

# STATISTIKENS FRAMSTÄLLNING

## Doktorsexaminerades villkor under forskarutbildningen och i arbetslivet

### Ämnesområde

Utbildning och forskning

### Statistikområde

Befolkningens utbildning

### Produktkod

UF0549

### Referenstid

Tiden efter doktorsexaminationen som skett under tiden 2007-2017

Tiden under utbildningen. Antagningsdatumerna för urvalet spänner över tiden 1975-2016.

Frågor om aktuella förhållanden avser mättillfället som var 2018-2019

### Kontaktuppgifter

<b>Statistikansvarig myndighet</b>	Statistiska centralbyrån
<b>Kontaktinformation</b>	Nikolaus Koutakis
<b>E-post</b>	<a href="mailto:nikolaus.koutakis@scb.se">nikolaus.koutakis@scb.se</a>
<b>Telefon</b>	010-479 6665

## Innehåll

1	Statistikens sammanhang.....	3
2	Undersökningsdesign .....	3
2.1	Målstorheter .....	3
2.2	Ramförfarande .....	3
2.3	Förfaranden för urval och uteslutning .....	3
2.3.1	Urvalsförfarande.....	3
2.3.2	Uteslutning från insamling (cut-off) .....	4
2.4	Insamlingsförfarande.....	4
2.4.1	Datainsamlingsmetoder .....	4
2.4.2	Mätning.....	4
2.4.3	Bortfallsuppföljning.....	5
2.5	Bearbetningar.....	5
2.6	Granskning.....	5
2.6.1	Granskning under insamlingen .....	5
2.6.2	Granskning av mikrodata .....	5
2.6.3	Granskning av makrodata.....	5
2.6.4	Granskning av redovisning .....	6
2.7	Skattningsförfarande .....	6
2.7.1	Principer och antaganden .....	6
2.7.2	Skattningsförfarande för målstorheter.....	8
2.7.3	Skattningsförfarande för tillförlitlighet.....	8
2.7.4	Röjandekontroll .....	8
3	Genomförande .....	9
3.1	Kvantitativ information.....	9
3.2	Avvikelser från undersökningsdesignen .....	9

## 1 Statistikens sammanhang

Undersökningen doktorsexaminerades beskriver sysselsättningen bland de som tagit ut en doktorsexamen mellan 2007 och 2017. Undersökningen redovisar statistik inom statistikområdet Befolkningens utbildning som fördjupar den officiella statistiken på området. Statistiken handlar om befolkningens utbildning, olika utbildningsgruppers övergång till och ställning på arbetsmarknaden samt nya frågeställningar på området som är aktuella i samhällsdebatten. Rapporterna publiceras främst i rapportserien Temarapporter. En liknande undersökning med samma målpopulation genomfördes under 2003, Forskarexaminerades arbetsmarknad, [www.scb.se/UF0522](http://www.scb.se/UF0522).

I detta dokument beskrivs översiktligt design och genomförande av Befolkningsstatistiken. Dokumentet kompletterar kvalitetsdeklarationen för Befolkningsstatistiken som finns tillgänglig på [www.scb.se/UF0549](http://www.scb.se/UF0549) under rubriken *Dokumentation*.

## 2 Undersökningsdesign

### 2.1 Målstorheter

De viktigaste målstorheterna i undersökningen, som alla baseras på urvalspersonernas egna bedömning är medelvärden för sexaminerade från olika forskningsämnesområden gällande:

- Huvudsaklig sysselsättning
- Överensstämmelse mellan utbildning och arbete
- Arbetets krav på färdigheter.
- Nöjdhet med utbildningens färdighetsträning.
- Effarenheter av handledares engagemang.
- Förekomst av kränkningar och sexuella trakasserier.

### 2.2 Ramförfarande

Rampopulationen för undersökningen är de som avslutat en doktorsexamen mellan 2007 och 2017 enligt SCB:s Universitets- och högskoleregister. Personerna var max 67 år vid slutet av 2018.

Observationsobjekten är de avgångna personer som har rapporterats in till Universitets- och högskoleregistret.

### 2.3 Förfaranden för urval och uteslutning

#### 2.3.1 Urvalsförfarande

Undersökningen är en stratifierad urvalsundersökning. Universitets- och högskoleregistret användes för att ta fram den aktuella rampopulationen som omfattade 22 890 personer efter de

avgränsningar som beskrivs i avsnitt 2.2 *Ramförfarande*. Ramen stratifierades på 12 stratum innan ett urval på 10.000 personer drogs. Stratifieringen skedde efter kön och forskningsutbildningsämne.

### 2.3.2 Uteslutning från insamling (cut-off)

Ingen del av målpopulationen utesluts från direktinsamlingen.

## 2.4 Insamlingsförfarande

### 2.4.1 Datainsamlingsmetoder

Undersökningen genomfördes som en post- och webbenkät under tiden 2018-10-26 – 2019-01-14. En så kallad webboffensiv strategi användes med totalt fyra utskick:

1. Informationsbrev med inloggningsuppgifter 26 oktober
2. Tack- och påminnelsekort (ToP) 9 november
3. Pappersenkät med inloggningsuppgifter 23 november
4. Påminnelsebrev med ny enkät 7 december

Förutom information från urvalspersoner så har registeruppgifter inhämtats, detta för att minimera uppgiftslämnarbördan men även för att minska effekten av minnesfel eftersom undersökningen för vissa urvalsperson rör förhållanden tio år tillbaka i tiden.

Följande interna register på SCB har använts:

- Registret över totalbefolkningen (RTB)
- Universitets- och högskoleregistret – doktorander och examina i högskoleutbildningen på forskarnivå

### 2.4.2 Mätning

Blanketten bestod av 43 numrerade frågor, av dessa hade åtta flera delfrågor totalt innehöll instrumentet 61 frågor.

Instrumentet hade frågor med fasta svarsalternativ samt några frågor där respondenten själv fick fylla i procentsatser, altså endast siffror. Många svarsalternativ hade även ett öppet alternativ där respondenten själv kan fylla i svar som saknas i de övriga svarsalternativen. Mätteknikern genomförde fem kognitiva intervjuer med personer som tillhör populationen för att förebygga tolkningsfel och missupfattningar.

I webbenkäten fanns inbyggda kontroller så att hoppsteg och filterfrågor kunde styras automatiskt utifrån respondentens svar. I den tryckta postenkäten anges istället anvisningar för vilka frågor som eventuellt ska hoppas över. I webbenkäten fanns även kontroller för att undvika så kallade dubbelmarkeringar där respondenten

markerat flera svar trots att frågan bara ska besvaras med ett alternativ.

### **2.4.3 Bortfallsuppföljning**

Bortfall definieras som ej inkomna svar från urvalspersoner under fältarbetstiden som omfattar tiden 2018-10-26 – 2019-01-14. För att minimera bortfallet har påminnelser samt enkäter skickats ut, förutom initialt informationsbrev med inloggningsuppgifter har tre olika påminnelser skickats via post. Påminnelser har endast skickats till dem som tillhör urvalet och vars svar inte registrerats.

## **2.5 Bearbetningar**

Nedan beskrivs härledningen av några bakgrundsvariabler som hämtats från Registret över totalbefolkningen (RTB) och Utbildningsregistret:

Svensk/ utländsk bakgrund: Elever med svensk bakgrund. Födda i Sverige med minst en förälder som också är född i Sverige.

Elever med utländsk bakgrund: Elever födda utomlands och elever födda i Sverige med båda föräldrarna födda utomlands.

Föräldrarnas utbildningsnivå

Högst gymnasial utbildning: Ingen av föräldrarna har uppgift om en högre utbildning än en gymnasial utbildning

Eftergymnasial utbildning: Minst en föräldrars högsta utbildning är eftergymnasial utbildning

## **2.6 Granskning**

### **2.6.1 Granskning under insamlingen**

Logiska kontroller har byggts in i webbformuläret så att felinmatningar av logisk karaktär ommöjliggjorts. Inscannade pappersenkäter saknar denna kontroll men granskas manuellt i förekommande fall.

### **2.6.2 Granskning av mikrodata**

Så kallade dubbelmarkeringar på frågor som endast tillåter ett svarsalternativ har granskats och kodats enligt särskilt formulär.

### **2.6.3 Granskning av makrodata**

Granskning av makrodata har gjorts med s.k. rimlighetsbedömningar samt i förekommande fall jämförelser med Forskarexaminerades arbetsmarknad.

#### 2.6.4 Granskning av redovisning

Inför publicering granskas samtliga delar av materialet efter laddning i SCB:s webbpubliceringsverktyg och efter laddning i statistikdatabasen.

SCB kontrollerar att alla tabeller och diagram finns med och att ingen av dem är tom eller innehåller obegripliga värden, till exempel interna koder. Rubriker och förklaringar till tabeller och diagram granskas så att de är korrekta. Man kontrollerar också att överensstämmelse råder där samma siffervärde eller text används på flera ställen.

Tabeller och figurer dubbekontrolleras. Kollegial granskning med seminarium av fem kolegor sker i två steg innan publicering.

### 2.7 Skattningsförfarande

#### 2.7.1 Principer och antaganden

För varje svarande person har en vikt beräknats. Syftet med detta är att kunna redovisa resultat för hela populationen och inte bara för de svarande. Vikten kallas därför även för uppräkningsstal. Detta förfarande, som utnyttjar hjälpinformation från registervariabler, används främst för att kompensera för den skevhet som kan förmodas uppstå till följd av de bortfalls- och täckningsproblem som finns, men även för att reducera den slumpmässiga osäkerhet: variansen som skattningsförfarandet innebär.

Vikterna har beräknats utifrån urvalsdesignen samt antaganden om objektbortfall och täckningsfel.

De hjälpvariabler som används i skattningsförfarandet har valts utifrån kriterier om hur de samvarierar med svarsbenägenhet och viktiga undersökningsvariabler. Nedan beskrivs hjälpvariablerna och de kategorier som används.

Vikterna beskrivs med formeln:

$$w_k = d_k \times v_k$$

där

$w_k$  = den totala vikten för objekt k

$d_k$  = designvikt

$v_k$  = kalibreringsvikt

Designvikten är den del av vikten som beror på urvalsdesignen. Kalibreringsvikten justerar för bortfall och täckningsbrister.

Hjälpvariabel	Kategorier
<b>Yrke</b>	Yrke som kräver högskoleutb. Yrke som ej kräver högskoleutb. Yrke okänt
<b>Inkomst</b>	Mindre än 100 000 100 000 -299 999 300 000 eller mer Anhöriga Flyktingskäl Övriga skäl
<b>Födelseland</b>	Norden EU28 (utom Norden) Övriga Europa Afrika Asien Övriga
<b>Kön</b>	Man Kvinna

Hjälpvariablerna *Grund för bosättning*, *Källa för uppgift om högsta utbildning*, *Födelseland* och *Invandringsår* är endast aktuella för utrikes födda personer. För inrikes födda har vissa utbildningsinriktningar grupperats ytterligare.

Utifrån de valda hjälpvariablerna beräknas därefter kalibrerade vikter. Vikterna har egenskapen att de vid skattning av populationstotalen, för var och en av de hjälpvariabler som används för att fastställa de kalibrerade vikterna, resulterar i en skattning som överens-stämmer exakt med den registrertotal som används för att fastställa de kalibrerade vikterna. För utrikes födda har vikterna justerats i avseende att kompensera för den övertäckning som urvalsramen innehåller. Detta innebär att de skattade populationstotalerna är lägre än motsvarande registrertotaler. Ett

antagande görs att övertäckningen i urvalsramen är lika stor som den övertäckning som observerats under insamlingen. Ingen justering för undertäckning har gjorts för utrikes födda.

### 2.7.2 Skattningsförfarande för målstorheter

För beräkning av skattningen av totaler används följande formel:

$$\hat{y} = \sum_r w_k y_k$$

där

$y_k$  = variabelvärde för objekt k

summering sker av de svarande (r)

För beräkning av skattningen av medelvärden används följande formel:

$$\hat{y} = \frac{\sum_r w_k y_k}{\sum_r w_k}$$

### 2.7.3 Skattningsförfarande för tillförlitlighet

Konfidensintervallen är beräknade enligt en metod som med 95 procents säkerhet täcker in det sanna värdet och beräknas enligt formeln

punktskattning  $\pm 1,96$  \* (skattning av estimatorns medelfel)

där såväl punkt- som medelfelsskattning beräknas med hjälp av ett av SCB egenutvecklat program, ETOS2.

### 2.7.4 Röjandekontroll

Ingen röjandekontroll har gjorts utifrån bedömningen att inga skyddsmetoder behövs. Undersökningen är en urvalsundersökning och statistiken som publiceras är inte på en så pass detaljerad nivå att det bedömts föreligga risk för röjande.

Vid publicering redovisas dock inte skattningar på enskilda frågor med färre än 30 svarande och inga enskilda cellvärden med färre tre svarande redovisas. Alltför osäkra skattningar redovisas inte heller.



### **3 Genomförande**

#### **3.1 Kvantitativ information**

Rampopulationen omfattade 22 890 personer. Urvalet bestod av 9 879 personer. Bortfallet uppgick till 40,14 procent. Övertäckningen i ramen bedtod av 51 personer. Den avgränsade populationen omfattar 22 890 personer.

#### **3.2 Avvikelser från undersökningsdesignen**

Inga avvikelser har gjorts.