

MOSTRATGE I ESTIMACIÓ DE POBLACIONS OCULTES

JOAN DEL CASTILLO

Universitat de Barcelona

En aquest article es fa una discussió detallada de les principals metodologies emprades per l'estudi de fenòmens marginals i ocults, com són el mostratge en bola de neu, el mètode d'estimació de captura-recaptura i el mètode nominatiu. A la secció 5 proposem un nou mètode per estimar el nombre de consumidors de cocaïna en una gran ciutat. Aquest nou mètode es basa en dues seleccions aleatòries i pot estimar l'ocultació del fenòmen.

Sampling and Estimation of Elusive Populations.

Keywords: Mostratge en bola de neu, mètode captura-recaptura, mètode nominatiu.

1. INTRODUCCIÓ

La investigació sobre fenòmens socials ha tingut en el nostre país un creixement espectacular els últims anys. Aquest interès creixent ha portat a una gran quantitat d'estudis que utilitzen una àmplia varietat de tècniques tant qualitatives com quantitatives. En aquest article ens proposem, en primer lloc, fer

—Joan del Castillo. Dep. Matemàtiques. Univ. Autònoma de Barcelona. 08193 Bellaterra, Barcelona.

—Aquest treball ha estat finançat parcialment per l'Institut Catòlic d'estudis socials de Barcelona, ICESB.

—Article rebut l'abril de 1991.

una discussió detallada sobre algunes tècniques d'estudi dels fenòmens marginals i ocults com són el mostratge per xarxes, el mostratge en bola de neu, el mètode d'estimació de captura-recaptura i el mètode nominatiu. En la secció 5 proposem un nou mètode d'estudi, que recull algunes de les idees dels mètodes anteriors, encaminat a conèixer el nombre de consumidors de cocaïna en una ciutat. Aquest mètode es basa en dues seleccions aleatòries que permeten estimar l'ocultació del fenomen i verificar les estimacions que s'en fan. A la vegada, per aquest procediment es pot obtenir una mostra aleatòria dels consumidors de cocaïna que pot ser utilitzada posteriorment per un estudi de les seves tipologies.

Entenem per fenomen *marginal* aquella característica que poseeixen un nombre petit dels individus d'una població. Serien fenòmens marginals les malalties poc freqüents, els individus sense residència fixa, els alcohòlics, els drogaaddictes, etc. Alguns d'aquests fenòmens presenten a més la característica de ser *ocultes*, és a dir, són amagats pels propis individus que en formen part.

L'estudi de les poblacions marginals i ocultes, és a dir, l'estudi dels individus que presenten característiques d'aquest tipus, presenta uns problemes específics. En primer lloc hi ha el problema de la seva *localització*. Un procediment habitual de mostratge ens proporciona una quantitat molt petita d'aquests individus simplement per la seva baixa proporció. A més cal afegir-hi que aquestes poblacions marginals poden estar formades per individus sense telèfon o sense domicili fix, el que dificulta més que formin part de la mostra. Finalment, si també es tracta d'un fenomen amb el problema de l'ocultació resulta encara més difícil arribar a entrevistar els individus que són objecte de l'estudi.

El segon problema en que ens trobem és el *problema de la quantificació*. Ja hem dit que els procediments habituals de mostratge ens proporcionen una quantitat molt petita d'individus de poblacions marginals i ocultes. A aquest fet s'afegeix el problema de l'esbiaixament. Les característiques de marginalitat i ocultació fan que la proporció dels individus de l'estudi que podem localitzar dintre de la mostra sigui inferior a la seva proporció en la societat que hem mostrat. Per aquesta raó sovint cal recórrer a procediments indirectes d'estimació que comentarem més endavant.

2. MOSTRATGE PER XARXES I CADENES

En els sistemes socials, a més de trobar-nos amb individus que poseeixen determinades *característiques*, molt aviat ens trobem amb *relacions* que s'estableixen entre individus. Un estudi sobre una societat es pot enfocar desde un punt de vista individual, és a dir, preocupar-nos per conèixer quines característiques

concretes poseeixen els seus individus, o bé desde un punt de vista col·lectiu, preocupant-nos per saber quines relacions hi ha establertes entre diferents individus. Un exemple de característica o atribut seria el sexe d'un individu que és una propietat que depèn exclusivament d'ell. Exemples de relacions serien els vincles matrimonials, la relació pare-fill, la relació professor-alumne, l'amistat... que sempre necessiten de dues o més persones per afirmar que existeixen.

L'anàlisi de xarxes es proposa estudiar relacions concretes que poden donar-se o no entre els individus d'una població. Un estudi d'aquest tipus, sens dubte d'innegable interès, és en general molt més complex que un estudi basat en atributs o característiques. Cal tenir present que en una població formada per n individus en la que és vol estudiar simplement una relació binària aquesta pot aparèixer en $n(n + 1)/2$ casos diferents, el que dóna una idea de la complexitat a la que ens referíem.

Juntament amb l'estudi de l'existència de relacions en una població, l'anàlisi de xarxes permet mesurar la intensitat de les mateixes. Les dades recollides en el treball de camp es representen, en el cas de relacions binàries, en una matriu de n^2 elements i a partir d'allí es pot estimar, per exemple, la mitjana de relacions per individu. Per a més informació es pot consultar Knoke i Kublinski (1990), Erickson i Nosanchuk (1983), O. Franch (1971). No obstant, sense pretendre entrar en l'anàlisi de xarxes, la idea de xarxa com a conjunt d'individus relacionats amb un individu concret per relacions d'amistat, de parentesc o simplement per habitar un mateix domicili pot ajudar-nos en l'estudi de les poblacions marginals i ocultes que ens proposem.

Sudman, Sirken i Cowan (1988) descriuen el *mostratge per xarxes* o *mostratge amb multiplicitat* com un procediment que permet augmentar la quantitat d'informació obtinguda en una entrevista. La idea d'aquest tipus de mostratge és senzilla: els individus entrevistats, a més de preguntar-los per qüestions personals, se'ls pregunta sobre els seus parents, companys de feina, o membres del grup social al que pertanyen. Naturalment, per evitar un esbiaixament dels resultats, cal ponderar aquestes respostes en funció del nombre de persones que formen la xarxa de relacions de l'entrevistat. Més endavant, en la secció 5, recollirem aquesta idea.

Relacionat amb la idea de xarxa podem considerar el *mostratge per cadenes*. Una cadena es construeix a partir d'un individu al que se li pregunta per la persona o persones que tenen més relació amb ell; a continuació es localitzen aquestes persones i es repeteix el procediment. El mostratge per cadenes ens dóna informació sobre l'estructura de xarxa de la població però té un objectiu diferent de l'anàlisi de xarxes. Aquest darrer pretén una observació completa d'una població tancada i generalment petita. El mostratge per cadenes pretén, en canvi, localitzar una mostra d'individus amb una certa característica dintre

d'una població gran. Hi ha, encara, una diferència més subtil. L'anàlisi de xarxes pretèn mesurar la intensitat de les relacions, és a dir, la quantitat d'individus que mantenen una, dues o més relacions. En canvi, el mostratge per cadenes s'interessa més per descriure el tipus de relació que per l'existència o no d'aquesta relació.

Un cas especialment important de mostratge per cadenes és el *mostratge en bola de neu* que descriurem a la secció següent. Per a més informació sobre el mostratge per cadenes pot consultar-se Erickson (1979).

3. MOSTRATGE EN BOLA DE NEU

El mostratge en bola de neu és un procediment molt utilitzat en estudis socials, encara que no sempre s'entèn el mateix sota aquest nom. Erickson (1979) el descriu com una cadena construïda de la següent manera: es comença en una primera mostra als individus de la qual se'ls demana que anomenin coneguts seus, que constituïran la segona ona d'entrevistats, a aquests se'ls hi formularan les mateixes preguntes que als primers per arribar a la tercera ona i així successivament.

Les referències més citades sobre el mostratge en bola de neu són L. Goodman (1962) i P. Biernaki and D. Waldorf (1981). El primer ens proporciona la base teòrica del mètode per estimar certs paràmetres de l'estructura social en estudi. El treball de Goodman presuposa que la mostra inicial, que ell anomena etapa zero, és una mostra aleatòria dels individus d'una població. També són fixos el nombre, k , d'individus anomenats per cada individu entrevistat, nombre que es mantindrà constant en cada una de les etapes o ones del mostratge, i el nombre, s , d'etapes que es duran a terme. Biernaki i Waldorf expliquen com portar a terme efectivament un mostratge en bola de neu, el seu treball està basat en un estudi exploratori d'ex-addictes a l'heroïna. En aquest treball, de tipus qualitatiu, no es comença amb una mostra aleatòria ni es manté constant el nombre de gent anomenada. Treballs posteriors com els de Kaplan, Korf i Sterk (1987) i Cohen (1989) només entrevisten una o dues persones sortejades aleatòriament entre els anomenats per cada entrevistat. Una altra diferència entre aquests treballs i el de Goodman és que prolonguen indefinidament les cadenes fins que aquestes es trenquen per la impossibilitat d'establir nous contactes. La longitud mitjana de les cadenes així obtingudes resulta ser molt baixa, en el cas de Cohen és de 2 contactes.

El principal avantatge d'aquest mètode de mostratge és que permet obtenir mostres d'una grandària considerable (entre 100 i 200 individus) en poblacions

marginals i ocultes, amb una certa facilitat. Això permet fer un bon estudi qualitatiu del fenomen en que estem interessats i, a la vegada, permet estudiar la xarxa de relacions socials dels entrevistats. Per aquest tipus d'estudis no és essencial començar amb una mostra aleatòria de la població, el que simplifica considerablement la feina i abarateix els costos.

Les dificultats es presenten quan hom vol fer una inferència des de la mostra a la població en general. Per tal de fer inferència és indispensable un procediment o altre d'aleatorització. Si es vol estimar, per exemple, el nombre de persones que consumeix un determinat tipus de droga no es pot fer a partir d'un mostratge en bola de neu inicialitzat de forma arbitrària. Cal, per tant, precisar en primer lloc els paràmetres que es volen estimar i procedir després a dissenyar un procediment per fer-ho.

4. MÈTODES D'ESTIMACIÓ

L'estimació de la grandaria de poblacions marginals i ocultes presenta sempre importants problemes. En el millor dels casos, suposant que es tracta d'una població marginal només en el sentit de grup petit de gent en el conjunt de la població en general, ja es presenta un primer problema. En les mateixes condicions de grandaria de mostra i de nivell de confiança, l'error relatiu que hom té en estimar el percentatge d'individus de la població marginal quan aquest és proper al 1% és deu vegades superior al cas estàndard $p = q = 0.50$. Tinguem present que Cohen (1989) estima en menys d'un 1% el percentatge de gent que diu haver provat algun cop la cocaïna en estudis fets a Alemanya.

Quan afegim als problemes estrictament tècnics els que es deriven del fet d'estudiar un fenomen marginal en el sentit social de la paraula, com és el cas de poblacions sense un domicili fix, i els problemes que es deriven d'estudiar un fenomen amb tendència a l'ocultació, resulta que els mètodes de mostratge usuals són d'elevat cost, poc precisos i, a més, esbiaixats. En el cas d'una població amb $p = 0.01$ necessitariem una mostra de 1.000 persones per localitzar-ne aproximadament 10 de les que volem estudiar. Al marge del problema d'estimar amb precisió el percentatge real, ens trobem amb la dificultat adicional que 10 persones són del tot insuficients per un estudi qualitatiu que probablement acompanyarà l'estudi quantitatiu.

4.1 Mostratge aleatori simple

De les consideracions generals que acabem de fer es segueix que el mostratge aleatori simple no és apropiat per estimar la incidència de poblacions marginals.

No obstant, aquest tipus de mostratge ens serveix per disposar d'una referència dels errors deguts a l'atzar que es produeixen en estimar la incidència d'una població marginal a partir d'una mostra, en funció de la grandaria d'aquesta. La següent taula ens dona els errors absoluts que s'obtenen en estimar la incidència d'una població marginal quan aquesta és propera al 1%, pels nivells de confiança del 95% i del 90%:

$p = 0.01$		
n	95%	90%
1000	0.006	0.005
500	0.009	0.007
250	0.012	0.010
125	0.017	0.015
100	0.019	0.016

Aquesta taula es calcula a partir de l'expressió de l'error absolut

$$e \leq z_{\alpha} \sqrt{p(1-p)} / \sqrt{n}$$

on $\phi(z_{\alpha}) = 1 - \alpha$ quan ϕ és la funció de distribució de la variable normal estandard.

4.2 El mètode de captura-recaptura

Diferents autors han proposat utilitzar el procediment d'estimació anomenat *captura-recaptura* per a l'estudi de poblacions marginals. Hartnoll, Lewis, Mitcheson i Bryer (1985) utilitzen aquest procediment per estimar els consumidors d'opiàcids en dos barris de Londres, comparant llistes d'addictes de diferents anys.

L'origen d'aquest mètode prové del camp de la Biologia i es basa en l'elecció de dues mostres independents. Anomenant N a la grandaria de la població que es vol conèixer, diguen n_1 i n_2 a les grandaries de les dues mostres i x el nombre d'individus que són de les dues mostres alhora, la proporció d'individus de la primera mostra dintre de la població n_1/N ha de ser aproximadament igual a la proporció d'aquests individus dintre de la segona mostra x/n_2 . Aleshores de $n_1/N = x/n_2$ obtenim l'estimació de la grandaria de la població $N = n_1 \cdot n_2/x$.

Sudman, Sirken i Cowan (1988) analitzen les hipòtesis necessàries per utilitzar aquest procediment. Les resumirem breument en els següents punts:

1. Cada individu de la població ha de tenir la mateixa probabilitat de figurar en la segona mostra, tant els que han estat escollits en la primera mostra com els que no ho han estat.
2. Cal suposar independència entre el fet d'observar un primer individu en les dues mostres i el fet d'observar individus successius. És a dir, la probabilitat d'observar un individu en les dues mostres no ha de venir condicionada pel fet d'haver-ne observat d'altres previament.
3. La grandària de la població d'estudi s'ha de mantenir constant en el temps en que es realitzen les observacions i les prediccions.

Aquestes hipòtesis són difícils d'assumir en estudis sociològics en que s'utilitzen procediments de mostratge diferents dels estàndard. Concretament, en utilitzar mostratge per cadenes, com en el cas del mostratge en bola de neu, es presenten algunes dificultats que passem a comentar.

En el mostratge de bola de neu és difícil assumir que tot individu de la població té la mateixa probabilitat de figurar a la mostra. Això és especialment greu si els individus de la primera fase o etapa zero de la bola de neu no són escollits aleatòriament i el nombre de punts inicials de les cadenes és petit. En general és erroni suposar que tots els individus de la població marginal en estudi estan lligats entre sí per alguna cadena. Per tant, cal procurar iniciar les cadenes en una primera mostra aleatòria i, si això no és possible, iniciar cadenes en llocs el més diversos possible.

Per tal d'utilitzar el procediment de captura-recaptura no és necessari suposar gaires condicions per a la primera mostra. La primera mostra serveix essencialment per "marcar" alguns individus de la població per tal d'observar posteriorment si apareixen en la segona mostra. Així doncs no hi ha inconvenient perquè la primera mostra sigui escollida pel procediment de bola de neu, amb començament aleatori o no aleatori. Les dificultats es presenten si hom pretèn escollir també la segona mostra pel procediment de bola de neu.

Per tal d'evitar les dificultats corresponents al primer punt ja hem dit que cal començar el mostratge en bola de neu en una primera mostra aleatòria però això no evita les dificultats derivades del segon punt. La dificultat més important que trobem en intentar aplicar el mètode de captura-recaptura amb dues mostres escollides pel mètode de bola de neu és el problema de la independència. Quan trobem un individu que forma part de les dues mostres, automàticament la gent relacionada amb ell, que pot formar part de la següent etapa del mostratge en

bola de neu, té una probabilitat molt més alta de formar part de les dues mostres que la gent que no està relacionada directament amb ningú escollit previament. Podríem dir que es produeix un efecte de "contaminació" que fa que la gent propera als individus de les dues mostres tinguin tendència a formar part de les dues mostres.

La dificultat apuntada en el tercer punt no és preocupant en aquest cas i es pot evitar realitzant l'estudi en un espai de temps el més curt possible i restringint les prediccions a aquest espai de temps.

4.3 Estimacions indirectes

Per estimar la grandària de certes poblacions marginals com els consumidors de certs tipus de drogues s'utilitzen molts cops estimacions indirectes. Les més freqüents són les estimacions fetes a partir de dades dels centres de salut públics sobre casos de abús de drogues, de mortalitat per reacció aguda adversa a drogues, casos de malalties relacionades amb la drogodependència com el cas del SIDA, etc. (veure el Sistema d'Informació de Drogodependència de Barcelona). Altres indicadors utilitzats són les quantitats de drogues decomisades per la policia. Per exemple, Cohen (1989) compara les quantitats de heroïna i cocaïna decomisada a Amsterdam entre 1977 i 1988. D'aquest tipus d'estudi es pot deduir quin tipus de droga és la més utilitzada o bé si el seu consum, relativament a les altres, és més important o menys. Sempre queda el problema de saber si un increment de droga decomisada està relacionat amb un augment del consum o amb un increment d'eficàcia policial.

Les dades relatives a l'assistència sanitària són realment d'interès per a certs tipus d'addicció. Els addictes a l'heroïna, per exemple, recorren freqüentment a la xarxa pública d'atenció sanitària, això fa que una proporció important d'aquests consumidors estigui controlada i a partir d'ella es poden fer extrapolacions. Altres tipus de drogues com la cocaïna semblen més difícils de detectar per aquest procediment. Potser els problemes de salut que es presenten en aquest cas són més difícils de relacionar amb el consum d'aquesta droga. El fet és que les dades mèdiques concorden menys amb les dades policials respecte a la importància d'aquest consum.

Una dada objectiva d'interès en l'addicció a l'heroïna és el nombre de xeringues que s'utilitzen per al seu consum. Segons dades del Col·legi de Farmacèutics a Catalunya es venen al detall 13.350 xeringues al dia del model utilitzat per al consum d'heroïna. D'aquestes es calcula que un 80% estan destinades a la droga. (Veure estudi sobre l'heroïna a Barcelona i Catalunya (1991)).

4.4 El mètode nominatiu

El mètode nominatiu és una tècnica d'estimació basada en la informació proporcionada pels individus d'una mostra sobre la xarxa formada pels seus coneguts que previament ha identificat. Hartnoll et al. (1985) utilitzen aquest mètode per estimar els addictes a opiàcids. A cada individu d'una mostra se li demana que *anomeni* els seus coneguts que són consumidors d'opiàcids i que indiqui quins d'aquests han acudit a centres de salut públics per demanar ajuda pels problemes derivats de la seva addicció. El cocient entre els que han acudit a centres de salut públics i els que no ho han fet s'utilitza per extrapolar el nombre de consumidors a partir del nombre de gent atesa en aquests centres de salut per problemes relacionats amb la drogaaddicció.

El mètode nominatiu és sens dubte una eina important però cal emprar-la amb prudència puig que està subjecte a diferents tipus d'errors. Cal en primer lloc que la mostra inicial sobre la que es basa el procés nominatiu sigui realment representativa. En segon lloc hi ha els problemes d'ocultació i de mala informació dels individus que anomenen els seus coneguts. Pot ocórrer, per exemple, que algú hagi acudit als centres de salut i que un amic seu no ho sàpiga. També es pot donar el cas que la gent que acudeix als centres de salut pública tinguin menys inconvenients en manifestar públicament la seva addicció a les drogues i ocultin menys la seva condició als seus coneguts.

5. PROPOSTA D'UN NOU MÈTODE DE MOSTRATGE

En aquesta secció presentem un nou mètode de mostratge.

Aquest mètode està pensat per estimar el nombre d'habitants de Barcelona que són o han estat consumidors de cocaïna. El mètode està basat en dues seleccions aleatòries amb un enfocament similar al mostratge per xarxes. En una primera etapa s'escull una mostra aleatòria del segment de població en estudi, pels procediments habituals. A cada un dels individus d'aquesta primera mostra se li demana que anomeni un determinat nombre d'amics seus. Es procedeix després a una aleatorització de la gent que ha estat anomenada i es constitueix així la segona mostra de l'estudi. Aquesta segona mostra serveix per verificar les estimacions sobre l'ús de la cocaïna fetes amb la primera mostra i permet estudiar els fenòmens d'ocultació que es donen en aquest tipus d'estudis.

5.1 Elecció de la primera mostra

En tot estudi quantitatiu que preten estimar paràmetres d'una població a partir d'una mostra cal un procés previ d'aleatorització. Per aquest estudi proposem o bé una aleatorització a partir de dades censals o bé un procés de rutes

l'interés social del treball. S'ha de resaltar la importància de l'anonimat: volem entrevistar alguns dels coneguts de l'entrevistat sense que ells sapiguen que han estat citats per alguna amistat. Tampoc volem saber quines de les persones citades han consumit cocaïna. Simplement volem passar exactament el mateix qüestionari a algunes de les persones citades.

Un cop fets els advertiments anteriors cal procedir, juntament amb l'entrevistat al sorteig de 2, 3 ó 4 de les persones citades. Aquest nombre serà el mateix en tots els casos, depenent del valor que decidim per a k .

Cal convèncer finalment a l'entrevistat perquè ens faciliti la localització de les persones sortejades algunes de les quals formaren part de la segona mostra del nostre estudi.

Naturalment apareixeran dificultats en aquesta fase de la investigació. Pot ser que l'entrevistat no vulgui facilitar-nos la localització dels seus coneguts o que, malgrat disposar de l'adreça no sigui possible trobar a la persona en qüestió. Per tal de mesurar si aquestes dificultats són degudes a l'ocultació associada als temes de drogues es convenient realitzar la tercera fase de l'entrevista en altres individus de la mostra inicial que responguim negativament a la pregunta a.2. Aquests entrevistats als que se'ls hi realitzarà la tercera part de l'entrevista també seran escollits aleatoriament.

Finalment, en acabar l'entrevista, cal preguntar al nostre entrevistat si ell ha consumit algun cop

b.1) Hash

b.2) Cocaïna

La resposta a aquestes preguntes ha de ser alguna de les tres possibilitats següents:

1. Assegura que no
2. Assegura que si
3. Prefereix no contestar

Potser fora interessant poguer fer la mateixa pregunta relativa a les persones citades per l'entrevistat. No obstant, mantenir l'anonimat pot afavorir la comunicació. La informació obtinguda en qualsevol cas és suficient pels nostres propòsits encara que en el primer cas es podria obtenir més informació amb menys costos.

5.5 Estimacions

En el nostre estudi disposem, en primer lloc, d'una mostra aleatòria formada pels n individus escollits inicialment. A partir d'aquesta mostra es poden obtenir dues primeres estimacions. En primer lloc podem calcular el percentatge d'individus que accepten haver consumit cocaïna i, en segon lloc, el percentatge d'individus sobre els que s'afirma que han consumit cocaïna. El primer resultat suposem que serà inferior al segon degut a que la ocultació del fenomen ha de ser més alta al fer la pregunta directament al interessat.

Associat a la mostra inicial tenim el conjunt, N , format per les kn persones anomenades en la primera fase. Aquests anomenats els classifiquem en tres categories. La primera categoria, N_{11} , la formaran els anomenats per una persona que respongui afirmativament a la qüestió a.2. La segona categoria, N_{12} , la formaran els nomenats per una persona que respongui negativament a la qüestió a.2 i afirmativament a la qüestió a.1. Finalment la tercera categoria, N_{13} , la formaran els anomenats per algú que respongui negativament a les dues qüestions.

La segona mostra estarà formada per individus del les tres categories anteriors. Els individus que s'escolleixin de la categoria N_{11} serviran per contrastar els resultats obtinguts amb la primera mostra. Els individus de les altres categories serviran per comparar les dificultats de localitzar i entrevistar les persones anomenades segons la categoria a la que pertanyen. Els anomenats de la categoria N_{12} presuposem que formen un nivell intermedi entre les altres dues i podrien, eventualment, suprimir-se segons les disponibilitats pressupostaries.

Naturalment no podrem conèixer tots els nivells d'ocultació. Per exemple, no podrem detectar aquella persona que no reconeix haver provat la droga ni als seus amics més propers. També hi ha el problema d'aquells casos en que els amics ho saben i no ho diuen. Aquest darrer cas pensem que es manifestarà en els problemes de localització dels amics i es podrà estudiar comparant les dificultats de localització en les tres categories que mencionarem anteriorment.

Finalment, per tenir una idea del nombre de gent que hauria de respondre afirmativament a la pregunta de conèixer algú que ha provat la cocaïna cal tenir present les dades seüents:

Suposem $n = 500$ i una incidència del fenomen d'un 6%. Suposem també que només un 50% dels casos són exposats pels amics. Aleshores si el nombre de gent anomenada per cada entrevistat és $k = 6$ cal esperar un 16% de gent, és a dir 80 persones, que afirmen conèixer algú que algun cop ha consumit cocaïna. Recomanariem escollir una mostra de la categoria N_{11} d'una grandaria igual al nombre de consumidors que es puguin esperar i que, a priori, estimem en 160. De les altres categories seria suficient escollir 80 persones de cada una.

6. BIBLIOGRAFIA

- [1] **Biernacki, P. & Waldorf, D.** (1981). "Snowball sampling". *Sociological methods and research*, 10, 141-163.
- [2] **Cohen, P.** (1989). Cocaine use in Amsterdam in non deviant subcultures. Universiteit van Amsterdam.
- [3] **Erickson, B. & Nosanchuk, T.** (1983). Applied network sampling social. North-Holland.
- [4] **Erickson, B.** (1979). "Some problems of inference from chain data", a Schuessler, K. (Ed.), in *Sociological Methodology*, Josey-Bass. London.
- [5] Estudi sobre l'heroïna a Barcelona i Catalunya. Federació de Barcelona de E.R.C. (1991).
- [6] **Franck, O.** (1978). Sampling and Estimation in large social networks, 1, 91-101.
- [7] **Goodman, L.** (1962). "Snowball sampling". *Annals of Mathematical Statistics*, 32, 245-268.
- [8] **Hartnoll, R.; Lewis, R.; Mitcheson, M. & Bryer, S.** (1985). "Estimating the prevalence of opioid dependence". In *The Lancet*.
- [9] **Kaplan, C.; Korf, D. & Sterk, C.** "Temporal and social contexts of Heroin-using populations". *The Journal of Nervous and Mental disease*, 175.
- [10] **Knoke, D. & Kublinski, J.** (1990). Network Analysis, SAGE Publications. London.
- [11] "Evolució indicadors SIDB". Sistema d'Informació de Drogodependències de Barcelona, Servei d'Informació Sanitària. Institut Municipal de la Salut.
- [12] **Sudman, S.; Sirken, H. & Cowan, C.** (1988). "Sampling rare and elusive populations". *Science*.