

2005-04-13

Enkät till målsman

Inledning

Enheten för utbildning och arbetsmarknadsstatistik vid Statistiska centralbyrån (SCB) har under februari – april 2005 genomfört en enkätundersökning till målsmän på uppdrag av Göteborgs universitet. Syftet med undersökningen är att undersöka skolans betydelse för eleverna.

Undersökningen genomfördes som en postenkät med tre påminnelser. Urvalet utgjordes av 10 021 elever, vilket motsvarar en enkät till målsman per elev. Totalt inkom 7 672 enkäter, vilket var 77 procent av urvalet. Frågeblanketterna registrerades via skanning. Från det skannade materialet framställdes en datafil.

Kerstin Båsjö har varit produktansvarig. Henrik Gustafsson har varit ansvarat för metodfrågor och viktning.

Population och urval

Populationen, d.v.s de objekt som man vill kunna dra slutsatser om, utgörs i denna undersökning av individer. För att kunna dra ett urval från populationen skapas en urvalsram som avgränsar, identifierar och möjliggör koppling till objekten i populationen. I denna undersökning är urvalet samtliga elever som ingår i elevpanel 6. I elevpanel 6 skapades urvalsramen utifrån SCB:s elevregister.

Urvalet av elever i elevpanel 6

Vid starten av elevpanel 6 drogs ett stratifierat 2-steps klusterurval om 10 352 personer med hjälp av ett av SCB egenutvecklat urvalsprogram. Ett stratifierat obundet slumpmässigt urval innebär att alla objekt inom respektive stratum har samma sannolikhet att komma med i urvalet.

Urvalet drogs i två steg. I steg 1 drogs 35 kommuner. I steg 2 drogs 10 352 elever fördelade på 227 rektorsområden.

Tabell 1 Population och urval (elevpanel 6, årskurs 3)

	Stratum 1	Stratum 1	Stratum 2	Stratum 2
Kommun	Population	Utvalda	Population	Utvalda
Stockholm	61	7	98	7
Uppsala	20	4	35	4
Linköping	14	4	20	4
Malmö	22	4	31	4
Göteborg	26	5	101	5

2005-04-13

	Stratum 1	Stratum 1	Stratum 2	Stratum 2
Kommun	Population	Utvalda	Population	Utvalda
Södertälje	10	4	19	4
Nacka	13	4	17	4
Jönköping	17	5	26	5
Kungsbacka	12	4	10	4
Västerås	19	5	21	5
Sundsvall	8	4	24	4

Haninge	13	4	6	4
Norrtälje	3	3	30	5
Nyköping	3	3	24	4
Partille	8	4	3	3
Uddevalla	9	4	9	4
Örnsköldsvik	11	4	2	2

Upplands-Väsby	10	6
Enköping	14	7
Staffanstorps	9	5
Stenungsund	3	3
Lidköping	14	7
Sandviken	17	7

Håbo	5	3
Tierp	4	4
Lomma	7	4
Hammarö	6	4
Forshaga	4	4
Säter	5	3

Rättvik	5	5
Hofors	5	5
Ånge	4	4
Åre	7	7
Vännäs	4	4
Arjeplog	1	1

Vid de insamlingar som gjorts av administrativa data i årskurs 3 - 5 har ett antal elever utgått ur undersökningen. Anledning till detta har bl.a. varit överflyttning till särskola och vägran.

För att få fram aktuella adressuppgifter till den nu genomförda enkäten till målsmän genomfördes en identifikationskontroll av personerna i urvalet mot befolkningsdata. Vid kontrollen visade det sig att 22 elever inte längre tillhörde populationen utan utgjorde s.k. övertäckning. Den vanligaste orsaken till övertäckning var att eleverna hade emigrerat.

2005-04-13

Sekretess och utlämnande

Frågeblanketten innehöll ett missivbrev där uppgiftslämnarna fick information om

- undersökningens bakgrund, syfte och att undersökningen
- att den genomfördes i samarbete mellan SCB och Göteborgs universitet
- att enkätdata kommer att kopplas samman med de uppgifter som finns för eleven
- att data kommer att hämtas från Registret över totalbefolkningen (RTB), Registerbaserad arbetsmarknadsstatistik (RAMS), Yrkesregistret, Skolverkets elevregister och Högskoleverkets universitets- och högskoleregister
- att det är frivilligt att delta i undersökningen
- att SCB levererar en avidentifierad datafil till Göteborgs universitet efter undersökningens slut.

För att SCB ska kunna lämna ut ett datamaterial så krävs ”informerat samtycke” av uppgiftslämnarna. Det innebär att de genom att besvara blanketten och skicka in den godkänner att deras svar kompletteras med bakgrundsvariabler och behandlas på det sätt som beskrivs i informationsbrevet.

SCB har gjort en intern sekretessprövning för utlämnandet av enkätdata. Dessutom har en sekretessöverenskommelse upprättats mellan Göteborgs universitet och SCB med avseende på hur det avidentifierade datamaterialet får hanteras.

Behandlingen av personuppgifter i undersökningen har anmälts till SCB:s juridiska kansli. Beslutet därifrån medger endast att registret (datamaterialet) bevaras på SCB till 2005-07-31 då uppgifterna kommer att förstöras.

Variabler

Göteborgs universitet har utformat frågorna i frågeblanketten.

Enkäten bestod av 15 numrerade frågor, flera av dem hade delfrågor vilket genererade ca 80 frågor.

Datainsamling

Frågeblanketterna sändes till urvalspersonerna via post. I informationsbrevet ombads de besvara frågorna och skicka tillbaka frågeblanketten till SCB. Det första utskicket genomfördes den 3 februari, sedan skickades tre påminnelser ut till dem som inte besvarat frågeblanketten. Den första påminnelsen bestod av ett tack- och påminnelsekort, den andra och den tredje påminnelsen var påminnelser med ny enkät. Insamlingen avslutades den 6 april.

Tabell 2 Beskrivning av inflödet i procent

2005-04-13

Efter första utsändning	28
Efter ToP	51
Efter enkätpåminnelse	65
Efter enkätpåminnelse	77

Bortfall

Bortfallet består dels av objektsbortfall som innebär att frågeblanketten inte är besvarad alls och av partiellt bortfall som innebär att vissa frågor på blanketten inte är besvarade. Om bortfallet skiljer sig från de svarande med avseende på undersökningsvariablerna så kan skattningarna som grundar sig på enbart de svarande vara skeva.

Objektsbortfall kan bland annat bero på att uppgiftslämnaren inte är villig att delta i undersökningen, att uppgiftslämnaren inte går att nå eller att uppgiftslämnaren är förhindrad att medverka, t ex. på grund av språksvårigheter. Objektsbortfallet i denna undersökning redovisas i tabell 3 nedan. Med "ej avhörda" menas att ingen uppgift om varför frågeblanketten inte är besvarad har lämnats. Med avböjd medverkan menas att SCB meddelats att uppgiftslämnaren inte vill medverka i undersökningen.

Tabell 3 Beskrivning av objektsbortfall

	Antal	Procent
Ej avhörda	2168	21,6
Tillfälligt utomlands	1	0
Postreturer	98	1,0
Avböjd medverkan	56	0,6
Språksvårigh	4	0
Övertäckning	22	0
Totalt	2 349	23

Partiellt bortfall kan bero på att frågan är svår att förstå, är känslig, att uppgiftslämnaren glömmer att besvara frågan, att lämpligt alternativ saknas eller att instruktionerna vid hoppfrågor misstolkas. Till partiellt bortfall räknas även dubbelmarkeringar och svar som inte kan tydas. I denna undersökning varierar det partiella bortfallet mellan under en procent till ca två procent för samtliga frågor utom fråga 2 och 3.

Fråga 2 handlar om föräldrarnas högsta utbildning. Det partiella bortfallet för mamman är ca 4 procent och för pappan ca 13 procent. Samma förhållande gäller i fråga 3 om huvudsaklig sysselsättning, ett högre partiellt bortfall för pappan, ca 3 procent resp. ca 11 procent.

2005-04-13

Databeredning

Samtliga inkomna frågeblanketter prickades av i dataregistreringsprogrammet RODE. Frågeblanketterna skannades med programvaran EYES&HANDS.

Kontroller genomfördes under och efter skanningen. Då kontrollerades bl.a. dubbelmarkeringar och att endast valida värden förekommer i materialet.

Svarsdatafilen kompletterades med vikter för uppräknig till populationsnivå.

Viktberäkning och estimation

Vikter finns med på svarsfilen så att resultat kan redovisas för hela populationen och inte bara för de svarande. Vikterna kompenserar för objektsbortfallet men inte för det partiella bortfallet. Om vikterna inte används så kan resultaten bli helt missvisande, speciellt som olika urvalspersoner haft olika sannolikhet att komma med i urvalet.

Vikterna i denna undersökning kan beskrivas med formeln:

$$W_k = w^1 w^2 w^3 w^4 w^5$$

Ursprungsvikten från elevpanel 6 består av $w^1 w^2 w^3 w^4$ medan w^5 är en kompensation för bortfallet i enkäten.

$$w^1 = \frac{N_{hij}}{N_{hijS}} \quad , \text{ antal elever i RO-området} /$$

antal elever i RO-området som svarat

$$w^2 = \frac{n_{hi}}{n_{his}} \quad , \text{ antal elever i utvalda RO-området i kommunen} /$$

antal elever i svarande RO-områden i kommunen

$$w^3 = \frac{N_{hl}}{n_{hl}} \quad , \text{ totalt antal RO-områden i resp. stratum} /$$

antal utvalda RO-områden

$$w^4 = \frac{N_{hi}}{n_{hi}} \quad , \text{ totalt antal kommuner i stratum} i /$$

antal utvalda kommuner

2005-04-13

$$w^5 = \frac{m_{hij}}{m_{hij5}} \quad , \text{ antal som fått enkäten i RO-området} /$$

antal som svarat på enkäten i RO-området

Sammanfattningsvis kan viktförfarandet beskrivas på följande sätt. Med utgång i den ordinarie vikten som räknats fram för EP6 har en justering gjorts för det bortfall som skett i samband med enkäten. Vikten för respektive rektorsområde har justerats med en form av "rak uppräknings". Kompensation har även gjorts för känd övertäckning.

Vikten som finns på kundfilen bör användas vid framställning av tabeller och diagram samt vid beräkning av statistiska mått.

Kvalitet

SCB tillämpar en bred definition på kvalitet. För att hålla en hög kvalitet och täcka kvalitetsdimensionerna används kontroller i undersökningens olika faser, allt från kundkontakter till dokumentation och uppföljning. I SCB:s MIS 2001:1 "Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik" ges en ingående beskrivning av SCB:s kvalitetsbegrepp och de olika kvalitetskomponenterna. Nedan beskrivs de begrepp som har betydelse för denna undersökning.

Innehåll

Statistiska mått

De statistiska storheterna (parametrarna) utgörs av totaler och kvoter. De senare består vanligtvis av procentuella andelar. Statistiken kan redovisas, förutom för hela målpopulationen, också på redovisningsgrupper.

Redovisningsgrupper

Ofta redovisas statistik inte bara för hela populationen utan också för delgrupper, redovisningsgrupper. Några frågor i enkäten är lämpliga att använda som redovisningsvariabler.

Referenstider

Referensperioden i den här undersökningen är februari - april 2005 för objekt och samtliga variabler.

Tillförlitlighet

2005-04-13

Ramtäckning

Täckningsfel, under- och övertäckning, innebär att urvalsram och population inte helt stämmer överens. Undertäckning innebär att vissa enheter som ingår i populationen saknas i urvalsramen. Övertäckning innebär att enheter som inte ingår i populationen ändå finns i urvalsramen. Ett sätt att minska täckningsfelen är att ha bra och uppdaterade register.

I denna undersökning visade sig 22 personer vara övertäckningsobjekt. Det var personer som emigrerat, saknade adressuppgift eller hade uppgift om sekretess i RTB. Undertäckning är vanligtvis svårare att upptäcka än övertäckning. I denna undersökning vet vi att det finns undertäckning i ramen. Den består av personer som har invandrat efter det att urvalet drogs i årskurs 3.

Urval

Denna kvalitetskomponent avser fel som uppkommer på grund av att endast ett urval av populationen undersöks. Urvalsfel är således den avvikelse mellan ett skattat värde (estimat) och det faktiska värdet (parametern) som beror på att man inte undersöker alla objekt i populationen. Urvalsfelens storlek minskar med en ökad urvalsstorlek.

Mätning

Ett fel som kan uppstå vid mätning är att lämnade uppgifter skiljer sig från faktiska uppgifter. Felet kallas mätfel och kan uppkomma då respondenten inte minns de faktiska uppgifterna, missförstår frågan eller medvetet svarar felaktigt.

Bearbetning

Vid den manuella och maskinella bearbetningen av datamaterialet kan så kallade bearbetningsfel uppstå. Exempel på bearbetningsfel är registreringsfel och kodningsfel.

Dessa fel kan förhindras och upptäckas i de kontroller som genomförs vid dataregistreringen. I den här undersökningen bedöms bearbetningsfel och registreringsfel vara försumbara.

Bortfall

Bortfallsfel inträffar om objekten i bortfallet och de svarande skiljer sig åt (har en annan fördelning) avseende undersökningsvariablerna.

En jämförelse mellan de svarande och urvalet finns beskriven under rubriken Bortfall.

Aktualitet

2005-04-13

Med statistikens aktualitet menas hur väl statistiken beskriver ”nuläget”. Faktorer som påverkar denna komponent är exempelvis framställningstid, punktlighet och hur snabbt verkligheten förändras.

SCB:s arbete omfattade enkätkonstruktion, tryckning, utsändning, insamling, databearbetning samt framställning av datafil. Mest tidskrävande var datainsamlingen som pågick i nio veckor, vilket behövdes för att få ett bra svarsinflöde.

Datamaterialet levererades vecka 15, enligt överenskommelse.

Jämförbarhet och sam användbarhet

Jämförbarhet över tiden

Vid flera tidigare tillfällen har Göteborgs universitet genomfört liknande undersökningar. Frågorna skiljer sig något åt mellan undersökningarna. I de tidigare undersökningarna finns inte någon vikt-variabel för uppräkningsnivå.

För att jämförelser ska kunna göras med tidigare undersökningar bör dessa förses med vikter. I annat fall kan jämförande analyser och slutsatser bli felaktiga.

Tillgänglighet och förståelighet

SCB levererar en datafil tillsammans med den här rapporten där arbetet beskrivs. Spridningen regleras i en sekretessöverenskommelse.

Att tänka på vid användandet av datamaterialet

SPSS-filen levereras med variabelnamn och lablar.

Överhoppade kryssfrågor (partiellt bortfall) är missing values. Hoppfrågor har fyllt ut med ”9”.

Observera att de bifogade vikterna ska användas vid analys och resultatframställning. Om urvalet är stratifierat kan datamaterialet inte ses som ett urval av oberoende lika fördelade observationer, som ofta antas i traditionell statistisk teori. Det är därför viktigt att tänka på att de flesta förprogrammerade dataprogram inte klarar av att analysera datamaterialet ifrån en urvalsundersökning på ett korrekt sätt. Var noga med att kontrollera att de dataprogram som används för analys behandlar vikterna på ett riktigt sätt. Att ta fram frekvenser med beaktande av vikter klarar däremot de flesta statistikprogram.