

Är icke-sannolikhetsurval aldrig representativa?

Jan Wretman

Webbpanelkommittén

Det kommer att handla om:

- Begreppet ”representativitet”.
- Bedömning av skattningars kvalitet, speciellt vid icke-sanolikhetsurval.
- Webbpanelundersökningar, speciellt marknadsundersökningar.

Vad är representativitet?

- Stickprovet "liknar" i någon mening den population man är intresserad av?
- Man kan på något sätt "generalisera" från stickprovet till populationen?

Vad betyder det?

Representativitet, exempel på olika språkbruk:

- Samma fördelning som i populationen m.a.p. vissa bakgrundsvariabler.
- Alla inressanta grupper i populationen finns representerade i stickprovet.
- Varje enskild medlem i stickprovet är en typisk representant för populationen.

(Representativitet, olika språkbruk, forts.)

- Alla i populationen har haft lika chans att komma med i stickprovet.
- Urvalet har gjorts enligt någon metod som möjliggör bra skattningar.

Hur används termen nu?

- Termen "representativitet" förefaller numera undvikas av professionella statistiker.
- Icke-statistiker, som gjort ett urval på något "icke-statistiskt" sätt, brukar ibland försvara sig genom att säga att urvalet i alla fall är "representativt". Oklart vad man då menar.

*Beträffande olika användningar av termen
"representativitet", se:*

- Tore Dalenius: *Statistisk tidskrift*, 1953.
- William Kruskal och Frederick Mosteller:
International Statistical Review, 1979-80 (fyra artiklar).

Vad har en undersökning för kvalitet?

- Alla slutsatser från urvalsundersökningar är osäkra. Hur osäkra? Hur mycket kan man lita på resultaten?
- Bryr sig användaren om hur stor osäkerheten är?
- Är kvalitetsnivån acceptabel i relation till undersökningens syfte?

Kvaliteten beror av många olika saker, bl.a.

1. Mätteknik:

- Är undersökningens syfte klart?
- Är frågorna relevanta för detta syfte?
- Kan den svarande förstå frågorna?
- Går frågorna att besvara?
- Är frågorna korrekt översatta?
- Kan vi lita på att svaren är riktiga?
- Kan eventuell intervjuare ha påverkat svaren?

2. Urval:

- Är undersökningspopulationen väldefinierad?
- Urvalsförfarande?
- Täckningsproblem?
- Bortfall?
- Hur har man fått den slutliga mängden av svarande?

3. Databearbetning:

- Databehandlingsfel?
- Hur har skattningarna beräknats?
- Vägning?

Temat för detta seminarium :

- Kvalitetsfrågor som har med *urval* att göra. (Vi tar inte upp mätteknik och databehandling.)
- Speciellt webbpanelundersökningar.

Olika typer av samhällsundersökningar, t.ex:

- Officiell statistikproduktion.
- Undersökningar inom samhällsvetenskaplig forskning.
- Opinionsundersökningar.
- Marknadsundersökningar.

Olika undersökningar kan ha olika kvalitetskrav:

- *Officiell statistik.* Ligger till grund för viktiga beslut. Höga kvalitetskrav. Felaktigheter framkallar kritik.
- *Marknadsundersökningar.* Viktiga för uppdragsgivaren? Vet uppdragsgivaren vad han ska använda resultaten till? Vaga kvalitetskrav. Felaktigheter kommer sällan att upptäckas. Ingen kommer att kritiseras.

Att bedöma skattningars kvalitet:

- Urval: Population \rightarrow Stickprov
- Estimation: Population \leftarrow Stickprov
- Hur stickprovet har erhållits är viktigt.

Två stora klasser av urvalsförfaranden:

- Sannolikhetsurval
- Icke-sannolikhetsurval

Sannolikhetsurval:

- Väldefinierad population.
- Varje medlem i populationen ska ha en känd sannolikhet (större än noll) att väljas ut.
- Vanligen görs urvalet från någon form av urvalsram.

Icke-sannolikhetsurval:

- Allt som *inte* är sannolikhetsurval, t.ex. självselektion, ”kvotaurval”, ”purposive selection”, snöbollsurval.

En skattnings kvalitet:

- Från givna stickprovdata har vi beräknat en skattning av en viss populationsparameter.
- Hur "bra" är denna skattning? Vad har den för "noggrannhet"? Ligger skattningens värde "nära" det för oss okända värdet på populationsparametern?

- FALL 1: *Strikt sannolikhetsurval*
- Vanligt att beräkna ett konfidensintervall: skattning \pm "felmarginal".
- Vi vet inte om det just erhållna intervallet täcker parametervärdet eller ej, men vi vet att intervallet har beräknats enligt en metod, som har följande attraktiva egenskap:

- I det långa loppet (d.v.s. en tänkt lång serie upprepade undersökningar) skulle vi få intervall som med viss önskad relativ frekvens (t.ex. 95%) täcker det sanna parametervärdet.
- Detta gäller oavsett hur populationen ser ut.

- Vi vet fortfarande inte om det just erhållna intervallet täcker det sanna parametervärdet eller ej.
- Men vi hoppas att vi inte har fått ett av dessa sällsynta misslyckade intervall, som inte täcker det sanna parametervärdet.

- Konfidensintervallet i sig är inget absolut mått på skattningens noggrannhet.
- Men man kanske kan säga att *ju kortare konfidensintervall, desto högre noggrannhet hos skattningen.*

FALL 2: *Icke-sannolikhetsurval.*

- Ingen teori eller metod finns för att förse skattningen med ett osäkerhetsintervall.
- Uttalanden om skattningens egenskaper måste bli baserade på diverse information och kunskaper som man råkar ha tillgång till, och som kan variera från fall till fall.

- Inga vedertagna numeriska kvalitetsmått.
- I stället för kvalitetsdeklaration kan man tänka sig en *beskrivning* av hur undersökningen har gått till från början till slut (inklusive vissa numeriska mått).
- *Beskrivningsmått* snarare än kvalitetsmått.
- *Flera mått* i stället för ett enda.

- Förmågan att ta fram en sådan här beskrivning (med tillhörande beskrivningsmått) är en sorts kvalitetsindikator i sig. Det visar om undersökaren har läget under kontroll.
- Användaren får dra slutsatsen att kvaliteten är den man kan förvänta sig under de beskrivna omständigheterna.

”Since no theoretical treatment is possible for nonprobability sampling methods, no general results can be obtained. Each application of a nonprobability sampling method must be evaluated individually.”

(Graham Kalton, *Encyclopedia of Statistical Sciences*, 1988)

I samband med en beskrivning kan man t.ex. fråga:

- Finns det skäl att tro att de svarande skiljer sig från populationen m.a.p. *undersökningsvariabelns* fördelning?
- Hur ser de svarandes fördelning ut m.a.p. vissa *bakgrundsvariabler*, jämfört med populationen?
- Hur var det med personers villighet att medverka i undersökningen?

Sannolikhetsurval kontra icke-sannolikhetsurval:

- Blir det bättre resultat med sannolikhetsurval än med icke-sannolikhetsurval?
- Inte säkert. Men vid icke-sannolikhetsurval saknas teori för att mäta osäkerheten.

- Undersökningar med icke-sannolikhetsurval visar sig ibland fungera bra, ibland mindre bra.
- När gäller det ena, och när gäller det andra?
- Skulle man kunna studera denna fråga med någon sorts vetenskaplig systematik?

Problem:

- Ett växande problem med *sannolikhetsurval* är det ökande bortfallet.
- Även om man från början hade sannolikhetsurval, men om bortfallet är stort, så får man i praktiken betrakta de svarande som ett icke-sannolikhetsurval från populationen.
- Vad hade man då för nytta av sannolikhetsurval?

Webbpanelundersökningar:

- Urval i två faser
- Urval: Population → Panel → Stickprov
- Estimation: Population ← Panel ← Stickprov

Två frågor om urvalet:

- Hur har panelen skapats?
- Hur har stickprovet för den specifika undersökningen valts från panelen?

FALL 1: *Panelen rekryterad genom icke-sannolikhetsurval.*

Ett *överdrivet* exempel (för att på ett naivt sätt tydliggöra vilka mekanismer som *kan* finnas):

- Skatta hur många procent i populationen som gillar en produkt P .

- Antag att alla som gillar P samtidigt är sådana som gärna är med i en panel, medan alla som inte gillar P är sådana som vägrar att vara med i en panel.
- Ett stickprov av personer från panelen kommer att innehålla enbart personer som gillar P .

- Då hjälper det inte ifall panelen har samma fördelning som populationen m.a.p. kön, ålder, utbildning etc.
- Det hjälper inte heller ifall urvalet från panelen till den specifika undersökningen har gjorts slumpmässigt.

(Exemplet slut)

Exempel på uttalanden (1):

- "... eftersom det inte går att göra slumpmässiga urval i en webbundersökning, så får man vara nöjd med att försöka göra det näst bästa så bra som möjligt. [Vi] har kontrollerat [vår] panel och anser att den är representativ, med undantag för personer som är äldre än 70-75 år."

(Referat i *Qvintensen* från Surveyföreningens seminarium om politiska opinionsundersökningar.)

Exempel på uttalanden (2):

- "Researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values. [...] Claims of "representativeness" should be avoided when using these sample sources."

(AAPOR, 2010)

FALL 2: Panelen rekryterad i samband med omnibusundersökning.

- Antag: Omnibus med sannolikhetsurval.
- Antag: Sannolikhetsurval från panelen.
- Problem: Stort bortfall.

Påhittat (men kanske inte helt orealistiskt)

räkneexempel:

Omnibus 40% svarande

Rekrytering 40% av dessa vill gå med i panel

Profilunders. 70% av de föregående blir kvar

Panelunders. 40% av de utvalda svarar

Räkneexempel, forts:

- Efter profilundersökningen kvarstår i panelen ung. 11% av det ursprungliga stickprovet.
($0,4 \times 0,4 \times 0,7 = 0,112$)

Räkneexempel, forts:

- Panelen kan i realiteten knappast ses som erhållen genom sannolikhetsurval från populationen.
- Sedan tillkommer bortfallet i den specifika undersökningen.

Räkneexempel, forts:

- Fast man eftersträvat formellt sannolikhetsurval i båda urvalfaserna, så kan man (p.g.a. det stora bortfallet) till slut ändå inte räkna som om det vore sannolikhetsurval.
- Man är hänvisad till samma sorts resonemang som när man drar slutsatser från icke-sannolikhetsurval.

Räkneexempel, forts:

- Om huvudprincipen är att undersökningen ska vara så billig som möjligt, finns det då anledning att lägga ner pengar på sannolikhetsurval, som ändå inte kommer att fungera?
- En svår fråga, som skulle behöva studeras mer systematiskt.

(Räkneexemplet slut)

Några avslutande frågor:

- Marknadsundersökningar via webbpanel =
= en sorts "lägre division" av undersökningar?
- Man får vad man betalar för?
- Man vet inte så mycket om resultatens kvalitet?
- Beställaren bryr sig inte om resultatens kvalitet?
- Bättre än ingenting?
- Hur ska man studera dessa saker?