



Nombrado Doctor Honoris Causa en el acto de apertura del curso 98/99

Excelentísimo Rector Magnífico de la Universidad, autoridades, profesoras y profesores,

En primer lugar, permítanme advertirles de lo poco conveniente que resulta ser estadístico si se desea alcanzar reconocimiento público y premios. De manera increíble, una reciente y muy divulgada lista de los cien científicos más importantes de este siglo no incluía ni a Fisher ni a Neyman, dos de nuestros predecesores cuyo trabajo ha sido extraordinariamente influyente en el mundo de las aplicaciones científicas. Por tanto, esta ceremonia de hoy no es en absoluto algo común para mí. Personalmente, me siento muy honrado y sólo puedo responder con mi más sincera gratitud a esta Universidad y a su Rector, al Departamento de Estadística y Econometría, y al Profesor Juan Romo que ha sido el artífice de esta visita.

Aun a riesgo de parecer presuntuoso, también considero que ésta es una hermosa ocasión para la Estadística. Hoy, cuando nos acercamos al final del siglo XX, querría decir algunas palabras sobre cómo ha sido para la Estadística y quizás incluso especular un poco sobre su futuro.

De hecho, éste ha sido un siglo extraordinario para la Estadística, en el que nos hemos convertido en la metodología elegida por docenas de disciplinas que van desde la Educación a la Psicología o a la investigación biomédica. Y, por supuesto, también a la Economía, aunque en este caso la elección se remonta al siglo XIX. Más recientemente, el pensamiento estadístico ha comenzado a desempeñar un papel más importante en las ciencias "duras", sobre todo en la Astronomía y en la Física de partículas.

El gráfico del triángulo que pueden ver es un intento de trazar el desarrollo de la Estadística durante el siglo XX a través de un viaje imaginario en un pequeño automóvil alrededor del triángulo. Los tres vértices, "Aplicaciones" en el de arriba, "Matemáticas" en el inferior izquierdo, y "Computación" en el inferior derecho, representan los tres polos principales del pensamiento estadístico. El viaje comienza en el vértice superior del triángulo en 1899, muy cerca del polo de las "Aplicaciones". En ese punto, la Estadística era más una colección de técnicas ingeniosas pero fragmentadas que una disciplina organizada de pensamiento. Existían algunas ideas importantes desarrolladas por figuras heroicas como Gauss, Laplace, Galton y Edgeworth, pero no había un núcleo central del área. Sí se lee la revista de la American Statistical Association o de la Royal Statistical Society de 1899, aparecen numerosos trabajos sobre índices económicos y resultados del censo, pero no muchos sobre pensamiento estadístico en general. En ese momento, se habría podido dudar de que alguna vez existiera una "teoría general de la estadística", que no fuera más que una teoría sobre cómo llevar a cabo experimentos de laboratorio.

Entre 1899 y 1900 nuestro pequeño automóvil hace de repente un notable giro en dirección al vértice de las Matemáticas. Este movimiento representa la publicación del artículo de Karl Pearson sobre la distribución ji-cuadrado. La sofisticación matemática del trabajo sobre la ji-cuadrado, junto con su utilidad metodológica, puso en marcha un tren de importantes desarrollos en los fundamentos matemáticos del pensamiento estadístico. Biometrika, la primera revista moderna de estadística, empieza a publicarse en 1902. El contraste t de Student le sigue pronto en 1908. El pequeño coche de la Estadística se mueve ahora sin cesar en dirección al vértice de las Matemáticas.

El siguiente movimiento es el crucial. El artículo de Fisher de 1925 estableció la teoría moderna de la estimación estadística: consistencia, suficiencia, verosimilitud, información de Fisher, eficiencia, y optimalidad del estimador de máxima verosimilitud se presentan en la forma en que hoy se utilizan. La gran hazaña de Fisher fue conseguir un patrón de optimalidad para la estimación estadística- una medida de cómo es el mejor procedimiento que se puede conseguir en cualquier problema dado de estimación. Los resultados de optimalidad constituyen un síntoma de madurez científica, y señalo 1925 como el año en que la Estadística pasó de ser un conjunto de técnicas ingeniosas a convertirse en una disciplina coherente.

El gran logro de Fisher desencadenó un excepcional interés por los resultados de optimalidad. En el gráfico, nuestro automóvil estadístico sigue moviéndose en dirección al vértice matemático. Hay muchos puntos importantes en el recorrido pero he marcado sólo dos de ellos: 1932 por los intervalos de confianza y el lema de Neyman-Pearson, y 1950 por el desarrollo de Wald de la teoría de la decisión.

Mi propia educación estadística, en Stanford a principios de los sesenta, se desarrolló dentro de una incómoda cercanía al vértice matemático. La teoría de la decisión es una maravillosa herramienta para explicar estadística a estadísticos profesionales, pero tiene una utilidad limitada para las aplicaciones. En algún momento de los años sesenta pareció que la Estadística giraba sobre sí misma, y estuvo en peligro de perder su base en el mundo científico. Pudimos haber seguido la ruta de la música "seria" moderna, que ha cedido la mayor parte de su audiencia en favor de competidoras menos serias pero más agradables.

Sin embargo, justo entonces, cuando parecía que la matematización de la Estadística podía dejarla fuera de circulación, nuestro viaje experimentó un brusco giro a la izquierda hacia el vértice computacional del triángulo. Este movimiento parece suficientemente obvio si se mira de forma retrospectiva: los modernos ordenadores nos proporcionan al menos un millón de veces más de capacidad de cómputo que la que tenían los estadísticos anteriores a la Segunda Guerra Mundial. La computación factible ha sido siempre el condicionante que ha limitado la teoría estadística genuinamente útil. Las teorías de Fisher fueron inteligentemente adaptadas a la potencia de las nuevas calculadoras mecánicas y, de hecho, la teoría de máxima verosimilitud fue criticada al principio por ser demasiado difícil desde el punto de vista computacional en comparación con el método de los momentos de Pearson. Como joven profesor titular visitante en Harvard en 1967, ví a Bill Cochran desarrollar una Tabla de Análisis de la Varianza 4 x 4 con interacciones en una calculadora mecánica Monroe, en la que se utilizaban todos sus recursos, incluyendo los misteriosos números rojos, pero semejante proeza estaba mucho más allá de las posibilidades de cualquiera de nosotros. La teoría estadística había perdido fuerza porque los límites de la computación mecánica restringían nuestro pensamiento a un repertorio limitado de técnicas útiles.

La respuesta de nuestro campo a las computadoras electrónicas fue más bien perezosa al principio. Es duro recordar lo difícil que resultaba programar algo, o conseguir que un programa funcionara, en aquellos días en que había que llevar cajas llenas de tarjetas al vestíbulo del centro de cálculo, hacer cola y esperar a un resultado favorable a la mañana siguiente. En estas circunstancias, los estadísticos utilizaban al principio los ordenadores sólo para desarrollar versiones mayores de procedimientos familiares: regresiones lineales con más predictores, tablas de contingencia con más entradas, etc. Entonces, en 1964, Tukey publicó un artículo visionario titulado "Sobre el futuro del análisis de datos". El nombre "análisis de datos", en vez de "estadística", me parece que se eligió para deshacerse del asidero más bien mohoso de la teoría estadística de los sesenta. El trabajo de Tukey apuntaba a una utilización mucho más interesante de la computación para forjar herramientas genuinamente nuevas y útiles para el análisis estadístico. El jackknife, un misterioso método nuevo para calcular errores estándar incluso para estimadores muy complicados, se invocaba como paradigma de desarrollos posteriores.

Los años entre 1900 y 1950 constituyen la "era de la teoría", en la que un grupo de genios, Student, Fisher, Neyman, los Pearson, Hotelling, Cramer y Rao, establecieron los fundamentos de la estadística moderna. Las ideas de preguerra no fueron descartadas- por el contrario, se amplificaron de forma asombrosa y en direcciones sorprendentes. En mi trabajo cotidiano de consulta en el Centro Médico de Stanford, realizo rutinariamente análisis que hubieran parecido imposibles hace treinta años: modelización con tasas de fallo proporcionales, muestreo de Gibbs, simulación de Monte Carlo mediante cadenas de Markov, modelos lineales generalizados, técnicas basadas en datos faltantes y, por supuesto, jackknife y bootstrap.

En mi gráfico, la Estadística en 1998 se ha movido hacia el centro del dibujo. Felizmente, la revolución computacional nos ha llevado más cerca del vértice de las aplicaciones. Nuestra amplia base metodológica nos permite ser de mucha más ayuda a nuestros colegas científicos. Hoy existe una auténtica vitalidad dentro de la Estadística que faltaba en los años sesenta, una vitalidad que procede de una interacción mucho más fuerte con numerosas áreas de aplicación. Pienso que a los científicos ahora les gusta más tratar con estadísticos porque no les forzamos a simplificar demasiado los problemas que nos presentan.

En el viaje de la Estadística he indicado el bootstrap mediante una pequeña estrella situada más o menos en 1979. Esta fue mi contribución principal a la "era metodológica". El bootstrap es un descendiente directo del jackknife de Tukey y Quenouille. Cuando yo era estudiante de doctorado en Stanford, Rupert Miller dió un importante seminario titulado "Un jackknife digno de confianza". Rupert estaba intentando desvelar parte del misterio de la habilidad del jackknife para convertir cálculos teóricos imposibles en otros computacionales sencillos. Sencillos para una computadora- el jackknife multiplica cualquier cálculo estadístico tradicional por lo que es un factor enorme para los estándares de preguerra.

Casualmente, en 1972 Rupert Miller y yo pasamos un año sabático en el Imperial College de Londres. Después de uno de los brillantes seminarios de Rupert, David Cox me preguntó si me parecía que la idea del jackknife contenía algo de interés. Tardé algún tiempo en darme cuenta de que esto era una indicación, pero entonces empecé a pensar en las bases del jackknife. Yo tenía una idea muy complicada para explicarlo que llamaba "la distribución combinacional". Sin embargo, cuanto más trabajaba en ello, más se simplificaba hasta que me quedé con algo que no parecía casi nada. Y eso era el bootstrap.

A dónde irá la Estadística? Las flechas con interrogantes que emanan de "1998" en mi triángulo dejan claro que no sé la respuesta. Una buena conjetura es que la revolución metodológica está todavía tomando velocidad, y convertirá a la Estadística en algo aún más útil y más empleado en las aplicaciones. El pensamiento estadístico entra en juego siempre que uno tiene que destilar información que llega en porciones muy pequeñas, ninguna de las cuales es suficiente para resolver la cuestión. Ahora, más áreas de investigación se hallan en esta situación: la Física de partículas es un buen ejemplo reciente de esto.

Por tanto, espero que el pensamiento estadístico juegue un papel incluso mayor en las ciencias y en las ciencias

sociales. Esto no garantiza por sí mismo que la Estadística vaya a prosperar. "El éxito siempre hace que parezca fácil el éxito", dijo un famoso militar americano, y el siglo XX ha sido ciertamente un tiempo de gran éxito para mi disciplina, que ha generado un torrente de nuevas ideas revolucionarias tanto para la teoría como para sus aplicaciones. Necesitamos ideas de la misma magnitud para tener éxito en el siglo XXI.

Creo que importantes y nuevas ideas estadísticas se están desarrollando en la actualidad, y que la Estadística en el siglo XXI crecerá para ser reconocida como uno de los pilares centrales del pensamiento científico, alineándose junto a la tradicional lógica matemática y a la moderna computación. Mi confianza en el futuro se basa principalmente en la cantidad y calidad de los nuevos investigadores que veo entrar en nuestra disciplina. Por primera vez. Asia ha comenzado a contribuir de forma importante. Y, lo que hoy particularmente me llama la atención es que Europa está volviendo a la importante posición que ocupaba en el mundo del pensamiento estadístico.

La historia de la estadística académica en los Estados Unidos es la de un lento pero firme crecimiento de los departamentos de estadística, con las poderosas universidades tradicionales como Harvard o Yale sucumbiendo ante escuelas más recientes como Stanford y Iowa y North Carolina. Quizá las nuevas disciplinas necesitan nuevas universidades. Hoy tengo un motivo especial para estar agradecido a esta Universidad y a su Departamento de Estadística y Econometría. Les deseo continua prosperidad y éxito intelectual, y gracias de nuevo por el honor de estar hoy aquí.

