#####  Mellanårsenkät 2010

 Teknisk rapport



## Inledning

Enheten för statistik om utbildning och arbete vid Statistiska centralbyrån (SCB) har under april - augusti 2010 genomfört en enkätunder­sökning på uppdrag av Kungliga tekniska Högskolan (KTH). Syftet var att undersöka hur studenterna upplever sin studiemiljö på KTH.

Undersökningen genomfördes som en postenkät med tre påminnelser. Populationen utgjordes av studerande på KTH som kommit ungefär halvvägs i sin utbildning. Totalt var det 1 009 personer som besvarade frågeblanketten, vilket var 52 procent av urvalet. Frågeblanketterna registrerades med skanning. Från det skannade materialet framställdes tabeller och diagram.

Sinisa Sauli har varit undersökningsledare, Johan Löfgren metodansvarig och Terttu Karlsson produktionsansvarig.

## Population och urval

Populationen, det vill säga de objekt som man vill kunna dra slutsatser om, utgjordes i undersökningen av studerande på KTH som kommit en bit in på sin utbildning. KTH skapade urvalsramen utifrån sitt register över studerande. Antalet individer i ramen var 1 937. Samtliga personer i ramen undersöktes (totalundersökning).

Aktuella adresser hämtades från Registret över totalbefolkningen (RTB).

## Sekretess och utlämnande

Med frågeblanketten följde ett informationsbrev där uppgiftslämnarna kunde läsa om undersökningens bakgrund, syfte och att undersökningen genomför­des i samarbete mellan KTH och SCB. Informationsbrevet informerade också om att uppgifter hämtades från KTH:s register, att svaren är skyddade av Personuppgiftslagen och Sekretesslagen, att insamlade uppgifter redovisas så att inga enskilda individers svar kan utläsas och att det var frivilligt att delta i undersökningen.

Behandlingen av personuppgifter i undersökningen har anmälts till SCB:s juridiska kansli.

## Variabler

Blanketten bestod av 22 numrerade frågor, flera av dem hade delfrågor vilket genererade totalt 69 frågor. Frågorna handlade bl.a. om studiemiljön på KTH, svårighetsgrad och tempo i utbildningen, studieuppehåll samt om negativt särbehandling vid KTH.

Jämfört med tidigare mellanårsenkäter har några frågor tillkommit, några har tagits bort och några har fått förändrade svarsalternativ. Många av frågorna är desamma som tidigare år.

Förutom de variabler som samlades in via frågeblanketten hämtades variabler från KTH:s register. Dessa är kön, ålder, skola och utbildningsprogram.

## Datainsamling

Frågeblanketterna sändes till urvalspersonerna via post. I informationsbrevet ombads de besvara frågorna och skicka tillbaka frågeblanketten till SCB. Det första utskicket genomfördes den 7 april, sedan skickades 3 påminnelser ut till dem som inte besvarat frågeblanketten, den 20 april, den 3 maj och den 19 maj. De tre påminnelserna innehöll ett påminnelsebrev och en ny enkät. Insamlingen avslutades den 16 juni 2010.

Aktuella adressuppgifter hämtades från Registret över totalbefolkningen.

**Tabell 1. Beskrivning av inflödet. Antal och andel.**

|  |  |  |
| --- | --- | --- |
|  | **Antal** | **Procent** |
| Efter första utskick | 412 | 21,3 |
| Efter enkätpåminnelse 1 | 313 | 16,2 |
| Efter enkätpåminnelse 2 | 165 | 8,5 |
| Efter enkätpåminnelse 3 | 119 | 6,1 |
| Totalt svar | 1 009 | 52,1 |
|  |  |  |
| Bortfall | 928 | 47,9 |
|  |  |  |
| Urval | 1 937 | 100,0 |

I 2004 och 2006 års undersökningar skickades två påminnelser. I båda dessa undersökningar blev svarsfrekvensen 56 procent. I 2008 års undersökning (tre påminnelser) blev svarsfrekvensen 57 procent.

## Bortfall

Bortfallet består dels av objektsbortfall, som innebär att frågeblanketten inte är besvarad alls, och av partiellt bortfall (uppgift saknas) som innebär att vissa frågor i 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. För att reducera bortfalls­skevheten har vikter beräknats med hjälp av kalibrering (se bilaga).

Objektsbortfallet i denna undersökning redovisas i tabell 2 nedan. Med ”ej avhörda” menas att ingen uppgift har lämnats om varför frågeblanketten inte är besvarad. Med ”avböjd medverkan” menas att SCB meddelats att uppgiftslämnaren inte vill medverka i undersökningen.

**Tabell 2. Beskrivning av objektsbortfall**

|  |  |
| --- | --- |
|  | Antal |
| Ej avhörda | 899 |
| Postreturer | 24 |
| Avböjd medverkan | 5 |
|  |  |
| Totalt | 928 |

I tabell 3 visas svarsandelen, som är komplementmängd till objektsbortfallet, i några utvalda grupper. I bilagan finns svarsandelar för ytterligare några grupper.

**Tabell 3. Svarsandel i olika grupper. Procent.**

|  |
| --- |
| Procent  |
| Kvinnor | 60,9 |
| Män | 48,4 |
|  |  |
| Civilingenjörer | 55,3 |
| Högskoleingenjörer | 47,2 |
| Arkitekter | 53,0 |
| Högskoleutbildning | 40,2 |
| Kandidatprogram | 46,6 |
|  |  |

Partiellt bortfall kan bero på att en fråga är svår att förstå, är känslig, att uppgiftslämnaren glömmer att besvara frågan eller att instruktionerna vid hoppfrågor misstolkas. Det partiella bortfallet anges som ”uppgift saknas” i tabellerna och diagram.

## Databeredning

Dataregistreringen genomfördes av enkätenheten vid SCB. Fråge­blanketterna skannades med programvaran ReadSoft DOCUMENTS for FORMS. Kontroller genomfördes under och efter registreringen. Då kontrollerades bl.a. att endast valida värden förekom i materialet.

Svarsdatafilen kompletterades med variabler från KTH samt med vikter för uppräkning till populationsnivå.

## Viktberäkning och estimation

Vikter har tagits fram 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.

Vikterna i denna undersökning kan beskrivas med formeln:



där wk = den totala vikten för objekt k

 dk=designvikt

 vk= kalibreringsvikt

Designvikten är den del av vikten som beror på urvalsdesignen. I denna undersökning är designvikten lika med 1 för alla i urvalet, eftersom det var en totalundersökning. I praktiken är det alltså bara kalibreringsvikten (vk) som används.

Vid bortfall kan det vara så att vissa grupper av urvalet svarar i större utsträckning än övriga, t.ex. kan kvinnor svara i högre grad än män. Om de grupper som svarat i högre grad har en annan fördelning på undersöknings­variablerna än övriga kan detta ha en snedvridande effekt på resultatet. För att kompensera för detta har kalibreringsvikter använts.

För mer utförlig beskrivning av kalibre­ringsestimatorn se Lundström och Särndal (2001): *Estimation in the Presence of Nonresponse and Frame Imperfections, Statistics Sweden.*

För beräkning av skattningen av totaler används följande formel:



där wk = den totala vikten för objekt k

 *yk =* variabelvärde för objekt k

summering sker av de svarande (r)

Och för beräkning av skattningen av medelvärden används följande formel:



där wk = den totala vikten för objekt k

*yk =* variabelvärde för objekt k

Kalibreringen och de tabeller som presenteras i denna rapport har tagits fram enligt de formler som presenterats ovan med hjälp av SAS och variansprogrammet ETOS.

## Tabellerna

Tabellerna har räknats upp till populationsnivå, vilket innebär att resultatet avser hela populationen och inte bara de svarande. Antalsuppgifterna är således skattningar av antal personer i populationen med den aktuella egenskapen. Att det uppräknade antalet individer är olika för olika frågor beror på hoppinstruktioner då olika delpopulationer avses. För de följdfrågor som föregås av hoppinstruktioner har selekteringar genomförts vid tabellframställandet.

Om en följdfråga är besvarad men dess styrfråga inte är besvarad och svaret på följdfrågan göt att ett visst svarsalternativ på styrfrågan är självklar så har hänsyn till detta tagits vid tabellframställningen.

I andelstabellerna är värdena avrundade till närmaste heltal och värden under 0,5 är därmed avrundade till noll.

Tabellerna har sekretessgranskats så att celler som bygger på färre än två svar har prickats. Mycket osäkra skattningar, där konfidensintervallet är större än andelsskattningen, har också prickats.

De tabeller som presenteras i denna rapport har tagits fram enligt de formler som presenteras i avsnitt ”Viktberäkning och estimation” med hjälp av SAS och variansprogrammet ETOS.

Högskoleingenjörsutbildningen datateknik och ekonom/hälsa ingår i tabellerna men baseras endast på tre svarande av ett sammanlagt urval på fem personer från de två programmen.

## 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 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

Materialet lämpar sig bäst att redovisas som totaler eller procentuella andelar för hela målpopulationen och fördelat på olika redovisningsgrupper.

## Redovisningsgrupper

Ofta redovisas statistiken inte bara för hela populationen utan också för delgrupper (redovisningsgrupper). Redovisningsgrupperna i den här undersökningen har avgränsats med hjälp av registervariabler (kön, ålder, program, skola) eller enkätfrågor (modersmål (fråga 2) och föräldrars högskole- eller universitetsutbildning (fråga 3)).

## Referenstider

Referensperioden (mätperioden) i undersökningen varierar mellan olika frågor, t.ex. vårterminen 2010 och ”de senaste 6 månaderna”. För objekten (studenterna) är referensperioden vårterminen 2010.

## Tillförlitlighet

## 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. KTH:s register bedöms vara av god kvalitet.

## Urval

Det gjordes en totalundersökning, vilket inte ger något urvalsfel. SCB har ändå valt att presentera osäkerhetsmått i form av konfidensintervall, för att visa att det finns osäkerhet i skattningarna. I de konfidensintervall (de urvalsosäkerheter) som redovisas utgörs urvalet av de svarande – som inte är slumpmässigt utvalda, vilket i teorin är en förutsättning.

## 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. För att försöka minimera mätfelen har kontroller genomförts, t.ex. kontroll av konsistens mellan olika frågor. Mätfelen för registervariabler som hämtats från KTH är okänt för SCB, men har antagits vara små.

## Bearbetning

Vid den manuella och maskinella bearbetningen av datamaterialet kan så kallade bearbetningsfel uppstå. Exempel på bearbetningsfel är registrerings­fel och kodningsfel. Risken för registreringsfel är större för öppna frågor än för dem med fasta svarsalternativ. Kontroller, s.k. verifiering, genomfördes vid dataregistreringen för att minimera felaktigt tolkade svar. I den här undersökningen bedöms registreringsfel vara försumbara. Kodning förekom inte.

## Bortfall

Bortfallsfel inträffar om objekten i bortfallet och de svarande skiljer sig åt (har en annan fördelning) avseende undersökningsvariablerna. En beskrivning av bortfallet finns i den bortfallsanalys som gjorts i samband med valet av hjälpinformation till kalibreringen, se bilaga.

## Aktualitet

SCB:s arbete omfattade enkätkonstruktion, tryckning, databearbetning samt framställning av tabeller och teknisk rapport. Materialet levererades i vecka 34, år 2010, enligt överenskommelse.

## Jämförbarhet och samanvändbarhet

## Jämförbarhet över tiden

Mellanårsenkäten 2004, 2005, 2006 och 2008 genomfördes av SCB. Tidigare år har den genomförts av annan producent.

Många frågor i frågeformuläret 2010 fanns med i tidigare års formulär. Samma skattningsmetod användes som i 2008 års enkät. Genom kalibre­ringsestimatorn har mer hjälpinformation har använts. SCB:s bedömning är att de fem senaste årgångarna kan jämföras med varandra.

## Tillgänglighet och förståelighet

SCB levererar tabeller och diagram tillsammans med den här rapporten där arbetet beskrivs. Användning och spridning regleras i SCB:s allmänna avtalsvillkor.

## Användning av tabellerna

För att kunna läsa tabellerna måste (oftast) inställningar i Excel ändras.

1. Excel 2003
För att tabellerna skall fungera måste säkerhetsnivån vara medel.
Detta kontrolleras genom att gå in under "Verktyg - Makro - Säkerhet - Säkerhetsnivåfliken"
När Excel-filen öppnas skall även makrot aktiveras.
2. Excel 2007
Välj Excel-alternativ – Säkerhetscenter – Inställningar för säkerhetscenter – Makroinställningar
Markera rutorna ”Aktivera alla makron” och ”Åtkomst till objektmodell för VBA-projekt betrodd”.

För att kunna se diagrammen måste man i tabellfliken välja ”Andel utan konfidensintervall” eller ”Antal utan konfidensintervall”.

Då statistiken redovisas på små grupper (olika program) och bortfallet var ganska högt bör viss försiktighet användas vid jämförelser mellan olika grupper. Små procentuella skillnader ligger ofta inom den s.k. felmarginalen.

## Övrigt

Bilaga: Kalibreringsrapport