



Nombrado Doctor Honoris Causa en 2006

Sr. Rector Magnífico de la Universidad Carlos III de Madrid Excmo. Sr. Don Gregorio Peces-Barba Martínez

Autoridades Académicas

Colegas y amigos

Quisiera expresar mi profunda emoción y agradecimiento por este título que me ha otorgado la Universidad Carlos III. Este título representa el mayor reconocimiento que un académico puede aspirar ya que esta otorgado por una institución del más alto prestigio como es la Universidad Carlos III y con la que me unen profundos lazos tanto de trabajo como afectivos.

Estos lazos se han ido fortaleciendo con mis continuas visitas al Departamento de Estadística de esta Universidad donde he realizado trabajos de investigación con el Dr. Peña y donde he tenido la oportunidad de dictar cursos para estudiantes de doctorado y hasta dirigir una tesis de doctorado. Durante todos estos años he visto crecer al Departamento de Estadística para convertirse en uno de los centros de investigación y enseñanza de la Estadística de mayor excelencia de Europa.

El hecho de que este reconocimiento provenga de una Universidad española tiene para mi un significado muy especial, ya que mi origen me liga con España de una manera muy especial y estrecha.

Primero como argentino estoy ligado a España por ser nuestra madre patria y como tal esa Nación influyó de una manera determinante en nuestra cultura y nuestra idiosincrasia. Por eso, cuando estoy en España me siento en mi casa. Por un lado encuentro costumbres que son muy similares a las argentinas reconociendo sus orígenes y por el otro recibo el afecto que me prodigan los españoles.

También me liga España mi origen judío sefardí. No puede menos que recordar que España es la patria de mis ancestros. Mis padres, ambos ciudadanos turcos nacieron y vivieron en Turquía hasta que de adultos emigraron a la Argentina. Ellos hablaban un único idioma: el ladino, también llamado judeo-español que no es otra cosa que el español antiguo. Durante más de 400 años los sefarditas conservaron el idioma español viviendo en el imperio turco. Una anécdota familiar que muestra hasta que punto estaban identificados los sefarditas de Turquía con el idioma español fue protagonizada por mi abuela. Mi abuela vivía en la ciudad de Gallípoli en Turquía en el barrio judío donde el único idioma que se hablaba era el ladino. Cuando mi abuela llegó a la Argentina, ella al observar que en Buenos Aires todos hablaban español, extrañada le preguntó a mi padre en ladino.

¿Dónde me trujiste, son todos judeos aquí?

Lo que en castellano actual sería:

¿Donde me trajiste, son todos judíos aquí?

Para mi abuela el idioma español era sinónimo de judeidad

Quisiera recordar brevemente como se desarrolló mi vocación científica que me llevó finalmente a investigar en el área de la Estadística Robusta.

Cuando realizaba mis estudios secundarios, eran varias las áreas del conocimiento que atraían mi atención pero la Matemática tenía una característica que la distinguía de otras ciencias. Sus enunciados o leyes eran verdades absolutas que no requerían para su validez ni la verificación experimental como en las Ciencias Naturales ni dependían de juicios de valor u opiniones subjetivas como en las Ciencias Sociales. En las ciencias sociales como la Historia,

Economía o Sociología podían convivir teorías diferentes y la aceptación de una a otra estaba muchas veces ligada a ideologías y no a factores objetivos. En las ciencias naturales la verificación de una teoría es generalmente experimental y por lo tanto pueden tenerse dudas sobre como se llevaron a cabo los experimentos que la validaron.

En cambio el criterio de verdad de una teoría matemática era únicamente lógico y ajeno a toda consideración ideológica, moral religiosa o experimental. En mi inseguridad adolescente la matemática era la única ciencia que me daba garantía de verdades absolutas y esa fue una de las razones que me hicieron elegirla para mis estudios universitarios.

Otra razón para elegir esta disciplina es el profundo sentido de belleza que se encuentra en la matemática. Esa belleza se manifiesta en los métodos de demostración, la armonía y profundidad de algunos de sus resultados y en la misma experiencia de hacer matemáticas.

Resulta muy difícil transmitir el sentido de la belleza de la Matemática. Cito un pensamiento del matemático húngaro Paul Erdős que decía

"Preguntar porque los números son bellos es como preguntar porque es bella la novena sinfonía de Beethoven. Si uno no se da cuenta solo porque es bella, nadie se lo podrá explicar."

Búsqueda de verdades absolutas y un profundo sentido estético de belleza. Estas fueron las principales razones que me llevaron a estudiar Matemáticas. Puedo decir que disfruté plenamente todos los años que estudié Matemáticas en la Universidad de Buenos Aires. Además de estudiar las diferentes teorías en los libros, desde muy temprano se nos planteaban en los cursos problemas a resolver que eran verdaderos desafíos a nuestra capacidad matemática y que nos sumergían en el mundo maravilloso de esta ciencia.

Quisiera agradecer y recordar a mis maestros que tanta importancia han tenido en mi formación en esta etapa de mi vida En particular debo mencionar a Micha Kotlar y Gregorio Kilmovsky.

Luego de finalizar la Licenciatura en Matemática vino un momento muy difícil. Tuve que decidir en que rama de las Matemática iba a encauzar mis estudios de doctorado. Por un lado la Matemática Pura me atraía especialmente por la libertad de elegir los problemas a investigar basado solamente en el interés matemático que ellos tenían sin considerar sus posibles aplicaciones. Pero por otro lado sentía la responsabilidad de trabajar en temas que fueran útiles a la sociedad en su conjunto y no solo a la comunidad matemática y que esa utilidad pudiera surgir en tiempos relativamente cortos. Es cierto que muchos de los temas que se desarrollaron como ramas de la Matemática pura sin tener en consideración ninguna posible utilidad, con el tiempo resultaron ser herramientas formidables para resolver importantes problemas prácticos. Un ejemplo es la Teoría de Números que se desarrolló como una rama totalmente abstracta de la Matemática sin tener en consideración ninguna aplicación. En el presente, la Teoría de Números se ha convertido en una herramienta fundamental para el desarrollo de la Criptografía.

Sin embargo no siempre ocurre así, y cuando ocurre las aplicaciones pueden aparecer luego de muchos años. En ese momento yo necesitaba que esa utilidad social estuviese garantizada y que surgiese luego de un tiempo corto.

Finalmente mi preocupación por la utilidad social de mi trabajo de investigación me llevó a elegir el área de Estadística para realizar mis estudios de Doctorado. La Estadística es una Ciencia que se nutre en gran medida de herramientas matemáticas, en especial de la Teoría de Probabilidades pero que no puede considerarse una rama de las Matemática, ya que tiene sus propios objetivos de estudio. Para realizar mis estudios de doctorado en Estadística elegí la Universidad de California, Berkeley. La razón de esta elección fue que en ese momento en esta Universidad se daba especial énfasis a los aspectos de la Estadística que requerían un mayor uso de las matemáticas. De esta manera mi transición de la Matemática a la Estadística fue menos traumática y pude continuar disfrutando de la belleza matemática mientras investigaba y resolvía problemas que consideraba tenían interés para la sociedad. Durante mis estudios de doctorado en Estadística en la Universidad de California, Berkeley tuvieron una especial influencia los doctores Bickel y Lehmann.

Allí aprendí una fórmula que continué durante toda mi carrera científica. Trabajar en problemas estadísticos cuya resolución requiriese utilizar mi entrenamiento matemático y que simultáneamente tuvieran una utilidad práctica para el análisis de datos.

Quisiera aquí recordar al profesor. Oscar Varsavsky de la Universidad de Buenos aires que tanta influencia tuvo en mi decisión de dedicarme a la Estadística.

Uno de los temas de investigación que más me ocuparon en mi carrera académica fueron los Métodos Estadísticos Robustos. En Estadística, los métodos clásicos se basan en suposiciones muy fuertes sobre la distribución de los datos, por ejemplo es muy común suponer que los datos son normales. Los procedimientos que se obtienen bajo estas hipótesis son óptimos cuando estas suposiciones se cumplen exactamente, pero son altamente sensibles y muy ineficientes frente a cualquier desviación de las mismas. Una pequeña desviación en la distribución de los datos puede alterar y corromper en alto grado las conclusiones obtenidas con estos procedimientos. En particular es muy frecuente que en los datos analizados haya unas pocas observaciones atípicas que llamaremos atípicos. Estos atípicos pueden ser debidos a distintas razones, como por ejemplo:

a errores en los instrumentos de medición,

a la variación las condiciones bajo las cuales se obtuvieron los datos (por ejemplo en fenómenos físicos pueden haber

variado la presión o temperatura),
en datos económicos a algún disturbio institucional como puede ser una guerra o catástrofe natural como un terremoto.

Cuando se usan los métodos clásicos, unos pocos atípicos, inclusive uno solo, pueden tener un efecto muy importante en los resultados del análisis estadístico y invalidar las conclusiones.

La primera defensa que se utilizó contra este peligro fue el desarrollo de procedimientos para detectar estos atípicos y suprimirlos del análisis. Una vez suprimidos, al resto de los datos se aplicaban los procedimientos clásicos. Las técnicas de detección de atípicos son en general exitosas cuando estos aparecen aislados. Pero pueden fracasar cuando los atípicos aparecen en grupos.

El enfoque de la estadística robusta es distinto, ya que reemplaza la detección de atípicos por el desarrollo de procedimientos que son muy poco afectados por la presencia de estos.

Para mostrar la diferencia entre el enfoque de detección de atípicos con el de la estadística robusta haremos un paralelismo con el problema de elegir una estrategia de lucha contra los hackers. En este caso hay dos estrategias posibles, tratar de detectar la presencia de hackers en la red y encarcelarlos de manera que no puedan actuar. Este sería un enfoque similar a la detección de atípicos. El segundo enfoque sería construir sistemas operativos tan seguros que ningún hacker pudiera penetrarlos. En ese caso no sería necesario detectar a los hackers. Uno podría dejarlos actuar libremente sabiendo que no representan un peligro porque sus intentos de penetrar la red fracasarán. Este sería un enfoque similar al de los procedimientos robustos cuando eliminan el efecto de los atípicos.

Como fundadores de la estadística robusta podemos mencionar tres grandes nombres John Tukey, Peter Huber y Frank Hampel. John Tukey (1962) fue el primero que puntualizó claramente los peligros de atenerse a un modelo exacto y señaló que muchos de los procedimientos que eran óptimos para datos normales eran muy inestables frente a la presencia de atípicos. Por esta razón propuso procedimientos alternativos como por ejemplo el uso de la mediana y el MAD en reemplazo de la media y la desviación típica. Huber (1964) construyó una sólida teoría matemática de la robustez y obtuvo los primeros procedimientos minimax óptimos. Hampel (1968) desarrolló dos conceptos que son básicos para cuantificar la robustez de los procedimientos estadísticos: la curva de influencia y el punto de ruptura.

He trabajado en el área de la estadística robusta en los últimos 35 años. Este número me asusta, porque representa más de la mitad de mi vida. En estos años tuve oportunidad de abordar diferentes problemas de esta área. Me referiré brevemente y solamente a uno de los temas que más me ocuparon: Cómo conciliar robustez y eficiencia cuando se ajusta un modelo de regresión.

Como dije anteriormente muchos de los procedimientos clásicos estadísticos están basados en la suposición de normalidad. Este es el caso del modelo de regresión, donde el procedimiento clásico de mínimos cuadrados corresponde al óptimo cuando los errores tienen distribución normal. El primer procedimiento altamente robusto para este modelo de regresión fue propuesto primero por Hampel y luego desarrollado incluyendo su implementación computacional por Rousseeuw en (1984). Este procedimiento que se denomina mediana de cuadrados es altamente robusto pero muy ineficiente cuando los datos son normales, es decir si los datos son efectivamente normales se requieren muchas más observaciones que en el caso de mínimos cuadrados para obtener la misma precisión. Una de mis contribuciones que considero más importantes fue desarrollar estimadores que simultáneamente fueran altamente robustos cuando las observaciones tienen atípicos y altamente eficientes cuando los errores fuesen normales sin atípicos, por ejemplo que tuvieran una eficiencia del 95%. El significado del 95 % de eficiencia es que para obtener la misma precisión que con el método de mínimos cuadrados bastaría tomar un 5% adicional de observaciones. Volviendo a las analogías este costo adicional del 5% puede interpretarse como una pequeña prima de seguro que se debe pagar para protegerse de la catástrofe que pueden provocar en el análisis unos pocos datos atípicos. En el año 1984 desarrollé los estimadores llamados del tipo MM y en 1987 en colaboración con el Dr. Zamar los estimadores del tipo tau. Como estas dos clases de estimadores eran de gran complejidad computacional y solo se podían aplicar para un número pequeño de variables explicativas, en un trabajo conjunto con el Dr. Peña del año 1985 desarrollamos una versión rápida de los tau-estimadores que permite ser aplicada a grandes conjuntos de datos. Los últimos trabajos en esta dirección consistieron en tratar de suprimir la prima del seguro o reducirla drásticamente, es decir tener un seguro casi gratis. Esto lo obtuvimos en el año 2002 donde conjuntamente con dos alumnos de doctorado, Daniel Gervini y Sonia Hernández, obtuvimos dos procedimientos para regresión altamente robustos y que tienen eficiencia del 100%.

Quiero agradecer a mis discípulos, muchos de ellos luego amigos, colegas y colaboradores cuyo interés y dedicación logró crear un excelente clima de trabajo y de mutuo estímulo que permitió desarrollar con éxito nuestras investigaciones. Quisiera mencionar a Ricardo Maronna, Rubén Zamar, Ricardo Fraiman, Graciela Boente, Nelly Ferretti, lamentablemente fallecida, Marta García Ben, Ana Bianco, Diana Kelmansky, Jorge Adrover, Matías Salibian Barrera, Elena Martínez, Nora Muler, Julia Villar y Sonia Hernández.

También quiero agradecer a mis distinguidos colegas Daniel Peña, Doug Martín, y Alfio Marazzi con los que compartí largas horas de fructífero trabajo y amistad.

Mi profundo agradecimiento al Dr. Peña y a los otros miembros del Dto. de Estadística que organizaron en estos dos últimos días un seminario en la Universidad Carlos III con motivo de mi investidura como Doctor Honoris Causa. Fue realmente muy emocionante encontrarme y compartir estos momentos tan importantes para mí con colegas con los que me ligan afectos e intereses científicos comunes. Mi profundo agradecimiento a todos los que participaron.

Finalmente debo agradecer muy especialmente a mi esposa Livia y a mi hija Paula por su continuo apoyo afectivo, y por haber demostrado una infinita comprensión y paciencia cuando en lugar de dedicarme a compartir con ellas momentos de esparcimiento, me ocupaba obsesivamente a resolver algún problema estadístico.

Víctor Yohai