



**ESTUDIO ESTADÍSTICO RETROSPECTIVO
DEL PROGRAMA DE ACOMPAÑAMIENTO
INTEGRAL DE LA UNIVERSIDAD
TECNOLÓGICA DE PEREIRA**

**Vicerrectoría de Responsabilidad Social y Bienestar
Universitario**

VICERRECTORÍA DE RESPONSABILIDAD SOCIAL Y BIENESTAR UNIVERSITARIO

Diana Gómez Botero

Vicerrectora de Responsabilidad Social y Bienestar Universitario

Ana Victoria Ospina

Coordinadora de Gestión Social

Genaro Daza Santacoloma

Angélica Rojas Muñoz

Laboratorio de Innovación y Desarrollo Sostenible

Con el acompañamiento de:

Álvaro Orozco Gutierrez

Profesor UTP

José Soto Mejía

Profesor UTP

Diciembre de 2017

Agradecimientos

De manera especial queremos agradecer el apoyo en el análisis, cálculos y evaluación de resultados del Ph.D. Andrés Álvarez Meza. También queremos agradecer de manera especial a los siguientes integrantes de la comunidad universitaria que amablemente nos brindaron acceso a la información disponible: jefe de la oficina de planeación Francisco Antonio Uribe Gómez, jefe de admisiones, registro y control académico Yetsika Natalia Villa Montes, jefe de gestión de tecnologías informáticas y sistemas de información Diana Patricia Jurado Ramírez, e igualmente a los profesionales Delany Ramírez, Iván Alexander Laverde Gaviria, Eliana Prado Brand, Angélica López Gómez y Julián Alejandro Piedrahita Monroy.

Índice general

1. Introducción	9
2. Análisis descriptivo de los datos muestrales	11
2.1. Descripción de los datos	11
2.1.1. Datos de registro y matrícula	11
2.1.2. Datos de atenciones del PAI	13
2.2. Descripción del conjunto poblacional	14
2.2.1. Población por género	14
2.2.2. Población por rango de edad	15
2.2.3. Población por programa académico	17
2.2.4. Población por facultad	20
2.2.5. Población por jornada	23
2.2.6. Población por naturaleza del colegio	25
2.2.7. Población por departamento	26
2.2.8. Población por capital	29
2.2.9. Población por estrato	30
2.2.10. Población por semestre de ubicación	31
2.3. Descripción del subconjunto poblacional PAI	33
2.3.1. Subconjunto PAI por género	33
2.3.2. Subconjunto PAI por rango de edad	34
2.3.3. Subconjunto PAI por programa académico	36
2.3.4. Subconjunto PAI por facultad	38
2.3.5. Subconjunto PAI por jornada	41
2.3.6. Subconjunto PAI por naturaleza del colegio	43
2.3.7. Subconjunto PAI por departamento	44
2.3.8. Subconjunto PAI por capital	46
2.3.9. Subconjunto PAI por estrato	47
2.3.10. Subconjunto PAI por semestre de ubicación	48
2.3.11. Comparación variables descriptivas	49
2.4. Deserción universitaria	50
2.4.1. Deserción de un período	50
2.4.2. Deserción de dos períodos	52
2.4.3. Deserción de un período en el subconjunto PAI	53

2.4.4.	Deserción de dos períodos en el subconjunto PAI	53
2.5.	Evaluación de probabilidades condicionales	54
2.5.1.	Análisis para el subconjunto poblacional: PAI	54
2.5.2.	Análisis para el subconjunto poblacional: Desertor un período	63
2.5.3.	Análisis para el subconjunto poblacional: Desertor dos períodos	72
3.	Análisis de relevancia de las características de los grupos objetivo	82
3.1.	Matriz de correlación	82
3.2.	Análisis de componentes principales	87
3.3.	Alineación centrada de kernel	90
4.	Evaluación de la relación costo-beneficio del PAI	99
4.1.	Simulación: Eliminación de atenciones del PAI	99
4.2.	Evaluación financiera del PAI	103
5.	Conclusiones y recomendaciones	105
5.1.	Conclusiones	105
5.2.	Recomendaciones	106
A.	Codificación de variables y categorías	108
A.1.	Codificación de la variable género	108
A.2.	Codificación de la variable edad	108
A.3.	Codificación de la variable programa académico	109
A.4.	Codificación de la variable facultad	112
A.5.	Codificación de la variable jornada	112
A.6.	Codificación de la variable naturaleza del colegio	112
A.7.	Codificación de la variable departamento	113
A.8.	Codificación de la variable capital	114
A.9.	Codificación de la variable atendido PAI	114
A.10.	Codificación de la variable atendido línea biopsicosocial	115
A.11.	Codificación de la variable atendido línea académica	115
A.12.	Codificación de la variable atendido línea socio-económica	115
A.13.	Codificación de la variable número de atenciones	116
A.14.	Codificación de la variable estrato	116
A.15.	Codificación de la variable semestre de ubicación	116
A.16.	Codificación de la variable créditos aprobados	116
A.17.	Codificación de la variable períodos matriculados	117
A.18.	Codificación de la variable apoyo económico	117
A.19.	Codificación de la variable promedio acumulado	117
A.20.	Codificación de la variable desertor de un período	117
A.21.	Codificación de la variable desertor de dos períodos	118
Bibliografía		119

Índice de figuras

2.1. Conformación de la población por género	15
2.2. Conformación de la población por rango de edad	17
2.3. Conformación de la población por programa académico	20
2.4. Conformación de la población por facultad	22
2.5. Conformación de la población por jornada	24
2.6. Conformación de la población por naturaleza del colegio	25
2.7. Conformación de la población por departamento	28
2.8. Conformación de la población por departamento excluyendo Risaralda .	28
2.9. Conformación de la población por capital	29
2.10. Conformación de la población por estrato	31
2.11. Conformación de la población por semestre de ubicación	32
2.12. Conformación del subconjunto PAI por género	34
2.13. Conformación del subconjunto PAI por rango de edad	36
2.14. Conformación del subconjunto PAI por programa académico	38
2.15. Conformación del subconjunto PAI por facultad	40
2.16. Conformación del subconjunto PAI por jornada	42
2.17. Conformación del subconjunto PAI por naturaleza del colegio	43
2.18. Conformación del subconjunto PAI por departamento	45
2.19. Conformación del subconjunto PAI por departamento excluyendo Risaralda	45
2.20. Conformación del subconjunto PAI por capital	46
2.21. Conformación del subconjunto PAI por estrato	47
2.22. Conformación del subconjunto PAI por semestre de ubicación	49
2.23. Conjuntos para evaluar permanencia y deserción de un período	51
2.24. Conjuntos para evaluar permanencia y deserción de dos períodos	52
3.1. Matriz de correlación para 2016-I	84
3.2. Matriz de correlación para 2016-II	86
3.3. Relevancia de variables según PCA para 2016-I	89
3.4. Relevancia de variables según CKA para 2016-I asociada a la deserción de 1 periodo	91
3.5. Relevancia de variables según CKA para 2016-I asociada a la deserción de 2 periodos	92

3.6. Relevancia de variables según CKA para 2016-II asociada a la deserción de 1 periodo	93
3.7. Clasificación automática de deserción empleando KNN para variables relevantes de CKA, 2016-I, deserción 1 periodo	95
3.8. Clasificación automática de deserción empleando KNN para variables relevantes de CKA, 2016-I, deserción 2 periodos	96
3.9. Clasificación automática de deserción empleando KNN para variables relevantes de CKA, 2016-II, deserción 1 periodo	97
4.1. Conformación de grupos de interés para análisis	100
4.2. Eliminación de atenciones PAI, 2016-I, deserción 1 periodo	101
4.3. Eliminación de atenciones PAI, 2016-I, deserción 2 periodos	102
4.4. Eliminación de atenciones PAI, 2016-II, deserción 1 periodo	102

Índice de tablas

2.1. Número de estudiantes matriculados por período	14
2.2. Conformación de la población por género	14
2.3. Conformación de la población por rango de edad	16
2.4. Conformación de la población por programa académico	19
2.5. Conformación de la población por facultad	21
2.6. Conformación de la población por jornada	23
2.7. Conformación de la población por naturaleza del colegio	25
2.8. Conformación de la población por departamento	27
2.9. Conformación de la población por capital	29
2.10. Conformación de la población por estrato	30
2.11. Conformación de la población por semestre de ubicación	32
2.12. Número atendidos en PAI por período	33
2.13. Conformación del subconjunto PAI por género	33
2.14. Conformación del subconjunto PAI por rango de edad	35
2.15. Conformación del subconjunto PAI por programa académico	37
2.16. Conformación del subconjunto PAI por facultad	39
2.17. Conformación del subconjunto PAI por jornada	41
2.18. Conformación del subconjunto PAI por naturaleza del colegio	43
2.19. Conformación del subconjunto PAI por departamento	44
2.20. Conformación del subconjunto PAI por capital	46
2.21. Conformación del subconjunto PAI por estrato	47
2.22. Conformación del subconjunto PAI por semestre de ubicación	48
2.23. Error de comparación de los histogramas de la población total y del subconjunto PAI	50
2.24. Permanencia y deserción de un período	52
2.25. Permanencia y deserción de un período para grupo PAI	53
2.26. Probabilidad condicional: $P(PAI Género)$	54
2.27. Probabilidad condicional: $P(PAI Rangodeedad)$	55
2.28. Probabilidad condicional: $P(PAI Programa)$	56
2.29. Probabilidad condicional: $P(PAI Facultad)$	57
2.30. Probabilidad condicional: $P(PAI Jornada)$	57
2.31. Probabilidad condicional: $P(PAI NaturalezaColegio)$	58
2.32. Probabilidad condicional: $P(PAI Departamento)$	59

2.33. Probabilidades condicionales $P(PAI Capital)$	60
2.34. Probabilidad condicional: $P(PAI Estrato)$	60
2.35. Probabilidad condicional: $P(PAI Semestre)$	61
2.36. Probabilidad condicional $P(PAI Periodos)$	61
2.37. Probabilidad condicional $P(PAI Creditos)$	62
2.38. Probabilidad condicional $P(PAI Promedio)$	63
2.39. Probabilidad condicional: $P(DES/NODES Genero)$	63
2.40. Probabilidad condicional: $P(DES/NODES RangoEdad)$	64
2.41. Probabilidad condicional: $P(DES/NODES Programa)$	65
2.42. Probabilidad condicional: $P(DES/NODES Facultad)$	66
2.43. Probabilidad condicional: $P(DES/NODES Jornada)$	66
2.44. Probabilidad condicional: $P(DES/NODES Naturaleza)$	67
2.45. Probabilidad condicional: $P(DES/NODES Departamento)$	68
2.46. Probabilidad condicional: $P(DES/NODES Capital)$	69
2.47. Probabilidad condicional: $P(DES/NODES Estrato)$	69
2.48. Probabilidad condicional: $P(DES/NODES Semestre)$	70
2.49. Probabilidad condicional: $P(DES/NODES Períodos)$	70
2.50. Probabilidad condicional: $P(DES/NODES Creditos)$	71
2.51. Probabilidad condicional: $P(DES/NODES Promedios)$	72
2.52. Probabilidad condicional: $P(DES/NODES Genero)$	72
2.53. Probabilidad condicional: $P(DES/NODES RangoEdad)$	73
2.54. Probabilidad condicional: $P(DES/NODES Programa)$	74
2.55. Probabilidad condicional: $P(DES/NODES Facultad)$	75
2.56. Probabilidad condicional: $P(DES/NODES Jornada)$	75
2.57. Probabilidad condicional: $P(DES/NODES NaturalezaColegio)$	76
2.58. Probabilidad condicional: $P(DES/NODES Departamento)$	77
2.59. Probabilidad condicional: $P(DES/NODES Capital)$	78
2.60. Probabilidad condicional: $P(DES/NODES Estrato)$	78
2.61. Probabilidad condicional: $P(DES/NODES Semestre)$	79
2.62. Probabilidad condicional: $P(DES/NODES PeriodosMatriculados)$	79
2.63. Probabilidad condicional: $P(DES/NODES CreditosAcumulados)$	80
2.64. Probabilidad condicional: $P(DES/NODES PromedioAcumulado)$	81
4.1. Resultados de deserción ante simulación de no atención PAI	103
4.2. Inversión de la UTP en el PAI para cada semestre de interés	104
4.3. Resumen de relación costo-beneficio del PAI	104
A.1. Codificación género	108
A.2. Codificación rango de edad	109
A.3. Codificación facultad	112
A.4. Codificación jornada	112
A.5. Codificación de la naturaleza del colegio	112
A.6. Codificación de capital	114

A.7. Codificación de atendido PAII	114
A.8. Codificación de atendido línea BPS	115
A.9. Codificación de atendido línea académica	115
A.10.Codificación de atendido línea socio-económica	115
A.11.Codificación de estrato	116
A.12.Codificación de apoyo económico	117
A.13.Codificación de desertor de un período	117
A.14.Codificación de desertor de dos período	118

Capítulo 1

Introducción

El Programa de Atención Integral (PAI) de la Universidad Tecnológica de Pereira (UTP) que desarrolla la Vicerrectoría de Responsabilidad Social y Bienestar Universitario se ocupa de aspectos biopsicosociales, económicos, académicos y normativos, en los que brinda apoyo a los estudiantes de la UTP para que así alcancen su egreso exitoso, otorgando mayor seguimiento y cobertura desde el trabajo mancomunado de sus dependencias, con miras a fortalecer el desarrollo de las distintas competencias que se requieran para la preparación profesional y personal.

La línea biopsicosocial, tiene por objetivo principal brindar acompañamiento y orientación al estudiante de manera personal, familiar y social, en temas que intervienen en la salud mental y emocional, fomentando estilos de vida saludables y aportando al desarrollo de competencias del estudiante.

La línea económica, esta concebida como una estrategias de atención al estudiante, donde se contemplan factores de riesgo que puedan afectar la permanencia del mismo en la Institución. Por medio de la gestión de acuerdos interinstitucionales y donaciones con empresas externas que posibiliten la entrega de bonos de alimentación , transporte, matrícula, así como reliquidación de matrícula y vinculación al servicio de salud.

La línea académica, permite orientar y remitir a los estudiantes que requieren de acompañamiento para lograr los objetivos de aprendizaje y rendimiento académico.

La línea normativa, buscar dar respuesta a las necesidades de información sobre admisiones, registro académico y documentación, así como la información oficial de los diferentes procesos y garantiza la aplicación adecuada de las normas constitucionales, legales, estatutarias y reglamentarias en los asuntos de competencia.

Es un objetivo de mayor importancia para el PAI reducir la deserción universitaria por medio del acompañamiento en múltiples frentes que se realiza a los estudiantes.

En tal sentido, es necesario, para la Vicerrectoría, identificar y evaluar de forma objetiva el impacto del PAI en la permanencia y egreso exitoso de los estudiantes de la UTP y adicionalmente cuantificar la relación existente entre el logro de los objetivos y los recursos empleados.

Dicha evaluación está enfocada inicialmente en el análisis cuidadoso de los estudiantes atendidos en la línea biopsicosocial, dada la alta demanda en temas

psicológicos, biológicos y condiciones sociales a los que la comunidad universitaria se ve expuesta actualmente y que pueden limitar de una u otra forma sus capacidades académicas y laborales.

En este documento se presenta un estudio estadístico retrospectivo del PAI de la UTP estructurado los siguientes capítulos:

1. **Análisis descriptivo de los datos muestrales.** Se presentan los indicadores generales de la deserción universitaria (semestral y anual) por grupos de interés relacionados con la labor de acompañamiento del PAI. Adicionalmente, son estimadas las probabilidades condicionales de ocurrencia de eventos de asistencia al programa y deserción universitaria.
2. **Análisis de relevancia de las características de los grupos objetivo.** Este capítulo muestra los resultados de tres tipos de análisis estadísticos, tales como: matriz de correlación, análisis de componentes principales y alineación centrada de kernel, con el fin de evidenciar cuáles son las características (variables) de mayor relevancia (importancia) en eventos tales como la deserción y la labor de acompañamiento del PAI.
3. **Evaluación de la relación costo-beneficio del PAI.** Muestra de manera cuantitativa los costos del promedio de la deserción universitaria, los costos de ejecución del PAI, y la relación de beneficio/pérdida al conseguir el egreso exitoso de los estudiantes.
4. **Conclusiones y recomendaciones.** Finalmente son planteadas las principales conclusiones de la evaluación objetiva de los datos y se presentan algunas recomendaciones de utilidad para fortalecer la visión y la estrategia de la UTP en la reducción de la deserción.

Capítulo 2

Análisis descriptivo de los datos muestrales

En este capítulo se describe de forma cuantitativa y gráfica las características de interés de la población de pregrado de la UTP. Asimismo, se presentan los indicadores generales de la deserción universitaria (semestral y anual) por grupos de interés y la relación que tiene esta cifra con la labor que desarrolla el PAI. Igualmente se estima y analiza un conjunto de probabilidades condicionales de ocurrencia de eventos asociados a la participación en el PAI, especialmente en la línea biopsicosocial y la deserción universitaria.

2.1. Descripción de los datos

La población objeto de análisis en este documento corresponde con los estudiantes de pregrado (tecnologías y carreras universitarias) de la UTP. En particular, para este trabajo, se tiene información estadística consolidada de datos de registro y matrícula de los periodos 2016-I, 2016-II y 2017-I. Adicionalmente, se tiene la información de asistencia al PAI con énfasis en la línea biopsicosocial de los mismos periodos temporales.

Es decir, para los análisis se tienen tres fuentes de información: registro y matrícula y atenciones del PAI. Además, la información está dada en los periodos de tiempo presentados anteriormente con el fin de poder establecer correlaciones y plantear hipótesis de dependencia entre las fuentes.

2.1.1. Datos de registro y matrícula

La información oficial de registro y matrícula cuenta con: datos de identificación personal de los estudiantes matriculados, datos asociadas al programa académico en el cual se matricula, variables demográficas, datos de ubicación y contacto, datos de educación secundaria, y algunos datos académicos.

Toda esta información fue depurada, homogeneizada y codificada para facilitar su tratamiento computacional. Así, las variables consideradas para el análisis son las siguientes:

- **ID3ro:** Código único interno de identificación de los estudiantes en la UTP.
- **Género:** Grupo al que pertenecen los seres humanos de cada sexo, entendido este desde un punto de vista sociocultural en lugar de exclusivamente biológico. Para el caso particular se tienen 2 categorías: masculino, femenino.
- **Edad:** Corresponde con los años cumplidos por el estudiante en el semestre que se analiza.
- **Rango de edad:** El rango de edad de cada estudiante se determina calculando, a partir de la fecha de nacimiento, la edad de cada estudiante en el semestre matriculado. En este estudio se consideran 11 rangos de edad de interés los cuales son descritos en la sección A.2. La variable rango de edad está calculada en años enteros cumplidos.
- **Programa:** Cada uno de los 63 programas académicos de pregrado que oferta o ha ofertado la UTP en sus diferentes modalidades para los semestres analizados. El detalle de estos programas se encuentra en la sección A.3.
- **Facultad:** Cada una de las 10 facultades de la UTP. Particularmente, facultad de bellas artes y humanidades, facultad de ciencias ambientales, facultad de ciencias básicas, facultad de ciencias de la educación, facultad de ciencias de la salud, facultad de ingeniería industrial, facultad de ingeniería mecánica, facultad de ingenierías, facultad de tecnología y facultad de ciencias agrarias y agroindustria.
- **Jornada:** Esta variable tiene 4 categorías, jornada diurna, jornada especial diurna, jornada nocturna, jornada especial nocturna.
- **Naturaleza del colegio:** Clasifica el colegio donde se graduó el estudiante de bachillerato en, oficial, privado o no aplica (NA).
- **Municipio del colegio:** Corresponde con el municipio donde está ubicado el colegio en el que se graduó como bachiller el estudiante. En caso que el estudiante provenga de un colegio ubicado fuera de Colombia, se etiqueta como *EXT*, y si la información se desconoce se clasifica como *NO DISPONIBLE*.
- **Departamento del colegio:** Corresponde al departamento en el que está ubicado el municipio del colegio en el que se graduó el estudiante. En caso que el estudiante provenga de un colegio ubicado fuera de Colombia, se etiqueta como *EXT*, y si la información se desconoce se clasifica como *NO DISPONIBLE*. El listado de los departamentos se presenta en la sección A.7.

- **Capital:** Define si el municipio del colegio es una ciudad capital de departamento o no.
- **Estrato:** La estratificación socioeconómica es una clasificación en estratos de los inmuebles residenciales que deben recibir servicios públicos. Es información brindada por el estudiante cuando realizó la última actualización de datos. Los estratos son: alto, bajo, bajo-bajo, medio, medio-alto, medio-bajo, o no disponible.
- **Semestre de ubicación:** Es el cálculo del semestre aproximado que esta cursando el estudiante con base en los créditos acumulados para el semestre de análisis.
- **Promedio acumulado:** Variable de tipo real continua que tiene el promedio acumulado calculado por registro y matrícula para el semestre en el que se analiza la información.

Es de aclarar que de la base de datos original se eliminaron registros que no eran adecuados para el análisis por que eran registros que correspondían a estudiantes de intercambio que solamente se matriculan en la UTP por unos pocos semestres.

2.1.2. Datos de atenciones del PAI

Adicional a las variables mencionadas antes, con relación al programa de atención integral se tiene la siguiente información:

- **Atendido PAI:** Variable dicotómica que identifica si el estudiante ha sido atendido en el PAI.
- **Atendido BPS:** Variable dicotómica que identifica si el estudiante ha sido atendido en el PAI, específicamente en la línea biopsicosocial.
- **Atendido académica:** Variable dicotómica que identifica si el estudiante ha sido atendido en el PAI, específicamente en la línea académica.
- **Atendido socio-económica:** Variable dicotómica que identifica si el estudiante ha sido atendido en el PAI, específicamente en la línea económica.
- **Número de atenciones:** Especifica cuantas veces en el semestre de análisis el estudiante fue atendido en el PAI.
- **Apoyo económico:** Variable dicotómica que identifica si el estudiante ha recibido apoyos económicos tales como, becas pa' pepas, becas talento, bono de alimentación, bono de donaciones, bono de matrícula, bono de transporte, DPS jóvenes en acción, monitoria social, plan padrino, reliquidacion de matrícula, Risaralda profesional, ser pilo paga, todos a la universidad.

2.2. Descripción del conjunto poblacional

Como se mencionó anteriormente, la población objeto de análisis de este documento son los estudiantes de pregrado de la UTP. El grupo de estudiantes de pregrado que realizaron el proceso de matrícula en cada semestre analizado conforman los conjuntos de análisis: M_{2016I} , M_{2016II} y M_{2017I} . El tamaño de estos conjuntos varía de semestre a semestre por los siguientes factores: ingreso de nuevos estudiantes, reingreso de estudiantes antiguos, graduación y deserción. La tabla 2.1 resume la cantidad de estudiantes matriculados en cada uno de los semestres de análisis.

Período	No. matriculados
2016-I	16.826
2016-II	15.998
2017-I	16.404

Tabla 2.1: Número de estudiantes matriculados por período

En las secciones siguientes se presenta la descripción cuantitativa y gráfica de la población objeto de análisis para cada una de las variables de interés señaladas con anterioridad.

2.2.1. Población por género

La población está conformada por género como se presenta en la tabla 2.2 y en la figura 2.1.

Género	2016-I	2016-II	2017-I
Masculino	9.704 (57,67 %)	9.228 (57,68 %)	9.469 (57,72 %)
Femenino	7.122 (42,33 %)	6.770 (42,32 %)	6.935 (42,28 %)

Tabla 2.2: Conformación de la población por género

La conformación por género es estable de semestre a semestre, con mayoría masculina que excede al género femenino en 15,3 % en promedio.

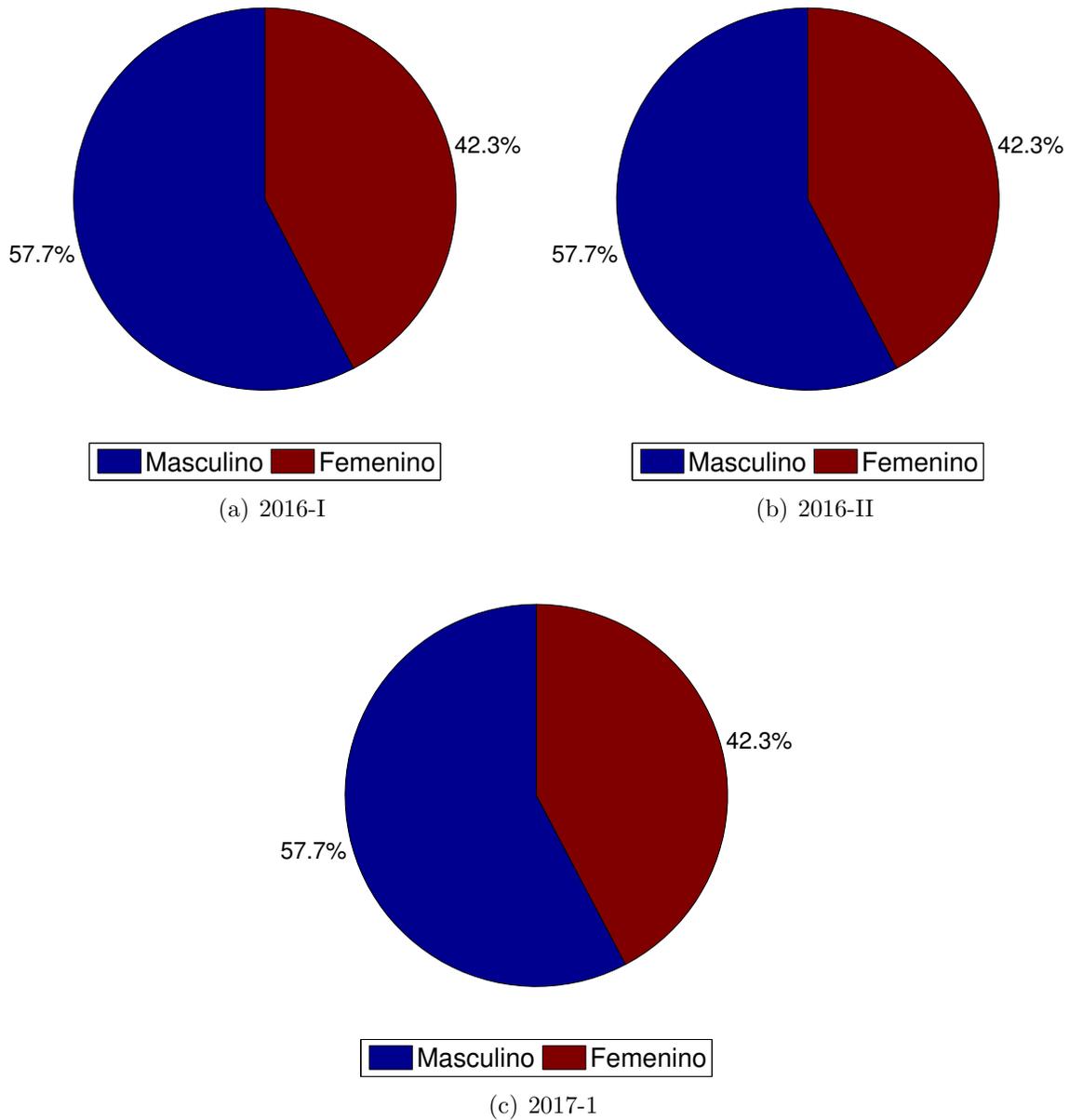


Figura 2.1: Conformación de la población por género

2.2.2. Población por rango de edad

La población está conformada por rango de edad como se presenta en la tabla 2.3 y en la figura 2.2. La edad promedio de la población en años es: $22,6 \pm 4,9$ para el período 2016-I, $22,69 \pm 5,01$ para el período 2016-2 y $22,45 \pm 4,92$ para el período 2017-I.

Rango edad	2016-I	2016-II	2017-I
0 a 10	0 (0%)	0 (0%)	0 (0%)
11 a 14	0 (0%)	0 (0%)	0 (0%)
15 a 17	1.042 (6,19%)	690 (4,31%)	1.067 (6,5%)
18 a 19	3.406 (20,24%)	3.387 (21,17%)	3.555 (21,67%)
20 a 21	4.036 (23,99%)	3.872 (24,2%)	3.828 (23,34%)
22 a 23	3.291 (19,56%)	3.226 (20,17%)	3.221 (19,64%)
24 a 25	1.933 (11,49%)	1.828 (11,43%)	1.836 (11,19%)
26 a 29	1.787 (10,62%)	1.705 (10,66%)	1.667 (10,22%)
30 a 35	885 (5,26%)	838 (5,24%)	788 (4,8%)
36 a 39	204 (1,21%)	200 (1,25%)	207 (1,26%)
40 y más	242 (1,44%)	252 (1,58%)	225 (1,37%)

Tabla 2.3: Conformación de la población por rango de edad

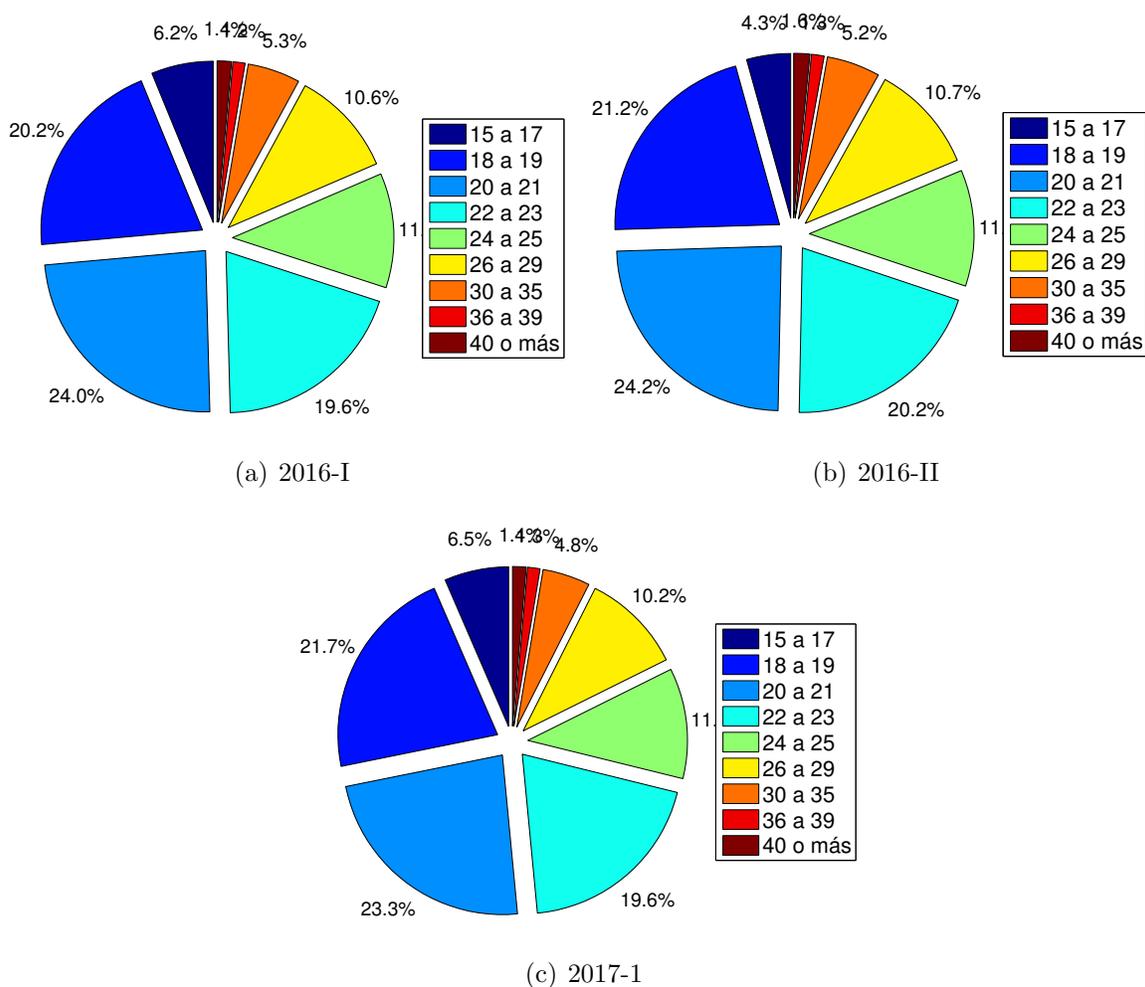


Figura 2.2: Conformación de la población por rango de edad

2.2.3. Población por programa académico

La población está conformada por programa académico como se presenta en la tabla 2.4 y en la figura 2.3. Los códigos de cada categoría se encuentran en la sección A.3.

El programa con más inscritos por semestre es P-24: ingeniería industrial, y el programa con menos estudiantes matriculados durante los semestres analizados son los programas P-23: Ingeniería en procesos sostenibles de las maderas y P-31: licenciatura en bilingüismo con énfasis en inglés, sin estudiantes matriculados en los semestres analizados.

Son notorios los programas que han crecido aceleradamente, en número de estudiantes matriculados, durante los semestres considerados, tales programas son, P-05: Administración del turismo sostenible jornada especial nocturna, P-11: Ingeniería de sistemas y computación (básicos ingeniería), P-12: Ingeniería de sistemas y computación convenio Cuba, P-14: Ingeniería eléctrica (básicos ingeniería), P-16:

Ingeniería electrónica jornada especial diurna, P-21: Ingeniería física (básicos ingeniería), P-22: Ingeniería en procesos agroindustriales, P-26: Ingeniería industrial (básicos ingeniería), P-29: Ingeniería mecánica (básicos ingeniería), P-37: Licenciatura en etnoeducación, P-62: Tecnología en gestión del turismo sostenible nocturna.

Por otra parte, se evidencia el cierre o reducción progresiva de estudiantes matriculados en los programas P-02: Administración del turismo sostenible - Ciclo I técnico profesional en procesos del turismo sostenible, P-03: Administración del turismo sostenible - Ciclo II tecnología en gestión del turismo sostenible. P-04: Administración turismo sostenible, P-19: Ingeniería en mecatrónica - Ciclo III ingeniería en mecatrónica jornada especial nocturna, P-33: Licenciatura en comunicación e informática educativa, P-34: Licenciatura en educación español y comunicación audiovisual, P-55: Tecnología industrial - Ceres Belén de Umbría, P-56: Tecnología industrial - Ceres Mistrató, P-57: Tecnología industrial - Ceres Quinchía, P-58: Tecnología industrial - Ceres Santuario.

Surge la hipótesis que reducción de matriculados en los programas P-02, P-03 y P-04 esté directamente relacionada con el aumento de matriculados en el programa P-05.

Programa	2016-I	2016-II	2017-I
P-00	798 (4,74 %)	769 (4,81 %)	748 (4,56 %)
P-01	130 (0,77 %)	104 (0,65 %)	130 (0,79 %)
P-02	240 (1,43 %)	1 (0,01 %)	0 (0 %)
P-03	62 (0,37 %)	0 (0 %)	0 (0 %)
P-04	44 (0,26 %)	22 (0,14 %)	0 (0 %)
P-05	0 (0 %)	15 (0,09 %)	55 (0,34 %)
P-06	0 (0 %)	1 (0,01 %)	1 (0,01 %)
P-07	722 (4,29 %)	688 (4,3 %)	708 (4,32 %)
P-08	1 (0,01 %)	1 (0,01 %)	0 (0 %)
P-09	748 (4,45 %)	740 (4,63 %)	756 (4,61 %)
P-10	298 (1,77 %)	299 (1,87 %)	290 (1,77 %)
P-11	0 (0 %)	0 (0 %)	31 (0,19 %)
P-12	0 (0 %)	0 (0 %)	46 (0,28 %)
P-13	1.007 (5,98 %)	979 (6,12 %)	995 (6,07 %)
P-14	0 (0 %)	0 (0 %)	22 (0,13 %)
P-15	291 (1,73 %)	226 (1,41 %)	197 (1,2 %)
P-16	0 (0 %)	23 (0,14 %)	68 (0,41 %)
P-17	311 (1,85 %)	317 (1,98 %)	286 (1,74 %)
P-18	141 (0,84 %)	153 (0,96 %)	200 (1,22 %)
P-19	190 (1,13 %)	167 (1,04 %)	125 (0,76 %)
P-20	351 (2,09 %)	352 (2,2 %)	373 (2,27 %)
P-21	0 (0 %)	0 (0 %)	17 (0,1 %)
P-22	0 (0 %)	0 (0 %)	57 (0,35 %)
P-23	0 (0 %)	0 (0 %)	0 (0 %)
P-24	1.239 (7,36 %)	1.212 (7,58 %)	1.225 (7,47 %)
P-25	823 (4,89 %)	813 (5,08 %)	828 (5,05 %)
P-26	0 (0 %)	0 (0 %)	29 (0,18 %)
P-27	943 (5,6 %)	895 (5,59 %)	936 (5,71 %)
P-28	3 (0,02 %)	2 (0,01 %)	0 (0 %)
P-29	0 (0 %)	0 (0 %)	22 (0,13 %)
P-30	695 (4,13 %)	681 (4,26 %)	661 (4,03 %)
P-31	0 (0 %)	0 (0 %)	0 (0 %)
P-32	355 (2,11 %)	318 (1,99 %)	356 (2,17 %)
P-33	719 (4,27 %)	662 (4,14 %)	637 (3,88 %)
P-34	6 (0,04 %)	4 (0,03 %)	0 (0 %)
P-35	540 (3,21 %)	539 (3,37 %)	553 (3,37 %)
P-36	371 (2,2 %)	306 (1,91 %)	242 (1,48 %)
P-37	0 (0 %)	51 (0,32 %)	103 (0,63 %)
P-38	158 (0,94 %)	141 (0,88 %)	169 (1,03 %)
P-39	224 (1,33 %)	179 (1,12 %)	240 (1,46 %)
P-40	417 (2,47 %)	378 (2,36 %)	421 (2,57 %)
P-41	852 (5,06 %)	815 (5,09 %)	814 (4,96 %)
P-42	23 (0,14 %)	23 (0,14 %)	23 (0,14 %)
P-43	18 (0,11 %)	18 (0,11 %)	17 (0,1 %)
P-44	161 (0,96 %)	164 (1,03 %)	155 (0,94 %)
P-45	20 (0,12 %)	20 (0,13 %)	20 (0,12 %)
P-46	26 (0,15 %)	44 (0,28 %)	40 (0,24 %)
P-47	734 (4,36 %)	723 (4,52 %)	736 (4,49 %)
P-48	465 (2,76 %)	483 (3,02 %)	494 (3,01 %)
P-49	397 (2,36 %)	429 (2,68 %)	377 (2,3 %)
P-50	464 (2,76 %)	430 (2,69 %)	416 (2,54 %)
P-51	0 (0 %)	24 (0,15 %)	14 (0,09 %)
P-52	179 (1,06 %)	171 (1,07 %)	185 (1,13 %)
P-53	608 (3,61 %)	543 (3,39 %)	525 (3,2 %)
P-54	0 (0 %)	33 (0,21 %)	0 (0 %)
P-55	30 (0,18 %)	0 (0 %)	0 (0 %)
P-56	34 (0,2 %)	0 (0 %)	0 (0 %)
P-57	19 (0,11 %)	0 (0 %)	0 (0 %)
P-58	35 (0,21 %)	0 (0 %)	0 (0 %)
P-59	86 (0,51 %)	57 (0,36 %)	45 (0,27 %)
P-60	510 (3,03 %)	450 (2,81 %)	439 (2,68 %)
P-61	338 (2,01 %)	223 (1,39 %)	244 (1,49 %)
P-62	0 (0 %)	310 (1,94 %)	333 (2,03 %)

Tabla 2.4: Conformación de la población por programa académico

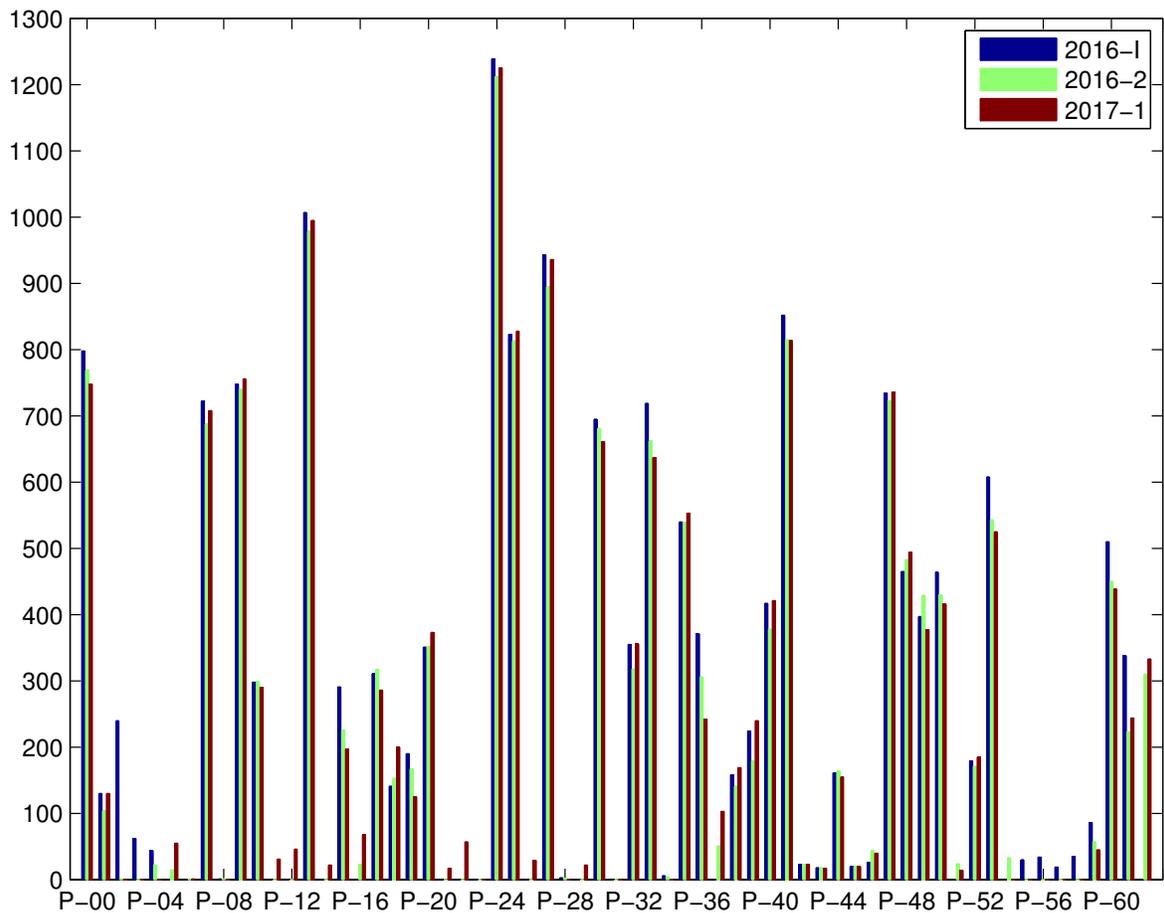


Figura 2.3: Conformación de la población por programa académico

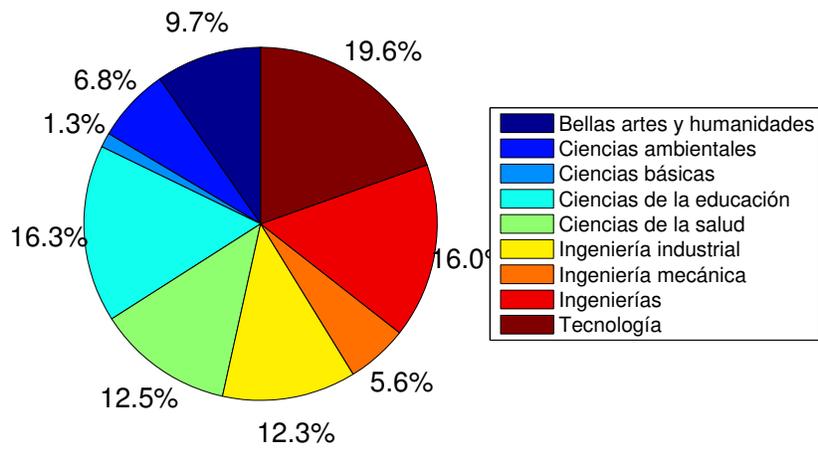
2.2.4. Población por facultad

La población está conformada por facultad como se presenta en la tabla 2.5 y en la figura 2.4.

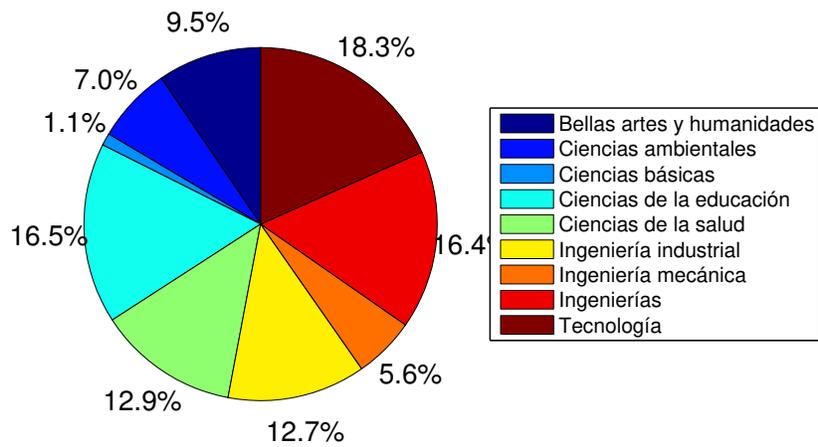
Al revisar la población por facultad, el mayor número de estudiantes matriculados lo tiene la facultad de tecnología. Sin embargo, esta facultad presenta una caída notoria en la cantidad de matriculados semestre tras semestre. Por otra parte, la facultad de ciencias agrarias es la que menos estudiantes matriculados tiene, factor atribuible a su reciente creación.

Facultad	2016-I	2016-II	2017-I
Bellas artes y humanidades	1.626 (9,66 %)	1.520 (9,5 %)	1.608 (9,8 %)
Ciencias ambientales	1.144 (6,8 %)	1.117 (6,98 %)	1.136 (6,93 %)
Ciencias básicas	224 (1,33 %)	179 (1,12 %)	240 (1,46 %)
Ciencias de la educación	2.736 (16,26 %)	2.646 (16,54 %)	2.604 (15,87 %)
Ciencias de la salud	2.100 (12,48 %)	2.065 (12,91 %)	2.123 (12,94 %)
Ingeniería industrial	2.062 (12,25 %)	2.025 (12,66 %)	2.082 (12,69 %)
Ingeniería mecánica	946 (5,62 %)	897 (5,61 %)	958 (5,84 %)
Ingenierías	2.695 (16,02 %)	2.619 (16,37 %)	2.795 (17,04 %)
Tecnología	3.293 (19,57 %)	2.930 (18,31 %)	2.801 (17,07 %)
Ciencias agrarias	0 (0 %)	0 (0 %)	57 (0,35 %)

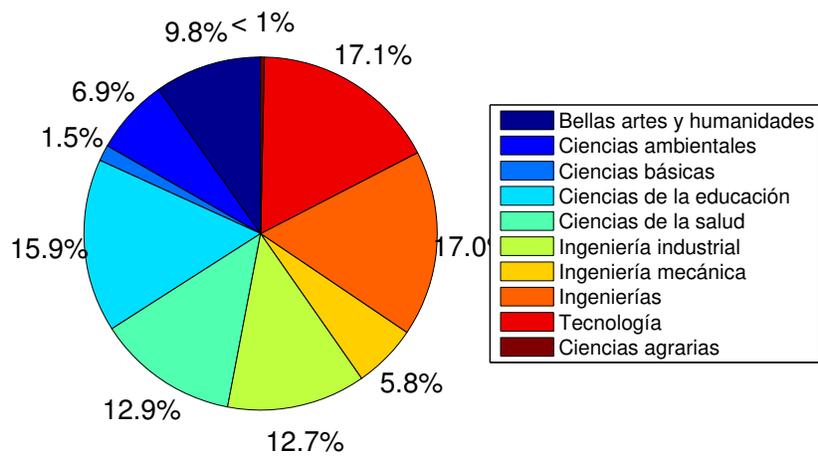
Tabla 2.5: Conformación de la población por facultad



(a) 2016-I



(b) 2016-II



(c) 2017-1

Figura 2.4: Conformación de la población por facultad

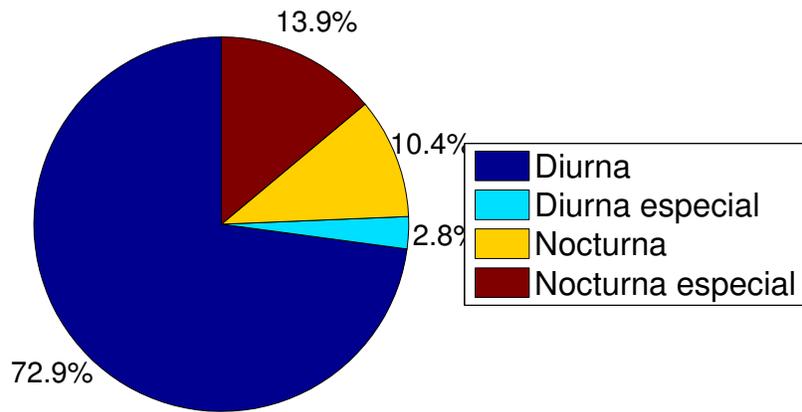
2.2.5. Población por jornada

La población está conformada por jornada como se presenta en la tabla 2.6 y en la figura 2.5.

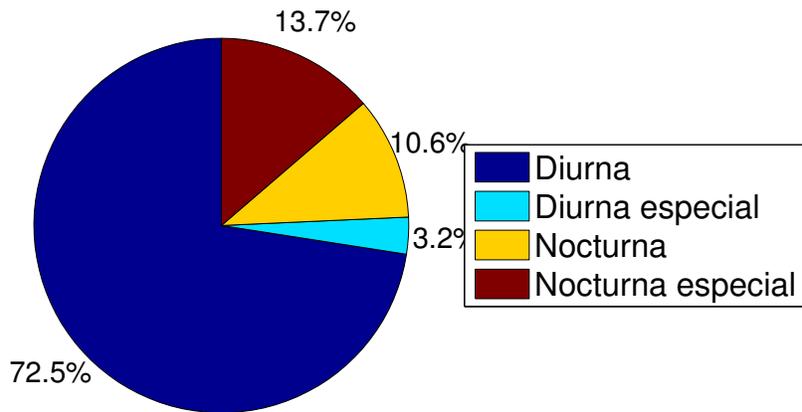
Es notorio que la jornada nocturna y jornada nocturna especial constituyen aproximadamente el 24 % de la población de la UTP. Además hay un comportamiento creciente de matriculados en las jornadas diurna especial y nocturna, mientras que la jornada diurna presenta un comportamiento decreciente.

Jornada	2016-I	2016-II	2017-I
Diurna	12.264 (72,89 %)	11.582 (72,55 %)	11.694 (71,49 %)
Diurna especial	465 (2,76 %)	506 (3,17 %)	562 (3,44 %)
Nocturna	1.755 (10,43 %)	1.692 (10,6 %)	1.815 (11,1 %)
Nocturna especial	2.342 (13,92 %)	2.185 (16,69 %)	2.287 (13,98 %)

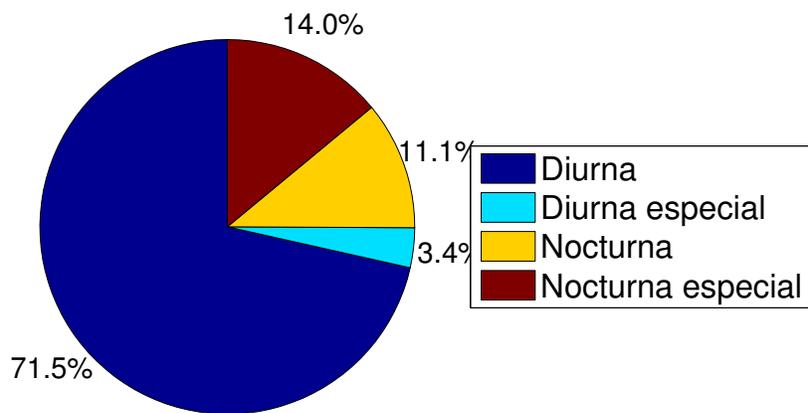
Tabla 2.6: Conformación de la población por jornada



(a) 2016-I



(b) 2016-II



(c) 2017-1

Figura 2.5: Conformación de la población por jornada

2.2.6. Población por naturaleza del colegio

La población está conformada por naturaleza del colegio como se presenta en la tabla 2.7 y en la figura 2.6.

Se aprecia que la conformación ha permanecido estable durante los semestres que se evalúan, siendo los estudiante matriculados en la UTP provenientes en su gran mayoría de colegios oficiales (promedio de 82,9%).

Naturaleza colegio	2016-I	2016-II	2017-I
N/A	22 (0,13 %)	7 (0,04 %)	12 (0,07 %)
Oficial	13.949 (82,9 %)	13.230 (82,7 %)	13.631 (83,1 %)
Privado	2.855 (16,97 %)	2.761 (17,26 %)	2.761 (16,83 %)

Tabla 2.7: Conformación de la población por naturaleza del colegio



Figura 2.6: Conformación de la población por naturaleza del colegio

2.2.7. Población por departamento

La población está conformada por departamento de procedencia como se presenta en la tabla 2.8 y en la figura 2.7. Adicionalmente, para una mejor comprensión de las gráficas en la figura 2.8 se repiten los resultados eliminando al departamento de Risaralda. Los códigos de cada categoría se encuentran en la sección A.7.

Departamento	2016-I	2016-II	2017-I
Amazonas	5 (0,03 %)	4 (0,03 %)	3 (0,02 %)
Antioquia	70 (0,42 %)	71 (0,44 %)	72 (0,44 %)
Arauca	5 (0,03 %)	6 (0,04 %)	6 (0,04 %)
Atlántico	17 (0,10 %)	23 (0,14 %)	22 (0,13 %)
Bogotá	180 (1,08 %)	179 (1,12 %)	171 (1,04 %)
Bolívar	16 (0,1 %)	16 (0,1 %)	14 (0,09 %)
Boyacá	18 (0,11 %)	14 (0,09 %)	13 (0,08 %)
Caldas	698 (4,18 %)	677 (4,23 %)	687 (4,19 %)
Caquetá	58 (0,35 %)	59 (0,37 %)	60 (0,37 %)
Casanare	8 (0,05 %)	11 (0,07 %)	10 (0,06 %)
Cauca	121 (0,72 %)	114 (0,71 %)	119 (0,73 %)
Cesar	13 (0,08 %)	13 (0,08 %)	10 (0,06 %)
Chocó	81 (0,48 %)	79 (0,49 %)	68 (0,41 %)
Córdoba	13 (0,08 %)	14 (0,09 %)	15 (0,09 %)
Cundinamarca	70 (0,42 %)	73 (0,46 %)	70 (0,43 %)
Guaviare	11 (0,07 %)	13 (0,08 %)	13 (0,08 %)
Huila	113 (0,68 %)	124 (0,78 %)	130 (0,79 %)
Guajira	4 (0,02 %)	4 (0,03 %)	4 (0,02 %)
Guainía	1 (0,01 %)	2 (0,01 %)	1 (0,01 %)
Magdalena	7 (0,04 %)	6 (0,04 %)	6 (0,04 %)
Meta	32 (0,19 %)	27 (0,17 %)	36 (0,22 %)
N. de Santander	23 (0,14 %)	24 (0,15 %)	25 (0,15 %)
Nariño	512 (3,06 %)	557 (3,48 %)	621 (3,79 %)
Putumayo	122 (0,73 %)	129 (0,81 %)	133 (0,81 %)
Quindío	364 (2,18 %)	358 (2,24 %)	367 (2,24 %)
Risaralda	12.182 (72,91 %)	11.378 (71,12 %)	11.647 (71 %)
Santander	21 (0,13 %)	19 (0,12 %)	20 (0,12 %)
San Andrés	35 (0,21 %)	45 (0,28 %)	43 (0,26 %)
Sucre	6 (0,04 %)	6 (0,04 %)	6 (0,04 %)
Tolima	146 (0,87 %)	150 (0,94 %)	159 (0,97 %)
Valle del Cauca	1.738 (10,4 %)	1.694 (10,59 %)	1.742 (10,62 %)
Vaupés	3 (0,02 %)	3 (0,02 %)	2 (0,01 %)
Ext	16 (0,1 %)	12 (0,08 %)	16 (0,1 %)
No disponible	117 (0,69 %)	94 (0,59 %)	93 (0,57 %)

Tabla 2.8: Conformación de la población por departamento

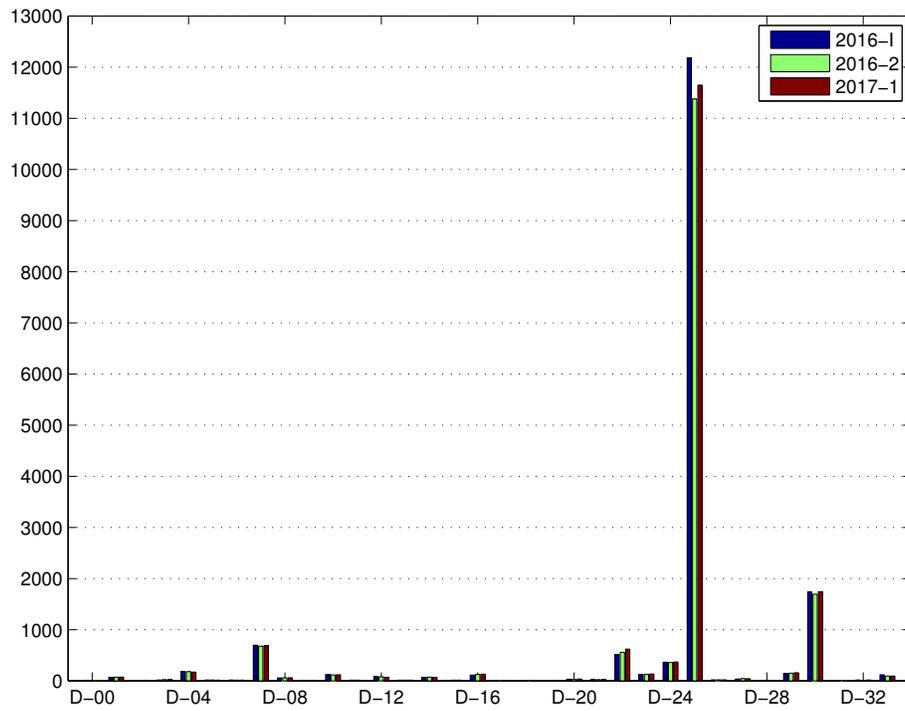


Figura 2.7: Conformación de la población por departamento

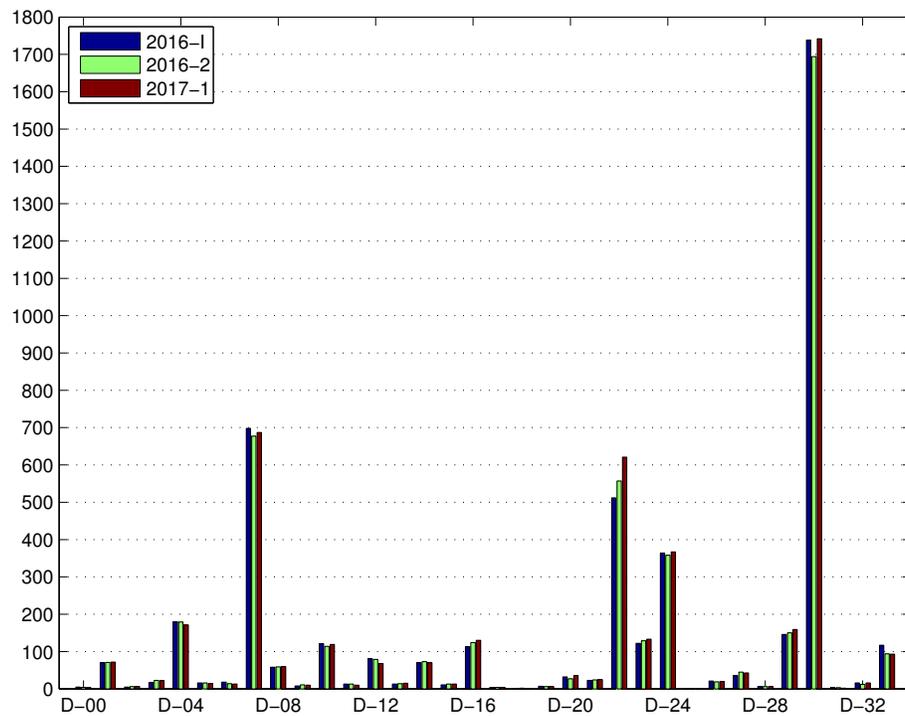


Figura 2.8: Conformación de la población por departamento excluyendo Risaralda

2.2.8. Población por capital

La población está conformada por capital de departamento como se presenta en la tabla 2.9 y en la figura 2.9. Es de notar que la mitad de los estudiantes matriculados de la UTP provienen de ciudades capitales.

Capital	2016-I	2016-II	2017-I
No	8.517 (50,62 %)	8.086 (50,54 %)	8.339 (50,84 %)
Sí	8.309 (49,38 %)	7.912 (49,46 %)	8.065 (49,16 %)

Tabla 2.9: Conformación de la población por capital

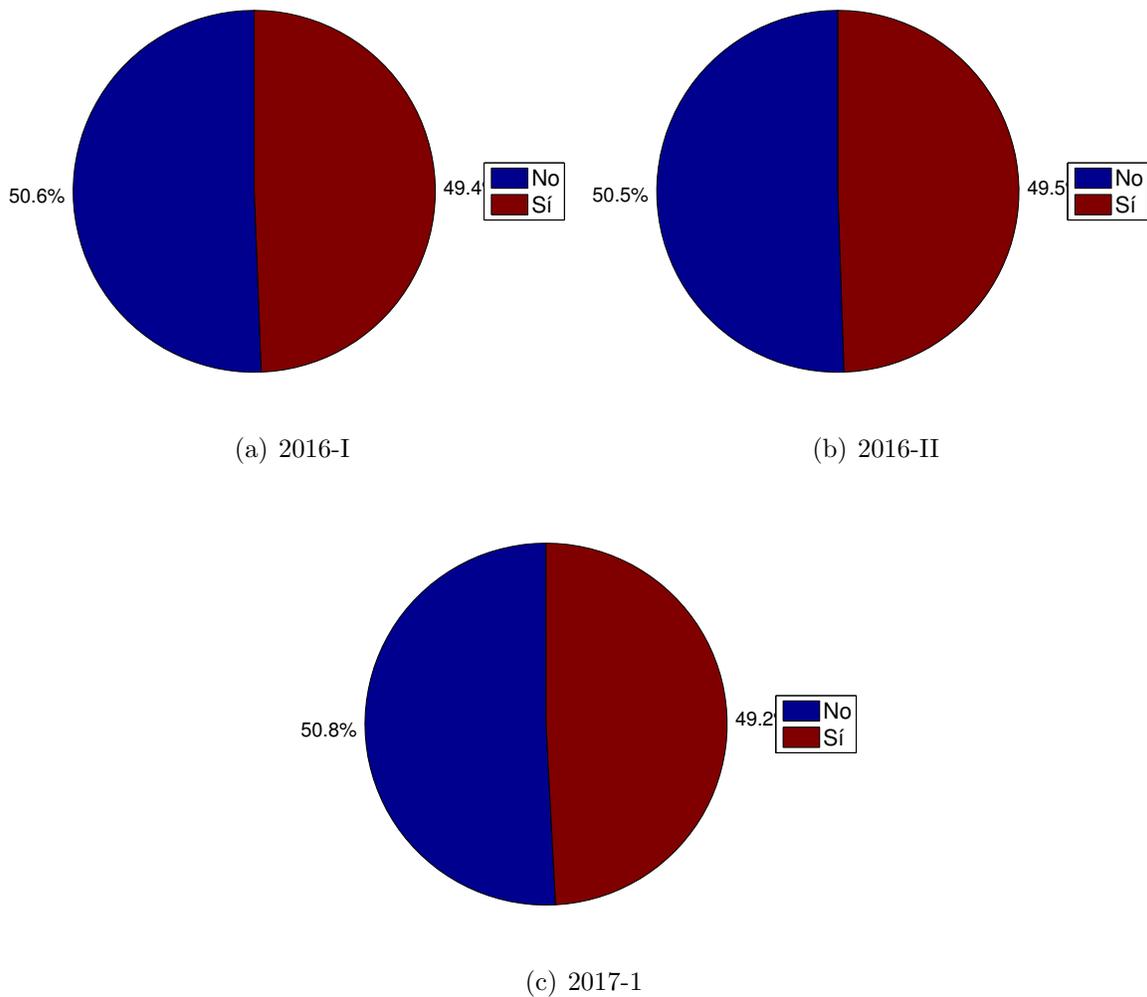


Figura 2.9: Conformación de la población por capital

2.2.9. Población por estrato

La población está conformada por estrato como se presenta en la tabla 2.10 y en la figura 2.10.

La UTP tiene una participación mayoritaria de estudiantes matriculados provenientes de estrato bajo. En particular, los estratos 1, 2 y 3 conforman en promedio el 88,75 % de los estudiantes matriculados.

Estrato	2016-I	2016-II	2017-I
Bajo-bajo	3.920 (23,3 %)	3.842 (24,02 %)	3.970 (24,2 %)
Bajo	7.363 (43,76 %)	7.078 (44,24 %)	7.335 (44,71 %)
Medio-bajo	3.601 (21,4 %)	3.254 (20,34 %)	3.327 (20,28 %)
Medio	1.509 (8,97 %)	1.411 (8,82 %)	1.343 (8,19 %)
Medio-alto	312 (1,85 %)	311 (1,94 %)	317 (1,93 %)
Alto	98 (0,58 %)	94 (0,59 %)	98 (0,6 %)
No disponible	23 (0,14 %)	8 (0,05 %)	14 (0,09 %)

Tabla 2.10: Conformación de la población por estrato

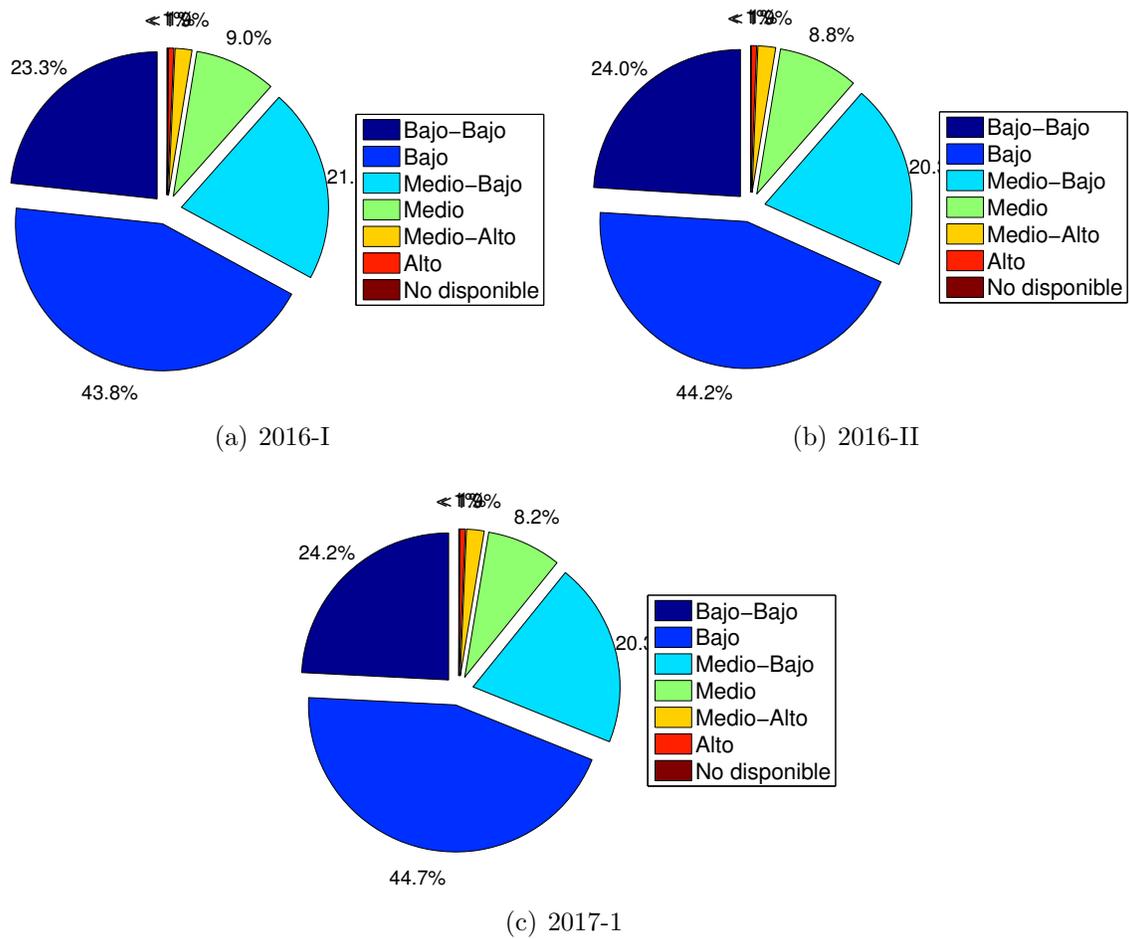


Figura 2.10: Conformación de la población por estrato

2.2.10. Población por semestre de ubicación

La población está conformada por semestre de ubicación como se presenta en la tabla 2.11 y en la figura 2.11.

Semestre	2016-I	2016-II	2017-I
1	1.140 (6,78 %)	1.179 (7,63 %)	1.222 (7,71 %)
2	2.322 (13,8 %)	1.861 (12,04 %)	2.451 (15,46 %)
3	1.647 (9,79 %)	1.748 (11,31 %)	1.555 (9,81 %)
4	1.639 (9,74 %)	1.373 (8,89 %)	1.583 (9,98 %)
5	1.226 (7,29 %)	1.447 (9,37 %)	1.193 (7,52 %)
6	2.030 (12,06 %)	1.592 (10,3 %)	1.655 (10,44 %)
7	1.220 (7,25 %)	1.147 (7,42 %)	1.059 (6,68 %)
8	993 (5,9 %)	1.218 (7,88 %)	1.116 (7,04 %)
9	968 (5,73 %)	943 (6,1 %)	1.177 (7,42 %)
10	2.432 (14,45 %)	2.378 (15,39 %)	2.252 (14,2 %)
11	215 (1,28 %)	200 (1,29 %)	235 (1,48 %)
12	361 (2,15 %)	300 (1,94 %)	273 (1,72 %)
13	70 (0,42 %)	65 (0,42 %)	83 (0,52 %)
No disponible	563 (3,35 %)	0 (0 %)	0 (0 %)

Tabla 2.11: Conformación de la población por semestre de ubicación

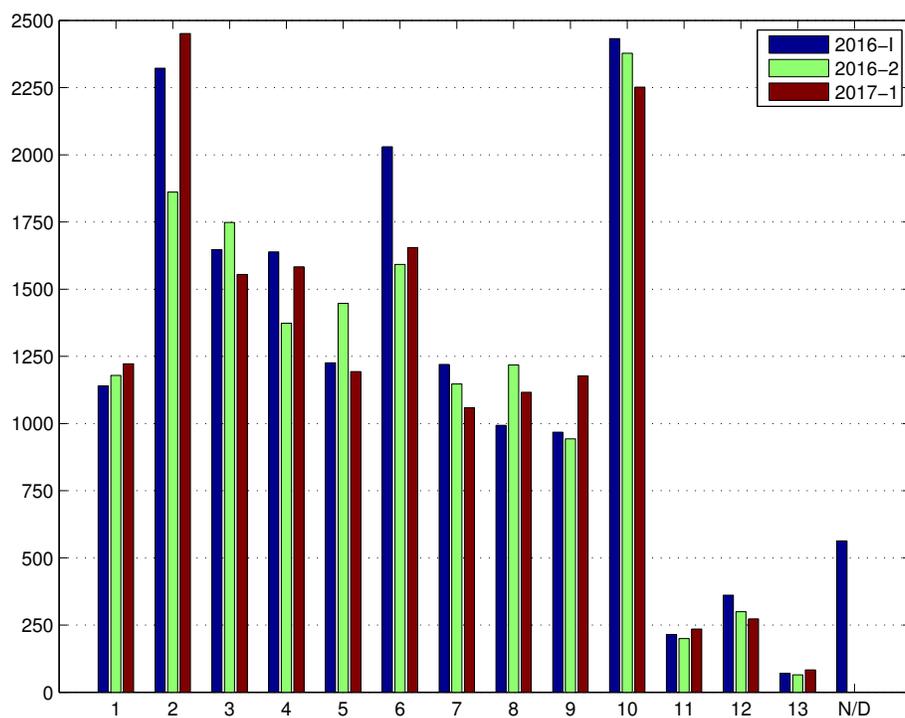


Figura 2.11: Conformación de la población por semestre de ubicación

2.3. Descripción del subconjunto poblacional PAI

En esta sección denotaremos al conjunto PAI como el subconjunto de la población de estudiantes matriculados que acudieron al Programa de Atención Integral (PAI), es decir, se tiene que: $PAI2016_I \subseteq M2016_I$, $PAI2016_{II} \subseteq M2016_{II}$ y $PAI2017_I \subseteq M2017_I$. Además, se presentan las estadísticas básicas descriptivas del subconjunto en mención. La tabla 2.12 muestra la cantidad de estudiantes atendidos en el PAI en los períodos 2016-I, 2016-II y 2017-I.

Período	Atendidos PAI	Atendidos/Población
2016-I	1.251	7,43 %
2016-II	1.581	9.88 %
2017-I	2.532	15,44 %

Tabla 2.12: Número atendidos en PAI por período

2.3.1. Subconjunto PAI por género

El subconjunto PAI está conformado por género como se presenta en la tabla 2.13 y en la figura 2.12.

Género	2016-I	2016-II	2017-I
Masculino	684 (54,68 %)	831 (52,56 %)	1.401 (55,33 %)
Femenino	567 (45,32 %)	750 (47,44 %)	1.131 (44,67 %)

Tabla 2.13: Conformación del subconjunto PAI por género

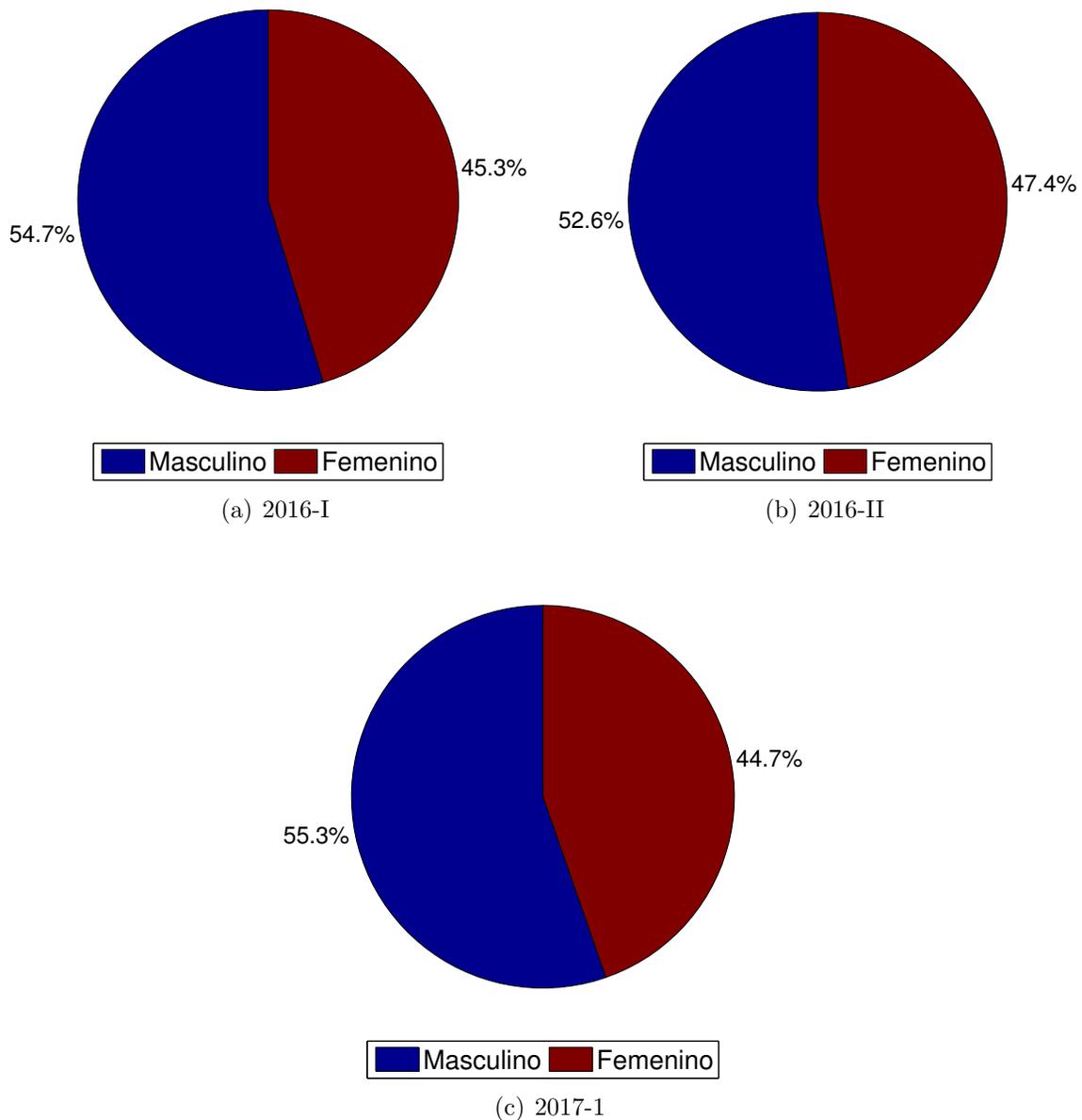


Figura 2.12: Conformación del subconjunto PAI por género

2.3.2. Subconjunto PAI por rango de edad

El subconjunto PAI está conformado por rango de edad como se presenta en la tabla 2.14 y en la figura 2.13. La edad promedio del subconjunto PAI en años es: $21,06 \pm 4,45$ para el período 2016-I, $20,69 \pm 4,25$ para el período 2016-2 y $19,86 \pm 3,85$ para el período 2017-I.

Rango edad	2016-I	2016-II	2017-I
0 a 10	0 (0%)	0 (0%)	0 (0%)
11 a 14	0 (0%)	0 (0%)	0 (0%)
15 a 17	231 (18,47%)	209 (13,22%)	718 (28,36%)
18 a 19	285 (22,78%)	595 (37,63%)	842 (33,25%)
20 a 21	276 (22,06%)	337 (21,32%)	398 (15,72%)
22 a 23	240 (19,18%)	176 (11,13%)	262 (10,35%)
24 a 25	87 (6,95%)	112 (7,08%)	131 (5,17%)
26 a 29	81 (6,47%)	83 (5,25%)	112 (4,42%)
30 a 35	39 (3,12%)	51 (3,23%)	50 (1,97%)
36 a 39	2 (0,16%)	6 (0,38%)	7 (0,28%)
40 y más	10 (0,8%)	12 (0,76%)	12 (0,47%)

Tabla 2.14: Conformación del subconjunto PAI por rango de edad

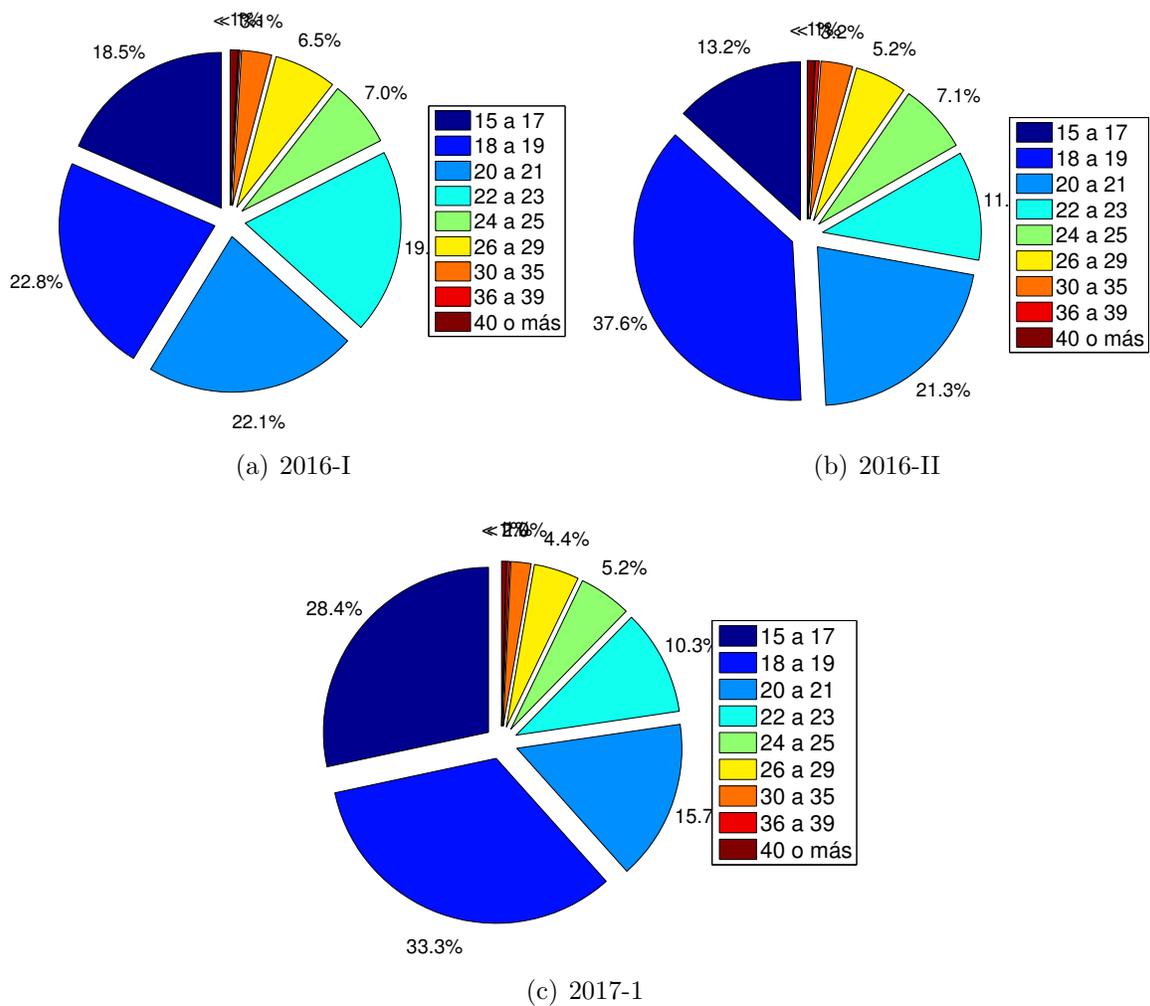


Figura 2.13: Conformación del subconjunto PAI por rango de edad

2.3.3. Subconjunto PAI por programa académico

El subconjunto PAI está conformado por programa académico como se presenta en la tabla 2.15. Los códigos de cada categoría se encuentran en la sección A.3.

Programa	2016-I	2016-II	2017-I
P-00	46 (3,68 %)	77 (4,87 %)	132 (5,21 %)
P-01	0 (0 %)	1 (0,06 %)	0 (0 %)
P-02	20 (1,59 %)	0 (0 %)	0 (0 %)
P-03	2 (0,16 %)	0 (0 %)	0 (0 %)
P-04	4 (0,32 %)	0 (0 %)	0 (0 %)
P-05	0 (0 %)	1 (0,06 %)	1 (0,04 %)
P-06	0 (0 %)	0 (0 %)	1 (0,04 %)
P-07	71 (5,68 %)	81 (5,12 %)	95 (3,75 %)
P-08	0 (0 %)	0 (0 %)	0 (0 %)
P-09	37 (2,96 %)	85 (5,38 %)	112 (4,42 %)
P-10	9 (0,72 %)	48 (3,04 %)	48 (1,9 %)
P-11	0 (0 %)	0 (0 %)	0 (0 %)
P-12	0 (0 %)	0 (0 %)	25 (0,99 %)
P-13	50 (4 %)	74 (4,68 %)	97 (3,83 %)
P-14	0 (0 %)	0 (0 %)	0 (0 %)
P-15	12 (0,96 %)	3 (0,19 %)	5 (0,2 %)
P-16	0 (0 %)	2 (0,13 %)	44 (1,74 %)
P-17	30 (2,4 %)	59 (3,73 %)	57 (2,25 %)
P-18	22 (1,76 %)	1 (0,06 %)	1 (0,04 %)
P-19	10 (0,8 %)	4 (0,25 %)	0 (0 %)
P-20	48 (3,84 %)	44 (2,78 %)	216 (8,53 %)
P-21	0 (0 %)	0 (0 %)	0 (0 %)
P-22	0 (0 %)	0 (0 %)	41 (1,62 %)
P-23	0 (0 %)	0 (0 %)	0 (0 %)
P-24	102 (8,15 %)	104 (6,58 %)	129 (5,09 %)
P-25	21 (1,68 %)	39 (2,47 %)	63 (2,49 %)
P-26	0 (0 %)	0 (0 %)	0 (0 %)
P-27	157 (12,55 %)	69 (4,36 %)	127 (5,02 %)
P-28	0 (0, %)	0 (0 %)	0 (0 %)
P-29	0 (0 %)	0 (0 %)	0 (0 %)
P-30	50 (4 %)	64 (4,05 %)	92 (3,63 %)
P-31	0 (0 %)	0 (0 %)	0 (0 %)
P-32	36 (2,88 %)	13 (0,82 %)	69 (2,73 %)
P-33	74 (5,92 %)	87 (5,5 %)	107 (4,23 %)
P-34	0 (0 %)	0 (0 %)	0 (0 %)
P-35	28 (2,24 %)	70 (4,43 %)	73 (2,88 %)
P-36	26 (2,08 %)	8 (0,51 %)	15 (0,59 %)
P-37	0 (0 %)	30 (1,9 %)	14 (0,55 %)
P-38	9 (0,72 %)	3 (0,19 %)	39 (1,54 %)
P-39	17 (1,36 %)	5 (0,32 %)	62 (2,45 %)
P-40	40 (3,2 %)	12 (0,76 %)	91 (3,59 %)
P-41	56 (4,48 %)	117 (7,4 %)	131 (5,17 %)
P-42	0 (0 %)	0 (0 %)	0 (0 %)
P-43	0 (0 %)	0 (0 %)	0 (0 %)
P-44	0 (0 %)	0 (0 %)	0 (0 %)
P-45	0 (0 %)	0 (0 %)	0 (0 %)
P-46	0 (0 %)	0 (0 %)	0 (0 %)
P-47	60 (4,79 %)	61 (3,86 %)	102 (4,03 %)
P-48	27 (2,16 %)	53 (3,35 %)	50 (1,97 %)
P-49	16 (1,28 %)	78 (4,93 %)	26 (1,03 %)
P-50	35 (2,8 %)	62 (3,92 %)	82 (3,24 %)
P-51	0 (0 %)	0 (0 %)	0 (0 %)
P-52	17 (1,36 %)	32 (2,02 %)	37 (1,46 %)
P-53	49 (3,92 %)	78 (4,93 %)	92 (3,63 %)
P-54	0 (0 %)	0 (0 %)	0 (0 %)
P-55	0 (0 %)	0 (0 %)	0 (0 %)
P-56	0 (0 %)	0 (0 %)	0 (0 %)
P-57	0 (0 %)	0 (0 %)	0 (0 %)
P-58	0 (0 %)	0 (0 %)	0 (0 %)
P-59	6 (0,48 %)	2 (0,13 %)	1 (0,04 %)
P-60	37 (2,96 %)	77 (4,87 %)	88 (3,48 %)
P-61	27 (2,16 %)	10 (0,63 %)	74 (2,92 %)
P-62	0 (0 %)	27 (1,71 %)	93 (3,67 %)

Tabla 2.15: Conformación del subconjunto PAI por programa académico

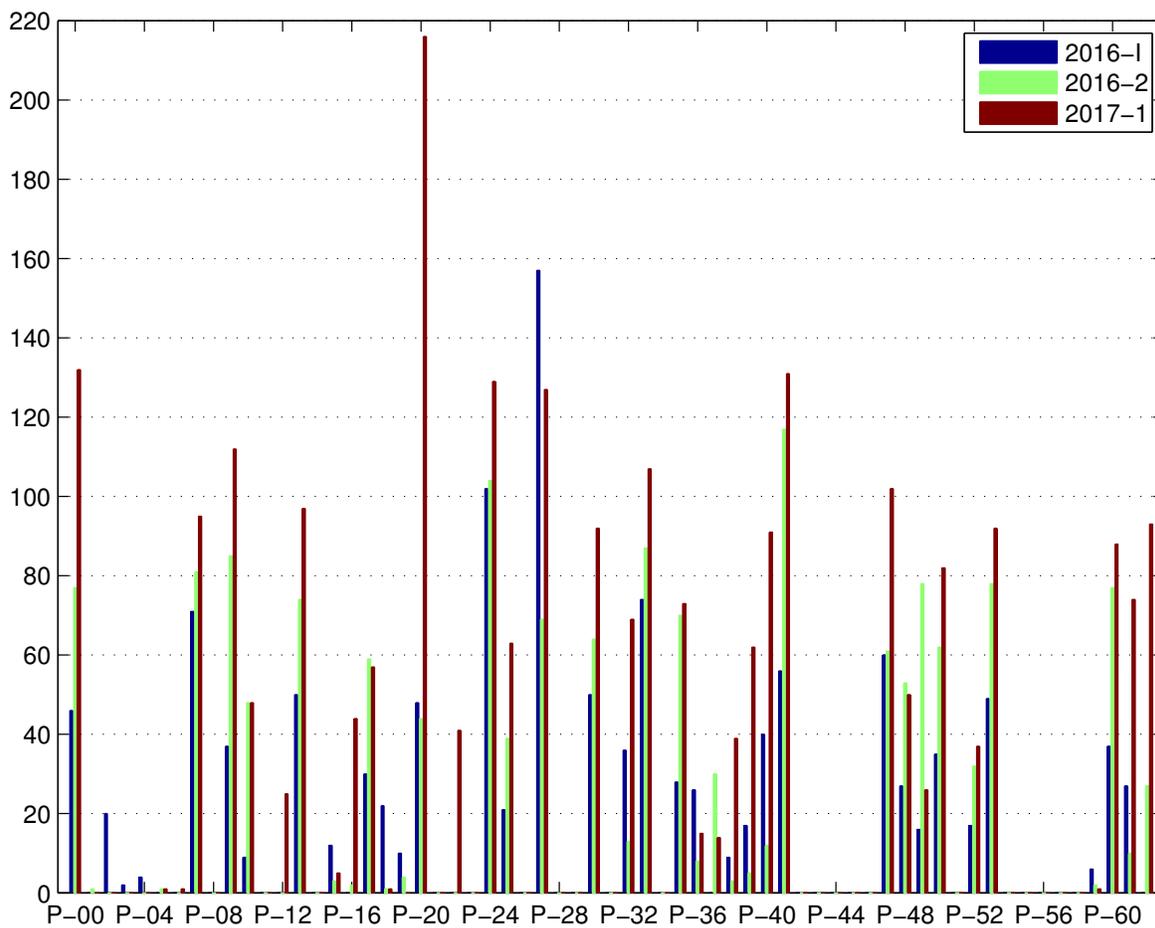


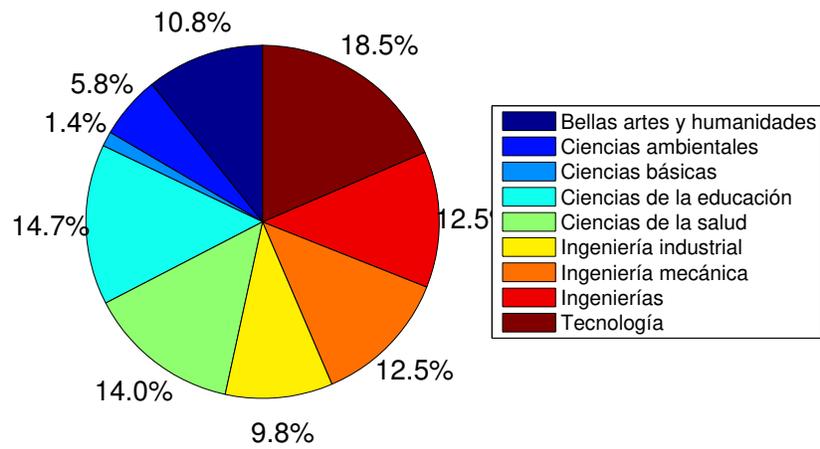
Figura 2.14: Conformación del subconjunto PAI por programa académico

2.3.4. Subconjunto PAI por facultad

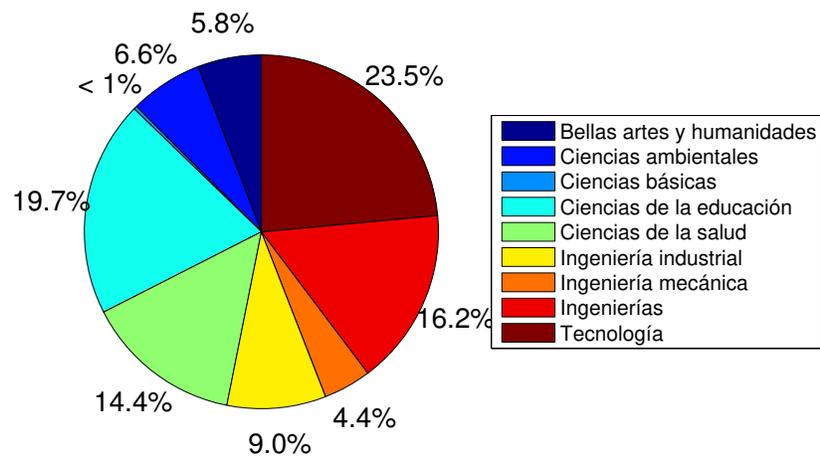
El subconjunto PAI está conformado por facultad como se presenta en la tabla 2.16 y en la figura 2.15.

Facultad	2016-I	2016-II	2017-I
Bellas artes y humanidades	135 (10,79 %)	92 (5,82 %)	292 (11,53 %)
Ciencias ambientales	72 (5,76 %)	105 (6,64 %)	226 (8,93 %)
Ciencias básicas	17 (1,36 %)	5 (0,32 %)	62 (2,45 %)
Ciencias de la educación	184 (14,71 %)	312 (19,73 %)	340 (13,43 %)
Ciencias de la salud	175 (13,99 %)	227 (14,36 %)	284 (11,22 %)
Ingeniería industrial	123 (9,83 %)	143 (9,04 %)	192 (7,58 %)
Ingeniería mecánica	157 (12,55 %)	69 (4,36 %)	127 (5,02 %)
Ingenierías	156 (12,47 %)	256 (16,19 %)	547 (21,6 %)
Tecnología	232 (18,55 %)	372 (23,53 %)	421 (16,63 %)
Ciencias agrarias	0 (0 %)	0 (0 %)	41 (1,62 %)

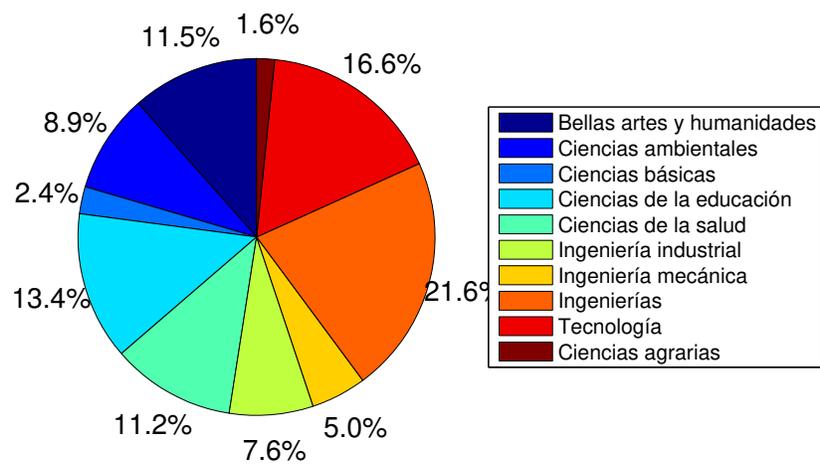
Tabla 2.16: Conformación del subconjunto PAI por facultad



(a) 2016-I



(b) 2016-II



(c) 2017-1

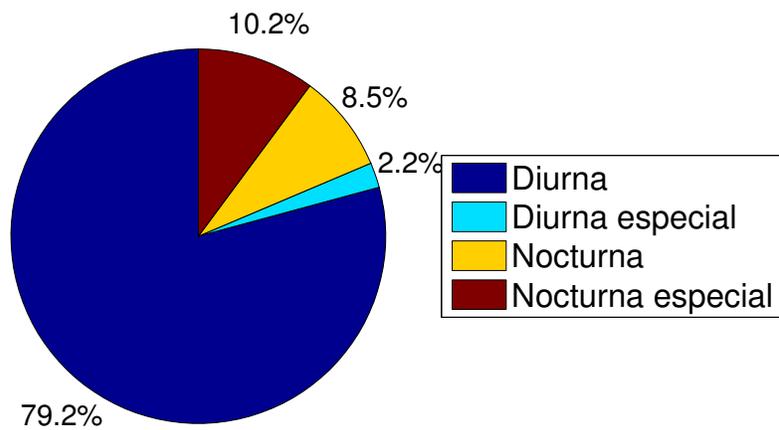
Figura 2.15: Conformación del subconjunto PAI por facultad

2.3.5. Subconjunto PAI por jornada

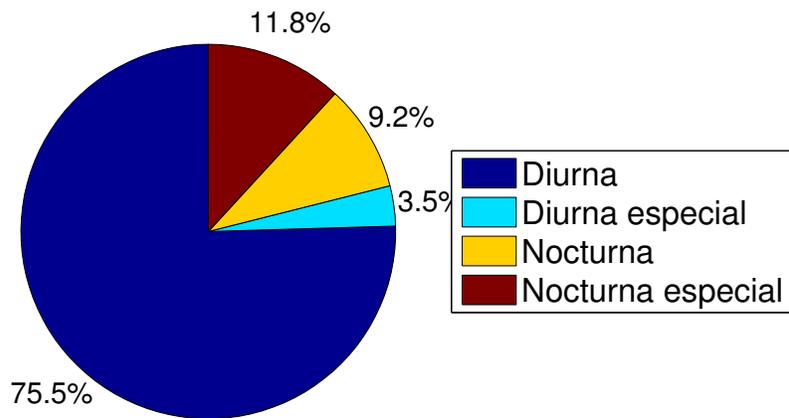
El subconjunto PAI está conformado por jornada como se presenta en la tabla 2.17 y en la figura 2.16.

Jornada	2016-I	2016-II	2017-I
Diurna	991 (79,22 %)	1.193 (75,46 %)	1.904 (75,95 %)
Diurna especial	27 (2,16 %)	55 (3,48 %)	94 (3,75 %)
Nocturna	106 (8,47 %)	146 (9,23 %)	297 (11,85 %)
Nocturna especial	127 (10,15 %)	187 (11,83 %)	212 (8,46 %)

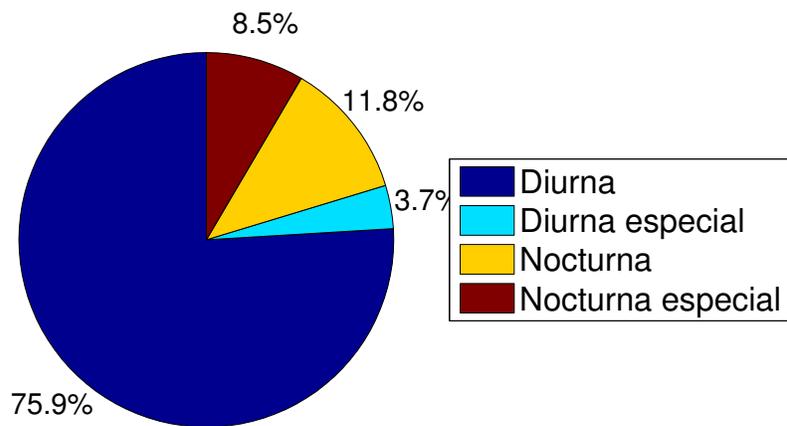
Tabla 2.17: Conformación del subconjunto PAI por jornada



(a) 2016-I



(b) 2016-II



(c) 2017-1

Figura 2.16: Conformación del subconjunto PAI por jornada

2.3.6. Subconjunto PAI por naturaleza del colegio

El subconjunto PAI está conformado por naturaleza del colegio como se presenta en la tabla 2.18 y en la figura 2.17.

Naturaleza colegio	2016-I	2016-II	2017-I
N/A	1 (0,08 %)	0 (0 %)	1 (0,04 %)
Oficial	1.049 (83,85 %)	1.312 (82,99 %)	2.155 (85,11 %)
Privado	201 (16,07 %)	269 (17,01 %)	376 (14,85 %)

Tabla 2.18: Conformación del subconjunto PAI por naturaleza del colegio

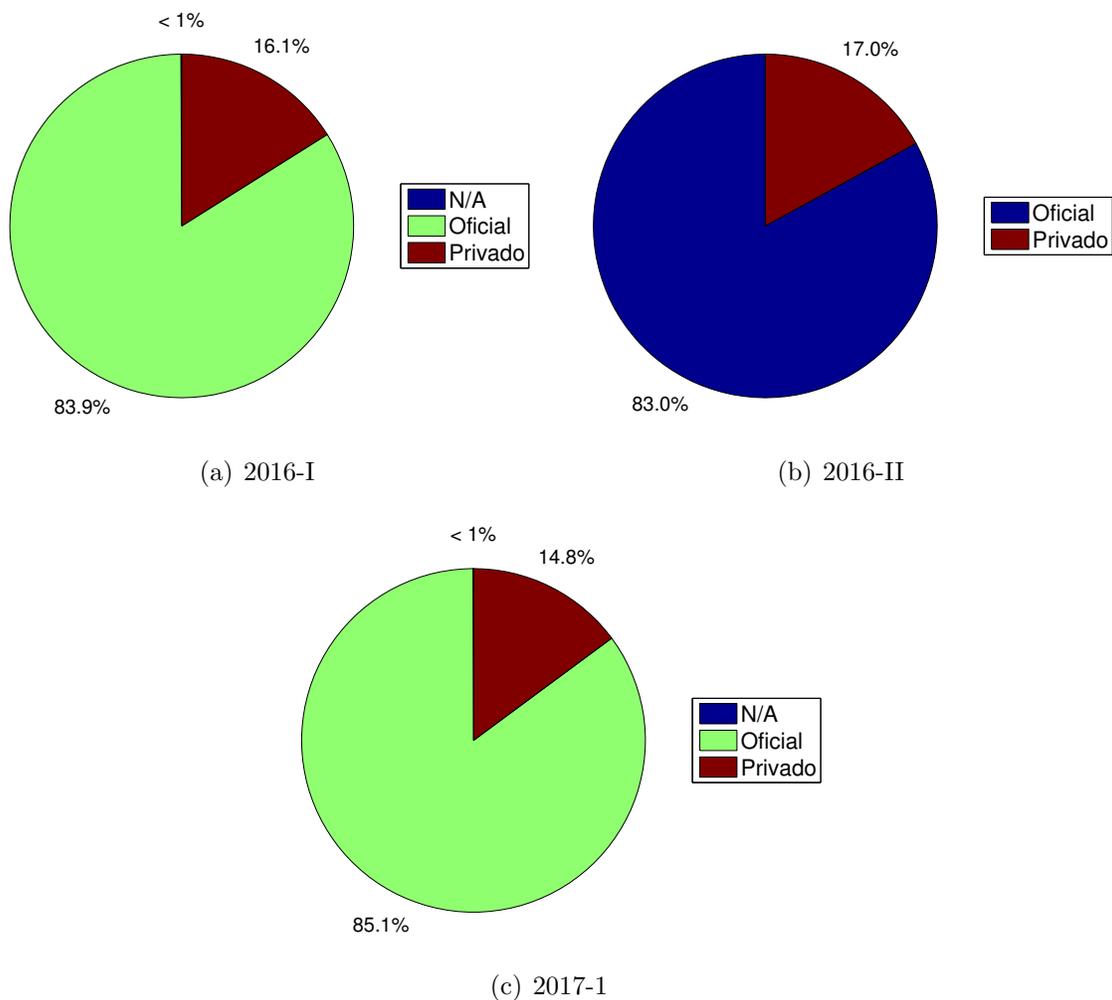


Figura 2.17: Conformación del subconjunto PAI por naturaleza del colegio

2.3.7. Subconjunto PAI por departamento

El subconjunto PAI está conformado por departamento de procedencia como se presenta en la tabla 2.19. Los códigos de cada categoría se encuentran en la sección A.7.

Departamento	2016-I	2016-II	2017-I
Amazonas	0 (0%)	0 (0%)	0 (0%)
Antioquia	5 (0,4%)	10 (0,63%)	7 (0,28%)
Arauca	4 (0,32%)	1 (0,06%)	1 (0,04%)
Atlántico	1 (0,08%)	3 (0,19%)	2 (0,08%)
Bogotá	19 (1,52%)	24 (1,52%)	28 (1,11%)
Bolívar	3 (0,24%)	0 (0%)	0 (0%)
Boyacá	0 (0%)	1 (0,06%)	3 (0,12%)
Caldas	39 (3,12%)	69 (4,36%)	108 (4,27%)
Caquetá	4 (0,32%)	7 (0,44%)	5 (0,2%)
Casanare	1 (0,08%)	2 (0,13%)	3 (0,12%)
Cauca	11 (0,88%)	4 (0,25%)	22 (0,87%)
Cesar	1 (0,08%)	0 (0%)	1 (0,04%)
Chocó	4 (0,32%)	3 (0,19%)	8 (0,32%)
Córdoba	2 (0,16%)	3 (0,19%)	2 (0,08%)
Cundinamarca	5 (0,4%)	10 (0,63%)	13 (0,51%)
Guaviare	2 (0,16%)	3 (0,19%)	3 (0,12%)
Huila	7 (0,56%)	18 (1,14%)	20 (0,79%)
Guajira	0 (0%)	1 (0,06%)	0 (0%)
Guainía	0 (0%)	2 (0,13%)	0 (0%)
Magdalena	3 (0,24%)	2 (0,13%)	3 (0,12%)
Meta	3 (0,24%)	1 (0,06%)	10 (0,39%)
N. de Santander	3 (0,19%)	28 (0,1756%)	6 (0,23%)
Nariño	49 (3,92%)	73 (4,62%)	110 (4,34%)
Putumayo	6 (0,48%)	14 (0,89%)	27 (1,07%)
Quindío	30 (2,4%)	33 (2,09%)	54 (2,13%)
Risaralda	886 (70,82%)	1.082 (68,44%)	1.802 (71,17%)
Santander	0 (0%)	5 (0,32%)	0 (0%)
San Andrés	1 (0,08%)	0 (0%)	2 (0,08%)
Sucre	0 (0%)	0 (0%)	3 (0,12%)
Tolima	16 (1,28%)	23 (1,45%)	27 (1,07%)
Valle del Cauca	136 (10,87%)	178 (11,26%)	252 (9,95%)
Vaupés	0 (0%)	0 (0%)	0 (0%)
Ext	1 (0%)	3 (0,19%)	2 (0,08%)
No disponible	10 (0,8%)	3 (0,19%)	8 (0,32%)

Tabla 2.19: Conformación del subconjunto PAI por departamento

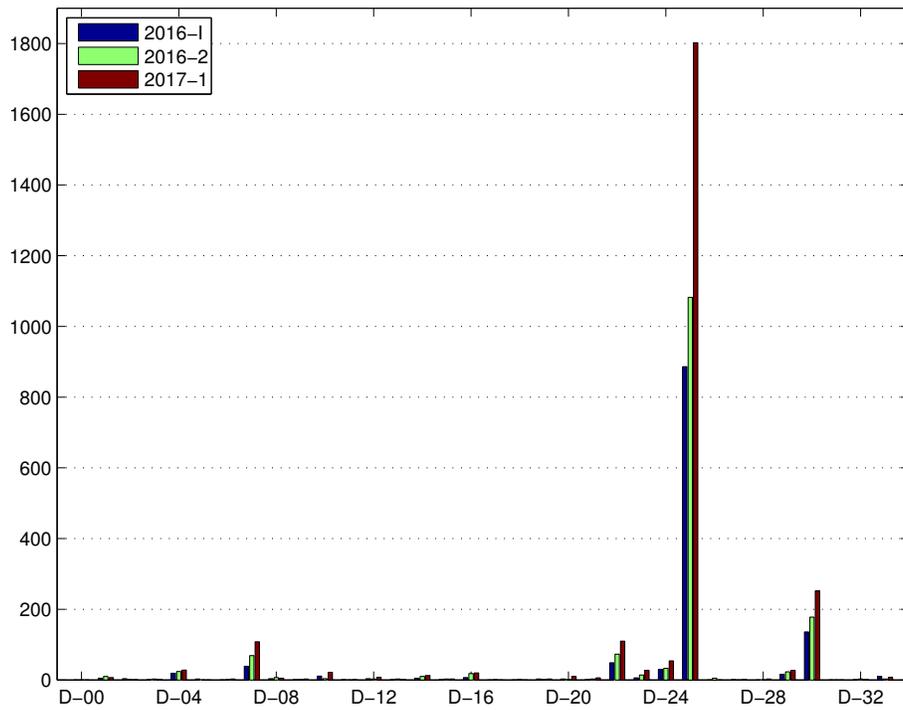


Figura 2.18: Conformación del subconjunto PAI por departamento

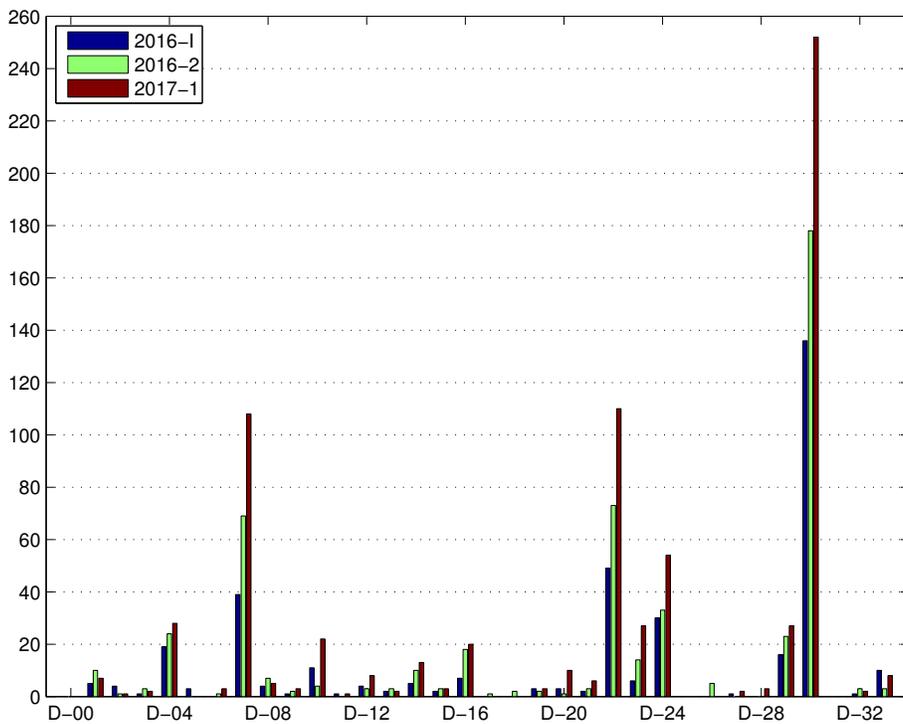


Figura 2.19: Conformación del subconjunto PAI por departamento excluyendo Risaralda

2.3.8. Subconjunto PAI por capital

El subconjunto PAI está conformado por capital de departamento como se presenta en la tabla 2.20 y en la figura 2.20.

Capital	2016-I	2016-II	2017-I
No	652 (52,12 %)	766 (48,45 %)	1.186 (46,84 %)
Sí	599 (47,88 %)	815 (51,55 %)	1.346 (53,16 %)

Tabla 2.20: Conformación del subconjunto PAI por capital

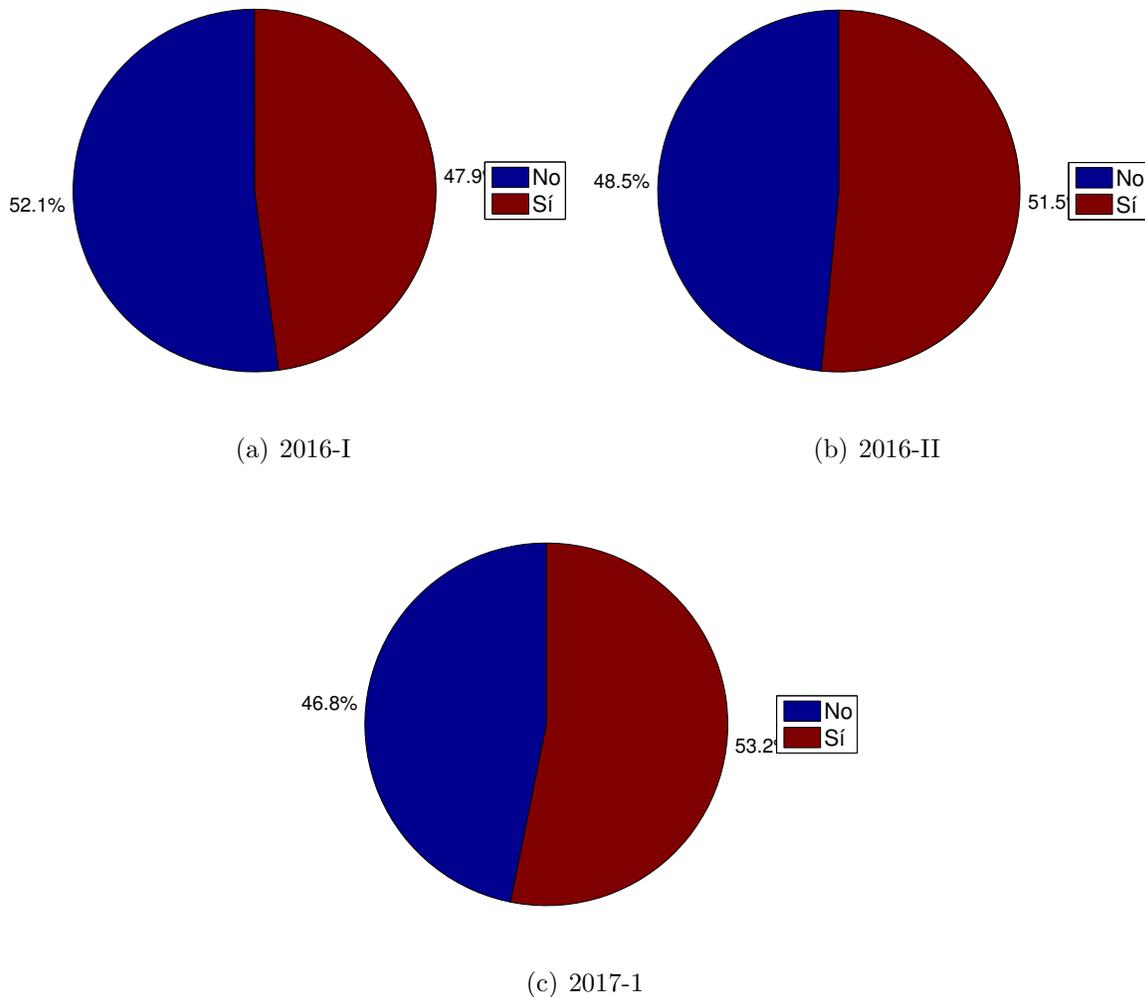


Figura 2.20: Conformación del subconjunto PAI por capital

2.3.9. Subconjunto PAI por estrato

El subconjunto PAI está conformado por estrato como se presenta en la tabla 2.21 y en la figura 2.21.

Estrato	2016-I	2016-II	2017-I
Bajo-bajo	317 (25,34 %)	432 (27,32 %)	654 (25,83 %)
Bajo	595 (47,56 %)	729 (46,11 %)	1.194 (47,16 %)
Medio-bajo	233 (18,63 %)	265 (16,76 %)	483 (19,08 %)
Medio	85 (6,79 %)	115 (7,27 %)	138 (5,45 %)
Medio-alto	11 (0,88 %)	33 (2,09 %)	51 (2,01 %)
Alto	9 (0,72 %)	6 (0,38 %)	10 (0,39 %)
No disponible	1 (0,08 %)	1 (0,06 %)	2 (0,08 %)

Tabla 2.21: Conformación del subconjunto PAI por estrato

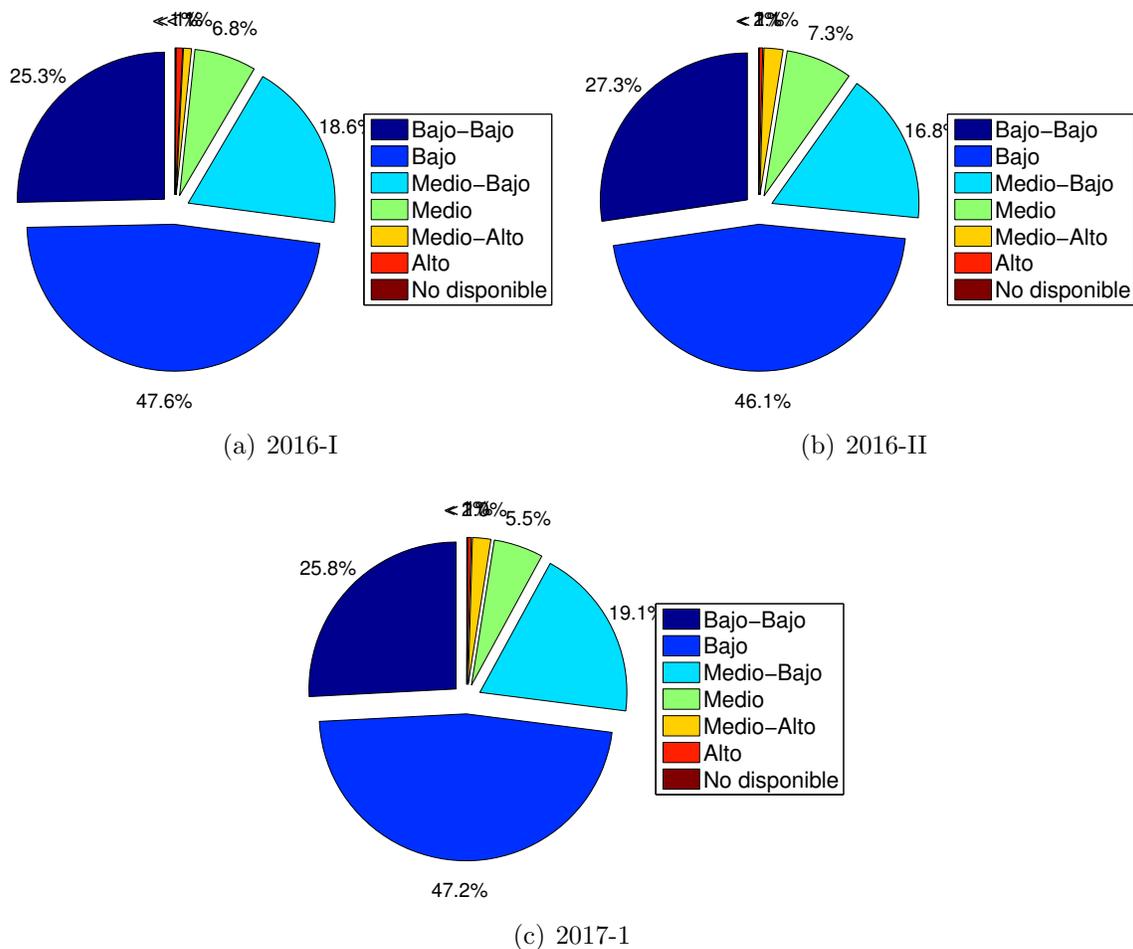


Figura 2.21: Conformación del subconjunto PAI por estrato

2.3.10. Subconjunto PAI por semestre de ubicación

La población está conformada por semestre de ubicación como se presenta en la tabla 2.22 y en la figura 2.22.

Semestre	2016-I	2016-II	2017-I
1	188 (15,67 %)	606 (38,33 %)	715 (28,24 %)
2	372 (31 %)	488 (30,87 %)	1.045 (41,27 %)
3	86 (7,17 %)	92 (5,82 %)	147 (5,81 %)
4	57 (4,75 %)	50 (3,16 %)	104 (4,11 %)
5	60 (5 %)	34 (2,15 %)	71 (2,8 %)
6	96 (8 %)	36 (2,28 %)	81 (3,2 %)
7	68 (5,67 %)	39 (2,47 %)	33 (1,3 %)
8	51 (4,25 %)	41 (2,59 %)	56 (2,21 %)
9	61 (5,08 %)	47 (2,97 %)	53 (2,09 %)
10	150 (12,5 %)	86 (5,44 %)	133 (5,25 %)
11	3 (0,25 %)	2 (0,13 %)	7 (0,28 %)
12	7 (0,58 %)	6 (0,38 %)	12 (0,47 %)
13	1 (0,08 %)	1 (0,06 %)	2 (0,08 %)
No disponible	0 (0 %)	53 (3,35 %)	73 (2,88 %)

Tabla 2.22: Conformación del subconjunto PAI por semestre de ubicación

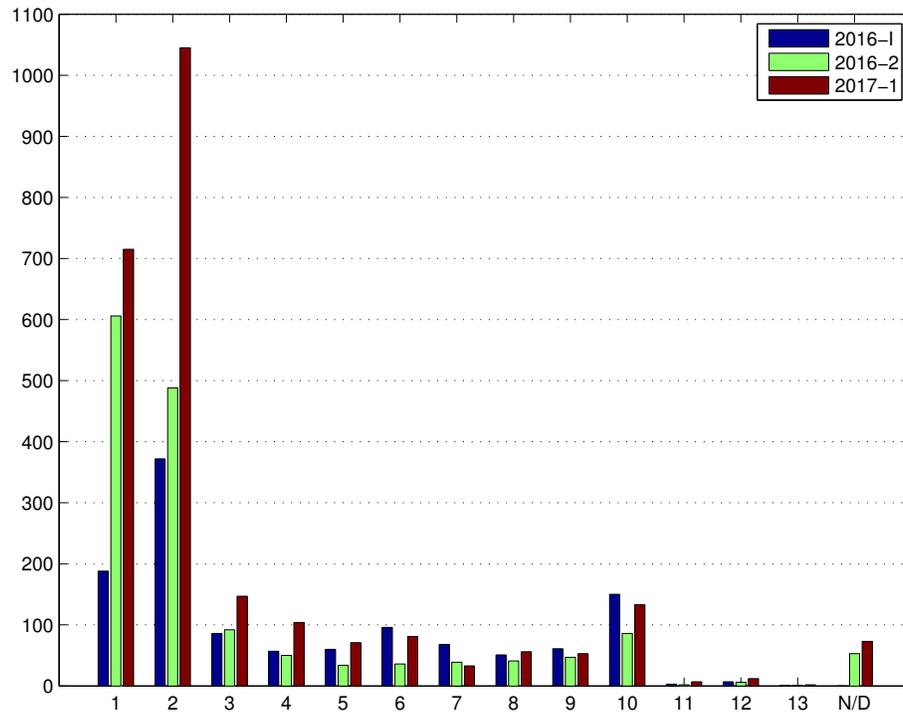


Figura 2.22: Conformación del subconjunto PAI por semestre de ubicación

2.3.11. Comparación variables descriptivas

En esta sección se compara el histograma porcentual del conjunto poblacional total con el del subconjunto PAI para cada una de las variables de interés en los períodos de análisis. Es decir, se calcula la norma cuadrática de la diferencia entre las variables para los conjuntos de matriculados y atendidos PAI, obteniendo un error para cada variable.

En tal sentido y como ejemplo, sea $\mathbf{g}_{M2016-I}$ el vector que contiene la participación porcentual de las categorías de la variable género para el conjunto poblacional de matriculados en 2016-I, y sea $\mathbf{g}_{PAI2016-I}$ el vector que contiene la participación porcentual de las categorías de la variable género para el subconjunto de atendidos PAI en 2016-I. Se procede primero a calcular la diferencia entre los vectores antes señalados y posteriormente se calcula su norma, tal como se presenta en la ecuación (2.1),

$$\mathbf{e}_{\mathbf{g}_{2016-I}} = \|\mathbf{g}_{M2016-I} - \mathbf{g}_{PAI2016-I}\|. \quad (2.1)$$

En la tabla 2.23 se presenta el cálculo de error para cada una de las variables en los períodos de análisis. En tal sentido, los valores más bajos indican que el subconjunto PAI se comporta como un muestreo uniforme del conjunto poblacional, por otra parte, a medida que los valores de error aumentan el subconjunto PAI tiene una distribución diferente a la del conjunto poblacional. Este último caso advierte de un posible conjunto

de variables relevantes para la prestación de la atención en el programa PAI puesto que identifican condiciones singulares por las cuales asisten los estudiantes al PAI.

Error variables	2016-I	2016-II	2017-I	Promedio
Género	4,24 %	7,24 %	3,38 %	5,74 %
Rango de edad	14,31 %	22,22 %	28,91 %	18,27 %
Programa	9,13 %	7,43 %	8,92 %	8,28 %
Facultad	8,63 %	8,27 %	8,15 %	8,45 %
Jornada	7,65 %	3,73 %	7,15 %	5,69 %
Naturaleza colegio	1,31 %	0,38 %	2,83 %	0,85 %
Departamento	2,32 %	3,21 %	1,07 %	2,76 %
Capital	2,12 %	2,96 %	2,65 %	2,54 %
Estrato	5,66 %	5,45 %	4,2 %	5,55 %
Semestre ubicación	21,28 %	40,65 %	37,22 %	30,97 %

Tabla 2.23: Error de comparación de los histogramas de la población total y del subconjunto PAI

2.4. Deserción universitaria

Según el Ministerio de Educación, la deserción estudiantil se entiende como el “estado de un estudiante que de manera voluntaria o forzosa no registra matrícula por dos o más períodos académicos consecutivos del programa en el que se matriculó; y no se encuentra como graduado, o retirado por motivos disciplinarios”. La deserción es el resultado del efecto de diferentes factores como individuales, académicos, institucionales, y socio-económicos.

En este documento se analiza la deserción tanto el caso en que el estudiante no registra matrícula en uno y en dos periodos consecutivos, a lo que se denomina deserción de un período y deserción de dos períodos, respectivamente. Adicionalmente, se realiza la evaluación pormenorizada de deserción de uno y dos períodos para el subconjunto PAI.

2.4.1. Deserción de un período

Para calcular la deserción de un período entre 2016-I y 2016-II se tienen los conjuntos: M_{2016_I} , $M_{2016_{II}}$, G_{2016_I} y $G_{2016_{II}}$; estos últimos corresponde al grupo de estudiantes graduados en el 2016-I y 2016-II, respectivamente. Los conjuntos mencionados se ejemplifican en la figura 2.23.

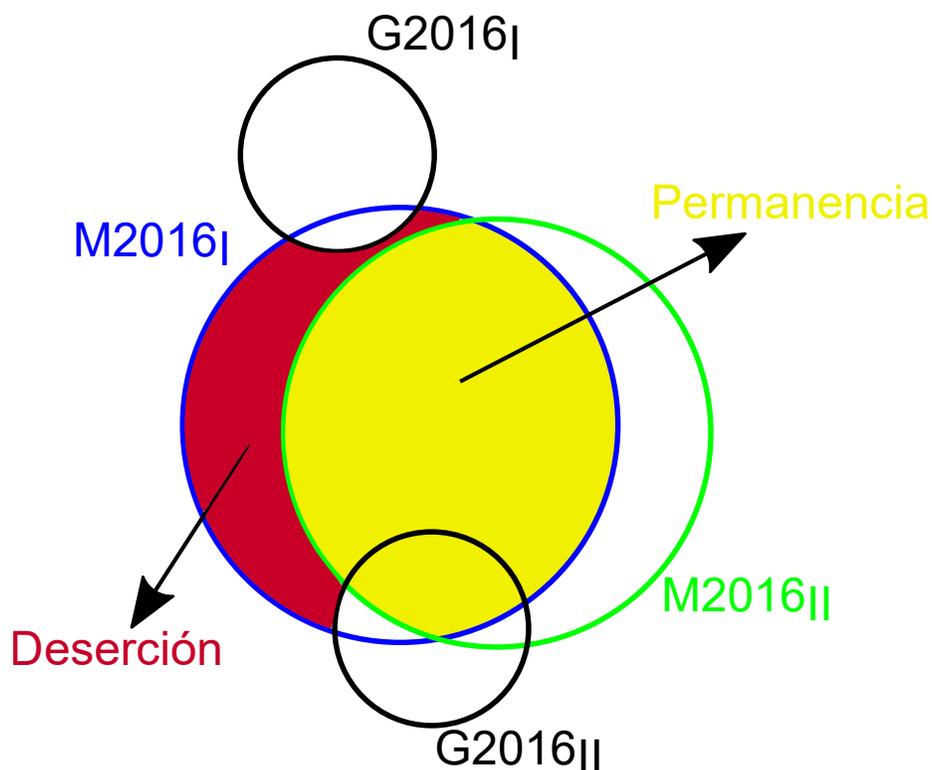


Figura 2.23: Conjuntos para evaluar permanencia y deserción de un período

A partir de estos conjuntos se puede calcular la permanencia (PER_a) como la intersección entre los conjuntos $M2016_I$ y $M2016_{II}$, que se muestra en la ecuación (2.2), y que corresponde con el conjunto de estudiantes que se matricularon en el 2016-I y también se matricularon en el 2016-II.

$$PER_a = M2016_I \cap M2016_{II} \quad (2.2)$$

y el número de elementos del conjunto PER_a es $n(PER_a) = 13.685$

Para establecer el conjunto deserción (DES_a), es necesario calcular la ecuación (2.3),

$$DES_a = ((M2016_I \setminus PER_a) \setminus G2016_I) \setminus G2016_{II} \quad (2.3)$$

y el número de elementos del conjunto DES_a es $n(DES_a) = 1.797$. Además la deserción como razón porcentual corresponde con

$$D_a = 100\% \times \frac{n(DES_a)}{n(M2016_I)} = 100\% \times \frac{1.797}{16.826} = 10,64\% \quad (2.4)$$

De igual manera se calcula la permanencia (PER_b) y la deserción (DES_b) entre los semestres 2016-II y 2017-I, obteniendo que $n(PER_b) = 13.354$, $n(DES_b) = 1.706$ y $D_b = 10,66\%$. Los cálculos anteriores se resumen en la tabla 2.24.

	2016-I a 2016-II	2016-II a 2017-I	Variación
$n(M)$	16.826	15.998	-4,92%
$n(PER)$	13.685	13.354	-2,32%
$n(DES)$	1.797	1.706	-5,06%
Deserción	10,64%	10,66%	0,19%

Tabla 2.24: Permanencia y deserción de un período

2.4.2. Deserción de dos períodos

Para calcular la deserción de dos períodos entre 2016-I y 2017-I se tienen en cuenta los conjuntos: $M2016_I$, $M2016_{II}$, $M2017_I$, $G2016 - I$, $G2016_{II}$ y $G2017 - I$. La representación gráfica de estos conjuntos se presenta en la figura 2.24.

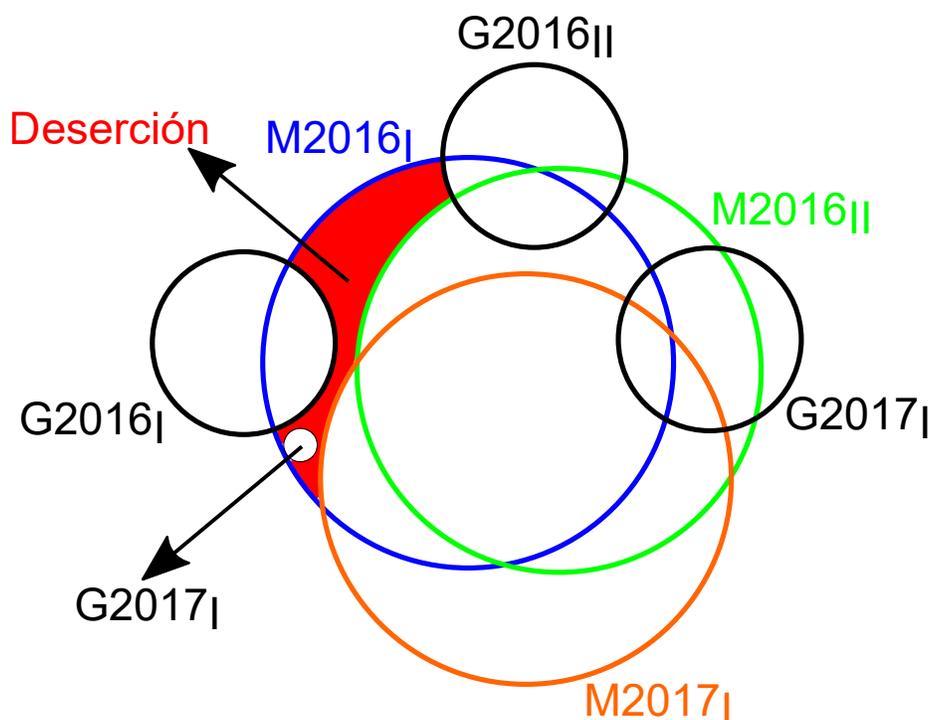


Figura 2.24: Conjuntos para evaluar permanencia y deserción de dos períodos

Para establecer el conjunto deserción de dos períodos (DES_c), es necesario calcular la ecuación (2.5),

$$DES_c = (((M2016_I \setminus M2016_{II}) \setminus M2017_I) \setminus G2016_I) \setminus G2016_{II} \setminus G2017_I \quad (2.5)$$

Así, el número de elementos del conjunto $n(DES_c) = 1.382$. Además la deserción como razón porcentual corresponde con

$$D_c = 100\% \times \frac{n(DES_c)}{n(M2016_I)} = 100\% \times \frac{1.382}{16.826} = 8,21\% \quad (2.6)$$

2.4.3. Deserción de un período en el subconjunto PAI

Ahora se considera en el análisis sólo el subconjunto poblacional PAI, a partir del cual se calcula el conjunto deserción de un período entre 2016-I y 2016-II, ver ecuación (2.7). En este caso se tienen en cuenta los conjuntos $PAI2016_I$, $M2016_{II}$, $G2016_I$ y $G2016_{II}$.

$$DES_d = (((PAI2016_I \setminus M2016_{II})) \setminus G2016_I) \setminus G2016_{II} \quad (2.7)$$

Así, el número de elementos del conjunto $n(DES_d) = 131$. Además la deserción como razón porcentual corresponde con

$$D_d = 100\% \times \frac{n(DES_d)}{n(PAI2016_I)} = 100\% \times \frac{131}{1.251} = 10,47\% \quad (2.8)$$

Para el período entre 2016-II y 2017-I, ver ecuación (2.9). En este caso se tienen en cuenta los conjuntos $PAI2016_{II}$, $M2017_I$, $G2016_{II}$ y $G2017_I$.

$$DES_e = (((PAI2016_{II} \setminus M2017_I)) \setminus G2016_{II}) \setminus G2017_I \quad (2.9)$$

Así, el número de elementos del conjunto $n(DES_e) = 255$. Además la deserción como razón porcentual corresponde con

$$D_e = 100\% \times \frac{n(DES_e)}{n(PAI2016_{II})} = 100\% \times \frac{255}{1.581} = 16,13\% \quad (2.10)$$

Los cálculos anteriores se resumen en la tabla 2.25.

	2016-I a 2016-II	2016-II a 2017-I	Variación
$n(PAI)$	1.251	1.581	26,38%
$n(DES)$	131	255	94,66%
Deserción	10,47%	16,13%	54,06%

Tabla 2.25: Permanencia y deserción de un período para grupo PAI

2.4.4. Deserción de dos períodos en el subconjunto PAI

Teniendo en consideración sólo el subconjunto poblacional PAI, para calcular la deserción de dos períodos entre 2016-I y 2017-I se tienen en cuenta los conjuntos: $PAI2016_I$, $M2016_{II}$, $M2017_I$, $G2016 - I$, $G2016_{II}$ y $G2017 - I$.

Para establecer el conjunto deserción de dos períodos (DES_f), es necesario calcular la ecuación (2.11),

$$DES_f = (((((PAI2016_I \setminus M2016_{II}) \setminus M2017_I) \setminus G2016_I) \setminus G2016_{II}) \setminus G2017_I \quad (2.11)$$

Así, el número de elementos del conjunto $n(DES_f) = 103$. Además la deserción como razón porcentual corresponde con

$$D_f = 100\% \times \frac{n(DES_f)}{n(PAI2016_I)} = 100\% \times \frac{103}{1.251} = 8,23\% \quad (2.12)$$

2.5. Evaluación de probabilidades condicionales

Con el fin de determinar la influencia que tienen las categorías dentro de cada una de las variables se calculan las probabilidades condicionales para el grupo PAI y para el grupo desertor (de uno y dos períodos), en los semestres de interés.

2.5.1. Análisis para el subconjunto poblacional: PAI

Se considera el subconjunto de los estudiantes que fueron atendidos en el PAI. Con este subconjunto se determina la probabilidad de ser atendido en el PAI dada la aparición de cada una de las categorías de las variables, es decir:

$$P(PAI | variable_i) = \frac{P(variable_i | PAI)P(PAI)}{P(variable_i)} = \frac{P(PAI \cap variable_i)}{P(variable_i)}. \quad (2.13)$$

Género

En la tabla 2.26 se presentan las probabilidades condicionales atención PAI dado el género del estudiante. Como se puede apreciar, la probabilidad de asistir al PAI del género femenino es mayor que la del masculino durante los tres semestres estudiados. Además, para el semestre 2017-I la probabilidad de asistir al PAI se incrementó, para ambos géneros con respecto al semestre anterior, en aproximadamente un 6%; y en aproximadamente un 8% con respecto al 2016-I.

Género	$P(PAI Género)$		
	2016-I	2016-II	2017-I
Masculino	7,05 %	9,01 %	14,8 %
Femenino	7,96 %	11,08 %	16,31 %

Tabla 2.26: Probabilidad condicional: $P(PAI|Género)$

Rango de edad

En la tabla 2.27 se muestran las probabilidades condicionales con respecto a la variable: rango de edad. En este sentido, es posible apreciar que la probabilidad de ser atendido en el PAI es mayor para el conjunto de estudiantes con edades entre los 15 a 17 años. Además, en el 2017-I el rango mencionado anteriormente tuvo un alza considerable, pasando del 22,17% en 2016-I al 67,29% en el 2017-I.

Rango Edad	$P(PAI \text{RangoEdad})$		
	2016-I	2016-II	2017-I
0 a 10	0 %	0 %	0 %
11 a 14	0 %	0 %	0 %
15 a 17	22,17 %	30,29 %	67,29 %
18 a 19	8,37 %	17,57 %	23,68 %
20 a 21	6,84 %	8,7 %	10,4 %
22 a 23	7,29 %	5,46 %	8,13 %
24 a 25	4,5 %	6,13 %	7,14 %
26 a 29	4,53 %	4,87 %	6,68 %
30 a 35	4,41 %	6,09 %	6,35 %
36 a 39	0,98 %	3 %	3,38 %
40 o más	4,13 %	4,76 %	5,33 %

Tabla 2.27: Probabilidad condicional: $P(PAI|Rangodeedad)$

Programa

La tabla 2.28 muestra las probabilidades condicionales del PAI relacionadas con el programa académico que cursan los estudiantes. Se puede notar que los programas para los que la probabilidad de participar en el PAI es mayor son los siguientes:

- Para el 2016-I: P-18 y P-27.
- Para el 2016-II: P-10, P-17, P-37, P-41, P-49, P-50, P-52, P-53 y P-60.
- Para el 2017-I: P-00, P-06, P-09, P-10, P-12, P-16, P-17, P-20, P-22, P-32, P-33, P-38, P-39, P-40, P-41, P-50, P-52, P-53, P60, P-61 y P-62.

De particular interés son los programas P-17, P-20 y P-33 que tienen una probabilidad mayor al 10 % en los tres semestres estudiados.

Programa	$P(PAI Programa)$		
	2016-I	2016-II	2017-I
P-00	5,76 %	10,01 %	17,65 %
P-01	0 %	0,96 %	0 %
P-02	8,33 %	0 %	0 %
P-03	3,23 %	0 %	0 %
P-04	9,09 %	0 %	0 %
P-05	0 %	6,67 %	1,82 %
P-06	0 %	0 %	100 %
P-07	9,83 %	11,77 %	13,42 %
P-08	0 %	0 %	0 %
P-09	4,95 %	11,49 %	14,81 %
P-10	3,02 %	16,05 %	16,55 %
P-11	0 %	0 %	0 %
P-12	0 %	0 %	54,35 %
P-13	4,97 %	7,56 %	9,75 %
P-14	0 %	0 %	0 %
P-15	4,12 %	1,33 %	2,54 %
P-16	0 %	8,7 %	64,71 %
P-17	9,65 %	18,61 %	19,93 %
P-18	15,6 %	0,65 %	0,5 %
P-19	5,26 %	2,4 %	0 %
P-20	13,68 %	12,5 %	57,91 %
P-21	0 %	0 %	0 %
P-22	0 %	0 %	71,93 %
P-23	0 %	0 %	0 %
P-24	8,23 %	8,58 %	10,53 %
P-25	2,55 %	4,8 %	7,61 %
P-26	0 %	0 %	0 %
P-27	16,65 %	7,71 %	13,57 %
P-28	0 %	0 %	0 %
P-29	0 %	0 %	0 %
P-30	7,19 %	9,4 %	13,92 %
P-31	0 %	0 %	0 %
P-32	10,14 %	4,09 %	19,38 %
P-33	10,29 %	13,14 %	16,8 %
P-34	0 %	0 %	0 %
P-35	5,19 %	12,99 %	13,2 %
P-36	7,01 %	2,61 %	6,2 %
P-37	0 %	58,82 %	13,59 %
P-38	5,7 %	2,13 %	23,08 %
P-39	7,59 %	2,79 %	25,83 %
P-40	9,59 %	3,17 %	21,62 %
P-41	6,57 %	14,36 %	16,09 %
P-42	0 %	0 %	0 %
P-43	0 %	0 %	0 %
P-44	0 %	0 %	0 %
P-45	0 %	0 %	0 %
P-46	0 %	0 %	0 %
P-47	8,17 %	8,44 %	13,86 %
P-48	5,81 %	10,97 %	10,12 %
P-49	4,03 %	18,18 %	6,9 %
P-50	7,54 %	14,42 %	19,71 %
P-51	0 %	0 %	0 %
P-52	9,5 %	18,71 %	20 %
P-53	8,06 %	14,36 %	17,52 %
P-54	0 %	0 %	0 %
P-55	0 %	0 %	0 %
P-56	0 %	0 %	0 %
P-57	0 %	0 %	0 %
P-58	0 %	0 %	0 %
P-59	6,98 %	3,51 %	2,22 %
P-60	7,25 %	17,11 %	20,05 %
P-61	7,99 %	4,48 %	30,33 %
P-62	0 %	8,71 %	27,93 %

Tabla 2.28: Probabilidad condicional: $P(PAI | Programa)$.

Facultad

En la tabla 2.29 se describen las probabilidades condicionales de haber sido atendido por el PAI dada la facultad a la que pertenezcan los estudiantes. Es pertinente resaltar que todas las facultades muestran una valores semejantes, con excepción de la facultad de ciencias agrarias y agroindustriales que no existía para los 2 primeros semestres y que en el semestre de apertura la mayoría de sus estudiantes fueron atendidos por el PAI. Finalmente, es preciso resaltar que la facultad de ingeniería industrial es la que presenta una menor influencia del PAI en promedio para los tres semestres estudiados.

Facultad	$P(PAI Facultad)$		
	2016-I	2016-II	2017-I
Bellas artes y humanidades	8,3%	6,05%	18,16%
Ciencias ambientales	6,29%	9,4%	19,89%
Ciencias básicas	7,59%	2,79%	25,83%
Ciencias de la educación	6,73%	11,79%	13,06%
Ciencias de la salud	8,33%	10,99%	13,38%
Ingeniería industrial	5,97%	7,06%	9,22%
Ingeniería mecánica	16,6%	7,69%	13,26%
Ingenierías	5,79%	9,77%	19,57%
Tecnología	7,05%	12,7%	15,03%
Ciencias agrarias y agroindustria	0%	0%	71,93%

Tabla 2.29: Probabilidad condicional: $P(PAI|Facultad)$.

Jornada

Con respecto a la variable Jornada, según la tabla 2.30, la probabilidad de ser atendido en el PAI dado que el estudiante pertenezca a la jornada diurna es mayor que cuando el estudiante pertenece a la jornada nocturna en aproximadamente 2% para los períodos 2016-I y 2016-II, sin embargo, para el 2017-I son estadísticamente iguales.

Jornada	$P(PAI Jornada)$		
	2016-I	2016-II	2017-I
Diurna	8,08%	10,3%	16,28%
Diurna especial	5,81%	10,87%	16,73%
Nocturna	6,04%	8,63%	16,36%
Nocturna especial	5,42%	8,56%	9,27%

Tabla 2.30: Probabilidad condicional: $P(PAI|Jornada)$.

Naturaleza del colegio

Con relación a la naturaleza del colegio, las probabilidades de ser atendido en el PAI dada la naturaleza del colegio de procedencia de los estudiantes han sido similares en cada uno de los semestres. Es de resaltar que para 2017-I hay un claro incremento en la probabilidad de todos los escenarios (ver tabla 2.31), incluyendo la atención de estudiantes con información desconocida.

Naturaleza colegio	$P(\text{PAI} \text{NaturalezaColegio})$		
	2016-I	2016-II	2017-I
N/A	4,55 %	0 %	8,33 %
Oficial	7,52 %	9,92 %	15,81 %
Privado	7,04 %	9,74 %	13,62 %

Tabla 2.31: Probabilidad condicional: $P(\text{PAI} | \text{NaturalezaColegio})$.

Departamento

Con base a los resultados presentados en la tabla 2.32, la probabilidad de ser atendido por el PAI es mayor para los departamentos de: Magdalena, Arauca y Guaviare para los tres semestres estudiados. Esto implica que la mayoría de los estudiantes provenientes de estos departamentos han tenido atención por el PAI. Ahora, la probabilidad de ser atendido en el PAI dado que el estudiante procede del departamento de Risaralda ha sido del 7,27 %, 9,51 % y 15,47 % para el 2016-I, 2016-II y 2017-I, respectivamente. Estas cifras son relativamente bajas dada la alta composición de estudiantes de la UTP que proceden de este departamento.

Departamento	$P(PAI Departamento)$		
	2016-I	2016-II	2017-I
Amazonas	0 %	0 %	0 %
Antioquia	7,14 %	14,08 %	9,72 %
Arauca	80 %	16,67 %	16,67 %
Atlántico	5,88 %	13,04 %	9,09 %
Bogotá	10,56 %	13,41 %	16,37 %
Bolívar	18,75 %	0 %	0 %
Boyacá	0 %	7,14 %	23,08 %
Caldas	5,59 %	10,19 %	15,72 %
Caquetá	6,9 %	11,86 %	8,33 %
Casanare	12,5 %	18,18 %	30 %
Cauca	9,09 %	3,51 %	18,49 %
Cesar	7,69 %	0 %	10 %
Chocó	4,94 %	3,8 %	11,76 %
Córdoba	15,38 %	21,43 %	13,33 %
Cundinamarca	7,14 %	13,7 %	18,57 %
Guaviare	18,18 %	23,08 %	23,08 %
Huila	6,19 %	14,52 %	15,38 %
Guajira	0 %	25 %	0 %
Guainía	0 %	100 %	0 %
Magdalena	42,86 %	33,33 %	50 %
Meta	9,38 %	3,7 %	27,78 %
N. de Santander	8,7 %	12,5 %	24 %
Nariño	9,57 %	13,11 %	17,71 %
Putumayo	4,92 %	10,85 %	20,3 %
Quindío	8,24 %	9,22 %	14,71 %
Risaralda	7,27 %	9,51 %	15,47 %
Santander	0 %	26,32 %	0 %
San Andrés	2,86 %	0 %	4,65 %
Sucre	0 %	0 %	50 %
Tolima	10,96 %	15,33 %	16,98 %
Valle del Cauca	7,83 %	10,51 %	14,47 %
Vaupés	0 %	0 %	0 %
Ext	6,25 %	25 %	12,5 %
No disponible	8,55 %	3,19 %	8,6 %

Tabla 2.32: Probabilidad condicional: $P(PAI|Departamento)$.

Capital

De la tabla 2.33 se observa un incremento a lo largo de los 3 semestres analizados de la cantidad de estudiantes atendidos por el PAI, sin marcadas diferencias entre la

probabilidad de ser atendido dado que el estudiante proceda de una ciudad capital o no.

Capital	$P(PAI Capital)$		
	2016-I	2016-II	2017-I
No	7,66 %	9,47 %	14,22 %
Sí	7,21 %	10,3 %	16,69 %

Tabla 2.33: Probabilidades condicionales $P(PAI|Capital)$.

Estrato

Para el 2017-I más del 16 % de los estudiantes de estrato Bajo-bajo y más del 10 % de estrato Alto son atendidos por el PAI, según se muestra en la tabla 2.34. Además, se observa que la probabilidad condicional aumenta considerablemente para el semestre 2017-I en todos los estratos. En particular, el estrato Medio-alto registra un incremento del 13 % aproximadamente entre el 2016-I y el 2017-I.

Estrato	$P(PAI Estrato)$		
	2016-I	2016-II	2017-I
Bajo-bajo	8,09 %	11,24 %	16,47 %
Bajo	8,08 %	10,3 %	16,28 %
Medio-bajo	6,47 %	8,14 %	14,52 %
Medio	5,63 %	8,15 %	10,28 %
Medio-alto	3,53 %	10,61 %	16,09 %
Alto	9,18 %	6,38 %	10,2 %
No disponible	4,35 %	12,5 %	14,29 %

Tabla 2.34: Probabilidad condicional: $P(PAI|Estrato)$

Semestre de ubicación

La mayoría de los estudiantes de los primeros 2 semestres son atendidos por el PAI, sin embargo, sólo el 5 % o menos de los estudiantes de los últimos semestres son atendidos en el programa (ver tabla 2.35).

Semestre	$P(PAI Semestre)$		
	2016-I	2016-II	2017-I
1	16,49	51,4 %	58,51 %
2	16,02 %	26,22 %	42,64 %
3	5,22 %	5,26 %	9,45 %
4	3,48 %	3,64 %	6,57 %
5	4,89 %	2,35 %	5,95 %
6	4,73 %	2,26 %	4,89 %
7	5,57 %	3,4 %	3,12 %
8	5,14 %	3,37 %	5,02 %
9	6,3 %	4,98 %	4,5 %
10	6,17 %	3,62 %	5,91 %
11	1,4 %	1 %	2,98 %
12	1,94 %	2 %	4,4 %
13	1,43 %	1,54 %	2,41 %
No disponible	9,06 %	9,69 %	13,27 %

Tabla 2.35: Probabilidad condicional: $P(PAI|Semestre)$

Número de períodos matriculados

En concordancia con la variable anterior, son atendidos aproximadamente el 32 % de los estudiantes en el rango de 0 a 3 períodos matriculados. Por otra parte, son atendidos menos del 7 % de los estudiantes en el rango de 8 a 11 semestres matriculados. Adicionalmente, los estudiantes con más de 15 períodos matriculados no son vistos por el PAI (ver tabla 2.36).

Períodos	$P(PAI Periodos)$		
	2016-I	2016-II	2017-I
0 a 3	11,74 %	22,44 %	32,02 %
4 a 7	4,92 %	2,85 %	5,29 %
8 a 11	6,38 %	3,89 %	6,31 %
12 a 15	4,37 %	3,37 %	4,14 %
16 a 19	2,78 %	2,47 %	2,98 %
20 a 23	3,76 %	0,83 %	3,85 %
24 a 27	0 %	0 %	0 %
28 a 31	9,09 %	0 %	0 %
32 o más	0 %	0 %	0 %

Tabla 2.36: Probabilidad condicional $P(PAI|Periodos)$

Créditos acumulados

De manera semejante a las dos variables antes analizadas, la variable créditos acumulados de la tabla 2.37 nos indica que la mayoría de estudiantes sobre los que ejerce atención el PAI son de primeros semestres, pero los estudiantes que van más adelante y tienen más créditos acumulados son un foco menor para el PAI.

Créditos	$P(\text{PAI} \text{Créditos})$		
	2016-I	2016-II	2017-I
0 a 10	17,39 %	48,25 %	55,46 %
11 a 20	16,9 %	27,71 %	45,42 %
21 a 30	4,99 %	5,64 %	12,42 %
31 a 40	5,53 %	4,73 %	9,1 %
41 a 50	2,79 %	4,68 %	7,02 %
51 a 60	4,37 %	3,04 %	5,97 %
61 a 70	4,12 %	2,19 %	5,07 %
71 a 80	4,34 %	2,97 %	4,87 %
81 a 90	6,88 %	2,5 %	5,99 %
91 a 100	6,59 %	2,14 %	4,53 %
101 a 110	5,74 %	2,14 %	4,71 %
111 a 120	3,45 %	3,94 %	2,43 %
121 a 130	5,26 %	4,22 %	4,54 %
131 a 140	6,39 %	3,06 %	7,38 %
141 a 150	6,34 %	4,55 %	5,78 %
151 a 160	6,52 %	3,72 %	3,89 %
161 a 170	5,39 %	3,56 %	5,03 %
171 a 180	4,71 %	3,44 %	6,18 %
181 a 190	4,49 %	2,92 %	5,22 %
191 a 200	5,39 %	2,11 %	3,87 %
200 o más	2,92 %	3,51 %	4,81 %

Tabla 2.37: Probabilidad condicional $P(\text{PAI} | \text{Creditos})$

Promedio acumulado por rangos

Coherente con los propósitos del PAI, la mayor cantidad de estudiantes que son atendidos en el programa son los de promedios más bajos (tabla 2.38). Los estudiantes del rango 0 a 0,99 son en su mayoría de primer semestre que no cuentan con un promedio registrado.

Promedio	$P(PAI Promedio)$		
	2016-I	2016-II	2017-I
0 a 0,99	14,07 %	23,94 %	56,96 %
1 a 1,99	12,61 %	34,96 %	6,77 %
2 a 2,99	8,19 %	18,7 %	7 %
3 a 3,99	6,5 %	7,73 %	5,64 %
4 a 5	7,33 %	7,98 %	7,65 %

Tabla 2.38: Probabilidad condicional $P(PAI|Promedio)$

2.5.2. Análisis para el subconjunto poblacional: Desertor un período

Se considera el subconjunto de los estudiantes que presentan deserción en un período, según como se determinó en la sección 2.4.1. Con este subconjunto se determina la probabilidad de ser desertor (o no desertor) dada la aparición de cada una de las categorías de las variables estudiadas. Por ejemplo:

$$P(Desert | variable_i) = \frac{P(variable_i | Desert)P(Desert)}{P(variable_i)} = \frac{P(Desert \cap variable_i)}{P(variable_i)}. \quad (2.14)$$

Género

La tabla 2.39 presenta las probabilidades condicionales de deserción en un período dado el género del estudiante. Se puede notar que las probabilidades de deserción del género masculino son mayores que las del femenino en los dos semestres analizados. Además, se evidencia una disminución considerable en la probabilidad de deserción entre el 2016-I y 2016-II, alrededor del 7 %, para ambos géneros.

Género	2016-I		2016-II	
	$P(DES G)$	$P(NODES G)$	$P(DES G)$	$P(NODES G)$
Masculino	18,02 %	81,98 %	11,71 %	88,29 %
Femenino	16,29 %	83,71 %	8,27 %	91,73 %

Tabla 2.39: Probabilidad condicional: $P(DES/NODES|Genero)$.

Rango de edad

En la tabla 2.40 se muestran las probabilidades condicionales con respecto a la variable rango de edad. De este modo, cabe resaltar que la probabilidad deserción para el 2016-I es mayor para el conjunto de estudiantes entre los 36 a 39 años, seguida por la del grupo de 30 a 35 años. Se observa que la probabilidad de deserción es mayor al 20 %

para estudiantes con edades superiores a los 24 años. Para el 2016-II las probabilidades de deserción disminuyeron de forma notoria en comparación con el 2016-I, y el rango de edad de mayor probabilidad de deserción es el de estudiantes de 40 años o más.

Rango Edad	2016-I		2016-II	
	$P(DES RE)$	$P(NODES RE)$	$P(DES RE)$	$P(NODES RE)$
0 a 10	0 %	0 %	0 %	0 %
11 a 14	0 %	0 %	0 %	0 %
15 a 17	11,61 %	88,39 %	9,57 %	90,43 %
18 a 19	12,45 %	87,55 %	11,46 %	88,54 %
20 a 21	10,88 %	89,12 %	7,85 %	92,15 %
22 a 23	17,68 %	82,32 %	7,81 %	92,19 %
24 a 25	24,83 %	75,17 %	10,07 %	89,93 %
26 a 29	27,2 %	72,8 %	14,13 %	85,87 %
30 a 35	28,81 %	71,19 %	16,23 %	83,77 %
36 a 39	31,86 %	68,14 %	13,5 %	86,5 %
40 o más	23,55 %	76,45 %	17,06 %	82,94 %

Tabla 2.40: Probabilidad condicional: $P(DES/NODES|RangoEdad)$.

Programa

Con respecto a la variable programa, la tabla 2.41 muestra las probabilidades condicionales de deserción. En este sentido, se puede evidenciar los programas para los que la probabilidad de deserción es mayor del 30 % son los siguientes:

- Para el 2016-I: P-01, P-04, P-08, P-34, P-55, P-56, P-57 y P-58.
- Para el 2016-II: P-16, P-34, P-51 y P-54.

De particular interés es el programa P-34 que tiene una probabilidad mayor al 30 % en los dos semestres estudiados.

Programa	2016-I		2016-II	
	$P(DES P)$	$P(NODES P)$	$P(DES P)$	$P(NODES P)$
P-00	17,04 %	82,96 %	7,54 %	92,46 %
P-01	35,38 %	64,62 %	0,96 %	99,04 %
P-02	17,08 %	82,92 %	0 %	100 %
P-03	8,06 %	91,94 %	0 %	0 %
P-04	45,45 %	54,55 %	13,64 %	86,36 %
P-05	0 %	0 %	6,67 %	93,33 %
P-06	0 %	0 %	0 %	100 %
P-07	17,45 %	82,55 %	8,28 %	91,72 %
P-08	100 %	0 %	0 %	100 %
P-09	13,5 %	86,5 %	9,32 %	90,68 %
P-10	24,16 %	75,84 %	25,08 %	74,92 %
P-11	0 %	0 %	0 %	0 %
P-12	0 %	0 %	0 %	0 %
P-13	12,61 %	87,39 %	7,25 %	92,75 %
P-14	0 %	0 %	0 %	0 %
P-15	22,68 %	77,32 %	11,95 %	88,05 %
P-16	0 %	0 %	39,13 %	60,87 %
P-17	16,4 %	83,6 %	17,67 %	82,33 %
P-18	7,09 %	92,91 %	8,5 %	91,5 %
P-19	30 %	70 %	13,17 %	86,83 %
P-20	13,96 %	86,04 %	11,93 %	88,07 %
P-21	0 %	0 %	0 %	0 %
P-22	0 %	0 %	0 %	0 %
P-23	0 %	0 %	0 %	0 %
P-24	12,11 %	87,89 %	6,35 %	93,65 %
P-25	19,93 %	80,07 %	15,13 %	84,87 %
P-26	0 %	0 %	0 %	0 %
P-27	14,53 %	85,47 %	7,26 %	92,74 %
P-28	0 %	100 %	0 %	100 %
P-29	0 %	0 %	0 %	0 %
P-30	17,99 %	82,01 %	9,69 %	90,31 %
P-31	0 %	0 %	0 %	0 %
P-32	15,77 %	84,23 %	9,12 %	90,88 %
P-33	18,5 %	81,5 %	9,52 %	90,48 %
P-34	33,33 %	66,67 %	50 %	50 %
P-35	15,74 %	84,26 %	12,06 %	87,94 %
P-36	23,18 %	76,82 %	13,4 %	86,6 %
P-37	0 %	0 %	29,41 %	70,59 %
P-38	20,25 %	79,75 %	8,51 %	91,49 %
P-39	20,09 %	79,91 %	13,41 %	86,59 %
P-40	13,91 %	86,09 %	11,9 %	88,1 %
P-41	16,55 %	83,45 %	9,2	90,8 %
P-42	0 %	100 %	0 %	100 %
P-43	0 %	100 %	5,56 %	94,44 %
P-44	0,62 %	99,38 %	5,49 %	94,51 %
P-45	0 %	100 %	0 %	100 %
P-46	0 %	100 %	9,09 %	90,91 %
P-47	10,35 %	89,65 %	2,9 %	97,1 %
P-48	11,61 %	88,39 %	3,93 %	96,07 %
P-49	18,39 %	81,61 %	9,09 %	90,91 %
P-50	17,89 %	82,11 %	18,14 %	81,86 %
P-51	0 %	0 %	41,67 %	58,33 %
P-52	24,02 %	75,98 %	7,02 %	92,98 %
P-53	19,9 %	80,1 %	9,76 %	90,24 %
P-54	0 %	0 %	96,97 %	3,03 %
P-55	100 %	0 %	0 %	0 %
P-56	100 %	0 %	0 %	0 %
P-57	100 %	0 %	0 %	0 %
P-58	100 %	0 %	0 %	0 %
P-59	22,09 %	77,91 %	17,54 %	82,46 %
P-60	23,73 %	76,27 %	18,22 %	81,78 %
P-61	23,08 %	76,92 %	8,52 %	91,48 %
P-62	0 %	0 %	14,84 %	85,16 %

Tabla 2.41: Probabilidad condicional: $P(DES/NODES|Programa)$.

Facultad

La tabla 2.42 presenta las probabilidades condicionales de deserción dada la facultad a la que pertenezcan los estudiantes. Cabe aclarar que todas las facultades muestran unos valores semejantes, con excepción de la facultad de ciencias agrarias y agroindustriales que no existía para los semestres estudiados y empezó labores en 2017-I. Es posible evidenciar una mayor probabilidad de deserción en la facultad de Tecnología para los semestres 2016-I y 2016-II, seguido por la facultad de Ciencias básicas.

Facultad	2016-I		2016-II	
	$P(DES F)$	$P(NODES F)$	$P(DES F)$	$P(NODES F)$
Bellas artes y hum.	16,73 %	83,27 %	10 %	90 %
Cienc. ambientales	17,66 %	82,34 %	9,67 %	90,33 %
Cienc. básicas	20,09 %	79,91 %	13,41 %	86,59 %
Cienc. de la educación	16,37 %	83,63 %	10,39 %	89,61 %
Cienc. de la salud	14,24 %	85,76 %	5,28 %	94,72 %
Ing. industrial	15,23 %	84,77 %	9,9 %	90,12 %
Ing. mecánica	14,48 %	85,52 %	7,25 %	92,75 %
Ingenierías	15,4 %	84,6 %	11,19 %	88,81 %
Tecnología	23,6 %	76,4 %	14,16 %	85,84 %
Cienc. agrar. y agroind.	0 %	0 %	0 %	0 %

Tabla 2.42: Probabilidad condicional: $P(DES/NODES|Facultad)$.

Jornada

De acuerdo a la variable jornada, ver tabla 2.43, se puede observar que las probabilidades de ser desertor con respecto a las 4 jornadas, son mayores para la jornada Nocturna y Nocturna especial en los dos semestres analizados. Por otra parte se evidencia una disminución cercana al 7 % en todas las jornadas, del semestre 2016-I al 2016-II.

Jornada	2016-I		2016-II	
	$P(DES J)$	$P(NODES J)$	$P(DES J)$	$P(NODES J)$
Diurna	16,4 %	83,6 %	8,92 %	91,08 %
Diurna especial	11,61 %	88,39 %	5,53 %	94,47 %
Nocturna	20,28 %	79,72 %	12,77 %	87,23 %
Nocturna especial	20,84 %	79,16 %	15,19 %	84,81 %

Tabla 2.43: Probabilidad condicional: $P(DES/NODES|Jornada)$.

Naturaleza del colegio

Con respecto a la variable naturaleza del colegio, los estudiantes con información desconocida presentan en el 2016-I una probabilidad de deserción mayor que los demás

con un 77,27% (ver tabla 2.44), no obstante para el 2016-II las probabilidades de deserción tienen un comportamiento similar pero menor al 2016-I y satisfactoriamente la probabilidad de deserción para los estudiantes con información desconocida se reduce en más de 60%.

Naturaleza colegio	2016-I		2016-II	
	$P(DES N)$	$P(NODES N)$	$P(DES N)$	$P(NODES N)$
N/A	77,27 %	22,73 %	14,29 %	85,71 %
Oficial	17,53 %	82,47 %	10,38 %	89,62 %
Privado	15,66 %	84,34 %	9,67 %	90,33 %

Tabla 2.44: Probabilidad condicional: $P(DES/NODES|Naturaleza)$.

Departamento

Ahora, respecto a la variable departamento, la tabla 2.45 muestra que la probabilidad de ser desertor es mayor al 20% para los departamentos de: Boyacá, Choco, Magdalena, Santander, Ext y No disponible en el 2016-I, y para los departamentos de: Amazonas, Córdoba, Guainía, Magdalena y Sucre en el 2016-II. En ambos semestres, el departamento del Magdalena registró la mayor probabilidad de deserción. Ahora, la deserción en el departamento de Risaralda fue de 18,13% y de 10,66% para el 2016-I y 2016-II, respectivamente, mostrando una disminución del 8% aproximadamente. Cabe resaltar que estas cifras son relativamente bajas comparadas con otros departamentos dada la alta composición de estudiantes de la UTP que proceden de Risaralda.

Departamento	2016-I		2016-II	
	$P(DES D)$	$P(NODES D)$	$P(DES D)$	$P(NODES D)$
Amazonas	40 %	60 %	25 %	75 %
Antioquia	17,14 %	82,86 %	8,45 %	91,55 %
Arauca	0 %	100 %	16,67 %	83,33 %
Atlántico	5,88 %	94,12 %	8,7 %	91,3 %
Bogotá	15,56 %	84,44 %	15,64 %	84,36 %
Bolívar	6,25 %	93,75 %	12,5 %	87,5 %
Boyacá	22,22 %	77,78 %	7,14 %	92,8 %
Caldas	16,05 %	83,95 %	10,04 %	89,96 %
Caquetá	15,52 %	84,48 %	6,78 %	93,22 %
Casanare	0 %	100 %	18,18 %	81,81 %
Cauca	12,4 %	87,6 %	6,14 %	93,86 %
Cesar	7,69 %	92,31 %	7,69 %	92,31 %
Chocó	25,93 %	74,06 %	16,46 %	83,54 %
Córdoba	15,38 %	84,62 %	21,43 %	78,57 %
Cundinamarca	17,14 %	82,86 %	12,33 %	87,67 %
Guaviare	9,09 %	90,91 %	15,38 %	84,62 %
Huila	7,96 %	92,04 %	8,87 %	91,13 %
Guajira	25 %	75 %	0	100 %
Guainía	0 %	100 %	50 %	50 %
Magdalena	28,57 %	71,43 %	33,33 %	66,67 %
Meta	25 %	75 %	0	100 %
N. de Santander	8,7 %	91,3 %	4,17 %	95,83 %
Nariño	8,4 %	91,6 %	5,39 %	94,61 %
Putumayo	9,84 %	90,16 %	9,30 %	90,7 %
Quindío	16,48 %	83,52 %	10,06 %	89,94 %
Risaralda	18,13 %	81,87 %	10,66 %	89,34 %
Santander	23,81 %	76,19 %	5,26 %	94,74 %
San Andrés	14,29 %	85,71 %	8,89 %	91,11 %
Sucre	0 %	100 %	33,33 %	66,67 %
Tolima	15,07 %	84,93 %	6,67 %	93,33 %
Valle Del Cauca	15,3 %	84,7 %	9,33 %	90,67 %
Vaupés	0 %	100 %	0 %	100 %
Ext	31,25 %	68,75 %	0	100 %
No Disponible	33,33 %	66,67 %	10,64 %	89,36 %

Tabla 2.45: Probabilidad condicional: $P(DES/NODES|Departamento)$.

Capital

De la tabla 2.46 se observa una disminución de la probabilidad de deserción con respecto a la variable capital cercana al 7 %, a lo largo de los semestres analizados, evidenciándose

un leve incremento en los estudiantes provenientes de ciudades capitales.

Capital	2016-I		2016-II	
	$P(DES C)$	$P(NODES C)$	$P(DES C)$	$P(NODES C)$
No	16,87 %	83,13 %	9,75 %	90,25 %
Sí	17,72 %	82,28 %	10,78 %	89,22 %

Tabla 2.46: Probabilidad condicional: $P(DES/NODES|Capital)$.

Estrato

Para el 2016-I las probabilidades más altas de deserción se registraron en los estratos Medio-bajo y Medio, con un 19,61 % y 18,29 %, respectivamente. Además, los estudiantes sin registro confiable del estrato mostraron la probabilidad más alta con un 73,91 %. Adicionalmente, para el 2016-II las probabilidades de deserción disminuyeron para todos los estratos, incluyendo la categoría no disponible, con la mayor probabilidad de deserción para el estrado Bajo-bajo, 11,19 % (ver tabla 2.47).

Estrato	2016-I		2016-II	
	$P(DES E)$	$P(NODES E)$	$P(DES E)$	$P(NODES E)$
Bajo-bajo	16,22 %	83,78 %	11,19 %	88,81 %
Bajo	16,57 %	83,43 %	10,3 %	89,7 %
Medio-bajo	19,61 %	80,39 %	9,47 %	90,53 %
Medio	18,29 %	81,71 %	9,92 %	90,08 %
Medio-alto	13,46 %	86,54 %	7,72 %	92,28 %
Alto	12,24 %	87,76 %	9,57 %	90,43 %
No disponible	73,91 %	26,09 %	12,5 %	87,5 %

Tabla 2.47: Probabilidad condicional: $P(DES/NODES|Estrato)$.

Semestre de ubicación

Un valor mayor al 30 % de estudiantes de primer semestre tienden a ser desertores de un periodo en los dos semestres analizados. Además, se observa un alto índice de deserción para estudiantes ubicados en semestres mayores a 10 en el 2016-I, aunque dichas cifras disminuyen notoriamente para el 2016-II. Finalmente, cabe anotar que estudiantes con registros de ubicación no disponible muestran las probabilidades de deserción más altas (ver tabla 2.48).

Semestre	2016-I		2016-II	
	$P(DES S)$	$P(NODES S)$	$P(DES S)$	$P(NODES S)$
0	0 %	0 %	0 %	0 %
1	40,61 %	59,39 %	33,6 %	66,41 %
2	12,32 %	87,68 %	14,13 %	85,87 %
3	10,2 %	89,8 %	9,32 %	90,68 %
4	7,5 %	92,5 %	6,55 %	93,45 %
5	6,77 %	93,23 %	5,94 %	94,06 %
6	17,93 %	82,07 %	5,72 %	94,28 %
7	2,79 %	97,21 %	3,57 %	96,43 %
8	4,33 %	95,67 %	2,87 %	97,13 %
9	3 %	97 %	2,33 %	97,67 %
10	31,17 %	68,83 %	5,85 %	94,15 %
11	25,12 %	74,88 %	3,5 %	96,5 %
12	43,21 %	56,79 %	10,67 %	89,33 %
13	67,14 %	32,86 %	0 %	100 %
No disponible	53,46 %	46,54 %	50,46 %	49,54 %

Tabla 2.48: Probabilidad condicional: $P(DES/NODES|Semestre)$.

Número de períodos matriculados

Con respecto a la variable número de períodos matriculados las mayores probabilidades de deserción se registran para estudiantes con 12 o más periodos matriculados para el 2016-I. Para el 2016-II, la probabilidad de deserción es mayor para estudiantes con 16 o más períodos matriculados. Debe resaltarse que las probabilidades de deserción disminuyen para notoriamente para todos rangos de un semestre a otro (ver tabla 2.49).

Períodos	2016-I		2016-II	
	$P(DES M)$	$P(NODES M)$	$P(DES M)$	$P(NODES M)$
0 a 3	16,2 %	83,8 %	15,49 %	84,51 %
4 a 7	12,09 %	87,91 %	7,33 %	92,67 %
8 a 11	16,74 %	83,26 %	5,4 %	94,6 %
12 a 15	32,78 %	67,22 %	9,13 %	90,87 %
16 a 19	38,36 %	61,63 %	15,28 %	84,72 %
20 a 23	36,84 %	63,16 %	18,32 %	81,67 %
24 a 27	29,41 %	70,59 %	17,24 %	82,76 %
28 a 31	27,27 %	72,73 %	18,18 %	81,82 %
32 o más	50 %	50 %	0 %	100 %

Tabla 2.49: Probabilidad condicional: $P(DES/NODES|Períodos)$.

Créditos acumulados

En concordancia a las dos variables antes analizadas, la variable créditos acumulados de la tabla 2.50 nos indica que la mayor probabilidades de deserción dado los créditos acumulados se registra en los rangos de 0 a 10 créditos y de 171 créditos en adelante para el semestre 2016-I. Nuevamente, los valores de probabilidades disminuyen en el 2016-II pero se mantiene el mayor registro, 34,36 %, para el rango 0 a 10.

Créditos	2016-I		2016-II	
	$P(DES CA)$	$P(NODES CA)$	$P(DES CA)$	$P(NODES CA)$
0 a 10	40 %	60 %	34,36 %	65,64 %
11 a 20	13,24 %	86,76 %	16,79 %	83,21 %
21 a 30	12,81 %	87,19 %	12,42 %	87,58 %
31 a 40	10,76 %	89,24 %	10,21 %	89,79 %
41 a 50	9,27 %	90,73 %	9,1 %	90,9 %
51 a 60	10,15 %	89,85 %	7,2 %	92,8 %
61 a 70	8,4 %	91,6 %	6,58 %	93,42 %
71 a 80	6,29 %	93,71 %	6,82 %	93,18 %
81 a 90	8,21 %	91,79 %	8,11 %	91,89 %
91 a 100	4,86 %	95,14 %	4,52 %	95,48 %
101 a 110	9,52 %	90,48 %	5,23 %	94,77 %
111 a 120	22,95 %	77,05 %	3,37 %	96,63 %
121 a 130	5,26 %	94,74 %	5,24 %	94,76 %
131 a 140	5,88 %	94,12 %	3,34 %	96,66 %
141 a 150	4,85 %	95,15 %	4,38 %	95,62 %
151 a 160	17,10 %	82,9 %	4,32 %	95,68 %
161 a 170	25,91 %	74,09 %	7,12 %	92,88 %
171 a 180	38,69 %	61,31 %	6,74 %	93,26 %
181 a 190	56,85 %	43,15 %	7,06 %	92,94 %
191 a 200	49,7 %	50,3 %	3,16 %	96,84 %
200 o más	38,6 %	61,4 %	3,51 %	96,49 %

Tabla 2.50: Probabilidad condicional: $P(DES/NODES|Creditos)$.

Promedio acumulado por rangos

Según la tabla 2.51, la mayor cantidad de estudiantes desertores según el promedio acumulado son los que presentan promedios entre 0 a 0,99, para los dos semestres estudiados. Además, promedios acumulados entre 1 a 2,99 tienen una probabilidad considerable de deserción en el 2016-I, aunque esta probabilidad disminuye casi a la mitad para el 2016-II.

Promedio	2016-I		2016-II	
	$P(DES PA)$	$P(NODES PA)$	$P(DES PA)$	$P(NODES PA)$
0 a 0,99	40,75 %	59,25 %	50,76 %	49,24 %
1 a 1,99	45,05 %	54,95 %	17,07 %	82,93 %
2 a 2,99	34,16 %	65,84 %	18,7 %	81,3 %
3 a 3,99	14,35 %	85,65 %	7,01 %	92,99 %
4 a 5	13,18 %	86,82 %	3,05 %	96,95 %

Tabla 2.51: Probabilidad condicional: $P(DES/NODES|Promedios)$.

2.5.3. Análisis para el subconjunto poblacional: Desertor dos períodos

Se considera el subconjunto de los estudiantes que desertaron dos períodos consecutivos, según como se determinó en la sección 2.4.2. Con este subconjunto se determina la probabilidad de ser desertor (o no desertor) dada la aparición de cada una de las categorías de las variables.

Género

La tabla 2.52 presenta las probabilidades condicionales de deserción en dos períodos según el género del estudiante. Se puede notar que las probabilidades de deserción son similares para ambos géneros.

Género	2016-I	
	$P(DES G)$	$P(NODES G)$
Masculino	14,94 %	85,06 %
Femenino	14,29 %	85,71 %

Tabla 2.52: Probabilidad condicional: $P(DES/NODES|Genero)$.

Rango de edad

En la tabla 2.53 se muestran las probabilidades condicionales con respecto a la variable rango de edad. En general, a medida que la edad aumenta, la probabilidad de deserción también lo hace. De este modo, la probabilidad deserción para el 2016-I es mayor para el conjunto de estudiantes entre los 36 a 39 años.

Rango de edad	2016-I	
	$P(DES RE)$	$P(NODES RE)$
0 a 10	0 %	0 %
11 a 14	0 %	0 %
15 a 17	10,08 %	89,92 %
18 a 19	10,22 %	89,78 %
20 a 21	8,70 %	91,3 %
22 a 23	15,25 %	84,75 %
24 a 25	21,52 %	78,48 %
26 a 29	23,11 %	76,89 %
30 a 35	25,08 %	74,92 %
36 a 39	28,43 %	71,57 %
40 o más	21,9 %	78,09 %

Tabla 2.53: Probabilidad condicional: $P(DES/NODES|RangoEdad)$.

Programa

Con respecto a la variable programa, la tabla 2.54 presenta las probabilidades condicionales de deserción. Se puede destacar que los programas con probabilidades de deserción mayores al 20 % son: P-01, P-04, P-08, P-10, P-15, P-19, P-34, P-36, P-52, P-55, P-56, P-57, P-58 y P-59.

Programa	2016-I	
	$P(DES P)$	$P(NODES P)$
P-00	15,16 %	84,84 %
P-01	34,62 %	65,38 %
P-02	14,17 %	85,83 %
P-03	6,45 %	93,55 %
P-04	45,45 %	54,55 %
P-05	0 %	0 %
P-06	0 %	0 %
P-07	14,54 %	85,46 %
P-08	100 %	0 %
P-09	11,76 %	88,24 %
P-10	21,14 %	78,86 %
P-11	0 %	0 %
P-12	0 %	0 %
P-13	9,24 %	90,76 %
P-14	0 %	0 %
P-15	20,62 %	79,38 %
P-16	0 %	0 %
P-17	13,18 %	86,82 %
P-18	4,26 %	95,74 %
P-19	28,42 %	71,58 %
P-20	11,4 %	88,6 %
P-21	0 %	0 %
P-22	0 %	0 %
P-23	0 %	0 %
P-24	9,85 %	90,15 %
P-25	15,67 %	84,33 %
P-26	0 %	0 %
P-27	11,45 %	88,55 %
P-28	0 %	100 %
P-29	0 %	0 %
P-30	15,83 %	84,17 %
P-31	0 %	0 %
P-32	12,11 %	87,89 %
P-33	16,69 %	83,31 %
P-34	33,33 %	66,67 %
P-35	13,33 %	86,67 %
P-36	21,83 %	78,17 %
P-37	0 %	0 %
P-38	16,46 %	83,54 %
P-39	15,63 %	84,38 %
P-40	12,47 %	87,53 %
P-41	15,61 %	84,39 %
P-42	0 %	100 %
P-43	0 %	100 %
P-44	0,62 %	99,38 %
P-45	0 %	100 %
P-46	0 %	100 %
P-47	8,58 %	91,42 %
P-48	8,39 %	91,61 %
P-49	16,12 %	83,88 %
P-50	13,79 %	86,21 %
P-51	0 %	0 %
P-52	22,91 %	77,09 %
P-53	17,76 %	82,24 %
P-54	0 %	0 %
P-55	80 %	20 %
P-56	100 %	0 %
P-57	73,68 %	26,32 %
P-58	91,43 %	8,57 %
P-59	20,93 %	79,07 %
P-60	18,24 %	81,76 %
P-61	19,23 %	80,77 %
P-62	0 %	0 %

Tabla 2.54: Probabilidad condicional: $P(DES/NODES|Programa)$.

Facultad

En la tabla 2.55 se presentan las probabilidades condicionales de deserción dada la variable facultad. En general, la probabilidad condicional de deserción son similares para todas las facultades. Aunque la facultad de Tecnología es la que presenta una mayor probabilidad del 20,1 %. La facultad de ciencias agrarias y agroindustriales que no existía en el semestre analizado.

Facultad	2016-I	
	$P(DES F)$	$P(NODES F)$
Bellas artes y hum.	14,27 %	85,73 %
Cienc. ambientales	15,65 %	84,35 %
Cienc. básicas	15,63 %	84,38 %
Cienc. de la educación	14,95 %	85,05 %
Cienc. de la salud	11,81 %	88,19 %
Ing. industrial	12,17 %	87,83 %
Ing. mecánica	11,42 %	88,58 %
Ingenierías	12,76 %	87,24 %
Tecnología	20,1 %	79,9 %
Cienc. agrarias y agroind.	0 %	0 %

Tabla 2.55: Probabilidad condicional: $P(DES/NODES|Facultad)$.

Jornada

Con respecto a la variable jornada, ver tabla 2.56, es posible evidenciar que las jornadas con mayor probabilidad de deserción son la Nocturna y Nocturna especial con 17,83 % y 17,85 %, respectivamente.

Jornada	2016-I	
	$P(DES J)$	$P(NODES J)$
Diurna	13,85 %	86,15 %
Diurna especial	8,39 %	91,61 %
Nocturna	17,83 %	82,17 %
Nocturna especial	17,85 %	82,15 %

Tabla 2.56: Probabilidad condicional: $P(DES/NODES|Jornada)$.

Naturaleza del colegio

Por otro lado la variable naturaleza del colegio, muestra como los estudiantes con información desconocida presentan una probabilidad de deserción mayor que los demás con un 77,27 % (ver tabla 2.57).

Naturaleza colegio	2016-I	
	$P(DES N)$	$P(NODES N)$
N/A	77,27 %	22,73 %
Oficial	14,88 %	85,12 %
Privado	13,17 %	86,83 %

Tabla 2.57: Probabilidad condicional: $P(DES/NODES|NaturalezaColegio)$.

Departamento

Referente a la variable departamento, la tabla 2.58 muestra que la probabilidad de ser desertor es mayor al 20 % para los departamentos de: Amazonas, Choco, Magdalena, Meta y No disponible en el 2016-I, y para los departamentos de: Amazonas, Chocó, Guajira, Magdalena y Meta. Respecto al departamento Risaralda, la probabilidad condicional registro valores de 15,41 %.

Departamento	2016-I	
	$P(DES D)$	$P(NODES D)$
Amazonas	40 %	60 %
Antioquia	11,43 %	88,57 %
Arauca	0 %	100 %
Atlántico	5,88 %	94,12 %
Bogotá	15 %	85 %
Bolívar	6,25 %	93,75 %
Boyacá	16,67 %	83,33 %
Caldas	13,04 %	86,96 %
Caquetá	12,07 %	87,93 %
Casanare	0 %	100 %
Cauca	10,74 %	89,26 %
Cesar	7,69 %	92,31 %
Chocó	23,46 %	76,54 %
Córdoba	0 %	100 %
Cundinamarca	14,29 %	85,71 %
Guaviare	9,09 %	90,91 %
Huila	6,19 %	93,81 %
Guajira	25 %	75 %
Guainía	0 %	100 %
Magdalena	28,57 %	71,43 %
Meta	25 %	75 %
N. de Santander	8,69 %	91,3 %
Nariño	6,45 %	93,55 %
Putumayo	8,19 %	91,8 %
Quindío	14,29 %	85,71 %
Risaralda	15,41 %	84,59 %
Santander	19,05 %	80,95 %
San Andrés	14,29 %	85,71 %
Sucre	0 %	100 %
Tolima	12,33 %	87,67 %
Valle Del Cauca	12,89 %	87,11 %
Vaupés	0 %	100 %
Ext	18,75 %	81,25 %
No Disponible	32,48 %	67,52 %

Tabla 2.58: Probabilidad condicional: $P(DES/NODES|Departamento)$.

Capital

La tabla 2.59, permite observar que independientemente del hecho que el estudiante provenga de una ciudad capital o no, la probabilidad de deserción es similar.

Capital	2016-I	
	$P(DES C)$	$P(NODES C)$
No	14,44 %	85,56 %
Sí	14,9 %	85,1 %

Tabla 2.59: Probabilidad condicional: $P(DES/NODES|Capital)$.

Estrato

Según la tabla 2.60, en el 2016-I las probabilidades más altas de deserción se registraron en los estratos Medio-bajo y Medio, con un 16,5 % y 16,24 %, respectivamente. Es importante centrar la atención en los estudiantes sin registro confiable del estrato puesto que presentan la probabilidad de deserción más alta (73,91 %).

Estrato	2016-I	
	$P(DES E)$	$P(NODES E)$
Bajo	14 %	86 %
Medio-bajo	16,5 %	83,5 %
Medio	16,24 %	83,76 %
Medio-alto	11,86 %	88,14 %
Alto	8,16 %	91,84 %
No disponible	73,91 %	26,09 %

Tabla 2.60: Probabilidad condicional: $P(DES/NODES|Estrato)$.

Semestre de ubicación

Un valor mayor al 36 % de los estudiantes del primer semestre tienden a ser desertores de dos períodos consecutivos. También se observa un alto índice de deserción para estudiantes ubicados en semestres mayores a 10. Cabe anotar que los estudiantes ubicados en semestre 13 y los “no disponible” muestran las probabilidades de deserción más altas (ver tabla 2.61).

Semestre	2016-I	
	$P(DES S)$	$P(NODES S)$
0	0 %	0 %
1	36,67 %	63,33 %
2	9,52 %	90,48 %
3	6,86 %	93,14 %
4	4,64 %	95,36 %
5	4,40 %	95,6 %
6	14,73 %	85,27 %
7	1,97 %	98,03 %
8	2,82 %	97,18 %
9	1,55 %	98,45 %
10	29,69 %	70,31 %
11	25,12 %	74,88 %
12	41,55 %	58,45 %
13	67,14 %	32,86 %
No disponible	43,87 %	56,13 %

Tabla 2.61: Probabilidad condicional: $P(DES/NODES|Semestre)$.

Número de períodos matriculados

Con respecto a la variable número de períodos matriculados se registran las mayores probabilidades de deserción para los rangos 12 a 15, 20 a 23, 24 a 27 y 28 a 31 (ver tabla 2.63).

Períodos	2016-I	
	$P(DES M)$	$P(NODES M)$
0 a 3	14,08 %	85,92 %
4 a 7	9,16 %	90,84 %
8 a 11	14,31 %	85,69 %
12 a 15	29,67 %	70,33 %
16 a 19	34 %	66 %
20 a 23	33,08 %	66,92 %
24 a 27	29,41 %	70,59 %
28 a 31	27,27 %	72,73 %
32 o más	0 %	100 %

Tabla 2.62: Probabilidad condicional: $P(DES/NODES|PeriodosMatriculados)$.

Créditos acumulados

De manera semejante a las dos variables antes analizadas, la variable créditos acumulados de la tabla 2.63 indica que las mayores probabilidades de deserción dado los créditos acumulados se registra en el rango de 0 a 10 créditos y de 171 créditos en adelante.

Créditos	2016-I	
	$P(DES CA)$	$P(NODES CA)$
0 a 10	35,74 %	64,26 %
11 a 20	10,55 %	89,45 %
21 a 30	9,01 %	90,99 %
31 a 40	7,27 %	92,73 %
41 a 50	5,98 %	94,02 %
51 a 60	6,73 %	93,27 %
61 a 70	4,89 %	95,11 %
71 a 80	4,8 %	95,2 %
81 a 90	5,8 %	94,2 %
91 a 100	3,47 %	96,53 %
101 a 110	6,78 %	93,22 %
111 a 120	19,59 %	80,41 %
121 a 130	4,02 %	95,98 %
131 a 140	4,2 %	95,8 %
141 a 150	2,61 %	97,39 %
151 a 160	15,65 %	84,35 %
161 a 170	23,58 %	76,42 %
171 a 180	38,3 %	61,7 %
181 a 190	55,06 %	44,94 %
191 a 200	49,7 %	50,3 %
200 o más	36,84 %	63,16 %

Tabla 2.63: Probabilidad condicional: $P(DES/NODES|CreditosAcumulados)$.

Promedio acumulado por rangos

Según la tabla 2.64, la mayor cantidad de estudiantes desertores dado el promedio acumulado son los que presentan promedios en el rango 0 a 0,99. Además, para promedios acumulados entre 1 a 2,99 se tiene una elevada probabilidad de deserción del 29,73 %.

Promedio	2016-I	
	$P(DES PA)$	$P(NODES PA)$
0 a 0,99	38,07 %	61,93 %
1 a 1,99	29,73 %	70,27 %
2 a 2,99	22,78 %	77,22 %
3 a 3,99	11,82 %	88,18 %
4 a 5	12,34 %	87,66 %

Tabla 2.64: Probabilidad condicional: $P(DES/NODES|PromedioAcumulado)$.

Capítulo 3

Análisis de relevancia de las características de los grupos objetivo

El objetivo fundamental del PAI es reducir los niveles de deserción de la UTP. En este sentido, es de mayor importancia determinar la eficacia del PAI en la consecución del mencionado objetivo. Para verificar la eficacia del PAI se analizan tres escenarios metodológicos que permiten establecer la existencia (o no) de una relación directa entre el servicio que brinda el PAI y la deserción estudiantil para los semestres analizados.

Se analizará el conjunto poblacional completo, sobre el siguiente grupo de características: género, edad, programa, facultad, jornada, naturaleza del colegio, municipio del colegio, departamento del colegio, capital, atendido PAI, atendido línea BPS, atendido línea académica, atendido línea socio-económica, número de atenciones del PAI, estrato, semestre de ubicación, créditos acumulados, períodos matriculados, apoyo económico, promedio acumulado, desertor de un período y desertor de dos períodos. Las variables son descritas en el apéndice A.

Las tres metodologías empleadas para establecer la existencia (o no) de una relación entre la atención del PAI y los niveles de deserción son los siguientes:

1. Matriz de correlación
2. Análisis de componentes principales
3. Alineación centrada de kernel

3.1. Matriz de correlación

La dependencia lineal entre dos variables, x_j y x_k , se analiza mediante el coeficiente de correlación lineal r_{jk} , tal como se presenta en la ecuación (3.1)

$$r_{jk} = \frac{s_{jk}}{s_j s_k}, \quad (3.1)$$

donde, s_{jk} es la covarianza entre las variables x_j y x_k , s_j es la varianza de x_j y s_k es la varianza de la variable x_k . El coeficiente de correlación cumple con lo siguiente: (1) $0 \leq |r_{jk}| \leq 1$, (2) si existe una relación lineal exacta entre las variables, entonces $|r_{jk}| = 1$, y (3) $|r_{jk}|$ es invariante ante las transformaciones lineales de las variables.

Por otra parte, la dependencia por pares entre las variables se mide por la matriz de correlación \mathbf{R} , la cual es cuadrada, simétrica y semidefinida positiva, con unos en la diagonal principal y fuera de ella los coeficientes de correlación lineal entre pares de variables [Peña2002], según la ecuación (3.2)

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ \vdots & \vdots & \dots & \vdots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix} \quad (3.2)$$

El cálculo de la matriz de correlación para el conjunto poblacional 2016-I se presenta en la figura 3.1, en esta matriz se tienen en cuenta las variables desertor de un período y desertor de dos períodos.

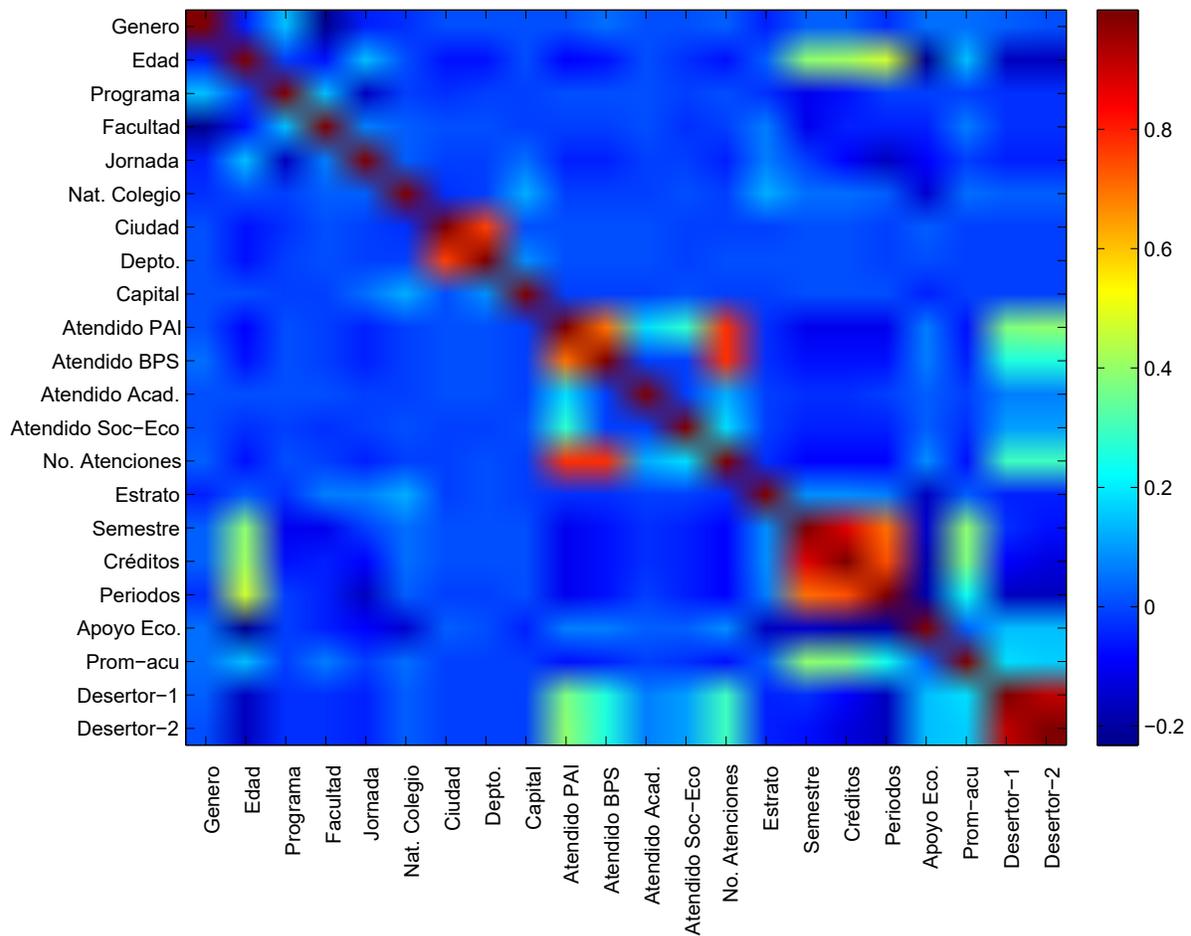


Figura 3.1: Matriz de correlación para 2016-I

Es importante a partir de la figura 3.1 analizar la relación existente entre las variables de deserción y cada una de las demás variables. Adicionalmente, aquellos valores de correlación mayores (bien sean positivos o negativos) indican que existe relación lineal entre las variables. Así pues, vemos claras correlaciones positivas entre los siguientes pares de variables:

- Número de atenciones PAI y Atendido PAI
- Número de atenciones PAI y Atendido línea BPS
- Atendido PAI y Atendido línea BPS
- Períodos matriculados y Semestre de ubicación
- Períodos matriculados y Créditos acumulados
- Créditos acumulados y Semestre de ubicación

También se aprecian correlaciones positivas pero un poco menores entre los siguientes pares de variables:

- Edad y Semestre de ubicación
- Edad y Créditos acumulados
- Edad y Períodos matriculados
- Promedio acumulado y Semestre de ubicación
- Promedio acumulado y Créditos acumulados

Con respecto a la deserción se pueden identificar las siguientes correlaciones positivas:

- Deserción de un período y Atendido PAI
- Deserción de un período y Atendido línea BPS
- Deserción de un período y Número de atenciones PAI
- Deserción de dos períodos y Atendido PAI
- Deserción de dos períodos y Atendido línea BPS
- Deserción de dos períodos y Número de atenciones PAI

El hecho de que estas correlaciones sean positivas está dado por el tipo de codificación empleado para las variables desertor de un período y desertor de dos períodos, en las cuales cero (0) significa que el estudiante sí desertó de la universidad y uno (1) que el estudiante no desertó de la universidad. En este caso la correlación positiva nos hace pensar que a mayor número de atenciones en las líneas de PAI, especialmente en la línea BPS, mayor es la no deserción de la universidad. El mismo análisis aplica para la variable deserción de dos períodos.

También se resalta la fuerte correlación positiva entre la variable desertor de un período y desertor de dos períodos, indicando que muchos de los estudiantes que desertan por un semestre terminan desertando por dos períodos consecutivos.

Para el escenario 2016-II, el cálculo de la matriz de correlación se presenta en la figura 3.2.

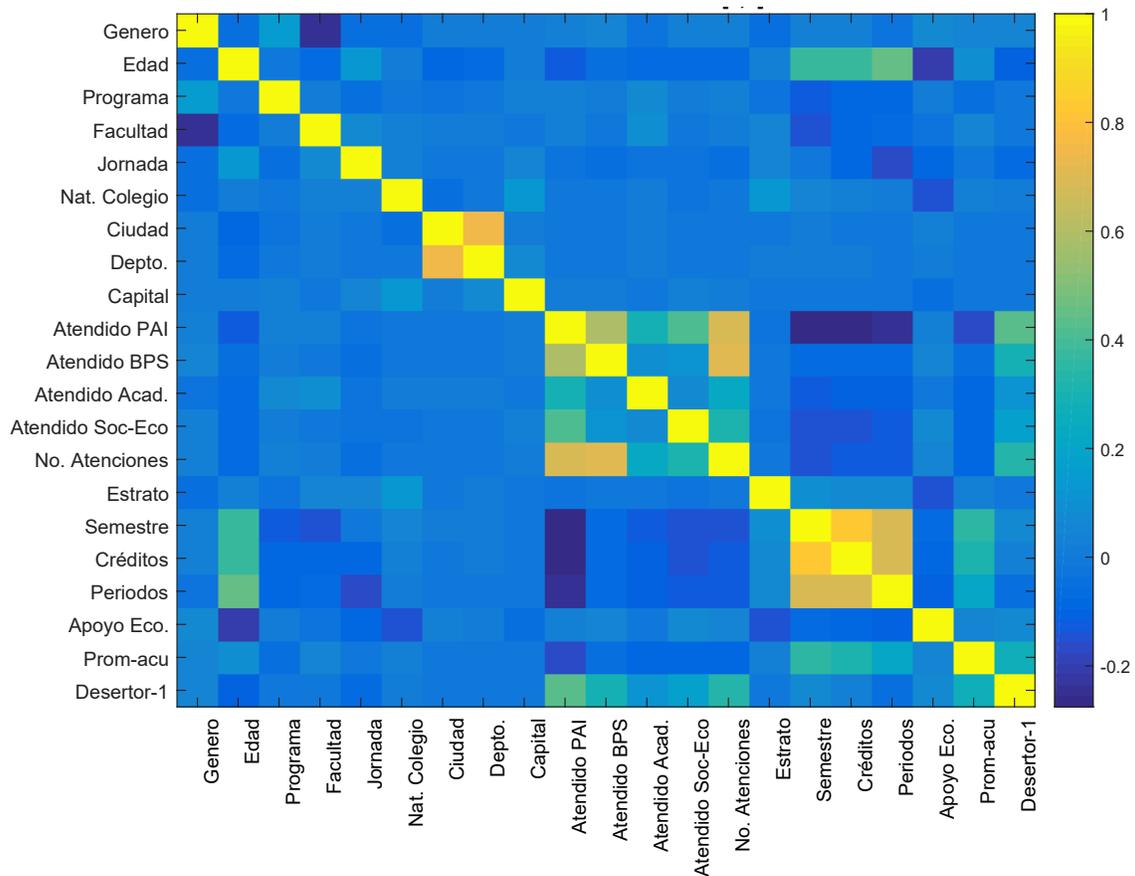


Figura 3.2: Matriz de correlación para 2016-II

Para el 2016-II, existen claras correlaciones positivas entre los siguientes pares de variables:

- Número de atenciones PAI y Atendido PAI
- Número de atenciones PAI y Atendido línea BPS
- Atendido PAI y Atendido línea BPS
- Ciudad y Departamento
- Períodos matriculados y Semestre de ubicación
- Períodos matriculados y Créditos acumulados
- Créditos acumulados y Semestre de ubicación
- Edad y Semestre de ubicación
- Edad y Créditos acumulados

- Edad y Períodos matriculados
- Promedio acumulado y Semestre de ubicación
- Promedio acumulado y Créditos acumulados

Con respecto a la deserción se pueden identificar las siguientes correlaciones positivas:

- Deserción de un período y Atendido PAI
- Deserción de un período y Número de atenciones PAI
- Deserción de un período y Promedio acumulado

Los análisis antes comentados señalan que efectivamente existe una relación entre la existencia del PAI y la deserción que indica que sin existencia del PAI la deserción podría ser mayor.

3.2. Análisis de componentes principales

En el análisis de componentes principales (PCA - *Principal Component Analysis*) se busca representar las variables originales en un número menor de componentes a través de combinaciones lineales. El modelo de transformación está dado por $\mathbf{Z} = \mathbf{X}\mathbf{W}$, donde \mathbf{Z} es una matriz ($n \times m$) de observaciones en el espacio transformado, \mathbf{X} es la matriz ($n \times p$) de observaciones en el espacio original y \mathbf{W} es la matriz ($p \times m$) de transformación que corresponde con los m vectores propios ($\mathbf{v}_i, i = 1, \dots, m$) de la matriz de covarianza (o correlación) asociados a los m valores propios ($\lambda_i, i = 1, \dots, m$) más grandes [Daza Santacoloma et al.2007]. El procedimiento se resumen en el algoritmo 1.

Algoritmo 1 – Análisis de componentes principales

Entrada: \mathbf{X} , m

- 1: Centralizar la matriz de datos \mathbf{X} .
 - 2: Obtener la matriz de covarianza $\mathbf{S} = (1/n)\mathbf{X}^\top\mathbf{X}$. También puede calcularse la matriz de correlación.
 - 3: Calcular los valores propios λ de la matriz \mathbf{S} y sus respectivos vectores propios \mathbf{v} .
 - 4: Ordenar de forma descendente los valores propios y por ende los vectores propios asociados.
 - 5: Proyectar los datos sobre las direcciones principales (primeros m vectores propios ordenados)
-

Adicionalmente, la técnica PCA no sólo realiza la reducción de dimensión del espacio original, sino que también permite la identificación y selección de las variables originales que de mayor forma contribuyen en la representación de los datos [Daza Santacoloma2006]. Para esta selección se analiza el vector de relevancia $\boldsymbol{\rho}$ dado por la ecuación (3.3)

$$\boldsymbol{\rho} = \sum_{i=1}^m |\lambda_i \mathbf{v}_i|, \quad (3.3)$$

los mayores valores dentro del vector $\boldsymbol{\rho}$ señalan cada una de las variables que más contribuyen en la representación del espacio reducido. El vector $\boldsymbol{\rho}$ puede ser estandarizado al rango $[0, 1]$ sin pérdida de generalidad ni de interpretabilidad.

Para los datos del periodo 2016-I, se presenta el respectivo vector de relevancia estandarizado en la figura 3.3, en la cual se observa el orden de importancia que la metodología le asigna a cada una de las variables del conjunto. Es de anotar que la técnica PCA da mayor importancia a las características que presenten mayor variabilidad, porque esto implica mayor cantidad de información aportada por la característica.

Con base en la gráfica 3.3, para PCA las 5 características más relevantes en el proceso de representación de la información son: 1) Apoyo económico, 2) Promedio acumulado, 3) Estrato, 4) Naturaleza del colegio y 5) Jornada.

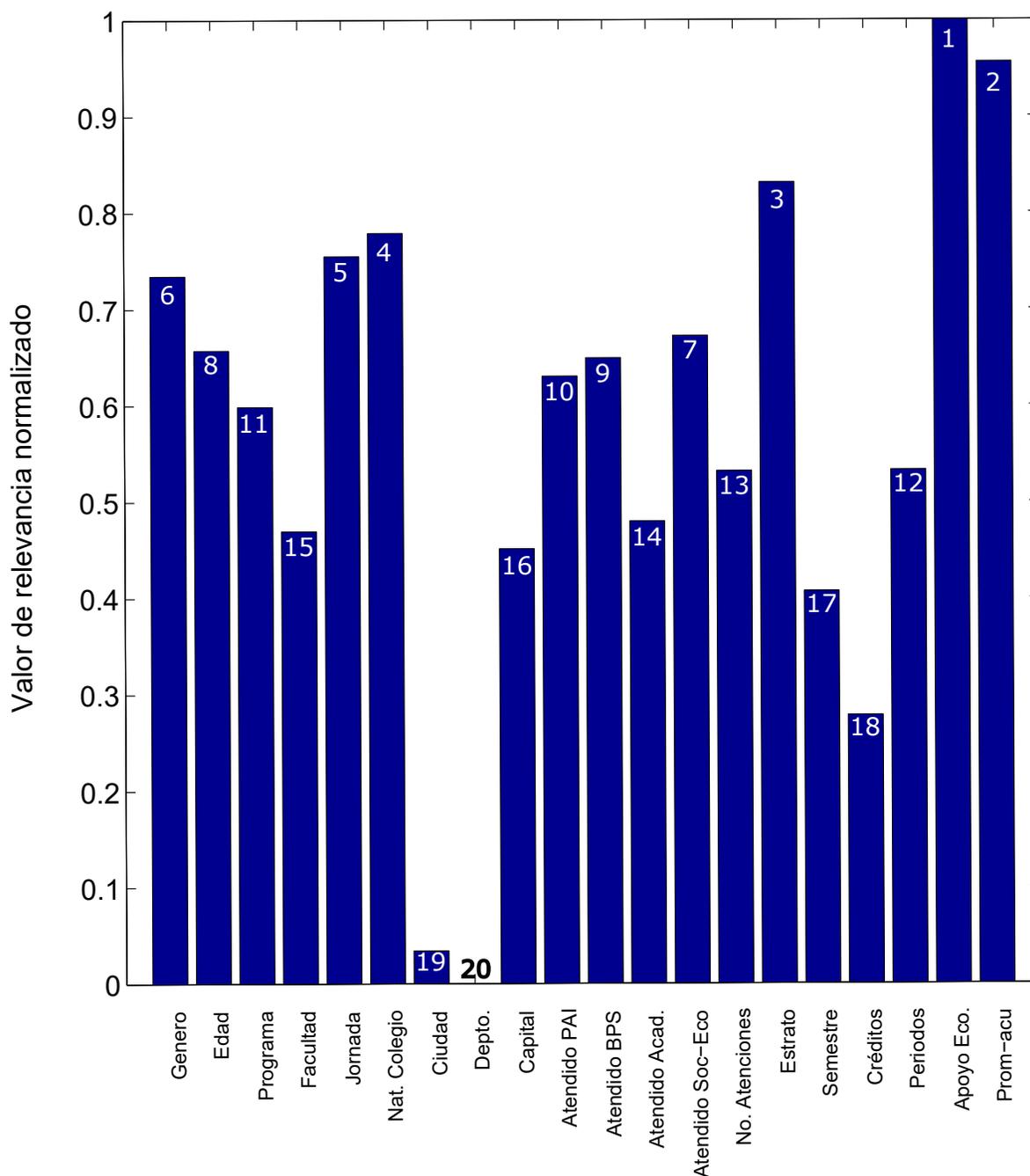


Figura 3.3: Relevancia de variables según PCA para 2016-I

Es de notar que bajo la técnica PCA la importancia del grupo de características no se puede asociar con el fenómeno de deserción universitaria, puesto que el modelo de transformación de componentes principales no ha recibido en ningún momento la información de deserción. La importancia de las variables seleccionadas radica, como se mencionó anteriormente, en la cantidad de información (variabilidad) que poseen, lo

cual es útil para reducir el conjunto de variables con la menor pérdida de información.

3.3. Alineación centrada de kernel

En aprendizaje de máquina, los métodos kernel son un tipo de algoritmos para el análisis de patrones. La tarea convencional del análisis de patrones es hallar y estudiar las diferentes clases de relaciones generales existentes en un conjunto de datos, tales como: agrupaciones, componentes principales, correlaciones, jerarquías, clasificaciones, etc. Para los algoritmos de este tipo sólo se requiere un *kernel* especificado por el usuario, por ejemplo: una función de similitud entre cada par de observaciones multidimensionales que posea el conjunto de datos. Los métodos kernel operan en espacios de características implícitos de alta dimensión sin tener que calcular las coordenadas de las observaciones en dicho espacio de manera directa, sino calculando los productos internos entre las imágenes de todos los pares de datos en el espacio de características.

Dado un conjunto de datos \mathbf{X} (observaciones medidas para cada una de las variables) junto con un vector de etiquetas \mathbf{l} que identifica si las observaciones corresponden con la categorización de desertor o no desertor, se pueden calcular las matrices kernel de caracterización $\mathbf{K}_{\mathbf{X}}$ y $\mathbf{K}_{\mathbf{l}}$. La primera contiene el análisis de similitud entre las observaciones y la segunda entre las etiquetas. Así, se aplican dos funciones kernel de forma secuencial para evaluar la información compartida entre el conjunto de observaciones y las etiquetas de deserción. Es importante determinar qué tan bien se acoplan las matrices kernel de observaciones y etiquetas, es decir, se debe determinar la alineación estadística entre ellas, lo que se conoce como alineación centrada de kernel (CKA - *Centered Kernel Alignment*) según la ecuación (3.4),

$$\hat{\rho}(\mathbf{K}_{\mathbf{X}}, \mathbf{K}_{\mathbf{l}}) = \frac{\langle \bar{\mathbf{K}}_{\mathbf{X}}, \bar{\mathbf{K}}_{\mathbf{l}} \rangle_{\text{F}}}{\sqrt{\langle \bar{\mathbf{K}}_{\mathbf{X}}, \bar{\mathbf{K}}_{\mathbf{X}} \rangle_{\text{F}} \langle \bar{\mathbf{K}}_{\mathbf{l}}, \bar{\mathbf{K}}_{\mathbf{l}} \rangle_{\text{F}}}}, \quad (3.4)$$

donde $\bar{\mathbf{K}}$ representa la matriz centrada de kernel y la notación $\langle \cdot, \cdot \rangle_{\text{F}}$ representa la norma de Frobenius para matrices. Por tanto la ecuación (3.4) es un estimador que permite cuantificar la similitud entre las observaciones de entrada y las etiquetas [Alvarez-Meza et al.2017, Cortes et al.2010].

Dado lo anterior, las figuras 3.4, 3.5 y 3.6 presentan la relevancia de las características observadas asociadas a la etiqueta de deserción para el semestre 2016-I con deserción de uno y dos períodos y para el semestre 2016-II para deserción de un período, respectivamente. En las figuras se observa el orden de importancia de las variables asociado al comportamiento de deserción.

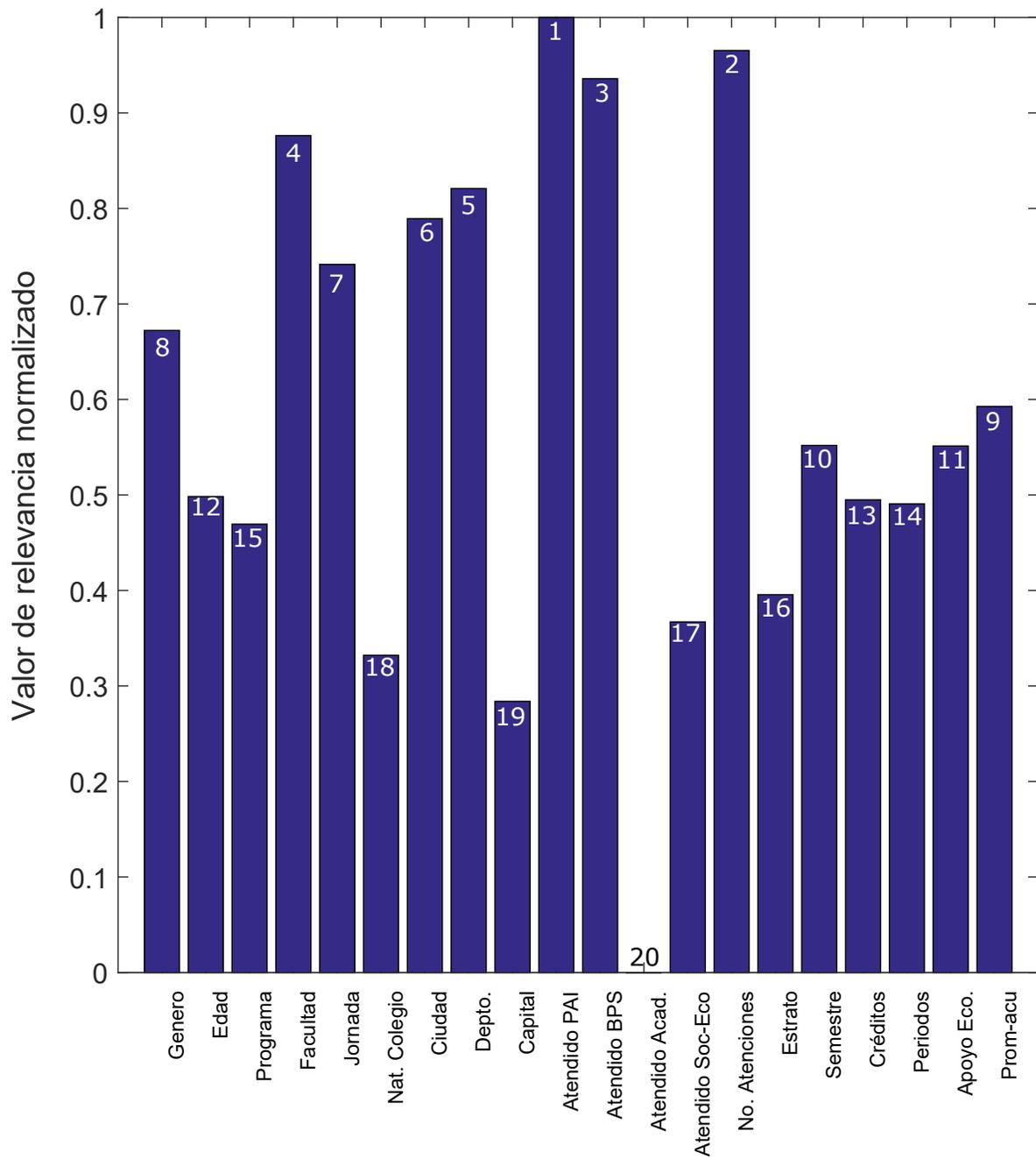


Figura 3.4: Relevancia de variables según CKA para 2016-I asociada a la deserción de 1 periodo

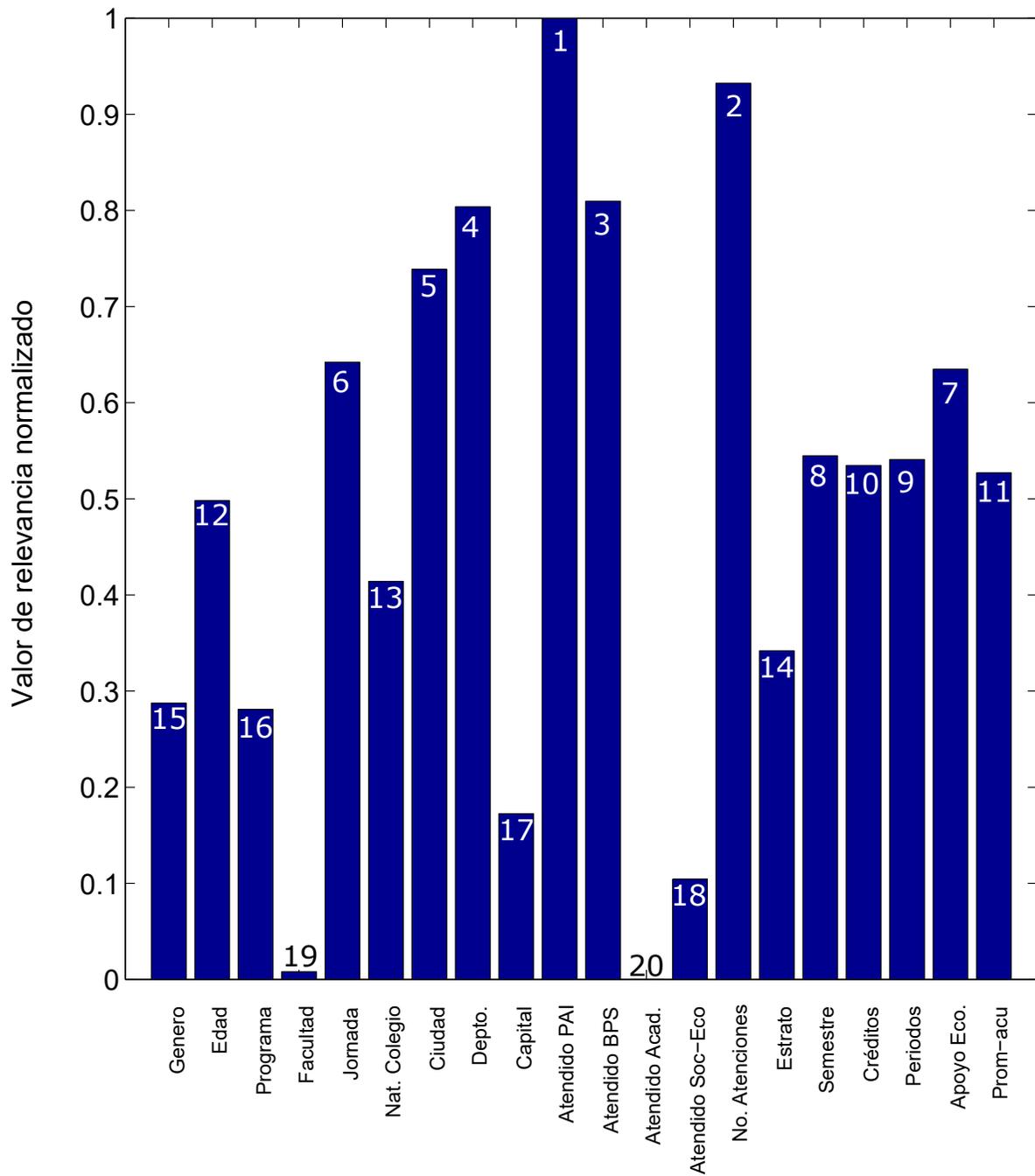


Figura 3.5: Relevancia de variables según CKA para 2016-I asociada a la deserción de 2 periodos

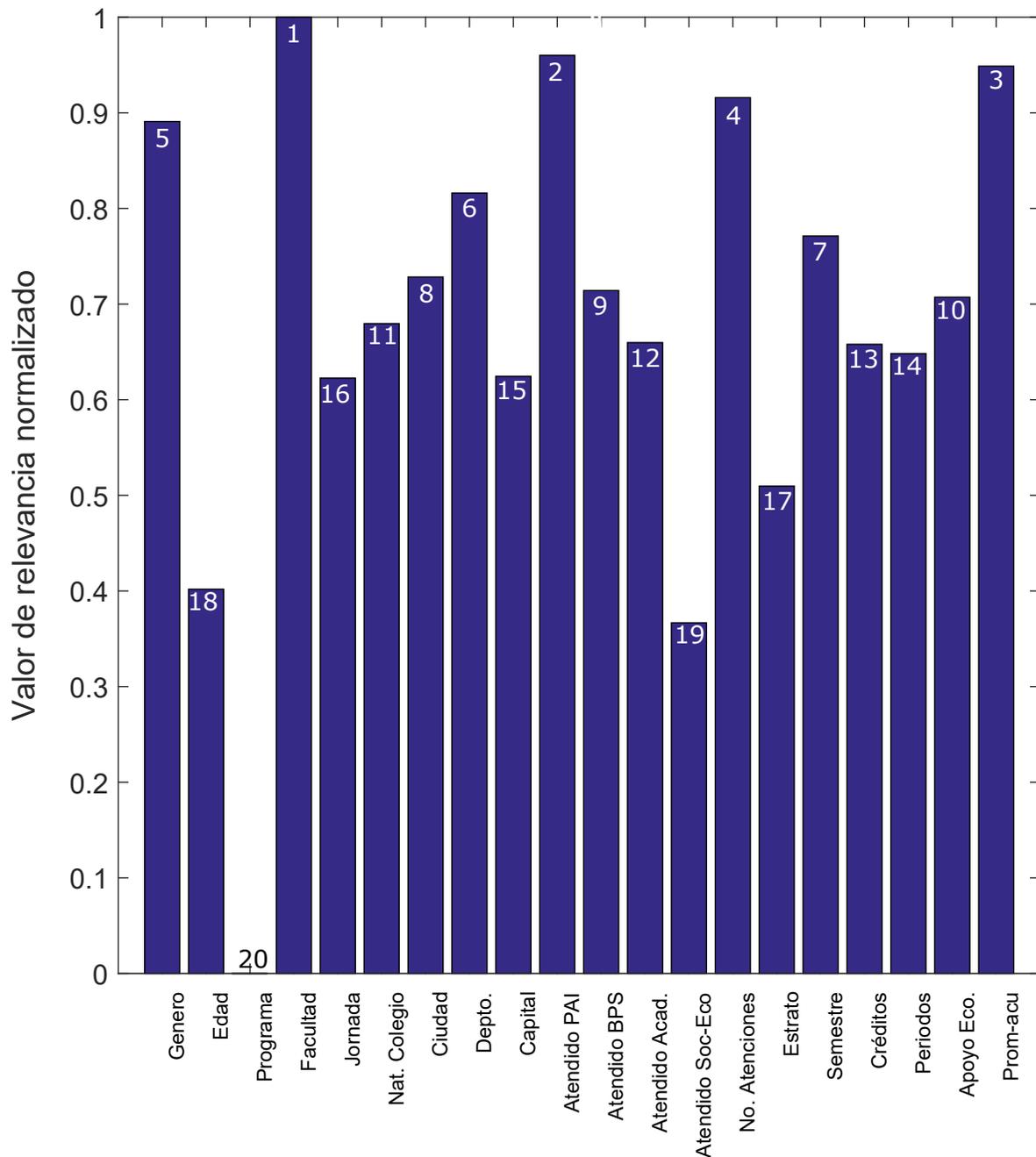


Figura 3.6: Relevancia de variables según CKA para 2016-II asociada a la deserción de 1 periodo

Con el fin de determinar cuántas características son las más importantes del conjunto ordenado por relevancia según CKA, se realiza una prueba de clasificación automática de las observaciones usando el clasificador de vecinos más cercanos de la siguiente forma: primero, con solo la variable atendido PAI (la más importante) se clasifican

los estudiantes en desertores y no desertores, y puesto que se conoce de antemano quienes en realidad fueron o no desertores se puede calcular un nivel de acierto para el clasificador, posteriormente se repite el proceso de clasificación empleando las dos variables más relevantes (atendido PAI y número de atenciones en el PAI) y se determina el acierto del clasificador, el proceso se repite una y otra vez agregando en cada ocasión la siguiente variable más relevante y determinando en cada momento el nivel de acierto. Los resultados de esta prueba son presentados en las figuras 3.7, 3.8 y 3.9. En las figuras mencionadas, se aprecia que el acierto de predicción (o clasificación de desertores y no desertores) aumenta a medida que son agregadas las variables en el orden de relevancia previamente determinado, y que luego de agregar entre 7 y 10 características se logra un nivel de acierto que tiene una tendencia a volverse constante a pesar de continuar aumentando el grupo de variables para clasificar. En este orden de ideas, para cada uno de los tres escenarios analizados las 10 variables más relevantes en relación con la condición de deserción son las siguientes:

2016-I, deserción 1 periodo: 1) Atendido PAI, 2) número de atenciones PAI, 3) atendido línea BPS, 4) facultad, 5) departamento, 6) ciudad, 7) jornada, 8) género, 9) promedio acumulado, y 10) estrato.

2016-I, deserción 2 períodos: 1) Atendido PAI, 2) número de atenciones PAI, 3) atendido línea BPS, 4) departamento, 5) ciudad, 6) jornada, 7) apoyo económico, 8) semestre de ubicación, 9) número de periodos matriculados y 10) créditos acumulados.

2016-II, deserción 1 período: 1) Facultad, 2) atendido PAI, 3) promedio acumulado, 4) número de atenciones PAI, 5) género, 6) departamento, 7) semestre de ubicación, 8) ciudad, 9) atendido línea BPS y 10) apoyo económico.

En los tres escenarios analizados se repiten como variables más relevantes las siguientes: atendido PAI, número de atenciones PAI, atendido línea BPS, departamento y ciudad.

La baja relevancia relativa de la variable atendido línea académica parece estar relacionado con el hecho de que para los semestres analizados existía un subreporte de las actividades que se realizaban en esta línea, es decir, no hay una medición adecuada para estos períodos de tiempo de la línea de apoyo académico.

Es importante destacar el hecho de que la relevancia de las variables es relativa entre ellas, por tanto las variables con valor de relevancia normalizado igual a cero no significa que no sean importantes, sino que en comparación con las otras variables tienen menor importancia al describir el fenómeno bajo análisis.

A pesar de no aparecer en los 10 primeros lugares de relevancia para los tres escenarios analizados, variables como: apoyo económico y promedio acumulado tiene un valor de relevancia normalizada alto en los escenarios.

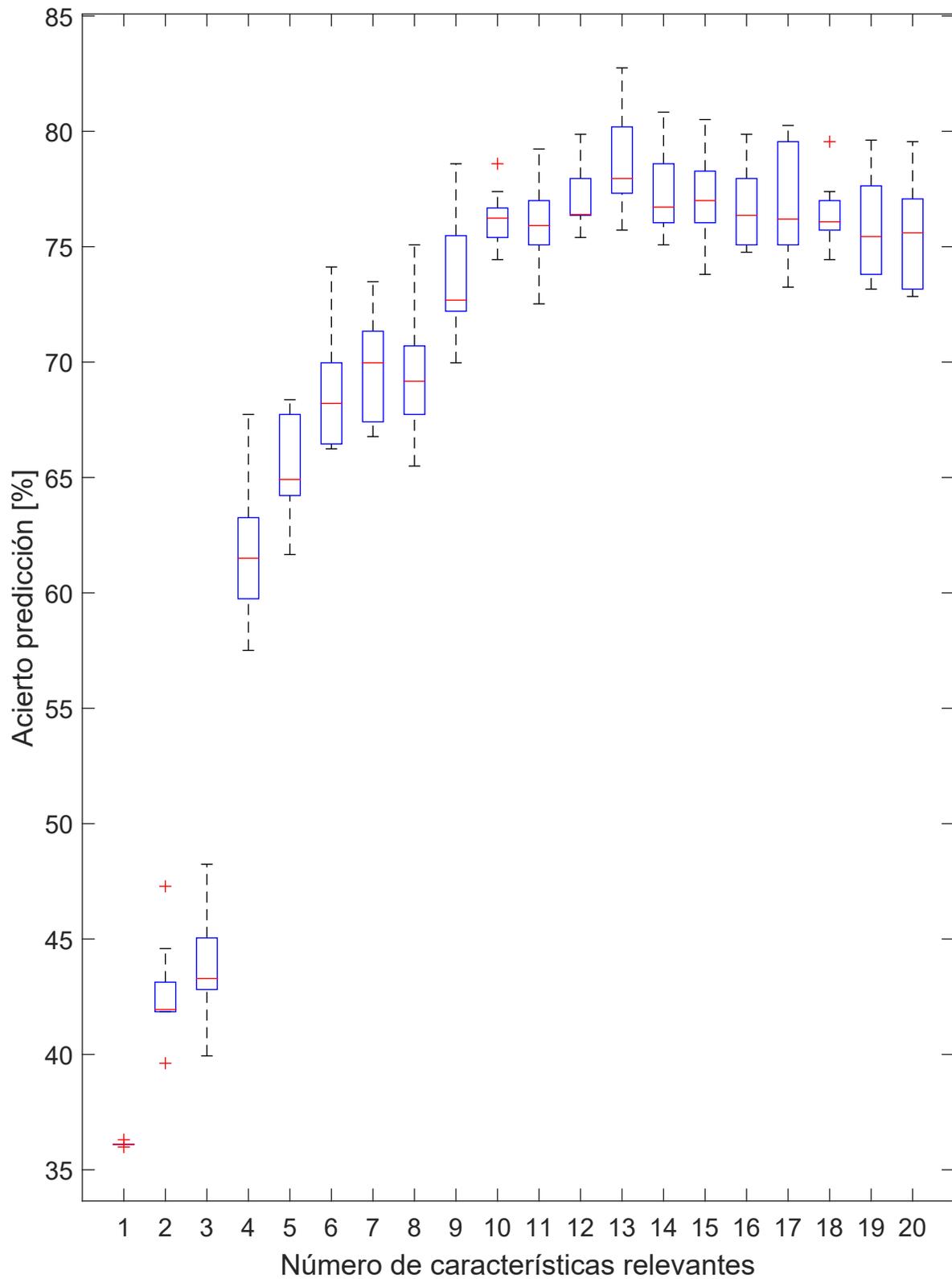


Figura 3.7: Clasificación automática de deserción empleando KNN para variables relevantes de CKA, 2016-I, deserción 1 periodo

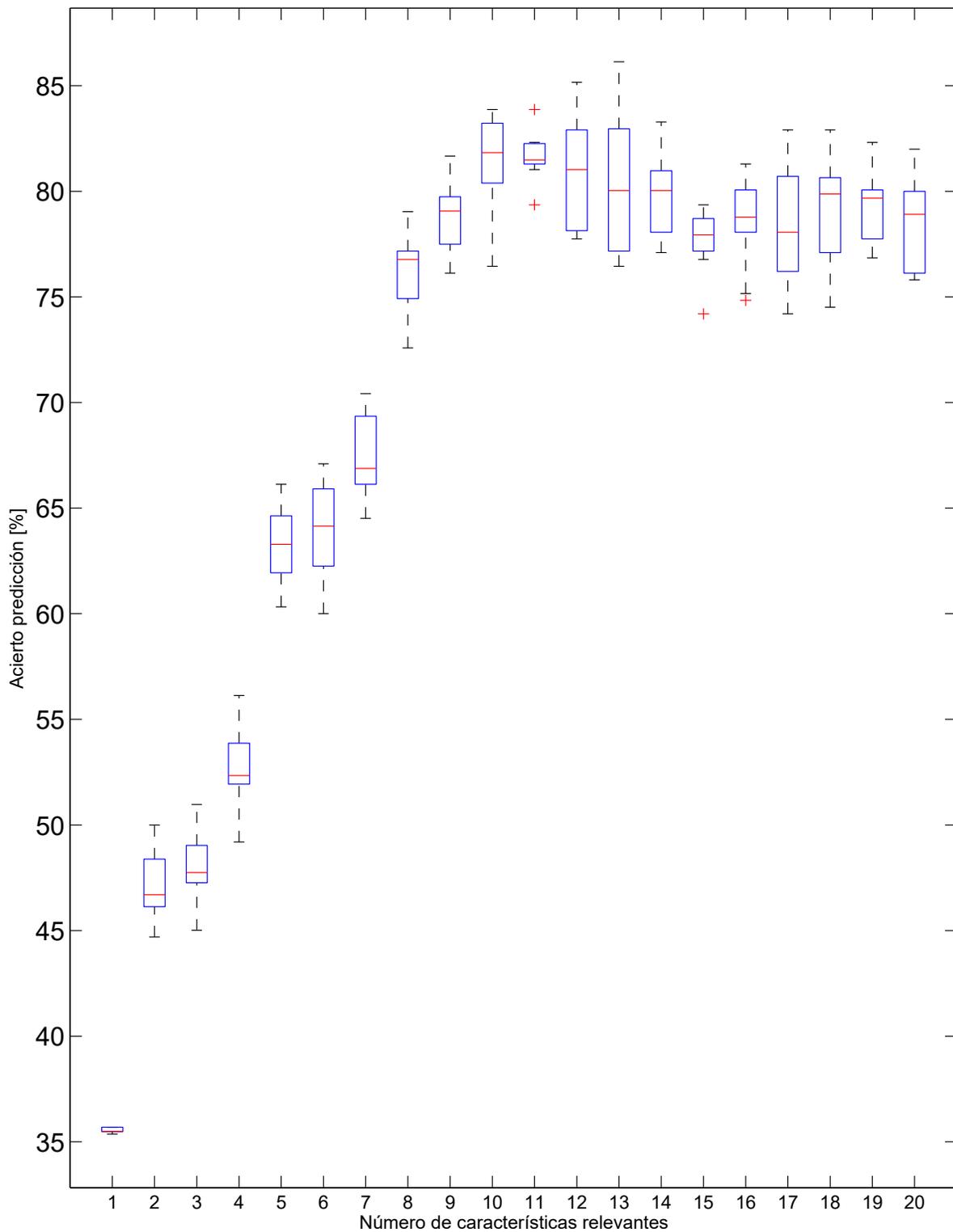


Figura 3.8: Clasificación automática de deserción empleando KNN para variables relevantes de CKA, 2016-I, deserción 2 periodos

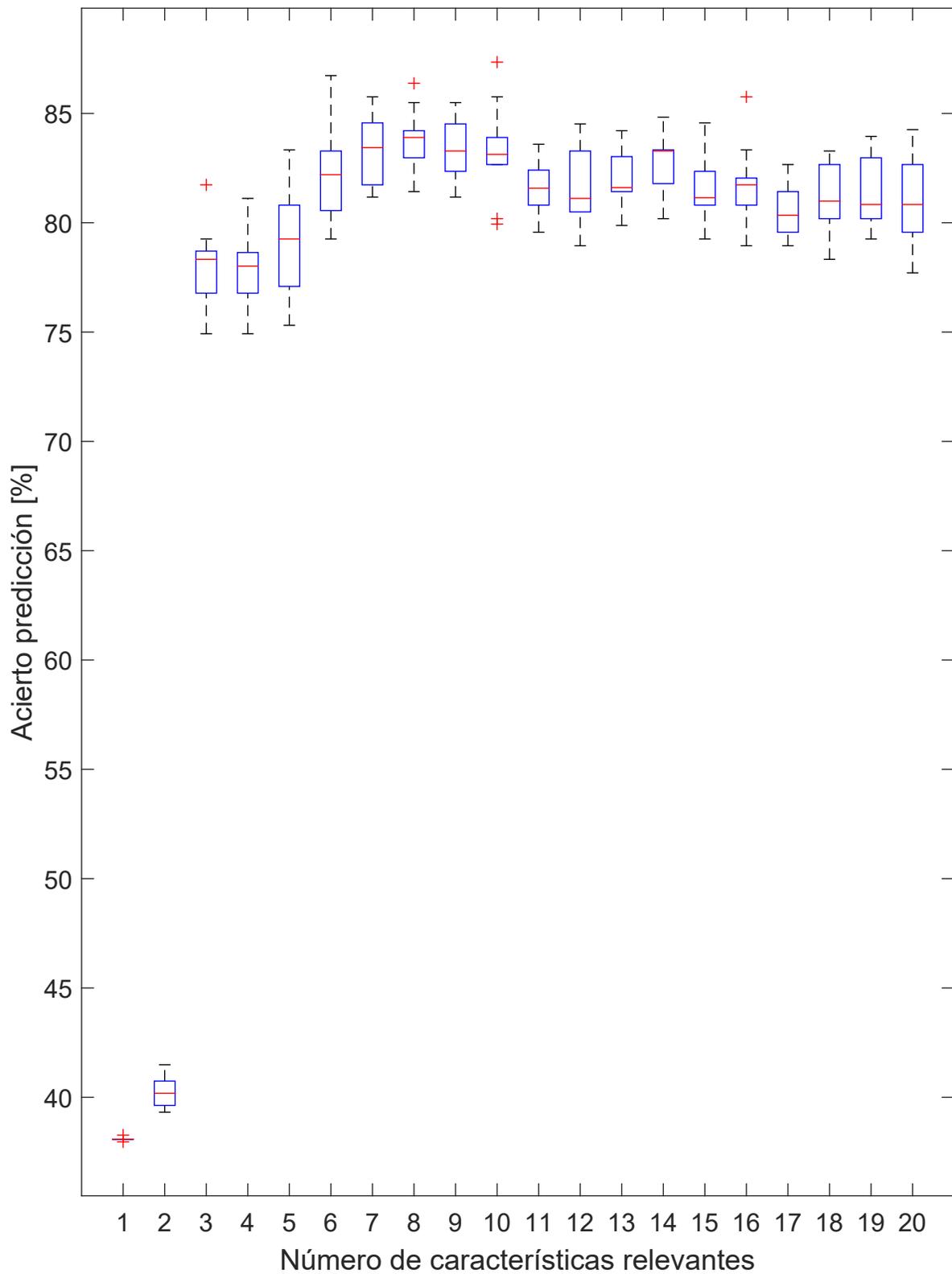


Figura 3.9: Clasificación automática de deserción empleando KNN para variables relevantes de CKA, 2016-II, deserción 1 periodo

Para los tres escenarios evaluados, los procesos automatizados de clasificación entre desertores y no desertores, a partir de las variables más relevantes jerarquizadas, alcanza más de un 75% de acierto.

Todo lo anterior permite señalar que existe una clara relación entre la atención de los estudiantes en el PAI y la deserción, especialmente la atención brindada en la línea BPS. Con estos resultados se demuestra que el PAI es un programa eficaz en el cumplimiento de su objetivo de evitar la deserción estudiantil de la UTP.

En el capítulo siguiente y a partir de los resultados hasta ahora obtenidos se procede a determinar el nivel de eficiencia en términos financieros del PAI.

Capítulo 4

Evaluación de la relación costo-beneficio del PAI

Con el fin de establecer las relaciones de costo-beneficio de la ejecución del PAI y su efectividad en la reducción de la deserción estudiantil, en este capítulo, se analizan diferentes experimentos de simulación de datos que permiten interpretar el nivel de eficiencia del programa y los costos ahorrados o causados por el mismo a la entidad.

4.1. Simulación: Eliminación de atenciones del PAI

Para cada uno de los tres escenarios que se analizan (semestre 2016-I deserción de un periodo, semestre 2016-I deserción de dos períodos, semestre 2016-II deserción de un período) se simula el efecto de eliminar las atenciones que brinda el PAI que evitan la deserción. En este sentido, se tiene grupo poblacional completo de estudiantes, el cual puede clasificarse en desertores y no desertores, y a la vez, estos se puede clasificar en: no atendidos PAI y atendidos PAI (ver figura [4.1](#)).

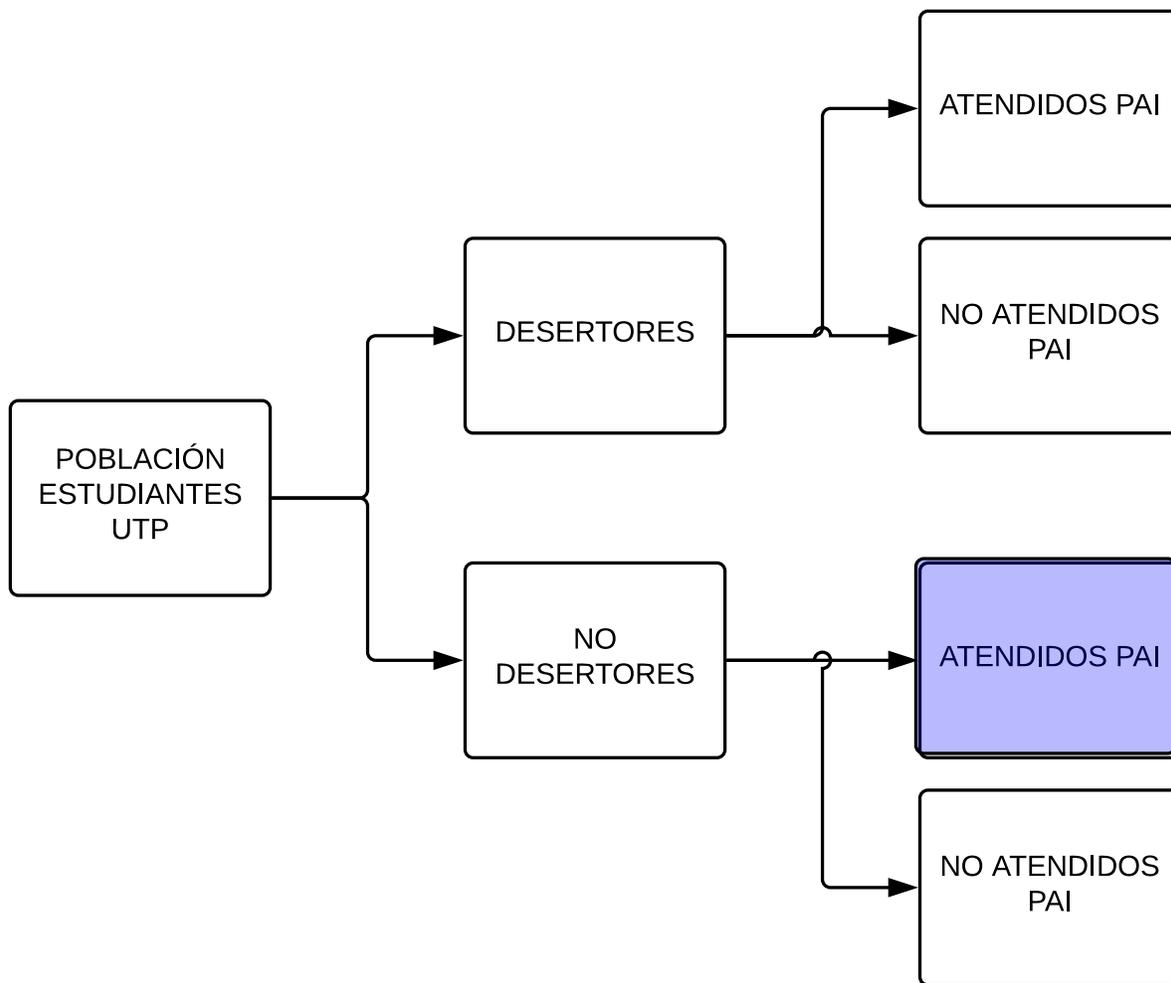
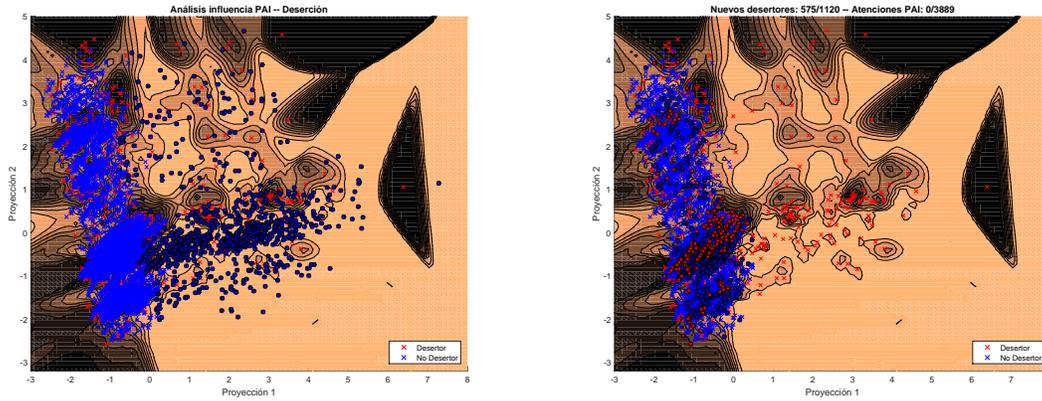


Figura 4.1: Conformación de grupos de interés para análisis

Tal como se resalta en la figura 4.1, el experimento de simulación se enfoca en determinar que ocurriría con los estudiantes atendidos en el PAI que no desertaron en caso que se eliminaran las atenciones del PAI. En tal caso existen dos posibilidades: 1) que no pase nada, es decir que el grupo se mantenga igual, 2) que los estudiantes se transformen en desertores, en esta situación quedaría nuevamente evidenciada la eficacia del PAI en evitar la deserción y además permite obtener una cifra de la eficiencia en tal labor. La dimensión del espacio de análisis de las variables es de 10 dimensiones, sin embargo, es posible observar lo que pasa en este experimento al proyectar las observaciones a un espacio de sólo 2 dimensiones por medio del análisis de componentes principales.

En la figura 4.2 se tiene la simulación para el semestre 2016-I deserción de un período, en particular, en la figura 4.2(a) se muestra la proyección del conjunto original de estudiantes y las curvas de nivel representan las zonas de no deserción y deserción, las áreas negras corresponden con ubicación de los desertores y las áreas claras con la ubicación de los no desertores. Sabemos que para el 2016-I deserción de un período, la cantidad de atenciones realizadas por el PAI a estudiantes no desertores fue de

3.889 atenciones (citas), en la figura 4.2(b) se simula la conformación de los grupos de deserción y no deserción ante la condición hipotética de que el PAI no hubiera realizado atenciones.



(a) Con atenciones PAI

(b) Sin atenciones PAI

Figura 4.2: Eliminación de atenciones PAI, 2016-I, deserción 1 periodo

Luego de la simulación se observa que sin las atenciones del PAI aparecerían 575 nuevos desertores, esto implica el crecimiento del desertores de 1.797 a 2.372 y significaría un cambio en la tasa de deserción de un período de 10,64 % a 14,1 %.

Para el 2016-I deserción de dos períodos se tiene la simulación de la figura 4.3. En la gráfica 4.3(a) se observa la proyección del conjunto original de estudiantes. Se sabe que para el 2016-I deserción de dos períodos, la cantidad de atenciones realizadas por el PAI a estudiantes no desertores fue de 3.976 atenciones (citas). En la figura 4.3(b) se simula la conformación de los grupos de deserción y no deserción ante la condición hipotética de que el PAI no hubiera realizado ninguna atención.

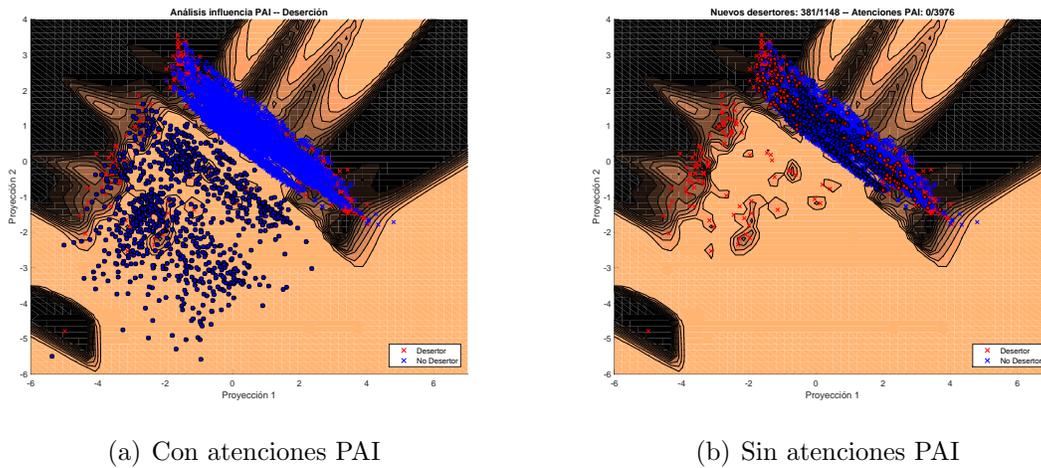


Figura 4.3: Eliminación de atenciones PAI, 2016-I, deserción 2 periodos

Luego de la simulación se observa que sin las atenciones del PAI aparecerían 381 nuevos desertores, esto implica el crecimiento del desertores de 1.382 a 1.763 y significaría un cambio en la tasa de deserción de un período de 8,21 % a 10,48 %.

Para el 2016-II deserción de un período se tiene la simulación de la figura 4.4. En la gráfica 4.4(a) se presenta la proyección del conjunto original de estudiantes. En 2016-II deserción de un período el PAI realizó 2.602 atenciones (citas) a estudiantes no desertores. En la figura 4.4(b) se simula la conformación de los grupos de deserción y no deserción ante la condición hipotética de que el PAI no hubiera realizado ninguna atención.

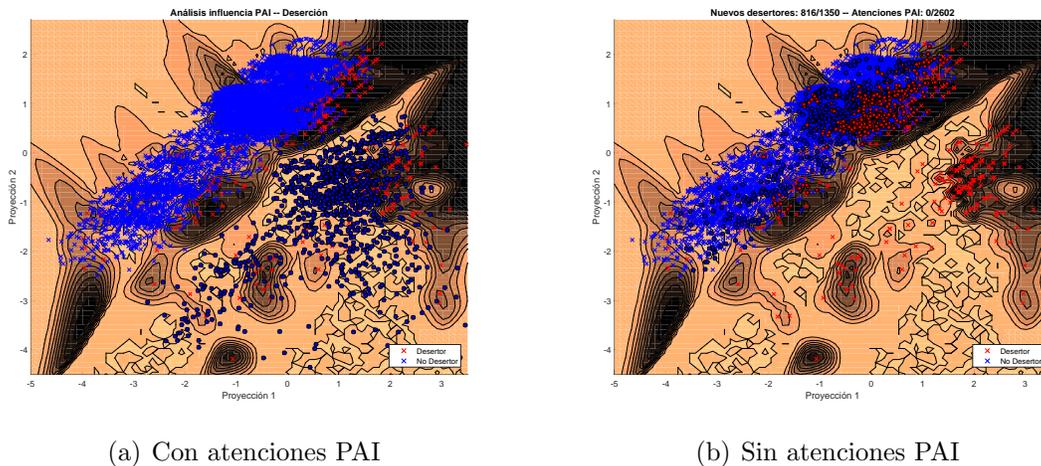


Figura 4.4: Eliminación de atenciones PAI, 2016-II, deserción 1 periodo

Luego de la simulación se observa que sin las atenciones del PAI aparecerían 816 nuevos desertores, esto implica el crecimiento del desertores de 1.706 a 2.522 y significaría un

cambio en la tasa de deserción de un período de 10,66 % a 15,76 %.

En la tabla 4.1 se sintetizan los resultados de la simulación de los niveles de deserción ante la condición hipotética de que el PAI no hubiera realizado ninguna atención.

	2016-I, 1 per.	2016-I, 2 per.	2016-II, 1 per.
No. nuevos desertores	575	381	816
Deserción original	10,64 %	8,21 %	10,66 %
Deserción simulada	14,1 %	10,48 %	15,76 %

Tabla 4.1: Resultados de deserción ante simulación de no atención PAI

A partir de los resultados se puede interpretar que la eficiencia del PAI es de evitar en aproximadamente 5 % la deserción de un período y en un 2 % la deserción de dos períodos.

4.2. Evaluación financiera del PAI

Con base en los resultados de cada una de las simulaciones presentadas en la sección 4.1, en esta sección se presenta la relación financiera de costo-beneficio del PAI. Para esto, primero, se identifica el listado de aquellos estudiantes que bajo la condición hipotética de no atención del PAI hubieran desertado. Adicionalmente, con base en información suministrada por el área financiera de la UTP se conocen los aportes totales del Estado entre los años 2.000 y 2.017, pero no el detalle de aporte por estudiante. En tal sentido, y sin pérdida de generalidad, se calcula la razón entre aportes anuales y el número de matriculados por año para obtener una cifra aproximada que es redondeada (por comodidad) a un valor de \$2.500.000, correspondiente al aporte del estado por cada matriculado.

Dado que se conoce la cantidad de semestres que ha matriculado cada uno de los estudiantes etiquetados como nuevos desertores, se pueden calcular los aportes promedio que ha dado el Estado por cada uno de esos estudiantes, según la siguiente ecuación:

$$\text{Aporte}_i = \text{Aporte promedio} \times \text{No. semestres matriculados}_i \quad (4.1)$$

donde $i = 1, 2, \dots, nd$ y nd es el número de nuevos desertores estimados en cada simulación.

Posteriormente se suman todos los aportes estatales dados a los estudiantes mencionados y se obtiene la cantidad total de aportes dada a los posibles desertores, ver ecuación (4.2).

$$\text{Aporte total} = \sum_{i=1}^{nd} \text{Aporte}_i \quad (4.2)$$

Por otra parte, se calcula el costo de ejecución del PAI en cada uno de los períodos analizados, el cual es la sumatoria de los honorarios de los profesionales del programa

más las becas y/o apoyos económicos brindados a los estudiantes que hayan sido beneficiados. De esta manera se puede cuantificar la inversión de la UTP en el PAI para cada semestre de interés en este estudio. Estos costos se resumen en la tabla 4.2.

Escenario	Apoyos económicos	Profesionales	Total
2016-I, des 1 per.	\$75.267.716	\$211.248.097	\$288.651.813
2016-I, des 2 per.	\$49.852.532	\$211.248.097	\$261.100.629
2016-II, des 1 per.	\$74.898.828	\$211.248.097	\$286.146.925

Tabla 4.2: Inversión de la UTP en el PAI para cada semestre de interés

Debe aclararse que la reliquidación de matrícula, la cual se considera un apoyo económico, no fue incluida en el cálculo de los costos del PAI debido a que no estaba disponible esta información.

En la tabla 4.3 se resumen los valores de inversión estatal susceptible a perderse en hipotéticos estudiantes desertores (cuando no hay atención PAI), el costo de ejecución del PAI, la diferencia entre estas cifras (Diferencia = Aporte total – Costo PAI) y el porcentaje relativo de inversión ($\% \text{ Inversión} = 100 \% \times \text{Costo PAI} / \text{Aporte total}$).

Escenario	Aporte total	Costo PAI	Diferencia	% Inversión
2016-I, des 1 per.	\$5.900.000.000	\$288.651.813	\$5.613.484.187	4,86 %
2016-I, des 2 per.	\$5.107.500.000	\$261.100.629	\$4.846.399.371	5,11 %
2016-II, des 1 per.	\$3.662.500.000	\$286.146.925	\$3.376.353.075	7,81 %

Tabla 4.3: Resumen de relación costo-beneficio del PAI

Los valores diferencia son altamente positivos, esto implica que la inversión que realiza la UTP en el PAI es muy pequeña en comparación con la posible pérdida de los aportes estatales en caso que los estudiantes desertaran. Para el 2016-II desertación un período, se aprecia de la tabla 4.3 la disminución de aportes totales por estudiantes simulados desertores, lo cual se explica por el hecho de que los hipotéticos estudiantes desertores (potenciales desertores) para el 2016-I son de semestres muy altos (muchos períodos matriculados), mientras que para 2016-II los potenciales desertores se encuentran en su mayoría en los primeros semestres. El cambio de configuración del grupo de posibles desertores está relacionado con el hecho de que la UTP ha estado apoyando la graduación de los estudiantes represados en últimos semestres, lo anterior hace que el subconjunto de estudiantes con muchos (más de 20) semestres matriculados se esté reduciendo en el grupo poblacional.

Capítulo 5

Conclusiones y recomendaciones

5.1. Conclusiones

A partir de las evaluaciones y simulaciones realizadas se puede concluir que el programa de atención integral (PAI) es eficaz en el cumplimiento de su objetivo de reducir la deserción de la Universidad Tecnológica de Pereira, es decir, tiene un impacto positivo en la permanencia y egreso exitoso de los estudiantes. La evidencia de esta afirmación se encuentra en los siguientes elementos:

- Se pudo demostrar la correlación existente entre las atenciones que realiza el PAI y la deserción estudiantil. Implicando que la reducción de la atención implica un aumento en la deserción.
- Se evidenció la alta relevancia relativa de las variables asociadas al PAI para describir el grupo poblacional.
- Cuando se simuló el caso hipotético en el que el PAI no realiza atenciones a los estudiantes, el índice de deserción crece en aproximadamente 4 % para deserción de un período y 2 % para la deserción de dos períodos.

Adicionalmente, al evaluar la relación de eficiencia entre la consecución de los objetivos del PAI y el gasto de recursos para tal fin, se pudo mostrar que: un escenario hipotético en el cual el PAI brindara ninguna atención a los estudiantes costaría a la Universidad, en el mejor de los casos, más de \$3.600.000.000. Por otra parte, la inversión semestral promedio que se realiza en el PAI para evitar la deserción y por ende la pérdida de los recursos mencionados es tan solo de \$278.633.122, es decir, la inversión con relación a la posible pérdida es de aproximadamente 7 % u 8 % (para el caso crítico).

En cuanto a las cifras que describen la población estudiantil desertora, cuantificadas a lo largo del informe, se puede determinar que:

1. La población desertora está especialmente ubicada en el rango de edad de mayores de 26 años.

2. El programa con mayor probabilidad de deserción para los semestres evaluados es Licenciatura en educación-español y comunicación audiovisual.
3. La facultad con mayor probabilidad de deserción es la de Tecnologías.
4. Los estudiantes con mayor probabilidad de deserción son aquellos que han matriculado menos de 3 semestres o que han matriculado más de 16 semestres.
5. Los estudiantes con promedio acumulado menor que 2 son los de mayor probabilidad de deserción.

Por otra parte, la población atendida por el PAI con respecto al grupo poblacional completo tiene el siguiente comportamiento:

1. Hay un marcado incremento en la cobertura de estudiantes por parte del PAI, con especial énfasis en la población de estudiantes que ingresan por primera vez a la Universidad.
2. Los estudiantes de menos de 17 años tienen una mayor probabilidad de asistencia al PAI.
3. Los programas: Ingeniería mecatrónica - Cicli I técnico profesional en mecatrónica jornada especial nocturna, Ingeniería física, Licenciatura en comunicación e informática educativa son los de mayor probabilidad de asistencia al PAI.
4. Las facultades de Ciencias agrarias y Ciencias básicas son las de mayor probabilidad de asistencia al PAI.
5. Los estudiantes con menos de 3 semestres matriculados son los que con mayor probabilidad son atendidos en el PAI.
6. Los estudiantes con promedio acumulado menor que 1 son los que más probabilidad de atención en el PAI tienen.

5.2. Recomendaciones

Con base en los datos presentados en este trabajo y en los análisis elaborados consideramos importante resaltar las siguientes recomendaciones:

- El Programa de Atención Integral (PAI) debe focalizar sus esfuerzos principalmente en la población con mayor probabilidad de deserción. En tal sentido es fundamental tener mayor participación en el PAI de los estudiantes con más de 16 semestres matriculados, con promedios entre 1 y 2, establecer un diálogo y un trabajo mancomunado con la facultad de Tecnologías y con el programa de Licenciatura en educación-español y comunicación audiovisual. En este sentido, se resalta que anexo a este informe se entrega un listado en excel que

permite identificar cada uno de los estudiantes que en la simulación son altamente sensibles a las atenciones del PAI en cuanto a su posible deserción y de esa forma concentrar esfuerzos en ellos.

- Es importante que la información de las variables analizadas sea de fácil acceso para la dependencia responsable del PAI con el fin de que de manera permanente pueda analizarse y reformular o afianzar las estrategias de trabajo con los estudiantes. En tal sentido es recomendable que exista un punto de información centralizado en la Vicerrectoría de responsabilidad social y bienestar universitario, es decir, una sistematización de la información.
- Es necesario evaluar algunas variables nuevas orientadas a determinar el riesgo de cada estudiante de la UTP de manera permanente y actualizada a lo largo de su vida universitaria, para ello puede diseñarse una aplicación para sistemas móviles que bajo algún incentivo fomente la actualización de datos y la respuesta de preguntas tipificadas para valorar riesgos.
- Debe fortalecerse la divulgación y difusión del PAI entre los estudiantes de todos los semestres, presentando claramente los alcances y actividades que realiza el programa.
- La información de cada una de las líneas del PAI: bipsicosocial, socio-economía y académica debe ser recolectada por la dependencia responsable pero reportarse de forma estandarizada y centralizada ante la Vicerrectoría a cargo del PAI.
- La información de todos los apoyos económicos brindada a los estudiantes debe centralizarse igualmente en el dependencia a cargo del PAI.
- Es importante la consolidación de fechas estandarizadas en la Universidad para el cierre o corte de información disponible, con el fin de obtener resultados informativos confiables y replicables a en todas las dependencias.
- Los problemas, dificultades, u obstaculos encontrados por el PAI deben reportase de forma sistematizada por el PAI ante la Vicerrectoría a cargo.

Apéndice A

Codificación de variables y categorías

Con el fin de simplificar la presentación de tablas y gráficas y mejorar la interpretabilidad del documento, los nombres de categorías de algunas variables fueron codificados. Adicionalmente, a las categorías de las variables les fueron asignados valores numéricos con el fin de realizar cálculos computacionales que permitieran establecer relaciones entre variables y dependencias.

A.1. Codificación de la variable género

En la tabla [A.1](#) se presenta la codificación de la variable género.

Categoría	Género	Valor numérico
G-00	Masculino	0
G-01	Femenino	1

Tabla A.1: Codificación género

A.2. Codificación de la variable edad

La edad es una variable discreta medida en años, esta variable pertenece al conjunto de los números naturales. Dada la alta dispersión de la edad en el grupo poblacional, se codifica por rangos según como se presenta en la tabla [A.2](#).

Categoría	Rango de edad (años)	Valor numérico
RE-00	0 a 10	0
RE-01	11 a 14	1
RE-02	15 a 17	2
RE-03	18 a 19	3
RE-04	20 a 21	4
RE-05	22 a 23	5
RE-06	24 a 25	6
RE-07	26 a 29	7
RE-08	30 a 35	8
RE-09	36 a 39	9
RE-10	40 y más	10

Tabla A.2: Codificación rango de edad

A.3. Codificación de la variable programa académico

A continuación se presenta la codificación de las categorías de la variable programa académico.

P-00 Administración ambiental.

P-01 Administración industrial.

P-02 Administración del turismo sostenible - Ciclo I técnico profesional en procesos del turismo sostenible.

P-03 Administración del turismo sostenible - Ciclo II tecnología en gestión del turismo sostenible.

P-04 Administración turismo sostenible.

P-05 Administración del turismo sostenible jornada especial nocturna

P-06 Artes plásticas.

P-07 Ciencias del deporte y la recreación.

P-08 Filosofía.

P-09 Ingeniería de sistemas y computación.

P-10 Ingeniería de sistemas y computación jornada especial nocturna.

P-11 Ingeniería de sistemas y computación (básicos ingeniería).

- P-12** Ingeniería de sistemas y computación convenio cuba.
- P-13** Ingeniería eléctrica.
- P-14** Ingeniería eléctrica (básicos ingeniería).
- P-15** Ingeniería electrónica jornada especial nocturna.
- P-16** Ingeniería electrónica jornada especial diurna.
- P-17** Ingeniería en mecatrónica - Ciclo I técnico profesional en mecatrónica jornada especial nocturna.
- P-18** Ingeniería en mecatrónica - Ciclo II tecnología en mecatrónica jornada especial nocturna.
- P-19** Ingeniería en mecatrónica - Ciclo III ingeniería en mecatrónica jornada especial nocturna.
- P-20** Ingeniería física.
- P-21** Ingeniería física (básicos ingeniería).
- P-22** Ingeniería en procesos agroindustriales.
- P-23** Ingeniería en procesos sostenibles de las maderas.
- P-24** Ingeniería industrial.
- P-25** Ingeniería industrial jornada especial nocturna.
- P-26** Ingeniería industrial (básicos ingeniería).
- P-27** Ingeniería mecánica.
- P-28** Ingeniería mecánica jornada especial nocturna.
- P-29** Ingeniería mecánica (básicos ingeniería).
- P-30** Licenciatura en lengua inglesa.
- P-31** Licenciatura en bilingüismo con énfasis en inglés.
- P-32** Licenciatura en artes visuales.
- P-33** Licenciatura en comunicación e informática educativa.
- P-34** Licenciatura en educación-español y comunicación audiovisual.
- P-35** Licenciatura en español y literatura.

- P-36** Licenciatura en etnoeducación y desarrollo comunitario.
- P-37** Licenciatura en etnoeducación.
- P-38** Licenciatura en filosofía.
- P-39** Licenciatura en matemáticas y física.
- P-40** Licenciatura en música.
- P-41** Licenciatura en pedagogía infantil.
- P-42** Licenciatura en pedagogía infantil - Ceres Belén de Umbría.
- P-43** Licenciatura en pedagogía infantil - Ceres Pueblo Rico.
- P-44** Licenciatura en pedagogía infantil - Ceres Puerto Caldas Pereira.
- P-45** Licenciatura en pedagogía infantil - Ceres Riosucio.
- P-46** Licenciatura en pedagogía infantil - Extensión San Andrés.
- P-47** Medicina.
- P-48** Medicina veterinaria y zootecnia jornada especial diurno.
- P-49** Química industrial.
- P-50** Tecnología eléctrica.
- P-51** Tecnología eléctrica energía Quindío.
- P-52** Tecnología en atención prehospitalaria jornada especial nocturna.
- P-53** Tecnología industrial.
- P-54** Tecnología industrial convenio cuba.
- P-55** Tecnología industrial - Ceres Belén de Umbría.
- P-56** Tecnología industrial - Ceres Mistrato.
- P-57** Tecnología industrial - Ceres Quinchía.
- P-58** Tecnología industrial - Ceres Santuario.
- P-59** Tecnología industrial jornada nocturna.
- P-60** Tecnología mecánica.
- P-61** Tecnología química.
- P-62** Tecnología en gestión del turismo sostenible nocturna.

A.4. Codificación de la variable facultad

En la tabla A.3 se presenta la codificación de la variable facultad.

Categoría	Facultad	Valor numérico
F-00	Bellas artes y humanidades	0
F-01	Ciencias ambientales	1
F-02	Ciencias básicas	2
F-03	Ciencias de la educación	3
F-04	Ciencias de la salud	4
F-05	Ingeniería industrial	5
F-06	Ingeniería mecánica	6
F-07	Ingenierías	7
F-08	Tecnología	8
F-09	Ciencias agrarias	9

Tabla A.3: Codificación facultad

A.5. Codificación de la variable jornada

En la tabla A.4 se presenta la codificación de la variable jornada.

Categoría	Jornada	Valor numérico
J-00	Diurna	0
J-01	Diurna especial	1
J-02	Nocturna	2
J-03	Nocturna especial	3

Tabla A.4: Codificación jornada

A.6. Codificación de la variable naturaleza del colegio

En la tabla A.5 se presenta la codificación de la variable naturaleza del colegio.

Categoría	Naturaleza del colegio	Valor numérico
NC-00	N/A	0
NC-01	Oficial	1
NC-02	Privado	2

Tabla A.5: Codificación de la naturaleza del colegio

A.7. Codificación de la variable departamento

A continuación se presenta la codificación de las categorías de la variable departamento de procedencia del estudiante.

D-00 Amazonas.

D-01 Antioquia.

D-02 Arauca.

D-03 Atlántico.

D-04 Bogotá.

D-05 Bolívar.

D-06 Boyacá.

D-07 Caldas.

D-08 Caquetá.

D-09 Casanare.

D-10 Cauca.

D-11 Cesar.

D-12 Choco.

D-13 Córdoba.

D-14 Cundinamarca.

D-15 Guaviare.

D-16 Huila.

D-17 Guajira.

D-18 Guainia.

D-19 Magdalena.

D-20 Meta.

D-21 Norte de Santander.

D-22 Nariño.

D-23 Putumayo.

D-24 Quindío.

D-25 Risaralda.

D-26 Santander.

D-27 San Andrés.

D-28 Sucre.

D-29 Tolima.

D-30 Valle del Cauca.

D-31 Vaupés.

D-32 Ext.

D-33 No disponible.

A.8. Codificación de la variable capital

En la tabla A.6 se presenta la codificación de la variable capital, la cual representa si el municipio de procedencia del estudiante es capital de departamento o no.

Categoría	Capital	Valor numérico
Cap-00	No	0
Cap-01	Sí	1

Tabla A.6: Codificación de capital

A.9. Codificación de la variable atendido PAI

En la tabla A.7 se presenta la codificación de la variable atendido PAI, la cual representa si el estudiante fue atendido o no por el PAI.

Categoría	Atendido PAI	Valor numérico
APAI-00	No	0
APAI-01	Sí	1

Tabla A.7: Codificación de atendido PAI

A.10. Codificación de la variable atendido línea biopsicosocial

En la tabla A.8 se presenta la codificación de la variable atendido línea biopsicosocial, la cual representa si el estudiante fue atendido o no en la línea biopsicosocial del PAI.

Categoría	Atendido BPS	Valor numérico
ABPS-00	No	0
ABPS-01	Sí	1

Tabla A.8: Codificación de atendido línea BPS

A.11. Codificación de la variable atendido línea académica

En la tabla A.9 se presenta la codificación de la variable atendido línea académica, la cual representa si el estudiante fue atendido o no en la línea académica del PAI.

Categoría	Atendido académica	Valor numérico
AAcad-00	No	0
AAcad-01	Sí	1

Tabla A.9: Codificación de atendido línea académica

A.12. Codificación de la variable atendido línea socio-económica

En la tabla A.10 se presenta la codificación de la variable atendido línea socio-económica, la cual representa si el estudiante fue atendido o no en la línea socio-económica del PAI.

Categoría	Atendido socio-económica	Valor numérico
ASE-00	No	0
ASE-01	Sí	1

Tabla A.10: Codificación de atendido línea socio-económica

A.13. Codificación de la variable número de atenciones

El número de atenciones es una variable discreta que determina el número de veces que un estudiante ha asistido al PAI. Esta variable pertenece al conjunto de los números naturales.

A.14. Codificación de la variable estrato

En la tabla [A.11](#) se presenta la codificación de la variable estrato.

Categoría	Estrato	Valor numérico
E-00	Alto	0
E-01	Bajo	1
E-02	Bajo-bajo	2
E-03	Medio	3
E-04	Medio-alto	4
E-05	Medio-bajo	5
E-06	Medio-bajo	6

Tabla A.11: Codificación de estrato

A.15. Codificación de la variable semestre de ubicación

El semestre de ubicación es una variable discreta que determina el semestre que para el sistema de registro y matrícula está cursando el estudiante con base en el número de créditos aprobados.

A.16. Codificación de la variable créditos aprobados

Créditos aprobados es una variable discreta que determina el número de créditos que ha aprobado el estudiante durante su permanencia en la UTP.

A.17. Codificación de la variable períodos matriculados

Períodos matriculados es una variable discreta que determina cuantos períodos semestrales ha matriculado el estudiante en la UTP.

A.18. Codificación de la variable apoyo económico

En la tabla A.12 se presenta la codificación de la variable apoyo económico, la cual representa si el estudiante recibe otras apoyos económicos por parte de la UTP tales como: bono de alimentación, bono de transporte, monitoria social, bono de matricula, reliquidación de matricula, entre otros.

Categoría	Apoyo económico	Valor numérico
AE-00	No	0
AE-01	Sí	1

Tabla A.12: Codificación de apoyo económico

A.19. Codificación de la variable promedio acumulado

Promedio acumulado es una variable numérica con representación en los reales que especifica el valor del promedio acumulado del estudiante según las respectivas ponderaciones otorgadas por los créditos.

A.20. Codificación de la variable desertor de un período

En la tabla A.13 se presenta la codificación de la variable desertor de un período, la cual indica si el estudiante desertó por un semestre de la UTP o no lo hizo: .

Categoría	Desertor 1 periodo	Valor numérico
DES1P-00	Sí	0
DES1P-01	No	1

Tabla A.13: Codificación de desertor de un período

A.21. Codificación de la variable desertor de dos períodos

En la tabla [A.14](#) se presenta la codificación de la variable desertor de dos períodos, la cual indica si el estudiante desertó por dos semestres consecutivos de la UTP o no lo hizo: .

Categoría	Desertor 2 periodos	Valor numérico
DES2P-00	Sí	0
DES2P-01	No	1

Tabla A.14: Codificación de desertor de dos período

Bibliografía

- [Alvarez-Meza et al.2017] Alvarez-Meza, A. M., Orozco-Gutierrez, A., and Castellanos-Dominguez, G. (2017). Kernel-based relevance analysis with enhanced interpretability for detection of brain activity patterns. *Frontiers in neuroscience*, 11:550.
- [Cortes et al.2010] Cortes, C., Mohri, M., and Rostamizadeh, A. (2010). Two-stage learning kernel algorithms. In *ICML*, pages 239–246.
- [Daza Santacoloma2006] Daza Santacoloma, G. (2006). Metodología de reducción de dimensión para sistemas de reconocimiento automático de patrones sobre bioseñales.
- [Daza Santacoloma et al.2007] Daza Santacoloma, G., Soto Mejía, J., and Castellanos Domínguez, G. (2007). Reducción de dimensión para el reconocimiento automático de patrones sobre bioseñales. *Scientia et Technica*, 13(37).
- [Peña2002] Peña, D. (2002). *Análisis de datos multivariantes*, volume 24. McGraw-Hill Madrid.