
Procediments i decisions per a determinar què s'observarà

PID_00248672

Gemma Sala
Rosa Borge Bravo
Albert Padró-Solanet i Grau

Índex

Introducció	5
1. Selecció de casos	7
1.1. Indeterminació	7
1.2. Biaixos de selecció de casos	8
1.3. Selecció de casos a partir de la variable independent	8
1.4. Selecció de casos a partir de la variable dependent	9
2. Selecció de variables	10
2.1. Error en la mesura de les variables	10
2.2. Biaix: omisió de variables rellevants	11
2.3. Ineficiència: inclusió de variables irrelevantes	11
2.4. Endogeneïtat: confusió de causes i de conseqüències	11
Activitats	13

Introducció

Es poden distingir tres posicions diferenciades quant a la relació entre els dos mètodes o paradigmes fonamentals de recerca social: el quantitatiu i el qualitatiu.

La primera posició és la purista, tant quantitativa com qualitativa (els extrems es toquen), i defensa que les dues metodologies són completament diferents, antagòniques, incommensurables i excloents. No té cap avantatge que l'investigador que fa servir un mètode tingui en compte el que es diria en l'altre mètode, perquè no li pot aportar res. Cada perspectiva parteix de concepcions radicalment diferents de la realitat social: el positivisme quantitatiu suposa la verificabilitat de la ciència, cosa que entra en contradicció amb el relativisme dels qualitativistes, que suposen que tota la realitat és socialment construïda.

La segona posició accepta que els dos mètodes són commensurables i tenen una mateixa lògica d'inferència científica. L'obra clau d'aquest punt de vista és la lògica de la recerca qualitativa de King, Keohane i Verba (2000). Suposa que els mètodes qualitatius s'apliquen en les fases inicials de la recerca científica en un camp, i es poden beneficiar d'aplicar l'esquema conceptual de la recerca quantitativa a la recerca qualitativa. Aquest és el punt de vista d'aquest mòdul didàctic, però ha estat criticat per alguns científics socials per reflectir un "imperialisme quantitativista", que suposa que la recerca quantitativa és superior a la qualitativa (Brady i altres, 2010, pàg. 3).

Finalment, la tercera posició, de vegades denominada dualista, està entremig de les dues anteriors. Els defensors d'aquesta posició creuen que les dues perspectives metodològiques es poden beneficiar l'una de l'altra, ja que totes dues tenen diferents fortaleses i febleses. Tal com diu el subtítol de l'obra que recull les majors contribucions en aquest debat (Brady i Collier, 2010), tot i que la recerca social empra "diferents eines", "comparteix" els mateixos "estàndards".

Referències:

Brady, H. E.; Collier, D. (2010). *Rethinking social inquiry: diverse tools, shared standards*. Lanham: Rowman & Littlefield.

Brady, H. E.; Collier, D.; Seawright, J. (2010). "Refocusing the Discussion of Methodology". *Rethinking social inquiry: diverse tools, shared standards*. Lanham: Rowman & Littlefield.

King, G.; Keohane, R. O.; Verba, S. (2000). *El diseño de la investigación social#: la inferencia científica en los estudios cualitativos*. Madrid: Alianza Editorial.

Si bé hi ha diversos mètodes de recerca social, tots comparteixen la mateixa lògica d'inferència. Cada mètode té les seves característiques pròpies des del punt de vista del nombre de casos, la mena de dades que fa servir, les preguntes que es formula o la relació amb la teoria. Però per a inferir coneixement general a partir de dades empíriques cal que tots compleixin certs requeriments

inferencials. Els apartats següents no es proposen una anàlisi intensiva de les característiques dels requeriments de la inferència, sinó que aspiren a proporcionar definicions concretes, com a suport marginal de les explicacions i els exemples de lectures obligatòries.

Per satisfer les regles de la inferència, el disseny de la investigació ha de tenir en compte certes consideracions en dos aspectes principals: la selecció de casos i la selecció de variables que ha d'observar. Pel que fa a la **selecció de casos**, la investigació ha d'evitar situacions com la supradeterminació, la multicolinealitat, i els biaixos que sorgeixen de l'observació de casos en determinades circumstàncies. Pel que fa a la **selecció de variables**, cal evitar errors en la mesura i en l'assignació de variables, la introducció de variables irrelevantes, o els problemes d'endogeneïtat. Els apartats següents expliquen en què consisteixen cada una d'aquestes consideracions. La seva rellevància consisteix en què si no les tenim en compte no podem estar segurs que les conclusions de la investigació siguin fiables.

Aquests requeriments de la inferència han estat tradicionalment aplicats al mètode estadístic. Recentment, però, es tendeix a considerar que també es poden aplicar als mètodes tradicionalment entesos com a qualitius. Des del moment que els diferents mètodes comparteixen la finalitat comuna d'arribar a conclusions generals, i científiques, tots ells han de satisfer les condicions que garanteixen que això sigui possible. L'aplicació de les regles quantitatives d'inferència a les investigacions de caràcter qualitatiu ha permès acabar amb la concepció tradicional sobre la naturalesa humanística, i exclusivament descriptiva de la investigació qualitativa. Aquesta "quantitativització" dels mètodes qualitius ha trobat resistències i crítiques entre els autors que defensen l'especificitat de cada mètode. Tot i així, es considera que mentre la recerca social estigui orientada a la construcció de teories generals ha de satisfer aquests requeriments. Només les investigacions atòriques, les interpretatives o les que aspiren a la formulació d'hipòtesis poden no tenir en compte les consideracions que es presenten en els apartats següents.

Bibliografia

Els exponents principals de l'extensió de les regles quantitatives d'inferència als mètodes qualitius són a: **G. King** i altres (2001). *El diseño de la investigación social*. Madrid: Alianza. Diversos capítols d'aquest llibre estan assignats com a lectures obligatòries en aquest curs.

Les reaccions a aquesta homogeneïtzació dels requeriments inferencials dels diversos mètodes d'investigació es troben en gran mesura compilades en el volum 89, núm. 2 de juny de 1995, de l'*American Political Science Review*. En aquest número, cinc politòlegs i sociòlegs de renom avaluen els avantatges i les possibilitats de defensar una única lògica inferencial per als dos tipus de mètodes. King, Keohane i Verba també responen al final del número. Teniu disponibles les lectures en PDF a l'aula.

1. Selecció de casos

Aquest apartat està dedicat a analitzar algunes de les consideracions o perills que s'han d'evitar a l'hora de triar els casos en què basarem la recerca.

1.1. Indeterminació

A vegades no es pot determinar si el fenomen que es vol explicar (la variable dependent) és el resultat d'un factor o un altre. Resulta impossible explicar d'on prové l'efecte causal que es vol explicar. Aquest problema s'anomena indeterminació. Quan hi ha indeterminació no es poden contrastar les hipòtesis causals.

La indeterminació es manifesta de diverses maneres: la sobredeterminació i la multicolinealitat.

Sobredeterminació

El problema de la sobredeterminació apareix quan el nombre de variables és més alt que el nombre de casos observats. Si el nombre de variables és més elevat que el nombre de casos, no es pot comprovar l'efecte causal de totes. Això és així perquè no hi ha prou casos com per a mantenir l'efecte de cada una de les variables controlades.

Exemple de sobredeterminació

Si dos casos observats es comporten de manera diferenciada en la variable dependent i presenten característiques diferents en dues variables independents (a i b), resulta impossible determinar quina d'aquestes causa la variació en el comportament de cada cas, o la proporció en la qual el causen, ja que tenim dues explicacions alternatives per al comportament diferent de la variable dependent. Tant pot ser que la variació de la variable dependent sigui a causa de la diferència en el valor de la variable explicativa a en cada cas o a l'efecte de diferents valors en la variable explicativa b . Per exemple, volem saber per què a Rússia va guanyar la revolució el 1917 (cas 2) i, en canvi, no el 1905 (cas 1). Una explicació és que el 1917 Rússia estava immersa en la Gran Guerra (variable a), l'altra explicació és que el 1917 hi havia hagut l'experiència de la revolta del 1905 (variable b). Com que tenim dues variables i dos casos que varien en totes dues, no podem determinar quina variable ha causat realment l'èxit de la revolució.

Aquest és un problema freqüent en els estudis comparats i s'anomena *problema de pocs casos-moltes variables*.

Multicol·linealitat

El problema de la multicol·linealitat es refereix a aquelles investigacions en les quals les variables estudiades estan correlacionades. És a dir, si bé les variables independents estan correlacionades amb la variable dependent (és a dir, determinats valors de la variable independent s'observen simultàniament amb determinats valors de la variable dependent moltes vegades) tal com s'espera

per a poder establir la causalitat, en els casos de multicol·linealitat les variables independents també estan correlacionades entre si, amb la qual cosa no es pot precisar si hi ha una veritable influència de cadascuna de les variables sobre la dependent. La multicol·linealitat és el mateix problema que la sobredeterminació, però vist des del punt de vista de l'estadística multivariant. Conseqüentment, el remei per als problemes de multicol·linealitat consisteix –com en el cas de la sobredeterminació– a recollir més casos.

1.2. Biaixos de selecció de casos

Els mètodes d'investigació social s'intenten aproximar als models experimentals que es desenvolupen en les ciències naturals. Si bé el mètode experimental es caracteritza per la selecció aleatòria de casos, en moltes subdisciplines de les ciències socials els han de triar segons criteris preestablerts per tal de poder observar la relació entre les variables estudiades. Així, doncs, se seleccionen els casos segons si presenten condicions que permetin contrastar les hipòtesis formulades. És en aquest procés de tria en què es poden escolar els principals biaixos en la recerca. La possibilitat de triar els casos a partir dels quals es basaran les nostres conclusions ens pot conduir inadvertidament a triar-los de manera que presentin les condicions que corroboren les nostres hipòtesis.

El millor criteri per a la selecció de casos és el que té en compte els valors de la variable independent en la selecció. En canvi, és corrent trobar casos en els quals la selecció de casos s'ha establert segons els valors de la variable dependent. És en aquests casos en els quals s'introdueixen els principals biaixos en la investigació.

1.3. Selecció de casos a partir de la variable independent

La selecció d'observacions segons la variable explicativa és la millor manera que els mètodes d'investigació social tenen per a aproximar-se a la investigació experimental. L'experimentació consisteix a sotmetre els grups observats (o casos) a diferents condicions de la variable independent (tractament enfront de control). Així, doncs, la selecció de casos segons la variable independent es pot entendre com l'aplicació de tractaments diferents dels casos observats.

L'estudi de l'efecte causal consisteix a saber si l'absència o presència de determinades condicions afecten el comportament de la variable dependent. Així, doncs, resulta evident que cal assegurar que els casos observats tinguin en compte la variació (o els diferents) valors de la variable independent.

Així, doncs, si volem conèixer l'efecte del partit polític en el govern sobre la política de reforma econòmica que s'ha dut a terme en diversos països de l'Europa de l'Est, caldrà que ens fixem en països on els partits en el govern són de dife-

rent signe polític. En altres paraules, per a determinar si una variable exerceix un efecte sobre una altra cal poder observar com varien els valors de la segona a mesura que els valors de la primera canvien.

1.4. Selecció de casos a partir de la variable dependent

La selecció de casos a partir dels valors de la variable dependent provoca biaixos en les conclusions de la investigació.

- Si es trien casos que tenen el **mateix valor en la variable dependent**, no es pot identificar què la fa canviar. Es corre el risc d'assignar la causalitat a qualsevol característica compartida pels casos observats. En altres paraules, ens portaria a creure que allò que es comú en tots els casos observats, genera resultats comuns. Per exemple, quan ens proposem esbrinar quines són les causes de l'esclat de guerres civils a diferents països africans i, per a fer-ho, triem els que han patit guerres civils, correm el risc de pensar que, per exemple, una elevada fragmentació ètnica o el baix finançament de l'exèrcit en són les causes si resulta que tots els casos que analitzem tenen aquestes característiques. Però aquesta conclusió és una fal·làcia. En la mesura que el valor de la variable dependent no varia no podem observar cap efecte. Per tant, és impossible determinar quines són les variables independents que expliquen el comportament de la dependent. Podria ser, per exemple, que hi hagués una elevada proporció de països sense guerres civils, però amb una elevada fragmentació ètnica, cosa que desmentiria la nostra conclusió. Per tal de mesurar efecte, cal observar variació en el comportament.
- Precisament el caràcter intuïtiu d'aquesta darrera proposició ha dut a molts investigadors a triar els casos d'estudi segons si **la variable dependent té valors diferents**. Si bé aquest procediment ens evita el problema del punt anterior, encara implica un biaix en la investigació. Si triem les observacions segons valors diferents en la variable dependent, podem identificar fàcilment com a variable independent allò que també presenta valors diferents entre els casos observats.

Lectura recomanada

Llegiu aquest capítol: G. King i altres (2000). "Determinar lo que se va a observar". A: *El diseño de la investigación social: la inferencia científica en los estudios cualitativos* (cap. 4). Madrid: Alianza.

2. Selecció de variables

Si bé l'apartat anterior resumeix a grans trets les consideracions que cal tenir en compte a l'hora de triar els casos de la investigació, les variables amb què volem explicar la realitat social també poden introduir problemes en la inferència.

D'una banda, poden tenir problemes relacionats amb la validesa i la fiabilitat de la mesura de les variables. Cal saber si les variables mesuren el que creiem que mesuren i que el que mesuren estigui ben mesurat. D'una altra banda, cal que les variables que estudiem siguin rellevants per a explicar el fenomen social que ens proposem entendre i que els donem el paper acertat en la seva relació amb les altres variables de l'estudi.

2.1. Error en la mesura de les variables

Un dels problemes freqüents en l'observació de la realitat social és la dificultat de traduir-la en categories o quantitats numèriques que la defineixin i que la facin analíticament observable. En aquest procés d'assignació de valors a les variables en el qual es poden infiltrar diversos errors en l'observació de la realitat social.

Les tres condicions fonamentals perquè els valors de les variables estudiades mesurin la realitat social d'una manera analíticament útil són:

- **Consistència o fiabilitat:** mesures diferents del mateix fenomen tenen els mateixos resultats. En altres paraules, que el criteri d'assignació de valors a les variables sigui homogeni a totes les observacions. La consistència també implica que els valors reflecteixin les similituds i les diferències que els casos presenten en la realitat. Això implica que els casos que presentin característiques similars, tinguin valors similars, i que els casos que tinguin característiques diferents, tinguin valors diferents en la mateixa proporció que les diferències que mostren en realitat.
- **Validesa:** els valors de les variables reflecteixen el que l'investigador creu que mesuren. És a dir, reflecteixen correctament la realitat perquè els indicadors de les variables representen i registren efectivament la propietat o característica de les unitats d'anàlisi que es pretén mesurar.
- **Completitud:** els conceptes i els valors analítics defineixen les característiques de cada cas de manera complerta. És a dir, que en el seu esforç per a simplificar la realitat social, les categories analítiques no perdin informació rellevant sobre els casos estudiats.

2.2. Biaix: omissió de variables rellevants

L'omissió de variables rellevants en l'observació pot esbiaixar els resultats de la investigació –és a dir, el valor de l'efecte causal entre dues variables al qual arribem a partir de l'estudi de certs casos no tendeix a centrar-se al voltant del valor real, si el poguéssim determinar.

Les variables rellevants són aquelles que d'acord amb la teoria causen un efecte significatiu en la variable dependent. Per tant, el fet que algunes de les variables que afecten el comportament de la variable dependent no estiguin considerades pot significar que l'estimació de l'impacte de la variable que volem estudiar estigui esbiaixada, per sobre o per sota de l'efecte real. No solament hi ha una proporció de l'explicació del comportament de la variable dependent que no està explicada, sinó que a més hem assignat la causalitat a terceres variables independents, magnificant-les. Així, doncs, l'omissió de variables independents rellevants implica un biaix en les conclusions de la recerca, perquè infla el pes explicatiu de les variables independents que sí s'han tingut en compte.

Tot i així, l'omissió de variables explicatives rellevants només produeix biaixos quan les variables incloses i les omeses estan correlacionades. Si aquestes variables no estan relacionades, l'omissió d'una d'elles no afecta a l'estimació correcta de l'efecte de la variable independent que volem estudiar.

2.3. Ineficiència: inclusió de variables irrellevants

També és problemàtica la inclusió de variables irrellevants en el nostre estudi. En concret, la inclusió de variables irrellevants produeix ineficiència –la variància del valor de l'efecte causal observat augmenta. És a dir, la possibilitat que la nostra recerca hagi estimat el valor real de l'efecte entre les variables que volem estudiar és inferior.

Passa el mateix amb l'omissió de variables rellevants, si les variables irrellevants són independents de les variables dependents, la inclusió de variables irrellevants no importa massa. Però si, per contra, hi ha correlació entre aquestes variables, els resultats de la investigació seran ineficients.

2.4. Endogeneïtat: confusió de causes i de conseqüències

L'endogeneïtat es refereix al fet que els valors de les variables explicatives són una conseqüència i no una causa de les variables explicades. En altres paraules, la relació entre les variables incloses a l'estudi és oposada a la que el disseny de la investigació li assigna. Aquest és un problema comú perquè en molts casos, l'assignació del paper de les variables en el procés causal en la realitat pot no estar clar o existir causalitat recíproca entre el que s'ha considerat variable

Lectura recomanada

Llegiu el capítol 5 de: G. King i altres (2000). "¿Qué evitar?". A: *El diseño de la investigación social: la inferencia científica en los estudios cualitativos* (cap. 5). Madrid: Alianza.

independent i variable dependent. En aquesta situació, tant les variables independents com les variables dependents tenen una relació que no reflecteix acuradament la seva relació causal.

Lectures obligatòries

Com a complement dels textos de King, Keohane i Verba, cal que estúdieu part del llibre d'**I. Lago** (2008). *La lógica de la explicación en las ciencias sociales* (pàg. 75-111). Madrid: Alianza. En aquest text teniu molts exemples que us podran il·lustrar els conceptes explicats en aquest mòdul.

A més, cal que torneu a estudiar i que tingueu present la lectura obligatòria del mòdul anterior: **E. Anduiza; I. Crespo; M. Méndez** (1999). "Las estrategias de investigación". A: *Metodología de la Ciencia Política* (capítol 3, pàg. 49-72). Madrid: CIS ("Cuadernos Metodológicos", núm. 28).

Activitats

A continuació es presenten diversos dissenys d'investigació. Avaluu els seus punts febles des del punt de vista de la possibilitat d'inferir coneixements generals a partir dels estudis proposats.

1. Reviseu la investigació d'Arend Lijphart, *Politics of Accommodation*, en la qual analitza la relació entre la presència de divisions socials segmentades i l'estabilitat democràtica a partir del cas d'Holanda.

Quin és el criteri de selecció de casos d'aquella investigació?

Considereu fins a quin punt es pot criticar aquesta investigació des del punt de vista de la selecció de casos segons la variable dependent?

2. El professor Garcia està interessat a saber fins a quin punt la repressió política en els països d'Amèrica Llatina ha afectat el creixement econòmic. Per tal d'estudiar el problema, Garcia ha recollit dades econòmiques dels deu països d'Amèrica Llatina que tenen l'índex de Producte Nacional Brut més alt. A l'hora en Garcia també es fixa en els indicadors de repressió política d'aquests països, i crea una escala de repressió en la qual zero indica repressió baixa i deu repressió alta. Aleshores demana a determinats especialistes que classifiquin el país del qual són experts en aquesta escala.

Quins són els defectes d'aquest disseny d'investigació?

Proposeu un disseny alternatiu que permetria una inferència menys problemàtica.

3. El professor Garcia estudia l'efecte de les estructures institucionals de certs països federals (on el poder polític està territorialment descentralitzat) en el ressorgiment de demandes regionals que reclamen més autonomia política. Per a analitzar aquesta relació, el professor se centra en dos països federals. En un les institucions federals es caracteritzen per un alt nivell de participació regional en les institucions de presa de decisions del govern central, i en l'altre no hi ha mecanismes de participació dels governs territorials en les institucions de presa de decisions centrals. Pel que fa a la variable dependent, el professor observa que en el primer país no hi ha demandes perifèriques que reclamen més autonomia política, mentre que en el segon hi ha regions determinades que han aixecat aquestes demandes i que, per tant, posen l'estabilitat del sistema federal a risc. La seva conclusió és que les estructures més participatives afavoreixen l'estabilitat del sistema federal i viceversa. També considera que la seva investigació evidencia que les teories anteriors que expliquen l'emergència de demandes perifèriques segons si existeixen cultures nacionals diferenciades a les regions no són vàlides.

Quines són les virtuts i els inconvenients d'aquesta recerca?

Fins a quin punt aquesta investigació permet contrastar els arguments de les teories anteriors que basen la inestabilitat federal en l'existència de cultures nacionals perifèriques diferenciades, com l'autor pretén?

Fins a quin punt el seu indicador d'estabilitat democràtica pot comportar problemes d'errors de mesura?

