

# Procedimientos de mejoramiento y recuperación de datos en un análisis de cohorte educativa.

Ríos, Graciela.

Cita:

Ríos, Graciela (2011). *Procedimientos de mejoramiento y recuperación de datos en un análisis de cohorte educativa. XI Jornadas Argentinas de Estudios de Población. Asociación de Estudios de Población de la Argentina, Neuquén.*

Dirección estable: <https://www.aacademica.org/xijornadasaepa/12>

ARK: <https://n2t.net/ark:/13683/eeQG/NT3>



Esta obra está bajo una licencia de Creative Commons.  
Para ver una copia de esta licencia, visite  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

*Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.*

**XI Jornadas Argentinas de Estudios de Población, Ciudad de Neuquén, 21-23 de septiembre de 2011. Sesión 2. ¿La información disponible es pertinente y suficiente para monitorear el cumplimiento de los derechos educativos?**

**Título de la ponencia: Procedimientos de mejoramiento y recuperación de datos en un análisis de cohorte educativa**

Graciela Ríos. [ggracielarios@gmail.com](mailto:ggracielarios@gmail.com); [gracielarios@arnet.com.ar](mailto:gracielarios@arnet.com.ar)

Universidad Nacional de Córdoba(UNC). Facultad de Derecho y Ciencias Sociales (FDyCS)

**Palabras clave:** *Análisis de cohorte educativa, Mejoramiento/recuperación de datos*

**Resumen**

En el trabajo sobre *procedimientos de mejoramiento y recuperación de datos en un análisis de cohorte* educativa desarrollamos aspectos metodológicos y conclusiones que forman parte de una investigación más amplia sobre factores sociodemográficos y rendimiento académico realizada sobre una cohorte de 2893 estudiantes de la carrera de Abogacía de la Universidad Nacional de Córdoba.

El rendimiento académico es uno de los indicadores de la desigualdad que afecta a los estudiantes para acceder y permanecer con igualdad en los estudios universitarios. La posibilidad de su análisis con los datos disponibles nos enfrenta ante la exigencia de minuciosos análisis de las bases de datos con que contamos y la realización de procedimientos de mejoramiento y recuperación de información.

Además de los procesos de mejoramiento de que fueron objeto los datos utilizados para el análisis de cohorte que describimos, el trabajo da lugar a una caracterización general de las posibilidades y recaudos a tener en cuenta en relación a las fuentes de datos más utilizadas en los estudios sobre población / educación.

Formulamos conclusiones en una doble dirección. Por un lado, presentamos las relaciones de asociación encontradas entre rendimiento académico –definido a través de variables construidas a tal fin- y factores socio demográficos e identificamos los grupos de estudiantes más vulnerables a los riesgos de fracaso académico. Por otro lado, realizamos sugerencias metodológicas en relación al análisis de cohortes educativas y procedimientos de manejo de los datos. Ambos

grupos de conclusiones tienden a apoyar medidas institucionales para una mayor igualdad de oportunidades educativas.

### **1.- Introducción**

En el trabajo sobre *procedimientos de mejoramiento y recuperación de datos en un análisis de cohorte* educativa desarrollamos aspectos metodológicos y algunas conclusiones que forman parte de una investigación más amplia sobre factores sociodemográficos y rendimiento académico en una cohorte de estudiantes de la carrera de Abogacía de la UNC. Esta carrera, en el contexto de la población estudiantil de la UNC representa un 10 % de su total de alumnos. Analizamos el rendimiento académico como uno de los indicadores de las condiciones de desigualdad que afectan a los estudiantes para poder acceder y permanecer con igualdad en los estudios universitarios. La posibilidad de su análisis con los datos disponibles nos enfrentó a la exigencia de minuciosos análisis de las bases de datos con que contábamos y la realización de procedimientos de mejoramiento y recuperación de información.

La ampliación de la cobertura por parte de los sistemas de educación superior, junto a los objetivos universales de mejoramiento de la calidad, promovieron el desarrollo de un campo de investigación orientado a analizar el rendimiento de la educación tanto a nivel macro sistémico como institucional. Si bien la expansión cuantitativa de la enseñanza superior se muestra a través de medidas optimistas, los sistemas educativos, incluido el argentino, enfrentan desafíos importantes en relación al cumplimiento de las metas y objetivos formulados por las políticas educativas. Se reconoce que, si bien todos los sistemas educativos han experimentado en las últimas décadas una expansión cuantitativa notable, el crecimiento de la educación superior es incompleto por insuficiencias de calidad de resultados, grado de inequidad con que acceden distintos estratos de la sociedad, bajos rendimientos que se manifiestan de distintas formas preocupantes (diferencia que existe entre los tiempos teóricos y los tiempos reales de duración de una carrera, el desgranamiento, el promedio de los egresados, el abandono de las carreras universitarias, etc...) y también por su pertinencia con respecto a los requerimientos del entorno económico y social (CEPAL, 1990 ).

En esta vasta problemática, para el rendimiento de los estudiantes, hay investigaciones que estudian la influencia de distintos tipos de factores -sociales, históricos, políticos, institucionales, la socialización temprana- que inciden en los rendimientos alcanzados. El conocimiento de los

factores que subyacen a estos problemas, y que afectan sobre todo a los grupos de mayor riesgo o vulnerabilidad, es fundamental para el diseño de políticas universitarias.

Se reconoce que los antecedentes familiares del estudiante juegan un rol fundamental en la determinación del éxito educativo. Los alumnos provenientes de hogares que en su proceso de socialización han desarrollado capacidades cognitivas, lingüísticas, valores y aspiraciones que concuerdan con las exigencias de la educación formal tienen mayores oportunidades de éxito. La socialización primaria recibida en la familia o la secundaria mediada por las instituciones escolares, son importantes al analizar el rendimiento del estudiante en la universidad. En la investigación que se realizó durante el seguimiento de una cohorte educativa, este aporte teórico tomó particular relevancia no solo en la formulación de hipótesis y construcción de variables, sino también en las conclusiones obtenidas que permiten confirmar la influencia del contexto social y familiar en el rendimiento académico.

En la investigación, los problemas se presentaron cuando la definición teórica y la construcción metodológica de las variables socio-emográficas y académicas que serían sometidas a pruebas empíricas se vieron condicionadas por las características que poseía la información disponible. Esta dificultad implicó no solo evaluar detenidamente los datos y realizar ajustes, sino también diseñar estrategias de recuperación. Las bases de datos de los estudiantes de Abogacía de los años 2000 a 2005, con las que se trabajó, contenían problemas que exigieron volver sobre fuentes y documentos originales de registro, tales como una base de datos complementaria sobre exámenes rendidos y exámenes aprobados año por año y certificados analíticos de estudios secundarios concluidos.

En esta presentación nos proponemos:

- Describir las características que presenta la información socio demográfica educativa en general, y los problemas de las fuentes de datos utilizadas para el análisis de una cohorte educativa en la Universidad Nacional de Córdoba.
- Mostrar cómo se resolvieron los problemas metodológicos derivados de la información disponible.
- Formular recomendaciones sobre los procesos de recolección y tratamiento de los datos en las instituciones.

En el desarrollo de la presentación se hará en primer lugar una caracterización de las principales fuentes de datos sobre población y educación, en segundo lugar se describirán los procedimientos seguidos y dificultades encontradas en el análisis de cohorte educativa, y por último se hará una sintética presentación de resultados obtenidos y recomendaciones.

## **2.- Fuentes de datos sobre población y educación. Límites y posibilidades.**

La información que utilizamos en las investigaciones sobre los fenómenos educativos desde una perspectiva poblacional, según la fuente de que se trate, tiene frecuentemente limitaciones que obligan a que la evaluación y ajuste de la misma se convierta en una actividad importante de la investigación, tanto en tiempo como en esfuerzos dedicados. Esto se debe, en parte, a la complejidad de los procesos educativos sujetos a permanentes cambios por lo que la información pierde validez rápidamente y a menudo deben utilizarse proyecciones como medida alternativa; pero también se debe a problemas ocasionados en los procesos de recolección y tratamiento de datos, sobre todo los intermedios que realizan las instituciones educativas. A nivel de las instituciones responsables de recibir los datos de primera mano y transformarlos en documentación estadística, muchas veces se producen alteraciones porque se consideran a los datos prioritariamente como insumos para la gestión cotidiana de tipo administrativo. De esta forma, el dato deja de tener una finalidad estadística, se distorsiona y a veces hasta se pierde definitivamente.

Por estas razones, realizamos una breve caracterización de las fuentes e información que generalmente se utilizan para estudiar indicadores población/educación y de los recaudos que conviene considerar cuando se las utiliza.

Una de las medidas más utilizadas para analizar la relación entre población y educación es la de *eficacia*<sup>1</sup>; trabajada empíricamente como *eficacia externa* e *eficacia interna* de los sistemas educativos. Entre los indicadores de *eficacia externa* contamos con las tasas brutas y netas de cobertura escolar e indicadores de educación societal tales como las tasas de alfabetización, el nivel educativo alcanzado por la población que ya no asiste al sistema escolar, el cálculo de la esperanza de vida escolar y la supervivencia escolar. Algunas de las fuentes cuya consulta está

---

<sup>1</sup> Por necesidades de extensión, no se definen los indicadores sino que se da prioridad en el desarrollo a las fuentes que pueden dar cuenta de ellos.

disponible en los sitios oficiales de organismos nacionales e internacionales y que permiten obtener información necesaria para estas medidas son:

- *Informes Estadísticos y publicaciones de Organismos Internacionales* tales como el Centro Latinoamericano de Estadísticas (CELADE), la Organización para la Cooperación y Desarrollo Económico (OCDE), Organización de las Naciones Unidas (ONU), Información Educativa para América Latina y el Caribe (IESALC) -órgano autónomo adscrito a la UNESCO-, Sistema de Información sobre Tendencias Educativas en América Latina (SITEAL) de la Organización de Estados Iberoamericanos. En estas fuentes es posible encontrar datos históricos, de actualidad, proyecciones estadísticas de poblaciones educativas y análisis comparados para países que integran esos organismos internacionales.
- *Censos Nacionales de Población y Vivienda*. Ofrecen datos sobre la población que asistió y asiste a los sistemas educativos según su estructura por niveles, ciclos, modalidades o dependencia, y organizados según distintas características demográficas, edad, sexo, provincias o regiones. A partir de estos datos es posible construir distintos tipos de indicadores educacionales de eficacia externa.
- *Censos Provinciales* que, si bien se realizan generalmente con finalidades económicas, se suelen incorporar preguntas sobre educación de la población.<sup>2</sup>
- *Encuesta permanente de hogares*, realizada en nuestro país por el Instituto Nacional de Estadísticas y Censos, en algunas de cuyas bases de datos puede contarse con alguna información educativa mínima.
- *Anuarios Estadísticos de la Secretaría de Políticas Universitarias* del Ministerio de Educación de la Nación, que contiene información sobre el sistema educativo de nivel universitario según estructura del sistema por instituciones, carreras, matrícula, egresados y docentes. La información permite la construcción de distintos tipos de indicadores, aunque estos datos no pueden ser directamente comparados con los que pertenecen al Censo de Población y Vivienda ni obtener conclusiones a partir de la comparación.
- *Anuarios Estadísticos* de la Dirección Nacional de Información y Evaluación Educativa (DINIECE), del Ministerio de Educación de la Nación, que poseen información sobre educación común correspondiente a los niveles inicial, primario, secundario, superior no

---

<sup>2</sup> Es el caso del Censo de Población que en el año 2008 se realizó en la provincia de Córdoba, con la finalidad de contar con datos actualizados sobre población para resolver la coparticipación económica, que integró en las cédulas censales alrededor de siete preguntas sobre educación.

universitario y regímenes especiales del sistema educativo nacional. La información se desglosa por estructura de los niveles y ciclos, instituciones, carreras, matrícula, egresados y docentes. Estos datos tampoco son comparables en forma directa con los que pertenecen al Censo de Población y Vivienda, ni se pueden extraer conclusiones de esa comparación,

- *Informes Estadísticos Universitarios* publicados por las Universidades.<sup>3</sup>
- *Informes sobre evaluaciones institucionales*. Se cuenta con ellos, fundamentalmente para los casos de carreras sometidas a procesos de evaluación y acreditación por parte del Ministerio de Educación de la Nación. El acceso a ellos puede ser de carácter restringido.
- *Informes estadísticos de los Ministerios de Educación Provinciales*. No todas las provincias argentinas tienen sus estadísticas publicadas. En caso de no estar publicadas, el acceso a los datos exige gestionar autorizaciones.
- Informes Estadísticos provenientes de distintas Oficinas Ministeriales.
- *Informes Estadísticos Privados* publicados por grupos consultores particulares y realizados con una determinada finalidad.

En relación a las fuentes mencionadas, el recaudo más importante a señalar desde el punto de vista de la investigación, es que estas fuentes difícilmente son comparables entre sí, en forma directa, ni se pueden extraer conclusiones a partir de esta comparación. Los tiempos de recolección, el alcance y cobertura de la información varía de una fuente a otra y exigen recaudos y análisis prolijos.

Otras fuentes de datos dan cuenta de los indicadores educativos denominados de *eficacia interna*. La eficacia interna de los sistemas educativos medidas de procesos internos de los sistemas educativos tales como la carrera académica de los estudiantes, ingresantes, tasas de graduación, desgranamiento, deserción y otras que evalúan la calidad de los procesos, por ej. evaluación del rendimiento, los conocimientos efectivamente alcanzados por los estudiantes,

---

<sup>3</sup> Por ej. la Universidad Nacional de Córdoba a través de su Programa de Estadísticas publica anualmente un informe completo que permite construir indicadores académicos y poblacionales. Estos Anuarios Estadísticos están disponibles en la página web y en forma impresa. Los datos están organizados en indicadores estratégicos de la función docente universitaria. Los niveles de agregación de datos están constituidos por Universidad, Área de Estudio, Dependencia académica, Carrera, asignatura, Cátedra. La partición corresponde a Grado, Pregrado y Posgrado. Los ámbitos son Alumnos, Docentes, Planes y programas. En el ámbito Alumnos, las variables son 'características sociodemográficas por alumno', 'rendimiento académico por alumno'; en el ámbito Docentes, las variables son 'características sociodemográficas por docente', 'categorización docente. Se puede conocer el comportamiento de cada sector, ya sea alumnos, docentes u otros, área respecto de los datos que aportan al total. También, se puede observar fácilmente cómo aporta cada unidad académica al total de la categoría 'sin datos'.

razones profesor – número de alumnos etc... Para estos indicadores las fuentes mencionadas anteriormente pueden ser útiles en menor medida ya que aquellos se tratan de datos sistémicos globales y su grado de discriminación interna permite elaboraciones más bien limitadas para este tipo de indicadores que suponen internarse en instituciones o poblaciones educativas acotadas. Si se desea profundizar en indicadores e eficacia interna hay que acceder a datos que realizan discriminaciones detalladas dentro de datos globales. Esta información si no está publicada exige autorización institucional o, aún contando con ella, requiere internarse en la vida cotidiana de las instituciones y contar con la buena disposición de los agentes responsables para atender el requerimiento de datos. Entre fuentes con las que se cuenta para la eficacia interna mencionamos<sup>4</sup>:

- *Registros Estadísticos internos* de las instituciones educativas, entre ellas de Áreas de Enseñanza de Unidades Académicas Universitarias. Se trata por ejemplo, de bases de datos elaboradas a partir de documentos originales.
- *Fuentes primarias* tales como las fichas de inscripción o matriculación que completan los estudiantes, planillas de asistencia de cada asignatura, planillas de notas de evaluaciones parciales, actas de exámenes, certificados analíticos de escolaridad secundaria. Si bien la primera fuente de estos registros son los originales en formato de su impresión en papel, para las universidades estatales, en mayor o menor medida hoy existe la posibilidad de acceder a todos los datos censales y académicos de alumnos y profesores a través del Sistema de Información Universitaria (SIU). Este tipo de fuente también suele tener características de documentación confidencial y su acceso puede estar restringido.

Los recaudos para este grupo de fuentes se refieren básicamente a:

*Errores de omisión:* para detectar estos errores es necesario el análisis de cada una de las variables con las que se desea trabajar. Estos problemas pueden ocurrir entre otras causas porque:

a) quienes responden originariamente al relevamiento de información no cumplen con todos los requerimientos de información dejando amplios campos de los formularios sin respuesta. La consecuencia de esto es que las estadísticas registran un número tan elevado de casos sin datos que vuelven imposible considerar esa variable a los fines del análisis; b) a pesar de que los datos

---

<sup>4</sup> Además, cuando se intentan enfoques centrados en el sujeto de la acción, está la posibilidad de la toma de entrevistas y encuestas.



han sido respondidos en los formularios, luego no aparecen en las estadísticas. Por ejemplo, esto ocurre con la edad del estudiante, su procedencia geográfica, información sobre su escolaridad secundaria etc... Por distintas razones, quien realiza el registro en el sistema informático omite cargas; tal vez porque las pautas que rigen la organización del trabajo cotidiano lleva al personal que manipula datos en los niveles intermedios de gestión a descuidar la finalidad estadística de la información para priorizar las urgencias de la administración.

*Errores de cobertura:* Las estadísticas educativas adolecen con frecuencia de este problema; para el caso de universidades hay algunas de ellas para las que no es posible conocer su matrícula, discriminación por carreras o su egreso.<sup>5</sup>

*Inconsistencia de la información:* Tiene su origen en declaraciones de respuestas confusas en los formularios, por ejemplo, el tipo de escuela a la que asistió, confusión entre orientaciones, especialidades y modalidades de los estudios de nivel medio etc... Entre la información inconsistente es común también que se registren datos ingresando códigos incorrectos o inexistentes.

Además, cuando se analizan poblaciones educativas, es necesario conocer cuál es el momento del período lectivo en la que ha sido relevada la información. En educación la información suele ser relevada con distintos fines en el mes de marzo, a mitad de año o en el mes de diciembre. Los datos de matriculados al mes de marzo guardan diferencia con los que pueden ser relevados a la mitad de año o hacia la finalización del mismo. La captación fiel del desgranamiento, la repetición o el abandono depende de estas fechas. Por lo tanto se debe tener sumo cuidado si se realizan comparaciones entre años lectivos de no estar comparando cifras que no lo son por pertenecer a distintos momentos del año lectivo.

### **3.- Posibilidades y problemas que presentaron los datos en el análisis de una cohorte educativa y resultados alcanzados**

En el seguimiento de una cohorte de estudiantes universitarios para establecer relaciones de dependencia entre factores socio demográficos y rendimiento académico, nos encontramos con

---

<sup>5</sup> En la información que se refiere a sistemas educativos provinciales de nivel inicial, primario, medio superior no universitario, el problema es más grave. Existen quedan agentes educacionales o servicios educativos que no son alcanzados por los registros. Un ejemplo de ellos es el número real de escuelas que existe en una jurisdicción, el número de alumnos, las horas docentes etc.

gran parte de los problemas antes mencionados. La cohorte estudiada fue la de ingresantes en el año 2000 a la carrera de Abogacía de la UNC. Los problemas de información incompleta o inconsistencia originados en distintos tramos del tratamiento de la misma impidieron completar y ajustar los datos para los 2893 estudiantes que componían la cohorte y obligaron a trabajar con una muestra. Construimos una muestra aleatoria estratificada de 490 estudiantes que se diseñó de modo de mantener un margen de error absoluto máximo de 0,04 en las estimaciones que se realicen para las variables de proporciones. Este margen de error se fijó en función del costo de recolección y posibilidades de manejo de datos para la reconstrucción de las variables incompletas en información. Los subgrupos de alumnos –o estratos- sobre los que se construyó la muestra responden a una característica del rendimiento del alumno antes que a un atributo personal. Se optó por esta alternativa luego de explorar las bases y observar que la población se caracteriza más por estos rasgos que por sexo o situación conyugal u otros. Se tomó, entonces, como punto de partida para definir los estratos la cantidad de materias aprobadas por el estudiante al sexto año de carrera académica (año 2005), situación que supone un determinado nivel de avance. Así, resultaron cuatro estratos: estudiantes con 30 y más materias aprobadas, estudiantes con 14 a 29 materias aprobadas, estudiantes con menos de 13 materias aprobadas y estudiantes inactivos.

Las fuentes de datos que utilizamos fueron: Anuarios Estadísticos del Programa de Estadísticas de la UNC, bases de datos de los estudiantes de Abogacía de la UNC desde el año 2000 hasta el año 2005<sup>6</sup>, base de datos complementaria sobre exámenes rendidos y exámenes aprobados, año por año, para cada alumno integrante de la muestra, certificados analíticos de estudios secundarios concluidos consultados en el Archivo de Abogacía de la FDyCS pertenecientes a los 490 estudiantes que conforman la muestra; en los certificados analíticos fue posible encontrar información referida a colegio secundario al que asistió el alumno, tipo de gestión y localización geográfica del mismo, orientación de los estudios y calificaciones obtenidas.

---

<sup>6</sup> Se trata de bases de datos por año académico que contienen información que el estudiante declara en la ficha de matriculación, referida a las siguientes variables: sexo, estado conyugal, fecha de nacimiento, hijos, con quien vive, tipo de gestión y orientación de la escuela secundaria a la que asistió, otros estudios de nivel superior, trabajo, rama de actividad, cantidad de horas que trabaja, relación del trabajo con la carrera, cómo costea sus estudios, obra social, práctica de deportes, lugar donde practica deportes, sobrevivencia del padre y madre, máximo nivel educativo alcanzado por el padre y madre, rama de actividad del padre y madre. Estas bases fueron objeto de análisis a fin de estimar la completitud de los datos y valorar en qué medida era viable la aplicación de métodos en estas condiciones o la alternativa de reconstruir variables a estos fines.

Para poder construir una base de datos en la que solo se encontrasen los estudiantes que ingresaron en el año 2000 con las variables completas en información, tuvimos que evaluar las bases correspondientes al año lectivo 2000, año de ingreso de la cohorte, y también las posteriores, hasta la del año 2005, último año teórico de cursado de la carrera; este procedimiento permitió completar información para cada alumno en las variables a considerar en el estudio. Además, tomando como referencia el número de matrícula del alumno, excluimos de la base a los estudiantes que pertenecen a otros años de ingreso y los de pase universitario. Como en la base de datos del año 2005, se encuentran los estudiantes matriculados en ese año, pero no los inactivos a ese año, en las bases de datos de los años 2000 a 2004 identificamos los estudiantes no matriculados en el año 2005 pero con actividad académica en años anteriores, y los incluidos en la base por ser integrantes de la cohorte. Los estudiantes que habiendo cursado Introducción a los Estudios de la Carrera de Abogacía (IECA) nunca llegaron a aprobarlo, por lo tanto no pueden ser considerados alumnos, no fueron incluidos en la base de datos de la cohorte. Esta breve descripción, da cuenta de los pasos intermedios, que exigieron tiempo de dedicación y prolijidad, para poder contar con una base sobre la que fuera posible trabajar.

Una vez construida la base de datos de 2893 casos, se continuaba evidenciando carencia de datos para algunas variables, por ejemplo para situación de convivencia, condición laboral, fecha de nacimiento, escuela de la que proviene y todo tipo de dato vinculado a su historia educativa previa. También observamos la concentración de respuestas en la misma categoría de una variable por ejemplo, estado conyugal y tenencia de hijos

#### *Alternativas de solución a los problemas de datos incompletos*

Describimos a continuación algunas de las estrategias seguidas para poder completar y mejorar los datos. La edad, variable importante en la investigación, se reconstruyó a partir de la fecha de nacimiento o, en casos en que ésta tampoco estuviera registrada, a partir del número de documento de identidad luego de construir una tabla de cálculo sometida a reiteradas pruebas. Luego se volvió a controlar cuando se consultaron los certificados analíticos de escolaridad secundaria.

Una cantidad de datos personales de los estudiantes o de sus padres para los que no había información completa (horarios y tipo de trabajo, rama de la actividad, relación del trabajo con la carrera), no se pudo recuperar al no estar disponibles los registros o fichas originales, por lo que

nos vimos obligados a descartar la utilización de estas variables. Los antecedentes educativos del estudiante (orientación de los estudios de nivel secundario, tipo de gestión de la escuela a la que asistió y promedio) tampoco ofrecían información utilizable, pero fue posible recuperarlos aunque con una inversión mayor de tiempo y esfuerzos. Iniciamos nuevamente el relevamiento de esta información a partir de los certificados analíticos de escolaridad secundaria, documentación que se conserva en el Archivo de la Facultad. Por la dimensión de la población estudiantil y las dificultades de organización del Archivo, tanto por su espacio físico como por los procedimientos de reserva de documentación, esta tarea sólo podía realizarse para la muestra de alumnos. La modalidad de archivo de esta documentación en la Unidad Académica ocasionó dificultades no esperadas y llevó casi dos meses de dedicación.<sup>7</sup> Estos inconvenientes fueron los que determinaron que trabajásemos finalmente con la muestra con margen de error 0,04 y no una más ambiciosa en precisión con riesgo de error 0,03 porque implicaban 762 estudiantes.

Estos procedimientos, ponen en evidencia que sólo la construcción de la base de datos correspondiente a la muestra requirió una tarea casi artesanal de máximo cuidado para cada caso.

De los certificados analíticos se relevaron los siguientes datos: *promedio alcanzado en la escuela secundaria* (en la mayor parte de los certificados hubo que calcularlo), *ubicación geográfica de la escuela secundaria*, *dependencia de la escuela*, *orientación de la escuela*.

La información referida a antecedentes educativos del estudiante permitió completar la base de datos correspondiente a la muestra de 490 estudiantes, conservando en ella las demás variables e información de la ficha de matriculación. Resultó así una matriz compuesta por las siguientes variables independientes: *sexo*, *edad al ingreso*, *máximo nivel educativo alcanzado por el padre*, *máximo nivel educativo alcanzado por la madre*, *situación de convivencia*, *forma de costear los estudios*, *promedio alcanzado en la escuela secundaria*, *localización geográfica de la escuela secundaria*, *dependencia de la escuela, según sea estatal o privada*, *orientación de la escuela*; y las siguientes variables dependientes: *promedio obtenido en la carrera universitaria*, *avance en*

---

<sup>7</sup> Los certificados analíticos no fueron encontrados dentro de un legajo personal sino archivados en cajas de orden alfabético que no siempre guardan relación con el apellido del alumno. Tampoco los certificados analíticos tienen consignado el número de matrícula universitaria del alumno, mientras los certificados se identifican por nombre de alumno, cada alumno se identifica en la base de datos por su número de matrícula o documento de identidad pero no por su nombre y apellido. Entonces, debimos encontrar primero el nombre y apellido de cada uno de los alumnos que caen en la muestra para vincularlo luego al número de matrícula que corresponde en la base; esta tarea, que parece simple en procedimiento, insumió tiempos considerables porque implicó entrar a cada historia académica individual, a través del SIU Guaraní.

*la carrera, eficiencia al primer ciclo y eficiencia al segundo ciclo* (para la construcción de éstas últimas se utilizó información sobre cantidad total de exámenes rendidos, cantidad total de materias aprobadas, exámenes rendidos año por año, exámenes aprobados año por año).

El *avance en la carrera* se define por la cantidad total de asignaturas que el estudiante que ingresó en el año 2000 tiene aprobadas en el año 2005, es decir al estar cursando su sexto año académico (aunque no en todos los casos esté cursando las asignaturas que corresponden al sexto año). Puede considerarse una variable que en buena medida refleja la capacidad de estudiante para avanzar en una progresión sistemática de contenidos y en su propia formación secuencial de integración de conocimientos. Para la aplicación de los modelos de regresión, esta variable se trabajó de manera normalizada en una escala de 0 a 1.

El *Promedio*, entendido como el cociente entre la sumatoria de todas las calificaciones obtenidas en los exámenes rendidos por el estudiante incluidos los aplazos y la cantidad de materias rendidas puede considerarse que es un indicador basado en datos concretos y que la diferencia en el promedio de dos alumnos, aún considerados iguales en otros aspectos, estaría indicando mayor incorporación de capital humano para quien obtuvo un promedio superior. (Gertel, Giuliadori, Cassini, González; 2007). Para aplicar la técnica de regresión lineal, el promedio se trabajó con escala normalizada de 0 a 1.

Las variables, *eficiencia total, eficiencia en el primer ciclo y eficiencia en el segundo ciclo*, procuran captar una medida de los resultados de aprendizaje efectivamente obtenidos. *Eficiencia total* considera la relación entre los intentos por aprobar (exámenes rendidos) y lo efectivamente aprobado sin considerar los tiempos en que se lo ha hecho. En el caso del estudiante cuyo desempeño académico se caracteriza por haber aprobado todas las veces que se presentó a rendir examen, la eficiencia es igual a 1. *Eficiencia al finalizar el primer ciclo y eficiencia al finalizar el segundo ciclo* son medidas que incorporan la variable tiempo al considerar la cantidad de materias que formalmente corresponde tener aprobadas a los tres años de carrera o a los seis años. Estas últimas resultan del cociente entre la cantidad de asignaturas aprobadas y la cantidad de asignaturas que debiera tener aprobadas de acuerdo a su año de ingreso. La eficiencia al primero y segundo ciclo de la carrera es una medida que nos aproxima a la situación de rendimiento del estudiante, a medida que va acumulando años de atraso en su cursado, el promedio o la eficiencia

total pierden poder descriptivo dado que puede ser un estudiante que posea promedio alto, o una eficiencia total cercana a 1 pero con muchos años de atraso en su carrera.<sup>8</sup>

Luego, para probar la relación entre las variables socio demográficas y académicas realizamos pruebas con modelos de regresión lineal.<sup>9</sup>

#### *Resultados obtenidos con las regresiones lineales*

A través de los modelos de regresión obtuvimos porcentajes de la variación de las variables dependientes explicados con las variables independientes tomadas como regresoras. Las variables dependientes consideradas son la aproximación más precisa al rendimiento que fue posible construir con los datos existentes y con los que fueron recogidos especialmente para completar debilidades de la base de datos; y las variables independientes seleccionadas son las más relevantes y confiables dentro del cuadro de disponibilidades y procedimientos seguidos.<sup>10</sup>

La Tabla I, ilustra las variables independientes utilizadas en las regresiones lineales y la forma como fue construida a partir de la información disponible y mejorada.

<b>Tabla I. Variables independientes utilizadas en las regresiones lineales. Tipo de variable y método de construcción</b>			
<b>Variable</b>	<b>Tipo de variable</b>	<b>Método de construcción</b>	<b>Signo esperado</b>
Sexo	Catagórica Binaria	Masculino (0) Femenino (1)	+
Edad al ingresar	Catagórica	21 y más años (0) 17 a 20 años (1)	+
Situación de convivencia	Catagórica Binaria	Otras situaciones (0) Vive con sus padres (1)	+
Dependencia de la escuela secundaria	Catagórica Binaria	Estatad (0) Privada (1))	+
Localización de la escuela secundaria	Catagórica binaria	Otras provincias o Interior de Córdoba (0) Ciudad de Córdoba (1)	+

<sup>8</sup> Para *eficiencia en el primer ciclo* se consideró el divisor 21, para *eficiencia en el segundo ciclo* de la carrera se consideró el divisor 42. El mapa curricular de la carrera de Abogacía incorporado en Anexo IV permite apreciar gráficamente la justificación de 21 materias aprobadas correspondientes al primer ciclo, y 42 materias aprobadas correspondientes al segundo ciclo.

<sup>9</sup> Mediante este método se puede analizar cómo los cambios en la/s variable/s predictora/s afectan a la variable respuesta mediante el ajuste de un modelo para la relación funcional entre ambas.

<sup>10</sup> A través de una matriz de correlaciones/probabilidades de las variables consideradas independientes, se controló la inexistencia de correlaciones significativas entre ellas.

Promedio general obtenido en la escuela secundaria	Numérica	Escala 0 a 1	A mayor promedio en la escuela secundaria, mayor promedio en la carrera universitaria
Orientación de la escuela secundaria	Catagórica Binaria	Otras orientaciones no afines (0) Orientaciones afines -Ciencias Humanas y Sociales (1)	+
Forma de costear sus estudios	Catagórica	No cuenta con aporte familiar para costear sus estudios (0) Cuenta con aporte familiar para costear sus estudios (1))	+
Trabajo	Catagórica Binaria	Trabaja (0) No Trabaja (1))	+
Máximo nivel educativo alcanzado por el padre	Catagórica Binaria	Secundario incompleto o menos <sup>11</sup> (0) Secundario completo y más (1))	+
Máximo nivel educativo alcanzado por la madre	Catagórica Binaria	Universitario incompleto (0) Universitario completo o más (+)	+

### *Síntesis de los resultados obtenidos en las regresiones lineales*

A continuación, para una mejor apreciación comparativa de los coeficientes que resultaron significativos en la explicación de cada una de las variables consideradas, en la Tabla II observamos los estimadores obtenidos con el signo correspondiente y el nivel de probabilidad asociado resultantes de la regresión lineal.

<b>Tabla II. Comparación entre los principales coeficientes que resultaron significativos para las variables del rendimiento académico</b>					
	<b>Promedio</b> <b>R<sub>2</sub>= 0,55</b>	<b>Avance en la carrera</b> <b>R<sub>2</sub>= 0,43</b>	<b>Eficiencia total</b>	<b>Eficiencia en el primer ciclo</b> <b>R<sub>2</sub>= 0,43</b>	<b>Eficiencia en el segundo ciclo</b> <b>R<sub>2</sub>= 0,44</b>
Sexo Femenino <sup>12</sup>		+ 0,01 (0,7688)	+0,01 (0,6267)		
Sexo Masculino	-0,03 (0,1511)			-0,01 (0,7435)	(0,9915)
Edad al ingresar 17 a 20 años	0,08 (0,0136)				

<sup>11</sup> Pruebas reiteradas de aplicación de modelos de regresión lineal indicaron que se obtenían mejores resultados construyendo la variable a partir de la escolaridad secundaria.

<sup>12</sup> Aunque los coeficientes obtenidos para la variable sexo no son significativos en la explicación, se los incluye debido a que se la ha considerado como criterio clasificatorio en las descripciones presentadas y constituye una variable demográfica importante para conocer el signo que tiene en las pruebas explicativas.

Situación de convivencia-Otras situaciones		+0,08 (0,0014)		+0,06 (0,0163)	+0,09 (0,0001)
Localización de la escuela secundaria en la ciudad de Córdoba	+0,06 (0,0044)	+0,06 (0,0186)		+0,07 (0,0049)	0,06 (0,0227)
Promedio de la escuela secundaria	+ 0,48 (0,0001)	+0,55 (0,0001)	+0,18 (0,0044)	+0,56 (0,0001)	+0,56 (0,0001)
Situación laboral No trabaja	+ 0,06 (0,0076)				
Madre con estudios de nivel superior completos	+0,06 (0,0018)	+0, 17 (0,0001)	+0,07 (0,0354)	+0,15 (0,0001)	+0,19 (0,0001)
Padre con estudios secundarios completos y más	+0,09 (0,0002)	+0,05 (0,0727)	+0,08 (0,0051)	+0,05 (0,0227)	+0,07 (0,0138)

Nota: entre paréntesis el nivel de probabilidad asociado a la estimación.

Distintos aspectos del contexto familiar y educativo del estudiante tales como el *promedio obtenido en la escuela secundaria* y el *máximo nivel educativo alcanzado por sus padres*, arrojaron coeficientes de alta significatividad para explicar el rendimiento académico. Siendo estas variables representativas de una situación sociodemográfica total del estudiante en esa etapa de su vida, y al alcanzar una fuerza condicionante tan importante, nos pone de manifiesto la importancia que poseen estos factores estructurales para definir desigualdades educativas. Nos alerta también sobre el alcance y los distintos flancos que deben contemplar las políticas y acciones institucionales que se orientan a paliar las desigualdades que afectan el rendimiento.

La *situación laboral del estudiante*, arrojó una relación significativa entre el estudiante que no trabaja y más elevados *promedios* en la carrera universitaria. Pero no mostró asociaciones con el *avance en la carrera* o con la *eficiencia*. Si el promedio de las notas en los exámenes habla de la calidad de los aprendizajes logrados, podría decirse entonces que un alumno que no trabaja está en condiciones de lograr mejores aprendizajes, aunque esto no necesariamente asegura la graduación en los tiempos teóricos del plan de estudios.

La *localización de la escuela secundaria en la misma ciudad sede de la Universidad*, se relaciona en mayor medida con el *promedio* obtenido y la *eficiencia en el primer ciclo de la carrera* pero se debilita su poder explicativo para el *avance* en la carrera y *eficiencia en el sexto año*. Si bien la ubicación geográfica de la escuela secundaria identificada con el lugar de procedencia, podría indicarnos que los estudiantes mejoran su promedio al haber proximidad entre su procedencia geográfica y la Universidad, también podría indicarnos una diferencia en los niveles de



preparación de los estudiantes según la escuela de la que provengan se encuentre en la ciudad de Córdoba o en el interior y otras provincias.

La *situación de convivencia del estudiante* no arrojó los resultados esperados en relación a que vivir con los padres podía incidir positivamente en su rendimiento universitario. Contrariamente, se encontró que existen situaciones veladas tras la categoría *otras situaciones* que serían las que configuran un contexto favorable a los estudiantes. De manera semejante, contar *con ayuda familiar para costear sus estudios*, no pudo encontrar significatividad en este trabajo pero deja el interrogante acerca de los procedimientos de recolección de información e instrumentos para acceder de una manera más sensible a este dato.

La *dependencia de la escuela secundaria según pertenezca al sector estatal o privado y la orientación de los estudios de nivel medio*, no evidenciaron ninguna relación con de las variables consideradas para el rendimiento.

Respecto de la variable *edad al ingresar*, no se obtuvieron estimadores de significación. La pertenencia de *sexo*, las pruebas no arrojan resultados significativos en los coeficientes de las regresiones como para explicar la variabilidad del rendimiento académico.

Los estudiantes inactivos, de los que no puede decirse que sean desertores, constituyen un grupo que para un conocimiento más profundo de los factores que los han llevado a esta situación académica debería emprenderse un estudio que posibilite la localización del joven y prevea instrumentos sensibles a la resistencia a dar respuestas.

#### **4.- Conclusiones**

Las conclusiones del trabajo se formularon en dos direcciones fundamentales. Por un lado, la identificación de rasgos sociodemográficos de los estudiantes de Abogacía que se asocian a su rendimiento académico, por otro lado los procedimientos metodológicos referidos al tratamiento de los datos educativos disponibles para la realización de estudios en esta línea.

En el primer grupo de conclusiones se puede decir que la bibliografía, los datos y las pruebas estadísticas realizadas permiten avanzar en el conocimiento e identificación de factores socio demográficos que inciden en el rendimiento académico de los estudiantes.

A través de la técnica de regresión lineal, se encontró que algunas variables socio demográficas de los estudiantes tienen una fuerza explicativa del 40% y hasta 55% de la variabilidad del

*promedio*, del *avance en la carrera* y de la *eficiencia en el primer ciclo de la carrera* y en el *segundo ciclo*. El porcentaje de las dimensiones del rendimiento académico que no se puede explicar a través de las variables socio demográficas está relacionado con percepciones y expectativas de los estudiantes que en alguna medida se identifican a través de la encuesta aplicada a estudiantes asistentes a clases, pero que no es desarrollada en esta presentación por cuestiones de espacio.

En síntesis, el *promedio obtenido en la escuela secundaria* y el *máximo nivel educativo alcanzado por el padre y la madre*, son los mejores predictores de su rendimiento académico en la carrera universitaria de Abogacía. Continúan en significatividad por su asociación la *situación laboral para el caso de quienes no trabajan*, la *localización de la escuela secundaria* en Córdoba Capital y la *situación de convivencia, con la observación* que vivir con los padres no necesariamente guarda relación con mejor rendimiento.

Estos datos sobre la población estudiantil de Abogacía y los factores asociados al rendimiento académico, indican que las acciones institucionales orientadas a prevenir el desgranamiento y mejorar el rendimiento deberían considerar prioritariamente al grupo de estudiantes con contextos familiares caracterizados por menores niveles de escolaridad de sus padres y bajos rendimientos en la escuela secundaria. De la misma manera, identificar al grupo de estudiantes de procedencia distinta a la misma ciudad de Córdoba. No obstante, el carácter estructural social y familiar de las condiciones que explican la variabilidad en el rendimiento académico indica que las desigualdades son profundas en su origen y exigirían acciones compensadoras de una magnitud que permitieran abarcar estos flancos.

En el segundo orden de conclusiones señalado, en el trabajo se pudo evaluar el estado de los datos educativos, sus posibilidades y límites para cada una de las variables demográficas que caracterizan al estudiante. La descripción de los procedimientos seguidos en relación a los datos, son una expresión de las características que tienen los análisis de cohortes educativas debido al estado de la información. A pesar de la idoneidad técnica con que los datos son elaborados en las oficinas responsables de las estadísticas universitarias, hasta llegar a ese punto, siguen un camino que debilita su validez y confiabilidad. Desde la respuesta a los cuestionarios, manipulación intermedia y registro informático, las rutinas cotidianas con frecuencia hacen que se pierda de

vista la importancia de los datos como insumos para las estadísticas, la gestión y la investigación y solo se vea en ellos la utilidad para la gestión de alumnos.

Ante los cuidados que se debieron tener al valorar los datos y mejorar la completitud y calidad de los mismos, surgen recomendaciones destinadas a los procesos intermedios de recepción y elaboración de la información que tiene lugar en las Unidades Académicas Universitarias.<sup>13</sup> Las recomendaciones se ubican en distintos niveles de trabajo institucional sobre: decisiones políticas de gestión, agentes institucionales responsables de los procesos intermedios de elaboración de la información, reserva de la documentación que es fuente primaria de información e ingreso de datos en el sistema informático.

Junto a las áreas de gestión académica debiera organizarse un sector de *información y evaluación de la educación entre* cuyas funciones prioritarias estaría la de promover la elaboración de información institucional sobre procesos educativos, producir diagnósticos y aportar permanentes insumos para la gestión. Es cada vez más necesaria un área de estas características que responda a requerimientos de mecanismos institucionales de control de gestión docente, monitoreo de planes de estudio y rendimiento académico. La toma de conciencia de la necesidad de estos mecanismos permitiría sostener actividades desarrolladas de manera especializada, y no como tareas adicionales que se suman coyunturalmente a agentes y áreas académicas para las que no está asignada esta función o no preparadas técnicamente para esto. El desarrollo actual de sistemas y tecnologías de información con grandes potencialidades, comunes a todo el sistema universitario, es un elemento favorable con el que ya se cuenta para conformar áreas de trabajo con personal especializado.

Otra medida que forma parte del nivel de decisiones institucionales es la de definir los perfiles laborales para los agentes responsables de los procedimientos de información y evaluación de la educación. El perfil formativo debe hacer posible el aprovechamiento de las posibilidades de los nuevos sistemas de información educativa además de organizar formas ágiles de respuesta a las necesidades de información de la institución.

---

<sup>13</sup> Las recomendaciones no pueden generalizarse a toda la Universidad, sino que se refieren exclusivamente a la FDyCS, carrera de Abogacía, más allá de que existan características compartidas con otras Unidades Académicas de las que no se puede dar cuenta en este trabajo.

En los agentes institucionales que son responsables de los primeros pasos del dato educativo hasta que se convierte en información estadística, se ubica una parte importante de las dificultades. La atención puesta en el cuidado de los datos como insumos de las decisiones institucionales, la gestión y la investigación es una dimensión del trabajo que está ajena porque priman las urgencias de la administración cotidiana. Como consecuencia de los errores que se suceden no solo se relega la importancia de los datos distorsionándolos sino que hasta se pierden definitivamente.

Otra recomendación tiene que ver con el reserbo de la información. La existencia de un legajo personal para cada estudiante dentro del cual se resguarde la documentación que contiene datos sobre los estudiantes no puede ser descartada porque existan archivos digitales. El reserbo de estos legajos debe contar necesariamente con espacios adecuados y personal preparado en archivología para un correcto ordenamiento y fácil recuperación.

Creemos, que ambos grupos de conclusiones son un aporte para las decisiones institucionales en relación a las finalidades de contribuir a una mayor igualdad de oportunidades educativas.

### **Bibliografía**

ARGENTINA; MINISTERIO DE EDUCACIÓN, CIENCIA Y TECNOLOGÍA. Secretaría de Políticas Universitarias. Anuarios de Estadísticas Universitarias 2000-2006.

ARGENTINA; MINISTERIO DE EDUCACIÓN, CIENCIA Y TECNOLOGÍA (1997) Secretaría de Políticas Universitarias. Sánchez Martínez Eduardo; Estadísticas y decisiones.

CU BALÁN, Guadalupe; ARAGON NAAL Faustina (2004). *El perfil sociodemográfico y su impacto en el rendimiento académico de los alumnos de la universidad autónoma de Campeche. México*. Publicación del Programa Institucional de tutorías de la UAC.

BRIGIDO, A.M. (2004). *La equidad en la educación argentina. Un análisis de las desigualdades en la distribución de la educación*. Ed. Universitas-Facultad de Filosofía y Humanidades de la UNC, Córdoba.

CÚ BALÁN Guadalupe; ARAGÓN NAAL Faustina (2002). *El perfil sociodemográfico y su impacto en el rendimiento académico*. Universidad Autónoma de Campeche. México.

DÍAZ ROSAS Francisco (1995). *La predicción del rendimiento académico en la universidad: un ejemplo de aplicación de la regresión múltiple*. Ediciones Universidad de Salamanca

DI GRESCIA Luciano (2004). *Acceso a la educación universitaria, evolución y determinantes para el caso argentino*. Anales de la Asociación Argentina de Economía Política

DI GRESCIA, PORTO, RIPANI. (2002). *Rendimiento de los Estudiantes de las Universidades Públicas Argentinas*. Documento de Trabajo N°45. Universidad Nacional de La Plata.

GERTEL, GIULIODORI, CASINI, GONZALEZ (2006). “*Rendimiento académico y éxito de los estudiantes de la Facultad de Ciencias Económicas de la UNC*”. En Porto, A.(ed). Mecanismos de admisión y rendimiento académico de los estudiantes universitarios. Cap II. Edit. de la Universidad de La Plata.

DiNIECE. (Área de Investigación. Proyecto DINIECE – UNICEF Seguimiento y monitoreo para el alerta temprana). Buenos Aires.

LANDI,J.; GIULIODORI, R. (2001). “*Graduación y Deserción en las universidades nacionales*”s. En Jozami; Sánchez; (comp.). *Estudiantes y profesionales en la Argentina. Una mirada desde la Encuesta Permanente de hogares*. Ed. Universidad de Tres de Febrero. Provincia de Bs.As.

LLACH, J; MONTOYA; R. (1999). *Educación para Todos*. IERAL. Córdoba.

MONTERO ROJAS, E., VILLALOBOS PALMA J. y VALVERDE BERMUDEZ, A. (2007). *Factores institucionales, pedagógicos, psicosociales y sociodemográficos asociados al rendimiento académico en la Universidad de Costa Rica: Un análisis multinivel*. Revista RELIEVE, v. 13, n. 2, p. 215-234.

ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT (OECD). (2008) *Education at a Glance*. OECD Indicators. Centre for Educational Research and Innovation. París

ELIEZER de los Santos. (2005) OEI. *Revista Iberoamericana de Educación*. (2005) Eliézer de los Santos. *Procesos de permanencia y abandono escolar en Educación Superior*. Universidad de Colima, México.