

## Att utforma och förbättra en statistisk undersökning

### INNEHÅLL

|     |   |    |
|-----|---|----|
| 1   | Introduktion .....  | 3  |
| 1.1 | Utformning är en investering.....                             | 3  |
| 1.2 | Syfte och målgrupper .....                                    | 3  |
| 1.3 | Utformning och några käpphästar .....                         | 4  |
| 1.4 | Processkartan och processmodeller .....                       | 6  |
| 1.5 | Endast en vägledning .....                                    | 7  |
| 1.6 | Struktur och läsanvisning .....                               | 7  |
| 2   | Grundläggande begrepp och termer .....                        | 9  |
| 2.1 | Begreppsbildning samt intresse, mål och observation .....     | 9  |
| 2.2 | Statistikens byggstenar och struktur .....                    | 11 |
| 2.3 | Skattning och statistisk inferens.....                        | 12 |
| 2.4 | Statistisk undersökning, process, produkt, system.....        | 16 |
| 2.5 | Statistikens kvalitet och dess komponenter .....              | 19 |
| 2.6 | Tillförlitlighet och osäkerhetskällor .....                   | 21 |
| 2.7 | Optimering, risker, kvalitetssäkring, kvalitetskontroll ..... | 23 |
| 2.8 | Metadata och processdata.....                                 | 26 |
| 3   | Väsentliga steg och osäkerhetskällor .....                    | 28 |
| 3.1 | En illustration och dess betydelse.....                       | 28 |
| 3.2 | Begrepp undersöks genom mätning av variabler .....            | 30 |
| 3.3 | Populationer av objekt som observeras och som skattas .....   | 32 |
| 3.4 | Sammanföra data – harmonisering m.m. ....                     | 35 |
| 3.5 | Osäkerhetskällor – några kommentarer .....                    | 35 |
| 3.6 | Betydelsen av definitioner och metoder .....                  | 36 |



|     |   |    |
|-----|---|----|
| 4   | Utforma och utvärdera – översikt .....                        | 38 |
| 4.1 | Inledning med delprocesser och kvalitetskomponenter .....     | 38 |
| 4.2 | Fastställ behov .....   | 39 |
| 4.3 | Designa och planera .....                                     | 41 |
| 4.4 | Skapa och testa .....   | 44 |
| 4.5 | Utvärdera och återkoppla.....                                 | 45 |
| 5   | Principer, resonemang och stöd för design.....                | 47 |
| 5.1 | Principer vid statistisk inferens.....                        | 47 |
| 5.2 | Ansatser för urval och estimation.....                        | 49 |
| 5.3 | Några resonemang för ramtäckning, urval och bortfall.....     | 52 |
| 5.4 | Kombinerade metoder vid direktinsamling.....                  | 54 |
| 5.5 | Svarsprocesser .....  | 55 |
| 5.6 | Pilotstudier, bl.a. experiment .....                          | 56 |
| 5.7 | Statistikproduktionsprocessen som helhet.....                 | 58 |
| 6   | Exempel på situationer och val.....                           | 62 |
| 6.1 | En ny eller annorlunda produktionsomgång .....                | 62 |
| 6.2 | Precisionskrav och prioriteringar .....                       | 65 |
| 6.3 | Design av hantering av outliers (i mikrodata).....            | 66 |
| 6.4 | Några situationer med modellbaserade skattningar .....        | 67 |
| 6.5 | Design av en löpande undersökning .....                       | 70 |
| 6.6 | Preliminär och slutlig statistik – löpande revideringar ..... | 71 |
| 6.7 | Ändringar av metoder och definitioner – tidsseriebrott .....  | 73 |
| 6.8 | Statistiksystem – påverkan på undersökningsdesign.....        | 75 |
| 7   | ”Modern” statistikproduktion.....                             | 78 |
| 7.1 | Statistikproduktion i en föränderlig värld .....              | 78 |
| 7.2 | Befintliga data och datalager.....                            | 79 |
| 7.3 | Standardisering .....   | 79 |
| 7.4 | Arkitektur och infrastruktur.....                             | 80 |
| 7.5 | Förbättringsarbete.....                                       | 81 |
| 8   | Referenser .....  | 84 |

# 1 Introduktion

## 1.1 Utformning är en investering

Att tänka efter före är oftast klokt. I statistiska undersökningar går det att vinna mycket på tydliga dialoger om t.ex. användarbehov, tänkta tabeller, variabler och datainsamlingsmetod(er): registerdata (i första hand) eller direktinsamling (väl vald). En genomtänkt utformning kan ha stor betydelse för kvalitet och kostnader. Det kan också finnas mycket att förbättra successivt i löpande undersökningar.

Statistikproduktion med insamling (från annan statistik eller tillgängliga register; eller direkt från uppgiftslämnare), granskning, skattning med mera löper smidigare och säkrare med ett gott förarbete än utan.

## 1.2 Syfte och målgrupper

Titeln "Att utforma och förbättra en statistisk undersökning" knyter an till processområdet *Utforma och utvärdera*. Avsikten är också att betona ett ständigt förbättringsarbete. Det sistnämnda avser inte bara löpande undersökningar utan även verksamhet med engångsundersökningar, oftast på uppdrag. Det är viktigt att notera att statistikproduktion och statistisk undersökning här (och även i många andra sammanhang) inkluderar t.ex. registerbaserad statistikproduktion, registerframställning och räkenskaper (som national- och miljöräkenskaper).

Denna vägledning ingår i Verksamhetsstödet och är ett komplement till dess korta instruktioner. Den vänder sig till *team* som arbetar löpande med en statistisk undersökning på ett sådant sätt att förbättringsarbete ingår med hjälp av olika kompetenser t.ex. ämne, datainsamling, metod och IT. Detta kallas *produktionsteam* i SCB:s modell för produktionsstöd. (Denna modell introduceras under 2016-2017 och följer den generella modellen pm3.) Utformning och förbättring kräver många kompetenser och sker därför lämpligen i ett lagarbete. Vägledningens avsnitt har något olika karaktär och riktar sig därmed till sinsemellan något olika läsare. Liksom i Verksamhetsstödet utgör innehållet något av en bruttolista. Var och en får välja det som är relevant för den egna situationen, och vid behov eller önskemål söka sig vidare genom referenser; texten ger en rad sådana.

Den första versionen av vägledningen färdigställdes i november 2015 till Verksamhetsstödet. En uppdatering görs nu till 2017 då en föreskrift om kvalitet för den officiella statistiken träder i kraft. Uppdateringen begränsas väsentligen till det nya kvalitetsbegreppet och dess effekter. Några fler aktualiseringar görs, t.ex. avseende texter i Verksamhetsstödet och för att förvaltningsmodellen FMOD tas ur bruk.

### 1.3 Utformning och några käpphästar

Det här avsnittet beskriver synsätt och aktiviteter i korta drag som en introduktion till kommande kapitel.

#### 1.3.1 Designarbete i korthet

Utan att ha definierat alla begrepp ges en introducerande beskrivning av designarbete genom punktlistan nedan. Betydelseerna av orden kund, användare, statistikanvändare och beställare är likartade. Ordet kund används här för att betona att det är en betalande kund, som därmed har stora möjligheter att påverka.

Design innebär i korthet att utarbeta hur en kommande omgång av en statistisk undersökning ska genomföras; en undersökning som görs en eller flera gånger.

- Resultatet är statistik eller observationsregister (eller båda).
- SCB-kompetenser av olika slag används under utformningen.
- Utformningsarbetet sker i samarbete eller samförstånd med kund, användare eller beställare.
- I detta arbete bevakas nedanstående krav och aspekter.
  - Kvaliteten blir tillräcklig för den avsedda användningen (en eller flera).
  - Genomförandet sker inom budget och är kostnadseffektivt (långsiktigt).
  - Hänsyn tas till uppgiftslämnare (börda). Det underlättar för uppgiftslämnarna, och det påverkar normalt de två föregående punkterna: kvaliteten blir högre och kostnaderna lägre.

Punkten om kvalitet uppvisar krav på undersökningens resultat, som oftast är statistik men kan vara observationsregister. Det är inte så att statistiken passar all användning, utan det finns ett (eller flera) uttalande ändamål för vilket kvaliteten är tillräcklig. Detta ställer krav också på en tillräcklig budget för att kunna uppfylla ändamålets kvalitetskrav. Om budgeten skulle vara otillräcklig får ändamålet omformuleras (eller budgeten höjas). Idealt ska den totala kostnaden minimeras givet kvalitetskraven, men formuleringen är försiktigare här. För en löpande undersökning behöver investeringar göras, och en långsiktig plan ska finnas utöver målmedvetenhet för enskilda omgångar. En statistisk undersökning ingår normalt i ett sammanhang där flera undersökningar och statistikprodukter tillsammans belyser ett ämnesområde eller liknande.

Ansvar beskrivs bl.a. i SCB:s arbetsordning. I denna vägledning förekommer ansvarsfrågor främst i avsnitt 7.5. Även där handlar det mer om *vad* än om *vem*.

### 1.3.2 Fem käpphästar för utformning

Som ett annat led i detta inledande kapitel följer nedan, i punktform, fem "käpphästar" som på olika sätt återkommer i den följande texten, dock utan att nämnas som sådana. De ska alla genomsyra resonemang då undersökningar utformas och förbättras.

- A   Precisera: Intresse, mål, observation.  
(Se avsnitt 2.1 med flera.)
  
- B   Precisera den statistiska inferens som görs: Vad är det som skattas, vad (vilka data) ingår i skattningarna, och vilka antaganden görs?  
(Se avsnitt 2.3 och många flera.)
  
- C   Tillförlitlighet: Arbeta med statistikens tillförlitlighet och bedöma dess storlek med hänsyn till alla osäkerhetskällor. Det är ett synsätt som ibland kallas *total survey error*. Det betyder här inte att andra kvalitetskomponenter är mindre viktiga.  
(Se avsnitt 2.6 med flera.)
  
- D   Förbättringsarbete:
  - Ha en plan för iakttagelser under produktionen och för särskilda studier samt för införande av ändringar och omläggningar.
  - Beakta tre K: Kvalitet, Kostnader, Kompetens.(Se avsnitt 2.7, 7.5 med flera.)
  
- E   Systemtänkande:
  - Statistikens kvalitet: design för sam användbarhet (definitioner, metoder) där register, ram, data, statistik och vidarebearbetad statistik (vilket inkluderar registerbaserad statistik) ingår; liksom ämne-metod-IT.
  - Arkitektur: IT, metod, Verksamhetsstödet, etc.(Se avsnitt 2.4.4, 6.8, 7.4 med flera.)

Fokus flyttas generellt allt mer från enskilda produkter till statistiksystem och från enskilda skräddarsydda verktyg mot en gemensam kvalitetssäkrad produktionsmiljö med gemensamma verktyg. Både metodologiskt och IT-mässigt är det motiverat att tala om och etablera en gemensam arkitektur för statistikproduktion. Förberedelser är ofta särskilt viktiga när något görs första gången, som när en löpande statistisk undersökning vidtar ändringar, t.ex. i sin datainsamlingsmetod eller i sin verktygsanvändning.

#### 1.4 Processkartan och processmodeller

SCB har en generell processkarta för all sin verksamhet, och inom den finns processen "Framställa och kommunicera statistisk information". Den beskrivningen, som togs fram först, kallas på SCB ofta statistikproduktionsprocessen eller processkartan. Grunden lades 2007 med hjälp av den processkarta som Nya Zeelands centralbyrå hade tagit fram 2006; troligen den första processkartan för statistikproduktion. Internationellt finns nu *Generic Statistical Business Process Model* (GSBPM); se UNECE (2013) för version 5.

Processkartan omfattar alla statistiska undersökningar som resulterar i statistik eller register, exempelvis: sammanställningar, registerbaserade och direktinsamlade, löpande och en enstaka gång, genom anslag och uppdrag, hela statistikproduktionsprocessen eller en del.

Verksamhetsstödet bygger på processkartan och beskriver vad som ska göras i den ordning som statistikproduktionsprocessen anger. Som alla modeller är processkartan en förenklad beskrivning. Tidsmässigt ser flödet något annorlunda ut, som följande två exempel illustrerar. Om problem uppstår, kan det finnas behov av att gå tillbaka. Delar av designprocessen görs, på grov nivå, redan i arbetet med att fastställa behov. Detta mer tidsenliga flöde synliggörs i en processmodellering. Denna finns på Inblick på en grov nivå (med ett modellbibliotek och beskrivningar). Den beskriver ett börnläge för huvuddelen av de statistiska undersökningar som SCB gör.

Det finns mycket som påverkar hur undersökningar utformas och genomförs. Det europeiska statistiska systemet, ESS, har stort inflytande på SCB:s statistik och statistikproduktion. Se t.ex. Eurostat (2011, 2012) och ROS (2013) för allmänna riktlinjer och kvalitets-säkring; riktlinjer för europeisk statistik som ofta kallas *Code of Practice* även på svenska. Det finns många förordningar för statistiska undersökningar och produkter eller motsvarande.

Processområdet *Utforma och utvärdera* omfattar fyra av de åtta processerna på den första nivån i statistikproduktionsprocessen: 1 *Fastställ behov*, 2 *Designa och planera*, 3 *Skapa och testa* samt 8 *Utvärdera och återkoppla*. Det handlar således om design i en vid mening och med ett cykliskt förfarande: utformning, utvärdering och återkoppling, utformning osv. (Processen *Stöd och infrastruktur* betecknas ofta, något oegentligt, som en nionde process på den första nivån.)

Även om en del ibland framställs som nytt, t.ex. processkartan, har mycket i synsätten funnits länge. Särndal, Swensson och Wretman (1992, sidan 20) återger en figur från 1970-talet av Tore Dalenius över processen *total survey design* med bl.a. behov, kravanalys, design, genomförande och användningar; därtill kontroll och utvärdering (se vidare avsnitt 5.7.1).

### 1.5 Endast en vägledning

Förut fanns "Design av statistiska undersökningar. En vägledning" som skrevs 2008–2009. Kopplingen från den vägledningen till processkartan och Verksamhetsstödet var svag, vilket har varit ett av flera skäl till den förnyelse som nu har gjorts. Den tidigare vägledningen har förstås varit till nytta på många ställen, men dispositionen och det mesta av texten är nya.

Verksamhetsstödet utgår från SCB:s processkarta för statistikproduktion. Det vänder sig för varje delprocess främst till utförarna av den delprocessen. Det beskriver gemensamma metoder, verktyg och rutiner, och det innehåller även en del förklarande bakgrundsdokument. Detta är ett sådant bakgrundsdokument, i första hand för *Designa och planera* men även för de övriga processerna i förbättringshjulet med utformning, utvärdering och återkoppling, utformning, osv.

Även om ordet vägledning inte längre finns i rubriken används benämningen nedan. Många önskemål har uttalats om att få en starkare handledning i hur arbete med utformning går till för nya undersökningar, översyner och löpande förbättringar. Dock finns det som regel inga enkla lösningar och svar på frågor om design av statistiska undersökningar. Det är lätt att peka på många ställningstaganden som behöver göras, t.ex. val av metoder och verktyg samt resursallokering. En mängd påverkande och samverkande faktorer behöver beaktas för att prioritera mellan eller väga samman olika önskemål. Det är svårt eller omöjligt att uttala generellt vad sådana ställningstaganden leder till.

### 1.6 Struktur och läsanvisning

Vägledningen är strukturerad med sex kapitel efter denna inledning, vilket nedanstående uppställning visar. Uppställningen är en kort sammanfattning och den indikerar för vilka läsare de olika kapitlen främst är tänkta. Varje kapitel är uppdelat i avsnitt.

#### 2 Grundläggande begrepp och termer

De valda begreppen och termerna används i de följande kapitlen utan förklaring. De är nyttiga att behärska för de flesta som arbetar med statistikproduktion. Texten innehåller mer än definitioner och är därmed även introducerande till fortsättningen.

Kapitlet är förklarande, och det vänder sig till den som är ny eller obekant med dessa begrepp och termer. Det gör inte anspråk på att vara fullständigt eller detaljerat.

#### 3 Väsentliga steg och osäkerhetskällor

Utgångspunkten för denna beskrivning, som på engelska kallas *survey life cycle*, är en bild med byggstenar i kon-

strukturen av statistik (och register). Bilden tjänar som utgångspunkt för en något metodologisk beskrivning med processteg och osäkerhetskällor.

Kapitlet vänder sig främst till metodintresserade.

#### 4 **Utforma och utvärdera – översikt**

Kapitlet är något instruerande till sin karaktär, men jämfört med Verksamhetsstödet är det mer resonerande, allmänt belysande och översiktligt. Påverkan på kvalitetskomponenter ingår. Det beskriver de fyra processerna i *Utforma och utvärdera* med tonvikt på designprocessen.

Kapitlet vänder sig till den som vill ha en översikt, t.ex. som ingång till Verksamhetsstödet.

#### 5 **Principer, resonemang och stöd för design**

Kapitlet ger exempel på teori och stöd som finns för utformning av några delprocesser. Det beskriver exempelvis lite om inbäddade experiment och nyttan med sådana.

Kapitlet vänder sig till den som vill få en bild av dessa principer och hjälpmedel inklusive teori.

#### 6 **Exempel på situationer och val**

Detta kapitel är av exemplifierande karaktär. Det tar upp några vanligt förekommande situationer och belyser ställningstaganden. Beskrivningarna är mer direkt tillämpbara än i föregående kapitel. Det ger dock inte omedelbara lösningar utan vägledningar.

Kapitlet vänder sig till den som är intresserad av en viss situation (bland de beskrivna) eller mer allmänt av angreppssätt.

#### 7 **”Modern” statistikproduktion**

Kapitlet ger en orientering om så kallad modernisering av statistikproduktion. Detta är inte nytt tankemässigt, men det är ett förändringsarbete som tar tid att införa. Det gäller exempelvis standardisering, bl.a. av statistikens innehåll och av verktyg. En del om arbetssätt ingår, t.ex. kvalitetssäkring och förbättringar; viss vidareutveckling men inte nyutveckling.

Kapitlet är en bred översikt som kan läsas av de flesta.

Avslutningsvis finns en referenslista i kapitel 8. Hänvisningar till den görs i texten; i allmänhet som exempel för den som vill läsa mer. Hänvisningar görs också till processkartan och Verksamhetsstödet med de numreringar och det innehåll som gäller då designvägledningen skrivs och uppdateras.



## 2 Grundläggande begrepp och termer

Detta kapitel innehåller beskrivningar av grundläggande och viktiga begrepp och termer. Dessa återkommer många gånger i vägledningen, och det är praktiskt med en separat beskrivning. Texten i detta kapitel visar samtidigt på sätt att resonera som återkommer i de följande kapitlen. Allt hänger ihop, och det finns ingen uppenbar ordningsföljd; det finns ofta ett behov av att hänvisa till något förklarande som kommer senare.

Kvalitetsbegreppet för officiell statistik fyller en viktig funktion i flera avseenden. Från och med 2017 gäller det begrepp som finns i SCB (2016b) och SCB (2016c). Mer om detta följer, bl.a. i avsnitt 2.5.

### 2.1 Begreppsbildning samt intresse, mål och observation

#### 2.1.1 Begrepp (koncept)

Det finns några grundläggande frågor då statistikbehoven ska fastställas i form av statistiska storheter och dessas byggstenar (se avsnitt 2.2).

- Vilka sakproblem ska belysas?
- Vilka statistiska storheter kan belysa sakproblemen?
- Vilka variabler behöver observeras (eller mätas) för vilka objekt för att de önskade statistiska storheterna ska kunna skattas?

Arbetet med att föreslå, definiera och operationalisera sakproblem, sakfrågor, statistiska storheter och variabler kallas för konceptualisering eller begreppsbildning. Detta arbete måste givetvis bedrivas i nära samverkan med kund eller representanter för de tilltänkta användarna av statistiken.

Det finns olika tillvägagångssätt för att behandla konceptualiseringsfrågorna vid design av en statistisk undersökning. Vissa betonar vikten av att arbeta med begrepp som beståndsdelar i teorier; teorierna leder till hypoteser och frågor som behöver testas empiriskt, och då måste frågeställningarna operationaliseras i termer av mätbara variabler, som i sin tur operationaliseras ytterligare till frågor i frågeformulär. Sådana ansatser kallas *top-down* eller teoridrivna.

En motsatt ansats, som kallas *bottom-up* eller empiri- eller datadriven, kan t.ex. innebära att man undersöker hur människor diskuterar olika problem och frågeställningar i naturligt språk, hur frågor i liknande undersökningar har formulerats, etc. Utifrån dessa empiriska underlag abstraherar man fram bakomliggande begrepp och möjliga samband mellan begreppen. I praktiken är det ofta ändamålsenligt att kombinera teori- och empiridrivna ansatser.

I vissa fall finns redan ramverk och samband beskrivna. Ett enkelt exempel är att befolkningsförändringar under en viss period beror av antalen födda, döda samt in- och utflyttade. Laux och Barham (2013) beskriver några sådana ramverk med flöden i tid och rum, vilket ger en god överblick och insikt om variabler. Till bilden kan även knytas källor i form av administrativa data och undersökningar med direktinsamling, som tillsammans leder till statistik.

En annan referens för den som vill läsa mer är Forbes och Brown (2012) som beskriver vikten av *conceptual thinking* samt hur man bör tänka och sätta begrepp i relation till ramverk. Ett par citat följer, där NSO betyder *National Statistical Office*.

*“Conceptual thinking is, therefore, the ability to reason in the abstract.”*

*“An advanced conceptual thinker in a NSO would be able to draft a conceptual framework that simplifies and redefines reality into a set of precise and possibly inter-related variables about which data can be collected. To reach this level of conceptual thinking requires both an understanding of some fundamental concepts and an ability to reason with these concepts. Good subject-matter knowledge is also important.”*

### 2.1.2 Intresse, mål och observation

Användarna har sina **intressen**, t.ex. vilken population som statistiken ska avse. Statistikproducenten gör avvägningar – med hänsyn till användarna, uppgiftslämnarbördan, statistikens kvalitet och kostnader – och sätter därefter upp **mål** för vad som ska åstadkommas. Dessutom kan det som ska redovisas i statistiken kanske inte mätas i just den formen, utan data behöver **observeras** på ett lite annat sätt innan det blir statistik.

De här tre aspekterna är väsentliga under utformningen, särskilt begreppsbyggnaden. I många undersökningar sammanfaller de, men när olikheter finns kan de behöva framhållas. Orden används som prefix för t.ex. variabel och population. Skillnaderna förklaras ytterligare nedan med variabler som exempel.

#### *Intressevariabler*

Utgångspunkten för en statistisk undersökning är ett sakproblem hos statistik användare. Dessa kan ha olika önskemål och krav på statistikens variabler. De efterfrågade variablerna kallas *intressevariabler*.

#### *Målvariabler*

Statistikproducenten måste ta ställning till möjligheterna att ta fram statistik som utgår från intressevariablerna. I bedömningen ingår statistikens kvalitet, kostnader och uppgiftslämnarbörda. En avvägning med val mellan intressevariabler kan behöva göras, exempelvis för att begränsa uppgiftslämnarbördan. Mättekniska bedömningar behöver göras av SCB, som en del i utredningen av möjligheterna att

samla in efterfrågad information. Efter det att SCB gjort mättekniska bedömningar och andra avvägningar väljs de variabler som skattningarna i statistiken ska avse. Dessa variabler kallas *målvariabler*.

#### *Observationsvariabler*

De variabler som ska hämtas från register eller samlas in direkt från uppgiftslämnare kallas *observationsvariabler*. Dessa variabler kan vara desamma som målvariablerna eller vara annorlunda. Det finns flera tänkbara skäl till skillnader. Det kan vara lättare och säkrare för uppgiftslämnaren att få svara på flera relativt enkla frågor än en enda komplex. En uppdelning på flera frågor görs då i frågeformuläret för att förebygga mätfel och mätosäkerhet. Målvariabeln härleds ur flera observationsvariabler genom en beräkning, t.ex. en summering.

Det finns fall där det bedöms att målvariabeln är omöjlig att besvara eller att den skulle leda till betydande mätosäkerhet och bortfall om den skulle användas i enkäten. Ett alternativ är då att använda en modell som beskriver målvariabeln med hjälp av andra variabler, som kan observeras. I modellen ingår vissa antaganden, t.ex. om samband, och en modell innebär naturligtvis osäkerhet. Alla modeller är förenklade beskrivningar av verkligheten. Den bedömning som görs är att modellen är tillräckligt bra för att använda i statistikproduktionen och bättre än alternativen.

## **2.2 Statistikens byggstenar och struktur**

I såväl planering som genomförande av en statistisk undersökning ingår arbete med

- objekt (av en eller flera typer) – t.ex. individer, företag, fastigheter
- population(er) – t.ex. individer bosatta i Sverige och byggnader yngre än tio år
- redovisningsgrupper (delpopulationer) – t.ex. individer efter kön och åldersgrupper
- variabler – t.ex. civilstånd, inkomster, omsättning, lönesumma, byggnadsår
- referenstider: tidpunkter eller tidsperioder för variabler, objekt etc.

Referenstiderna ses ibland som separata och ibland som en integrerad del i t.ex. variabler och populationer.

En statistisk undersökning går ut på att skaffa kunskap om värdena på en eller flera statistiska storheter. En **statistisk storhet** bildas typiskt genom att

- ett statistiskt mått (t.ex. summa, medelvärde, antal) sammanfattar
- enskilda värden för variabler (t.ex. omsättning, kostnader, lön) hos
- objekten (t.ex. företag, fastigheter, individer) i en redovisningsgrupp
- för en eller flera referenstider.

Det är vanligt att tabellceller och marginaler i en statistisk tabell motsvarar statistiska storheter och att nedbrytningarna visar indelningar i redovisningsgrupper eller referenstider. Idealt skulle värdena på de statistiska storheterna kunna beräknas utan osäkerhet och fel. I praktiken resulterar dock en statistisk undersökning normalt i statistikvärden som har formen av skattningar, förenade med osäkerheter.

Användarnas önskemål och statistikproducentens val beträffande de statistiska storheterna särskiljs genom att använda prefixen intresse respektive mål framför orden (avsnitt 2.1). Målstorheterna byggs upp av målvariabler, målobjekt, målpopulationer etc. När valet är givet eller inga missförstånd uppstår, kan prefixet mål utelämnas. Det är t.ex. vanligt att säga population i stället för målpopulation.

### 2.3 Skattning och statistisk inferens

Termen estimation används enstaka gånger i designvägledningen, särskilt i den etablerade sammansättningen "urval och estimation". Annars används skattning och skattningsförfarande.

#### 2.3.1 Inledning med en illustration

**Statistisk inferens** är enligt Nationalencyklopedin en induktiv vetenskap där man drar slutsatser ur empiriska data under en osäkerhet orsakad av slumpmässighet i data. Denna definition framhåller erfarenhet, osäkerhet och resonemang; det är inte en säker härledning (deduktion).

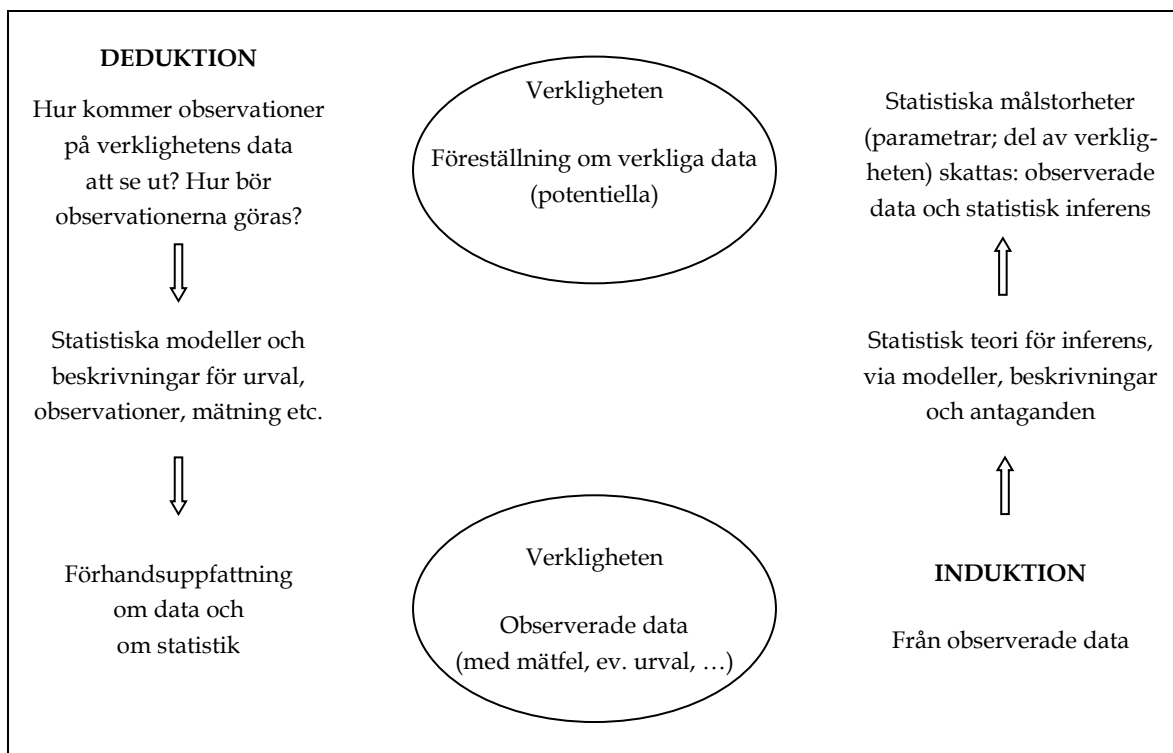
Den statistik som SCB tar fram och som karaktäriseras i avsnitt 2.2 är i första hand deskriptiv. Det hindrar inte att det är en form av statistisk inferens – mer eller mindre explicit – där statistikproducenten går från empiriska data till mer allomfattande skattningar. Osäkerhet finns genom flera källor som avsnitt 2.6 pekar på. Det är önskvärt att – med fokus på den resulterande statistiken – bedöma och presentera osäkerheten för statistikanvändarna för att deras vidarebearbetningar, tolkningar och slutsatser ska bli väl underbyggda och riktiga. SCB

presenterar ibland kvantitativ information som stöd för analys av skillnader, t.ex. osäkerhetsmarginaler och konfidensintervall.

Den osäkerhet som finns i data påverkar dels hur skattningar bör utformas, dels hur osäkra (tillförlitliga) skattningarna blir. Det är väsentligt vad som ska skattas. (Det är typiskt statistiska målstorheter, men det kan vara någon annan form av parameter; mer om detta, t.ex. i kapitel 5.) Det kan ligga nära till hands att tänka att statistisk inferens enbart gäller vid urvalsundersökningar, men den begränsningen finns inte. Även i andra fall vill statistikanvändare kunna dra slutsatser, t.ex. om likheter och olikheter vid jämförande studier mellan grupper och över tid.

Resonemang om skillnader och observerade skillnaders betydelse förs för övrigt även internt, typiskt i samband med inbäddade experiment, se avsnitt 5.6.2.

Figur 2.1 Statistisk inferens, en illustration



Figur 2.1 (och den mer omfattande och något abstrakta versionen figur 5.1 i kapitel 5) är inspirerade av figurer om inferens i Barnett (1999, sidan 8), Makar och Rubin (2009, sidan 85) samt Kass (2011, sidorna 5 och 6). Figur 2.1 har två ovaler som båda avser data i verkligheten. Den övre ovalen "innehåller" faktiska data som helhet utan praktiska begränsningar. Exempelvis finns både intresse- och mål-

variabler där, oavsett om de är observerbara eller inte. Den nedre ovalen innehåller observerade data.

Till vänster i figuren illustreras "förflyttningen" från den övre ovalen till den nedre. I den ingår mer eller mindre kända förfaranden som ramframställning, urvalsdragning (när sådan görs), mätförfarande, etc. Detta kan ses som en form av härledning (deduktion) i förväg av data som kommer att observeras med hänsyn tagen till de samband och osäkerheter som finns. Här ingår t.ex. sannolikhets teori som används i förhandskalkyler av resursallokering, bortfall, mätosäkerhet etc. eller för ett urvalsförfarande. Syftet är att i förväg bedöma hur den statistiska undersökningen bör utformas och hur tillförlitlig statistiken kan väntas bli.

Till höger i figuren illustreras den omvända förflyttningen, från observerade data till verklighetens helhet med de utsedda statistiska målstorheterna. Denna induktion "räknar i motsatt riktning" jämfört med vad som sker på den vänstra sidan. Den beaktar således de osäkerheter som finns och söker skattningar av målstorheterna, skattningar med goda egenskaper som ingen eller liten bias (skevhet, systematisk avvikelse; se vidare avsnitt 2.6) och liten varians. För att kunna prata om bias och varians måste det finnas en beskrivning av systematiska och slumpmässiga avvikelser.

Vad som orsakar de systematiska och slumpmässiga avvikelserna, och hur stora dessa är, beror på den statistiska undersökningens utformning och genomförande. (Osäkerhetskällor och felmodeller beskrivs kort i avsnitt 2.6 och mer ingående i kapitel 3.) Skattningar utformas genom att aggregera (mikro)data till statistikvärden (makrodata) enligt någon princip som hänger samman med relationerna mellan de data som man har och den statistiska målstorhet som man ska skatta.

För en statistiker kan det vara naturligt att se data som genererade på något sätt sådant att sannolikheter ingår. Det kan vara en beskrivning av ett urvalsförfarande. Det kan vara en beskrivning med svarsbenägenheter eller av mätvärden jämfört med sanna värden. Det kan vara en modell med slumpmässiga (stokastiska) inslag som beskriver hur vissa variabelvärden genererats utifrån andra. Vidare finns det olika sätt att beskriva osäkerheter i data och att göra generaliseringar – olika statistiska skolor. Skillnaderna är påtagliga i teorin. De kan vara stora också i praktiken, när det gäller hur undersökningar och skattningar bör utformas, men de behöver inte vara det. Många gånger överensstämmer synen både på vad som är betydelsefullt och på vad som varierar mycket. Skattningarna är i vissa situationer desamma, medan tillförlitligheten beräknas på olika sätt.

### 2.3.2 Utgångspunkter för skattning och statistisk inferens

Ett i officiell statistik vanligt förekommande synsätt på skattning kallas designbaserat. Då är det den slumpmässiga urvalsdragningen med urvalssannolikheter som dominerar utformning av skattningar och beräkning av tillförlitlighet.

Det finns en utvidgning där modeller tas till hjälp utan att frånga grundprincipen. Den metoden kallas modellassisterad estimation. Kalibrering är ett typiskt exempel; en eller flera hjälpvariabler från register inkluderas i skattningen. En ursprunglig tanke med förfarandet var att minska variansen. I dag används förfarandet ofta i syfte att minska risken för att skattningar snedvrids på grund av ökande bortfallsandelar. Hjälpvariablerna är, enkelt uttryckt, kraftfulla om de "fångar" variationer i undersökningsvariabeln och i svarsbenägenheten.

Designbaserade skattningar för små redovisningsgrupper har en hög varians. Det är då vanligt att i stället använda modellbaserade skattningar, dvs. skattningar som är beroende av modellantaganden. Andra situationer där modellbaserade inslag förekommer, åtminstone delvis, är vid imputering och vid tidiga skattningar.

När statistik tas fram från administrativa data, kan modellering och omräkningar/justeringar behövas, t.ex. för att härleda eller modellera statistiska variabler baserat på de administrativa variabeldefinitionerna och för att räkna om från administrativa objekt till de objekt som definieras för statistiken (mer om detta i avsnitten 3.3 och 3.4).

Det är även vid totalundersökningar och registerbaserad statistikproduktion viktigt att tydligt uttala målet för skattning och inferens, t.ex. målpopulationen. Det kan finnas täckningsbrister, saknade data och andra osäkerhetskällor att hantera i skattningen och i tillförlitlighetsbedömningen. I sådana skattningssituationer, med totalundersökning eller administrativa data (eller båda), finns (ännu) inte allmänt vedertagna inferensprinciper.

Betydelsen för statistikens kvalitet av vissa moment, bl.a. matchning, diskuteras dock. Se t.ex. Bakker, van der Heijden och Scholtus (2015) samt flera artiklar i det specialnumret av tidskriften *Journal of Official Statistics* (JOS) om täckningsproblem för administrativa källor.

Ett par referenser för den som vill läsa med ett mer framtida perspektiv är följande. I en holländsk rapport diskuterar Buelens, Boonstra, van den Brakel och Daas (2012) hur tänkesättet för statistisk inferens går från designbaserat till modellbaserat och vidare till vad de kallar algoritmisk inferens. Det sistnämnda är aktuellt t.ex. med så kallade *big data*. Little (2012) beskriver det som han kallar en kalibrerad bayesiansk ansats i en artikel som åtföljs av en diskussion.

Några olika typer av statistisk inferens som är vanliga i officiell statistik beskrivs i flera avsnitt, t.ex. 5.1, 5.2 och 6.4.

## 2.4 Statistisk undersökning, process, produkt, system

### 2.4.1 Allmänt om statistisk undersökning

I en statistisk undersökning ingår, i stora drag, att fastställa behov och designa undersökningen med hänsyn till krav och restriktioner, att samla in och bearbeta data, att analysera data, att redovisa resultat (statistik eller mikrodata) samt att utvärdera produktionsprocessen.

En statistisk undersökning producerar således statistisk information utifrån insamlade uppgifter enligt en specificerad design. Undersökningen syftar till att tillgodose informationsbehov om en population. Behov och design fastställs vanligen i samarbete med användare/kund(er).

Det finns flera sätt att karaktärisera en undersökning inklusive resultatet:

- Engångsundersökning *eller* upprepad undersökning. De senare är ofta löpande med fasta intervaller, t.ex. varje månad, varje år eller vart femte år.
- En undersökning som samlar in data, ofta mikrodata, från en eller flera källor (se punkterna nedan) *eller* en undersökning som hämtar data, ofta makrodata (statistik), från flera undersökningar och sammanställer eller vidarebearbetar dessa data.
- Direktinsamling *eller* användning av register; en kombination är vanligt förekommande:
  - Data samlas in från uppgiftslämnarna och specifikt till undersökningen.
  - Undersökningen använder administrativa data eller andra insamlade data som finns tillgängliga internt eller externt.
- Datainsamlingsmetod vid direktinsamling, som enkät per post eller webb, intervju genom telefon eller besök, filöverföring eller observation i fält.
- Urvalsundersökning eller totalundersökning:
  - I en urvalsundersökning ingår ett urval (t.ex. 2 000 individer av 7,4 miljoner i Sveriges befolkning i åldern 17–74 år) i datainsamlingen.
  - I en totalundersökning ingår **alla** objekt (t.ex. alla kommuner) i datainsamlingen.

Ytterligare karaktäriseringar finns. Det kan vara praktiskt att skilja på situations- och händelsebaserade undersökningar, där de senare utmärks av att det är en händelse (t.ex. en födelse eller en rapporterad trafikolycka) som skapar ett objekt. I registervård görs fortlöpande



uppdateringar, oftast med hjälp av flera källor. Detta är en form av undersökning, som ofta är registerbaserad men som kan innehålla direktinsamling. När ett objekt uppstår till följd av en händelse, är rapporteringssystemet viktigt: hur fullständigt och snabbt det är.

#### **2.4.2 Observationsregister**

Observationsregister är datamaterial som innehåller mikrodata som den statistiska undersökningen har samlat in. Observationsregistren uppstår och används under statistikproduktionsprocessen. Det slutliga observationsregistret är den sista versionen. Detta sparas, och det är ett resultat från den statistiska undersökningen. Det används för att framställa statistik.

#### **2.4.3 En undersökning kan vara "begränsad"**

En statistisk undersökning behöver inte innefatta alla delar av statistikproduktionsprocessen. Den måste dock innehålla såväl utformning (med användarbehov, design, planering samt skapande och test) som något genomförande (insamling, bearbetning, analys eller spridning).

En statistisk undersökning behöver inte ta fram statistik utan kan resultera i mikrodata. Dessa data kan sedan vidarebearbetas, t.ex. av en extern kund eller av andra på SCB, och resultera i en ny uppsättning mikrodata, i mer aggregerade data eller i statistik. De båda sistnämnda kategorierna kallas även makrodata.

När en statistisk undersökning begränsas behöver dess kopplingar till andra undersökningar tydliggöras; beroenden i båda riktningarna. Den som hämtar data eller statistik från någon annan behöver kvalitetsinformation om det som hämtas. BaR – Beskrivning av register – är ett exempel, speciellt för register, på att tillhandahålla information. Denna form av dokumentation är relativt ny; den infördes i Verksamhetsstödet 2014 och har successivt börjat användas internt.

Vanliga exempel på "begränsade" undersökningar är enkätundersökning med leverans av mikrodata, framställning av interna register och sammanställning av register till nya register (integrationsregister).

#### **2.4.4 Undersökning, process, produkt, system**

Införandet av ett processororienterat arbetssätt på SCB har bland mycket annat medfört att processkartor används. Se Inblick allmänt och se Verksamhetsstödet och processkartan där för statistikproduktion. Det har också medfört ett delvis nytt språkbruk, ofta parallellt med det tidigare. Det gäller t.ex. för det nyare ordet statistikproduktionsprocess och det mer etablerade statistisk undersökning som lever kvar. De kan i allmänhet användas synonymt. En del uppfattar statistisk

undersökning som smalare, t.ex. endast vid direktinsamling, men så är det inte enligt nu gällande terminologi på SCB.

Begreppet produkt används i olika sammanhang och ofta men inte alltid med ett tydliggörande prefix. En statistisk undersökning leder till en slutprodukt, och det finns två typer av sådana slutprodukter, nämligen (slutliga) observationsregister och statistik.

I systemet för Sveriges officiella statistik är statistikprodukt ett etablerat begrepp; ibland används i stället det kortare ordet produkt. I årsrapporten från Rådet för den officiella statistiken (ROS) för 2015 (SCB, 2016a) står bl.a. "Begreppet statistikprodukt används för att beskriva vilken statistik som produceras. I princip kan en statistikprodukt likställas med en statistisk undersökning." Normalfallet är att det finns ett "ett-till-ett" förhållande mellan en statistikprodukt och en undersökning. Ibland ger en undersökning upphov till flera statistikprodukter. Omvänt kan flera undersökningar bidra till en statistikprodukt. Hur de enskilda undersökningarna och produkterna inom ett ämne avgränsas är en praktisk fråga.

Begreppet produkt används även administrativt – t.ex. i produktdata-basen – och för ekonomiska produkter. Det finns en roll produkt-ansvarig kopplad till sådana användningar.

Det finns numera en strävan att se mindre till enskilda produkter som skräddarsys för specifika behov och mera till statistiska system där flera (eller snarare många) statistiska undersökningar (statistikprodukter) ingår. Detta gäller både nationellt och internationellt. Starka skäl för sådan systemsyn är krav på eller önskemål om jämförbarhet och sammanvändbarhet av statistik. Samordning av statistiska undersökningar kan ha fler positiva effekter, t.ex. lägre uppgiftslämnarbörda och lägre kostnader.

Det ekonomisk-statistiska systemet är ett exempel på ett system som har funnits länge och är rätt väl etablerat med tre huvudnivåer. Företagsdatabasen FDB är en samordnande grund för de följande nivåerna med undersökningar baserade på mikrodata och vidarebearbetande undersökningar med bl.a. nationalräkenskaper (NR). Krav på samordning kommer från NR, och samordningen byggs från grunden, t.ex. i ramframställningen som är enhetlig och i många fall samordnad.

Ytterligare aspekter gäller produktionssystem; hur de är byggda med avseende på metoder, verktyg, data, IT etc. i sig själva och i förhållande till varandra. Sådana arkitekturfrågor diskuteras vidare nedan, särskilt i avsnitt 7.4, men även i övrigt och ofta utan ordet arkitektur.

Frågan om hur ett system sätts samman av undersökningar har många delar. Hur kan och ska frågeformulär samordnas med var-

andra och med administrativa uppgifter? Kan och ska uppgiftslämnare som är med i flera undersökningar få sammansatta frågeformulär? Kan systemet och undersökningarna (ofta med urval) byggas i moduler som kombineras när skattningar tas fram? Det finns många vinster att göra, men det kräver mycket förberedelser. Se vidare nedan, främst avsnitt 6.8.

Ytterligare en term är **omgång**. Den används här främst i sammansättningen produktionsomgång. I det enskilda fallet är det viktigt att precisera vad som ingår i en insamlingsomgång, beräkningsomgång, redovisningsomgång etc.

## 2.5 Statistikens kvalitet och dess komponenter

### 2.5.1 Definition av kvalitet

Det finns många besläktade varianter på definitioner av statistikens kvalitet. Några exempel:

- Alla sammantagna egenskaper hos en produkt som ger dess förmåga att tillfredsställa uttalade eller underförstådda behov.
- Kvaliteten hos en produkt (vara eller tjänst) är dess förmåga att tillfredsställa, eller helst överträffa, kundernas behov och förväntningar.
- *Fitness for use* (använt av Joseph M. Juran).
- *Fitness for purpose* (använt av W. Edwards Deming), som betonar att kvaliteten ska sättas i relation till den avsedda användningen; en eller flera användningar för en eller flera användare eller kunder. I det svenska statistiska systemet finns följande riktlinje: "Myndigheten avgör om statistikens kvalitetsnivå är tillräcklig, i relation till hur den är avsedd att användas." (se t.ex. SCB 2016a, Bilaga 7).
- *Quality is a multi-dimensional concept and encompasses all aspects of how well statistics are fit for their purpose* (Eurostat, 2014).

Användningar av statistiken kan behöva prioriteras sinsemellan, t.ex. om kostnadsrestriktionerna är starka eller om användningarna innebär olika omfattning eller krav på statistiken. (Mer om detta på flera ställen nedan.)

En annan definition anger kvaliteten som en funktion av storleken på det totala felet för en skattning (*total survey error*). Det är kvalitet i den begränsade meningen tillförlitlighet. Se vidare avsnitt 2.6.

### 2.5.2 Huvudkomponenter i kvalitetsbegreppet för officiell statistik

I det svenska statistiska systemet (och likartat i många andra statistiska system; se t.ex. Eurostat (2014) för det europeiska statistiska systemet,

ESS) beskrivs statistikproduktens kvalitet fr.o.m. 2017 med nedanstående fem huvudkomponenter (SCB 2016b, 2016c):

- Relevans
- Tillförlitlighet
- Aktualitet och punktlighet
- Tillgänglighet och tydlighet
- Jämförbarhet och sammanvändbarhet

Huvudkomponenterna delas i sin tur in i underkomponenter och, i förekommande fall, delkomponenter, se vidare nedan, särskilt avsnitt 2.5.3 och 2.6. Kvalitetsbegreppet används vid utveckling, framställning och spridning av officiell statistik, bl.a. i kvalitetsdeklarationer avseende statistikprodukter.

Kvalitetskomponenterna avser att beskriva kundens uppfattning, men det är främst producenter som har arbetat med dessa begrepp/ramverk/strukturer för kvaliteten med komponenter eller dimensioner på ett par nivåer.

Tillförlitlighet är den kvalitetskomponent som normalt kräver mest uppmärksamhet i designarbetet. (På engelska är termen *accuracy*. Den översätts i andra sammanhang ofta med noggrannhet. Tillförlitlighet valdes 1994 för det svenska systemet av bedömda användarskäl. Det kvarstår 2016.)

Kunden uttrycker sina krav på kvalitet i olika avseenden och ska vid behov få sakkunnig hjälp att precisera kraven. I kunddialogen ingår avvägningar mellan kvalitet och kostnader. Det gäller att vara lyhörd och ge konstruktiva förslag. Kanske kan kvaliteten för kunden höjas betydligt till en liten merkostnad. Kvalitetskomponenten framställningstid, som i SCB:s kvalitetsbegrepp finns under *Aktualitet och punktlighet*, kan antingen vara fastlagd och därmed utgöra en restriktion eller ingå explicit i den avvägning som ska göras (mer om detta nedan under rubriken optimering). I det senare fallet kan avvägningar mellan tillförlitlighet och aktualitet vara en del i kunddialogen och design-/optimeringsarbetet. En annan tänkbar avvägning är aktualitet mot kostnader. Det kan i vissa fall vara så att kunden finner det värt en merkostnad att få resultat något tidigare.

Kvaliteten i output – mikrodata och statistik – beror på kvaliteten i valda processer och i utförandet. Design handlar mycket om val av metoder, verktyg och arbetssätt samt om resursfördelning. En aspekt i detta arbete är att vara medveten om vilka risker som finns och vilka konsekvenser de kan medföra – och planera därefter. Om exempelvis tidshållning är viktig ska tidsplanen innehålla goda marginaler.

Dokumentation av avvägningar, av statistikproduktionsprocessen och av dess output är väsentlig, se avsnitt 2.8.

### 2.5.3 Huvudkomponenten Relevans och statistikens ändamål

Huvudkomponenten *Relevans* har två underkomponenter, som i sin tur har två respektive fem delkomponenter:

- Ändamål och informationsbehov
  - Statistikens ändamål
  - Statistik användares informationsbehov
- Statistikens innehåll
  - Objekt och population
  - Variabler
  - Statistiska mått
  - Redovisningsgrupper
  - Referenstider

Kvalitetskomponenten *Ändamål och informationsbehov* får med synsättet *fitness for purpose* en särställning, inte minst *Statistikens ändamål*. Det är väsentligt att formulera statistikens ändamål och att relatera statistikens kvalitet till ändamålet. Alla informationsbehov kan normalt inte tillgodoses utan prioriteringar måste göras.

Ett sätt att konkretisera ändamålet och kvalitetskraven är att sätta upp en *tabellplan* – ett antal tabeller som speglar de statistiska målstorheter som ska skattas. Tabellplanen visar således statistikens avsedda innehåll. Den blir samtidigt ett hjälpmedel för att ta ställning till krav på andra kvalitetskomponenter och till restriktioner. Följer statistiken EU-regleringar? Kommer tabellerna i tabellplanen att kunna redovisas med hänsyn till tillförlitlighet och statistiksekretess? Kommer statistiken att kunna sammanvändas med annan statistik som det finns önskemål om? Kommer den att tillgodose användarnas behov av aktualitet? Detta är frågor både i designarbetet och vid utvärdering.

### 2.6 Tillförlitlighet och osäkerhetskällor

Tillförlitligheten är en kvalitetskomponent som det finns stora möjligheter att påverka i designarbetet. I kvalitetsbegreppet har huvudkomponenten *Tillförlitlighet* tre underkomponenter, av vilka en är indelad i sex delkomponenter. De tre underkomponenterna belyser olika aspekter.

Den andra delkomponenten handlar om osäkerhetskällor som påverkar tillförlitligheten totalt, dvs. den första delkomponenten. Att inkludera osäkerhetskällor innebär att ha ett producentperspektiv, eftersom det underlättar för producenten vid beskrivning av tillförlitlighet. För många användare är den "totala" tillförlitligheten det

som är intressant. Indelningen kan dock underlätta bedömningar för en del statistik användare. Producenten behöver kunskap i sitt arbete med förbättringar för att lägga resurser där de gör mest nytta. Relationer mellan osäkerhetskällor kan variera över tid, t.ex. om bortfallsandelen ökar eller om ett register som används till ramen förändras.

Den tredje underkomponenten handlar inte direkt om skattningars tillförlitlighet. Den sätter i stället preliminära statistikvärden i relation till motsvarande slutliga statistikvärden. Den kvalitetsinformationen ger användarna information om hur stora löpande revideringar de kan vänta sig. Se främst avsnitt 6.6.

- Tillförlitlighet totalt
- Osäkerhetskällor
  - Urval
  - Ramtäckning
  - Mätning
  - Bortfall
  - Bearbetning
  - Modellantaganden
- Preliminär statistik jämförd med slutlig

Slumpmässiga och systematiska osäkerheter är viktiga att hålla isär. Strävan bör vara att hålla de systematiska avvikelserna små (helst obefintliga) med hänsyn tagen till kostnader. För urvals- och skattningsförfaranden finns en teori och tydliga kriterier för önskvärda egenskaper. Detta gäller inte bara urvalsundersökningar utan generellt. Två typiska sådana kriterier är:

- ingen bias (ingen systematisk avvikelse, ingen skevhet)
- liten varians.

Det är vanligt att "mjuka upp" kriteriet ingen bias till försumbar eller liten bias.

Ofta görs en sammanvägning av de två kriterierna ovan. Det ofta använda sammanfattande mått som kallas medelkvadratavvikelse eller medelkvadratfel är summan av skattningens (mer tekniskt uttryckt: estimatorns) kvadrerade bias och dess varians. På engelska heter medelkvadratfelet *Mean Squared Error*, vilket ofta förkortas MSE. Denna akronym förekommer också i svensk text. Man talar även om *Root Mean Squared Error (RMSE)*, som är kvadratroten ur MSE. Det gäller således för varje skattning att

$$\text{MSE} = \text{RMSE}^2 = B^2 + S^2, \text{ där } B \text{ är bias och } S \text{ är standardavvikelse.}$$

Går varians och bias att beräkna? Särskilt för många vanligt förekommande urvals- och skattningsmetoder finns formler som hänger samman med urvalsförfarandet. I andra situationer (med andra urvalsmetoder eller med totalundersökningar) kan det vara betydligt

svårare. Medelkvadratfelet utgör ett tänkbart sätt att göra en sammanvägning. Ingen eller obetydlig bias kan dock vara ett krav.

När det finns formler behövs siffervärden att sätta in i formlerna, vilket är en utmaning, särskilt för nya undersökningar. Se vidare avsnitt 2.7.1 om optimering samt avsnitt 5.6 om pilotstudier.

Kriteriet för den bästa lösningen (ett optimum, se avsnitt 2.7) blir, med denna sammanvägning, att åstadkomma minsta medelkvadratfelet givet budgeten. Detta mått på osäkerhet kan tas fram för en eller flera skattningar (variabler), för en eller flera redovisningsgrupper, för nivå- eller förändringsskattningar etc. Därmed måste kriteriet för optimeringen (den bästa lösningen) preciseras: t.ex. om en eller flera skattningar ska ingå samt om nivå- och förändringsskattningar är lika viktiga. Valet sker efter en gradering av skattningarnas betydelse för kund/användare.

Det totala felet i en skattning – *total survey error* – har diskuterats mycket i både teori och praktik. Två ganska nya artiklar på detta tema har underrubrikerna "*Design, implementation, and evaluation*" (Biemer, 2010) och "*Past, present, and future*" (Groves och Lyberg, 2010). Lyberg (2012) ger en bred diskussion av kvalitet med en tillbakablick. Det totala felet är intressant i sig. En nedbrytning på osäkerhetskällor (felkällor) kan bidra till dels en förståelse för vad som är stort och smått, dels kunskap om vad som samverkar. Det finns felmodeller, som kan vara användbara. Det kan dock vara svårt att isolera och mäta komponenterna i sådana felmodeller. Det finns många bias- och varianskomponenter. Många fel är svåra att kvantifiera. Marella (2007) ger ett exempel på en modell som beaktar både fel (urval, bortfall, mätning) och kostnader.

## **2.7 Optimering, risker, kvalitetssäkring, kvalitetskontroll**

De fyra begreppen i rubriken beskrivs nedan och de är besläktade. Undersökningens design påverkar både kvalitetssäkring och kvalitetskontroll – och omvänt bör kvalitetssäkring och kvalitetskontroll påverka designen. Det finns även en likhet med tillräcklig kvalitet, där syftet är en viss kvalitetsnivå.

### **2.7.1 Optimering**

En optimal lösning är den som är bäst i något avseende. Det handlar för statistiska undersökningar ofta om kvalitet eller kostnader eller någon relation däremellan.

En optimal design är då den design som ger högsta kvalitet för en given kostnad eller, omvänt, den design som uppnår en viss kvalitet till lägsta kostnad. Detta är det klassiska effektivitetskriteriet för planering av en undersökning. Kvalitet är ett mångfacetterat begrepp och behöver delas upp i sina komponenter. Avvägningar mellan

kvalitetskomponenter som "drar åt olika håll" ingår ofta i optimeringsdiskussioner och designarbete. Tillförlitlighet och aktualitet är en sådan konflikt. Tillförlitligheten kan kanske ökas genom en intensiv bortfallsuppföljning som dock tar tid. Andra konflikter finns t.ex. mellan reduktion av bortfall och reduktion av mätosäkerhet.

Optimeringsmålet kan, som exempel, vid urvalsdesign vara högsta möjliga tillförlitlighet för en given urvalsstorlek eller minsta möjliga urvalsstorlek eller kostnad för ett givet precisionskrav. I exempel som dessa två kan det gå att räkna fram en optimal lösning när andra faktorer och förutsättningar för undersökningen har lagts fast, som objekt, population, variabler etc. Det är naturligtvis i allmänhet enklare att räkna på en avgränsad del av undersökningen än på hela. Ändå ska hela undersökningen beaktas, och det får göras med tidigare erfarenheter som utgångspunkt. Ny information kan behöva samlas in för att möjliggöra beräkningar, t.ex. genom en pilotstudie (avsnitt 5.6.1).

I praktiken finns knappast en strikt optimal design. Det finns många påverkande faktorer, och det är inte möjligt att härleda konsekvenserna för kvalitet och kostnader. Strävan får främst vara att åstadkomma en god design som beaktar viktiga faktorer och undviker inflytelserika förbiseenden.

Det finns rätt få publicerade konkreta studier om kvalitet och kostnader. Några exempel på studier, för den som vill läsa mer:

- Linacre och Trewin (1993) är ett klassiskt exempel, där många val och kostnadsuppgifter ingår.
- Marella (2007) är ett annat exempel som är inriktat på felstorlekar till följd av urval, bortfall och mätning samt allokering för att minimera dessa för en given total kostnad.
- Steel och Clark (2014) gör en studie där de visar att skillnaden i kostnad mellan objekt ska vara rätt stor för att det ska löna sig att beakta den i urvalsdesignen – de pekar dessutom på en effekt av datainsamlingsmetod som intressant.

På senare tid har många studier gjorts om kostnader och bortfall med inriktning på strategier och metoder för datainsamling. Se t.ex. avsnitt 5.3.4, där det finns referenser.

### 2.7.2 Riskhantering

I riskhantering ingår att identifiera och värdera risker. I statistikproduktionen kan det t.ex. vara resursbrist, lågt inflöde eller oroande stora avvikelser i makrogranskningen. I riskvärderingen ingår sannolikheter och konsekvenser. I SCB:s riktlinjer för riskhantering, SCB (2014a), beskrivs riskvärderingen som en bedömning av dels sannolikheten att risken inträffar, dels konsekvenserna om den inträffar.



Värderingarna – ofta på en skala från 1 till 5 för var och en – multipliceras. Hanteringen prioriterar främst utifrån det erhållna värdet.

I Verksamhetsstödet finns skrivningar om detta, i nuvarande version t.ex. som nedan (mer utförligt i process 2.8.2 *Planera och boka in resurser*). Särskilt den andra punkten visar på vikten av dokumentation: i planeringen, under genomförandet och vid åtgärder.

- Identifiera kritiska moment i produktionen, t.ex. moment där det finns risk för störningar eller där konsekvenserna av fel är särskilt allvarliga. Beakta dessa för att om möjligt undvika att problem måste lösas under tidspress, och identifiera redan i förväg var behov av extra insatser eller speciell kompetens kan uppstå. Exempelvis ska det finnas kapacitet (tid och resurser) för att kunna ta hand om brister som upptäcks i makrogranskning och inför publicering/redovisning.
- Oavsett om det finns ett avtal (motsvarande) eller inte, så föreligger en rapporteringsskyldighet: vid avvikelse eller risk för avvikelse från gjorda specifikationer ska processutföraren i fråga rapportera till de berörda på egen och andra enheter, däribland produktansvarig (motsvarande). Parterna diskuterar lämpliga åtgärder för att komma till rätta med avvikelsen. Vid oklarheter i åtgärder eller finansiering lyfts frågan till närmaste ansvariga chefer.

Det är ofta extra känsligt när en löpande statistisk undersökning genomför ändringar, särskilt om det är flera ändringar samtidigt och om många organisatoriska enheter är inblandade. Det gäller att vara förutseende, som framgår ovan och i nästa avsnitt som handlar om kvalitetssäkring.

### 2.7.3 Kvalitetssäkring

Storbritanniens *Office for National Statistics* (ONS) förklarar i sin kvalitetsterminologi kvalitetssäkring med att förutse och undvika problem (*anticipating and avoiding problems*). En annan beskrivning som kommer från en ISO-standard säger att kvalitetssäkring handlar om nedanstående två aspekter.

- Angreppssätt och metoder för att uppnå avsedd kvalitet.
- Inge tilltro till att kvalitetskrav kommer att uppfyllas.

För att statistiken ska hålla den kvalitet som har avtalats, förutskickats eller utlovats behövs en god och realistisk planering, styrning under statistikproduktionen samt bedömningar och kontroller av kvaliteten i processer och i slutprodukten.

Att använda utprovade arbetssätt, metoder, checklistor etc. har flera positiva effekter. Det är t.ex. lättare att förutse slutproduktens kvalitet

och att undvika situationer som gör att den önskade kvaliteten inte uppnås. Gemensamma metoder, verktyg och arbetssätt bidrar därmed till kvalitetssäkringen av enskilda statistikprodukter.

Medan kvalitetssäkring är allt som görs för att få en bra kvalitet, är kvalitetskontroll en kontroll av att det blev som det var tänkt, se nästa avsnitt.

#### 2.7.4 Kvalitetskontroll

ONS beskriver kvalitetskontroll som att ta hand om observerade problem (*responding to observed problems*). Fokus ligger på det som kan mätas och bedömas som acceptabelt eller inte. Granskning är en viktig form av kvalitetskontroll, t.ex. produktionsgranskningen.

Kvalitetskontroll används för att följa upp att inplanerade arbetssätt, metoder, verktyg etc. används, fungerar som avsett och resulterar i avsedd kvalitet. Kontrollerna kan vara av olika slag. Det kan vara en fråga om att kontrollera att designspecifikationer följs, t.ex. att översättning av surveymaterial har skett på anvisat sätt. Det kan också vara fråga om beräkningar av felfrekvenser i produktionsprocesserna. Ibland består kontrollen av att värden på processvariabler plottas i styrdiagram så att det blir möjligt att kontrollera variationer i processen. Vilken typ av kontroller som ska användas beror på situationen. (Se avsnitt 2.8 om processdata.)

Det är viktigt att kvalitetskontrollen används för att styra upp processer som inte fungerar som tänkt. Exempel på åtgärder är processförbättringar, extra träning av intervjuare och kodare eller modifieringar i designen som en följd av värdena på vissa processdata.

Kontroller görs bl.a. i produktionsprocess 6, *Analysera*, där makrodata granskas, slutliga observationsregister fastställs, samt tolkningar och förklaringar tas fram. Slutkontroll görs i process 7, *Redovisa och kommunicera*, genom slutgranskning vid publicering och leverans av statistik.

#### 2.8 Metadata och processdata

Användare behöver dokumentation av olika slag för att kunna förstå och använda mikrodata och statistik på ett riktigt sätt. En användare behöver i första hand information om data, så kallade metadata, t.ex. definitioner av variabler samt kvalitetsinformation. Metadata har två huvudsyften:

- (i) Att underlätta för användaren att hitta relevant statistik (eller relevanta data) givet ett sakproblem.
- (ii) Att hjälpa användaren att tolka och analysera statistiken (eller data). Olika användare har olika behov av metadata beroende på användning och kompetens.

Metadata används även internt i produktionsprocesserna för att styra processerna, se vidare nedan. En producent behöver detaljerad information om processerna, så kallade processdata eller paradata. Det finns ingen skarp gräns mellan metadata (data om data) och processdata (data om processen). Ibland betraktas processdata som ett slags metadata.

Data om produktionsprocessen och hur denna fungerar bör tas fram och sparas av flera skäl. Processdata bidrar till information om produktkvalitet och processkvalitet. De är ett av flera underlag för utvärdering och återkoppling till kommande produktionsomgångar. Detta gäller främst men inte enbart för löpande produkter – lärdomar kan göras också för likartade produkter. Processdata kan (liksom metadata) även användas dynamiskt för förändringar och ingripanden under en produktionsomgång.

Ur den stora mängden tänkbara processdata gäller det att välja ut och spara dem som är mest användbara för att förbättra processernas kvalitet och effektivitet. En sådan användning är att justera processens "intensitet"; t.ex. satsa mer eller mindre på bortfallsuppföljning beroende på resultatet av analysen av processdata. En annan användning av processdata är att förändra hela processen; att exempelvis inte intensifiera bortfallsuppföljningen utan försöka möta ökande bortfall med att underlätta uppgiftslämnandet.

Arbetet med att ta fram dokumentation och metadata för produkter och processer upplevs ofta som resurskrävande. Det är viktigt att designa processerna så att dokumentation, metadata och processdata i görligaste mån genereras automatiskt som biprodukter från processerna. Dokumentation bör genereras, skrivas och sparas fortlöpande, inte skjutas upp till dess att en undersökning är färdigdesignad eller en produktionsprocess är genomförd.

Det är viktigt att dokumentera inte bara genomförandet av produktionsprocessen utan också förberedelsearbetet, dvs. kundens krav och önskemål, vilka val som har gjorts samt skälen till dessa val.

Uttrycket metadata driven statistikproduktion har funnits länge. Det finns goda skäl att i stället prata om design driven statistikproduktion. I detta ligger "tänk först, gör sedan". Det handlar om val och allokeringar samt om att styra produktionen med metadata. Det finns fördelar med att ha parametrar som sätts en enda gång som input till ett system eller beräkningsprogram, hellre än att ändra en parameter manuellt (t.ex. referensperiod).

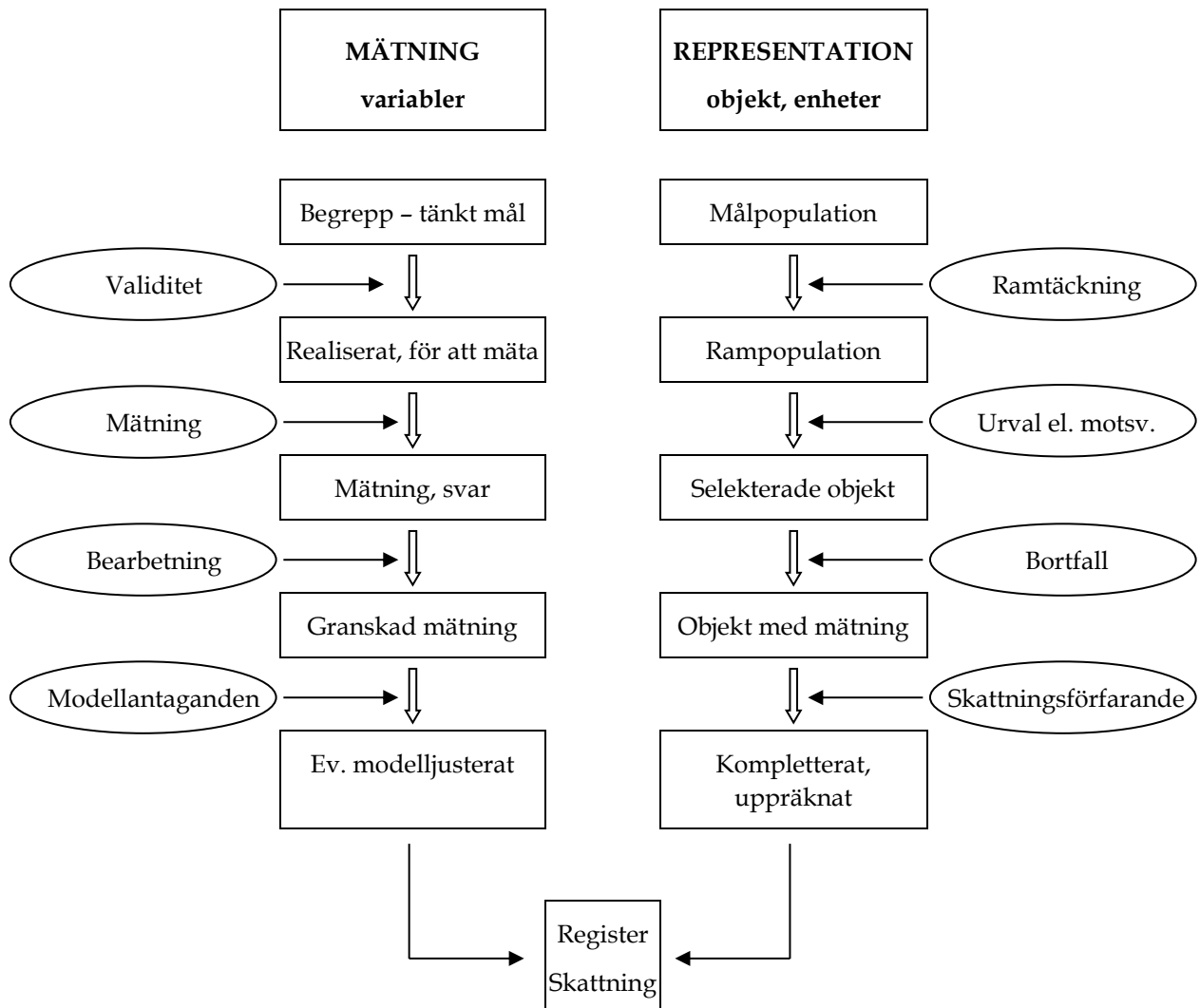
### 3 Väsentliga steg och osäkerhetskällor

Avsikten med detta kapitel är att beskriva var osäkerhetskällor finns och därmed hur osäkerheter uppstår och kan motverkas. Även kvalitetskomponenterna jämförbarhet och sam användbarhet ingår i någon mån. Utgångspunkt för diskussionen är en figur, som har blivit mycket använd.

#### 3.1 En illustration och dess betydelse

Figur 3.1 finns i många varianter. Den källa som brukar anges är boken *Survey Methodology* av Groves et al. (se Groves, Fowler, Couper, Lepkowski, Singer och Tourangeau, 2009, sidan 48). Ett par modifieringar har gjorts här, se vidare nedan.

Figur 3.1. Osäkerhetskällor i statistikproduktionsprocessen



Benämningen "survey life cycle" är vanligt förekommande på engelska, bl.a. tillsammans med metadata och dokumentation. Bilden visar ett flöde, och den har två samverkande delar. Det är mätning med variabler på ena sidan, och på andra sidan finns det som kallas representation med objekt eller statistiska enheter samt populationer.

Figur 3.1 visar en grundsituation med en enda objekttyp. Den kan tyckas begränsad till direktinsamling för urval, men så är inte fallet. Det kan vara insamling (hämtning) från redan befintliga data, och det kan vara en selektion som tar med allt. Det finns varianter och vidareutvecklingar av figuren för att bättre beskriva situationer med administrativa data och situationer med matchning. Det beskrivs vidare nedan.

Figur 3.1 avviker något från originalet i ordval och stegen på slutet. Avsikten är att figuren ska passa bättre ihop med processkartan och med designvägledningen i övrigt. Ordet fel undviks på den övergripande nivån (mätning i stället för mätfel etc.). Möjligheten att bygga in modellbaserade skattningar av variabelvärden (på mikronivå) öppnas i sista steget på vänstersidan. Högersidan inkluderar både uppräknings-/skattningsförfarandet (som i originalbilden) och möjligheten att stanna vid ett register som slutprodukt.

Figuren börjar med begreppsbildning och mål, som sedan realiserar. Användarnas intressen ligger steget före figurens flöden. Det kan finnas (eller snarare uppstå) brister på båda sidor i förhållande till statistikens ändamål och statistikanvändarnas informationsbehov när målen utformas.

När utvidgningar görs för att belysa situationer med administrativa data kan det finnas flera aspekter att lyfta fram. Ingående begrepp styrs av de administrativa definitionerna, men de kan eventuellt transformeras. Även objekt kan behöva modifieras eller transformeras till statistiska enheter. Objekten kan anges som åtkomliga/tillgängliga ("accessible") respektive nådda/selekterade ("accessed"), vilket i figur 3.1 motsvarar rampopulationen respektive selekterade objekt.

Figur 3.1 visar, som sagt, en grundsituation med en enda objekttyp och mätning. Därefter följer ofta matchningar eller anpassningar till statistiken (eller båda), vilket kan påverka såväl variabler som objekt. Zhang (2012) ger en mer komplex figur som beskriver detta, vilket berörs nedan, främst i avsnitt 3.4.

## 3.2 Begrepp undersöks genom mätning av variabler

### 3.2.1 Inledning om figurens vänstra sida

Den ena huvuddelen av figuren handlar om variabler och mätning. Det kan gälla t.ex. välfärd, produktivitet eller nybyggnation. Diskussioner med användare om statistikbehov är inte med i figuren, utan de har redan förts. Det är bestämt vilka begrepp – koncept – som ska användas i statistiken. Steget från intressevariabler till målvariabler är taget; målvariablerna är definierade. Inför detta steg och samtidigt som diskussioner har förts med användare om intresse- och målvariabler, så har operationaliseringen förberetts, och där ingår observationsvariabler.

Efter detta följer flera steg till det slutliga observationsregistret och statistiken; mikro- och makronivåerna.

### 3.2.2 Validitet

Validitet kan användas som en måttstock på hur väl ett utformat, konkretiserat begrepp fångar ett efterfrågat, tänkt, kanske mer abstrakt begrepp. Termen validitet används inte så ofta på SCB. Som framgår i avsnitt 2.1 används snarare intresse, mål och observation.

Här handlar validitet om operationaliseringen, från mål till observationer, och hur väl målet kan skapas ur observationerna.

Steket från det avsedda begreppet till det praktiska genomförandet är inte trivialt, och det kan ge upphov till avvikelser – även i fall då observationsvariabeln till synes överensstämmer med målvariabeln. Kanske fångar observationsvariabeln inte hela begreppet, kanske något förbiseende görs, t.ex. för att något är självklart för statistikproducenten men inte för uppgiftslämnaren.

Det är rätt vanligt förekommande att observationsvariablerna väljs så att de inte är identiska med målvariablerna men ändå leder fram till dessa. Det kan vara en konstruktion som ska underlätta för uppgiftslämnaren, t.ex. genom filterfrågor eller genom att efterfråga delar i en summa var för sig. I sådana fall *härleds* målvariabeln, den beräknas genom en räkneregeln.

Det finns också situationer där målvariabeln beräknas med hjälp av modellantaganden, t.ex. om samband mellan variabler. Det kan gälla vid direktinsamling om målvariabeln är alltför svår eller betungande för uppgiftslämnaren och en modellbaserad skattning bedöms ge tillräckligt god kvalitet för målvariabeln. Detta kan också gälla då register med t.ex. administrativa data används.

Den första pilen i figur 3.1 visar att målvariablerna är väsentliga att formulera och operationalisera.

### 3.2.3 Mätning

I steget från den efterfrågade uppgiften till uppgiftslämnarens svar (motsvarande) kan mätfel uppstå till följd av missförstånd, misstag eller bristande kunskap. Det gäller för administrativa data, och det gäller vid direktinsamling. Datainsamlingsmetoden kan ha betydelse. Det kan t.ex. vara skillnad mellan att höra och att läsa en fråga. Vidare kan syftet med insamlingen påverka uppgiftslämnarens ansträngningar att lämna ett riktigt svar.

Såväl de ursprungliga mätfelen (vid uppgiftslämnandet) som de återstående (efter granskning) är intressanta. De ger information om eventuella problem i mätningen respektive om kvaliteten i data (tillsammans dessutom om granskningen, se nedan). Sådan information kan användas för undersökningen i sig och även för kommande undersökningar.

När administrativa eller andra befintliga data används är kunskapen om mätningen och eventuella svårigheter för uppgiftslämnarna mindre än vid direktinsamling. Den ska ändå uppmärksammas och beaktas.

### 3.2.4 Bearbetning

I detta steg ingår bearbetning på mikronivå, med kodning av öppna svar, granskning och rättning (t.ex. efter återkontakt) eller annan ändring, sambearbetningar med mera.

En jämförelse mellan ogranskade data (dvs. data före granskning) och granskade data visar vad som har ändrats, och effekter kan bedömas.

Justeringar på mikronivå kan vara mindre eller större. De kan behövas vid kombinerad datainsamling och när data integreras. Det kan finnas olikheter i objekt att beakta, se vidare avsnitten 3.3 och 3.4. Det kan vara motiverat att se integration av data som ett mer omfattande steg.

### 3.2.5 Modellantaganden

Om målvariabeln är alltför svår att mäta kan det vara motiverat att använda modeller. I detta steg (avsnitt) avses i första hand modellantaganden om mikronivån. Ett exempel från miljöstatistik är att samla in data om jordbruk och därifrån göra modellbaserade beräkningar av varje objekts påverkan på miljön genom utsläpp till luft och vatten.

### 3.2.6 Överväganden

Figuren visar vilka osäkerhetskällor som finns. Vid direktinsamling används ämneskunskap och mätteknisk kunskap för att bedöma vad som är rimligt att mäta med tanke på uppgiftslämnarna. Lämpliga

målvariabler formuleras med hänsyn till användarintressen, mät-  
möjligheter och eventuella modeller för olika slag av omräkning,  
transformation och "förädling".

När register (redan befintliga, t.ex. från administrativa data) används  
finns vissa principiella skillnader, främst att variablerna i registret är  
givna. Viss granskning är redan gjord, men med den ursprungliga  
användningen som utgångspunkt. Granskning behövs därför, och  
beräkning av nya variabler kan vara aktuell. Det sistnämnda gäller  
även vid direktinsamling.

Variabler kan inte betraktas isolerat från objekten. Kanske kan mät-  
ningen underlättas om fler objekttyper än målobjektens typ(er) an-  
vänds. Det kan hänga samman med vilka registeruppgifter som finns  
och med uppgiftslämnare vid direktinsamling.

### **3.3 Populationer av objekt som observeras och som skattas**

#### **3.3.1 Inledning om figurens högra sida**

Den andra huvuddelen av figuren handlar om objekt och populatio-  
ner. Det kan finnas flera typer av objekt, t.ex. individer och hushåll  
eller företag med arbetsställen. Även här kan intresse, mål och obser-  
vation behöva särskiljas. Det kan finnas "fel" i själva objekten, se  
avsnitt 3.4.

#### **3.3.2 Ramtäckning: rampopulation jämfört med målpopulation**

Målpopulationen är den population som statistiken avser. Det är vik-  
tigt att vara tydlig med den i både presentation och design. Referens-  
tiden för målpopulationen är ett exempel på detta. En statistisk  
undersökning kan ha flera målpopulationer, t.ex. om det finns flera  
typer av målobjekt.

En ram används för att såvitt möjligt "hitta" målpopulationen och  
"nä" den i datainsamlingen (register eller direktinsamling). Objekt-  
typen är för många undersökningar densamma i ramen som i mål-  
populationen, men det måste inte vara så. Två exempel på olikheter  
är (i) individer i ramen och hushåll i målpopulationen och (ii) företag  
i ramen och varor i målpopulationen. Det kan behövas flera steg från  
ramens objekt för att nå fram till "rätt" objekt, målobjekten.

Till ramen hör en rampopulation som har samma objekttyp som mål-  
populationen. Rampopulationen behöver beskrivas, särskilt då flera  
objekttyper ingår. Det görs med hjälp av det eller de steg som tas från  
ramens objekt. I exemplen ovan är rampopulationen de hushåll re-  
spektive de varor som ramen kan leda fram till.

Skillnader mellan rampopulationen och målpopulationen – som blir  
tydliga genom att dessa har samma objekttyp – är brister i täckning.



Rampopulationen kan ha över- och undertäckning: rampopulationen innehåller objekt som inte ingår i målpopulationen respektive saknar objekt som ingår i målpopulationen. (För företag är nedlagda och inte verksamma företag exempel på övertäckning. Se även avsnitt 5.3.1.)

Ramen bör, så långt det är möjligt, väljas och avgränsas så att rampopulationen har en god överensstämmelse med målpopulationen. Det har betydelse för statistikens kvalitet, och det underlättar skattningsförfarandet.

Det kan finnas dubletter (vilket har viss likhet med övertäckning), och det kan finnas brister då objekt identifieras. Det sistnämnda kan bli märkbart vid dataintegration, t.ex. genom att företag som till synes är desamma i två datamaterial i själva verket avser olika verksamheter. Detta kan bero på skillnader i avgränsning.

### 3.3.3 Urvals- och selektionsförfaranden

Det är inte säkert att alla objekt i rampopulationen ingår i själva undersökningen. Några exempel är följande. Det kan vara en urvalsundersökning, t.ex. att ett urval av objekt dras slumpmässigt från ramen. Även mer subjektiva förfaranden förekommer vid urval, t.ex. för varor vid prismätning då en ram saknas. Det kan vara en företagsundersökning där de minsta företagen medvetet tas bort genom så kallad *cut-off* och skattningen är modellbaserad, se avsnitt 6.4.1.

Ytterligare ett exempel, om än inte så vanligt, är arbete med en delmängd av ett befintligt register för mera detaljerade analyser, som sedan ska generaliseras. Selektionen kan t.ex. hänga samman med andra tillgängliga data.

### 3.3.4 Bortfall: saknade data

Vid direktinsamling är bortfall vanligt. Det gäller särskilt i frivilliga undersökningar. Bortfall förekommer även i register, t.ex. administrativa data. Saknade data (*missing data*) kan vara en mer naturlig benämning.

Det är intressant att förstå eller ana vilken mekanism som orsakar om det blir svar eller bortfall. Med ökad kunskap om mekanismen följer ökade möjligheter att motverka bortfallet eller hantera dess konsekvenser i skattningsförfarandet. Benämningen svarsbenägenhet används ibland för något som kan liknas vid en sannolikhet för svar på objektsnivå.

Efter detta steg finns en uppsättning objekt med mätning till grund för den avslutande behandlingen. Det kan finnas både partiellt bortfall och objektbortfall. I det förra fallet saknas värde(n) på en eller flera, men inte samtliga, variabler för ett svarande objekt. I det senare fallet saknas variabelvärden helt för objektet eller så var variabelvärdena så ofullständiga eller av så låg kvalitet att de var oanvändbara.

### 3.3.5 Skattningsförfarande

Detta steg går från de objekt som har mätning till slutresultatet som är ett slutligt observationsregister eller statistik (eller båda). Det kan handla om att komplettera informationen på mikronivå eller om att ta fram statistik genom skattning. Det senare kan vara en uppräknings från svarande till population. Återigen är det viktigt att notera att detta inte gäller enbart urvalsundersökningar utan även andra undersökningar.

Ett skattningsförfarande görs enligt någon princip (designbaserad, modellassisterad, modellbaserad etc.) och med hänsyn tagen till de ingående osäkerhetskällorna, åtminstone de betydelsefulla. Det är önskvärt att beräkna, uppskatta eller bedöma osäkerheten i den resulterande skattningen. Det är ju ett mål att denna osäkerhet (bias och varians, om en uppdelning är möjlig) är liten. Utformningen ska formulera principen och sträva mot målet; olika möjligheter prövas och den bästa väljas. Figur 3.1 kan vara ett stöd och ett åskådliggörande i detta arbete. Det finns dock aspekter som inte är omedelbart synliga, t.ex. kan valet av datainsamlingsmetod ha stor betydelse.

Även den vänstra delen av figur 3.1 ingår i skattningsförfarandet. I en "traditionell" urvalsundersökning är det naturligt att lägga ett steg med skattning från svarande till population på den högra sidan. Ett viktsystem med hänsyn till urvalsförfarande, bedömning av bortfall och användning av hjälpinformation (t.ex. kalibrering) är betydelsefullt. Ett steg Modellantaganden finns i figur 3.1 för att markera att även variabelvärden kan påverkas i ett skattningsförfarande med modeller.

Ett skattningsförfarande innebär en statistisk inferens; jämför avsnitten 2.3 och 5.1.

### 3.3.6 Överväganden

Ett objekt kan vara intressant i sig eller vara ett sätt att få tag i den sökta informationen, när register och ramar saknas. Exempelvis kan information om varor kanske inte fås "direkt", eftersom ramar saknas. Mätningar kan i stället fås via t.ex. tillverkande företag eller säljande butiker.

Tillgång till register och ramar med god täckning är önskvärt. Identitetsbegrepp är också viktigt, särskilt om data ska integreras (matchas). Vissa objekt är stabila medan andra är föränderliga, som hushåll och företag (t.ex. modifierad verksamhet, delvis andra arbetsställen).

När register (t.ex. från administrativa data) används är situationen principiellt likartad den vid direktinsamling. Termerna kan vara något annorlunda, som redan har nämnts. En vanlig skillnad ligger i

urvals-/selektionsförfarandet, eftersom befintliga data ofta har karaktären totalundersökning. Det är dock viktigt att vara observant på täckningsfrågor och på att data kan saknas.

Det kan också finnas skillnader mellan administrativa objekt och statistiska objekt. Det kan gälla definitioner av objekttyper som har något olika syften eller uppstår av olika skäl. Det kan gälla för enskilda objekt, t.ex. till följd av olika system för inrapportering/mätning. Det kan krävas en "transformation" för att komma från det ena till det andra. Det kan också behövas "transformationer" eller kompletteringar vid matchningar mellan register, särskilt om dessa kommer från olika källor eller tidsperioder. Återigen är täckningsfrågor väsentliga. Vad är mål och vad finns observerat? Många gånger finns ingen skarp gräns mellan täckningsbrist och saknade objekt. Zhang (2012) ger exempel.

### **3.4 Sammanföra data – harmonisering m.m.**

Figur 3.1 är en enkel illustration av en statistisk undersökning. Vad händer t.ex. om flera datainsamlingsmetoder ingår eller om flera datakällor ska integreras? Steg och tillvägagångssätt kan behöva byggas ut, och benämningar justeras. Något av detta har redan nämnts.

Zhang (2012) har en mer komplex variant av figur 3.1 på mikronivå. Det finns flera datakällor. Objekt i sådana källor transformeras till statistiska objekt (jämför FDB:s objekttyper; det finns administrativa och statistiska). Det finns inte bara täckningsbrister på den högra sidan, utan fel kan uppstå till följd av svårigheter med objekten, t.ex. för att de är sammansatta eller på annat sätt svåra att identifiera.

På vänstersidan kan det också behövas "transformationer", men där av variabler, som ska harmoniseras eller justeras på annat sätt till en helhet. Det har redan nämnts för kombinerad datainsamling.

Med statistiska system kommer data och undersökningar att (behöva) integreras alltmer – undersökningar kommer då också att designas mer för det.

### **3.5 Osäkerhetskällor – några kommentarer**

Figur 3.1 visar var de olika osäkerhetskällorna väsentligen finns. Den kan vara ett stöd i arbetet med att motverka eller reducera fel och osäkerheter samt att kompensera för dem i skattningsförfarandet.

Jämför listan över osäkerhetskällor i avsnitt 2.6 med de två sidornas steg i figur 3.1. Se även beskrivningar i kapitel 4 som pekar på viktiga samband mellan designen av delprocesser och kvalitetskomponenterna, t.ex. i figur 4.2.

För objekt finns ramtäckning, urval eller annan selektion och bortfall samt skattningsförfarandet. För variabler finns validitet (för begrepp), mätning (mätfel), bearbetning samt modellantaganden. En del av dessa steg med osäkerhetskällor finns på båda sidor, även om de bara är ritade på en sida, t.ex. bearbetning och modellantaganden. Skattningsförfarandet kan ses som en "motverkande konsekvens" till och på grund av osäkerhetskällorna. Det bidrar dock i sig till osäkerhet, beroende på hur framgångsrikt det är i sina antaganden.

På senare år har det konstaterats och debatterats, särskilt för undersökningar riktade till individer och hushåll, att osäkerheten till följd av bortfall inte nödvändigtvis har ett starkt samband med bortfallsnivån, se t.ex. Groves och Peytcheva (2008). En intensiv jakt för att få in svar från bortfallsobjekten kan bli dyrbar med liten effekt, och mätfelen kan öka, om jakten leder till att svars kvaliteten blir dålig för de in jagade objekten. Effekten av bortfallet beror på flera faktorer, bl.a. skattningsförfarandets (estimatorns) förmåga att gå från de svarande objekten till populationen, t.ex. med hjälp av kalibrering.

Se även avsnitt 3.6 som bl.a. beskriver systematiska och slumpmässiga skillnader.

Att jämföra data kan ge kunskap direkt eller uppslag till fortsatta jämförande studier. Det kan vara en jämförande analys i efterhand av direktinsamlade data och administrativa. Det kan vara jämförelser mellan undersökningar med likartade variabler eller mellan undersökningar med olika frekvens.

Det är naturligt att skattningsförfarandet "motverkar" osäkerheten, om det går, t.ex. genom att kompensera för bortfall eller för systematiska avvikelser i mätningen.

### 3.6 Betydelsen av definitioner och metoder

När är statistik jämförbar och sammanvändbar med annan statistik? Två nyckelord är **definitioner** och **metoder**; definitioner omfattar då även begrepp och klassifikationer. Motsvarande kvalitetsfråga finns för mikrodata (register), och svaret är likartat.

För att kunna göra jämförelser över tid eller mellan redovisningsgrupper behöver de definitioner som ingår i jämförelsen vara desamma: de som hålls konstanta, medan andra (t.ex. referenstid eller region) varieras. Motsvarande gäller även sammanvändbarhet av statistik; för att relationen mellan utgifter och inkomster ska kunna studeras med hjälp av statistik från två separata undersökningar, krävs t.ex. att undersökningarna definierar hushåll på samma sätt. Detta är kanske tydligast för jämförbarhet över tid: om en variabeldefinition ändras blir det normalt ett tidsseriebrott som behöver

hanteras, t.ex. mäts eller på något sätt överbryggas, för att det även i fortsättningen ska gå att göra meningsfulla jämförelser över tid.

Metoder kan ha en systematisk inverkan, t.ex. kan det finnas systematiska skillnader mellan olika datainsamlingsmetoder eller mellan olika skattningsmetoder. Faktorer som kan ha betydelse är vilka ramar som används och vid vilka tidpunkter ramarna framställs. Effekter av systematiska skillnader kan vara svåra att bedöma. De kan vara olika för nivåskattningar och förändringsskattningar.

Om det finns en slumpmässig variation i de statistikvärden som ska jämföras eller sammanvändas blir analysen mindre skarp; ju större variation desto mindre skärpa. Det är dock en helt annan typ av försvårande omständighet än systematiska skillnader av okänd storlek.

Det är således viktigt att – om möjligt – hålla isär vad som är systematiskt och vad som är slumpmässigt. Se vidare bl.a. avsnitt 6.7.

Om två eller flera datamaterial ska matchas är det på samma sätt viktigt att se dels hur definitioner, begrepp och klassifikationer är utformade, dels vilka skillnader i metoder som eventuellt finns.

## 4 Utforma och utvärdera – översikt

Syftet med det här kapitlet är att:

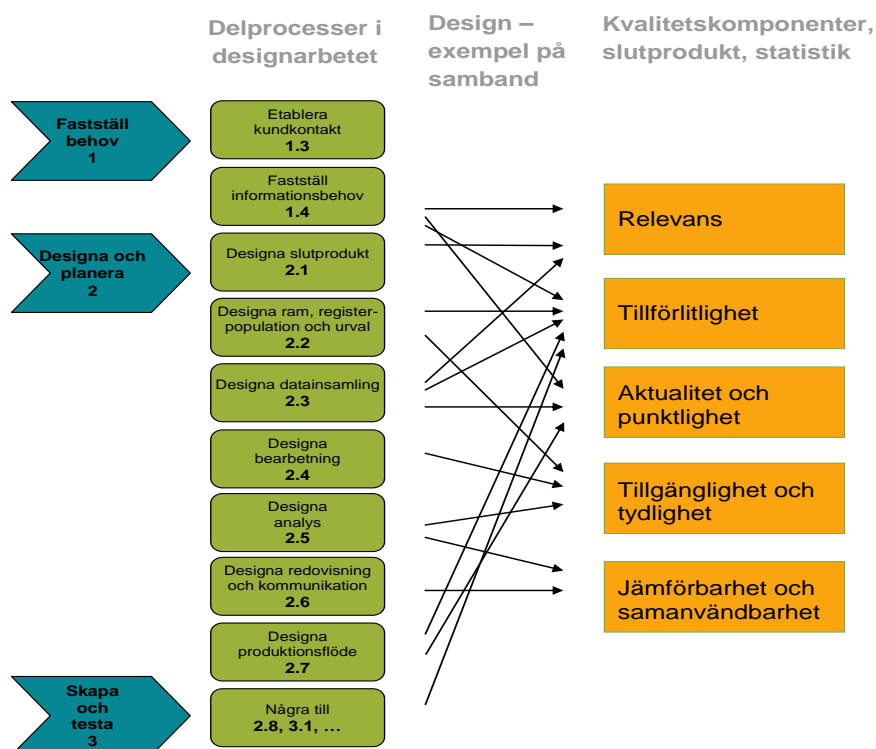
- Ge en översiktlig beskrivning av de tre inledande utformningsprocesserna 1-3 och den avslutande processen 8.
- Beskriva delprocessers koppling till statistikens kvalitet.
- Lägga en grund för förbättringsarbete.

Se även avsnitt 5.7 för olika aspekter då helheten designas.

### 4.1 Inledning med delprocesser och kvalitetskomponenter

Termen *utforma* används som samlingsbenämning på de tre förberedande processerna i statistikproduktionsprocessen: 1 *Fastställ behov*, 2 *Designa och planera* samt 3 *Skapa och testa*. Dessa tre processer utgör tillsammans med process 8 *Utvärdera och återkoppla* ett förbättringshjul – inte bara för en löpande undersökning utan även för engångsundersökningar från vilka det går att lära för framtida likartade undersökningar. De fyra processerna beskrivs kort i avsnitten 4.2–4.5.

Figur 4.1 Samband mellan designprocesser och statistikens kvalitet; exempel (där Relevans främst handlar om statistikens innehåll)



Figur 4.1 har syftet att visa komplexiteten i sambanden mellan undersökningsdesign och statistikens kvalitet. Pilarna visar exempel på samband mellan enskilda designprocesser och huvudkomponenterna i statistikens kvalitet. Betoningen ligger på design, men inte bara process 2 utan även processerna 1 och 3 har tagits med.

Är de fem huvudkomponenterna för statistikens kvalitet lätta eller svåra att:

- mäta eller åtminstone uppskatta?
- förklara?

Det är relativt lätt i båda avseendena för de tre huvudkomponenterna Relevans (med statistikens innehåll), Aktualitet samt Tillgänglighet och tydlighet. Tillförlitligheten totalt är svår i båda avseendena; med bättre kvantitativa mått skulle det också vara lättare att förklara. Jämförbarhet och sammanvändbarhet intar ett mellanläge. Det är lätt att konstatera och beskriva vissa brister, t.ex. tidsseriebrott till följd av en lagändring eller olikheter i täckning som omöjliggör en sammanvändning. Det kan dock vara svårt att veta om och hur mycket en metodändring betyder och att förutse alla önskvärda sammanvändningar.

Oavsett hurdan statistisk undersökning det är (t.ex. registerbaserad eller direktinsamling; hel eller del) finns det tre huvudsituationer för utformning:

- A. En ny statistisk undersökning, som ska göras en eller flera gånger. En utredning behövs där informationsbehov, undersökningsdesign och produktionssystem ingår.
- B. Omdesign; göra en översyn av befintlig(a) undersökning(ar). En utredning behövs; jämfört med föregående punkt finns praktisk erfarenhet att dra nytta av.
- C. Ständiga förbättringar; ett arbete med utvärdering och återkoppling. Det är ett löpande arbete.

## 4.2 Fastställ behov

### 4.2.1 Allmänt

Syftet med processen *Fastställ behov* i statistikproduktionsprocessen är att identifiera behov som finns eller kan komma att uppstå hos kunder och användare; behov som kan tillgodoses.

Enklast – i bemärkelsen renodlad – är situationen vid en direktbeställning av ett uppdrag av typen *ad hoc*. Där diskuterar SCB som producent direkt med en enda kund som har ett mål för undersökningen, ett mål som är mer eller mindre tydligt formulerat. Uppgiften i diskussionen är att gemensamt klara ut vilket sakproblem som ska belysas och hur detta sakproblem översätts till en statistikprodukt vars egenskaper preciseras i dialog.

En liknande situation är när en SAM (statistikansvarig myndighet) beställer en statistisk undersökning för att producera statistik inom sitt ansvarsområde. SAM har rollen som beställare/kund i diskussionen med SCB. I SAM-rollen ingår arbete med att väga samman olika användarbehov, att göra avvägningar och prioriteringar samt att definiera behovet av officiell statistik på området. I både denna och den föregående situationen har SCB som producent en enda kund/användare att diskutera med. Denna kan förväntas ha en god kännedom om statistikbehoven.

En annorlunda situation råder på de statistikområden där SCB är SAM. Kraven på SCB är då mer omfattande, för SCB ska kunna väga och sammanjämka olika användarintressen till en helhet för statistiken inom ansvarsområdet. Någon form av behovsanalys behövs för detta. Här är en löpande omvärldsbevakning ett viktigt instrument. Oftast finns någon form av infrastruktur etablerad t.ex. i form av användarråd eller regelbunden "uppsökande verksamhet" (möten med viktiga användare eller användargrupper). Till viktiga användare räknas EU och andra internationella byråer/organisationer som kan ha tvingande krav eller rekommendationer för statistiken.

Det är i samtliga dessa fall viktigt att diskutera och prioritera produktens kvalitet med dess komponenter och att då beakta kostnader. Det behövs uppenbarligen för nya statistiska undersökningar och vid omläggningar, men det är väsentligt också vid löpande justeringar. Ett växelvis samspel mellan fastställande av behov och undersökningsdesign är viktigt avseende både kvalitet och kostnader. Denna process för behov har kopplingar till alla kvalitetskomponenter.

Kvalitetsbegreppets komponenter ger en struktur för att fånga, diskutera och prioritera olika användarbehov. Användarna och deras värderingar ändras över tid, och producenten bemödar sig om att anpassa sin produkt till önskemål och omvärldsförändringar; se avsnitt 7.1.

Det är viktigt att beakta kostnader för hela statistikproduktionsprocessen. Det gäller t.ex. förberedelsearbete och för löpande undersökningar ingår underhåll och förbättringar. Det finns rutiner att följa; se Verksamhetsstödet, process 1.5 *Förhandla och avtala*.

#### **4.2.2 Konceptualisering (begreppsbildning)**

Grundläggande frågor vid begreppsbildning rör byggstenarna i intresse- och målstorheter, ofta gäller det särskilt objekt och variabler som ska observeras/mätas. Arbetet med att operationalisera sakfrågor i statistiska storheter och variabler kallas för konceptualisering, se avsnitten 2.1 och 3.2. Detta arbete, som givetvis måste bedrivas i nära samverkan med kund eller representanter för de tilltänkta användarna av statistiken, måste ha kommit en bra bit på väg, innan det är meningsfullt att detaljutforma mätinstrument m.m.



## 4.3 Designa och planera

### 4.3.1 Allmänt

Syftet med processen *Designa och planera* är att utifrån ett fastställt behov göra val av tillvägagångssätt, metoder och verktyg samt att fördela de tillgängliga resurserna över produktionsprocessen. Designprocessen har karaktären överväganden vid skrivbordet; i den efterföljande processen 3 skapas och testas produktionsflödet och dess delar. Processen *Designa och planera* avser all produktion i statistikproduktionsprocessen.

Alla delprocesserna ska gås igenom vid design- och planeringsarbetet. Det finns många ömsesidiga beroenden att beakta. Utförandet av de följande genomförandeprocesserna ska förberedas så väl som möjligt i tidigt skede, före det operativa produktionsarbetet. Designarbetet ska i första hand fokusera på det som är betydelsefullt, och i andra hand finslipa detaljer. Det ska se till att processerna utformas så att de samverkar effektivt och kvalitetsmålen uppnås.

Resursfördelningen är betydelsefull att avväga i designen för att nå en så långt möjligt optimal och konkurrenskraftig relation mellan kvalitet och kostnad i undersökningen. Metoder finns för vissa moment inom delprocesser, t.ex. urvalsallokering. Omprövningar kan behövas under designarbetet, särskilt för nya statistiska undersökningar. En del designarbete görs redan då behoven diskuteras och fastställs, inklusive planering. Även arbetet i process 3 *Skapa och testa* kan leda till omprövningar i process 2.

### 4.3.2 Välja och allokeras

Oavsett hurdan undersökning ingår det att välja och allokeras:

- **Välja**, exempelvis:
  - Datakälla – befintlig(a) eller direktinsamling?
  - Datainsamlingsmetod
  - Urvalsmetod
  - Skattningsmetod (ska beaktas för all statistik)

Kommentarer till val:

- Vissa val har "förstagångskaraktär", men de kan omprövas, t.ex. datakälla eller datainsamlingsmetod.
- Andra val kan (enkelt) ändras i en löpande undersökning, t.ex. kontaktstrategi eller skattningsmetod.
- Om informationsbehovet kan tillgodoses genom befintliga data ska dessa användas och inte en direkt datainsamling.
- Valen ska fungera väl tillsammans, t.ex. finns en stark koppling mellan metoderna för urval och estimation.
- Skattningsmetoden beror av många faktorer, t.ex. hur data har "uppkommit" samt principer och antaganden som behövs för att komma från data till statistik.

- **Allokera**, fördela resurser/insatser, exempel:
  - Bortfallsuppföljning kontra urvalsstorlek kontra granskning kontra ...
  - Avstämning av källor kontra modellering kontra granskning kontra ...
  - Personresurser – över processer och med hänsyn till olika kompetenser ...

Kommentarer till allokering:

- Allokeringen är föremål för omprövning.
- Samla in och använd relevanta processdata om kvalitet och kostnader.
- Allokering görs för statistiksystem, produkt, processer, kompetenser; i verksamhetsplanering och fortlöpande.

### 4.3.3 Designprocesser och kvalitetskomponenter

Flertalet av designprocesserna har en direkt koppling till någon av de kommande genomförandeprocesserna. Här nedan ges en kort översikt över de åtta delprocesserna i *Designa och planera* med viktiga frågeställningar och med exempel på vilka kvalitetskomponenter som de främst påverkar. Resonemangen är en viss konkretisering av figur 4.1 och samtidigt ännu mer av en förenkling.

Figur 4.2 har formen av en tabell (matris) med avseende på dels delprocesser med viktiga frågeställningar, dels huvudkomponenter i kvalitetsbegreppet. För tillförlitlighet har viktiga osäkerhetskällor tagits med. Det är både för frågeställningar och för kvalitetskomponenter kraftiga "urval" som gjorts för att få en koncis beskrivning. Sammanfattningsvis, och med det omvända perspektivet, framgår vad som är viktigt för de olika kvalitetskomponenterna.

- Relevans med statistikens innehåll:
  - Slutprodukt.
- Tillförlitlighet:
  - Variabler och mätning.
  - Objekttyp(er) och ram(ar).
  - Bearbetning och modellantaganden.
- Aktualitet och punktlighet:
  - Befintliga data eller direktinsamling.
  - Databeskrivningsmetod vid direktinsamling.
  - Tidsplanering.
- Tillgänglighet och tydlighet:
  - Val av redovisningsformer och presentationsformer.
  - Utformning av tolkningar och förklaringar.
- Jämförbarhet och sammanvändbarhet:
  - Definitioner; statistikens innehåll
  - Metoder; systematiska skillnader

Figur 4.2 Huvudsakliga frågeställningar och påverkan på kvalitetskomponenter

| Delprocess                        | Frågeställningar   | Kvalitetskomponenter                          |   |   |   |                                   |
|-----------------------------------|--|---|---|---|---|-----------------------------------|
|                                   |  | Relevans med Statistikens innehåll            | Tillförlitlighet                                    | Aktualitet och punktlighet                                  | Tillgänglighet och tydlighet                                    | Jämförbarhet och sam användbarhet |
| Designa (planera)                 | Huvudsakliga   |   |   |   |   |                                   |
| Slutprodukt                       | Tabellplan.<br>Slutliga observationsregister.  | Val styr.                                     |   |   |   | Val påverkar genom definitioner.  |
| Ram, registerpopulation och urval | Beakta målpopulationen.<br>Sök ram; ett eller flera steg. Ev. urval.   |   | Ramtäckning,<br>ev. urval.                          |   |   |                                   |
| Datainsamling                     | Befintliga register?<br>Datainsamlingsmetod(er),<br>kontaktstrategier,<br>skiss av mätinstrument.                      | Val påverkar.                                 | Mätning,<br>bortfall,<br>modell-<br>antaganden.     | Befintliga data? Datainsamlingsmetod(er)?<br>Tidsplanering. |   | Val påverkar genom metoder.       |
| Bearbetning                       | Bortfallskompensation,<br>hjälpinformation,<br>kodning, granskning,<br>imputering, komplettering,<br>vikter, outliers. |   | Bearbetning,<br>bortfall,<br>modell-<br>antaganden. |   |   | Val påverkar genom metoder.       |
| Analys                            | Skattningar,<br>säsongrensning,<br>outputgranskning,<br>röjandekontroll.   | Val kan öka.<br>Röjandekontroll kan begränsa. |   |   | Tolkning,<br>förklaring.  |                                   |
| Redovisning och kommunikation     | Publicera?<br>Redovisningsformer?<br>Förberedelser?<br>Gallring och arkivering.  |   |   |   | Redovisningsformer.<br>Publicering,<br>tolkning,<br>förklaring. |                                   |
| Produktionsflöde                  | Metoder, verktyg,<br>arbetsformer.   |   | Kan påverka.  | Kan påverka.  |   |                                   |
| Produktionsomgång (plan)          | Rutiner, resurser, extra marginaler?   |   | Påverkar.   | Kan påverka.  |   |                                   |

Redan i denna förenklade uppställning är det tydligt att designen har stor betydelse för genomförandet och för den resulterande kvaliteten. Statistikproduktionen bör vara vad som kan kallas designdriven.

I registerbaserad statistikproduktion brukar termerna ram och ram-population sällan användas fast de principiellt är relevanta. Termen registerpopulation har därför införts: en registerpopulation skapas som är nära målpopulationen, se Verksamhetsstödet process 2.2.

#### 4.4 Skapa och testa

Syftet med process 3 *Skapa och testa* är att se till att mätinstrument, metoder, verktyg och produktionsflöde finns och fungerar, i första hand i den aktuella produktionsomgången.

Att skapa kan här, t.ex. för ett verktyg, innebära att ta fram något nytt, alternativt att vidareutveckla, anpassa eller konfigurera något befintligt. Standardverktyg ska användas om det är möjligt. Anpassningar kan behövas, och viss utveckling från tidigare produktionsomgångar kan vara motiverad. Den utveckling som sker inom ramen för en produktionsomgång är normalt liten – större utveckling sker i form av utvecklingsprojekt.

Mätinstrument är förstås undersökningsspecifika, men de påverkas av standarder, variabelsamordning m.m. De hänger nära samman med designen av datainsamlingen, som i sin tur hänger samman med diskussioner och preciseringar av statistikbehov.

Att skapa och att testa är två olika aktiviteter med ett gemensamt syfte. Det kan vara praktiskt, eller till och med nödvändigt, att tester görs parallellt med skapandet. Test ger ofta återkopplingar. Det kan behövas flera vändor innan mätinstrumentet eller verktyget uppfyller de kvalitetskrav som ställts.

Det kan vara en investering att göra en pilotstudie av svårförutsägbara delar, eller av hela undersökningen, innan denna sjsätts. Det kan likartat vara en investering att göra ett inbäddat experiment innan en undersökning ändras. Se avsnitten 5.6.1 respektive 5.6.2.

Ytterligare aspekter här är produktionsflöde, system och kvalitets-säkring. Det är väsentligt hur produktionsflödet ser ut och vilka kopplingar som finns mellan det egna produktionssystemets komponenter och till andra produktionssystem. Det gäller, t.ex. när korrigeringar kan behöva göras, att kunna återskapa tidigare produktionsomgångar och resultat. Det är en form av spårbarhet. Det gäller att ha en genomtänkt plan för vad som ska sparas och hur. Ett annat exempel på vad som ska sparas är ogranskade data (dvs. data före granskning), eftersom dessa krävs för att kunna studera effekter av granskning.

Sammanfattningsvis innehåller process 3 *Skapa och testa* de tre huvuddelarna nedan som är viktiga för kvalitetssäkring och särskilt för de två kvalitetskomponenterna tillförlitlighet och aktualitet.

- Mätinstrument, och därmed variabler och mätning.
- Verktyg och produktionsflöde praktiskt, vilket kan beskrivas som möjliggörare och kvalitetssäkring.
- Pilotstudier och experiment.

#### 4.5 Utvärdera och återkoppla

Syftet med process 8 *Utvärdera och återkoppla* är att ge förutsättningar för ett systematiskt och effektivt förbättringsarbete, särskilt inom undersökningen – men också för andra undersökningar och arbetet med gemensamma metoder, verktyg etc. Syftet är att

- ge en bedömning av statistikproduktionsprocessens resultat jämfört med plan samt av effekter av vidtagna åtgärder
- föreslå åtgärder för att förbättra statistikproduktionsprocessen
- återkoppla till berörda för att bidra till ett systematiskt och effektivt förbättringsarbete.

Återkopplingen är intressant för processutförarna, t.ex. datainsamling, granskning och analys. Den är intressant också för produktansvarig tillsammans med andra kompetenser i ett *team* (avsnitt 7.5.2). Det generella arbetet med processer behöver också återkoppling, t.ex. från rapporter med utvärdering och åtgärdsförslag. Processvariabler kan också vara en informationskälla.

Redan i arbetet med att utforma undersökningen ska en (grov) plan tas fram som visar vilken information som ska mätas och sparas till utvärderingen.

Processvariabler ger kvantitativ information, och de kan mätas under processens förlopp eller efter avslutad produktionsomgång. Syftet här är uppföljning och utvärdering. (Processvariabler kan dessutom användas för styrning och kvalitetskontroll under statistikproduktionsprocessen, de fyra genomförandeprocesserna 4–7.)

Erfarenheterna kan vara sammanfattande kvantitativa mått, t.ex. på bortfall eller täckning. I många fall kan erfarenheterna vara av kvalitativ karaktär, t.ex. synpunkter från kunder, processutförare och uppgiftslämnare. Det kan exempelvis vara avvikelser från undersökningsplanen och orsakerna, upplevelser och synpunkter från processutförare och uppgiftslämnare angående t.ex. tolkningsproblem eller uppgiftslämnarbördan, eller synpunkter från intervjuare. Två exempel är loggbok och datainsamlingsdokumentation, se Verksamhetsstödet process 2.8.2 respektive 4.4.9.

En granskardebriefing är en gruppdiskussion mellan personer som arbetar med granskningen i en specifik undersökning om erfarenheter. Metoden används främst för företagsundersökningar och undersökningar om offentlig sektor. Den kan ge information om

mätning, granskning och uppgiftslämnarbörda. En granskar-debriefing kan t.ex. användas som underlag till en efterföljande expertgranskning av mätinstrumentet. Den kan även användas i kombination med t.ex. kognitiva test för att ge en bild av vilka problem som finns i mätinstrumentet och var fokus för testet ska ligga. Motsvarande diskussion kan hållas mellan intervjuare. De har ofta en god uppfattning om vad som fungerar och vad som är problem med ett mätinstrument.

I vissa fall kan speciella kundstudier övervägas som ett sätt att skaffa sig utvärderingsinformation.

En viktig typ av processvariabler är variabler som avser kostnader eller tidsåtgång för olika produktionsmoment. Sådana variabler kan dessutom flera likartade undersökningar dra nytta av. Processdata avser ofta, men inte bara, datainsamlingen, som är en del av statistikproduktionsprocessen där det kan finnas mycket att lära och förbättra. Det kan t.ex. gälla kontaktstrategi och adaptiv design, se t.ex. avsnitten 5.4 respektive 5.3.4.

Den erhållna produktkvaliteten ska jämföras med den designade och planerade; så gott det går. Vissa kvalitetskomponenter är enkla att mäta och studera (som aktualitet), medan andra är svåra (som tillförlitlighet); se avsnitt 4.1.

Kunskapen som fås genom de olika informationskällorna ska dokumenteras och göras tillgänglig för processutförare och andra som är berörda. Slutligen ska observationerna sammanställas och gemensamt analyseras som underlag för förbättringsåtgärder.

Rapporten från utvärderingen är i en löpande undersökning ofta en uppdatering av en tidigare rapport. Den föreslår åtgärder och anger för varje åtgärd förväntad effekt, hur effekten ska mätas, effektens uppskattade kostnad och dess prioritet. Den noterar även, utan detaljer, åtgärder som är önskvärda men orealistiska i ett kortare tidsperspektiv samt förslag för processer och andra undersökningar.

En konkret åtgärdsplan formuleras med utarbetade förslag i första hand på kort sikt. Se även avsnitt 7.5.

Process 8 *Utvärdera och återkoppla* kan, särskilt för löpande undersökningar, även beskrivas på följande sätt. Arbeta vid behov med kvalitetskrav som inte uppfylldes. Ta ställning till tänkbara effektiviseringar och möjliga kvalitetsförbättringar inom budget (ev. höjd budget, om det har motiverats). Beakta möjligheten till generellt lärande från genomförandet och eventuella studier (experiment, kvalitetsstudier etc.).

## 5 Principer, resonemang och stöd för design

En sammanhållen teori för design av statistiska undersökningar finns inte (ännu). Däremot finns en rad olika teorier – med olika syften och för olika delprocesser – som kan användas var och en för sig eller tillsammans i olika delprocesser. Det kan gälla val av metod(er) för en process eller för flera processer tillsammans. Det kan gälla styrning och uppföljning av processer. Kapitel 5 beskriver sådan teori med resonemang och stöd. Det efterföljande kapitlet 6 är likartat, men det är mer inriktat på praktiska situationer. Båda beskriver angreppssätt och metoder.

Kapitel 5 har sju avsnitt som varierar i karaktär och typ av innehåll. De har flera inbördes samband, vilket framgår i avsnitten. Hänvisningar ges till processkartan och Verksamhetsstödet. Allt kan dock inte listas varje gång. Det gäller särskilt processerna 2.1 *Designa slutprodukt* och 2.7 *Designa produktionsflöde*, som är mer övergripande än de övriga designprocesserna och därmed ofta återkommande. Det gäller även process 1.4 *Fastställ informationsbehov* och process 2.8 *Planera produktionsomgång*.

### 5.1 Principer vid statistisk inferens

Syftet med detta avsnitt är dels att fördjupa det tidigare avsnittet 2.3 och dess figur 2.1, dels att lägga en grund för avsnitt 5.2 om urval och estimation. I det här avsnittet finns ingen begränsning till urvalsundersökningar. I förhållande till processkartan är frågeställningarna övergripande; de ingår bl.a. i flera av designens delprocesser.

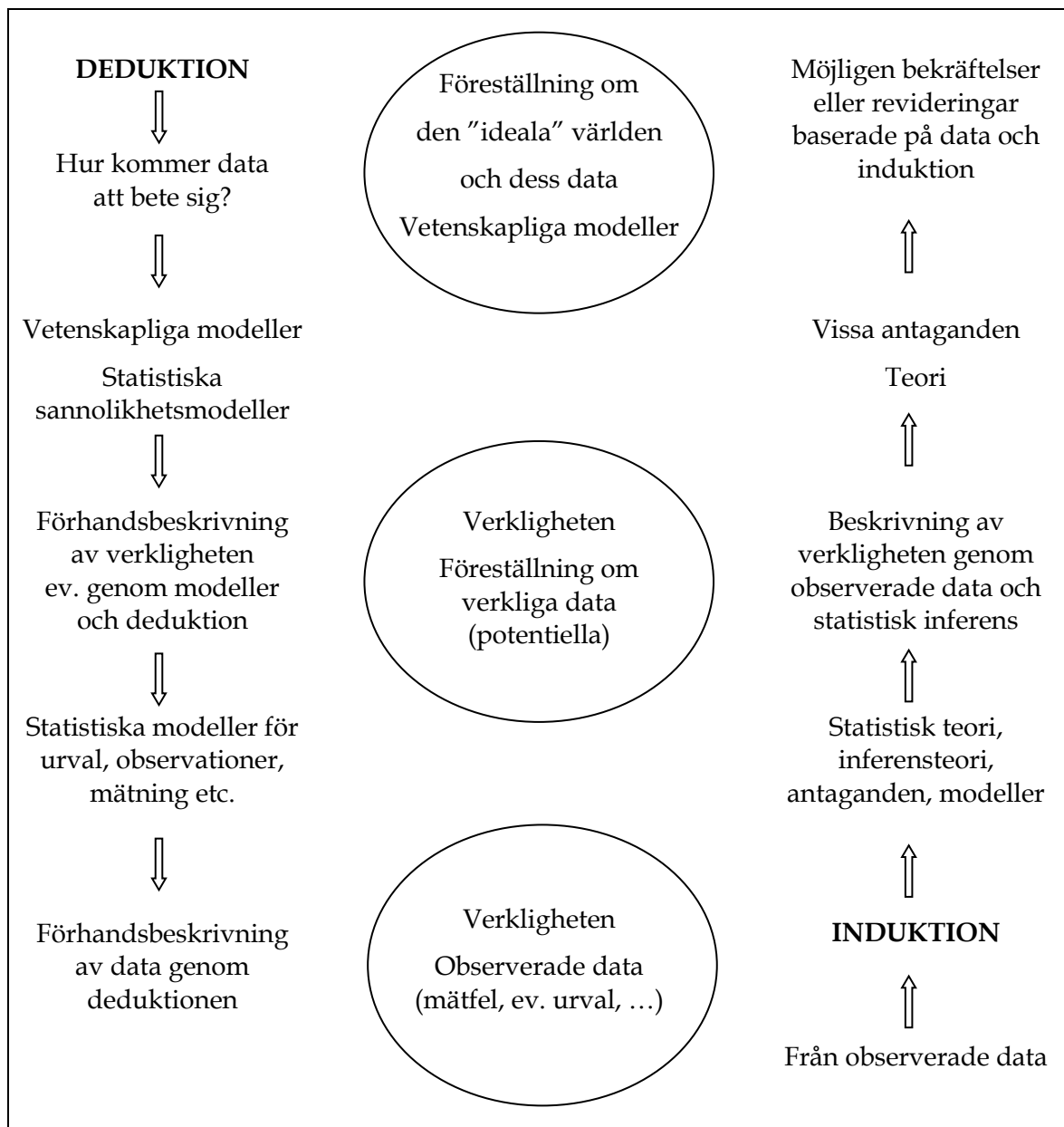
Statistisk teori har två huvuddelar, nämligen sannolikhetslära och statistisk inferensteori (med skattning och test). Detta kan ses som statistiska beräkningar åt "två håll" som figurerna 2.1 och 5.1 visar. Det ena hållet räknar "nedåt" och det andra "uppåt". Det ena bedömer och beskriver i förväg hur det kan tänkas bli, och det andra går från observerade data till något mer allmängiltigt.

Den statistik som SCB framställer avser ofta populationer, t.ex. befolkning, fastigheter eller företag i Sverige. När statistikanvändare gör jämförelser och andra analyser kan de vara inriktade på sådana ändliga populationer och på förhållanden i dessa populationer. Det kan gälla samband av olika slag. Ibland använder de modeller i och för analyser. Det kan t.ex. vara en ekonometrisk modell eller en regressionsmodell med ett slumpmässigt (stokastiskt) inslag i modellen. Sådana modeller kan ha en vidare karaktär än just den ändliga populationen. Det har ofta även modeller som används i modellassisterade och modellbaserade skattningar.

Figur 5.1 är mer omfattande än figur 2.1 genom att ytterligare en del har tillkommit utöver de tidigare. Denna översta del är mer abstrakt.

Här ingår t.ex. en tänkt slumpmekanism som en beskrivningsmodell för den ändliga populationen. Den översta ovalen är övergripande eller vetenskaplig, och den avser en abstrakt tänkbar värld. Denna kan vara intressant i sig eller vara ett hjälpmedel. Den mellersta ovalen avser verkligheten som helhet och den nedersta ovalen avser verkligheten med observerade data – som förut i figur 2.1.

Figur 5.1 Statistisk inferens, inklusive deduktion och induktion





Liksom förut finns det en vänstersida som går från det tänkta till förhandsbeskrivningen av data (sannolikhetsteorin) och en högersida som går åt andra hållet, dvs. från observerade data genom en statistisk inferens. Denna kan avse den mellersta ovalen eller den översta, beroende på statistikanvändarens intresse. SCB producerar i allmänhet skattningar avseende den mellersta ovalen.

Den vänstra sidan av figuren fyller en viktig funktion i utformningen av en statistisk undersökning. Det är tydligt för designen av urval (nästa avsnitt), men det gäller även andra osäkerhetskällor. Ju mer förhandskunskap, desto bättre möjligheter att göra en god utformning med hänsyn till tillförlitlighet och risker. Med ett sannolikhetsurval finns en slumpmekanism som statistikproducenten råvar över (det finns flera fördelar). Inbäddade experiment görs också kontrollerat efter förhandskalkyler.

Bortfallsmekanismen (på vänstersidan) styrs inte av statistikproducenten, men med en förhandskunskap om den mekanismen kan motverkande förhandsåtgärder och bortfallsuppföljning planeras med hänsyn till kvalitet och kostnader. Även skattningen (den "återvändande" högersidan) bör på något sätt beakta bortfallsmekanismen och dess troliga snedvridande effekter.

## 5.2 Ansatser för urval och estimation

### 5.2.1 Inledning

För urval och estimation finns teorier med tydliga kriterier, t.ex. en målsättning av typen minimera variansen eller minimera medelkvadratfelet. För många situationer med sannolikhetsurval och designbaserad inferens finns även formler som gör det möjligt att beräkna vad som är bäst eller åtminstone bra. Detta område är mer välutvecklat än många andra i teoretiskt avseende. Läs exempelvis i SCB (2008), Särndal, Swensson och Wretman (1992) eller Lohr (2010). I processkartan är, naturligt nog, främst processerna 2.2 *Designa ram, registerpopulation och urval* samt 2.5 *Designa analys* aktuella här.

Steget från urvalsdata till skattningar av populationsvärden har diskuterats mycket. Det finns flera principiella skolor. Ett par nämns här. Synsättet har betydelse även för hur urvalet ska utformas; urval och estimation ses tillsammans – detta kallas ofta för en strategi för urvalsundersökningen. Beskrivningar av olika synsätt på urval och estimation ges t.ex. av Brewer (1999, 2013) och Chambers (2011); Brewer (2013) är lite mer historisk och lättläst, och de båda andra artiklarna är lite mer tekniska. Båda författarna är positiva till en användning av modeller. Det finns många olika tillvägagångssätt, såväl i urvalsdesignen som i utformningen av estimator. Några specifika situationer beskrivs i kapitel 6.

### 5.2.2 Designbaserad ansats

En principiell ansats för att bilda skattningar från urvalsdata kallas designbaserad inferens; urvalsdesignen är grunden. I skattningen som räknar upp från urval till population multipliceras variabelvärden med inverterade urvalssannolikheter. Ett objekt som i ett slumpmässigt urval dras med exempelvis sannolikheten 0,1 multipliceras med 10. Det kan sägas representera tio objekt: sig självt och nio andra objekt. När skattningens osäkerhet beräknas så är det utifrån den slumpmekanism som urvalsdragningen innebär. Det finns teori för detta; hur urval och skattning ska utformas för att ge ingen eller liten skevhet och liten varians.

### 5.2.3 Designbaserad ansats med modellassistans

Modeller som beskriver samband mellan variabler kan användas i skattningar. Hur väl modellen beskriver verkligheten påverkar skattningens tillförlitlighet. Det måste finnas en princip för modellen användningen. Kalibrering är ett exempel på vad som brukar kallas en modellassisterad ansats. Urvalsmekanismen är då den viktiga för skattningen, medan hjälpinformation och modell ger ett stöd som kan förbättra precisionen, särskilt om modellen är en god beskrivning av variabelsamband.

I praktiken kompliceras en urvalsundersökning av flera osäkerhetskällor, t.ex. bortfall. Redan det gör att det grundläggande antagandet med urvalsmekanismen som "allena rådande" för skattningens utformning är otillräckligt. Det kan vara en möjlighet att arbeta med en så kallad svarsbenägenhet som varierar mellan olika grupper och bygga in denna benägenhet i skattningsförfarandet. En modellassisterad ansats använder antaganden och modeller som assistans till utgångspunkten design och urvalssannolikheter. Hjälpvariabler i en kalibrering är, som noterats i avsnitt 2.3.2, kraftfulla om de "fångar" variation i undersökningsvariabeln och i svarsbenägenheten. Det finns mer att beakta, se t.ex. SCB (2008).

En situation med bortfall skiljer sig från en "ren" urvalssituation genom att antaganden om mekanismen för svar/bortfall ingår. Det innebär att skattningsförfarandet blir beroende av en sådan modell. Exempelvis antas att svarsbenägenheten inte påverkas av undersökningsvariablernas värden; däremot kan hjälpvariablerna ingå i modellen, t.ex. genom svarshomogenitetsgrupper.

### 5.2.4 Modellbaserad ansats

Alternativt, jämfört med synsätt där undersökningsdesignen har stor betydelse, kan en modell som beskriver data med en inneboende slumpmekanism vara styrande för bedömningen av skattningens tillförlitlighet (osäkerhet). Modellen påverkar även skattningen i sig.

Några typiska situationer där modeller får en starkare roll än i designbaserade och modellassisterade skattningar:

- Små redovisningsgrupper; kallas ofta "small area" men det är inte bara för små geografiska områden. De enskilda skattningarna tar stöd av modeller för större områden eller grupper.
- Skattningar som görs mycket tidigt, då data är ofullständiga, och en modell behövs för att "fylla ut".
- En undersökning med så kallad *cut-off*, då t.ex. små företag inte undersöks alls. Skattningen bygger på en modell, typiskt skattningar av modellparametrar med hjälp av något större företag samt registerinformation om de små företagen.

Detta är tre exempel på vad som kallas modellbaserade skattningar. Tillförlitligheten beror även av modellen. Se avsnitt 6.4 för fler exempel och detaljer.

#### 5.2.5 Urval och modeller

I en urvalsundersökning ska strategin med urvals- och estimationsförfarande väljas med omsorg. Det kan, även med en designbaserad ansats, finnas skäl att använda en modell i designsteget för att förbättra allokeringen. Ett sådant exempel, som avser stratifiering, ges av SCB (2008, sidan 54).

Även om sannolikhetsurval har fördelar kan det bli nödvändigt med icke-sannolikhetsurval, t.ex. för att det inte finns en ram. Det är exempelvis typiskt för prismätningar (avsnitt 3.3.3).

En amerikansk förening, nämligen *American Association for Public Opinion Research* (AAPOR), har tagit fram rapport om icke-sannolikhetsurval som har sammanfattats i en artikel 2013 av Baker et al. (Baker, Brick, Bates, Battaglia, Couper, Dever, Gile och Tourangeau, 2013). Några noteringar:

- Det finns flera skäl till icke-sannolikhetsurval och många typer; webbpaneler med medverkan på eget initiativ är ett exempel.
- Syftet med undersökningen (*fit for purpose*), kvalitetskomponenter och kostnader är viktigt.
- Öppenhet om hur undersökningen är gjord är viktigt. Användare uppmanas vara observanta.
- Modellantaganden behövs i en undersökning, även med sannolikhetsurval. Modellerna bör hantera både urvals- och skattningsstegen.

Efter artikeln, som är ganska kort, följer fem kommentarer och ett svar.

### 5.3 Några resonemang för ramtäckning, urval och bortfall

Det här avsnittet kan ses som en fortsättning på det förra, en fortsättning som tar upp tre osäkerhetskällor. Särskilt ramtäckning är dock inte begränsad till att gälla urvalsundersökningar. I processkartan är främst de två processerna 2.2 *Designa ram, registerpopulation och urval* samt 2.5 *Designa analys* aktuella.

#### 5.3.1 Ramtäckning

Övertäckning och undertäckning har olika karaktär.

Objekt som är övertäckning ingår i rampopulationen men inte i målpopulationen. Att avgöra att så är fallet för ett objekt kan vara både lättare (med hjälp av svar eller andra uppgifter än ramens) och svårare (vid bortfall).

Objekt som är undertäckning ingår i målpopulationen men inte i rampopulationen. Undertäckningen är oftast okänd initialt, men information kan tillkomma med senare versioner av register och ramar.

I vissa fall kan studier av täckningsbrister göras i efterhand, t.ex. kan brister som beror på eftersläpning i rapportering uppskattas. Sådana erfarenheter kan ingå i skattningsförfarandet – med viss försiktighet eftersom de förstås inte behöver vara konstanta utan kan variera, t.ex. med konjunkturcykeln. Kalibrering kan vara en möjlighet.

Skattningsförfarandet ska ta ett samlat grepp om täckningsbrister: tydliggöra rimliga antaganden och vad de innebär. Detta görs ofta samtidigt med kompensationen för bortfall (avsnitt 5.3.2). Ett antagande om relationen mellan över- och undertäckning (t.ex. likhet) kan vara en möjlighet, om det bedöms rimligt. Information genom senare versioner av register kan vara ett annat tänkbart tillvägagångssätt.

Dubletter är en täckningsbrist (med viss likhet med övertäckning) som man, beroende på objekttyp, kan behöva vara observant på för att identifiera.

#### 5.3.2 Bortfallskompensation

Bortfall – saknade värden – förekommer på objektsnivå och partiellt. Det senare innebär att ett eller flera men inte alla variabelvärden saknas för ett objekt; det finns användbara data för objektet.

Vid partiellt bortfall är det vanligt att imputera värden på objektsnivå i stället för de saknade. Bortfall på objektsnivå kan hanteras genom vägning eller imputering. Att använda vägning för bortfallskompensation är vanligt för urvalsdata som ändå ska vägas ("räknas upp"); det kan användas även i andra fall. Imputering på objektsnivå förekommer i varierande omfattning. Valet kan bero bl.a. på tillgänglig information, objektets storlek, och om det är en totalundersökning.

Vilket förfarande som än väljs för att hantera bortfallet (kompensera för det) innebär det någon form av antaganden. Dessa kan vara explicita eller implicita. En fördel med explicita antaganden är att det blir lättare att bedöma rimligheten och påverkan på statistikens kvalitet. Det gäller även kalibrering. Sådana antaganden skrivs i form av en modell, exempelvis samband mellan variabler eller samband mellan värden vid olika tidpunkter. Det kan vid imputering t.ex. vara ett gruppmedelvärde eller en framskrivning av objektets variabelvärde föregående gång.

Valet mellan imputering och vägning kan bero av flera faktorer. Den resulterande tillförlitligheten är en viktig sådan. Det kan vid en imputering vara enklare att beakta flera möjligheter och ha en successiv regel för metodvalet. Att skatta tillförlitligheten kan vara svårare vid imputering. Om de imputerade värdena används som svar kan det bli en "optimistskattning" (t.ex. vid medelvärdesimputering).

Några metoder för imputering finns i Verksamhetsstödet, process 5.3 *Imputera för bortfall* (och design i process 2.4).

### 5.3.3 Balans och representativitet

Representativitet är en gammal tanke som framförts bl.a. av norrmannen Kiaer 1895 för urval (långt före ansatsen med slumpmässiga urval och designbaserade skattningar), se t.ex. beskrivningen av Brewer (2013). Tanken om representativitet har senare återkommit och gjorts mer stringent.

Det kan vara balans i urvalet (jämfört med populationen) som eftersträvas, se t.ex. Deville och Tillé (2004) eller Chambers (2011). En av tankarna med sådan representativitet är att även ett slumpmässigt urval kan bli extremt i något avseende som påverkar skattningen olyckligt; sådana extremer bör på något sätt undvikas.

En annan typ av representativitet gäller svarsmängden jämfört med urvalet. Denna typ kan mätas genom så kallade R-indikatorer; se avsnitt 5.3.4 om adaptiv design, som ger referenser. Syftet är att motverka att bortfallet snedvrider skattningarna. Även mått på balans, som jämför medelvärden för svarsmängd och urval, kan användas.

### 5.3.4 Adaptiv design

Termen *responsive design* (eller responsiv design på svenska) används i flera sammanhang, t.ex. för webbdesign och för undersökningsdesign. Här diskuteras undersökningsdesign.

På engelska används de två uttrycken *responsive design* och *adaptive design* med likartade betydelser. Det senare är mer generellt. Det förra var det som introducerades först av de båda uttrycken för design av statistiska undersökningar. Datainsamlingen har med en sådan design flera faser, där den första är lärande; det finns likheter med en

pilotstudie men utan uppehåll före huvudundersökningen. Särskilt benämningen *responsive design* förekommer på SCB, men här har i stället svenska ord valts: adaptiv design.

En design som är adaptiv – anpassningsbar – har en från början inbyggd förmåga att anpassa sig till omständigheterna under datainsamlingen och att differentiera mellan objekt i en direktinsamling. En adaptiv design ställer krav på processdata och på en utformning av de val och modifieringar som kan och ska göras. Det kan t.ex. gälla val av datainsamlingsmetod eller prioritering av vissa objekt i bortfallsuppföljningen. Syftet är att öka kvaliteten för en viss kostnad eller att minska kostnaderna.

Den adaptiva designen kan vara helt inriktad på kontaktförsök i bortfallsuppföljningen. Med kombinerad datainsamling kan den adaptiva designen handla om att börja med billigare datainsamlingsmetoder och sedan övergå till dyrare, som t.ex. telefon. För att en adaptiv design ska vara meningsfull måste det finnas resurser (personer och budget) så att förändringar är möjliga.

Mindre formella adaptiva metoder har ”alltid” funnits, t.ex. att öka insatsen om inflödet är lågt eller för strata med inga eller mycket få svarande objekt. Det är en fördel med en formalisering som visar på tidpunkter för omprövning eller de händelser som styr eller påverkar de aktiviteter som utförs. Denna formalisering till en adaptiv design ställer krav på förutseende av vad som kan inträffa.

Pilotstudier (avsnitt 5.6.1) kan ge kunskap och innebära en mer stabil design med mindre behov av att införa ändringar under genomförandet. De kan ge mer utrymme än en adaptiv design för att beakta gjorda erfarenheter.

Adaptiv design och balansindikatorer är, liksom kombinerad datainsamling, ett område där många studier görs för närvarande. Några referenser för den som vill läsa mer är Schouten, Calinescu och Luiten (2013), Beaumont, Bocci och Haziza (2014) samt Roberts, Vandenplas och Stähli (2014). Den sistnämnda har med kostnader. Lundquist och Särndal (2013) studerar en svensk situation, för undersökningen om levnadsförhållanden (ULF); mått på representativitet och balans ingår. Dessa fyra studier avser individundersökningar, vilket är vanligast. Ouwehand och Schouten (2014) studerar däremot R-indikatorer för företagsundersökningar; korttidsstatistik med direktinsamling eller administrativa data.

#### **5.4 Kombinerade metoder vid direktinsamling**

För datainsamling finns teorier och kunskaper om för- och nackdelar med olika enskilda metoder för direktinsamling (enkät, telefonintervju, besöksintervju etc.) i olika situationer (t.ex. ämne, kostnader och

tidsåtgång). Exempelvis är telefonintervjuer en datainsamlingsmetod som kan genomföras snabbt men som inte lämpar sig för alla typer av frågor och frågekonstruktioner.

En kontaktstrategi behövs redan för *en* datainsamlingsmetod. Det gäller t.ex. hur och när uppgiftslämnarna ska kontaktas, med vilket material och genom vilka kommunikationskanaler. Se Verksamhetsstödet, process 2.3.3.

Kombinationer av datainsamlingsmetoder vid direktinsamling har ökat av flera skäl, t.ex. för att minska kostnader och för att underlätta för uppgiftslämnare och få mindre bortfall. Benämningen *mixed mode* är relativt väl inarbetad. Den kan användas eller översättas med **kombinerade datainsamlingsmetoder** (men inte med blandad datainsamling, som låter mindre genomtänkt).

Det är mycket att beakta då en kombinerad datainsamling utformas.

- Ska frågeformuleringarna vara desamma eller ska de anpassas till var och en av datainsamlingsmetoderna?
- Kan det bli systematiska skillnader, och hur ska de i så fall hanteras?
- Ska metoderna användas parallellt med valfrihet eller sekventiellt med de(n) billigare först?
- Det finns exempel på att valfrihet mellan flera insamlingsmetoder har gett lägre svarsandel än en enda metod. (Referensen Persson (2013) finns i Verksamhetsstödet .)
- Går det att separera mätfel och metodeffekt? (Det är svårt vid valfrihet.)

När datainsamlingsmetoder kombineras ska det finnas en genomtänkt strategi, se Verksamhetsstödet, process 2.3 *Designa datainsamling*. Det finns även referenser där för den som vill läsa mer, i första hand Persson (2013). Det här är ett område som tilldrar sig stort intresse, och det skrivs mycket. Två artiklar som är belysande och något tekniska är Vannieuwenhuyze och Revilla (2013) samt Hox, De Leeuw och Zijlmans (2015). Även boken av Groves et al. (2009) och boken *Designing and conducting business surveys* av Snijkers, Haraldsen, Jones och Willimack (2013) har information.

## 5.5 Svarsprocesser

Teorier om hur svarsprocesser går till för olika typer av undersökningar och uppgiftslämnare har utvecklats inom beteendevetenskaperna. Som ofta skedde utvecklingen tidigare för individundersökningar än för företagsundersökningar.

Svarsprocessen beskrivs ofta med följande fyra steg: 1) förståelse av frågan, 2) återhämtning från minnet, 3) bedömning, 4) lämnande av

svaret. Svarsprocessen i en undersökning kan vara komplicerad och varje steg ställer kognitiva krav på uppgiftslämnaren. I alla stegen kan också misstag och mätfel uppstå. Källorna till mätfel kan delas in i fyra grupper: de som beror på respondenten, de som beror på instrumentet (frågeformuläret, frågorna, brev och instruktioner), de som beror på intervjuaren och de som beror på datainsamlingsmetoden.

Mättekniker utnyttjar teorier och erfarenheter för att undvika eller reducera mätfel (t.ex. minska svarsfelet till följd av svåra ord och minnesfel). Svarsprocessen eller mätprocessen kräver uppmärksamhet i utformningen vare sig den handlar om att hämta uppgifter från register, visuellt observera eller ställa frågor.

Det är uppenbart viktigt att ha kunskap om svarsprocesser och underlätta uppgiftslämnandet på olika sätt. Detta är avgörande för att förebygga mätfel och har även betydelse för att minska bortfall, både partiellt bortfall och objektbortfall. Risken att en uppgiftslämnare avbryter sitt svarande ökar när den kognitiva bördan är hög. Utöver väl designade mätinstrument kan kombinerade datainsamlingsmetoder vara en sådan underlättande metod, se avsnitt 5.4.

I processkartan är process 3.1 *Skapa mätinstrument* närliggande, och det finns information i Verksamhetsstödet.

För den som vill läsa mer om svarsprocesser finns t.ex. böcker av Tourangeau, Rips och Rasinski (2000), främst för individundersökningar, samt Snijkers et al. (2013), främst för företagsundersökningar.

## **5.6 Pilotstudier, bl.a. experiment**

Pilotstudier är hjälpmedel för att göra val och ställningstaganden. De finns i Verksamhetsstödet beskrivna i process 3.6 *Genomför pilotstudie*. Experiment är en form av pilotstudie.

### **5.6.1 Pilotstudier**

För en effektiv design behövs, paradoxalt nog, tillgång till information som normalt inte finns förrän undersökningen är avslutad. Det kan gälla information om populationens utseende, hur uppgiftslämnarna reagerar på specifika frågor eller hur ny teknik påverkar svarsprocessen. Denna paradox kan hanteras med pilotstudier av olika slag. Syftet med en pilotstudie är att *i förväg* inhämta kunskaper som möjliggör en bättre huvudundersökning. Det finns olika typer av pilotstudier, och valet beror av syfte med studien och styrkan i informationsbehoven.

En pilotstudie kan vara formell eller informell. Skillnaden ligger i möjligheten att generalisera resultaten från studien till en population.



En formell studie kan vara en separat undersökning eller ett planerat experiment, utformade så att resultaten kan generaliseras.

En informell studie saknar generaliseringsmöjligheten. Den är mer explorativ. En genomtänkt studie på ett litet väl planerat subjektivt urval kan ge mycket information till en låg kostnad, t.ex. avslöja svagheter som kan rättas till. En informell studie kan således vara mycket kraftfull. Om en ansats inte fungerar under gynnsamma omständigheter, kan den inte användas alls.

För både formella och informella studier gäller att de kan avse hela produktionsprocessen eller enbart en del.

Mycket behöver prövas i praktiken, t.ex. begrepp inom nya ämnesområden, ny teknik och förändringar i mätinstrument. Det gäller både nya undersökningar och ändringar i löpande undersökningar, det senare för att undvika eller kontrollera tidsseriebrott. Inför aldrig något nytt utan någon form av test. Det kan vara ett så kallat användbarhetstest, där t.ex. ett framtida verktyg testas av potentiella användare som får pröva hur lättbegripligt och praktiskt det är. Det är viktigt att försöka förutse risker och förebygga dem, t.ex. genom att först göra ett experiment eller annan pilotstudie. Det kan vara en ändring i datainsamlingsmetod eller i kontaktstrategi.

Hur pilotstudier görs finns beskrivet t.ex. i boken av Snijkers et al. (2013, särskilt sidorna 275–276).

### **5.6.2 Experiment, särskilt inbäddade**

Avsikten med ett experiment är normalt att pröva en hypotes, t.ex. att valet av datainsamlingsmetod inte påverkar uppgiftslämnarnas svar. Det kan gälla en engångsundersökning eller en löpande undersökning där tidsserier är viktiga.

Typiskt för ett experiment i en urvalsundersökning är att urvalet slumpmässigt delas in i två eller flera delurval enligt en experimentell design. Det finns normalt en kontrollgrupp och en eller flera behandlingsgrupper (delurval). Behandlingen(-arna) kan t.ex. gälla förbrev, datainsamlingsmetoder eller kontaktstrategier. Försiktighet måste iaktas vid planering för att undvika att även andra faktorer påverkar och för att se till att experimentet är genomförbart i praktiken. Dessutom ska hypoteser formuleras i förväg, och beräkningar ska göras för att se vilka skillnader som kan betraktas som signifikanta med den befintliga urvalsstorleken. Annars kan experimentet visa sig vara en besvikelse utan slutsatser.

Ett huvudmål för en upprepad undersökning är att mäta förändringar i populationsstorheter, så jämförbarhet över tid är mestadels viktigt. Ändringar i produktionen kan vara motiverade av flera skäl, t.ex. indikationer om önskvärda eller möjliga förbättringar från

tidigare produktionsrundor eller andra interna förändringar i produktionsmiljön. Kostnadseffektivitet är alltid viktigt, och små förbättringar kan innebära betydande besparingar i det långa loppet för en undersökning som genomförs många gånger.

Konsekvenser av ändringar bör dock – så långt möjligt och rimligt – förutses innan ändringarna införs. Löpande undersökningar ger goda möjligheter, eftersom jämförande experiment kan göras inom undersökningen: så kallade inbäddade experiment.

Det är naturligtvis en fördel om resultat av experiment kan generaliseras från en undersökning till andra undersökningar. Det kräver dock att antaganden kan göras om att dessa undersökningar är rimligt lika i de avseenden som är väsentliga för experimentet. Nyttan blir större om även andra kan använda resultatet, och det kan vara viktigare att göra sådana experiment. Det är angeläget med dokumentation och att den är allmänt tillgänglig.

Några referenser för den som vill läsa mer, utöver det metodstöd som finns i Verksamhetsstödet, är följande. En metodbeskrivning av inbäddade experiment ges av van den Brakel och Renssen (2005), som även ger fler referenser. En annan beskrivning med både ett experiment och ett tidsserieperspektiv ges av van den Brakel, Smith och Compton (2008). De beskriver kvalitetsrutiner för undersökningsomläggningar. En praktisk studie för att hantera uppgiftslämnarbörda presenteras av Hedlin, Lindkvist, Bäckström och Erikson (2008). Ett praktiskt experiment avseende för- och nackdelar med förtryckning ges av Holmberg (2004).

## **5.7 Statistikproduktionsprocessen som helhet**

Till de återkommande orden för design hör helhetssyn; en helhetssyn som beaktar statistikens användning och kvalitet, kostnader och uppgiftslämnarbörda. Perspektivet är inte bara den enskilda produkten utan det sammanhang/system som den ingår i. Det här avsnittet behandlar några aspekter på statistikproduktionsprocessen (mer kommer i kapitel 6). De är sammanvägningen av delar till en helhet, kvalitetssäkring, kvalitetskontroll samt utvärdering och återkoppling. Jämför avsnitt 2.7.

### **5.7.1 Sammanvägning och helhetsdesign**

Uttrycket *total survey error* används för att betona att alla bidrag till osäkerhet ingår, se avsnitt 2.6. Tore Dalenius arbetade mycket med fel i undersökningar och kontroller. Han införde uttrycket *total survey design* (Dalenius, 1974). En bild, som återges av Särndal, Swensson och Wretman (1992, sidan 20), har likheter med processkartan. Denna figur visar en process med behov, kravanalys, design, genomförande och användningar. Både design och genomförande omfattar kontroll

och utvärdering. Fokus ligger på fel i undersökningen, dvs. på kvalitetskomponenten tillförlitlighet.

Ofta ägnas mycket av designarbetet åt tillförlitlighet, men de övriga kvalitetskomponenterna är också viktiga. Aktualitet är ett exempel – påtaglig och ofta efterfrågad, t.ex. för ekonomisk korttidsstatistik. I detta exempel kan en avvägning mellan aktualitet och tillförlitlighet behövas (se avsnitt 6.6). Ett annat exempel är jämförbarhet och sammanvändbarhet, där samordning kan behövas, bl.a. av innehåll och utformning av rampopulation.

Figur 4.2 visar delprocesser i designen med huvudsakliga frågeställningar och påverkan på kvalitetskomponenter. Det är många val och allokeringar som tillsammans ska ge ett sökt "optimum". Detta kan t.ex. vara uttryckt i termer av minimikrav på vissa komponenter och bästa möjliga för andra komponenter. Kostnader kan ha ett tak eller vara det som optimeras. Figur 3.1 visar osäkerhetskällor och steg i produktionsprocessen. Den kan vara ett stöd: var finns viktiga osäkerhetskällor och moment som kräver särskilda insatser?

Kunskaper kan komma från pilotstudier, experiment, tidigare produktionsomgångar eller liknande undersökningar. De kan vara en hjälp att sortera bort mindre lämpliga val och plocka fram tänkbara val för fortsatta överväganden. Studien av Linacre och Trewin (1993) är ett exempel med bedömningar av medelfel och kostnader i en figur, där en liten uppsättning goda val synliggörs för fortsatta överväganden.

Det finns just ingen teori. Kalkyler baseras på uppskattningar av den valda estimatorns tillförlitlighet (med hänsyn till alla osäkerhetskällor, så gott det rimligen går) och kostnader. Känslighetsanalys kan vara ett hjälpmedel. I en sådan analys studeras för ett antagande (en variabel) i taget hur mycket en liten avvikelse i antagandet påverkar resultaten. Om känsligheten är stor, är det klokt att försöka förbättra kalkylen/uppskattningen. I utformningen ska både "optimum" och säkerhet beaktas.

Processerna 2.2–2.6 designar efterföljande genomförandeprocesser. I process 2.7 *Designa produktionsflöde* hanteras helheten i designen. Information i Verksamhetsstödet underlättar arbetet genom en lista över viktiga ställningstaganden. Omprövningar av gjorda val kan behövas, och allokeringar kan behöva justeras.

Produktionsflöde ska ses i en vid mening som omfattar metoder, verktyg, arbetssätt, data etc. Det hanteras i designprocessen samt i process 3 *Skapa och testa*. Se vidare t.ex. avsnitt 6.1.

### 5.7.2 Förberedelser – kvalitetssäkring

Som avsnitt 2.7.3 beskriver handlar kvalitetssäkring om det som görs i förväg för kvaliteten; om att förutse och undvika problem.

Tidsoptimism är vanligt förekommande. En bedömning av tidsåtgången för implementering av en metod eller ett verktyg är ett exempel. Bedömningen måste inkludera att arbetet sällan löper friktionsfritt som om det vore i en isolerad värld. Det är svårt att förutse alla detaljer, och när något händer finns kanske inte de resurser som krävs omedelbart tillgängliga.

En adaptiv design kan vara motiverad för att kunna modifiera genomförandet på ett förberett sätt, t.ex. när något i undersökningen är nytt.

Test är en viktig del i kvalitetssäkringen. I tidsnöd händer det ofta att testarbetet får stå tillbaka. I sådana fall händer det också ofta att det blir problem längre fram – problem som det då tar längre tid att lösa än om de upptäckts vid det bortvalda testtillfället.

En annan del är att automatisera manuella moment. Det kan vara information som flyttas, t.ex. i samband med redovisningen. Särskilt när det är bråttom blir det lätt fel.

I process 3 skapas/anpassas/justeras och testas produktionsflödet för att sedan driftsättas. Verksamhetsstödet anger fyra övergripande aktiviteter enligt nedan. Se kapitel 6, särskilt avsnitt 6.1, för detaljer.

- Kontrollera kopplingar mellan delprocesser, verktyg etc.
- Sammanställ underlag för test.
- Testa produktionsflöde.
- Stäm av personresurser, kapacitet i verktyg (t.ex. med tanke på svarstider) och att arbetsrutiner finns.

### 5.7.3 Granskning – kvalitetskontroll

Granskning är en del av kvalitetskontrollen, framför allt kvalitetskontroll av datainsamlingen.

Utformningen av granskning ingår i designen. Det har funnits (och finns förmodligen fortfarande) problem med övergranskning. En sådan felaktig resursallokering ska ändras, normalt så att mer resurser läggs tidigare i produktionsprocessen. Teorin, metoderna och verktygen för granskning har utvecklats på senare år.

Granskningsinsatserna ska kopplas till användningen av statistiken och mikrodata. Det är de kvalitetskrav som de prioriterade användningarna ställer som ska styra granskningens utformning och omfattning. Ett sådant statistiskt synsätt har resulterat i metoder som selektiv granskning och makrogranskning. Läs mer t.ex. i SCB (2002) med titeln Guide till granskning och Verksamhetsstödet, process 5.2.

Erfarenheter i granskning kan och ska löpande undersökningar dra nytta av i kommande produktionsomgångar. Granskardebriefing är en metod för detta, där erfarenheter från granskningsarbetet fångas upp.

Hur ska en avvägning mellan granskningsinsatser göras? Det är en svår designfråga där teori saknas i nuläget.

Verksamhetsstödet innehåller underlag, som ska anpassas till den aktuella situationen, för makrogranskning (process 6.2 *Granska makro-data*) och slutgranskning inför redovisning (underlag i process 7.3 *Redovisa slutprodukt till kund*). Dessa anpassningar ska om möjligt göras redan under utformningen.

#### **5.7.4 Utvärdering och återkoppling**

Process 8 *Utvärdera och återkoppla* beskrivs i avsnitt 4.5 och nämns här mest för fullständighetens skull. Se även avsnitt 7.5 om förbättringsarbete.

Grundtanken är uppenbar: observera och dra nytta framåt. Observationerna kan ske under produktionsprocessen eller genom separata studier. Upplägget ingår redan i utformningen, t.ex. genom process-data eller experiment. I en löpande undersökning kan det, särskilt för korttidsstatistik, finnas flera tidsperspektiv samt koppling till verksamhetsplaneringen. Det kan t.ex. vara resursallokeringen eller tidsplaneringen. Det gäller att mäta tillräckligt, t.ex. kostnader för olika moment. En adaptiv design kan ha fördelar, bl.a. genom flexibilitet och successiva bedömningar.

## 6 Exempel på situationer och val

Detta kapitel innehåller åtta avsnitt som belyser situationer av olika karaktär. Det finns många beröringspunkter med kapitel 5, men avsnitten i kapitel 6 är mer konkreta och praktiska (fast utan att vara detaljerade). Hänvisningar ges till processkartan och Verksamhetsstödet, men de mer övergripande processerna upprepas inte varje gång.

### 6.1 En ny eller annorlunda produktionsomgång

Det här avsnittet avser

- löpande undersökningar som görs första gången eller som gör något utöver rutinen, t.ex. en ändring eller en korrigerings
- engångsundersökningar.

Det finns många typer av risker i statistikproduktionen. En av dem är då något litet, till synes enkelt, görs – och någonting förbises. Det kan vara alternativa vägar i programkod, där någon väg missas. Det kan vara beroenden till andra komponenter i statistikproduktionen; egna eller andras.

Det gäller att både förbereda och testa med eftertanke. Det är ofta lönsamt att ha mer än en person involverad för att öka säkerheten; väl valda punktinsatser och inte dubbelarbete. Ambitionsnivån för test ska stå i proportion till risken, dvs. sannolikheten för fel och felens konsekvenser.

Redan en ny produktionsomgång kan innebära ändringar eller justeringar. Avsnitten 6.1.1–6.1.2 är inriktade på produktionsflödet (jämför avsnitten 5.7.1–5.7.2), medan avsnitten 6.1.3–6.1.4 beskriver kvalitets-säkring vid mer påtagliga ändringar respektive nyheter. Avsnitt 6.1.5 beskriver korrigeringar.

Med det fokus som detta avsnitt har är den främsta hänvisningen i processkartan till process 3.5 *Testa verktyg och produktionsflöde*.

#### 6.1.1 Produktionsflödet – tonvikt på val, tid och resurser

Det finns en beskrivning i avsnitt 4.3.2 med de två huvudorden **välja** och **allokera** som i första hand syftar på metoder och resurser.

Stäm av att alla kvalitetskomponenter är beaktade och att formulerade kvalitetskrav (på register och statistik) är uppfyllda, så långt det går att bedöma i designsteget. Stäm samtidigt av att eventuella ytterligare samordningsbehov är beaktade, t.ex. tidpunkter för genomförande. Är designen ”optimal” och säker?

Gör – med tanke på kvalitetssäkring och kvalitetskomponenten aktualitet – en bedömning av vilka moment som kan automatiseras

hellre än att göras manuellt. Kontrollera att SCB-gemensamma metoder och verktyg har valts, där sådana finns.

Inkludera det utrymme i tid som behövs för att skapa och testa mätinstrument, verktyg och produktionsflödet. Det gäller engångsundersökningar och löpande undersökningar med ändringar eller nyheter.

Studera resursfördelningen igen, både "produktionsideal", så långt de är kända utifrån tidigare erfarenheter, och tillgängliga resurser. Finns det några känsliga punkter att bevaka? Är det lämpligt att lägga in en eller flera hållpunkter i produktionsprocessen? Det kan vara särskilt lämpligt vid en större omläggning. Ska det finnas adaptiva inslag i designen?

Tydliggör överlämnanden - vad och när - mellan delprocesser. Det kan vara särskilt viktigt mellan organisatoriska enheter, internt och externt. Stäm av att metoder, IT-verktyg, andra verktyg, arbetsrutiner och personresurser finns på plats när de behövs.

Se till att designen av delprocesser och produktionsflödet finns dokumenterad med överväganden och val.

#### **6.1.2 Produktionsflödet – tonvikt på verktyg och flöden**

Gå igenom designen av produktionsflödet och rita flödet om det inte redan är gjort. Kontrollera att eventuella ändringar är beaktade i produktionsflödet. Beskriv kopplingar mellan delprocesser så att flödet av data, processdata och metadata kan ske smidigt. Se till att processdata och metadata som kommer att behövas under produktionen och för utvärdering genereras.

Fastställ vilka metoder, verktyg (IT-verktyg, arbetsrutiner, mallar etc.) och arbetssätt som är klara för användning. Konstatera om några verktyg behöver anpassas eller utvecklas. Större insatser läggs lämpligen utanför produktionsflödet, medan mindre insatser kan innefattas. Avgör vad i produktionsflödet som behöver testas. Iterera mellan momenten skapa och testa tills alla ställda krav är uppfyllda.

Användbarhet är en viktig aspekt. Skriv vid behov en användarhandledning för t.ex. verktyg eller uppdatera den befintliga. Se till att utbildningsmaterial finns för verktyg och för produktionsflödet om och där det behövs.

Se till att tillräckligt mycket sparas av programversioner och data så att databearbetningar kan upprepas. Slutliga data ska vid behov kunna återskapas. Se Verksamhetsstödet, process 3.3 *Skapa produktionsflöde*.

### 6.1.3 Kvalitetssäkring av en ändring

Att göra en ändring i ett produktionssystem är alltid ett riskmoment. Det är lätt att skriva fel, att missa en variant bland några stycken, att missa att en annan SCB-enhet använder data på ett annat sätt, etc.

Verksamhetsstödet kompletterar SiP (Systemutveckling i Praktiken) genom att ge ett stöd för mindre ändringar, vidareutvecklingar och tillägg. Det är ett upplägg med arbete i fyra delar – i korthet:

- Förarbete.
- Genomförande av ändringar, tillägg etc.
- Genomförande av test.
- Efterarbete.

Nedanstående moment är exempel på sådant som är viktigt.

- Gå igenom och beakta beroenden med andra (del)system.
- Genomföra ändringar, kontrollera och dokumentera.
- Genomföra test enligt plan (ska finnas, se Verksamhetsstödet, process 3.5).
- Kontrollera att dokumentationen är gjord. Informera berörda.

Det är många gånger en fördel om en annan person än den som ändrar är med och föreslår vilka test som ska göras. Det minskar risken för att glömma en variant eller ett beroende.

### 6.1.4 Implementering av en nyhet

När något nytt görs är det på ett sätt lättare att tänka på alla möjligheter än vid en ändring. Samtidigt kan det vara svårare att testa, eftersom det kanske inte finns data till ett realistiskt test i förväg. Strukturen är dock densamma som ovan med fyra delar.

- Förarbete.
- Genomförande av ändringar, tillägg etc.
- Genomförande av test.
- Efterarbete.

Det är viktigt att testa i förväg i en separat miljö – och inte vänta till skarpt läge.

### 6.1.5 Korrigeringar

En korrigering innebär en ändring utöver gängse rutiner. Sådana gängse rutiner med preliminär och slutlig statistik samt löpande revideringar beskrivs i avsnitt 6.6. Det kan hända att ett misstag (vid framställningen av statistiken eller hos uppgiftslämnaren) upptäcks efter ordinarie redovisning. Det kan hända att ny information tillkommer. Konsekvenserna för statistiken avgör om och när en korrigering ska göras, och om det i så fall är i en särskild redovisning eller inom ramen för en senare reguljär redovisning. Både preliminär och slutlig statistik kan bli föremål för korrigering.



En separat utredning behövs. Gäller korrigeringen en enda referens-tid eller flera i en tidsserie? Är en eller flera statistikprodukter berörda? När ska redovisningen ske? Resonemang och tillvägagångssätt finns i Verksamhetsstödet, särskilt i process 7.3.11.

Om en särskild redovisning ska göras krävs en extra produktionsomgång med design. Om redovisningen sker vid en ordinarie redovisning kan produktionsomgången behöva designas för mer omfattande beräkningar än normalt (fler referenstider). Själva redovisningen kräver extra kommentarer; beakta statistiksekretessen.

## 6.2 Precisionskrav och prioriteringar

Detta avsnitt avser främst urvalsundersökningar, och information finns särskilt i process 2.2 *Designa ram, registerpopulation och urval*, men även i process 2.5 *Designa analys*. Några frågeställningar rör även totalundersökningar, t.ex. frekvens och prioriteringar som påverkar resursallokeringar.

### 6.2.1 Prioriteringar inom en redovisning – och urvalsallokering

Statistiken från en statistisk undersökning innehåller i allmänhet många statistikvärden: tabellceller i en eller vanligen flera statistiska tabeller. Det finns oftast flera variabler och flera indelningsgrunder (spridningsvariabler). En intressant fråga – oavsett om budget eller kvalitetskrav har störst inverkan – är då vad som ska skattas ”så bra som möjligt”, dvs. hur precisionskrav ska formuleras.

Är variablerna (om det finns flera) lika viktiga eller inte? Är aggregat viktigare än delposter (tabellens total, marginaler, ”innanmäte”)? Ska precisionskravet vara i samma enhet som skattningen eller ska det avse ett relativt tal? Är det ”nuläget” (nivåskattningen) som är viktigt eller förändringen sedan en tidigare tidsperiod? Ska en sammanvägning göras för att tillgodose flera krav? Användarsynpunkter är viktiga.

I en urvalsundersökning kan det vara ett precisionskrav eller en prioritering som är styrande för urvalsallokeringen och som är väsentlig för skattningens utformning. I SCB (2008) finns både formler och exempel, och illustrationer av olika prioriteringar finns bl.a. i dess exempel 4.2.8 och 4.2.9. Metodstatistiker kan och ska vara med och göra beräkningar. Det är dock viktigt att komma ihåg att urval inte är den enda osäkerhetskällan.

Ytterligare två referenser för den som vill läsa mer om urvalsallokering är följande; de är tekniska. Ballin och Barcaroli (2013) beskriver hur arbete med optimal stratifiering och allokering kan ske samtidigt. Choudhry, Rao och Hidiroglou (2012) beskriver hur urvalet bör allokeras med hänsyn tagen till de redovisningsgrupper

som skattningarna ska avse. Det är inte trivialt att beräkna ett optimum. Några exempel ges i artikeln.

Även i andra undersökningar än de som har urval med direktinsamling är det viktigt att bedöma tillförlitligheten i förväg, så gott det går med ett rimligt arbete. Är undersökningens upplägg tillräckligt bra för att tillgodose användarbehoven?

### 6.2.2 Förändringstakt och undersökningsfrekvens

Följande frågor är väsentliga att ställa för löpande undersökningar med tanke bl.a. på hur ofta de bör genomföras. Om verkligheten förändras långsamt behövs hög tillförlitlighet i skattningar (t.ex. ett stort urval) för att kunna konstatera skillnaden mellan tidpunkter; att den uppmätta skillnaden är signifikant.

- Hur snabbt förändras verkligheten?
- Vilken precision ska finnas i t.ex. förändringsskattningar?
- Hur ofta bör undersökningen genomföras?

Om undersökningen görs ofta hinner det inte hända så mycket – om den görs sällan kan det dröja innan förändringar märks. Det gäller att hitta en avvägning och presentera statistiken i väl valda ordalag (inte överdriva eller övertolka skillnader mellan osäkra statistikvärden).

### 6.3 Design av hantering av outliers (i mikrodata)

En outlier är ett objekt med ett värde som avviker kraftigt från den stora mängden i den delpopulationen. Det kan vara ett företag som klassificerades som litet, på basis av tillgänglig registerinformation vid urvalsallokeringen, men som klassificeras som stort på basis av den information som inhämtas i undersökningen. Det kan vara en variabel som är snedfördelad i sig, t.ex. individers förmögenhet eller företags byggnadsinvesteringar. Värdet är korrekt men i ett urval med uppräknings blir det "alltför" betydelsefullt. Skattningen har ingen bias (inget systematiskt fel), men variansen är hög. Det kan vara motiverat att ändra skattningen så att medelkvadratfelet blir mindre (kanske minimeras); variansen minskas då till priset av en viss bias.

Det är förstås ännu bättre om situationen inte uppstår. Det går att påverka delvis men inte helt. Ett sätt är att i designen undvika mycket höga uppräkningsstal; det kan vara bättre att använda *cut-off* och modellskattning.

Det finns i princip två metoder för att hantera outliers i mikrodata: minska vikten eller modifiera värdet. Verksamhetsstödet innehåller viss information, men ganska lite. En etapp i ett utvecklingsprojekt har genomförts och redovisats i början av 2016. Fortsatt arbete pågår.

Därefter kan det komma stöd för valet av metod, som görs i process 2.5 *Designa analys*.

Modifiering av värde kan t.ex. göras genom så kallad winsorisering. Då dämpas, som exempel, ett högt observerat värde  $y$  till en gräns  $z$  före multiplikationen med uppräkningsfaktorn  $f$ . Produkten  $f \cdot y$  med stort genomslag ersätts därigenom med produkten  $f \cdot z$ . Det finns en variant där i stället  $y + (f-1) \cdot z$  används, eftersom värdet  $y$  är korrekt för det observerade urvalsobjektet. Hur kraftig winsoriseringen ska vara beror på den valda skattningsprincipen och på fördelningen hos observationsvariabeln, enligt erfarenheten.

Betydelsen varierar med redovisningsgrupp; ett värde kan vara (allt-för) inflytelserikt på en detaljerad nivå utan att vara särskilt märkbart för hela populationen. En möjlighet är att använda den ordinarie skattningen på totalnivå, dämpa inflytandet på den detaljerade nivån, ta skillnaden mellan den ordinarie och den dämpade skattningen samt "fördela ut" den skillnaden på populationen i övrigt.

För den som vill läsa mer finns t.ex. nedanstående tre referenser som kommit nyligen.

- Mulry, Oliver och Kaputa (2014): artikeln studerar två metoder för värdemodifiering och illustrerar med en detaljhandelsundersökning i USA.
- Martinoz, Haziza och Beaumont (2015): artikeln ger en metod för att bestämma tröskeln för winsorisering och för att hantera både redovisningsgrupper och deras total så att det blir samstämmighet.
- Beaumont, Haziza och Ruiz-Gazen (2013): den rätt teoretiska artikeln studerar robust estimation för ändliga populationer.

En äldre referens är Kokic och Bell (1994). Den metoden har testats på SCB med gott resultat då.

Hantering av outliers är ett exempel på värdet av ett underbyggt ställningstagande i designen. Att i skattningssteget behandla en upptäckt outlier som bortfall eller att "flytta" den till ett eget stratum med vikten 1 är i allmänhet inte att rekommendera med tanke på bias (om än något beroende på hanteringen i övrigt, t.ex. av vikter).

#### 6.4 Några situationer med modellbaserade skattningar

I officiell statistik är en kombination av slumpmässigt urval och designbaserade skattningar ofta använd. Det finns dock situationer där den slumpmässiga osäkerheten anses vara oacceptabelt stor, och andra typer av skattningar är då vanligt förekommande. Dessa kan grundas på en modell med en slumpmekanism, varvid modellen påverkar både skattningen och bedömningen av skattningens till-

förlitlighet (osäkerhet). Några typiska situationer där modeller ofta får en starkare roll än i designbaserade och modellassisterade skattningar ges nedan. Se även avsnitt 5.2, som lägger en gemensam resonerande grund för dessa tillämpningsexempel. De processer som är aktuella är 2.2 *Designa ram, registerpopulation och urval* samt genomgående, som utgångspunkt, processerna 1.4 *Fastställ informationsbehov* och 2.1 *Designa slutprodukt*.

#### 6.4.1 Undersökning med cut-off

En undersökning kan använda så kallad *cut-off*; ett typiskt exempel vid direktinsamling är att små företag inte ingår i datainsamlingen, men de ingår i målpopulationen och skattningarna. En modell bildas genom att en relation antas gälla mellan en undersökningsvariabel  $y$  och en eller flera hjälpvariabler  $x$  som finns i ett register också för de små företagen. Parametrar i modellen för relationen kan skattas med hjälp av undersökningens direktinsamling för ganska små företag (strax ovanför gränsen för *cut-off*, dvs. tröskelvärdet) och modellen antas gälla även under denna gräns för *cut-off*. Skattningen sker i allmänhet på aggregerad nivå, men en annan möjlighet är imputeringar för enskilda objekt.

Ibland beskrivs *cut-off* som en medveten undertäckning, men det finns skillnader genom att det finns raminformation och att det är ett kalkylerat val.

Förr var det ofta en fast gräns för *cut-off*, t.ex. vid 10 anställda oavsett bransch. Numera är det rätt vanligt att gränsen beräknas utifrån en andel som baseras på en viktig undersökningsvariabel, t.ex. att 95 procent av omsättningen ska ligga ovanför gränsen. Det behövs mer designarbete än att ta en så enkel regel; variansen för estimatormåste studeras. Denna varians beror av flera faktorer, bl.a. av modellen och hur många företag det finns i olika storleksklasser.

För den som vill läsa mer finns t.ex. Benedetti, Bee och Espa (2010), som ger en formaliserad beskrivning av bl.a. stratifiering med stöd för val av en lämplig gräns för *cut-off*.

#### 6.4.2 Små redovisningsgrupper

Skattningsförfaranden i situationer med så små redovisningsgrupper att en designbaserad ansats inte kan tillämpas, kallas ofta *small area estimation*, men de gäller inte bara för geografiska områden. De enskilda skattningarna brukar bestå av två delar som kombineras.

- En urvalsundersökning görs och används för att skatta parametrar för större grupper/områden.
- Det finns registerinformation för små grupper/områden.

Parameterskattningar och registerinformation kombineras, t.ex. beräknas skattningar för kommuner från designbaserade skattningar av relationer på länsnivå och kommunspecifik information.

Ett annat exempel är branscher, där det i den europeiska statistiken kan finnas redovisning på en delbransch som i svensk statistik är liten och inte efterfrågad. Ska den delbranschen undersökas som en "vanlig" bransch eller ska den modellskattas inom ramen för en större bransch som undersöks? I det senare fallet kan urvalet slumpmässigt innehålla något enstaka eller inget urvalsundersökt objekt. Registerinformation för den lilla delbranschen och en sambandsmodell baserad på den större branschen får då användas. Valet mellan undersökning och modellskattning är en avvägning med hänsyn tagen till tillförlitlighet, uppgiftslämnarvärda och kostnader. Om antalet urvalsobjekt ses som fixt kan nyttan vara större om objekten dras i en annan bransch än den lilla.

#### 6.4.3 Skattningar som görs tidigt med begränsad information

I vissa fall behöver skattningar göras mycket tidigt i produktionsomgången, då data är ofullständiga. Det kan vara en undersökning med direktinsamling eller en undersökning baserad på administrativa data. Någon form av antagande behövs då om data som saknas: i förhållande till helheten eller i förhållande till det som har kommit in. I löpande undersökningar kan kunskap byggas upp över tid. Det finns dock en risk att antagandena inte är stabila över tid, t.ex. om det finns ett konjunkturberoende. Se även avsnitt 6.6 om preliminär statistik.

I ett ESSnet-projekt om administrativa data – se ESSnet Admin Data (2011) och ESSnet Memobust (2014) för en översikt av skattningar – finns en struktur avseende administrativa data med två huvudfall enligt nedan.

- A. Nästan kompletta data: Använd dessa och imputering.
- B. Inga eller begränsade administrativa data: Använd administrativa data från tidigare period(er) tillsammans med antingen regressionsteknik eller någon form av *benchmarking* eller *nowcasting* där kortperiodisk statistik och årsstatistik ingår tillsammans med antaganden om tidsserien, t.ex. för förändringsmönster; detta för att skatta denna referensperiod.

Motsvarande resonemang kan föras även vid direktinsamling.

#### 6.4.4 Glesare undersökningstillfällen?

Det kan finnas skäl att glesa ut datainsamlingen i en undersökning. Är det ändå möjligt att ta fram skattningar? Kanske, om modellbaserad statistik accepteras. Det finns likheter med huvudfallet B i avsnitt 6.4.3. Det är nödvändigt att använda en modell, t.ex. en regressions-

modell där en oberoende variabel  $x$  finns för den aktuella perioden och skattningar för en variabel  $y$  tas fram utan att undersöka  $y$  den perioden.

Det är skillnad mellan situationer med *vissa data* för den aktuella referensperioden och situationer *utan data* för den aktuella referensperioden. I det senare fallet tar SCB bara fram statistik (prognoser och motsvarigheter) i utpekade fall, t.ex. för befolkning, utbildning och arbetsmarknad.

### 6.5 Design av en löpande undersökning

Typiskt för en löpande undersökning är bl.a. nedanstående punkter som ska beaktas i designen.

- Information från tidigare produktionsomgångar kan vara till nytta, såväl i kontakt med återkommande uppgiftslämnare (vanligt i t.ex. företagsundersökningar och panelundersökningar) och granskning på mikro- och makronivå som i resursallokering. Även imputeringsförfaranden kan bygga på sådan information, t.ex. genom ett tidigare värde i kombination med ett utvecklingstal som baseras på likartade, svarande, objekt. Det kan vara bättre än en medelvärdesimputering; det får bedömas i varje enskilt fall.
- Vissa undersökningar använder så kallad förtryckning vid direktinsamling; se Holmberg (2004) för en SCB-studie. Det innebär att uppgifter som objektet (t.ex. ett företag) har lämnat för en eller flera tidigare referensperioder förtrycks i en postenkät (motsvarande vid webb). Det kan vara ett stöd för uppgiftslämnaren, men det riskerar att konservera ett tidigare felaktigt svarsbeteende. Det kan omvänt leda till att tidigare misstag upptäcks av uppgiftslämnaren (en annan eller samma). – Tidigare lämnade uppgifter kan användas också vid intervju, t.ex. genom att ställa frågan om den uppgiften fortfarande är aktuell.
- Det finns vinster att göra genom att redan före den första omgången beakta i designen att undersökningen är löpande. Det gäller t.ex. vid utformningen av urval och vid utformningen av de processdata som ska både samlas in och användas.
- Kvalitetskraven avser i många fall (t.ex. inom ekonomisk statistik) förändringsskattningar snarare än nivåskattningar. Det ska beaktas i utformningen (jämför avsnitt 6.2).
- Tidsserier är ofta viktiga. Tidsseriebrott bör i allmänhet undvikas men kan vara befogade. Metodändringar kan vara motiverade för att minska systematiska fel (eller slumpmässiga fel,

vilka dock ofta är mindre störande). Vid ändring bör storleken på tidsseriebrottet mätas eller uppskattas, t.ex. genom en så kallad dubbelmätning eller genom en särskild studie (exempelvis ett experiment, se avsnitt 5.6.2). Kvalitetskrav på jämförbarhet och sammanvändbarhet innebär krav på "likadana" definitioner och metoder. När verkligheten ändras, så ändras också krav, behov och önskemål på statistik. Samtidigt brukar det finnas önskemål om att bevara och inte bryta tidsserier. I någon mån kan tidsserier skrivas bakåt och framåt i tiden med hjälp av varandra (länkning); se avsnitt 6.7.

- Process 8 *Utvärdera och återkoppla* är en återkommande rutin, där processdata och förbättringsmöjligheter studeras. Det är viktigt att ha lämpliga processdata sparade och tillgängliga.

Urvalen i SAMU-systemet är samordnade mellan undersökningar (positivt för samordning eller negativt för att sprida uppgiftslämnarbördan) och över tid (positivt för bättre precision). Detta fungerar särskilt väl då det finns gott om objekt i ett urvalsstratum (motsvarande) så att en spridning är möjlig. Det är normalt en fördel för både en uppgiftslämnare och SCB om uppgiftslämnaren är med i en undersökning under en sammanhängande period och sedan inte alls under en följande period, detta hellre än att hoppa in och ut.

### 6.6 Preliminär och slutlig statistik – löpande revideringar

Tillförlitlighet och aktualitet är två kvalitetskomponenter som "drar åt olika håll". Ett sätt att tillgodose båda slagen av önskemål är att redovisa först preliminär statistik och sedan slutlig statistik. Den preliminära statistiken är mer aktuell och mindre tillförlitlig; den kan också vara mindre detaljerad med hänsyn till den lägre tillförlitligheten. Det finns sedan möjlighet att hinna granska mer och att samla in mer data genom samma eller eventuellt någon ytterligare källa till den slutliga statistiken.

Preliminär statistik revideras en eller flera gånger. Revideringars storlek ger viss information om tillförlitligheten. Den storleken avser dock en skillnad mellan skattningar, medan tillförlitlighet avser närhet till sant värde. Användarna ska få information om hur stora revideringar har varit i genomsnitt under en tidsperiod. Vanligen används absolutbelopp, dvs. skillnader utan tecken, i genomsnittet. I genomsnittet av skillnader med tecken blir det en utjämning och en optimistisk bild. Denna skillnad, där tecken ingår, kan dock indikera en systematik med över- eller underskattningar.

Både den preliminära och den slutliga statistiken kan och ska ha tillräcklig kvalitet; var och en för sitt ändamål.

Hur tidigt kan den (första) preliminära statistiken komma, hur många gånger ska den revideras och vid vilka tillfällen – och när ska den slutliga statistiken komma? Det är viktigt att känna till användarbehoven. Det är en fördel med samordning mellan undersökningar, däribland NR. Det kan finnas EU-förordningar att beakta. Det kan vara praktiskt att revidera samtidigt som en senare referensperiod publiceras (redovisas) för första gången.

Ett nyckelord kan då vara **informationsvärde**. Det finns en sådan grundprincip för korrigeringsprocessen i Verksamhetsstödet process 7.3.11.1, resonemang av Meinke och Schmidt (2014) samt likartade sammanställningar av IMF och av OECD (2007).

Hur mycket ökar en revidering informationsvärdet för statistikanvändare som därigenom ska ändra/uppdatera sina beräkningar/analyser/slutsatser?

En avvägning ska således göras mellan å ena sidan förbättringen av statistikens kvalitet och å andra sidan merarbetet för användarna. Det finns ännu ingen konkretisering av detta informationsvärde. Det är önskvärt att ha något mätbart. Det finns likartat ett engelskt uttryck "news or noise" som betonar att revideringar ska ge ny information och inte vara brus.

Det är viktigt för statistikproducenten att analysera revideringar fortlöpande och ställa exempelvis nedanstående frågor.

- Finns det något systematiskt mönster?  
Är det t.ex. vanligare med underskattningar än överskattningar?  
Finns det en samvariation med exempelvis konjunkturcykeln?
- Går skattningarna att förbättra?
- Hur stort är informationsvärdet för användarna?  
Finns det skäl att ändra antalet revideringar och tidsschemat?

En genomgång av revideringars antal och storlek i svensk ekonomisk statistik gjordes av Köll (2007); en genomgång som ledde till rekommendationer. Ett viktigt budskap då var att inte revidera (så) många gånger.

Till viktiga processer hör 1.4 *Fastställ informationsbehov* och 2.1 *Designa slutprodukt*. Jämför avsnitt 6.5 om en löpande undersökning samt avsnitt 6.1.5 om korrigeringar. Notera även att det i kvalitetsbegreppet finns en underkomponent Preliminär statistik jämförd med slutlig, avsnitt 2.6.



## 6.7 Ändringar av metoder och definitioner – tidsseriebrott

### 6.7.1 Inledning

Många statistikanvändare vill ha långa tidsserier, med jämförbarhet över tiden. Samtidigt sker många förändringar i omvärlden, och metodförbättringar av olika slag blir önskvärda. Onödiga tidsseriebrott ska förstås undvikas. Ändringar och omläggningar i statistiken ska förberedas och göras kontrollerat. Diskussioner med användarna ingår. Tidsseriebrott diskuteras bl.a. i avsnitt 6.5 om löpande undersökningar. Två huvudfall – se avsnitt 3.6 – beskrivs i avsnitten 6.7.2 och 6.7.3.

För den som vill läsa mer finns t.ex. projektrapporten SCB (2007), som beskriver tidsseriebrott och diskuterar hur jämförbarhet över tiden kan kvalitetssäkras. Många exempel ingår, bl.a. på länkning.

Processerna 1.4 *Fastställ informationsbehov* och 2.1 *Designa slutprodukt* är återigen viktiga; även processerna 2.5 *Designa analys* och 3.6 *Genomför pilotstudie* (med olika former av pilotstudier, bl.a. experiment).

### 6.7.2 Ändring av en metod

Vid enstaka tillfällen ändrar en löpande undersökning en metod, t.ex. en skattningsmetod eller en datainsamlingsmetod. Detta görs t.ex. för att öka statistikens tillförlitlighet, minska uppgiftslämnarbördan eller sänka kostnaderna. Det finns en risk för ett tidsseriebrott genom en systematisk skillnad. Detta bör studeras i förväg, t.ex. genom en metodutredning eller ett inbäddat experiment, se avsnitt 5.6.2. Diskussioner med statistikanvändare ska föras, och användarna ska få förhandsinformation i god tid, t.ex. i användarråd och på SCB:s webbplats. Vid tidsseriebrott är det ofta önskvärt att förlänga den nya tidsserien bakåt med hjälp av den gamla, att länka de två tidsserierna. Se t.ex. SCB (2007).

Ett inbäddat experiment har likheter med vad som ofta kallas dubbelmätning. En samtidig insamling med både gammal och ny metod eller med både gammalt och nytt begrepp kan användas för att mäta skillnader. Det kan t.ex. gälla en ändrad datainsamlingsmetod, övergång från direktinsamling till administrativa data, ett ändrat begrepp eller en ändrad frågeformulering. På det här sättet går det att få ett mått på tidsseriebrottets storlek för den eller de perioder som dubbelmätningen görs. För den som vill läsa mer finns t.ex. en (rätt teknisk) artikel skriven av van den Brakel et al. (2008), som går igenom olika möjligheter och exemplifierar med holländska undersökningar.

En annan typ av angreppssätt är att göra en successiv infasning av det nya. Det innebär att exempelvis en ny urvals- eller insamlingsmetod introduceras successivt under en period (en del av urvalet i taget) så att det blir en liten skillnad varje gång. Därmed sprids ett

eventuellt tidsseriebrott ut över tiden. Statistikproducenten kan, beroende på upplägget, göra vissa jämförelser medan infasningen pågår. Det kan vara svårt att särskilja effekter av metodbytet, dvs. tidsseriebrottets storlek.

En ytterligare möjlighet, som bör undvikas, är att inte ha någon gemensam period alls. Se t.ex. van den Brakel et al. (2008), som beskriver hur en tidsseriemodell då kan användas för att ta hänsyn till såväl metodövergång som förändringar i den verklighet som mäts. Förfarandet innebär ett modellberoende.

### **6.7.3 Ändring av begrepp, definition eller klassifikation**

Som avsnitt 6.7.1 pekar på behövs avvägningar mellan långa tids-serier och utvecklad statistik, som beaktar omvärldsförändringar och förskjutningar i användarnas intressen. Detta kräver genomtänkta bedömningar och användarkontakter.

När begrepp, definitioner och klassifikationer ändras eller tillkommer, får statistiken ett delvis nytt innehåll med nya redovisningsgrupper. Då finns ingen motsvarande gammal tidsserie. Det brukar finnas önskemål om bakåträkning av statistiken för den nya/ändrade definitionen eller klassifikationen. Det är en form av länkning med "många-till-många" tidsserier.

Bakåträkning av en ny näringsgrensindelning är ett illustrativt exempel på en "omfördelning" av statistiken från en till en annan uppsättning redovisningsgrupper. Det är vanligt med en makroansats, som använder relationer avseende en viss period för vilken information om båda klassifikationerna (gammal och ny) finns. Som exempel är FDB dubbelkodat ett år, och det finns administrativa data om t.ex. löner och omsättning. En matris med lämplig detaljeringsgrad (inte alltför fin) tas fram och används som fördelningsnyckel bakåt i tiden, så långt som det är lämpligt. De gamla redovisningsgrupperna "omfördelas" till den nya indelningen under antagandet att fördelningsnyckeln är oförändrad. Här krävs en diskussion av rimligheten i antagandet: Hur långt tillbaka kan det användas och på hur detaljerade nivåer?

Den som vill läsa mer kan t.ex. se SCB (2007), SCB (2010, Bilaga 4) och SCB (2012, avsnitt 3.3). I rapporten SCB (2007) om kvalitetssäkring av jämförbarhet över tiden finns många exempel på hantering av tids-serier inklusive information om tidsseriebrott i situationer av denna typ, t.ex. ändrade regionala indelningar och ändrade klassifikationer för yrke, utbildning m.m. EU-harmonisering i arbetskraftsundersökningarna (AKU) är ett exempel; ett krävande arbete där även säsongrensning och metodbyten ingick.

## 6.8 Statistisksystem – påverkan på undersökningsdesign

I ett statistisksystem behövs dels ett gemensamt agerande som drivs från en högre nivå, dels ett förändringsarbete på de enskilda undersökningarna för att tillgodose systemets uttalade behov. Ändamålen för systemets statistikprodukter ska, såvitt möjligt, utgöra en helhet.

Det här avsnittet har en framåtblickande karaktär. Beskrivningar och ansatser finns ännu inte i Verksamhetsstödet.

### 6.8.1 Inledning med Strategi 2020

Det finns olika skäl till och former av samordning. Det kan gälla kvalitetskomponenter (statistikens innehåll, sammanvändbarhet med flera), kostnader och uppgiftslämnarbördor. I det europeiska statistiksystemet finns en rörelse från en harmonisering av statistiken mot en samordning som avser hela statistikproduktionen med ramverk, metoder och verktyg. Riktlinjer för europeisk statistik och Ramverket för kvalitetssäkring (Eurostat, 2011 respektive 2012) betonas mycket starkt. Den svenska motsvarigheten, ROS (2013), är en handledning med europeiska principer, indikatorer och kvalitetssäkringsmetoder samt svenska exempel. ESS-visionen 2020 beskrivs t.ex. av Eurostat (2015).

I SCB:s Strategi 2020 står nedanstående som en av fyra prioriterade vägar för att nå de övergripande målen (SCB, 2014b).

- Samordnad statistik

Med ett enhetligt, samordnat statistisksystem med konsistenta data arbetar SCB effektivt för att möta användarnas behov och krav. Den statistiska informationen är ordnad i ett strukturerat datalager för att möjliggöra flexibel sammanvändning. Genom analyser av det statistiska materialet förbättrar SCB ständigt statistikens kvalitet och användbarhet.

I den konkretisering av visionen som presenterades i maj 2015 konstateras att SCB:s samlade statistikutbud är grundat på tre pelare: ett ekonomisk-statistiskt system, ett socialstatistiskt system och ett miljöstatistiskt system. De tre systemen bygger på gemensamma standarder i form av begrepp, definitioner och klassificeringar. Helheten prioriteras framför enskilda produkter.

### 6.8.2 Kvalitet – särskilt innehåll, jämförbarhet och sammanvändbarhet

I ett statistisksystem finns (idealt) en samlad prioritering och samordning av statistisksystemets **innehåll** så att statistikprodukterna tillsammans utgör en helhet som omfattar väsentliga delar av ämnesområdet: såväl utan väsentliga luckor som utan (omotiverat) överlapp.

Systemets statistikprodukter är sinsemellan **jämförbara** och **sammanvändbara** om det inte finns goda skäl för avsteg. Om det finns moti-

verade ändringar som förväntas medföra **tidsseriebrott** genomförs de samordnat. I statistikprodukter som skattar samma målstorheter – om det förekommer – överensstämmer statistikvärdena numeriskt, bortsett från med hänsyn till slumpvariation försumbara skillnader. Dessa målsättningar kan vidareutvecklas, se nedan.

### 6.8.3 Design – definitioner och metoder i undersökningar

Jämförbarhet och sammanvändbarhet beror av definitioner (statistiska målstorheter) och metoder, se t.ex. avsnitten 3.6 och 4.3.3. Definitioner och metoder väljs i designen, som således är ett medel att uppnå målen för statistiksystemet.

De statistiska målstorheterna beror av en rad definitioner och klassifikationer, inklusive statistikens byggstenar: objekt, populationer, redovisningsgrupper, variabler och referensperioder samt sammanfattande mått. I en jämförelse eller sammanvändning ska vissa byggstenar vara lika, medan andra byggstenar är de som jämförs eller analyseras (t.ex. tidsperioder, länder, sektorer eller relationen mellan två variabler). Samordning av befintliga undersökningar gäller t.ex. variabeldefinitioner, populationsavgränsningar och referensperioder.

Metoder ska väljas så lika som möjligt och rimligt i avseenden som kan ha påverkan, särskilt bör systematisk inverkan undvikas. Metodvalen kan t.ex. gälla ramframställning inklusive tidpunkter, mätinstrument, urvalsmetod, skattningsförfarande och hjälpinformation.

I nationalräkenskaperna (NR) ska "allt" vara med – en och endast en gång. Det är lätt att se några stora utmaningar i detta, även för de statistiska undersökningar som NR använder. Om statistik för två sektorer produceras genom två olika undersökningar är det viktigt att dessa är samstämmiga i sin syn på och implementering av t.ex. populationsavgränsningar, klassificeringar och referensperioder, detta inklusive ramframställning.

Det finns en inneboende svårighet i att få korttids- och årsstatistik att "stämma". De genomförs vid olika tider, under respektive efter kalenderåret, och därmed har de olika aktualitet i sin information, t.ex. i ramar. Det har betydelse då populationerna är snabbt föränderliga.

En möjlighet vore att låta årsstatistiken använda korttidsstatistikens äldre information (med praktiska svårigheter vid t.ex. omorganiseringar). En annan möjlighet skulle kunna vara att anpassa korttidsstatistiken till årsstatistiken – genom så kallad *benchmarking* – när den senare blir tillgänglig. Ett sådant förfarande med årsanpassning har förekommit för månadsstatistik. NR anpassar kvartalsräkenskaperna till årsräkenskaperna; en intern numerisk konsistens. Anpassningen kan ha olika principer för vad som ska bevaras, t.ex. förändringsmönstret mellan närliggande perioder.

SAMU-systemet innebär (se avsnitt 6.5) inte bara en samordning av urval, utan det underlättar också samordning av populationsavgränsningar och ramar. Därmed kan dubbelräkningar och luckor till följd av olika definitioner eller registerversioner (eller båda) undvikas. Även tidpunkterna för raminformationen har betydelse.

#### 6.8.4 Ytterligare aspekter på system

**Uppgiftslämnarbördan** blir enklare att beakta i ett system än i separata undersökningar. Det gäller variabelsamordning, antal undersökningar, insamlingstidpunkter m.m. SAMU-systemet ger funktionalitet i några för uppgiftslämnarbördan viktiga avseenden. Vissa av dem gäller också för separata undersökningar.

Även indelningen i undersökningar kan vara motiverad att studera och överväga. Det som har tillkommit successivt är kanske inte optimalt nu. Flera länder, bl.a. i det europeiska statistiksystemet, har arbetat med administrativa data och med samordning av dessa och andra undersökningar, med direktinsamling, till ett system. Variabler är viktiga. Smith (2014) beskriver arbete i Storbritannien, och Bakker, van Rooijen och van Toor (2014) i Nederländerna; i båda fallen avses det socialstatistiska systemet. Den senare artikeln beskriver statistiska och mer tekniska aspekter med register och matchningar.

Karlberg, Reis, Calizzani och Gras (2015) beskriver vad de kallar en verktygslåda för en design med moduler och en analys som väger samman undersökningsresultat. Det finns intressanta inslag, men systembygget får inte bli alltför komplext.

**Konstruktionen av ett system** innehåller många avvägningar, där bl.a. kvalitet och kostnader ingår; vidare internationella förordningar/riktlinjer/rekommendationer samt datakällor. Administrativa data och befintliga data har stor inverkan på uppbyggnaden.

Man kan tänka i moduler eller block av undersökningar som bidrar till systemet. Det kan vara olika källor, variabelgrupper, olika undersökningsfrekvens etc. Vissa variabler undersöks varje år, som exempel, medan det för andra kan räcka med exempelvis vart tredje år.

Om systemet är väl byggt och organiserat kan skattningar skapas genom att kombinera data från flera moduler. Eurostat har börjat beskriva, något visionärt, hur ett socialstatistiskt system med moduler skulle kunna se ut.

Ett statistiksystem bör normalt innebära att det sammantaget blir en **rationalisering** av statistikproduktionen. Det blir färre eller inga skraddarsydda lösningar att underhålla. Det blir mer av gemensamma komponenter och system. Detta är på intet sätt trivialt att åstadkomma. Det kan – för att ta två exempel – handla om samgranskning och sammansatta skattningar.

## 7 "Modern" statistikproduktion

Det här kapitlet beskriver vad som ibland kallas modern statistikproduktion. Det finns invändningar mot ordet modern. De karaktäristika som betonas är inte särskilt nya – åtminstone de flesta har diskuterats under ganska lång tid. Eftersom den utveckling och implementering som fordras är tidskrävande, kommer det att dröja till det moderna är uppnått.

Det är dessutom så att statistikproduktionen påverkas av att verkligheten förändras, likaså av statistik användarnas prioriteringar. Detta innebär att flera typer av förändringar pågår samtidigt.

De sinsemellan olika aspekter som beskrivs i avsnitten 7.1–7.5 påverkar alla design och designarbete. Det gäller bl.a. kvalitetskomponenter, sätt att producera statistik, verktyg och arbetssätt. Det som står i avsnitten 7.1–7.4 kan ses som något "visionärt" eller framåtblickande. En del är på plats, medan annat är ett börläge att sträva mot. Dessa avsnitt har många beröringspunkter.

Avsnitt 7.5 skiljer sig från de övriga i kapitlet genom att vara tydligt närliggande; ett nuläge att arbeta både i och för.

### 7.1 Statistikproduktion i en föränderlig värld

Det här avsnittet ger en kort beskrivning av de omställningar som pågår allmänt på statistiska centralbyråer och de inriktningar som finns i strategier. Det är till stor del omvärldens förändringar, krav och förväntningar som är orsakerna.

Statistikens innehåll behöver ständigt förnyas och utvecklas för att beskriva nya fenomen och aktuella frågeställningar. Detta ska förutses och tillgodoses, helst med sådan framförhållning att det finns tidsserier för den nu intressanta verkligheten. Statistiken ska vara lätt att hitta och att förstå. Den ska enkelt kunna vidarebearbetas och användas tillsammans med såväl annan statistik som andra data. Hög aktualitet efterfrågas, förstärkt av det stora informationsflödet i samhället.

Det kan vara svårt för en användare att se och förstå skillnader mellan officiell statistik och andra data. Användarkontakter och kvalitetsbeskrivningar av statistiken blir än mer viktiga att ha och ge.

Produktionsmöjligheter påverkas också av den stora mängd data som finns tillgänglig. Administrativa data har använts länge, och andra källor prövas också, t.ex. i prismätningar som kan använda kassaregisterdata. Det blir allt svårare och dyrare att samla in data direkt. I många undersökningar ökar bortfallet trots större insatser, vilket är en utmaning att hantera. Det är överhuvudtaget starkt tryck på effektiviseringar; att åstadkomma mer till en lägre kostnad.

En nyligen publicerad artikel har den talande titeln: "*From multiple modes for surveys to multiple data sources for estimates*"; Citro (2014). Även i USA, med mycket svagare registertradition än i de nordiska länderna, går tankarna i sådana banor. Artikeln kommenterar såväl osäkerhetskällor och svag kunskap om t.ex. mätfel som paradigmskifte vad gäller synen på statistiska undersökningar. Fördelar med att kombinera datakällor betonas.

Ett produktionssystem för en undersökning byggs numera normalt inte isolerat och helt skräddarsytt. Det är önskvärt att ha komponenter som enkelt och säkert kan sättas samman till en ny eller ändrad statistikprodukt. Det finns uttryck för detta som *assemble to order* och *plug and play*. Ett sådant angreppssätt ställer krav på både metoder och verktyg.

Det internationella samarbetet ökar och breddas. Det gäller klassifikationer (sedan länge), metoder, verktyg, synen på statistikproduktion, arkitektur, modeller av olika slag, etc.

## 7.2 Befintliga data och datalager

Det finns flera goda skäl att använda befintliga data hellre än att göra en ny insamling av samma eller likartade data. Uppgiftslämnarbördan minskar, liksom kostnaderna. Administrativa data och andra befintliga register återanvänds. Flera aspekter måste beaktas vid återanvändning. Det ska t.ex. vara meningsfullt, juridiskt tillåtet och tekniskt möjligt. Sådana aspekter blir tydliga när flera datamängder ska integreras. Är exempelvis objekten desamma, populationerna avgränsade på samma sätt, etc.

DuR står för Datalagerutveckling och Registersamordning. Ett projekt för att strukturera och samordna SCB:s datalager startade 2012 efter en lång förberedelse. Projektet är nedlagt, men DuR-visionen lever, och sådant arbete bedrivs inom andra projekt. Frågor om begrepp och termer, informationsstruktur och informationsutbyte samt konsistens och sammanvändbarhet blir tydliga och viktiga. Det behöver finnas "data om data" (metadata) och en strategi, t.ex. för vad som sparas och för ändringar och spårbarhet.

## 7.3 Standardisering

Vid standardisering skapas det, allmänt sett, gemensamma konventioner eller sätt att arbeta för att något ska fungera smidigt, t.ex. en del tillsammans med andra delar. En standard kan vara en gemensam lösning på ett återkommande problem.

Vad kan och bör standardiseras i statistikproduktion – och varför? Fokus ligger, som så ofta, på statistikens kvalitet, kostnader och uppgiftslämnarbörda. Standardiseringen påverkar bl.a. innehåll med

begrepp och klassifikationer. En smidig kommunikation av data utan risk för missförstånd är ett annat skäl för standardisering; här ingår dataformat, metadata med mera. Statistikens kvalitet påverkas även av metoder och verktyg (i vid mening), se t.ex. avsnitten 3.6 och 6.7. Kostnader minskar i många fall genom standardisering, men de kan öka om standarden innebär en ambitionshöjning.

Standardisering övertolkas ibland till att betyda en och endast en metod för urval (skattning, granskning och så vidare), oberoende av förutsättningarna. Det är en utmaning att hitta en balans mellan standardisering och flexibilitet, t.ex. då ett gemensamt verktyg ska konstrueras och byggas på ett sådant sätt att det är både funktionellt och användarvänligt i många vanligt förekommande situationer. Vilka situationer, metoder etc. som verktyget ska hantera kräver utredning och prioritering bland många önskemål.

Ett par exempel, bland många, för den som vill läsa mer om statistikbyråers arbete med standardisering är följande. Merad och Brodie (2011) beskriver hur kortperiodiska brittiska företagsundersökningar integrerades. Godbout (2011) beskriver en utveckling av bearbetningen i företagsundersökningar vid Statistics Canada, inklusive återkoppling till tidigare delprocesser. Exempelen innefattar standardisering i flera avseenden, såsom statistikens innehåll, metoder och verktyg.

#### **7.4 Arkitektur och infrastruktur**

Exempel på resurser, utöver personella och ekonomiska resurser, som SCB förfogar över är den tekniska infrastrukturen, Verksamhetsstödet och det processororienterade arbetssättet. Arkitektur och infrastruktur innebär både möjligheter och vissa begränsningar i planeringen. De innebär en språngbräda för nya produkter, som inte behöver utvecklas från grunden. Dessa nya statistikprodukter kan dra nytta av och bygga på standardlösningar och standardkomponenter i den befintliga arkitekturen.

Statistikproduktionen ändras vid många statistikbyråer successivt från skräddarsydda "stuprör" för enstaka undersökningar/produkter mot en arkitektur med återanvändning av data, gemensamma verktyg, datalagring, statistiksystem med tjänster etc. Standarder förenklar utbyte av data och metadata, t.ex. mellan källor, undersökningar och länder. Självklart måste arkitekturen vara väl utformad och planerad för framtiden. Metodik är en viktig del, för att förutse framtida behov av statistik och data, framtida källor och insamlingsmöjligheter, statistisk inferens etc. Jämför avsnitt 7.3, och se exemplen där och nedan.



Tidskriften *Journal of Official Statistics* (JOS) ägnade sitt första nummer 2013 åt system och arkitekturer för högkvalitativ statistikproduktion. Flera nationella statistikproduktionssystem och pågående förändringar beskrivs. Eltinge, Biemer och Holmberg (2013) presenterar ett potentiellt ramverk, som inkluderar t.ex. (i) undersökning, kvalitet, kostnader och användarnytta, (ii) integration av systemarkitektur med modeller för total undersökningskvalitet och adaptiv totaldesign, (iii) eventuell användning av begrepp från GSBPM och GSIM samt (iv) rollen för *governance*/styrningsprocesser i den praktiska implementeringen. Akronymen GSIM står för *Generic Statistical Information Model*.

Även begreppet och termen metodarkitektur används, se artiklarna i det nämnda JOS-numret och fjärde numret 2014 av *Statistical Journal of the IAOS: Journal of the International Association for Official Statistics* (SJIAOS), där sex artiklar på sidorna 369–398 beskriver metodarkitektur och tillämpningar på några centralbyråer, inklusive en sammanfattande diskussion av Karlberg och Radermacher (2014).

Nya Zeelands centralbyrå, som troligen var först med en processkarta, har ändrat mycket i sin produktionsmiljö, se exempelvis bidragen i JOS och SJIAOS.

## 7.5 Förbättringsarbete

### 7.5.1 Inledande motivering

En viktig del i modern statistikproduktion – och påtaglig sedan många år – är ett synsätt med ständiga förbättringar. Kvalitetshöjning och kostnadssänkning är två huvudskäl till förbättringsarbete. Minskad uppgiftslämnarvärden är ytterligare ett skäl. Användarnas behov och önskemål ska sammanställas och prioriteras tillsammans med tillgänglig budget nu och, för löpande undersökningar, även på sikt. Slutproduktens kvalitet (och kostnader) beror i sin tur av processkvaliteten (och av den institutionella miljön, som inte diskuteras här).

För statistikproduktionsprocessen generellt pågår utveckling, förbättring och implementering av gemensamma metoder, verktyg och arbetssätt. Varje löpande undersökning observerar och förbättrar sin verksamhet, t.ex. resursallokeringen. Den bevakar vad som händer i Verksamhetsstödet och implementerar nya och vidareutvecklade metoder, verktyg och arbetssätt när det är relevant.

En utvärdering av en undersökning kan i många fall flera dra nytta av, t.ex. likartade engångsuppdrag. Även det gemensamma arbetet med processer kan dra nytta av erfarenheter från enskilda undersökningar, t.ex. för arbetsrutiner, mallar och andra verktyg.

### 7.5.2 Produktionsstöd med team

Det är naturligt med årliga planer som tas fram under den allmänna planeringscykeln (verksamhetsplaneringen). Planerna kan behöva uppdateras under året. Nya behov kan tillkomma och inbördes prioriteringar kan ändras. Det kan komma nya (versioner av) metoder och verktyg, som är relevanta och viktiga. Det är lämpligt att också ha en mer långsiktig planering med en tidshorisont på exempelvis 2-5 år, ofta på en grövre nivå.

De olika kompetenserna som finns i produktionsteamet ska tas tillvara i utvärderingar, slutsatser och återkopplingar. Det gäller både mindre justeringar och större ändringsförslag. Kompetenserna ska hjälpas åt med förslag och prioriteringar. Det totala utrymmet styrs i första hand från och på en högre organisatorisk nivå i linjen genom avvägningar.

Det är viktigt att de olika rollerna/kompetenserna känner ett ansvar för såväl det egna området som helheten. Det gäller både genomförande (löpande produktion) och förbättringsarbete med vidare- och nyutveckling. Utformning och förbättring av en statistisk undersökning kräver både specifik kunskap av många olika slag och en helhetssyn på t.ex. statistikens kvalitet och kostnader.

### 7.5.3 Översyner – plan med delar

För en befintlig löpande statistisk undersökning handlar den årliga utformningen många gånger mest om justeringar på basis av processdata och åtgärdsplanen från utvärderingen av tidigare omgångar. Det är ett viktigt förbättringsarbete, och det ska vara tydligt vem som ansvarar för vad, se vidare avsnitt 7.5.4.

Det kan handla om förskjutningar i tidsplaneringen eller resursallokeringen, t.ex. för att öka inflödet eller förbättra granskningen. Både kvalitetssäkring och kvalitetskontroll behövs. Blev produktkvaliteten den avsedda? Tillförlitligheten är normalt svår att mäta, men uppskattningar ska göras, åtminstone av osäkerhetskällorna i grova drag. Om brister framkommer är de viktiga att hantera – och på något sätt åtgärda i en löpande undersökning. Jämför avsnitt 4.5.

Justeringar av produktionsprocessen räcker dock inte – det är motiverat att med viss regelbundenhet göra en mer genomgripande översyn med tanke på t.ex. samordning och effektivisering.

Det kan vara praktiskt att dela upp en översyn av statistikproduktionen i delar som går igenom successivt under en period av flera år för att få en jämnare belastning. Det är dock inte alldeles enkelt att göra en uppdelning eftersom allt hänger samman. Om dokumentationen är ofullständig ska den kompletteras, t.ex. så att metodmässigt "svarta lådor" öppnas och klagörs. Åtgärdsplanen kan indikera vad

som bör prioriteras. Nya eller utbyggda gemensamma metoder och verktyg kan också påverka när och hur en översyn lämpligen görs.

Om det finns idéer och förslag till förbättringar kan det vara lämpligt att göra ett inbäddat experiment, se avsnitt 5.6.2. Frågeställningar kan då vara om det blir tidsseriebrott, om det blir en kvalitetshöjning, om det blir en kostnadsminskning, etc.

Att införa en ändring innebär ett merarbete som kan vara betydande. Det gäller att planera väl och att beakta de risker som tillkommer. Beakta, som alltid, helheten, så att en förbättring på ett ställe inte medför en oväntad försämring någon annanstans. Se vidare t.ex. avsnitten 6.7 och 6.8.

#### **7.5.4 Planering, resurser och åtaganden (motsvarande)**

Oavsett om en undersökning görs en eller flera gånger krävs en god planering så att den tänkta designen kan realiseras. Detta måste säkerställas i grova drag redan då informationsbehoven fastställs och avtal (i någon form) skrivs.

Personresurser med rätt kompetens ska finnas i rätt tid. Metoder och verktyg ska vara valda och färdigutformade, dvs. vara både ny- eller vidareutvecklade och testade. Detta gäller även produktionsflödet, så att alla delar hänger samman och fungerar smidigt tillsammans. Det ska finnas tillräckliga marginaler och en beredskap att agera om något skulle inträffa som gör att åtgärder måste vidtas. Det kan t.ex. vara en försenad registerleverans, ett oväntat högt granskningsbehov eller ett resultat som ser orimligt ut.

Dokumentation av planering och målsättningar på kort och lite längre sikt behövs, särskilt när flera avdelningar är inblandade. Arbetsfördelningen med åtaganden och olika leveranser behöver beskrivas övergripande och detaljerat. Det gäller t.ex. mellan datainsamlings- och ämnesenhet samt mellan NR- och ämnesenhet. Se Verksamhetsstödet process 2.8.2 för upplägg och mallar. På en detaljerad nivå gäller det data: leveranser eller hämtningar av preciserade data vid preciserade tidpunkter.

När en löpande undersökning gör ändringar, är det extra viktigt att planera med marginal och att informera alla berörda så att det finns en beredskap att agera. Det tar i allmänhet längre tid när något görs för första gången jämfört med en inarbetad rutin. Ett litet misstag kan få stora följdverkningar. Är inflödet det normala? Ger granskningen normala indikationer? Det kan vara klokt med en tidig rimlighetsbedömning, t.ex. genom en makrogranskning.

## 8 Referenser

Baker, R., Brick, J.M., Bates, N.A., Battaglia, M., Couper, M.P., Dever, J.A., Gile, K.J. och Tourangeau, R. (2013). Summary Report of the AAPOR Task Force on Non-probability Sampling, *Journal of Survey Statistics and Methodology*, **1**, 90–105 (och Rejoinder 137–143).

Bakker, B. F. M., van der Heijden, P.G.M. och Scholtus, S. (2015). Preface to Special Issue on Coverage Problems in Administrative Sources. *Journal of Official Statistics*, **31**, 349–355.

Bakker, B.F.M., van Rooijen, J. och van Toor, L. (2014). The System of social statistical datasets of Statistics Netherlands: An integral approach to the production of register-based social statistics. *Statistical Journal of the IAOS*, **30**, 411–424.

Ballin, M. och Barcaroli, G. (2013). Joint determination of optimal stratification and sample allocation using genetic algorithm. *Survey Methodology*, **39**, 369–393.

Barnett, V. (1999). *Comparative Statistical Inference*. Third Edition. Wiley.

Beaumont, J.-F., Bocci, C. och Haziza D. (2014). An Adaptive Data Collection Procedure for Call Prioritization. *Journal of Official Statistics*, **30**, 607–621.

Beaumont, J.-F., Haziza, D. och Ruiz-Gazen, A. (2013). A unified approach to robust estimation in finite population sampling. *Biometrika*, **100**, 555–569.

Benedetti, R., Bee, M., och Espa, G. (2010). A Framework for Cut-off Sampling in Business Survey Design. *Journal of Official Statistics*, **26**, 651–671.

Biemer, P.P. (2010). Total Survey Error: Design, Implementation, and Evaluation. *Public Opinion Quarterly*, **74**, 817–848.

Brewer, K.R.W. (1999). Design-based or Prediction-based Inference? Stratified Random vs Stratified Balanced Sampling. *International Statistical Review*, **67**, 35–47.

Brewer, K. (2013). Three controversies in the history of survey sampling. *Survey Methodology*, **39**, 249–262.

Buelens, B., Boonstra, H.J., van den Brakel, J. och Daas, P. (2012). Shifting paradigms in official statistics. From design-based to model-based to algorithmic inference. Discussion paper (201218), Statistics Netherlands.

Chambers, R.L. (2011). Which sample survey strategy? A review of three different approaches. *Pakistan Journal of Statistics*, **27**, 337–357. – Länk (2015-05-18): [http://www.pakjs.com/journals/27\(4\)/27\(4\)2.pdf](http://www.pakjs.com/journals/27(4)/27(4)2.pdf)

- Choudhry, G.H., Rao, J.N.K. och Hidiroglou, M.A. (2012). On sample allocation for efficient domain estimation. *Survey Methodology*, **38**, 23–29.
- Citro, C.F. (2014). From multiple modes for surveys to multiple data sources for estimates. *Survey Methodology*, **40**, 137–161.
- Dalenius, T. (1974). *Ends and means of total survey design*. Forskningsprojektet "Fel i undersökningar". Stockholms universitet.
- Deville, J.-C. och Tillé, Y. (2004). Efficient balanced sampling: the cube method. *Biometrika*, **91**, 893–912.
- Eltinge, J.L., Biemer, P.P. och Holmberg, A. (2013). A Potential Framework for Integration of Architecture and Methodology to Improve Statistical Production Systems. *Journal of Official Statistics*, **29**, 125–145.
- ESSnet Admin Data (2011). WP 4: *Timeliness of Administrative data*. Admin Data Wiktionary. Se t.ex. *Deliverable 4.2 & 4.3* – Länk (2015-06-01): <http://essnet.admindata.eu/WorkPackage?objectId=4255>
- ESSnet Memobust (2014). *Theme Module: Estimation with Administrative Data*. – Länk (2015-06-01): <http://www.crosportal.eu/content/estimation-administrative-data-theme>
- Eurostat (2011). Riktlinjer för europeisk statistik. (European Statistics Code of Practice). Reviderad version 2011. – Länk (2015-05-18): <http://ec.europa.eu/eurostat/documents/3859598/6223797/10425-SV-SV.PDF>
- Eurostat (2012). *Quality Assurance Framework of the European Statistical System. Version 1.1*. Eurostat webpage.– Länk (2015-05-18): [http://ec.europa.eu/eurostat/documents/64157/4392716/qaf\\_2012-en.pdf/8bcff303-68da-43d9-aa7d-325a5bf7fb42](http://ec.europa.eu/eurostat/documents/64157/4392716/qaf_2012-en.pdf/8bcff303-68da-43d9-aa7d-325a5bf7fb42)
- Eurostat (2014). ESS Handbook for Quality Reports. Eurostat Manuals and guidelines. – Länk (2015-05-18): <http://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf/18dd4bf0-8de6-4f3f-9adb-fab92db1a568>
- Eurostat (2015). The ESS Vision 2020. (Odaterad, hämtad på Eurostats hemsida 2015, troligen från 2014). – Länk (2015-06-09): <http://ec.europa.eu/eurostat/documents/10186/756730/ESS-Vision-2020.pdf/8d97506b-b802-439e-9ea4-303e905f4255>
- Forbes, S. och Brown, D. (2012). Conceptual thinking in national statistics offices. *Statistical Journal of the IAOS*, **28**, 89–98.

- Godbout, S. (2011). Standardization of post-collection processing in Business Surveys at Statistics Canada. *Proceedings of Statistics Canada Symposium 2011*. Länk (2015-11-06):  
[http://publications.gc.ca/collections/collection\\_2013/statcan/11-522-x/CS11-522-2011-eng.pdf](http://publications.gc.ca/collections/collection_2013/statcan/11-522-x/CS11-522-2011-eng.pdf)
- Groves, R.M., Fowler, F.J. Jr., Couper, M.P., Lepkowski, J.M., Singer, E. och Tourangeau, R. (2009). *Survey Methodology*. Second Edition. Wiley.
- Groves, R.M. och Lyberg, L. (2010). Total Survey Error: Past, Present, and Future. *Public Opinion Quarterly*, **74**, 849–879.
- Groves, R.M. och Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias. A Meta-Analysis. *Public Opinion Quarterly*, **72**, 167–189.
- Hedlin, D., Lindkvist, H., Bäckström, H. och Erikson, J. (2008). An Experiment on Perceived Survey Response Burden among Businesses. *Journal of Official Statistics*, **24**, 301–318.
- Holmberg, A. (2004). Pre-printing Effects in Official Statistics: An Experimental Study. *Journal of Official Statistics* **20**, 341–355.
- Hox, J.J., De Leeuw, E.D. och Zijlmans, E.A.O. (2015). Measurement equivalence in mixed mode surveys. *Frontiers in Psychology*, **6**, Article 87, 1–11. – Länk (2015-06-19):  
<http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00087/full>
- Karlberg, M. och Radermacher, W. (2014). Discussion: Methodology Architecture – an area under construction. *Statistical Journal of the IAOS*, **30**, 393–398.
- Karlberg, M., Reis, F., Calizzani, C. och Gras, F. (2015). A toolbox for a modular design and pooled analysis of sample survey programmes. *Statistical Journal of the IAOS*, **31**, 447–462.
- Kass, R.E. (2011). Statistical Inference: The Big Picture. *Statistical Science*, **26**, 1–9.
- Kokic, P. N. och Bell, P. A. (1994). Optimal Winsorizing Cutoffs for a Stratified Finite Population Estimator. *Journal of Official Statistics*, **10**, 419–435.
- Köll, J. (2007) Mot mer samordnade revideringar i ekonomisk statistik. Slutrapport 2007-11-26. SCB. – Länk (2015-06-01):  
[http://www.scb.se/statistik/\\_publikationer/NR9999\\_2010A01\\_BR\\_21\\_NR\\_FT1002.pdf](http://www.scb.se/statistik/_publikationer/NR9999_2010A01_BR_21_NR_FT1002.pdf)
- Laux, R. och Barham, C. (2013). Can't see the wood for the trees? Using statistical frameworks to draw the bigger picture. *Statistical Journal of the IAOS*, **29**, 209–221.

- Linacre, S.J. och Trewin, D.J. (1993). Total Survey Design – Application to a Collection of the Construction Industry. *Journal of Official Statistics*, **9**, 611–621.
- Little, R.J. (2012). Calibrated Bayes, an Alternative Inferential Paradigm for Official Statistics. *Journal of Official Statistics*, **28**, 309–372; inklusive en diskussion med sex inlägg och ett svar.
- Lohr, S.L. (2010). Sampling: Design and Analysis. Second Edition. Duxbury Press.
- Lundquist, P. och Särndal, C.-E. (2013). Aspects of Responsive Design with Applications to the Swedish Living Conditions Survey. *Journal of Official Statistics*, **29**, 557–582.
- Lyberg, L. (2012). Survey Quality. *Survey Methodology*, **38**, 107–130.
- Makar, K. och Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, **8**, 82–105.
- Marella, D. (2007). Errors Depending on Costs in Sample Surveys. *Survey Research Methods*, **1**, 85–96.
- Martinoz, C.F., Haziza, D. och Beaumont, J.-F. (2015). A method of determining the winsorization threshold, with an application to domain estimation. *Survey Methodology*, **41**, 57–77.
- Meinke, I. och Schmidt, P. (2014). Good governance in statistics: The development of a revision policy for the German Federal Statistical Office in the European context. Bidrag till konferensen Q2014. – Länk (2015-06-01):  
[http://www.q2014.at/fileadmin/user\\_upload/MeinkeSchmidt\\_paper\\_revision\\_policy.pdf](http://www.q2014.at/fileadmin/user_upload/MeinkeSchmidt_paper_revision_policy.pdf)
- Merad, S. och Brodie, P. (2011). Standardizing UK sub-annual Business Surveys. *Proceedings of Statistics Canada Symposium 2011*. Länk (2015-11-06):  
[http://publications.gc.ca/collections/collection\\_2013/statcan/11-522-x/CS11-522-2011-eng.pdf](http://publications.gc.ca/collections/collection_2013/statcan/11-522-x/CS11-522-2011-eng.pdf)
- Mulry, M.H., Oliver, B.H., och Kaputa, S. H. (2014). Detecting and Treating Verified Influential Values in a Monthly Retail Trade Survey. *Journal of Official Statistics*, **30**, 721–747.
- OECD (2007). Data and Metadata Reporting and Presentation Handbook. Länk (2015-06-01): <http://www.oecd.org/std/37671574.pdf>
- Ouwehand, P. och Schouten, B. (2014). Measuring Representativeness of Short-Term Business Statistics. *Journal of Official Statistics*, **30**, 623–649.

- Persson, A. (2013). Mixed mode – datainsamlingsstrategier & blankett-design. SCB. Länk (2015-11-04):  
<http://verksamhetsstod/sites/vstod/statprod/process2/Beskrivningssidor/Mixed%20mode%20-%20datainsamling-%20och%20blankettdesign.pdf>
- Roberts, C., Vandenplas, C. och Stähli, M.E. (2014). Evaluating the Impact of Response Enhancement Methods on the Risk of Nonresponse Bias and Survey costs. *Survey Research Methods*, **8**, 67-80.
- ROS (2013). Riktlinjer för europeisk statistik (Code of Practice) – en handledning och exempelsamling. – Länk (2015-05-18):  
<http://www.scb.se/Grupp/OmSCB/Verksamhet/ROS/Dokument/Riktlinjer-for-europeisk-statistik-Code-of-Practice.pdf>
- SCB (2002). Guide till granskning. CBM 2002:1. – Länk (2015-05-18):  
[http://www.scb.se/statistik/\\_publikationer/OV9999\\_2003M00\\_BR\\_X97%3%96P0201.pdf](http://www.scb.se/statistik/_publikationer/OV9999_2003M00_BR_X97%3%96P0201.pdf)
- SCB (2007). Kvalitetssäkring av jämförbarhet över tiden. Slutrapport 2007-05-09 från projektet Lotta P1 Kvalitetssäkring av jämförbarhet över tid.
- SCB (2008). Urval – från teori till praktik. Handbok 2008:1. – Länk (2015-05-18):  
[http://www.scb.se/statistik/\\_publikationer/ov9999\\_2007a01\\_br\\_x99br0801.pdf](http://www.scb.se/statistik/_publikationer/ov9999_2007a01_br_x99br0801.pdf)
- SCB (2010). Långa tidsserier i nationalräkenskaperna. Slutrapport 2010-11-03. – Länk (2015-06-01 till ViP): [Slutrapport Långa tidsserier i NR\\_ver 2\\_101103](#)
- SCB (2012). Dokumentation av införandet av SNI 2007. Bakgrundsfakta till Ekonomisk statistik 2012:1. – Länk (2015-06-01):  
[http://www.scb.se/statistik/\\_publikationer/NV0101\\_2012A01\\_BR\\_X100BR1201.pdf](http://www.scb.se/statistik/_publikationer/NV0101_2012A01_BR_X100BR1201.pdf)
- SCB (2014a). Riktlinjer för riskhantering vid SCB. – Länk (2015-06-15 till Inblick; Dnr 2014/28):  
[http://inblick.scb.intra/globalassets/dokument/staben/policy\\_riktlinjer/isik/riktlinjer-for-riskhantering-2014-till-gd-foredragning-rev.pdf](http://inblick.scb.intra/globalassets/dokument/staben/policy_riktlinjer/isik/riktlinjer-for-riskhantering-2014-till-gd-foredragning-rev.pdf)
- SCB (2014b). STRATEGI 2020. – Länk (2015-06-09 till Inblick):  
<http://inblick.scb.intra/scbs-verksamhet/Fakta-om-SCB/strategi-2020/>
- SCB (2016a). Sveriges officiella statistik – årlig rapport för 2015. Statistiska centralbyrån, Sekretariatet vid Rådet för den officiella statistiken. – Länk (2016-12-19):  
[http://www.scb.se/Statistik/\\_Publikationer/OV9999\\_2015A01\\_BR\\_X43BR1601.pdf](http://www.scb.se/Statistik/_Publikationer/OV9999_2015A01_BR_X43BR1601.pdf)



- SCB (2016b). Statistiska centralbyråns föreskrifter (SCB-FS 2016:17) om kvalitet för den officiella statistiken – Länk (2016-12-12):  
<http://www.scb.se/Grupp/OmSCB/Verksamhet/Foreskrifter/Tillganglighet-SOS/SCB-FS-2016-17.pdf>
- SCB (2016c). Kvalitet för den officiella statistiken – en handbok. – Länk (2016-12-20; intern respektive extern):  
<http://verksamhetsstod/sites/vstod/statprod/process9/Beskrivningssidor/Kvalitet-for-den-officiella-statistiken-en%20handbok.pdf>  
<http://www.scb.se/Grupp/OmSCB/Verksamhet/ROS/Dokument/Kvalitet-for-den-officiella-statistiken-WEB.PDF>
- Schouten, B., Calinescu, M. och Luiten, A. (2013). Optimizing quality of response through adaptive survey designs. *Survey Methodology*, **39**, 29–58.
- Smith, P.A. (2014). Redesign in a statistical office – sociodemographic statistics in the Office for National Statistics. *Statistical Journal of the IAOS*, **30**, 431–435.
- Snijkers, G., Haraldsen, G., Jones, J., och Willimack, D.K. (2013). *Designing and conducting business surveys*. John Wiley and Sons, Inc.
- Steel, D.G. och Clark, R.G. (2014). Potential gains from using unit level cost information in a model-assisted framework. *Survey Methodology*, **40**, 231–242.
- Särndal, C.-E., Swensson, B. och Wretman, J. (1992). *Model Assisted Survey Sampling*. New York: Springer Verlag.
- Tourangeau, R., Rips, L.J. och Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge University Press, Cambridge, UK.
- UNECE (2013), Generic Statistical Business Process Model. Version 5.0 – December 2013. The United Nations Economic Commission for Europe (UNECE). – Länk (2015-05-18):  
<http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0>
- Van den Brakel, J.A. och Renssen, R.H. (2005). Analysis of Experiments Embedded in Complex Sampling Designs. *Survey Methodology*, **31**, 23–40.
- Van den Brakel, J.A., Smith, P.A. och Compton, S. (2008). Quality procedures for survey transitions – experiments, time series and discontinuities. *Survey Research Methods*, **2**, 123–141.
- Vannieuwenhuyze, J.T.A. och Revilla, M. (2013). Evaluating Relative Mode Effects on Data Quality in Mixed-Mode Surveys. *Survey Research Methods*, **7**, 157–168.
- Zhang, L.-C. (2012). Topics of statistical theory for register-based statistics and data integration. *Statistics Neerlandica*, **66**, 41–63.