

EL VOCABULARIO ESTADISTICO

La conveniencia de utilizar un vocabulario uniforme se ha dejado sentir en la Estadística como en otras muchas técnicas. Pero esta conveniencia se convierte en necesidad en el caso de la Estadística, pues su desarrollo ha avanzado en los últimos lustros con una velocidad pocas veces igualada en la historia del saber, y este desarrollo se está llevando a cabo primordialmente en un solo idioma: el inglés.

El Instituto Interamericano de Estadística (I. A. S. I.) —entidad radicada en Washington, en estrecha relación con la Pan American Union, para fomentar, con fines de aplicación práctica, los estudios estadísticos en América— abordó este problema del vocabulario estadístico en su II Congreso Interamericano de Estadística, celebrado en Bogotá en enero de 1950. La labor previamente realizada por el I. A. S. I. y la que resultó de dicho Congreso se han condensado en la publicación de un *Vocabulario Estadístico*¹, que, a más de unas «Notas explicativas», redactadas en los cuatro idiomas, contiene una «Lista principal», formada por unos 1.800 términos ingleses y sus equivalentes en español, portugués y francés, y un «Índice alfabético» en cada una de estas tres últimas lenguas, en el que cada término lleva su correspondiente traducción inglesa.

Para nosotros no deja de ser lamentable que en el Congreso de Bogotá se haya afirmado (Resolución núm. 12) que, «en lo que se refiere a la traducción francesa, se estima conveniente advertir que se presentan los términos propuestos solamente sobre bases preliminares», y no haya creído, en cambio, necesario el Congreso tener en cuenta la opinión de españoles ni portugueses para establecer los vocabularios de los correspondientes idiomas.

Se ha de reconocer, sin embargo, que en la elección de vocablos se ha procedido, salvo algunas excepciones, con sumo acierto, y que el estilo en que está escrita la traducción castellana de la Introducción es bastante correcto y, desde luego, muy superior a lo que por desgracia se encuentra con tanta frecuencia en la producción científica publicada allende los mares.

¹ *Statistical Vocabulary-Vocabulario Estadístico-Vocabulário Estatístico-Vocabulaire Statistique*, First edition; Inter American Statistical Institute, Washington, junio 1950; XVIII + 117 págs.

Señala el II Congreso Interamericano como normas para fijar las traducciones:

a) Reducir a una sola palabra, o al menor número de ellas, cada término propuesto;

b) Conservar, cada vez que fué posible, la unidad etimológica en las lenguas española, portuguesa y francesa;

c) Respetar los términos ya consagrados por el uso en cuanto ello era compatible con los puntos a) y b), arriba indicados.»

El vocabulario abarca la teoría estadística, los términos de muchas de sus aplicaciones e incluso de ciencias conexas, sobre todo de la Economía.

Desgraciadamente deja sin resolver bastantes problemas difíciles de terminología, por ejemplo: *batch*, *confounding*, *closeness*, *probit*. Resuelve otros de manera muy deficiente, entre ellos: *bias*, conservado en castellano como neologismo; *tetrad*, conservado asimismo en vez de admitir, como es lógico, «tétrada»; *average*, traducido correctamente, en general, por «promedio», pero en algunas locuciones por «media»; *run*, que se vierte por «repetición», y finalmente *standard*, traducido innecesariamente por el neologismo «estándar», cuando existe la palabra «típico», que ajusta perfectamente en todas las locuciones en que aquélla entra. (Hemos de reconocer, sin embargo, que tampoco en España hay acuerdo sobre este término, pues se ha dado ya carácter oficial a la traducción de *standard* por «norma», en otra rama de la técnica.) Señalamos asimismo como demasiado numerosas las equivalencias dadas a *random* («accidental», «al azar», «aleatorio», «casual»), en contra del requisito establecido en a). Si hubiese empleado «varianza» y «covarianza», en lugar de «variancia» y «covariancia», se tendría la doble ventaja de una mayor eufonía y cumplirse mejor la condición c).

Finalmente, a nuestro juicio, hay traducciones completamente inaceptables; así, *acceptance sampling* por «muestreo de ensayos», que, sin embargo, ha sido vertido al portugués por «amostragem para aceitação»; ¿por qué no se ha considerado aquí la unidad etimológica y se ha dicho «muestreo para aceptación», como sería correcto?

En España el tema del vocabulario estadístico ha preocupado singularmente al señor Cansado, quien a la vista de una edición provisional del vocabulario del Instituto Interamericano, publicada en 1948, se ha ocupado en repetidas ocasiones del tema, hasta llegar a la publicación de un vocabulario² de unas 1.000 palabras, incluidos derivados y locuciones, con la atención centrada en la estadística teórica.

En los dos trabajos que comentamos se abusa de la inclusión de

² ENRIQUE CANSADO. «Vocabulario estadístico inglés-español». *Trabajos de Estadística*, vol. I, cuaderno II; Madrid, 1950.

locuciones innecesarias por haberse traducido ya todos sus términos, como *coefficient of correlation*, ociosa por figurar también las traducciones de *coefficient* y *correlation*. Aunque sea evidente que lograr un vocabulario completo es tarea casi imposible, sin embargo, en estas dos publicaciones faltan muchos términos, de los cuales —y limitándonos al campo común a ambas— nos permitimos señalar algunos a continuación:

Abruptness, available information.

Centroid, chunk, communality, complexity, confluence analysis, cross-ever (design).

Dummy treatments, dummy variable.

Factorization, functor.

Half-plaid square, heterotyposis.

Irrelevant information, isostatistical, iterative estimation.

Latent roots, loadings.

Missing data.

Path-regression.

Quasi-latins design.

Recovery of information, relevant information, reversal.

Single component, specification, split-plot design, stretching-scores, superefficient, switchback trial.

Ultimate distribution.

Traits.

El trabajo del señor Cansado es muy meritorio, y desde el punto de vista puramente estadístico no cabría oponerle graves objeciones. No obstante, un vocabulario es ante todo idioma, y bajo este aspecto nuestra discrepancia con él es profunda.

A nuestro juicio presenta como error fundamental el abuso en la utilización de neologismos y arcaísmos, que llegan a transformar el lenguaje estadístico en algo completamente ininteligible para el no iniciado en tal vocabulario. Así, nos ofrece palabras como «abarcamiento» para *coverage*, «diferecer» para *difference*, «allegamiento» para *gathering*, «sobrevisión» para *survey*, «tarja» para *tally*, «dócima» para *test* y tantas otras completamente incomprensibles, no ya para el lector corriente, sino para los propios especialistas.

Hay que considerar que en otros idiomas, y singularmente en inglés, «el hombre de la calle» puede leer un párrafo de un trabajo estadístico y saber al menos de qué se trata, aunque no conozca el exacto sentido técnico de los términos empleados. Con el vocabulario del señor Cansado alguien se podría dirigir a una autoridad administrativa, especializada o no en estadística, con un párrafo como el siguiente:

«*Reporte.*—Se ha realizado esta sobrevisión con un abarcamiento excesivo. La muestra se ha afijado con inadecuación y el allegamiento de datos peca de incompletitud. Las dócimas indican gran inacuracidad.»

En lenguaje más vulgar esto quiere decir:

«Informe.—El estudio cubre un campo demasiado amplio. La muestra no se ha repartido debidamente y los datos obtenidos son muy incompletos. Las pruebas estadísticas indican gran imprecisión.»

Y no creemos que pueda alegarse que por necesitar las palabras en la técnica un sentido estricto y determinado deban desecharse todos los términos del lenguaje vulgar para evitar torcidas interpretaciones, pues ¿qué sería entonces del vocabulario de todas nuestras ciencias? ¿Cómo conservar denominaciones tales como «producto», en el Análisis; punto, recta y plano, en Geometría; fuerza y masa, en Física, etc., etc., cuyo sentido es tan distinto al de la vida ordinaria?

La antedicha oscuridad queda agravada por el afán de traducir palabra por palabra todas las locuciones y de conservar una sola raíz para todos los derivados de una raíz inglesa, lo que lleva a decir «estrictéz» en vez de «lo estricto», «determinántico» y no «de los determinantes», «diferecer» en lugar de «calcular diferencias» y «aleatorización» en vez de «sorteo».

El señor Cansado afirma en su introducción (pág. 222): «Debo advertir que las traducciones que figuran en este *Vocabulario* son el resultado de una labor ejecutada con escrupuloso cuidado, después de estudiar y comparar numerosos diccionarios españoles e ingleses, así como gran número de libros y artículos de estadística escritos en lengua inglesa.» Estamos completamente seguros de que esto es cierto, ya que no ignoramos los amplios conocimientos y la infatigable y honrada forma de trabajar del señor Cansado. Nos extraña por eso que no haya examinado la escasa literatura estadística española o no haya querido conservar en muchos casos palabras ya consagradas por el uso, tales como «equilibrio» para *balance*, «conjunción» para *confounding*, «celosía» para *lattice*, etcétera, no mejoradas al decir «contrabalance», «confusión» y «retículo», respectivamente.

Queremos también señalar la necesidad que a nuestro juicio existe de hacer intervenir en la formación del vocabulario a los técnicos del idioma, para evitar al menos los errores gramaticales que con tanta facilidad cometemos los no especialistas, tales como supresión de artículos («análisis de varianza» por «análisis de la varianza»), trasposición indebida de adjetivo y sustantivo («mínimos cuadrados» por «cuadrados mínimos», censurado ya hace largo tiempo por Rey Pastor; «grandes números» por «números grandes», «máxima verosimilitud» por «verosimilitud máxima»), sufijos contrarios al uso («binomial» por «binómico»), modificación del caso de la declinación (*operating characteristics curve*, traducido por «curva característica operante» en lugar de «curva de las características operantes») y utilización de barbarismos (*distribution-free methods*, traducido por «métodos a libre distribución» en lugar de «métodos independientes de la distribución»).

Finalmente anotamos, entre otros muchos aciertos de la obra del señor Cansado, algunos términos elegidos por él que mejoran las traducciones del Instituto Interamericano, cuales son *best estimate* = «estimación óptima» (y no «la mejor e.»), *cross-classification* = «clasificación cruzada» (y no «combinada»), «decila» (y no «decil»), «escedasticidad» (y no «sedasticidad»), «curtosis» (y no «kurtosis»), «muestreo polifásico» (y no «por fases múltiples»), «correlación serial» (y no «c. en serie»), «estadígrafo» (neologismo) para *statistics*, entre otras muchas.

No son a nuestro juicio despreciables estas cuestiones terminológicas. Dejando a un lado que a ellas puede ir en parte unido el progreso de la estadística, hay que tener en cuenta que ésta es una técnica que ha entrado muy directamente al servicio de la sociedad, y es menester que esa técnica gane entre nosotros la atención y el interés de todos para aplicarla a los problemas de la vida nacional.

ANGEL ANÓS