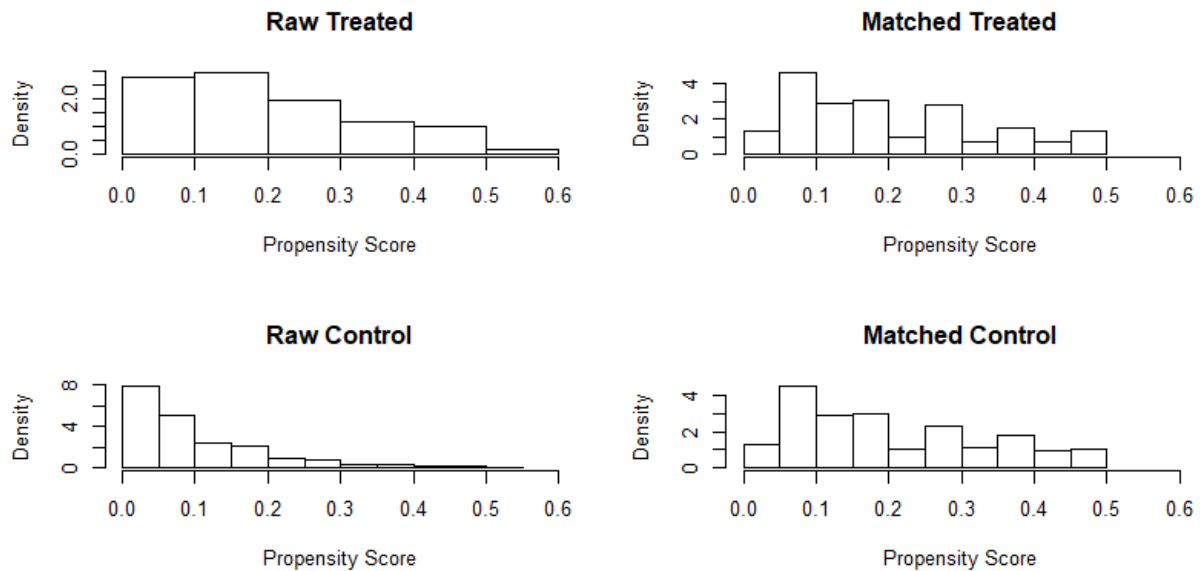


Figura 76. Histograma de puntajes de propensión para álgebra I antes y después del pareamiento.

1ª cohorte de 2015

La Tabla 50 muestra los resultados comparativos de los dos grupos después del pareamiento. Se observa un impacto negativo aunque no significativo del programa SEA en álgebra I para los estudiantes de esta cohorte.

Tabla 50.

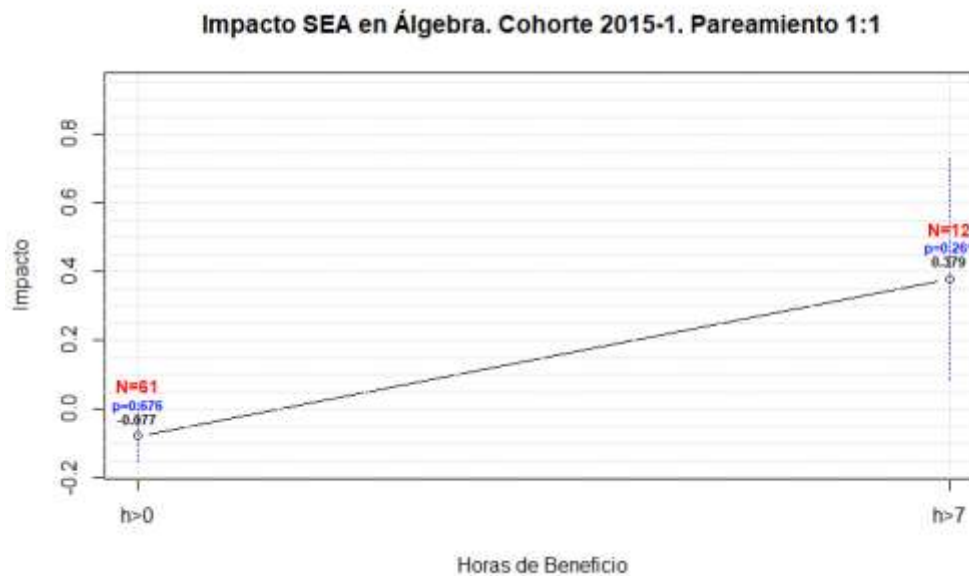
*Comparación de estudiantes con beneficio o no del SEA en álgebra I (después de pareamiento).*

1ª cohorte de 2015

	Sin Beneficio	Con Beneficio	p
--	---------------	---------------	---

n	60	60	
Nota (mean (sd))	3.23 (1.00)	3.17 (1.01)	0.751
Horas (mean (sd))	0.00 (0.00)	5.50 (7.43)	<0.001
Prog (%)			NaN
Civ	11 (18.3)	11 (18.3)	
Ele	5 (8.3)	6 (10.0)	
Elo	4 (6.7)	3 (5.0)	
Fis	2 (3.3)	4 (6.7)	
Geo	3 (5.0)	1 (1.7)	
Ind	4 (6.7)	4 (6.7)	
IQu	6 (10.0)	5 (8.3)	
LMa	0 (0.0)	0 (0.0)	
Mat	0 (0.0)	1 (1.7)	
Mec	2 (3.3)	2 (3.3)	
Met	7 (11.7)	9 (15.0)	
Pet	7 (11.7)	4 (6.7)	
Qui	6 (10.0)	7 (11.7)	
Sis	3 (5.0)	3 (5.0)	
Edad (mean (sd))	17.92 (1.21)	17.83 (0.85)	0.663
Genero = Mujer (%)	32 (53.3)	28 (46.7)	0.584
Matemática (mean (sd))	76.45 (7.50)	74.80 (7.94)	0.244
Lenguaje (mean (sd))	70.08 (7.21)	69.68 (8.71)	0.785
Inglés (mean (sd))	69.33 (13.34)	67.93 (14.22)	0.579
Física (mean (sd))	74.03 (6.29)	73.08 (6.31)	0.410
Química (mean (sd))	73.90 (6.69)	72.92 (6.65)	0.421
Biología (mean (sd))	73.90 (6.69)	72.97 (6.52)	0.440
Riesgo_Economica = MEDIO ALTO (%)	14 (23.3)	17 (28.3)	0.677
Riesgo_Social (%)			0.531
ALTO	2 (3.3)	2 (3.3)	
BAJO	33 (55.0)	25 (41.7)	
MEDIO	19 (31.7)	25 (41.7)	
MEDIO ALTO	6 (10.0)	8 (13.3)	
Riesgo_Salud (%)			0.496
BAJO	48 (80.0)	46 (76.7)	
MEDIO	11 (18.3)	14 (23.3)	
MEDIO ALTO	1 (1.7)	0 (0.0)	
Riesgo.Academico (%)			0.895
ALTO	14 (23.3)	15 (25.0)	
BAJO	2 (3.3)	1 (1.7)	
MEDIO ALTO	36 (60.0)	33 (55.0)	
MEDIO BAJO	6 (10.0)	9 (15.0)	
MUY ALTO	2 (3.3)	2 (3.3)	
Riesgo_Cognitiva (%)			0.482
ALTO	16 (26.7)	21 (35.0)	
BAJO	1 (1.7)	2 (3.3)	
MEDIO	43 (71.7)	37 (61.7)	
Ben (%)			<0.001
(-1,0]	60 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	48 (80.0)	
(7,47]	0 (0.0)	12 (20.0)	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
distance (mean (sd))	0.21 (0.13)	0.21 (0.13)	0.942

Los resultados promedio después de realizar 20 corridas aleatorias variando el umbral en el número de horas de beneficio a partir del cual se considera como realmente beneficiado del SEA se muestra en la Figura 77. Aunque se aprecia un impacto significativo al 26% cuando el estudiante toma más de 8 horas de beneficio, este se basa en un grupo de 12 estudiantes solamente por lo que no se puede considerar concluyente.



*Figura 77.* Impacto del programa SEA sobre álgebra I en función del número de horas de beneficio. 1ª cohorte de 2015

En síntesis, la técnica de pareamiento muestra un impacto significativo sobre la nota de álgebra I atribuida al uso del programa SEA a un nivel de significancia del 26% en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2015. Este impacto resulta de 0.4 décimas en la nota. Sin embargo, este se basa en un grupo de 12 estudiantes solamente. También se encontró por esta técnica que, con menos de 8 horas en el semestre, el programa SEA no tiene efecto significativo en la nota de álgebra (valor  $p = 0.676$ )

**3.2.2.4 Impacto del SEA sobre repitentes en álgebra lineal. Cohorte 2015-1.** En esta cohorte se encontraron 168 estudiantes que cursaron álgebra dos veces. La técnica de diferencias en diferencias supone que las tendencias de la nota son similares en los grupos de beneficiados y no beneficiados antes de la intervención y que los únicos factores que explican las diferencias en la nota entre ambos grupos, aparte del efecto del programa SEA, son constantes en el tiempo. Para asegurar la no violación del supuesto anterior, la técnica de diferencias en diferencias se debe condicionar a que el grupo de beneficiados haya recibido beneficios los dos semestres que cursó la materia y el grupo de no beneficiados no haya recibido beneficios ninguno de los dos semestre. Esto reduce el grupo a 145 ninguno de los cuales ninguno recibió beneficios del SEA, lo que imposibilita el uso de la técnica de diferencias en diferencias en este caso.

### **3.2.3 Impacto en cohorte 2016-1**

**3.2.3.1 Impacto del SEA sobre Cálculo I. Cohorte 2016-1.** De los 946 estudiantes de Ciencia e Ingeniería en esta cohorte, solo 649 (69%) contiene la información completa, 161 de los cuales reportan beneficio del SEA en cálculo I mientras que los restantes 488 no. La distribución de beneficios del SEA se aprecia en la Figura 78

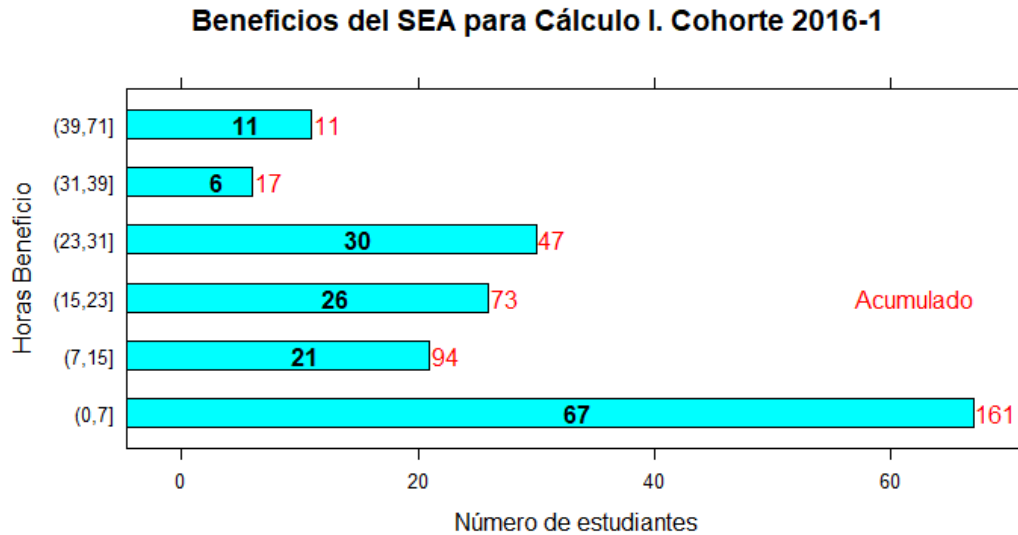


Figura 78. Beneficios del SEA en cálculo I. 1ª cohorte de 2016.

La Tabla 51 muestra los estadísticos más importantes de los dos grupos. Se observa una diferencia negativa entre el grupo de beneficiados y no beneficiados del SEA aunque ésta no resulte significativa al 5%. Sin embargo, también se aprecia una diferencia significativa en otras co-variables, lo que puede confundir el verdadero impacto del programa SEA sobre cálculo I para los estudiantes de esta cohorte.

Tabla 51.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2016*

	Sin Beneficio	Con Beneficio	p
n	488	161	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
Nota (mean (sd))	2.71 (0.96)	2.64 (0.85)	0.400
Horas (mean (sd))	0.00 (0.00)	15.10 (13.71)	<0.001
Prog (%)			<0.001
Civ	54 (11.1)	12 (7.5)	
DIn	25 (5.1)	7 (4.3)	
Ele	22 (4.5)	18 (11.2)	

Elo	34 (7.0)	10 (6.2)	
Fis	3 (0.6)	0 (0.0)	
Geo	38 (7.8)	5 (3.1)	
Ind	53 (10.9)	9 (5.6)	
IQu	46 (9.4)	12 (7.5)	
LMa	29 (5.9)	9 (5.6)	
Mat	14 (2.9)	13 (8.1)	
Mec	48 (9.8)	11 (6.8)	
Met	25 (5.1)	14 (8.7)	
Pet	26 (5.3)	10 (6.2)	
Qui	17 (3.5)	20 (12.4)	
Sis	54 (11.1)	11 (6.8)	
Edad (mean (sd))	18.05 (1.76)	17.80 (1.47)	0.096
Genero = Mujer (%)	145 (29.7)	75 (46.6)	<0.001
Matemática (mean (sd))	82.56 (10.31)	78.14 (9.62)	<0.001
Ingles (mean (sd))	73.19 (12.90)	67.64 (12.95)	<0.001
Naturales (mean (sd))	75.43 (9.03)	72.41 (7.81)	<0.001
Lectura (mean (sd))	67.13 (7.22)	65.59 (7.20)	0.019
Riesgo_Economica = BAJO (%)	432 (88.5)	145 (90.1)	0.694
Riesgo_Social (%)			0.725
ALTO	29 (5.9)	7 (4.3)	
BAJO	399 (81.8)	135 (83.9)	
MEDIO	60 (12.3)	19 (11.8)	
Riesgo_Salud = MEDIO (%)	483 (99.0)	161 (100.0)	0.442
Riesgo.Academico (%)			0.628
ALTO	373 (76.4)	117 (72.7)	
BAJO	45 (9.2)	17 (10.6)	
MEDIO	70 (14.3)	27 (16.8)	
Riesgo_Cognitiva (%)			0.007
ALTO	120 (24.6)	48 (29.8)	
BAJO	77 (15.8)	10 (6.2)	
MEDIO	291 (59.6)	103 (64.0)	
Ben (%)			<0.001
(-1,0]	488 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	67 (41.6)	
(7,15]	0 (0.0)	21 (13.0)	
(15,23]	0 (0.0)	26 (16.1)	
(23,31]	0 (0.0)	30 (18.6)	
(31,39]	0 (0.0)	6 (3.7)	
(39,71]	0 (0.0)	11 (6.8)	

El modelo de propensión se resume en la Tabla 52. En este modelo se encuentran que las variables que más afectan la propensión del estudiante a inscribirse en el programa SEA para recibir beneficios en cálculo I fueron: el programa académico, la edad, el género, los puntajes en matemáticas e inglés de las pruebas SABER 11 y el riesgo económico.

Tabla 52.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2016 (Modelo depurado.)*

```
Beneficio ~ Prog + Edad + Genero + Matemática + Ingles + Naturales
+ Lectura + Riesgo_Economica + Riesgo_Social + Riesgo_Salud + Riesgo
.Academico + Riesgo_Cognitiva"
```

```
Call:
glm(formula = Beneficio ~ . - Nota - Horas - Ben, family = binomial,
     data = Datos.Evaluacion2)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.63848	-0.73890	-0.53080	-0.00032	2.41936

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-7.961949	595.348377	-0.013	0.9893
ProgDIn	-0.646378	0.591102	-1.094	0.2742
ProgEle	0.909966	0.509306	1.787	0.0740
ProgElo	0.225136	0.526626	0.428	0.6690
ProgFis	-14.001265	796.779022	-0.018	0.9860
ProgGeo	-0.351946	0.593100	-0.593	0.5529
ProgInd	-0.673657	0.510480	-1.320	0.1870
ProgIQu	0.015255	0.478886	0.032	0.9746
ProgLMa	-0.786238	0.643846	-1.221	0.2220
ProgMat	0.707191	0.642177	1.101	0.2708
ProgMec	0.394308	0.491127	0.803	0.4221
ProgMet	0.390055	0.543014	0.718	0.4726
ProgPet	0.486887	0.522660	0.932	0.3516
ProgQui	1.065129	0.524729	2.030	0.0424 *
Progsis	-0.459088	0.529445	-0.867	0.3859
Edad	-0.162982	0.083281	-1.957	0.0503 .
GeneroMujer	0.982882	0.236251	4.160	3.18e-05 ***
Matemática	-0.029525	0.014308	-2.064	0.0391 *
Ingles	-0.027794	0.009458	-2.939	0.0033 **
Naturales	-0.009219	0.016124	-0.572	0.5675
Lectura	-0.007202	0.016513	-0.436	0.6628
Riesgo_EconomicaBAJO	0.631990	0.343681	1.839	0.0659 .
Riesgo_SocialBAJO	0.355984	0.492261	0.723	0.4696
Riesgo_SocialMEDIO	0.265096	0.557531	0.475	0.6344
Riesgo_SaludMEDIO	13.984302	595.343054	0.023	0.9813
Riesgo_AcademicoBAJO	-0.407758	0.375555	-1.086	0.2776
Riesgo_AcademicoMEDIO	-0.162105	0.280829	-0.577	0.5638
Riesgo_CognitivaBAJO	-0.546768	0.418517	-1.306	0.1914
Riesgo_CognitivaMEDIO	0.071238	0.236552	0.301	0.7633

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 727.15 on 648 degrees of freedom
Residual deviance: 626.38 on 620 degrees of freedom
AIC: 684.38
```

Number of Fisher Scoring iterations: 14

La Figura 79 y la Tabla 53 resumen los resultados del pareamiento encontrándose 3 estudiantes con beneficios y 36 no beneficiados que fueron descartados del pareamiento debido a que se encontraban fuera del rango común.

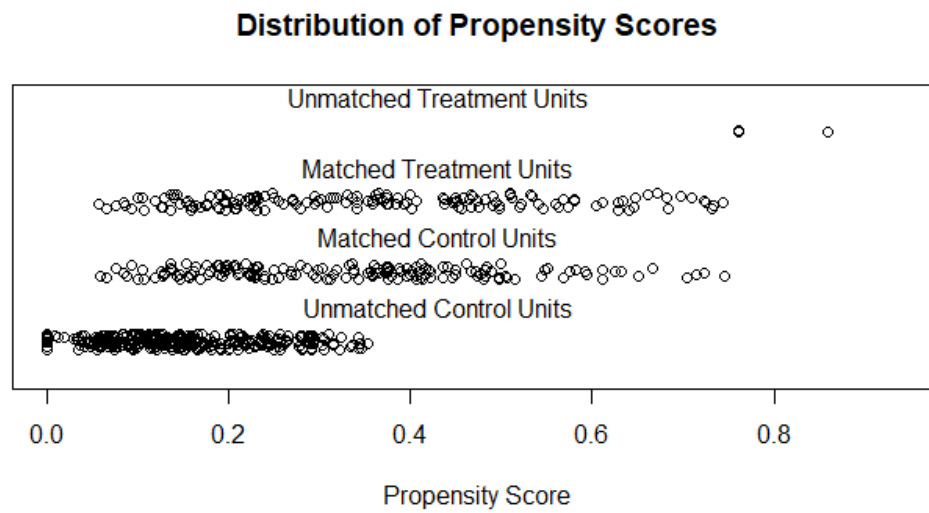


Figura 79. Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2016

Tabla 53.

*Resultados finales del pareamiento para beneficiados en cálculo I. 1ª cohorte de 2016 de Ciencias e Ingeniería.*

	Control	Treated
All	488	161
Matched	158	158
Unmatched	294	0
Discarded	36	3

La Figura 80 muestra la distribución de los puntajes de propensión entre el grupo de beneficiados y no beneficiados del SEA antes y después del pareamiento.

Se observa que el pareamiento lleva a grupos más homogéneos con los que se obtiene un mejor estimador del contrafactual y por ende una mejor evaluación del programa SEA sobre cálculo I en los estudiantes de esta cohorte.

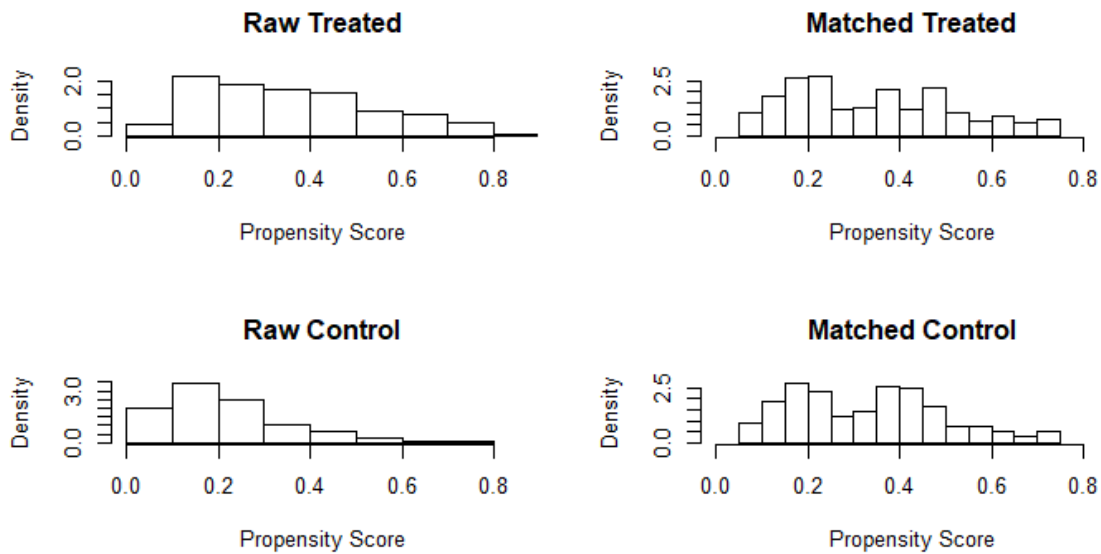
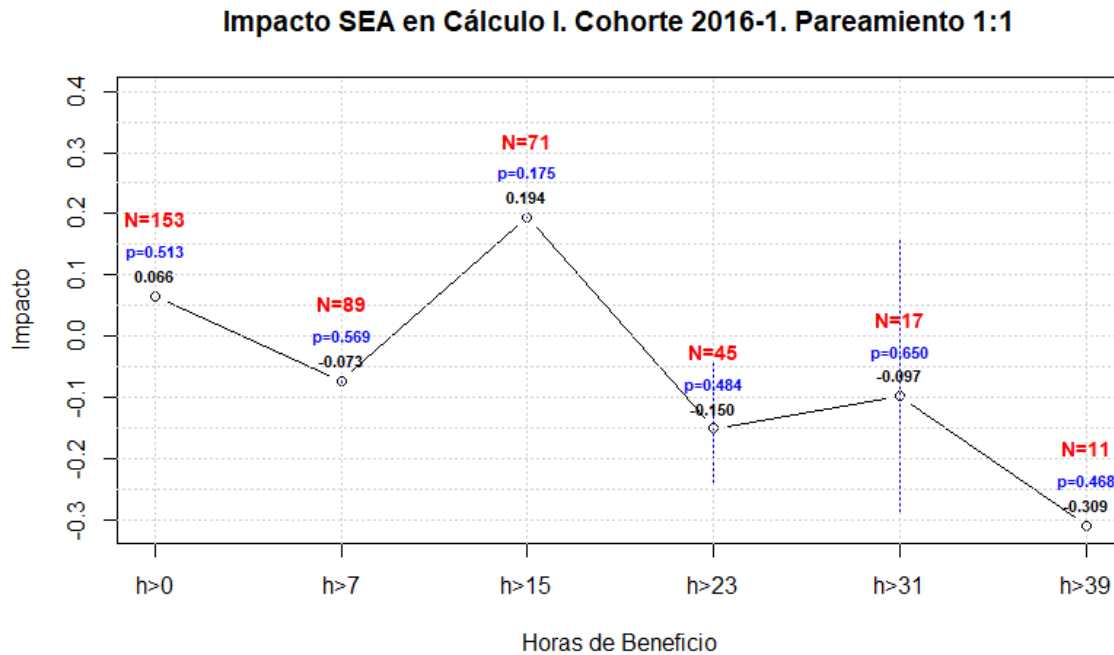


Figura 80. Histograma de puntajes de propensión para cálculo I antes y después del pareamiento.

1ª cohorte de 2016

Finalmente, la Figura 81 muestra la forma como cambia la evaluación en función del umbral a partir del cual se considera beneficiado del programa SEA.

En ningún caso se encontró un impacto significativo al 5% aunque después de 16 horas de beneficio se observa un pico de casi 0.2 décimas de beneficio debidas al programa SEA y significativas al 18% aunque decae por encima de 24 horas de beneficio.



*Figura 81.* Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio. 1ª cohorte de 2016

En síntesis, se encontró que la técnica de pareamiento no muestra un impacto significativo del programa SEA al 5% en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2016. Se observa una tendencia decreciente del impacto al aumentar el número de horas con un pico máximo de 0.2 décimas con más de 17 horas de beneficio y significativa al 18%

**3.2.3.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2016-1.** Solo se encontraron 6 estudiantes beneficiados del SEA con los que se puede emplear la técnica de diferencias en diferencias. La distribución del número de estudiantes en función de la cantidad de horas de beneficio recibidas se muestra en la Figura 82.

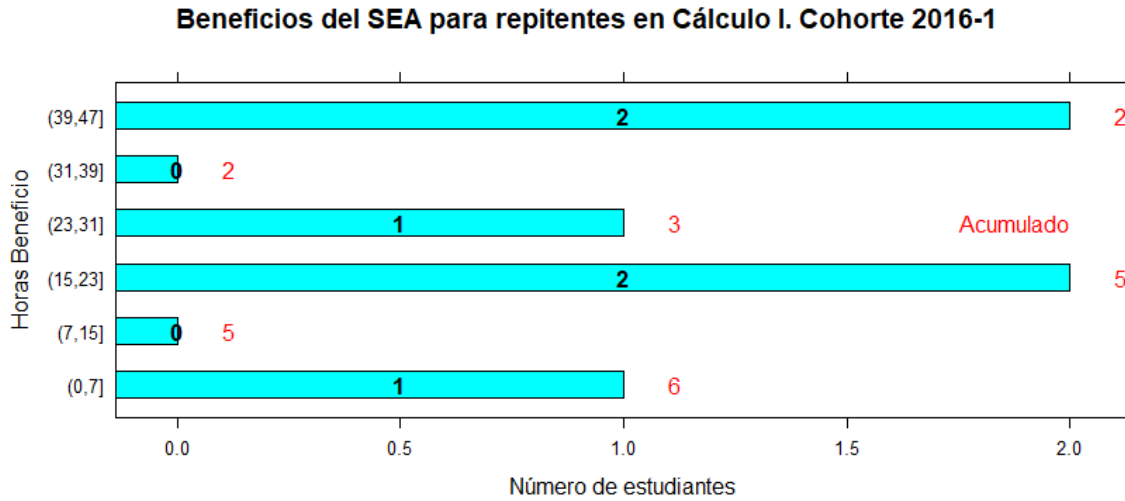


Figura 82. Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2016.

La Figura 83 muestra que el cambio promedio en nota del grupo de beneficiarios (0.83 décimas) es superior al cambio en el grupo de no beneficiarios (0.58 décimas) lo que sugiere un impacto positivo del programa SEA de 0.25 décimas la segunda vez que ve la materia, es decir cuando repite la materia. Sin embargo, este no resulta significativo desde el punto de vista estadístico como se logra apreciar en la Tabla 54.

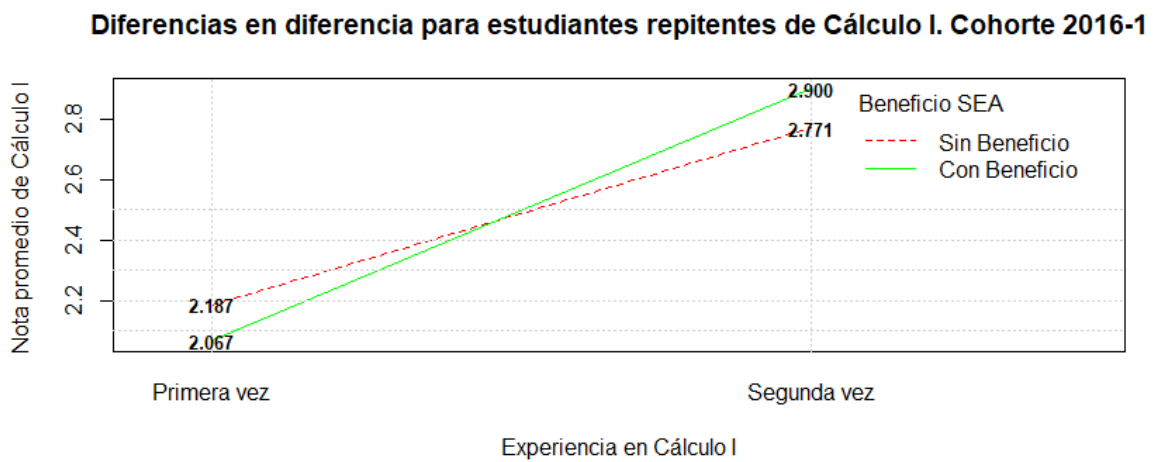


Figura 83. Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2016.

Tabla 54.

*Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2016*

Estimate	Std. Error	t value	Pr(> t )
0.2496296	0.4246814	0.5878044	0.5569057

**3.2.3.3 Impacto del SEA sobre Álgebra Lineal. Cohorte 2016-1.** En este caso, se encontraron 647 estudiantes con información completa que cursaron la materia: 120 con beneficio del SEA en álgebra y 527 sin él. La distribución de estudiantes por horas de beneficio recibidas se aprecia en a la Figura 84.

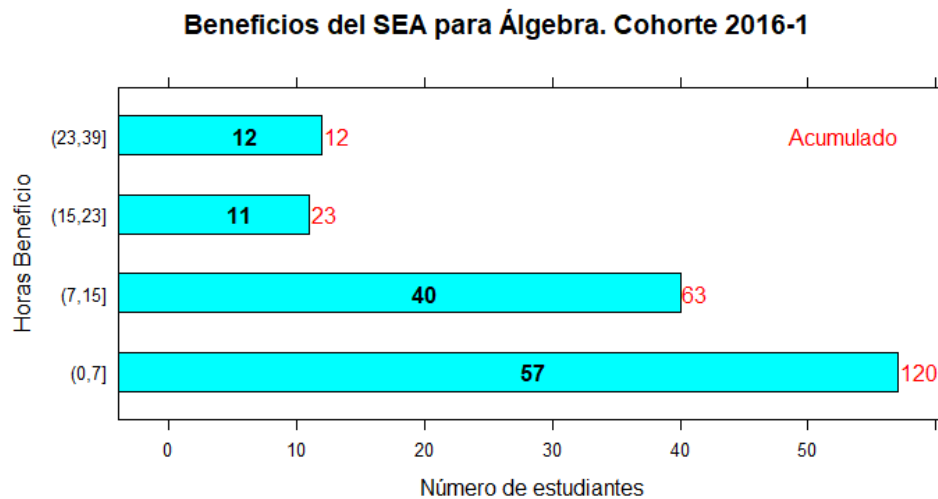


Figura 84. Beneficios del SEA en álgebra I. 1ª cohorte de 2016.

La Tabla 55 muestra la comparación entre el grupo de beneficiados y no beneficiados del programa SEA en álgebra I. Se observa diferencias significativas en el programa académico, el género, los el puntaje en matemáticas y naturales de las pruebas SABER 11 y en el riesgo cognitivo.

Tabla 55.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en álgebra I. 1ª cohorte de 2016*

	No Beneficiados	Beneficiados	p
n	527	120	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
Nota (mean (sd))	3.22 (0.97)	3.23 (0.83)	0.980
Horas (mean (sd))	0.00 (0.00)	9.97 (8.24)	<0.001
Prog (%)			<0.001
Civ	61 (11.6)	5 (4.2)	
Ele	31 (5.9)	10 (8.3)	
Elo	36 (6.8)	8 (6.7)	
Fis	32 (6.1)	2 (1.7)	
Geo	34 (6.5)	8 (6.7)	
Ind	36 (6.8)	25 (20.8)	
IQu	46 (8.7)	12 (10.0)	
LMa	31 (5.9)	8 (6.7)	
Mat	13 (2.5)	13 (10.8)	
Mec	50 (9.5)	8 (6.7)	
Met	36 (6.8)	3 (2.5)	
Pet	34 (6.5)	2 (1.7)	
Qui	27 (5.1)	11 (9.2)	
Sis	60 (11.4)	5 (4.2)	
Edad (mean (sd))	18.10 (2.25)	17.80 (0.99)	0.158
Genero = Mujer (%)	148 (28.1)	57 (47.5)	<0.001
Matemática (mean (sd))	82.46 (10.15)	79.12 (10.95)	0.001
Inglés (mean (sd))	72.06 (13.19)	70.76 (13.53)	0.334
Naturales (mean (sd))	75.26 (8.76)	73.59 (9.13)	0.063
Lectura (mean (sd))	66.62 (7.30)	67.37 (8.36)	0.328
Riesgo_Economica = BAJO (%)	472 (89.6)	103 (85.8)	0.312
Riesgo_Social (%)			0.956
ALTO	29 (5.5)	6 (5.0)	
BAJO	436 (82.7)	99 (82.5)	
MEDIO	62 (11.8)	15 (12.5)	
Riesgo_Salud = MEDIO (%)	522 (99.1)	120 (100.0)	0.622
Riesgo_Academico (%)			0.539
ALTO	404 (76.7)	87 (72.5)	
BAJO	51 (9.7)	12 (10.0)	
MEDIO	72 (13.7)	21 (17.5)	
Riesgo_Cognitiva (%)			0.043
ALTO	136 (25.8)	28 (23.3)	
BAJO	81 (15.4)	9 (7.5)	
MEDIO	310 (58.8)	83 (69.2)	
Ben (%)			<0.001
(-1,0]	527 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	57 (47.5)	
(7,15]	0 (0.0)	40 (33.3)	
(15,23]	0 (0.0)	11 (9.2)	
(23,39]	0 (0.0)	12 (10.0)	

El modelo de regresión logística con las variables que mejor explican la propensión de un estudiante a inscribirse en el programa SEA para recibir beneficios en álgebra se aprecia en la Tabla 56.

Tabla 56.

*Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2016 (Modelo depurado.)*

```
Call:
glm(formula = Beneficio ~ Prog + Genero + Matemática + Lectura +
     Riesgo_Cognitiva, family = binomial, data = Datos.Evaluacion2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.3785 -0.6479 -0.4494 -0.2795  2.9732

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      -2.27790    1.67032  -1.364  0.172647
ProgEle           1.23161    0.62506   1.970  0.048794 *
ProgElo           1.06997    0.63132   1.695  0.090114 .
ProgFis          -0.23129    0.88242  -0.262  0.793234
ProgGeo           1.32813    0.62143   2.137  0.032580 *
ProgInd           1.90253    0.54581   3.486  0.000491 ***
ProgIQu           1.08226    0.57698   1.876  0.060692 .
ProgLMa           0.91747    0.69311   1.324  0.185601
ProgMat           2.47394    0.68740   3.599  0.000319 ***
ProgMec           0.90977    0.61651   1.476  0.140032
ProgMet          -0.14925    0.79517  -0.188  0.851119
ProgPet          -0.32383    0.87521  -0.370  0.711379
ProgQui           1.28764    0.62757   2.052  0.040191 *
ProgSis          -0.16684    0.68448  -0.244  0.807428
GeneroMujer       0.56299    0.24188   2.328  0.019934 *
Matemática       -0.03429    0.01403  -2.445  0.014501 *
Lectura           0.03146    0.01723   1.826  0.067782 .
Riesgo_CognitivaBAJO -0.24628    0.44500  -0.553  0.579968
Riesgo_CognitivaMEDIO 0.50026    0.26974   1.855  0.063650 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 620.59  on 646  degrees of freedom
Residual deviance: 536.31  on 628  degrees of freedom
AIC: 574.31

Number of Fisher Scoring iterations: 5
```

La Tabla 57 y la Figura 85 muestran los resultados del pareamiento resaltando que fue necesario descartar 10 estudiantes beneficiados por encontrarse fuera del rango común de propensión.

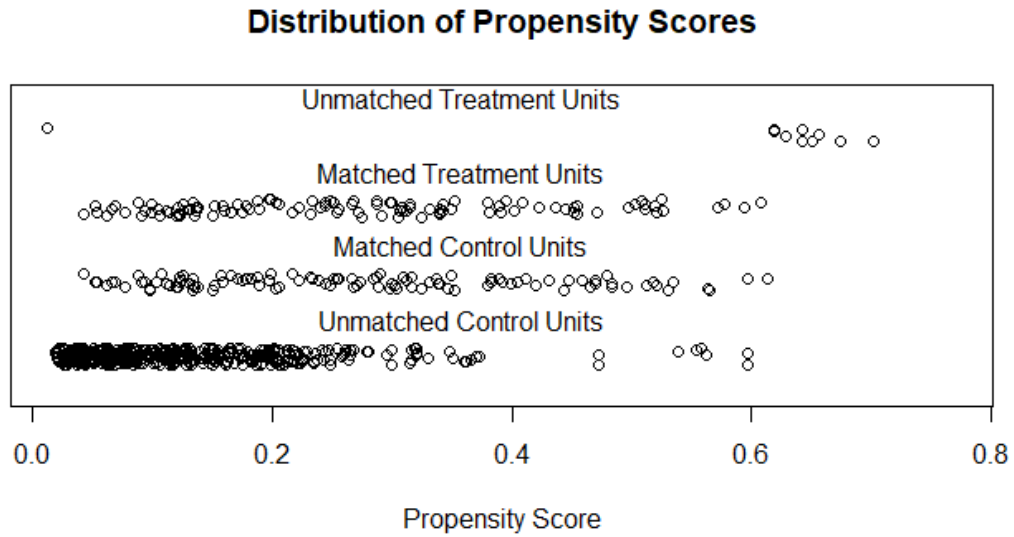


Figura 85. Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2016

Tabla 57.

Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2016 de Ciencias e Ingeniería.

	Control	Treated
All	527	120
Matched	110	110
Unmatched	417	0
Discarded	0	10

Figura 86 muestra la homogeneidad de los grupos comparados antes y después del pareamiento. Se destaca la similitud de los dos grupos después del pareamiento (derecha) comparada con la distribución de puntajes de propensión antes del pareamiento (izquierda)

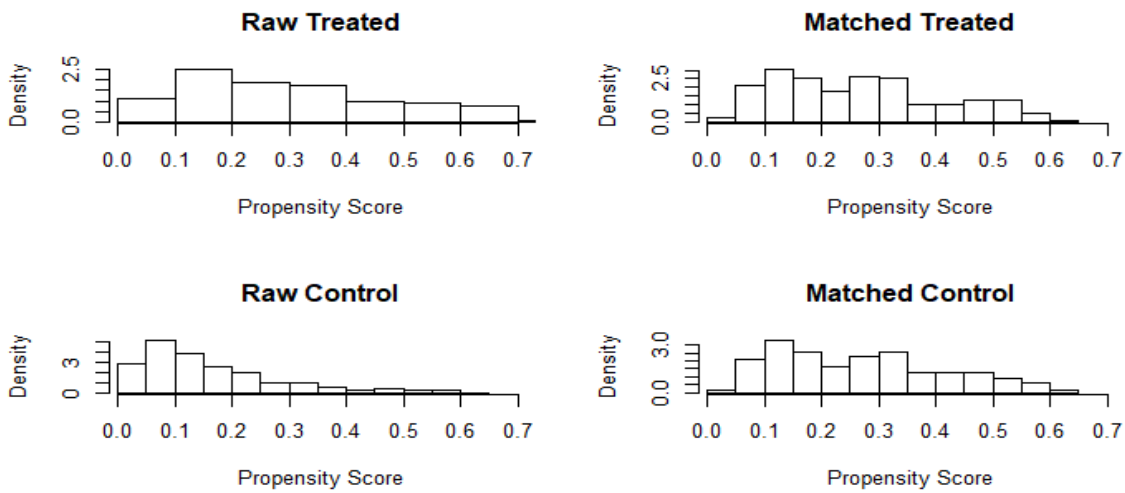


Figura 86. Histograma de puntajes de propensión para álgebra I antes y después del pareamiento. 1ª cohorte de 2016

Finalmente, la Figura 87 muestra que el impacto del programa SEA sobre álgebra es de 0.5 décimas pero solo resulta significativo si se reciben entre 8 y 15 horas. Menos de 7 horas no tiene impacto significativo y más de 16 tampoco lo tiene.

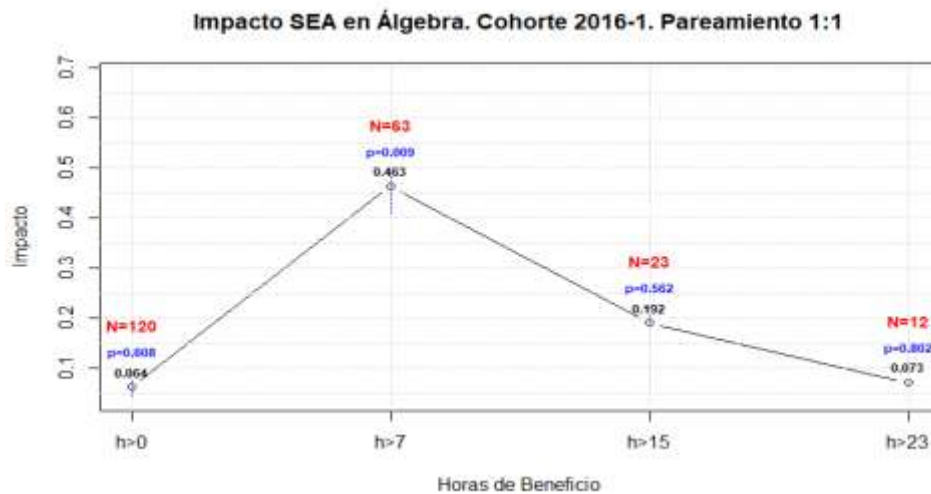


Figura 87. Impacto del programa SEA sobre álgebra I en función del número de horas de beneficio. 1ª cohorte de 2016

**3.2.3.4 Impacto del SEA sobre repitentes en álgebra lineal. Cohorte 2016-I.** En esta cohorte solo se encontraron 10 estudiantes beneficiados del SEA como se muestra en la Figura 88.

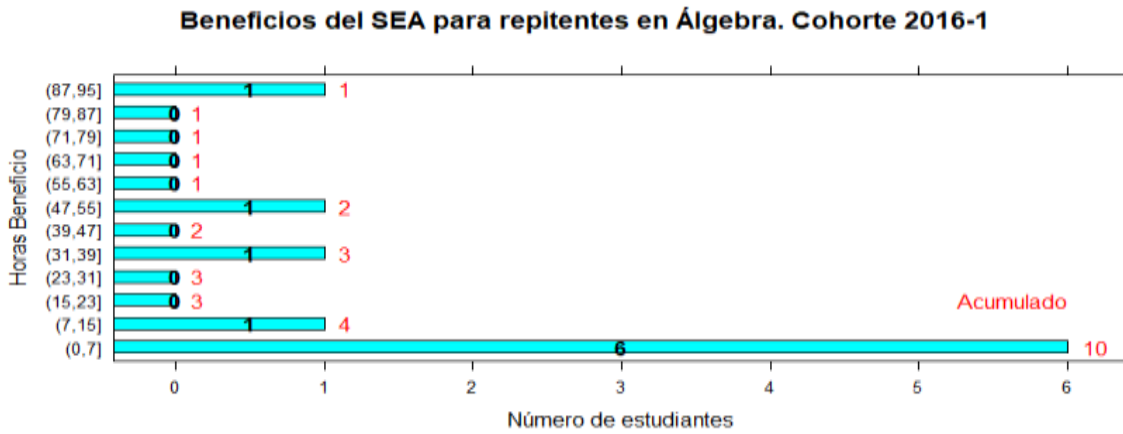


Figura 88. Beneficios del SEA para repitentes en álgebra I. 1ª cohorte de 2016.

La Figura 89 muestra el cambio promedio de la nota en el grupo de beneficiados y no beneficiados del SEA en álgebra. Se observa que el cambio en la nota promedio de álgebra es levemente mayor para beneficiados que para no beneficiados.

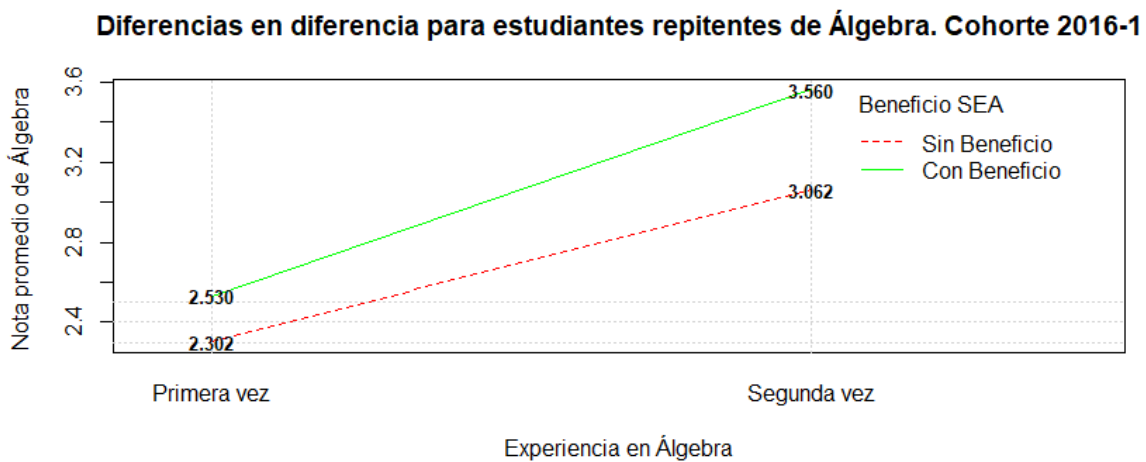


Figura 89. Diferencias en diferencias a repitentes en álgebra I. 1ª cohorte de 2016

El estimado del impacto del programa SEA se muestra en la Tabla 58. Se encontró un impacto positivo pero solo significativo al 36 %.

Tabla 58.

*Estimador del impacto del programa SEA sobre repitentes de álgebra I usando diferencias en diferencias. 1ª cohorte de 2016*

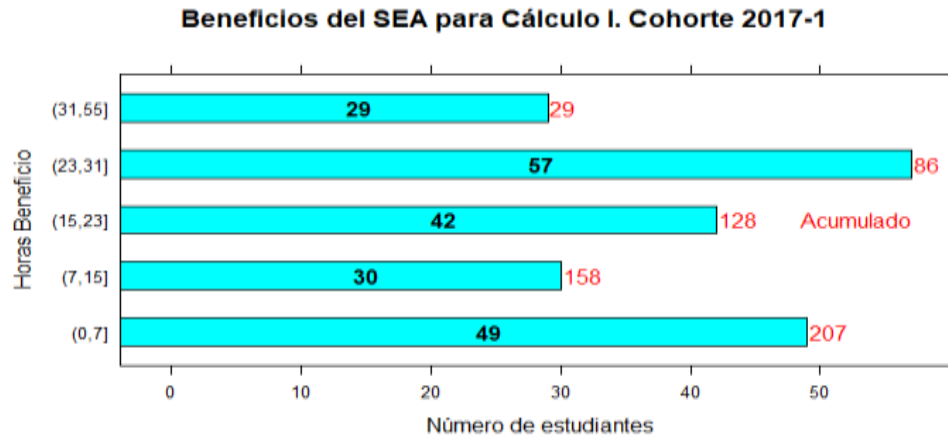
Estimate	Std. Error	t value	Pr(> t )
0.2702597	0.2941903	0.9186563	0.3589588

### 3.2.4 Impacto en cohorte 2017-1

**3.2.4.1 Impacto del SEA sobre Cálculo I. Cohorte 2017-1.** Para los estudiantes de esta cohorte se encontraron 669 (72%) estudiantes con información completa, 207 con beneficio y 462 sin beneficio del SEA en cálculo I.

La distribución por horas de beneficio se aprecia en la Figura 90. Se observa que la mayor cantidad de estudiantes (57) se encuentra entre 24 y 31 horas que corresponde a promedio de 3 y 4 horas de beneficio a la semana.

También se aprecia una mayor cobertura global de programa SEA, comparado con los estudiantes de las cohortes anteriores consideradas en el presente estudio, para la misma asignatura.



*Figura 90.* Beneficios del SEA en cálculo I. 1ª cohorte de 2017.

La Tabla 59 muestra la comparación entre el grupo de beneficiados y no beneficiados del SEA en diferentes características que, así como el SEA, también pueden afectar el desempeño del estudiantes en cálculo I. Específicamente, se encuentran diferencias significativas en el programa académico (valor  $p = 0.003$ ), la edad (valor  $p = 0.004$ ), los puntajes en matemáticas (valor  $p = 0.008$ ), Ciencias Naturales (valor  $p = 0.01$ ), lectura crítica (valor  $p = 0.021$ ), inglés (valor  $p = 0.077$ ) y el puntaje global ( $<0.001$ ) de las pruebas SABER 11.

No se encontraron diferencias significativas en las dimensiones de riesgo consideradas por el programa SEA, excepto la dimensión económica, la cual resulta significativa, encontrándose que la proporción estudiantes en riesgo bajo es estadísticamente superior en el grupo de no beneficiados (88.4% con valor  $p$  de 0.007)

Tabla 59.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en cálculo I. 1ª cohorte de 2017*

	No Beneficiado	Beneficiado	p
n	462	207	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
Nota (mean (sd))	2.85 (0.99)	2.91 (0.80)	0.474
Horas (mean (sd))	0.00 (0.00)	18.85 (11.47)	<0.001
Prog (%)			0.003
Civ	50 (10.8)	19 (9.2)	
DIn	23 (5.0)	5 (2.4)	
Ele	32 (6.9)	15 (7.2)	
Elo	29 (6.3)	27 (13.0)	
Fis	4 (0.9)	0 (0.0)	
Geo	26 (5.6)	16 (7.7)	
Ind	49 (10.6)	21 (10.1)	
IQu	33 (7.1)	18 (8.7)	
LMa	12 (2.6)	15 (7.2)	
Mat	11 (2.4)	6 (2.9)	
Mec	54 (11.7)	17 (8.2)	
Met	29 (6.3)	10 (4.8)	
Pet	28 (6.1)	16 (7.7)	
Qui	24 (5.2)	12 (5.8)	
Sis	58 (12.6)	10 (4.8)	
Edad (mean (sd))	17.99 (2.01)	17.57 (0.87)	0.004
Genero = Hombre (%)	325 (70.3)	128 (61.8)	0.037
Matemática (mean (sd))	76.58 (7.11)	75.06 (6.13)	0.008
Ciencias Naturales (mean (sd))	72.81 (5.52)	71.65 (5.11)	0.010
Lectura Critica (mean (sd))	69.74 (6.02)	68.57 (6.33)	0.021
Inglés (mean (sd))	72.81 (10.23)	71.28 (10.67)	0.077
Global (mean (sd))	361.43 (21.79)	354.69 (23.39)	<0.001
Riesgo_Economica = BAJO (%)	437 (94.6)	183 (88.4)	0.007
Riesgo_Social (%)			0.529
ALTO	24 (5.2)	8 (3.9)	
BAJO	358 (77.5)	157 (75.8)	
MEDIO	80 (17.3)	42 (20.3)	
Riesgo_Salud = MEDIO (%)	459 (99.4)	205 (99.0)	1.000
Riesgo_Academico (%)			0.308
ALTO	307 (66.5)	125 (60.4)	
BAJO	32 (6.9)	18 (8.7)	
MEDIO	123 (26.6)	64 (30.9)	
Riesgo_Cognitiva (%)			0.218
ALTO	104 (22.5)	52 (25.1)	
BAJO	72 (15.6)	22 (10.6)	
MEDIO	286 (61.9)	133 (64.3)	
Ben (%)			<0.001
(-1,0]	462 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	49 (23.7)	
(7,15]	0 (0.0)	30 (14.5)	
(15,23]	0 (0.0)	42 (20.3)	
(23,31]	0 (0.0)	57 (27.5)	
(31,55]	0 (0.0)	29 (14.0)	

La Tabla 60 resume los estadísticos más importantes del modelo de propensión depurado empleando el algoritmo de regresión por pasos. Se destaca que todas las variables incluidas en el modelo contribuyen significativamente a explicar la propensión de un estudiante de esta cohorte a tomar beneficios del SEA en cálculo I (valor  $p < 0.05$ ). Sin embargo, las más influyentes fueron el programa académico (valor  $p = 3.09e-05$ ), la edad (valor  $p = 0.00240$ ) y el puntaje global del estudiante en la pruebas SABER 11 (valor  $p = 0.00332$ ). La única dimensión de riesgo significativa fue el riesgo económico (valor  $p = 0.0181$ )

Tabla 60.

*Resultados del modelo de propensión para beneficiados del programa SEA en cálculo I. 1ª cohorte de 2017 (Modelo depurado.)*

```
Call:
glm(formula = Beneficio ~ Prog + Edad + Genero + Global + Riesgo_Eco
nomico,
     family = binomial, data = Datos.Evaluacion2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6624  -0.8649  -0.6676   1.1645   2.3203

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      10.131133   2.431592   4.166 3.09e-05 ***
ProgDIn           -0.882147   0.577937  -1.526  0.12692
ProgEle           0.157486   0.432450   0.364  0.71573
ProgElo           0.814552   0.400221   2.035  0.04183 *
ProgFis          -15.236128  670.904190 -0.023  0.98188
ProgGeo           0.333874   0.428114   0.780  0.43547
ProgInd           0.039598   0.384552   0.103  0.91799
ProgIQu           0.128258   0.410583   0.312  0.75475
ProgLma           0.132954   0.564515   0.236  0.81381
ProgMat          -0.387649   0.653274  -0.593  0.55292
ProgMec          -0.296926   0.398440  -0.745  0.45614
ProgMet          -0.737288   0.508379  -1.450  0.14698
ProgPet          -0.068830   0.446773  -0.154  0.87756
ProgQui          -0.281224   0.484267  -0.581  0.56143
ProgSis          -1.073665   0.456334  -2.353  0.01863 *
Edad              -0.244493   0.080527  -3.036  0.00240 **
GeneroHombre     -0.424112   0.198636  -2.135  0.03275 *
Global            -0.015283   0.005204  -2.937  0.00332 **
Riesgo_EconomicoBAJO -0.802344   0.339461  -2.364  0.01810 *
```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

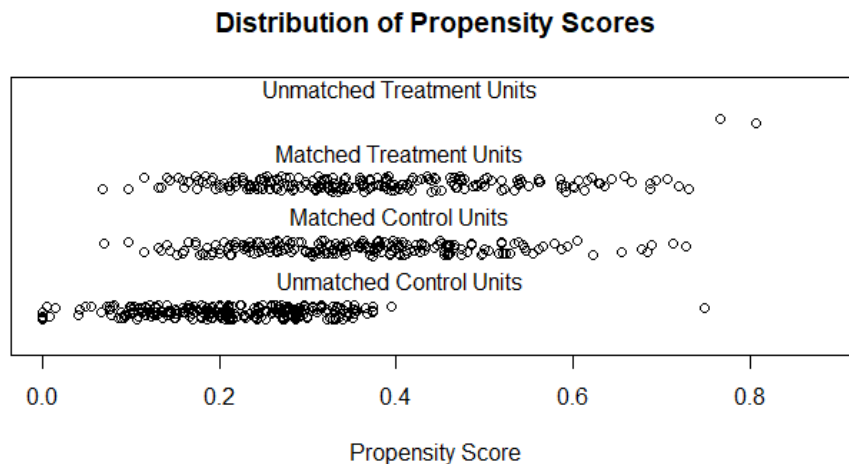
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 827.73  on 668  degrees of freedom
Residual deviance: 760.24  on 650  degrees of freedom
AIC: 798.24

Number of Fisher Scoring iterations: 14

```

La Figura 91 ilustra gráficamente los resultados del pareamiento, el diagrama de puntos de la parte superior muestra los estudiantes beneficiados que no fueron tenidos en cuenta por estar fuera el rango del común. Los dos diagramas del centro muestran los estudiantes pareados y en la parte inferior se parecían los estudiantes sin beneficios que no fueron tenidos en cuenta el proceso de pareamiento.



*Figura 91.* Pareamiento por puntajes de propensión para cálculo I. 1ª cohorte de 2017

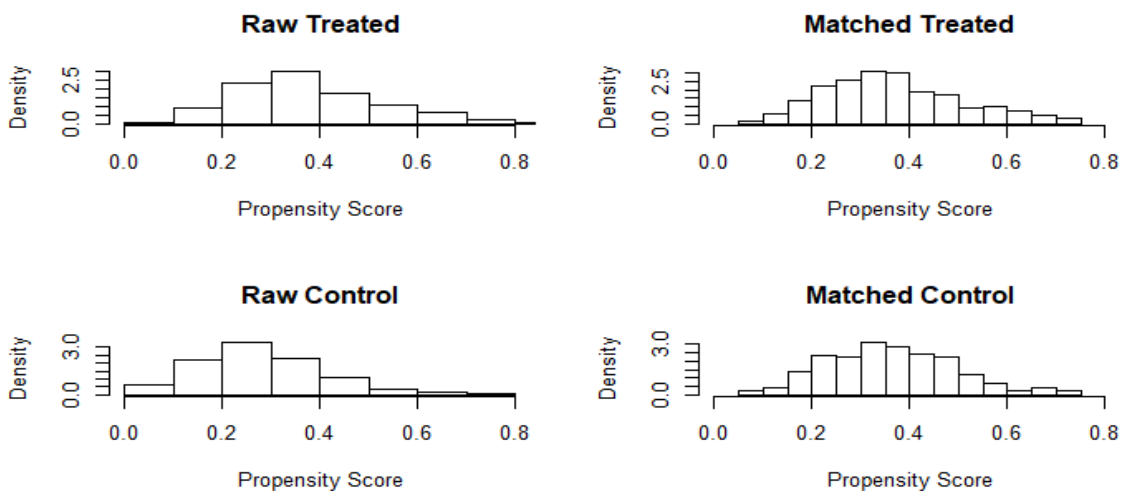
La Tabla 61 muestra los resultados finales del aplicar la técnica de pareamiento, se observan los dos estudiantes beneficiados descartados del pareamiento así como 12 sin beneficio que se descartaron del proceso de pareamiento por estar fuera del rango común de propensión.

Tabla 61.

*Resultados finales del pareamiento para beneficiados en cálculo I. 1ª cohorte de 2017 de Ciencias e Ingeniería.*

	Control	Treated
All	462	207
Matched	205	205
Unmatched	245	0
Discarded	12	2

Como etapa final del diagnóstico del pareamiento, en la Figura 92 se muestra el histograma de puntajes de propensión entre el grupo de beneficiados y no beneficiados antes y después del pareamiento. Antes del pareamiento (histogramas a la izquierda) se observa un sesgo a la izquierda mayor en el grupo de control (no beneficiado) con respecto al grupo de tratamiento (beneficiado) aspecto que se corrige después del pareamiento (histogramas a la derecha), lo que indica una homogeneidad en los dos grupos comparados en términos de puntajes de propensión y por ende, en las co-variables subyacentes en el modelo de propensión.



*Figura 92. Histograma de puntajes de propensión para cálculo I antes y después del pareamiento.*

1ª cohorte de 2017

La Tabla 62 muestra la comparación entre el grupo de beneficiados y no beneficiados del SEA en cálculo I par esta cohorte, después del pareamiento. Se encontró un impacto positivo del programa SEA sobre el desempeño de los estudiantes de esta cohorte en cálculo I de 0.13 décimas en la nota promedio el cual resultó significativo al 12%. No se encontró cambio significativo en las demás co-variables que también pueden afectar el desempeño del programa SEA en cálculo I.

Tabla 62.

*Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2017*

	No Beneficiado	Beneficiado	p
n	205	205	
Nota (mean (sd))	2.78 (1.00)	2.91 (0.80)	0.121
Horas (mean (sd))	0.00 (0.00)	18.94 (11.47)	<0.001
Prog (%)			NaN
Civ	23 (11.2)	19 (9.3)	
DIn	3 (1.5)	5 (2.4)	
Ele	13 (6.3)	15 (7.3)	
Elo	22 (10.7)	27 (13.2)	
Fis	0 (0.0)	0 (0.0)	
Geo	17 (8.3)	16 (7.8)	
Ind	26 (12.7)	21 (10.2)	
IQu	18 (8.8)	18 (8.8)	
LMa	10 (4.9)	13 (6.3)	
Mat	8 (3.9)	6 (2.9)	
Mec	17 (8.3)	17 (8.3)	
Met	9 (4.4)	10 (4.9)	
Pet	16 (7.8)	16 (7.8)	
Qui	15 (7.3)	12 (5.9)	
Sis	8 (3.9)	10 (4.9)	
Edad (mean (sd))	17.56 (0.97)	17.59 (0.86)	0.788
Genero = Hombre (%)	127 (62.0)	127 (62.0)	1.000
Matemática (mean (sd))	75.56 (6.54)	75.12 (6.12)	0.488
Ciencias Naturales (mean (sd))	72.00 (4.72)	71.74 (5.04)	0.585
Lectura Critica (mean (sd))	69.17 (6.37)	68.64 (6.26)	0.399
Inglés (mean (sd))	72.22 (10.28)	71.40 (10.66)	0.426
Global (mean (sd))	357.19 (21.63)	355.14 (23.00)	0.353
Riesgo Economica = BAJO (%)	189 (92.2)	183 (89.3)	0.394
Riesgo Social (%)			0.688
ALTO	10 (4.9)	8 (3.9)	
BAJO	160 (78.0)	156 (76.1)	
MEDIO	35 (17.1)	41 (20.0)	
Riesgo Salud = MEDIO (%)	203 (99.0)	203 (99.0)	1.000
Riesgo Academico (%)			0.376

ALTO	136 (66.3)	125 (61.0)	
BAJO	11 (5.4)	17 (8.3)	
MEDIO	58 (28.3)	63 (30.7)	
Riesgo_Cognitiva (%)			0.478
ALTO	44 (21.5)	51 (24.9)	
BAJO	29 (14.1)	22 (10.7)	
MEDIO	132 (64.4)	132 (64.4)	
Ben (%)			<0.001
(-1,0]	205 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	48 (23.4)	
(7,15]	0 (0.0)	29 (14.1)	
(15,23]	0 (0.0)	42 (20.5)	
(23,31]	0 (0.0)	57 (27.8)	
(31,55]	0 (0.0)	29 (14.1)	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
distance (mean (sd))	0.36 (0.13)	0.37 (0.14)	0.318

La Figura 93 muestra el resultado de 20 corridas de pareamiento modificando el umbral a partir del cual se considera como beneficiado un estudiante. La primera categoría ( $h > 7$ ) considera como beneficiario del SEA todo estudiantes que tenga al menos una hora de beneficio, la segunda categoría ( $h > 7$ ) corresponde a más de 7 horas de beneficio por semestre o media hora a la semana en promedio; las demás categorías se interpretan de forma análoga.

Se destaca que, en la medida que aumenta el número de horas de beneficio, el impacto del programa aumente y empieza a ser significativo por encima de 8 horas de beneficio en el semestre (media hora a la semana en promedio). Se mantiene significativo y creciente hasta 31 horas pero decae levemente y deja de ser significativo al 5% por encima de 32 horas de beneficio.

En síntesis, la técnica de pareamiento muestra un impacto significativo sobre la nota de cálculo I atribuida al uso del programa SEA a un nivel de significancia inferior al 5% en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2017 siempre y cuando el número de horas de beneficio sea al menos de 8 horas en el semestre o lo que es lo mismo, media hora a la semana en promedio. El impacto aumenta en función del número de horas llegando

a un valor máximo de 0.37 décimas de diferencia en la nota con más de 24 horas de beneficio en el semestre (hora y media de beneficio a la semana en promedio) y se mantienen en este valor con más de 32 horas de beneficio con 18% de significancia.

También es importante mencionar que la variabilidad de las diferentes corridas (línea azul punteada de la Figura 93) aumenta con el número de horas de beneficio; sin embargo, este aumento se debe a la disminución del número de estudiantes beneficiados en el pareamiento y no a la cantidad de horas de beneficio recibidas.

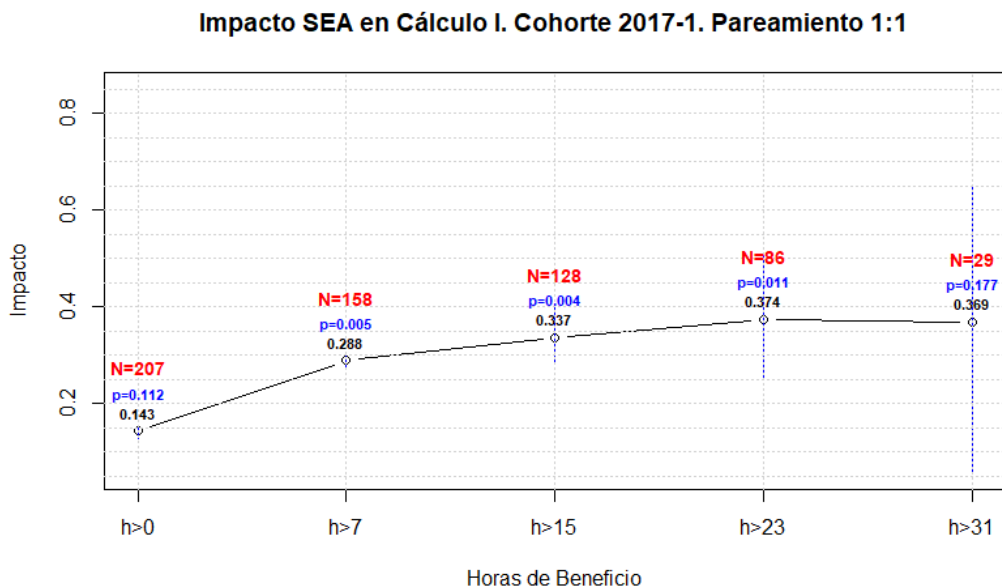


Figura 93. Impacto del programa SEA sobre cálculo I en función del número de horas de beneficio.

1ª cohorte de 2017

**3.2.4.2 Impacto del SEA sobre repitentes en cálculo I. Cohorte 2017-1.** La distribución de estudiantes de esta cohorte que recibieron beneficio del SEA los dos semestres que cursaron la materia dos veces se muestra en la Figura 94.

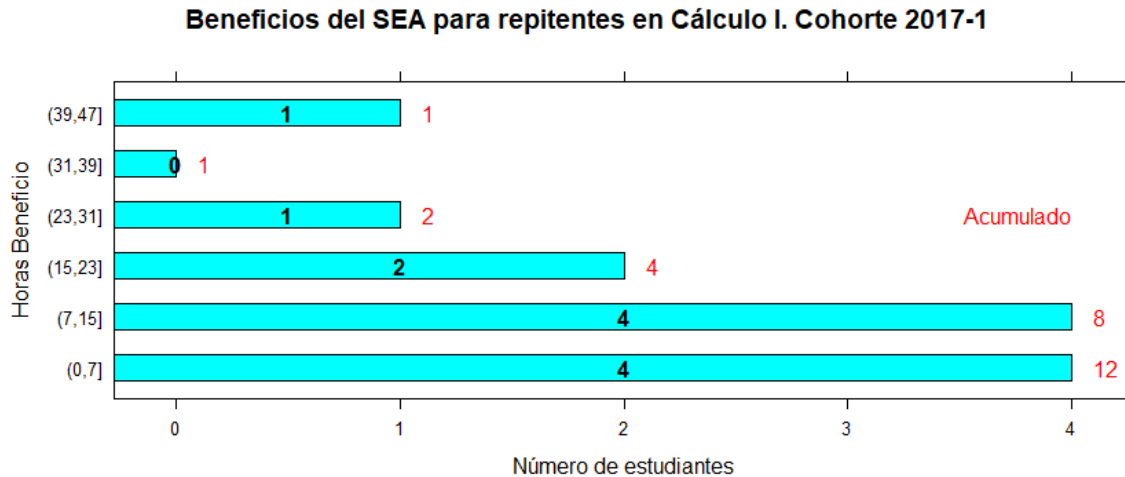


Figura 94. Beneficios del SEA para repitentes en cálculo I. 1ª cohorte de 2017.

La Figura 95 muestra que el cambio promedio en nota en el grupo de beneficiarios (0.82 décimas) es superior al cambio en el grupo de no beneficiarios (0.60 décimas) lo que sugiere un impacto positivo del programa SEA de 0.22 décimas la segunda vez que ve la materia, es decir cuando repite la materia. Sin embargo, este no resulta significativo desde el punto de vista estadístico como se logra apreciar en la Tabla 63, impacto que resulta similar al encontrado para los estudiantes del primera cohorte del 2016 (ver sección 0, página 204)

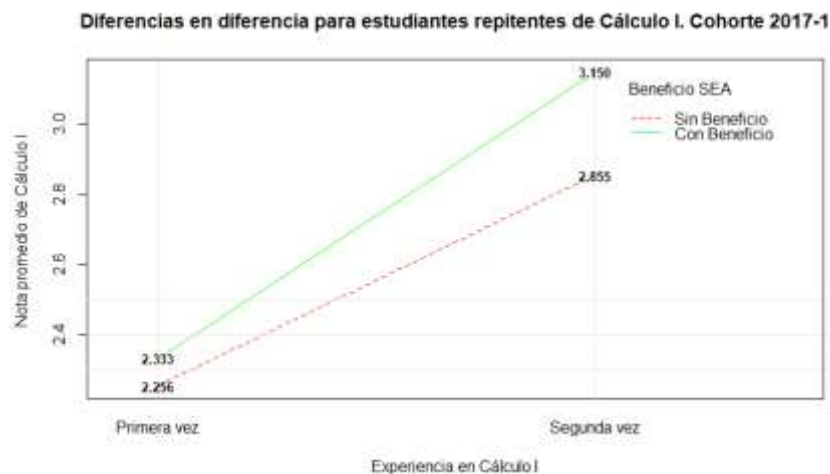


Figura 95. Diferencias en diferencias a repitentes en cálculo I. 1ª cohorte de 2017

Tabla 63.

*Estimador del impacto del programa SEA sobre repitentes de cálculo I usando diferencias en diferencias. 1ª cohorte de 2017*

Estimate	Std. Error	t value	Pr(> t )
0.2173469	0.2753322	0.7893989	0.4304746

**3.2.4.3 Impacto del SEA sobre Álgebra Lineal. Cohorte 2017-I.** De los 924 estudiantes de esta cohorte solamente 658 (71%) presentan información completa, incluidas las notas de álgebra lineal: 106 beneficiados y 552 no beneficiados. La distribución detallada por horas de beneficio se muestra en la Figura 96.

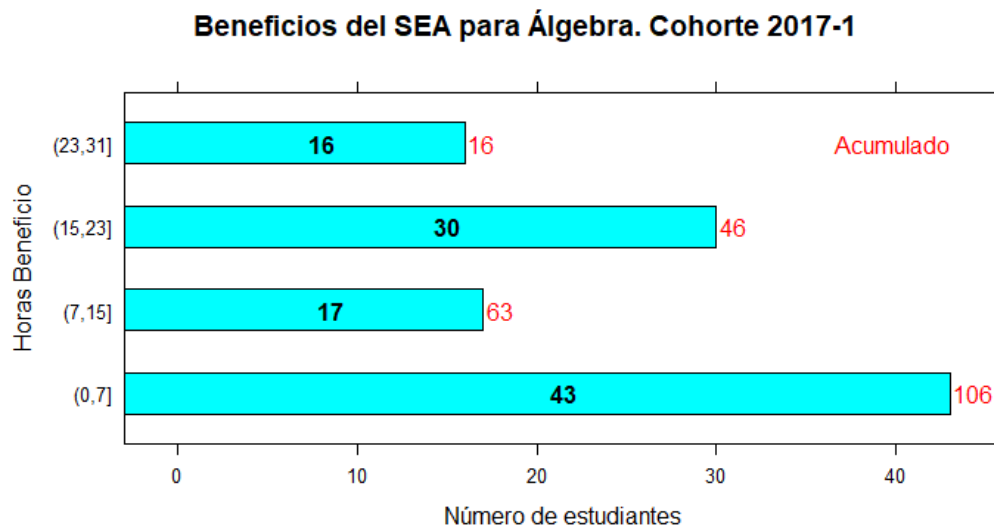


Figura 96. Beneficios del SEA en álgebra I. 1ª cohorte de 2017.

La Tabla 64 muestra la comparación entre el grupo de beneficiados y no beneficiados en álgebra de la primera cohorte del 2017. Se observa una diferencia de 0.05 unidades entre el grupo de beneficiados y no beneficiados aunque no significativa al 5%. Sin embargo, esta diferencia no solo se puede atribuir al beneficio recibido del SEA si no a otras co-variables o características previas

del estudiantes, que pueden inflar o de confundir el efecto real del programa. Específicamente, se encuentra que también existen diferencias significativas en el programa (valor  $p < 0.01$ ) la edad (valor  $p = 0.063$ ), el género (valor  $p = 0.002$ ) y el puntaje en las pruebas SABER 11 (valor  $p < 0.05$  en todos los casos) así como el riesgo económico (valor  $p = 0.001$ ) y el académico (valor  $p = 0.013$ ).

Tabla 64.

*Comparación del grupo de Beneficiados y no Beneficiados del SEA en álgebra I. 1ª cohorte de 2017*

	No Beneficiado	Beneficiado	p
n	552	106	
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)	<0.001
Nota (mean (sd))	3.29 (0.84)	3.35 (0.91)	0.534
Horas (mean (sd))	0.00 (0.00)	12.51 (8.86)	<0.001
Prog (%)			<0.001
Civ	60 (10.9)	8 (7.5)	
Ele	42 (7.6)	4 (3.8)	
Elo	43 (7.8)	12 (11.3)	
Fis	22 (4.0)	1 (0.9)	
Geo	35 (6.3)	8 (7.5)	
Ind	64 (11.6)	5 (4.7)	
IQu	43 (7.8)	8 (7.5)	
LMa	9 (1.6)	17 (16.0)	
Mat	15 (2.7)	2 (1.9)	
Mec	61 (11.1)	10 (9.4)	
Met	31 (5.6)	8 (7.5)	
Pet	37 (6.7)	7 (6.6)	
Qui	30 (5.4)	6 (5.7)	
Sis	60 (10.9)	10 (9.4)	
Edad (mean (sd))	17.90 (1.85)	17.56 (0.94)	0.063
Genero = Hombre (%)	396 (71.7)	59 (55.7)	0.002
Matemática (mean (sd))	76.34 (6.98)	74.73 (6.89)	0.029
Ciencias Naturales (mean (sd))	72.79 (5.39)	71.10 (5.52)	0.003
Lectura Critica (mean (sd))	69.69 (6.06)	68.13 (6.59)	0.017
Inglés (mean (sd))	73.09 (10.04)	68.59 (11.36)	<0.001
Global (mean (sd))	361.09 (22.11)	351.71 (25.03)	<0.001
Riesgo Economica = BAJO (%)	521 (94.4)	90 (84.9)	0.001
Riesgo Social (%)			0.253
ALTO	30 (5.4)	3 (2.8)	
BAJO	427 (77.4)	79 (74.5)	
MEDIO	95 (17.2)	24 (22.6)	
Riesgo Salud = MEDIO (%)	547 (99.1)	106 (100.0)	0.709
Riesgo Academico (%)			0.013

ALTO	368 (66.7)	56 (52.8)	
BAJO	36 (6.5)	13 (12.3)	
MEDIO	148 (26.8)	37 (34.9)	
Riesgo_Cognitiva (%)			0.554
ALTO	129 (23.4)	30 (28.3)	
BAJO	78 (14.1)	14 (13.2)	
MEDIO	345 (62.5)	62 (58.5)	
Ben (%)			<0.001
(-1,0]	552 (100.0)	0 (0.0)	
(0,7]	0 (0.0)	43 (40.6)	
(7,15]	0 (0.0)	17 (16.0)	
(15,23]	0 (0.0)	30 (28.3)	
(23,31]	0 (0.0)	16 (15.1)	

El modelo de propensión que incluye las variables más significativas (basadas en el criterio de AIC) se resume en la Tabla 65. Se destaca que las variables más influyentes en el modelo de propensión fueron el programa académico, la edad, genero, el puntaje de inglés en las pruebas SABER 11 así como los factores de riesgo económico y de salud.

Tabla 65.

*Resultados del modelo de propensión para beneficiados del programa SEA en álgebra I. 1ª cohorte de 2017 (Modelo depurado.)*

```
Call:
glm(formula = Beneficio ~ Prog + Edad + Genero + Ingles + Riesgo_Eco
nomica +
    Riesgo_Salud, family = binomial, data = Datos.Evaluacion2)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.8724	-0.5915	-0.4682	-0.3139	2.7826

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-8.45481	629.39788	-0.013	0.989282
ProgEle	-0.34737	0.66189	-0.525	0.599707
ProgElo	0.80991	0.51862	1.562	0.118366
ProgFis	-1.22217	1.10665	-1.104	0.269424
ProgGeo	0.42760	0.55654	0.768	0.442307
ProgInd	-0.69466	0.60901	-1.141	0.254016
ProgIQu	0.12974	0.55166	0.235	0.814072
ProgLMa	1.85398	0.62302	2.976	0.002922 **
ProgMat	-0.44957	0.89767	-0.501	0.616497
ProgMec	0.25775	0.51949	0.496	0.619780
ProgMet	0.15098	0.59362	0.254	0.799241

ProgPet	0.02500	0.58868	0.042	0.966120
ProgQui	0.02096	0.60901	0.034	0.972545
ProgSis	0.22192	0.52619	0.422	0.673213
Edad	-0.22811	0.10824	-2.107	0.035088 *
GeneroHombre	-0.84232	0.25331	-3.325	0.000884 ***
Inglés	-0.03295	0.01409	-2.339	0.019342 *
Riesgo_EconomicaBAJO	-0.73658	0.38068	-1.935	0.052999 .
Riesgo_SaludMEDIO	14.21364	629.39367	0.023	0.981983

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 580.99 on 657 degrees of freedom  
 Residual deviance: 508.14 on 639 degrees of freedom  
 AIC: 546.14

Number of Fisher Scoring iterations: 14

La Figura 97 muestra gráficamente los resultados del pareamiento y la Tabla 66 los muestra los numéricamente. Se observa la exclusión de 12 estudiantes no beneficiados y 2 beneficiados por encontrarse fuera del rango común de propensión; así como 434 no tenidos en cuenta en el pareamiento 1:1.

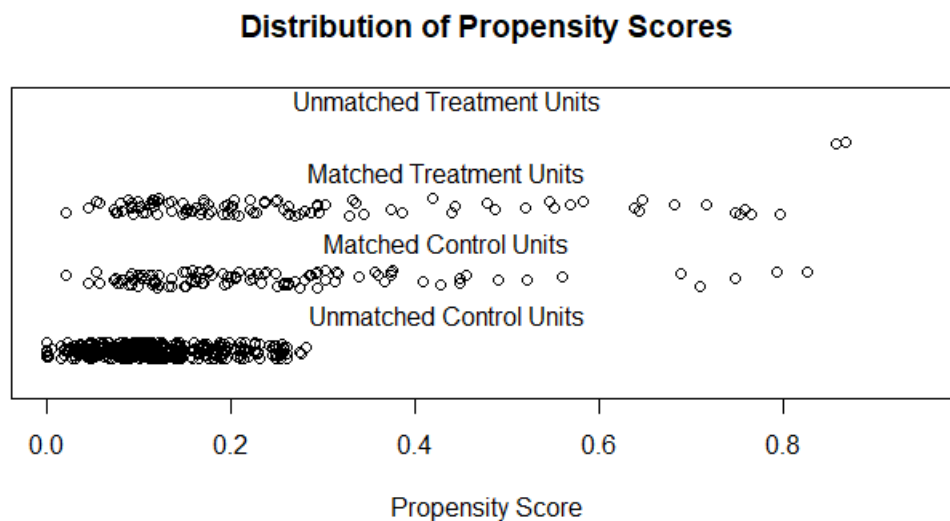


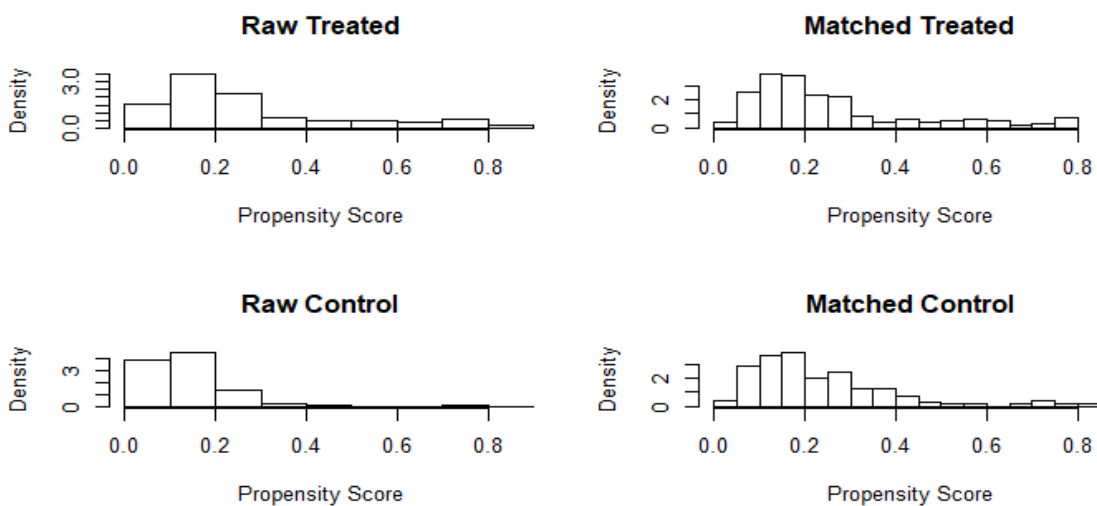
Figura 97. Pareamiento por puntajes de propensión para álgebra I. 1ª cohorte de 2017

Tabla 66.

*Resultados finales del pareamiento para beneficiados en álgebra I. 1ª cohorte de 2017 de Ciencias e Ingeniería.*

	Control	Treated
All	552	106
Matched	104	104
Unmatched	434	0
Discarded	14	2

El diagnóstico final de calidad del pareamiento se aprecia en la Figura 98, donde se aprecia que después del pareamiento se comparan dos grupos homogéneos en todas las propiedades excepto por el hecho que uno de ellos se benefició del programa SEA en álgebra I y el otro no lo hizo; los dos histogramas a la derecha muestran la distribución de puntajes de propensión entre el grupo de beneficiados y no beneficiados del SEA en álgebra después del pareamiento, distribuciones que resultan muy similares; contrario a lo que se observa con las dos distribuciones de la izquierda, que muestran los datos originales, antes del pareamiento.



*Figura 98.* Histograma de puntajes de propensión para álgebra I antes y después del pareamiento.

1ª cohorte de 2017

En síntesis, se puede decir que el pareamiento resultó efectivo y por tanto se puede emplear de manera apropiada para estimar un contrafactual de adecuado que reduzca el sesgo de selección en la evaluación del impacto del programa SEA en álgebra, para los estudiantes de la primera cohorte del 2017.

La comparación entre el grupo de beneficiados y no, después del pareamiento, se muestra en la Tabla 67. Se encontró una diferencia no significativa en la nota promedio de álgebra entre el grupo de beneficiados y el contrafactual obtenido empleando el método de pareamiento 1:1. Se observa que en las demás co-variables no existe una diferencia significativa por lo que los cambios en la nota de álgebra se pueden atribuir, sin mayor riesgo, al efecto del programa SEA.

Tabla 67.

*Comparación de estudiantes con beneficio o no del SEA en cálculo I (después de pareamiento). 1ª cohorte de 2017*

	No Beneficiado	Beneficiado	p
n	104	104	
Nota (mean (sd))	3.27 (0.79)	3.36 (0.92)	0.457
Horas (mean (sd))	0.00 (0.00)	12.52 (8.91)	<0.001
Prog (%)			0.906
Civ	7 (6.7)	8 (7.7)	
Ele	4 (3.8)	4 (3.8)	
Elo	20 (19.2)	12 (11.5)	
Fis	2 (1.9)	1 (1.0)	
Geo	5 (4.8)	8 (7.7)	
Ind	7 (6.7)	5 (4.8)	
IQu	6 (5.8)	8 (7.7)	
LMa	8 (7.7)	15 (14.4)	
Mat	2 (1.9)	2 (1.9)	
Mec	10 (9.6)	10 (9.6)	
Met	10 (9.6)	8 (7.7)	
Pet	9 (8.7)	7 (6.7)	
Qui	7 (6.7)	6 (5.8)	
Sis	7 (6.7)	10 (9.6)	
Edad (mean (sd))	17.51 (0.99)	17.57 (0.94)	0.667
Genero = Hombre (%)	63 (60.6)	59 (56.7)	0.673
Matemática (mean (sd))	73.88 (5.93)	74.77 (6.95)	0.325
Ciencias Naturales (mean (sd))	70.69 (5.61)	71.24 (5.48)	0.477
Lectura Critica (mean (sd))	68.34 (6.58)	68.15 (6.64)	0.842
Inglés (mean (sd))	68.47 (10.06)	68.81 (11.37)	0.821
Global (mean (sd))	349.58 (23.89)	352.16 (25.03)	0.447
Riesgo Economica = BAJO (%)	87 (83.7)	90 (86.5)	0.697

Riesgo_Social (%)				0.376
ALTO	6 (5.8)	3 (2.9)		
BAJO	81 (77.9)	78 (75.0)		
MEDIO	17 (16.3)	23 (22.1)		
Riesgo_Salud = MEDIO (%)	104 (100.0)	104 (100.0)		NA
Riesgo.Academico (%)				0.619
ALTO	62 (59.6)	55 (52.9)		
BAJO	11 (10.6)	13 (12.5)		
MEDIO	31 (29.8)	36 (34.6)		
Riesgo_Cognitiva (%)				0.678
ALTO	25 (24.0)	30 (28.8)		
BAJO	13 (12.5)	14 (13.5)		
MEDIO	66 (63.5)	60 (57.7)		
Ben (%)				<0.001
(-1,0]	104 (100.0)	0 (0.0)		
(0,7]	0 (0.0)	42 (40.4)		
(7,15]	0 (0.0)	17 (16.3)		
(15,23]	0 (0.0)	29 (27.9)		
(23,31]	0 (0.0)	16 (15.4)		
Beneficio (mean (sd))	0.00 (0.00)	1.00 (0.00)		<0.001
distance (mean (sd))	0.24 (0.16)	0.26 (0.19)		0.372

La Figura 99 muestra los resultados de aplicar la técnica de pareamiento 20 veces, variando el umbral en el número de horas recibidas de beneficio a partir del cual se considera beneficiario del SEA. Se observa un aumento del impacto al aumentar el número de horas de beneficio; sin embargo solo resulta significativo a partir de 8 horas de beneficio.

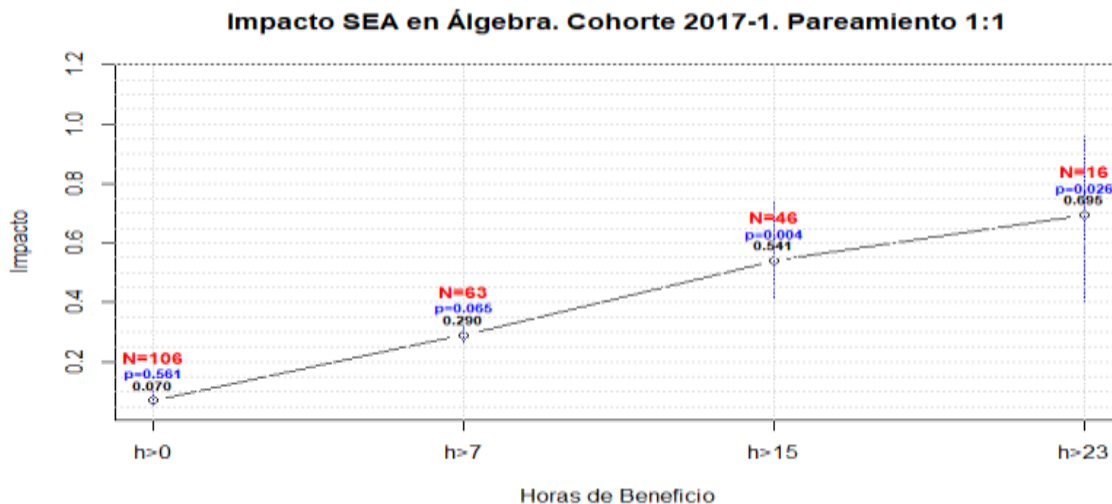


Figura 99. Impacto del programa SEA sobre álgebra lineal en función del número de horas de beneficio. 1ª cohorte de 2017

En síntesis, la técnica de pareamiento muestra un impacto significativo sobre la nota de álgebra atribuida al uso del programa SEA a un nivel de significancia inferior al 6.5% en los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2017 siempre y cuando, el número de horas de beneficio sea al menos de 8 horas en el semestre o lo que es lo mismo, media hora a la semana en promedio.

El impacto tiende a aumentar de manera significativa a partir de 8 horas de beneficio recibidas pasando de 0.3 décimas de beneficio con más de 8 horas por semestre (media hora semanal de beneficio en promedio), a 0.5 décimas con más de 16 horas (una hora en promedio a la semana) y 0.7 décimas con más de 24 horas de beneficio. Sin embargo, así como se aprecia un impacto creciente con el número de horas de beneficio, también se observa una incertidumbre cada vez mayor, debido a la reducción en el número de estudiantes en los que se basa el pareamiento conforme aumenta el umbral a partir del cual se considera beneficiario del SEA, pasando de 106 (con  $h > 0$  como umbral) hasta 16 (con  $h > 23$ , es decir más de 23 horas de beneficio como umbral para considerarlo beneficiario del SEA).

**3.2.4.4 Impacto del SEA sobre repitentes en álgebra lineal. Cohorte 2017-1.** Solo se encontraron tres estudiantes de esta cohorte que recibieron beneficio las dos veces que cursaron la materia. Se observa un impacto positivo de 0.43 décimas aunque este no resultó significativo al 5% como se observa en la Figura 100 y la Tabla 68

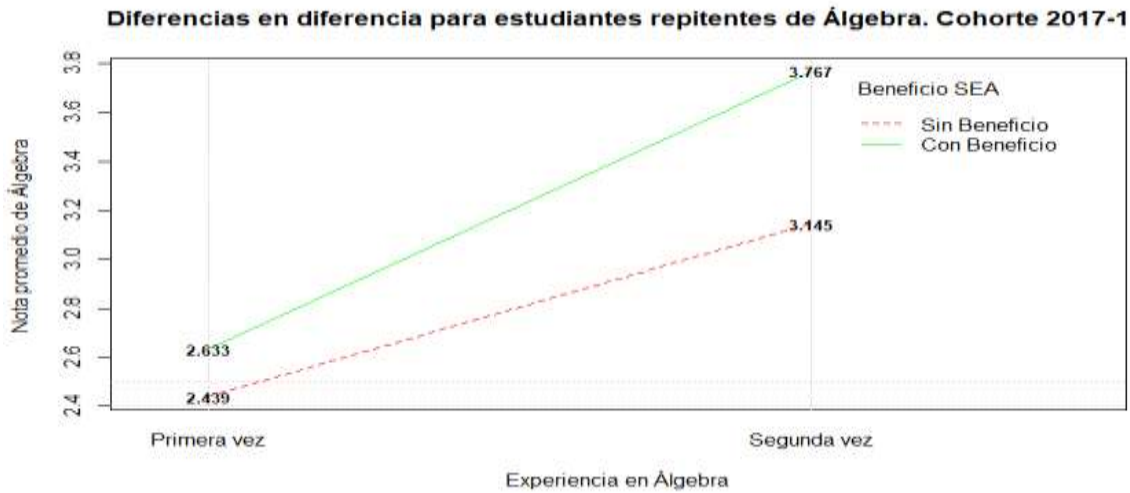


Figura 100. Diferencias en diferencias a repitentes en álgebra I. 1ª cohorte de 2017.

Tabla 68.

*Estimador del impacto del programa SEA sobre repitentes de álgebra I usando diferencias en diferencias. 1ª cohorte de 2017*

Estimate	Std. Error	t value	Pr(> t )
0.4278788	0.5508093	0.7768184	0.4380928

**3.2.4.5 Análisis consolidado y comparativo de las cohortes estudiadas.** En esta sección se realiza un análisis comparativo de los semestres analizados tratando de encontrar patrones generales en los modelos de propensión y determinar los factores que más afecta la inscripción de los estudiantes en el programa SEA según la asignatura.

La Tabla 69 muestra las variables que resultaron significativas en los modelos de propensión de las asignaturas del matemáticas (cálculo I, II y III y álgebra lineal) para estudiantes de la primera cohorte del 2014. Se observa, en general, que en el primer semestre las variables que más afectan la propensión (probabilidad de participar en el programa) son el programa académico, el

género (siendo mayor la participación de mujeres) y el puntaje en matemáticas, física y lenguaje de las pruebas SABER 11. En las materias del segundo semestre se aprecia también la influencia de la edad, y el desempeño académico en semestres anteriores así como las horas de beneficio recibidas en asignaturas vistas con anterioridad.

Tabla 69.

*Resumen de variables significativas en modelos de propensión, cohorte 2014-1*

Cohorte 2014-I / Semestre		1	1	2	3
Información	Asignatura	Álgebra	Cálculo I	Cálculo II	Cálculo III
Estudiantes	Totales	921	921	921	921
	Beneficiados	31	59	28	15
	No Beneficiados	806	371	277	250
General	Programa		*		
	Edad				
	Género	*		**	
Saber 11	Matemática	***	***		
	Física		*		
	Química				
	Biología			*	
	Lenguaje	*			
Caracterización SEA	R. Académico				
	R. Salud				*
	R. Económico				
Rendimiento Académico	N. Calculo I				
	Promedio Ac.				
Beneficio SEA	H. Álgebra				
	H. Cálculo I				
	H. Cálculo II				*
	H. Cálculo III				
Mayor Significancia			*	**	***



La Tabla 71 muestra los aspectos más relevantes del uso de la técnica de pareamiento en los semestres analizados. Se observa que el programa SEA impacta positivamente y de forma significativa a partir del 2017-1 en cálculo y álgebra, los semestres anteriores no se observó un impacto significativo e incluso en muchos casos, el impacto resulta negativo (destacados en rojo en la Tabla 71 ).

Se observa que solo en 2017 se contó con la información suficiente y completa para aplicar la técnica de pareamiento encontrando un impacto positivo y significativamente distinto de cero en el que se aprecia una tendencia lógica del impacto en función de la cantidad de horas de beneficio recibidas (ver Figura 93 y Figura 99). Es importante tener en cuenta que rechazar la hipótesis nula siempre se puede considerar una decisión fuerte ya que se puede estimar con mayor precisión la probabilidad de equivocarse con esta decisión (Error tipo I o valor P). Entre tanto, no rechazar la hipótesis nula,  $H_0$ , se considera una decisión débil ya que estimar la probabilidad de equivocarse (error tipo II) es función del tamaño de la muestra y del valor verdadero del impacto el cual se desconoce véase (Montgomery & Runger, 1996, pág. 379).

Por lo anterior, en lugar de decir que no existe impacto del programa SEA en las asignaturas y cohortes en los que el valor P dio por debajo del 5% se debe indicar que no hay evidencia suficiente, con la información disponible, para rechazar la hipótesis nula que el impacto es igual a cero.

Tabla 71.

*Resultados relevantes de evaluación de impacto usando el método de pareamiento.2014-2017*

Cohorte I	Semestre	Asignatura	H Beneficio	# Beneficiados	Impacto	valor P
2014-I	1	Álgebra	> 0	31	0.01	0.764
			≥ 8	11	-0.14	0.620
2014-I	1	Cálculo I	> 0	59	-0.04	0.835
			≥ 8	32	0.08	0.707
			≥ 16	16	-0.29	0.390
			≥ 24	9	0.36	0.375
2014-I	2	Cálculo II	> 0	28	-0.12	0.517
			≥ 8	12	-0.15	0.712
2014-I	3	Cálculo III	> 0	15	-0.23	0.510
2015-I	1	Cálculo I	> 0	116	0.01	0.917
			≥ 8	80	0.09	0.561
			≥ 16	69	0.19	0.257
			≥ 24	57	0.18	0.321
2016-I	1	Cálculo I	> 0	153	0.07	0.513
			≥ 8	89	-0.07	0.569
			≥ 16	71	0.19	0.175
			≥ 24	45	-0.15	0.484
2017-I	1	Cálculo I	> 0	207	0.14	0.112
			≥ 8	158	0.29	0.005
			≥ 16	128	0.34	0.004
			≥ 24	86	0.37	0.011
2015-I	1	Álgebra	> 0	61	-0.08	0.679
			≥ 8	12	0.38	0.261
2016-I	1	Álgebra	> 0	120	0.06	0.608
			≥ 8	63	0.46	0.009
			≥ 16	23	0.19	0.562
			≥ 24	12	0.07	0.802
2017-I	1	Álgebra	> 0	106	0.07	0.561
			≥ 8	63	0.29	0.065
			≥ 16	46	0.54	0.004
			≥ 24	16	0.70	0.026

Por lo anterior, solo en la primera cohorte del 2016 y más aún en la primera cohorte del 2017, donde el tamaño de la muestra es mayor, se encuentra un impacto significativo del programa SEA sobre cálculo I y álgebra lineal que resulta ser una conclusión fuerte desde el punto de vista

estadístico y con un patrón lógico del impacto en función del número de horas de beneficio recibidas. No obstante es importante tener en cuenta que para esta cohorte, el impacto en cálculo I tiende a estabilizarse en 0.3 décimas con más de 8 horas de beneficio, impacto que no mejora sustancialmente al aumentar el número de horas de beneficio recibidas.

La Tabla 72 resume los resultados del uso de la técnica de diferencias en diferencias sobre los repitentes en las asignaturas de cálculo y álgebra. Se destacan que en ningún caso el impacto dio significativamente distinto de cero, tal vez debido a la cantidad tan reducida de estudiantes con sobre los que se aplicó la técnica de diferencias en diferencias.

Tabla 72.

*Resumen de resultados claves del uso del método de Diferencias en Diferencias en estudiantes repitentes de cálculo y álgebra lineal. 2014-2017*

Asignatura	Cohorte I	# Beneficiados	Impacto	valor P
<b>Cálculo</b>	2014-I	7	-0.04	0.992
	2015-I	15	-0.01	0.969
	2016-I	6	0.25	0.557
	2017-I	12	0.22	0.431
<b>Álgebra</b>	2014-I	3	0.17	0.755
	2015-I	-	-	-
	2016-I	10	0.27	0.359
	2017-I	3	0.42	0.440

#### 4. Conclusiones

Las variables que más influyen la participación de los estudiantes en el programa SEA (propensión), son el programa al que pertenecen, el género y el puntaje de la prueba SABER 11, especialmente en matemáticas (Tabla 70).

La caracterización realizada por el programa SEA muestra concordancia con la propensión en las dimensiones económica y cognitiva, pero no con la dimensión académica lo que sugiere un replanteamiento de esta prueba de caracterización (Tabla 70).

Para los estudiantes de ciencia e ingeniería que ingresaron a la UIS de la primera cohorte del 2014:

- La técnica de pareamiento no mostró un impacto significativo sobre el desempeño académico del SEA, en las asignaturas de matemáticas (cálculos, álgebra y ecuaciones diferenciales - Tabla 71). Se observa un leve aumento del impacto por encima de 24 horas de beneficio para cálculo I, aunque no resultó significativo (Figura 48).
- La técnica de diferencias en diferencias no mostró impacto significativo repitentes de cálculo I (Tabla 23), álgebra (Tabla 29) y cálculo II (Tabla 35) del SEA. No obstante, con un número tan reducido de estudiantes beneficiados en cada caso, estos resultados pueden considerarse no concluyente, además esta misma condición impidió el análisis de las demás asignaturas de matemáticas (Tabla 72).

Para los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2015:

- En Cálculo I, se encontró un impacto de 0.2 décimas en la nota promedio, atribuidas al uso del programa SEA, significativo al 26%, con un número de horas de beneficio superior a 16 (Figura 71). No se observó un impacto significativo del SEA en los estudiantes repitentes de cálculo I (Tabla 46).
- En álgebra, el impacto del SEA solo empieza a ser positivo (0.4 décimas) con más de 8 horas de beneficio por semestre, significativo al 26% (Figura 77). No se pudo aplicar la técnica en los repitentes de álgebra por falta de información completa.

Para los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2016:

- En cálculo I, se encontró un impacto del programa SEA de 0.2 décimas en la nota promedio, significativo al 18%, con un número de horas de beneficio entre 16 y 23 horas por semestre (Figura 81). También se encontró un impacto positivo de 0.25 décimas del programa SEA sobre los repitentes en cálculo I aunque no resultó significativo (Tabla 54).
- En álgebra, el impacto llegó a ser de 0.46 décimas y significativo al 0.9%, en el rango 8 a 15 horas de beneficio (Figura 87). También se encontró un impacto positivo de 0.27 décimas en la nota aunque solo significativo al 36% (Tabla 58).

Para los estudiantes de ciencias e ingeniería que ingresaron a la UIS en la primera cohorte del 2017:

- En cálculo I, se encontró un aumento progresivo del impacto del SEA con el número de horas de beneficio, con un valor máximo en 0.37 décimas en la nota promedio, con 24 horas de beneficio en el semestre (Figura 93).
- En álgebra, se encontró un aumento progresivo del impacto del SEA con el número de horas de beneficio, con un valor máximo en 0.7 décimas en la nota promedio, con 24 horas de beneficio en el semestre (Figura 99).
- En los repitentes de cálculo I se encontró un impacto de 0.22 décimas (Tabla 63) y en álgebra de 0.42 décimas (Tabla 68), pero no significativo desde el punto de vista estadístico.

## 5. Recomendaciones

Se recomienda extender el uso de las técnicas de evaluación impacto (pareamiento y diferencias en diferencias) para evaluar el impacto del programa SEA en otras asignaturas cubiertas por el programa para reducir la subjetividad en la toma de decisiones. No obstante, con base en la experiencia del presente estudio, se sugiere realizar este ejercicio solamente a partir del 2017 donde se tuvo una cobertura suficiente y la información completa que permiten aplicar las técnicas de evaluación de impacto de forma apropiada.

Se recomienda incluir variables adicionales en los modelos de propensión como el estrato socio-económico del estudiante, el año y colegio en el que se graduó de bachiller, la formación académica

de los padres, los ingresos netos del núcleo familiar, ya que pueden contribuir a mejorar la estimación del contrafactual y en consecuencia el pareamiento en estudiantes de primer semestre.

Se recomienda usar de forma combinada la técnica de diferencias en diferencias con la técnica de pareamiento para evaluar el impacto del programa SEA sobre los repitentes en todas las asignaturas, especialmente en las de mayor mortandad académica.

Es muy importante que la evaluación de impacto se pueda complementar con un análisis costo beneficio y un análisis costo efectividad como lo sugieren Gertler y cols. (2017). Esto permite a los responsables del programa tomar decisiones bien fundamentadas sobre las inversiones a realizar y las modificaciones a realizar en el programa.

### Referencias Bibliograficas

- Rico, J., Rodríguez Hernández, A., & Vilalta Alonso, J. (2012). EMPLEO DE LA REGRESIÓN LOGÍSTICA ORDINAL PARA LA PREDICCIÓN DEL RENDIMIENTO ACADÉMICO. *REVISTA INVESTIGACION OPERACIONAL*, 252-267.
- Bernal, R., & Peña, X. (2016). *Guía práctica para evaluación de impacto*. Bogotá: Universidad de los Ande. Obtenido de <https://ediciones.uniandes.edu.co/paginas/DetalleLibro.aspx?lid=144>
- Botello, I. C., & Parada R., S. E. (2013). Tutorías entre pares: una oportunidad de formación para futuros profesores de matemáticas. *REVISTA CIENTÍFICA / ISSN 0124 2253/*.
- Cabezas Baños, C. M. (2010). *Análisis de las cifras de desercion en pregrado de la universidad industrial de santander en el periodo 2003-2006*. Bucaramanga: Universidad Industrial de Santander.
- Cardozo Ortiz, C. E. (2010). *PAMRA : reconstruccion de realidades, sueños y vidas desde una mirada colectiva*. Bucaramanga: Universidad Industrial de Santander.
- Carvajal Olaya, P., Trejos Carpintero, A., & Corzo Isaza, C. A. (5 de 2006). Estudio de causas de deserción de los estudiantes de la universidad tecnológica de Pereira entre enero/2000 – diciembre/2004 utilizando la técnica de análisis de correspondencias simples. *Scientia et Technica Año XII. No 30*.
- Cendales, B., Vargas-Trujillo, E., & Barbosa, C. (2013). Factores psicológicos asociados al desempeño académico en los cursos universitarios de estadística: diferencias por sexo y área de titulación. *Avances en Psicología Latinoamericana/Bogotá (Colombia)*, 63-375.

Centro Interuniversitario de Desarrollo. (1971). *CINDA*. Recuperado el 29 de 03 de 2017, de CINDA: <http://www.cinda.cl/htm/es.htm>

Centro Interuniversitario de Desarrollo. (2006). *Repitencia y Deserción Universitaria en América Latina* (1 ed.). Santiago, Chile: Centro Interuniversitario de Desarrollo, CINDA.

Cepeda Gómez, F. (2004). *Actualización e implantación del sistema de información del programa de asesorías para el mejoramiento del rendimiento académico pamra de la división de bienestar universitario*. Bucaramanaga: Universidad Industrial de Santander.

Colaboradores de Wikipedia. (28 de 4 de 2017). *Wikipedia*. Recuperado el 30 de 6 de 2017, de Wikipedia:

[https://es.wikipedia.org/w/index.php?title=Valor\\_at%C3%ADpico&oldid=98711428](https://es.wikipedia.org/w/index.php?title=Valor_at%C3%ADpico&oldid=98711428)

Colaboradores de Wikipedia. Coeficiente de correlación de Pearson [en línea]. Wikipedia, La enciclopedia libre. (16 de 07 de 2017). *Wikipedia, La enciclopedia libre*. Obtenido de Wikipedia, La enciclopedia libre.: [https://es.wikipedia.org/w/index.php?title=Coeficiente\\_de\\_correlaci%C3%B3n\\_de\\_Pearson&oldid=98139012](https://es.wikipedia.org/w/index.php?title=Coeficiente_de_correlaci%C3%B3n_de_Pearson&oldid=98139012).

Colombia Aprende. (2015). *Colombia Aprende. La red del conocimiento*. Recuperado el 28 de 03 de 2016, de <http://www.colombiaaprende.edu.co/html/micrositios/1752/w3-article-345093.html>

Corengia, A., Pita, M., & Centeno, A. (2012). LA PREDICCIÓN DE RENDIMIENTO ACADÉMICO Y DESERCIÓN EN ESTUDIANTES UNIVERSITARIOS. *ISSN: 2223-7666 (Impresa), ISSN: 2223-7666 (Digital)*.

Díaz M., L. G. (2002). *Estadística multivariada: inferencia y metodos*. Bogotá: Panamericana Formas e Impresos S.A.

- División de Publicaciones UIS. (2015). Reglamento académico-estudiantil de pregrado. (UIS, Recopilador) Bucaramanga, Santander, Colombia. Obtenido de <https://www.uis.edu.co/webUIS/es/acercaUis/reglamentos/reglamentoPregrado.pdf>
- Durán, D. (2013). *Estadísticas e Indicadores de Deserción Estudiantil*. Recuperado el 02 de 04 de 2017, de Ministerio de Educación Nacional, <http://www.mineduacion.gov.co/>: [http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702\\_archivo\\_pdf\\_indicadores\\_permanencia.pdf](http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702_archivo_pdf_indicadores_permanencia.pdf)
- Espinoza, E. (2006). Impacto del maltrato en el rendimiento académico. *Revista electronica de Investigación Psicoeducativa*, 221-238.
- García Jiménez, M. V., Alvarado Izquierdo, J. M., & Jiménez Blanco, A. (2000). La predicción del rendimiento académico: regresión lineal versus regresión logística. *Psicothema*, 12, 248-252.
- Gertler, P. J., Martínez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2011). *La evaluación de impacto en la práctica*. Washington: Banco Mundial.
- Gertler, P. J., Martínez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2017). *La evaluación de impacto en la práctica* (2 ed.). Washington, DC: Banco Mundial. Obtenido de <http://www.worldbank.org/ieinpractice>
- Gutiérrez Pulido, H., & De la Vara Salazar, R. (2008). *Análisis y Diseño de Experimentos*. Mexico: Mc Graw Hill.
- Johnson, D. E. (2000). *Métodos Multivariados Aplicados al Análisis de Datos*. Mexico: Internatinal Thonson Editores.

- Lizarazo Velasco, D. R. (2006). *Factores que inciden en el rendimiento académico de los estudiantes beneficiarios del programa pamra, en el segundo semestre académico de 2003*. Bucaramanga: Universidad Industrial de Santander.
- Martínez Rosas, A. K., Caceres Martinez, J., & Ramirez Prada, P. C. (2011). *Sistema de informacion para el soporte de las actividades del programa de asesoria para el mejramiento del rendimiento académico (pamra) de bienestar universitario*. Bucaramanga: Universidad Industrial de Santander.
- Mazza, M. N.-A. (2007). *Análisis multivariado del desempeño académico de estudiantes universitarios de Química*. Madrid: Universidad Autónoma de Madrid.
- Mendoza Mendoza, A., & Herrera Acosta, R. (2013). Propuesta para la predicción del rendimiento académico de los estudiantes de la universidad del atlántico, basado en la aplicación del análisis discriminante. *Word Enginnering Education Forum, Cartagena*.
- Ministerio de Educación Nacional. (2008). *Disersion Estudiantil en la Educación Siperior Colombiana*. Bogotá: Ministerio de Educación Nacional.
- Ministerio de Educación Nacional. (11 de 2012). *Ministerio de Educación Nacional*. Recuperado el 30 de 03 de 2017, de MEN: [https://www.google.com.co/url?sa=t&rc t=j&q=&esc=s&source=web&cd=6&ved=0ahUKEwjE4LfUjf\\_SAhWIRCYKHc1gCEcQFggxMAU&url=http%3A%2F%2Fwww.mineduacion.gov.co%2F1621%2Farticles-298484\\_ar\\_chi vo\\_xls\\_diciembre\\_2012.xls&usg=AFQjCNHIBCD0RAhWrKtIErmv8qc DVCT0pw&bvm=bv](https://www.google.com.co/url?sa=t&rc t=j&q=&esc=s&source=web&cd=6&ved=0ahUKEwjE4LfUjf_SAhWIRCYKHc1gCEcQFggxMAU&url=http%3A%2F%2Fwww.mineduacion.gov.co%2F1621%2Farticles-298484_ar_chi vo_xls_diciembre_2012.xls&usg=AFQjCNHIBCD0RAhWrKtIErmv8qc DVCT0pw&bvm=bv).
- Ministerio de Educación Nacional. (2015). *Del problema de la deserción estudiantil a la apuesta por la permanencia y la graduación*. Bogotá: Ministerio de Educación Nacional.

- Ministerio de Educación Nacional. (2015). *Estadísticas deserción y graduación 2015*. Recuperado el 28 de 03 de 2016, de [http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-357549\\_recurso\\_3.pdf](http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-357549_recurso_3.pdf)
- Ministerio de Educación Nacional. (2015). *Estrategias para la Permanencia en Educación Superior: Experiencias Significativas*. Bogotá: Sanmartín Obregón & Cía Ltda.
- Ministerio de Educación Nacional. (2016). *Factores determinantes de la deserción*. Bogotá: Ministerio de Educación Nacional.
- Ministerio de Educación Nacional. (28 de 03 de 2016). *Sistema de Información SPADIES*. (MEN, Ed.) Recuperado el 28 de 03 de 2016, de <http://www.mineducacion.gov.co/sistemasdeinformacion/1735/w3-propertyname-2895.html>
- Miñano Pérez, P. (2009). *Un Modelo Causal-Explicativo sobre la incidencia de las variables Cognitivo-Motivacionales en el rendimiento Académico*. Universidad de Alicante.
- Montgomery, D. C., & Runger, G. (1996). *Probabilidad y Estadística Aplicadas a la Ingeniería*. México: McGraw-Hill.
- Olaya, P. C., Mosquera M, J. C., & Artamonova, I. (2009). MODELOS DE PREDICCIÓN DEL RENDIMIENTO ACADEMICO EN MATEMÁTICAS I EN LA UNIVERSIDAD TECNOLÓGICA DE PEREIRA. *Scientia et Technica Año XV(43)*, 258-263.
- Osma Castellanos, W., Mojica Perdomo, A., & Rivera Flórez, T. (2014). Factores asociados al rendimiento en las pruebas saber pro en estudiantes de ingeniería civil en universidades colombianas. *Innovaciencia facultad cienc. exactas fis. naturales*.
- Perez Pulido, M. O., & Roa Fuentes, Hernan Arturo. (2008). *Estudio de la relacion del rendimiento academico de los estudiantes de primer semestre de las ingenierias en la uis con los*

- resultados de las pruebas de estado a través del análisis de componentes principales y tablas de contingencia.* Bucaramanga: Universidad Industrial de Santander.
- Pérez Pulido, M., & Roa Fuentes, H. (2008). *Estudio de la relación del Rendimiento académico de los estudiantes de primer semestre de las Ingenierías en la UIS con los resultados de las pruebas de estado a través del Análisis por Componentes Principales y tablas de contingencia.* Bucaramnaga: Universidad Industrial de Santander.
- Porcel, E. A., Dapozo, G. N., & López, M. V. (2011). Predicción del Rendimiento Académico de Alumnos del Primer año de Universidad mediante Redes Neuronales. *Inevstigación Operativa.*
- Porcel, E., Dapozo, G., & López, M. (2010). Predicción del rendimiento académico de alumnos de primer año de la FACENA (UNNE) en función de su caracterización socioeducativa. *Revista Electrónica de Investigación Educativa.*
- Quin Alonso, I. A. (2008). *Factores que inciden en el bajo rendimiento académico de los estudiantes condicionales, repitiendo materias y por fuera de la universidad industrial de santander.* Bucaramanga: Universidad Industrial de Santander.
- Socha Pedraza, C. M. (2009). *Estudio sobre las motivaciones de desercion estudiantil en la universidad industrial de santander.* Bucaramanga: Universidad Industrial de Santander.
- Socha Pedraza, C. M. (2009). *Estudio sobre las motivaciones de deserción estudiantil en la Universidad Industrial de Santander.* Bucaramanga: Universidad Industrial de Santander.
- Torrado Arenas, D. M., Manrique Hernández, E. F., & Ayala Pimentel, J. O. (2009). La tutoría entre pares: una estrategia de enseñanza y aprendizaje de histología en la Universidad Industrial de Santander . *médicas uis.*

Universidad Industrial de Santander. (01 de 01 de 2014). *Sistema de Apoyo a la Excelencia Académica*. (U. I. Santander, Editor) Recuperado el 28 de 03 de 2016, de SEA-UIS: <https://www.uis.edu.co/webUIS/es/estudiantes/excelenciaAcademica/>

Universidad Industrial de Santander SEA-UIS. (01 de 01 de 2014). *Sistema de Apoyo a la Excelencia Académica, SEA*. (U. I. Santander, Editor) Recuperado el 28 de 03 de 2016, de SEA-UIS: <https://www.uis.edu.co/webUIS/es/estudiantes/excelenciaAcademica/>

Vázquez, C., Cavallo, M., Aparicio, S., Muñoz, B., Cynthia M., R., Luciana I., R., . . . Escobar, M. (2012). FACTORES DE IMPACTO EN EL RENDIMIENTO ACADÉMICO UNIVERSITARIO. UN ESTUDIO A PARTIR DE LAS PERCEPCIONES DE LOS ESTUDIANTES. *Decimoséptimas Jornadas "Investigaciones en la Facultad" de Ciencias Económicas y Estadística*.

Wikipedia. (27 de 8 de 2017). *Colaboradores de Wikipedia. Test exacto de Fisher [en línea]*. *Wikipedia, La enciclopedia libre, 2017*. Obtenido de [https://es.wikipedia.org/w/index.php?title=Test\\_exacto\\_de\\_Fisher&oldid=98806305](https://es.wikipedia.org/w/index.php?title=Test_exacto_de_Fisher&oldid=98806305)

Wikipedia, La enciclopedia libre. (9 de 11 de 2017). *Criterio de información de Akaike. (2017, 5 de noviembre)*. Obtenido de [https://es.wikipedia.org/w/index.php?title=Criterio\\_de\\_informaci%C3%B3n\\_de\\_Akaike&oldid=103160566](https://es.wikipedia.org/w/index.php?title=Criterio_de_informaci%C3%B3n_de_Akaike&oldid=103160566)

## Apéndices

En esta sección de anexos se incluyen los script en R más relevantes empleados en el presente trabajo. Los cuales pueden servir como guía para estudios similares.

### Apéndice A. Script principal en R

```
# Borra las variables anteriores
rm(list = ls())

# Cambia el Directorio de Trabajo
setwd('D:/OneDrive - Ecopetrol S.A/RStudio/Especializacion')

# Ejecutar Script de R con acentos y caracteres especiales sin problemas
Run.File=function (RFile) source(RFile, encoding = 'UTF-8')

#Impota datos 2014 desde MS Excel y genera variables adicionales
Run.File('0000-Funciones.R') # Carga librerías y funciones propias requeridas
Run.File('0001-Importar_Notas.R') # Notas de estudiantes 2014-1 analizados hasta 2016-2
Run.File('0002-Importar_Beneficiarios.R') # Beneficiarios SEA Todos
Run.File('0003-Importar_Caracterizados.R') # Caracterizados SEA 2014-1
Run.File('0004-ASAE.R') # Agrega datos Académico y cognitivo
Run.File('0005-Datos_Evaluacion_Impacto.R') # Datos para evaluación de impacto de 2014-1
#Run.File('0006-Datos_A_Excel.R') # Datos a Excel.

# Análisis descriptivo de los datos
Run.File('001-Hist.Genero.R') # Histograma de estudiantes por género 2014-1
Run.File('002-Dist.Genero.R') # Distribución por género y carrera 2014-1
Run.File('005-Prop.Genero.R') # Proporción por género y carrera 2014-1

Run.File('008-Hist.Edad.R') # Histograma de estudiantes por edad 2014-1
Run.File('009-Dist.Edad.R') # Distribución por edad 2014-1
Run.File('012-Prop.Edad.R') # Proporción por edad 2014-1

Run.File('013-Cajas.Prom.Sem.R') # Cajas de promedio semestral por carreras 2014-1
Run.File('017-Cajas.Prom.Acu.R') # Cajas de promedio acumulado por carreras 2014-1
Run.File('021-Cond.Sem.R') # Condicionalidad calculada del promedio acumulado 2014-1
Run.File('027-Cond.Sem.Fg.R') # Condicionalidad de DB notas UIS 2014-1
Run.File('033-Continuidad.R') # Continuidad de DB notas UIS
Run.File('039-Cajas_Mat.R') # Cajas de notas asignaturas de matemáticas (Calculos, Algebra y ED)
Run.File('044-Cajas_SABER11.R') # Cajas SABER 11 de matemáticas en DB notas UIS
Run.File('053-Benef_SEA.R') # Beneficiarios SEA totales, en Matemáticas(Calculos, Algebra y ED),
con cero Horas por carrera
Run.File('055-Cajas_SEA_Mat.R') # Horas Beneficiarios SEA en Matemáticas (Calculos, Algebra y ED) y
Run.File('067-Dist_SEA_Mat.R') # Beneficiados SEA por asignatura mat. en cada carrera
Run.File('071-Cajas_Sub_Prog_SEA.R') # Horas Beneficiarios por sup-programa SEA (ASAE, PAMRA, MIDAS, FPC) y
Carrera
Run.File('073-Sub_Programa_SEA.R') # Beneficiarios sup-programa SEA (ASAE, PAMRA, MIDAS, FPC) por Carrera
Run.File('077-Sub_SEA_por_Asignatura.R') # Beneficios Sub-Programa SEA (ASAE, PAMRA, MIDAS, FPC) por
asignatura en Matemáticas
Run.File('081-Caracterizados_SEA.R') # Beneficios Sub-Programa SEA (ASAE, PAMRA, MIDAS, FPC) por
asignatura en Matemáticas
Run.File('091-SABER_Vs_Rendimiento.R') # Graficas bivariadas de SABER Vs Rendimineto
Run.File('094-SEA Vs Calculo I.R') # Nota Calculo I Vs Horas de Beneficio SEA
Run.File('095-SEA Vs Calculo II.R') # Nota Calculo II Vs Horas de Beneficio SEA
Run.File('096-SEA Vs Algebra.R') # Nota Álgebra Vs Horas de Beneficio SEA
Run.File('097-SEA Vs Calculo III.R') # Nota Calculo I Vs Horas de Beneficio SEA
Run.File('098-SEA Vs Ecuaciones.R') # Nota Ecuaciones Diferenciales Vs Horas de Beneficio SEA
Run.File('099-Caracterización SEA.R') # Caracterización Vs Riesgo
Run.File('100-Analisis_Global_PorCarreras.R') # Análisis global por Carreras
```

```
Run.File('110-Importar_Datos_2015-6.R')      # Importa datos 2015-2016
Run.File('111-Importar_Datos_2017.R')      # Importa datos 2017
Run.File('112-Unificación de Datos.R')     # Unifica bases de datos para gráficas comparativas
Run.File('113-Importar_Notas2014a2017.R')  # Notas de estudiantes 2014 a 2017
Run.File('114-Notas_Cohorte_Asignatura.R') # DB Notas por cohorte y asignatura
Run.File('115-Beneficio_SEA_Por_Estudiante.R')

# Análisis descriptivo 2014 a 2017 comparativo
Run.File('120-Distribucion_Por_Genero.R')
Run.File('130-Distribucion_Por_Edad.R')
Run.File('140-Cajas_Promedio_Semestral.R')
Run.File('145-Cajas_CalculoI.R')
Run.File('150-Cajas_Algebra.R')
Run.File('155-Cajas_MatematicasSABER11.R')
Run.File('160-Cajas_BeneficioCalculoI.R')
Run.File('165-Cajas_BeneficioAlgebra.R')
Run.File('170-Condicionabilidad_Por_Ano.R')
Run.File('180-Beneficiados_SEA_CalculoI.R')
Run.File('190-Beneficiados_SEA_Algebra.R')

# Pareamiento 2014-1
Run.File('200-Pareamiento 2014-1 (Calculo I).R')
Run.File('200-Pareamiento 2014-1 (Algebra).R')
Run.File('200-Pareamiento 2014-1 (CalculoII).R')
Run.File('200-Pareamiento 2014-1 (CalculoIII).R')
Run.File('200-Pareamiento 2014-1 (Ecuaciones).R')

# Pareamiento 2015-1
Run.File('201-Pareamiento 2015-1 (Calculo I).R')
Run.File('201-Pareamiento 2015-1 (Algebra).R')

# Pareamiento 2016-1
Run.File('202-Pareamiento 2016-1 (Calculo I).R')
Run.File('202-Pareamiento 2016-1 (Algebra).R')

# Pareamiento 2017-1
Run.File('203-Pareamiento 2017-1 (Calculo I).R')
Run.File('203-Pareamiento 2017-1 (Algebra).R')

# Diferencias 2014-1
Run.File('210-Diferencias 2014-1(Algebra).R')
Run.File('210-Diferencias 2014-1(Calculo I).R')
Run.File('210-Diferencias 2014-1(Calculo II).R')
Run.File('210-Diferencias 2014-1(Calculo III).R')
```

## Apéndice B. Funciones en R y paquetes empleados

```

#Función para guardar figuras como png en el directorio actual
guardar_png=function (g, nombre=paste(g$ylab, "Vs", g$xlabel),
                        w = 2300/2, h = 1400/2, bg = "white"){
  #print(g)
  archivo=paste('./Figuras/',nombre, '.png',sep = '') # Nombre del archivo con extensión png
  png(archivo, width = w, height = h, bg = bg) # Abre el dispositivo gráfico
  print(g) # Muestra la figura en el dispositivo grafico actual
  dev.off() # Manda al dispositivo gráfico
}

#Cortes multiples de 8 requeridos para las horas de beneficio SEA recibidos
cortes.de.8=function(Horas, est.min=5, fusionar=TRUE,cortes=c(-1,0,8*(1:20)-1) ){

  # Cortes menores al máximo número de horas recibidas
  n=sum(cortes<max(Horas))+1
  cortes=cortes[1:n]
  if (fusionar){
    tbl=table(cut(Horas,breaks = cortes))
    #fusiona clases con menos de una cantidad minima de estudiantes, est.min
    pos=c(TRUE,tbl>est.min)
    cambiar=!pos[n]
    pos[n]=TRUE
    cortes=cortes[pos]
    if (cambiar){
      n=length(cortes)-1
      cortes=cortes[-n]
    }
  }
  # Cortes de interés
  cortes
}

#Estadísticos por grupo
estad=function(x) {
  m=mean(x)
  me=mean(x,trim = 0.5)
  s=sd(x)
  v=var(x)
  cv=s/m
  y= c(m,me,s,v,cv )
  names(y)=c('Media','Mediana', 'Desviacion', 'Varianza','Cv')
  y
}

# Diagrama de cajas para promedio general de notas
fig.cajas=function(cex=3, # Tamaño de letra
                  d=datos, # datos
                  tit=titulo, # Título principal
                  xl=xlabel, # Etiqueta del eje x
                  pos.lab=1, # Posición de la etiqueta descriptiva (x,s)
                  mark.x=0:5, # Posición de Marcas en eje x
                  delta.x=(mark.x[2]-mark.x[1]), # Delta de marcas en eje x
                  lim.x=c(mark.x[1]-2*delta.x,range(mark.x)[2]+delta.x), # Rango de visualización eje x
                  vl=mark.x, # líneas Verticales
                  layout=c(2,3) #Layout
){
  bwplot(Prog~Nota|Periodo, data =d,
         main=list(label=tit, cex=cex),
         xlab=list(xl, cex=cex),
         ylab=list('Programa académico',cex=cex),
         scales=list(x=list(cex=cex,at=mark.x, lim=lim.x), y=list(cex=cex)),
         par.strip.text=list(cex=cex,font=2),
         par.settings = list(
           box.rectangle = list(fill='green', col='black',lwd=2),
           box.umbrella = list(col='black',lwd=2),
           box.dot = list( col='black',lwd=2),
           plot.symbol=list(col='black',pch=1,cex=0.6*cex)
         ),
         pch='|', #Marca de la mediana
         layout = layout,
         as.table=TRUE, # Orden como tabla
         #index.cond=list(c(5,3,4,1,2)),#this provides the order of the panels
         panel = function(x,y,...) {

```

```

#panel.grid(h=0, v = -1)
panel.abline(h=1:16,v=v1,lty=2, col='lightgray', lwd=0.3)
panel.bwplot(x,y,...)
m=tapply(x, y, FUN = mean, na.rm = TRUE) # Media de cada carrera
s=tapply(x, y, FUN = sd, na.rm = TRUE) # Desviación estándar de cada carrera
s=s/m
#s=sprintf('%0.0f%%',s)
#et=sprintf('%0.0f,%0.0f',m,s)

n=length(s)
for (k in 1:n){
  if (k==pos.lab)
    et=substitute(bold( group("(",list(z,y),")") == group("(",list(bar(x),c[v]),
")")),list(z=round(m[[k]],1), y=round( s[[k]],2) ))
  else
    et=substitute(bold( group("(",list(bold(z),y),")")),list(z=round(m[[k]],1),
y=round(s[[k]],2) ))
  ltext(adj=0,x=lim.x[1]+delta.x/10, y=k, labels=et, cex=0.7*cex, font=2, col='blue')
  #print(et)
}
#ltext(adj=0,x=21, y=seq_along(m), labels=et, cex=cex-1, font=2, col='blue')
panel.points(x=m, y=seq_along(m), pch = '*', col='red', cex=1.8*cex) # Media de cada carrera
panel.abline(v=mean(x, na.rm = TRUE),lty =2, col='red', lwd=2) # Media de grupo
)
}

#Gráfica de Barras generica (Cantidades)
fig.barras=function(tbl,cex=3, tit=titulo, xl=xlabel, col=colores, col2=colores2, xc=50, yc=15, xmin=10){
  barchart(tbl,
    main=list(label=tit, cex=cex),
    xlab=list(xl, cex=cex),
    ylab=list('Programa académico', cex=cex),
    scales=list(x=list(cex=cex), y=list(cex=cex)),
    par.strip.text=list(cex=cex, font=2),
    par.settings=simpleTheme(col=col, border='black'), # Color y borde de barras
    auto.key = list(columns=dim(tbl)[3],font=2, rectangles = TRUE, points=FALSE, cex=cex),
    stack = TRUE,
    layout = c(2, 3),
    as.table=TRUE,
    #index.cond=list(c(5,6,3,4,1,2)),#this provides the order of the panels
    panel=function(x, y, subscripts, groups, stack) {
      panel.grid(h = 0, v = -1)
      panel.barchart(x=x, y=y,subscripts= subscripts, groups =groups, stack=stack)
      g=groups[subscripts]
      total=ave(x, y, FUN=cumsum)
      m=ave(total, y, FUN=max)
      tg=aggregate(x~groups[subscripts], FUN=sum)
      names(tg)=c('g','x')
      tg$g=100*tg$x/sum(tg$x)
      xx=total-x/2
      et=sprintf('%0.0f(%0.0f%%)',x,100*x/m)
      lab=ifelse(x<xmin, ifelse(x<2,'',x),et)
      ltext(xx, y=y, labels=lab, cex=0.8*cex, font=2)
      id=(m==total)
      ltext(-1,adj=1, y=y[id], labels=total[id], cex=0.9*cex, font=2, col='blue')
      et=sprintf('%0.0f Estudiantes',sum(total[id]))
      xp=xc
      yp=yc
      ltext(xp, yp, adj=0, labels=et, cex=cex, font=2, col='blue')
      nc=dim(tg)[1] # Número de categorías
      for (k in 1:nc){
        yp=yp-0.7
        ltext(xp, yp, adj=0, labels=sprintf('%0.0f %s (%0.0f%%)',tg$x[k],tg$g[k],tg$g[k]),
cex=cex, font=2, col=col2[k])
      }
    }
)
}

fig.barras2=function(tbl,cex=3, tit=titulo, xl=xlabel, col=colores,
  col2=colores2, xc=65, yc=15, dy=0.7, xmin=10,
  colum=dim(tbl)[2], layout=c(1,1)){
  barchart(tbl,
    main=list(label=tit, cex=cex),
    xlab=list(xl, cex=cex),
    ylab=list('Programa académico', cex=cex),
    scales=list(x=list(cex=cex), y=list(cex=cex)),
    par.strip.text=list(cex=cex, font=2),
    par.settings=simpleTheme(col=col, border='black'), # Color y borde de barras
    auto.key = list(columns=colum,font=2, rectangles = TRUE, points=FALSE, cex=cex),

```

```

stack = TRUE,
layout = layout,
as.table=TRUE,
#index.cond=list(c(5,6,3,4,1,2)),#this provides the order of the panels
panel=function(x, y, subscripts, groups, stack) {
  panel.grid(h = 0, v = -1)
  panel.barchart(x=x, y=y,subscripts= subscripts, groups =groups, stack=stack)
  g=groups[subscripts]
  total=ave(x, y, FUN=cumsum)
  m=ave(total, y, FUN=max)
  tg=aggregate(x~groups[subscripts], FUN=sum)
  names(tg)=c('g','x')
  tg$p=100*tg$x/sum(tg$x)
  xx=total-x/2
  et=sprintf('%0.0f(%0.0f%%)',x,100*x/m)
  lab=ifelse(x<xmin , ifelse(x<0.2*xmin, '',x),et)

  lab=ifelse(x==m , ifelse(x<0.2*xmin, '',x),lab)

  ltext(xx, y=y, labels=lab, cex=0.8*cex, font=2)
  id=(m==total)
  ltext(-1,adj=1, y=y[id], labels=total[id], cex=0.9*cex, font=2, col='blue')
  et=sprintf('%0.0f Total',sum(x))
  xp=xc
  yp=yc
  ltext(xp, yp, adj=0, labels=et, cex=cex, font=2, col='blue')
  nc=dim(tg)[1] # Número de categorías
  for (k in 1:nc){
    yp=yp-dy
    et=sprintf('%0.0f %s (%0.0f%%)',tg$x[k],tg$g[k],tg$p[k])
    ltext(xp, yp, adj=0, labels=et, cex=cex, font=2, col=col2[k])
  }
}
)
}

#Gráfica de Barras generica (Cantidades)
fig.barras.prop=function(tbl,cex=3, tit=titulo, xl=xlabel, col=colores, xmin=2){
  barchart(tbl,
    main=list(label=tit, cex=cex),
    xlab=list(xl, cex=cex),
    ylab=list('Programa académico', cex=cex),
    scales=list(x=list(cex=cex), y=list(cex=cex)),
    par.strip.text=list(cex=cex, font=2),
    par.settings=simpleTheme(col=col, border='black'), # Color y borde de barras
    auto.key = list(columns=dim(tbl)[3],font=2, rectangles = TRUE, points=FALSE, cex=cex),
    stack = TRUE,
    layout = c(2, 3),
    as.table=TRUE,
    #index.cond=list(c(5,6,3,4,1,2)),#this provides the order of the panels
    panel=function(x, y, subscripts, groups, stack) {
      panel.grid(h = 0, v = -1)
      panel.barchart(x=x, y=y,subscripts= subscripts, groups =groups, stack=stack)
      g=groups[subscripts]
      total=ave(x, y, FUN=cumsum)
      m=ave(total, y, FUN=max)
      tg=aggregate(x~groups[subscripts], FUN=sum)
      names(tg)=c('g','x')
      tg$p=100*tg$x/sum(tg$x)
      xx=total-x/2
      et=sprintf('%0.0f',x)
      lab=ifelse(x<xmin, '',et)
      ltext(xx, y=y, labels=lab, cex=0.8*cex, font=2)
    }
  )
}

fig.barras.prop2=function(tbl,cex=3, tit=titulo, xl=xlabel, col=colores, xmin=2, colum=4, layout = c(1, 1)){
  barchart(tbl,
    main=list(label=tit, cex=cex),
    xlab=list(xl, cex=cex),
    ylab=list('Programa académico', cex=cex),
    scales=list(x=list(cex=cex), y=list(cex=cex)),
    par.strip.text=list(cex=cex, font=2),
    par.settings=simpleTheme(col=col, border='black'), # Color y borde de barras
    auto.key = list(columns=colum,font=2, rectangles = TRUE, points=FALSE, cex=cex),
    stack = TRUE,
    layout = layout,
    as.table=TRUE,
    #index.cond=list(c(5,6,3,4,1,2)),#this provides the order of the panels

```

```

panel=function(x, y, subscripts, groups, stack) {
  panel.grid(h = 0, v = -1)
  panel.barchart(x=x, y=y, subscripts= subscripts, groups =groups, stack=stack)
  g=groups[subscripts]
  total=ave(x, y, FUN=cumsum)
  m=ave(total, y, FUN=max)
  tg=aggregate(x~groups[subscripts], FUN=sum)
  names(tg)=c('g', 'x')
  tg$P=100*tg$x/sum(tg$x)
  xx=total-x/2
  et=sprintf('%0.0f', x)
  lab=ifelse(x<xmin, '', et)
  ltext(xx, y=y, labels=lab, cex=0.8*cex, font=2)
}
)
}

# Cambia los parámetros graficos de lattice permanentemente (preferible no usar)
cambiar.parametros.lattice=function(cex=3){
  #Tamaño de parámetros graficos para el dispositivo actual
  trellis.par.set(par.main.text = list(cex = 1.2*cex))
  trellis.par.set(par.xlab.text = list(cex = cex))
  trellis.par.set(par.ylab.text = list(cex = cex))
  trellis.par.set(axis.text = list(cex = cex))
  trellis.par.set(plot.symbol = list(cex = 0.5*cex))
  trellis.par.set(add.text = list(cex = cex))

  #Tipo de letra font= 2 es negrilla
  trellis.par.set(add.text = list(font = 2))

  #Ancho de linea
}

primer.nota=function(materia){
  n=names(Est.2014.1.CeI)
  notas=subset(Est.2014.1.CeI, select=grepl(paste(materia, '_', sep = ''), n))
  nm=notas[,1]
  for (k in 2:dim(notas)[2]){
    pos.na=is.na(nm)
    nm[pos.na]=notas[pos.na,k]
  }
  nm
}

#names(Est.2014.1.CeI)

primer.nota.con.periodo=function(materia){
  materias=c('CalculoI', 'Algebra', 'CalculoII', 'CalculoIII', 'Ecuaciones')
  cod_materia=c(20252, 22979, 20253, 20254, 20255)[materia==materias]

  a=c(2014, 2014, 2015, 2015, 2016, 2016)
  p=c(1, 2, 1, 2, 1, 2)
  s=1:6
  n=names(Est.2014.1.CeI)
  pos.materia=grepl(paste(materia, '_', sep = ''), n)
  pos.materia[1]=TRUE
  notas=subset(Est.2014.1.CeI, select=pos.materia)
  nm=data.frame(Codigo=notas[,1], Nota=notas[,2], Ano=a[2-1], Periodo=p[2-1], Semestre=s[2-1])
  for (k in 3:dim(notas)[2]){
    pos.na=is.na(nm$Nota)
    nm[pos.na,]=data.frame(Codigo=notas[pos.na,1], Nota=notas[pos.na,k], Ano=a[k-1], Periodo=p[k-1], Semestre=s[k-1])
  }

  # Agrega Beneficio SEA en asignatura
  ben=subset(Ben.SEA.2014.1.CeI,
             Codigo_asignatura == cod_materia & igual.prog.notas & Horas>0,
             select = c('Codigo', 'Ano_SEA', 'Periodo_SEA', 'Prog_SEA', 'Sub_Programa_SEA', 'Horas'))

  # Agrega el semestre
  ben$Semestre=ben$Ano_SEA-2014+ben$Periodo_SEA

  # Agrega beneficios acumulados
  nm$Horas=0
  nm$Num.Semestres.Ben=0
  for (k in 1:dim(nm)[1]){
    n=nm[k,]
    benk=ben[ben$Codigo==n$Codigo & ben$Semestre<=n$Semestre,]

```

```

nm$Horas[k]=sum(benk$Horas)
if (nm$Horas[k]>0){
  nm$Num.Semestres.Ben[k]=length( aggregate(Horas~Semestre, data=benk, FUN=sum)$Horas )
}
}
nm$Beneficio=ifelse(nm$Num.Semestres.Ben==0,0,1)
#Cambia los nombres
names(nm)=paste(names(nm),materia, sep = ".")
nm
}

lista.notas=function(materia){
materias=c('CalculoI','Algebra','CalculoII','CalculoIII','Ecuaciones')
cod_materia=c(20252,22979,20253,20254,20255)[materia==materias]
s=1:6
pos.materia=grep(paste(materia,'_', sep = ''), names(Est.2014.1.CeI))
pos.materia[1]=TRUE # Para incluir el código

# Notas de la materia escogida
notas=subset(Est.2014.1.CeI, select=pos.materia)

m=dim(notas)[1]
n=dim(notas)[2]

notas$Nota1=NA
notas$Nota2=NA
notas$Nota3=NA

for (k in 2:n){
  pos=is.na(notas$Nota1) & !is.na(notas[,k])
  notas$Nota1[pos]=notas[pos,k]
  notas$Semestre1[pos]=k-1
  if (k<n){
    pos.rep= notas$Nota1<3 & !is.na(notas$Nota1) & !is.na(notas[,k+1])
    notas$Nota3[pos.rep]=notas[pos.rep,k+1]
    notas$Semestre3[pos.rep]=k
    pos.rep2=pos.rep & is.na(notas$Nota2)
    notas$Nota2[pos.rep2]=notas[pos.rep2,k+1]
    notas$Semestre2[pos.rep2]=k
  }
}

# Agrega Beneficio SEA para la asignatura
ben=subset(Ben.SEA.2014.1.CeI,
  Codigo_asignatura == cod_materia & igual.prog.notas & Horas>0,
  select = c('Codigo','Ano_SEA','Periodo_SEA','Prog_SEA','Sub_Programa_SEA','Horas'))

ben$Semestre=ben$Ano_SEA-2014+ben$Periodo_SEA

notas$Horas1=NA
notas$Horas2=NA
notas$Horas3=NA

notas$Num.Sem.Ben1=NA
notas$Num.Sem.Ben2=NA
notas$Num.Sem.Ben3=NA

# Agrega beneficios recibidos del SEA en la materia
for (k in 1:m){
  ben1=ben[ben$Codigo==notas[k,'Codigo'] & ben$Semestre<=notas[k,'Semestre1'],]
  s=sum(ben1$Horas, na.rm = TRUE)
  if (s>0) {
    notas$Horas1[k]=s
    notas$Num.Sem.Ben1[k]=length(aggregate(Horas~Semestre, data=ben1, FUN=sum)$Horas)
  }
  if (!is.na(notas[k,'Semestre2'])){
    ben2=ben[ben$Codigo==notas[k,'Codigo'] & ben$Semestre<=notas[k,'Semestre2'],]
    s=sum(ben2$Horas, na.rm = TRUE)
    if (s>0){
      notas$Horas2[k]=s
      notas$Num.Sem.Ben2[k]=length(aggregate(Horas~Semestre, data=ben2, FUN=sum)$Horas)
    }
  }
  if (!is.na(notas[k,'Semestre3'])){
    ben3=ben[ben$Codigo==notas[k,'Codigo'] & ben$Semestre<=notas[k,'Semestre3'],]
    s=sum(ben3$Horas, na.rm = TRUE)
    if (s>0){
      notas$Horas3[k]=s
      notas$Num.Sem.Ben3[k]=length(aggregate(Horas~Semestre, data=ben3, FUN=sum)$Horas)
    }
  }
}

```

```

    }
}

#Ajusta la nota 3
pos=is.na(notas$Nota3) & !is.na(notas$Nota1)
notas$Nota3[pos]=notas$Nota1[pos]

#Ajusta los beneficios acumulados
notas$Horas1[is.na(notas$Horas1)]=0

pos=is.na(notas$Horas2) & !is.na(notas$Horas1)
notas$Horas2[pos]=notas$Horas1[pos]
notas$Horas2[is.na(notas$Horas2)]=0

pos=is.na(notas$Horas3) & !is.na(notas$Horas2)
notas$Horas3[pos]=notas$Horas2[pos]
notas$Horas3[is.na(notas$Horas3)]=0

# Dummy si recibio o no beneficio
notas$Ben1[!is.na(notas$Nota1)]=0
notas$Ben2[!is.na(notas$Nota2)]=0
notas$Ben3[!is.na(notas$Nota3)]=0

notas$Ben1[notas$Num.Sem.Ben1>0]=1
notas$Ben2[notas$Num.Sem.Ben2>0]=1
notas$Ben3[notas$Num.Sem.Ben3>0]=1

# Agrega el nombre de la materia
names(notas)=paste(names(notas),materia, sep = ".")
notas
}

#Histograma bivariado
#Adaptado del foro en: https://www.r-bloggers.com/5-ways-to-do-2d-histograms-in-r/
hist2=function(x,y,xlab='x', ylab='y', n=25){

  # Color housekeeping
  rf = colorRampPalette(rev(brewer.pal(11,'Spectral'))))
  rr = rf(32)
  #rr=rev( gray.colors(32))

  #Histogramas
  hx <- hist(x, breaks=n, plot=F)
  hy <- hist(y, breaks=n, plot=F)
  top <- max(hx$counts, hy$counts)

  k <- kde2d(x, y, n=n)

  # margins
  oldpar = par()
  par(mar=c(2,2,1,1))
  layout(matrix(c(2,0,1,3),2,2,byrow=T),c(3,1), c(1,3))
  image(k, col=rr) #plot the image
  #plot(x,y, pch=16, col='black', cex=0.5)
  par(mar=c(0,1,2,0))
  barplot(hx$counts, axes=F, ylim=c(0, top), space=0, col='red')
  mtext(xlab, side = 1)
  par(mar=c(1,0,0.5,0))
  barplot(hy$counts, axes=F, xlim=c(0, top), space=0, col='red', horiz=T)
  mtext(ylab,side = 2)
  par=oldpar
}

#Paquetes usados
# install.packages("readxl")           # Importar datos de MS Excel a R (solo la primera vez)
library(readxl)
#install.packages('plyr')              # Agregación de datos y aplicación de funciones por grupos.
library('plyr')
#install.packages('dplyr')             # It's the next iteration of plyr
library(dplyr)
#install.packages('reshape2')         # Realizar transformaciones a los datos.(comando melt y dcast). Apila los
# resultados
library('reshape2')
#install.packages('lattice')          # Paquete para gráficas sofisticadas en R
library(lattice)
#install.packages('latticeExtra')     # Paquete para gráficas sofisticadas en R que extiende lattice
library(latticeExtra)
#ninstall.packages('XLConnect')       # Envía o exporta tabla a MS Excel

```

```

library(XLConnect)
#install.packages('abind')           # Combine multi-dimensional arrays
library(abind)
#install.packages('nortest')         # 5 pruebas para normalidad de datos
library(nortest)
#install.packages('optmatch')       # Functions for Optimal Matching
library(optmatch)
#install.packages('MatchIt')        # Para realizar pareamiento
library(MatchIt)
#install.packages('tableone')       # Tablas vistasas
library(tableone)
#install.packages("knitr")          # Tablas vistasas
library(knitr)
#install.packages("RColorBrewer")   # Colores empleado en histogramas bibariados
library(RColorBrewer)
#install.packages("MASS")           # comando kde2d empleado en histogramas bibariados
library(MASS)

#Funciones no propias

# Uso de R en el libro: Design of Observational Studies. Paul R. Rosenbaum. Springer 2010. Chapter 13: Matching
in R (Pgs 237-252)

mahal=function(z,X){
  X<-as.matrix(X)
  n<-dim(X)[1]
  rownames(X)<-1:n
  k<-dim(X)[2]
  m<-sum(z)
  cv<-cov(X)
  out<-matrix(NA,m,n-m)
  Xc<-X[z==0,]
  Xt<-X[z==1,]
  rownames(out)<-rownames(X)[z==1]
  colnames(out)<-rownames(X)[z==0]
  library(MASS)
  icov<-ginv(cv)
  for (i in 1:m) {
    out[i,]<-mahalanobis(Xc,Xt[i,],icov,inverted=T)
  }
  out
}

smahal=function(z,X){
  # Rank based Mahalanobis distance.
  X<-as.matrix(X)
  n<-dim(X)[1]
  rownames(X)<-1:n
  k<-dim(X)[2]
  m<-sum(z)
  for (j in 1:k) X[,j]<-rank(X[,j])
  cv<-cov(X)
  vuntied<-var(1:n)
  rat<-sqrt(vuntied/diag(cv))
  cv<-diag(rat)*%*cv*%*diag(rat)
  out<-matrix(NA,m,n-m)
  Xc<-X[z==0,]
  Xt<-X[z==1,]
  rownames(out)<-rownames(X)[z==1]
  colnames(out)<-rownames(X)[z==0]
  library(MASS)
  icov<-ginv(cv)
  for (i in 1:m) {
    out[i,]<-mahalanobis(Xc,Xt[i,],icov,inverted=T)
  }
  out
}

addcaliper=function(dmat,z,p,caliper=0.2,penalty=1000){
  sdp<-sd(p)
  adif<-abs(outer(p[z==1],p[z==0],"-"))
  adif<-(adif-(caliper*sdp))*adif+(caliper*sdp)
  dmat<-dmat+adif*penalty
  dmat
}

addalmostexact=function(dmat,z,f,mult=10){
  penalty<-mult*max(dmat)

```

```
mismatch<-outer(f[z==1],f[z==0],"!=")  
dmat<-dmat+mismatch*penalty  
dmat  
}
```

## Apéndice C. Evaluación de impacto por pareamiento en R

```

# Evaluación de impacto en cohorte 2014-1. Cálculo I

# Diferencias entre beneficio del semestre y acumulado (no hay diferencias)
table(with(Est.2014.1.CeI,Horas.Acum.CalculoI.1-Horas.Total.CalculoI.1))

# Diferencias con los datos anteriores
td=table(with(Est.2014.1.CeI,Horas.CalculoI-Horas.Total.CalculoI.1))
td

# Estudiantes con diferencias mayores que cero = 13
sum(td[names(td) != 0])

# Datos a considerar
materia='Cálculo I'
periodo='2014-1'
Est.2014.1.CeI$Horas=Est.2014.1.CeI$Horas.Total.CalculoI.1
Est.2014.1.CeI$Nota=Est.2014.1.CeI$Nota.CalculoI.1

# Cambia los NA a cero
Est.2014.1.CeI$Horas[is.na(Est.2014.1.CeI$Horas)]=0

# Halla los cortes
cortes=cortes.de.8(Est.2014.1.CeI$Horas)
tbl=table(cut(Est.2014.1.CeI$Horas,breaks = cortes))
tbl

# Estudiantes que tomaron más horas que el límite
n=length(tbl)
tbl.ben=cumsum(tbl[n:1])[n:1]
tbl.ben

# Distribución de horas de beneficio
barchart(tbl[-1],
  main=sprintf('Beneficios del SEA para %s. Cohorte %s',materia,periodo),
  xlab='Número de estudiantes',
  ylab='Horas Beneficio',
  panel=function(x, y) {
    panel.barchart(x=x, y=y)
    n=length(x)
    total=cumsum(x[n:1])[n:1]
    ltext(x=x/2, y=y, labels=x, font=2)
    ltext(x=x, y=y,adj=0, labels=total)
  })

# Filtra datos necesarios para evaluación de impacto
Datos.Evaluacion=
  subset(Est.2014.1.CeI,
    select = c(Nota,Horas,
              Prog,Edad,Genero,
              Matematica,Lenguaje,Ingles,Fisica,Quimica,Biologia,
              Riesgo_Economica,Riesgo_Social,Riesgo_Salud,
              Riesgo_Academico,Riesgo_Cognitiva))

# Elimina registros incompletos
pos=complete.cases(Datos.Evaluacion) # Posición de registros completos

# Cantidad de registros completos
sum(pos)
100*sum(pos)/length(pos)

# cantidad de Incompletos
sum(!pos)
100*sum(!pos)/length(pos)

# Toma solo los estudiantes con información completa
Datos.Evaluacion=Datos.Evaluacion[pos,]
dim(Datos.Evaluacion)

# Borra los niveles no usados del programa
Datos.Evaluacion$Prog= droplevels(Datos.Evaluacion$Prog)

```

```
# Tabla de beneficiados y no beneficiados completos
#debug(cortes.de.8)
#undebug(cortes.de.8)
cortes=cortes.de.8(Datos.Evaluacion$Horas)
tbl=table(cut(Datos.Evaluacion$Horas,breaks = cortes))
tbl

# Tabla de beneficios acumulada
n=length(tbl)
tbl.ben=cumsum(tbl[n:1])[n:1]
tbl.ben

# Muestra las dos tablas combinadas
tbl2=rbind(tbl, tbl.ben)
row.names(tbl2)=c('# Beneficiados', 'Total')
kable(tbl2,digits = 0, align = 'c')

# Distribución de beneficiarios finales con datos completos
barchart(tbl[-1],
  main=sprintf('Beneficios del SEA para %s. Cohorte %s',materia, periodo),
  xlab='Número de estudiantes',
  ylab='Horas Beneficio',
  panel=function(x, y) {
    panel.barchart(x=x, y=y)
    n=length(x)
    total=cumsum(x[n:1])[n:1]
    ltext(x=x/2, y=y, labels=x, font=2)
    ltext(x=x+0.2, y=y,adj=0, labels=total, col='red')
    ltext(x=max(x),y=4,adj=1, labels='Acumulado', col='red')
  })

# Realiza un Numero de Corridas (NC) de sensibilidad y guarda en data.frame 'd'
Run.File('200-Escenarios_Pareamiento.R')

# Grafica Nota promedio en función de las horas de beneficio
Run.File('200-Impacto_Vs_Horas_Beneficio.R')
```

## Apéndice D. Evaluación de impacto por diferencias en diferencias en R

```
#####
###      Técnica de diferencias en diferencias para evaluar el impacto del programa      #####
#####
#####      SEA sobre los repitentes Cálculo I      #####
#####

# Datos a considerar
materia='Cálculo I'
periodo='2014-1'

#Primera vez que ve la materia
Est.2014.1.CeI$Horas.1=Est.2014.1.CeI$Horas.Total.CalculoI.1
Est.2014.1.CeI$Nota.1=Est.2014.1.CeI$Nota.CalculoI.1

#Segunda vez que ve la materia
Est.2014.1.CeI$Horas.2=Est.2014.1.CeI$Horas.Total.CalculoI.2
Est.2014.1.CeI$Nota.2=Est.2014.1.CeI$Nota.CalculoI.2

# Cambia los NA a cero
Est.2014.1.CeI$Horas.1[is.na(Est.2014.1.CeI$Horas.1)]=0
Est.2014.1.CeI$Horas.2[is.na(Est.2014.1.CeI$Horas.2)]=0

# Filtra datos necesarios para evaluación de impacto
Datos.Evaluacion=
  subset(Est.2014.1.CeI,
    select = c(Codigo,Nota.1,Nota.2,Horas.1,Horas.2,
              Prog,Edad,Genero,
              Matematica,Lenguaje,Ingles,Fisica,Quimica,Biologia,
              Riesgo_Economica,Riesgo_Social,Riesgo_Salud,
              Riesgo_Academico,Riesgo_Cognitiva),
    subset = (Nota.1<3 & !is.na(Nota.2)) # Solo repitentes
  )

# Horas de beneficio en los dos semestres
Datos.Evaluacion$Horas.12=with(Datos.Evaluacion,Horas.1 + Horas.2)

# Cambio de la nota
Datos.Evaluacion$Cambio=with(Datos.Evaluacion,Nota.2 - Nota.1)

# Conviente en categorico Beneficiados y no Beneficiados
#Beneficiario el que recibio al menos una hora de beneficio los dos semestres que vió la materia
Datos.Evaluacion$Ben=ifelse(Datos.Evaluacion$Horas.1>0 & Datos.Evaluacion$Horas.2>0,1,0)

# No beneficiario el que no recibio beneficio ningun semestre en el que vió la materia
Datos.Evaluacion$NoBen=ifelse(Datos.Evaluacion$Horas.1==0 & Datos.Evaluacion$Horas.2==0,1,0)

#Porcentaje de datos validos para diferencias en diferencias
pos=with (Datos.Evaluacion, Ben==1 | NoBen==1)
sum(pos)/length(pos)

# Filtra los datos de interés
Datos.Evaluacion2=Datos.Evaluacion[pos,]

# Compara perdida de datos
dim(Datos.Evaluacion)
dim(Datos.Evaluacion2)

# Tabla de beneficiados y no beneficiados completos
cortes=cortes.de.8(Datos.Evaluacion2$Horas.2, fusionar = FALSE)
tbl=table(cut(Datos.Evaluacion2$Horas.2,breaks = cortes))
tbl

# Tabla de beneficios acumulada
n=length(tbl)
tbl.ben=cumsum(tbl[n:1])[n:1]
tbl.ben

# Muestra las dos tablas combinadas
tbl2=rbind(tbl, tbl.ben)
row.names(tbl2)=c('# Beneficiados', 'Total')
kable(tbl2,digits = 0, align = 'c')
```

```

# Distribución de beneficiarios finales de repitentes
barchart(tbl[-1],
  main=sprintf('Beneficios del SEA para repitentes en %s. Cohorte %s',materia, periodo),
  xlab='Número de estudiantes',
  ylab='Horas Beneficio',
  panel=function(x, y) {
    panel.barchart(x=x, y=y)
    n=length(x)
    total=cumsum(x[n:1])[n:1]
    ltext(x=x/2, y=y, labels=x, font=2)
    ltext(x=x+0.1, y=y,adj=0, labels=total, col='red')
    ltext(x=max(x),y=3,adj=1, labels='Acumulado', col='red')
  })

# Modelo de Cambio de Nota Vs Beneficio primera y segunda vez
summary(lm(Cambio~Ben,data=Datos.Evaluacion2))
boxplot(Cambio~Ben,data=Datos.Evaluacion2, xaxt = "n",
  main=sprintf('Diagrama de cajas para cambio de nota de %s. Cohorte %s',materia, periodo),
  xlab='', ylab='Cambio de nota promedio')
tbl=with(Datos.Evaluacion2,table(Ben))
axis(1, at=1:2, labels=c(sprintf('Sin Beneficio, n=%i',tbl[[1]]),sprintf('Con Beneficio, n=%i',tbl[[2]])))#
Personaliza marcas de eje x

# Prueba t pareada
with(Datos.Evaluacion2,t.test(Nota.2,Nota.1,paired = TRUE)) # Prueba t pareada global
with(Datos.Evaluacion2,t.test(Cambio)) # Prueba t pareada (equivalente)

# Prueba de la comparación de las dos. El impacto es la diferencia de los dos cambios
t.test(Cambio~Ben,data=Datos.Evaluacion2)

# Apila los datos para construir modelo de diferencias
datos2=reshape(Datos.Evaluacion2,direction = 'long',
  idvar = c('Codigo','Ben'),
  #drop = c('Cambio'),
  varying = c(2,3),v.names='Nota',
  times=c(1,2),timevar='Periodo',
  new.row.names=seq( 2*dim(Datos.Evaluacion2)[1]))

#View(datos2)

# Modelo para estimar el impacto usando diferencias en diferencias
modelo=lm(Nota~Ben*Periodo, data = datos2)
s=summary( modelo)

#Estimación de impacto Inteacción entre Ben:Periodo en posición 4
s$coefficients[4,]

# Plots the mean (or other summary) of the response for two-way combinations of factors,
# thereby illustrating possible interactions.
# Two-way Interaction Plot

# Cambia etiquetas para construcción de gráficas
datos2$Periodo=factor(datos2$Periodo,labels = c('Primera vez','Segunda vez'))
datos2$Ben=factor(datos2$Ben,labels = c('Sin Beneficio', 'Con Beneficio'))

tbl=aggregate(Nota~Periodo+Ben, data= datos2, FUN = mean)

with(datos2,
  interaction.plot(Periodo,Ben,Nota,
    main= sprintf('Diferencias en diferencia para estudiantes repitentes de %s. Cohorte
%s',materia, periodo),
    xlab = sprintf('Experiencia en %s',materia),
    ylab = sprintf('Nota promedio de %s ',materia),
    type="l",
    col=c("red","green"), ## Colors for levels of trace var.
    pch=c(19, 17), ## Symbols for levels of trace var.
    fixed=TRUE, ## Order by factor order in data
    trace.label='Beneficio SEA')
)
abline(h = seq(2.1,2.5,0.1), v = c(1,2), col = "lightgray", lty = 3)
text(x=tbl$Periodo,y=tbl$Nota,labels=sprintf('%0.3f',tbl$Nota),
  #col=c("red","red","green","green"),
  cex=0.8,font = 2)

# Estimador del Impacto de dos formas equivalentes
s$coefficients[4,]
(tbl$Nota[[4]]-tbl$Nota[[3]])-(tbl$Nota[[2]]-tbl$Nota[[1]])

```

```
(tbl$Nota[[4]]-tbl$Nota[[3]])
(tbl$Nota[[2]]-tbl$Nota[[1]])

# Segunda forma opcional

# with(datos2,
#       interaction.plot(Ben,Periodo,Nota,
#                         main= sprintf('Diferencias en diferencia para estudiantes repitentes de %s. Cohorte
#s',materia, periodo),
#                         xlab = '',
#                         type="l",
#                         col=c("red","green"),      ### Colors for levels of trace var.
#                         pch=c(19, 17),           ### Symbols for levels of trace var.
#                         fixed=TRUE,             ### Order by factor order in data
#                         ylab = sprintf('Nota promedio de %s ',materia),
#                         trace.label='Primera vez = 1')
# )
#
#
#
# abline(h = seq(2.1,2.5,0.1), v = c(1,2), col = "lightgray", lty = 3)
# text(x=tbl$Ben,y=tbl$Nota,labels=sprintf('%0.1f',tbl$Nota),
#       #col=c("red","red","green","green"),
#       cex=0.8,font = 2)

# Interaction Plot in ggplot2
#https://stackoverflow.com/questions/7323246/interaction-plot-in-ggplot2
```