

CRITERIS ESTADÍSTICS EN EL RECALL I LA INTERPRETACIÓ DE DADES SOCIOLINGÜÍSTIQUES

*Sessió a càrrec de J. OLIVERAS, E. BONET
i J. CAPELLADES (20 de març de 1991)**

1. CRITERIS GENERALS

Quan tenim unes dades estadístiques a les mans com les de coneixement del català o les de l'ús de diverses llengües, ens plantejem, conscientment o inconscientment, dues gestions fonamentals. La primera és buscar les conseqüències que en podem treure més enllà de la lectura directa; la segona és saber si les interpretacions que en fem són correctes o no. Per ambdós temes, diguem-ne el d'indagació i el de rigor, la teoria i els mètodes estadístics ens donen respostes i ens obren investigacions.

Destaquem d'entrada que, atesos uns resultats, les conseqüències que en podem treure i les interpretacions que en podem fer depenen de la classe de dades i del tipus de recull que s'ha emprat. Així la producció i la interpretació de dades es relacionen íntimament. Aquesta propietat exigeix d'una banda que els organismes i els responsables de la producció estadística hagin de tenir una cura molt especial en les operacions de disseny i en les de camp; i d'altra banda que els organismes i els responsables de les anàlisis i dels estudis posteriors hagin de conèixer molt bé el tipus i les fonts de les dades. Sempre és oportú de recordar aquesta darrera recomanació, i en aquesta sessió pot ser interessant d'exposar breument algunes distincions bàsiques sobre resultats estadístics.

1.1. Distincions sobre els qüestionaris

Neutralitat. La preparació dels qüestionaris és una operació estadística molt delicada, ja que aquests han de ser neutres; és a dir, no han d'introduir apreciacions de valors, no han d'incidir, orientar o modificar l'opinió de la gent i no han d'afavorir unes respostes enfront de les alternatives. Respecte a la neutralitat, a l'altre extrem del ventall, hi trobem aquelles campanyes publicitàries que malauradament prenen les enquestes com a pretext per a acostar-se a la gent, com a disfressa manyaga per a persuadir els consumidors i fins i tot com a sistema de vendes o de proselitisme.

Proves objectives. Uns qüestionaris es basen en proves objectives i d'altres en criteris personals, i per tant subjectius, dels enquestats. Les preguntes de coneixement del català incloses en el cens corresponen a aquest darrer cas. Però no hem de confondre la distinció entre l'objectivitat i la subjectivitat entesa com a criteri personal amb la que hi ha entre claredat i la vaguetat de les preguntes.

Definició i vaguetat. El fet de respondre sí o no a la qüestió d'entendre el català correspon al criteri de cada ciutadà, però el nivell de comprensió que es requereix i que defineix la pregunta és precís i ben explicat.

Coneixement i ús. Les preguntes del cens lingüístic són de coneixement i no d'ús de la llengua catalana.

* Jordi Oliveras és director de l'Institut d'Estadística de Catalunya; Eduard Bonet és president del Consell Català d'Estadística; i Joaquim Capellades és subdirector d'Estadístiques Demogràfiques i Socials (Institut d'Estadística de Catalunya).

Homogeneïtat. La disciplina de l'operació de camp, amb les explicacions, les normes, la formació dels agents i les inspeccions, té cura, entre d'altres aspectes, que tot-hom entengui les preguntes de la mateixa manera. Així, en els censos posteriors de coneixement del català, va servir d'experiència el fet que, en el cens de 1975, gent de la plana de Vic, que parlaven sempre en català, van contestar que no el sabien i gent del cinturó industrial de Barcelona, que quasi no l'entenien gens, van contestar que sí.

Forma de passar els qüestionaris. La qualitat dels resultats estadístics es relaciona fortament amb la manera com es passen els qüestionaris. Així, no és correcte fer la pregunta de coneixement de català en un grup o en una escola i demanar que es contesti col·lectivament alçant la mà.

1.2. Distincions sobre l'exhaustivitat

En tots els treballs estadístics d'entrada s'ha de definir la població que es vol estudiar i després s'ha de determinar si es farà un estudi exhaustiu o per mostres.

En els estudis exhaustius les normes, la disciplina i el control per a assegurar aquesta propietat són molt importants. Moltes vegades queden fora de l'enquesta parts substantives de la població, que formen grups amb qualitats específiques com, per exemple, la de no tenir telèfon. En el límit, de vegades s'ha patit la il·lusió o el miratge de l'exhaustivitat quan només s'ha preguntat als amics i als coneguts.

En els estudis exhaustius les dades poden ser desagregades en diferents grups com per exemple, segons l'edat, el sexe, la comarca o el lloc de naixement. En tots aquests grups òbviament es continua tenint dades exhaustives, i la comparació entre elles sempre té sentit. L'única limitació és de no fer grups massa petits, ja que d'una banda poden obrir conflictes amb el secret estadístic i d'altra banda presenten uns percentatges sense gaire interès.

1.3. Distincions sobre les mostres

En els estudis per mostres a les observacions anteriors, s'hi afegeixen la qüestió conceptualment i pràcticament difícil de tenir mostres representatives i no i ha d'establir la relació o aproximació que hi ha entre els percentatges o valors que calculem en les mostres i per tant coneixem i els corresponents en la població, que són els que ens interessen, però que desconeixem.

Les mostres aleatòries o a l'atzar s'escullen per un sistema equivalent a un sorteig de manera que tots els elements de la població tenen la mateixa probabilitat de formar part de la mostra.

En aquest cas les diferències que hi ha entre els percentatges o valors de les mostres i els corresponents de la població poden estudiar-se matemàticament a través de lleis de probabilitat. D'aquest càlcul, en destacarem els conceptes i les propietats següents:

— No podem assegurar que els percentatges de la població i la mostra coincideixin. Si per casualitat és així, tampoc no ens n'adonem.

— Les diferències entre els percentatges de la mostra i els de la població tant poden ser positius com negatius, de manera que les possibilitats s'equilibren, segons un concepte que l'estadística precisa una mica més i que dóna lloc a la noció de mostres sense biaix enfront de les esbiaixades.

— Les diferències entre els percentatges de la població i de la mostra tendeixen a disminuir i acostar-se a zero en augmentar la mida de la mostra, és a dir, el nombre d'elements que la formen.

— Si es fixa un grau d'aproximació o interval de confiança per la diferència entre el percentatge de la població i el de la mostra, es pot calcular el nombre d'elements de la mostra que són necessaris per a assolir aquest grau. Tanmateix cal destacar que, d'aquesta aproximació, no en tenim una certesa completa, sinó que hem d'admetre un risc o probabilitat de fallar. Usualment el risc és del cinc per cent, i es diu que es té un nivell de confiança del noranta-cinc per cent.

— L'estudi de la mostra per subgrups de població s'ha de fer amb molta precaució, ja que normalment les dades són insuficients per a admetre desagregacions.

— En el cas que els percentatges de dos grups de la població siguin iguals, els percentatges corresponents en les mostres poden presentar diferències d'una certa importància; però s'han de considerar no significatives.

Totes aquestes propietats matemàtiques aclareixen els conceptes de mostra representativa, d'estimació sense biaix i de diferències no significatives, els quals requereixen ser enunciat amb fórmules matemàtiques.

Les mostres per quotes. Es formen de manera que en uns grups prèviament controlats de la població com els de sexe, edat, comarca o situació professional, els percentatges de la mostra coincideixen amb els de la població. Aquesta condició tota sola no assegura la representativitat de les mostres per d'altres propietats com les de coneixement del català. La situació ideal és fer un sistema de quotes de manera que en cada grup els elements s'escullin a l'atzar.

Mostres sense regles. Moltes vegades les mostres es fan sense regles prèvies per a triar els elements. Aquests procediments són molt dubtosos i no tenen cap garantia científica encara que es prengui un gran nombre de persones. En un cas extrem s'acaba preguntant a molta gent, però sempre entre amics i coneguts, o entre grups de persones molt afins.

1.4. *Distincions sobre l'oficialitat i els censos*

Cal distingir entre les estadístiques oficials i les no oficials. Les primeres compleixen una sèrie de normes professionals i tècniques, són fetes sota la responsabilitat d'una institució estadística i prenen un valor oficial. Entre les estadístiques oficials hi ha tant els estudis exhaustius dels censos com els estudis per mostres.

Els censos. Malgrat que en la nostra pràctica habitual donem el nom de cens a moltes llistes, com d'artistes o de teatres, i tothom diu que fa un cens, des d'un punt de vista oficial els censos han de complir una sèrie de normes establertes per les Nacions Unides, la Comunitat Europea i els estats. En els censos, els ciutadans tenen l'obligació de facilitar la informació; l'administració ha de mantenir el secret estadístic i el respecte a la intimitat; el recompte ha de ser exhaustiu sobre un territori; l'operació ha de ser feta per un organisme estadístic oficial i els resultats han de ser aprovats i publicats oficialment per l'òrgan de govern que determina la llei.

Les mostres oficials. Han de ser dissenyades per un organisme estadístic oficial, han de ser aprovades per l'òrgan de govern que determina la llei i s'han de publicar oficialment amb totes les precisions metodològiques necessàries per a interpretar-les correctament.

1.5. *Precaucions sobre les fonts*

Per als estudis i les anàlisis que es fan d'unes dades estadístiques s'han de conèixer a fons tots els aspectes metodològics de la seva producció i s'han d'esmentar les seves fonts originals i distingir-les d'altres estudis consultats.

2. ALGUNS CRITERIS SOBRE EL CENS DE CONEIXEMENT DE LA LLENGUA CATALANA DE 1991

El Cens de Coneixement del Català es fa a través de les preguntes incloses en el Cens Nacional de Població de 1991. Així, aquesta vegada no s'ha seguit el padró municipal d'habitants, que ha quedat reduït a les dades bàsiques d'empadronament. Aquesta actuació es basa en la bona col·laboració i els bons acords entre l'Institut d'Estadística de Catalunya i l'Institut Nacional d'Estadística. L'Institut d'Estadística de Catalunya, en col·laboració amb la Direcció General de Política Lingüística, ha preparat les preguntes i la redacció dels fulls censals bilingües. L'INE porta a terme tota l'operació de camp i després la totalitat dels qüestionaris passarà a l'Institut d'Estadística de Catalunya per a la producció de l'arxiu informàtic i l'elaboració de les taules sobre el coneixement de la llengua catalana. Finalment els resultats del Cens Lingüístic seran aprovats pel Consell Executiu del Govern de la Generalitat i publicats oficialment en el DOG.

Per tant, el procés de preparació del Cens de Coneixement de la Llengua Catalana, de 1991, amb la realització de camp a través del Cens Nacional de la Població de l'INE i l'aprovació del Govern de la Generalitat, obeeix totes les normes i formalitats demanades pels organismes internacionals i té un valor oficial reconegut universalment.

Aquestes qualitats es donen també en el Cens de Coneixement de la Llengua Aranesa, que a través de qüestionaris trilingües es fa per primera vegada a la Vall d'Aran.

3. ALGUNES CONSIDERACIONS SOBRE L'INSTITUT D'ESTADÍSTICA DE CATALUNYA

Aquesta sessió de treball representa el primer contacte i la primera col·laboració de l'Institut d'Estadística de Catalunya, que va iniciar les seves activitats el 1990, amb el Grup Català de Sociolingüística.

El senyor Jordi Oliveras, director de l'Institut d'Estadística de Catalunya, remarca algunes línies de treball de l'Institut i destaca la voluntat d'aprofundir els temes i les orientacions següents:

— La col·laboració amb l'INE, per no duplicar esforços i costos i per facilitar a Catalunya moltes més dades estadístiques.

— La coordinació del sistema estadístic de Catalunya basada en principis metodològics i actuacions oficials.

— L'elaboració rigorosa del Cens de Coneixement de la Llengua Catalana.

— La col·laboració de l'Institut, com a organisme tècnicament independent i neutre, amb la Direcció General de la Política Lingüística, l'Institut de Sociolingüística de Catalunya i el Grup Català de Sociolingüística, sobretot en els aspectes de facilitar treballs i resultats, de donar un suport tècnic i de fomentar la formació estadística.